DISS. ETH No. 21112

MODELLING THE SPACE-TIME STRUCTURE OF PRECIPITATION AND ITS IMPACT ON BASIN RESPONSE

A dissertation submitted to
ETH ZURICH

for the degree of
Doctor of Sciences

presented by

ATHANASIOS PASCHALIS
Dipl. Civil Engineering, National Technical University of Athens

born 14 September 1986
citizen of Greece

accepted on the recommendation of
Prof. Dr. Paolo Burlando, examiner
Prof. Dr. Peter Molnar, co-examiner
Prof. Dr. Efi Foufoula-Georgiou, co-examiner
Dr. Simone Fatichi, co-examiner

Zurich, 2013

ABSTRACT

An improved understanding of the spatio-temporal statistical structure of precipitation across scales is a prerequisite for making connections with the physics of precipitation formation and for developing advanced stochastic modelling techniques useful in operational hydrology and especially in natural hazard risk management. The aim of this thesis is to investigate the precipitation structure in space and time for the orographically complex area of the European Alps and develop appropriate stochastic simulation tools for its modelling. With these tools, the impact of the spatial and temporal structure of precipitation on the response of river basins is investigated.

The first part of the thesis focuses on the description of the statistical structure of the precipitation process in space and time. The statistical theories based on the notion of scale invariance are presented, and their applicability to precipitation time series and radar precipitation fields are assessed. Several issues concerning the parameter estimation are raised and the influence of the measurement errors of rain-gauges and radars are quantified. After establishing optimal parameter estimation techniques in terms of bias and robustness, an extensive data analysis of the precipitation records of the Federal Office of Meteorology and Climatology of Switzerland (MeteoSwiss) consisting of long and reliable point scale measurements (68 rain-gauges with 10 minutes temporal resolution [1981-2009]) and radar measurements (7 years of 5 min data) is conducted. The general applicability of the theory of scale invariance is tested by the data analysis and limitations are presented. Connections between the physical mechanisms that trigger precipitation processes and their respective statistical structure are identified and discussed.

The second goal of this thesis is to provide appropriate stochastic modelling techniques for precipitation. For temporal rainfall, an extensive intercomparison of the state of the art modelling approaches is presented. Their main capabilities and deficiencies are presented and new modelling techniques are proposed. For temporal rainfall, new modelling techniques of combining various classes of stochastic models are developed. For space-time rainfall a new stochastic model (STREAP) is developed that captures most of the key structural and dynamic characteristics of rainfall as captured by weather radars. The model is shown to outperform other traditional and well established space-time stochastic models for rainfall based on Poisson processes.

Finally, the sensitivity of basin response to precipitation structure is investigated through an extensive numerical experiment that combines the new stochastic rainfall model (STREAP) and a detailed fully-distributed hydrological model (TOPKAPI-ETH). Different response patterns dependent on various structural and kinematic patterns of the precipitation fields are identified and their respective discharge generating mechanisms are studied. This study leads to results which demonstrate which parameters of the precipitation structure influence flood discharge and can potentially lead to an optimal decision of the precipitation monitoring network density for flood prediction.

ZUSAMMENFASSUNG

Ein besseres Verständnis der räumlichen und zeitlichen statistischen Struktur des Niederschlags über verschiedene Skalen ist eine Voraussetzung für die Verknüpfung mit der Physik der Niederschlagsbildung und für die Entwicklung fortgeschrittener stochastischer Modelliertechniken, welche nützlich sind für angewandte hydrologische Fragestellungen und vor allem im Risikomanagement von Naturgefahren. Das Ziel dieser Forschungsarbeit ist die Niederschlagsstruktur in Raum und Zeit in einem orographisch komplexen Gebiet der europäischen Alpen zu untersuchen und geeignete stochastische Simulationstechniken zu entwickeln, um diese Prozesse zu modellieren. Damit wird der Einfluss der räumlichen und zeitlichen Struktur des Niederschlags auf das Abflussverhalten im Einzugsgebiet untersucht.

Der erste Teil der Forschungsarbeit fokussiert auf die Beschreibung der statistischen Struktur des Niederschlagsprozesses in Raum und Zeit. Die statistischen Theorien, welche auf der Notation der Skaleninvarianz basieren, werden präsentiert, und deren Anwendbarkeit im Bereich von Niederschlagszeitreihen und Radar Niederschlagsfeldern wird abgeschätzt. Verschiedene Probleme bezüglich der Parameterschätzung werden aufgezeigt, und der Einfluss des Messfehlers der Pluviometer und des Radars wird quantifiziert. Nach der Festlegung der optimalen Technik zur Parameterschätzung in Bezug auf den systematischen Messfehler und die Robustheit wird eine ausführliche Datenanalyse der Niederschlagsmessreihen des Bundesamtes für Meteorologie und Klimatologie der Schweiz (MeteoSchweiz) bestehend aus langen und verlässlichen Punktmessungen (68 Pluviometer mit 10 min Auflösung [1981-2009]) und Radarmessungen (7 Jahre mit 5 min Auflösung) durchgeführt. Die allgemeine Anwendbarkeit der Theorie der Skaleninvarianz wird durch die Datenanalyse getestet, und die Einschränkungen werden aufgezeigt. Die Verknüpfung zwischen den physikalischen Mechanismen, welche den Niederschlagsprozess auslösen, und der entsprechenden statistischen Struktur wird identifiziert und diskutiert.

Das zweite Ziel dieser Forschungsarbeit ist die Entwicklung von geeigneten stochastischen Modellierungstechniken für den Niederschlag. Für zeitlichen Niederschlag wird ein ausführlicher Vergleich der Modellierungsansätze gemäss dem aktuellen Stand der Forschung präsentiert. Ihre wichtigsten Einsatzmöglichkeiten und ihre Defizite werden aufgezeigt, und eine neue Modellierungstechnik wird vorgeschlagen. Für zeitlichen Niederschlag werden basierend auf einer Kombination verschiedener Klassen von stochastischen Modellen neue Modellierungstechniken entwickelt. Für Raum-Zeit Niederschlag wird ein neues stochastisches Modell (STREAP) entwickelt, welches die wichtigsten strukturellen und dynamischen Charakteristiken des Niederschlags gemäss Wetterradarbeobachtungen erfasst. Es wird aufgezeigt, dass das Modell andere traditionelle und etablierte stochastische Raum-Zeit Modelle, welche auf dem Poisson Prozess basieren, übertrifft.

Zum Schluss wird anhand eines ausführlichen numerischen Experiments, welches das neue stochastische Niederschlagsmodell (STREAP) mit einem detaillierten räumlich diskretisierten hydrologischen Modell (TOPKAPI-ETH) kombiniert, die Sensitivität der Antwort des Einzugsgebiets auf die Niederschlagsstruktur untersucht. In Abhängigkeit von verschiedenen strukturellen und kinematischen Mustern des

Niederschlagsfelds werden verschiedene Abflussmuster identifiziert, und der entsprechende Abflussbildungsmechanismus wird untersucht. Diese Studie führt zu Resultaten, welche aufzeigen welche Parameter der Niederschlagsstruktur den Hochwasserabfluss beeinflussen, und kann daher potentiell als Entscheidungshilfe für die Optimierung der Dichte des Niederschlagsmessnetzes für Hochwasservorhersage eingesetzt werden.


Precipitation
Random
Stochastic
Modelling
structure
Fractal
Statistical
Spatiotemporal
multi-scaling
Scaling
time-series
Orography
discharge
response
convective
stratiform
Bartlett-Lewis
Markov
Switzerland
measurement
stricture
Multiplicative
risk
rainfall
process
basin
Alps
comparison
STREAP
flooding
advection
gauge
snow
Neyman-Scott
bias
hydrology
Cascade
Poisson
variability
radar
complex
mesoscale
impact
data

ACKNOWLEDGEMENTS

Approaching the end of my thesis I would like to thank all those people that helped me during the last four years of my studies and made this thesis possible.

First of all I would like to thank my advisors Prof. Paolo Burlando and Prof. Peter Molnar who offered me continuous advice and support. This thesis would be impossible without their help and enthusiasm. I also would like to thank Dr. Simone Fatichi. His ideas led to a great improvement of this thesis. I would also like to thank him for his guidance that helped me me to broaden my knowledge in the field of Geosciences.

Moreover, I would like to thank all the people that have helped me with discussions and suggestion concerning my PhD research. In particular I would like to thank Prof. Efi Foufoula-Georgiou who accepted my invitation to serve as an examiner of my thesis and also provided me with a great feedback for my research during my visit to the Saint Anthony Falls Laboratory. In addition, I would like to thank all my colleagues from the HWRM group of the Institute of Environmental Engineering. Their help, comments and suggestions have been very useful during my PhD research.

I would also like to express my sincere gratitude to Prof. Demetris Koutsoyiannis my professor from the National Technical University of Athens that first initiated my strong interest in hydrology and especially in probability theory and stochastic processes.

Last but not least I would like to thank my family and my friends for their continuous support and encouragement.

Finally I would like to acknowledge the financial support of the Swiss National Science Foundation. Funding for this research was provided by the SNSF grant 200021- 120310.

Zurich, 2013

CONTENTS

LIST OF FIGURES

LIST OF TABLES

1

GENERAL INTRODUCTION

Cloud formation and precipitation are fundamental processes studied both in meteorology and hydrology. Precipitation, in particular, is the main driving force for the entire hydrological cycle. Moreover it is the cause of flooding, one of the most disastrous natural hazards. According to the National Oceanic and Atmospheric Administration (NOAA [1]) the average annual losses due to flooding only in the US are estimated about 2 billion US$. More importantly, flooding is also a major threat for human life worldwide (figure 1.1(a)). In the US where advanced flood protection infrastructure exists there are 97 fatalities a year on average due to flooding [2].An interesting characteristic of flood risk is that in contrast to all other water related threats, it is not substantially economically driven (figure 1.1(b)) [*Koutsoyiannis*, 2011]. This suggest that even developed countries with advanced flood protection infrastructure can be highly vulnerable and require investments in order to achieve a reliable civil protection system.

The problem of flooding is enhanced in mountainous areas which are prone to flash-floods that are difficult to predict [*Doswell et al.*, 1996]. A striking example were the floods of August 2005 in Switzerland that according to the Federal Department of the Environment, Transport, Energy and Communications (BAFU [3]) caused 6 casualties and a total financial loss over 2.5 billion Swiss francs (CHF).

These are some of the major reasons why a very good understanding of the general structure of precipitation and a good quantification of its statistical properties are crucial not only for scientific but also for engineering practice. The study of this process has thus attracted the interest of the scientific and engineering community worldwide for centuries.

These concerns form the main scientific question of this thesis which will address the statistical description of the precipitation process in space and time, its modelling, and its impact on the basin response and flood generation.

1.1 Precipitation as a Natural Process

Seeking to understand and describe the precipitation process in a formal way has been subject of several scientific fields such as atmospheric thermodynamics, cloud microphysics etc.

[1] http://www.noaa.gov/
[2] http://www.nws.noaa.gov/hic/flood_stats/recent_individual_deaths.shtml
[3] http://www.bafu.admin.ch/

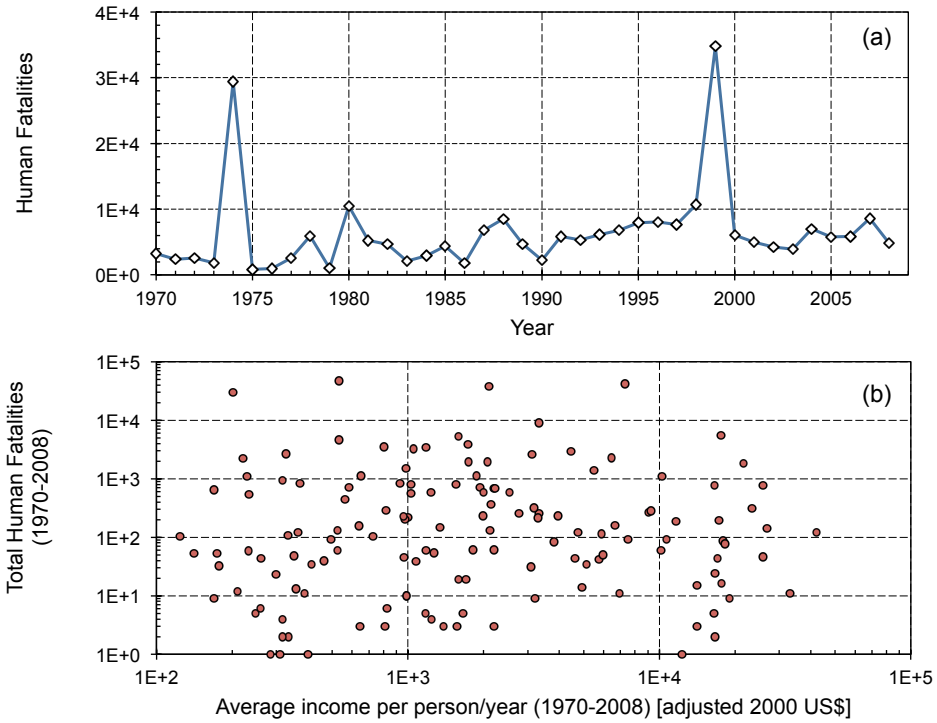


FIGURE 1.1: Data concerning human fatalities due to flooding (1970-2008). (a) Total number of fatalities per year. (b) Total number of fatalities in relation to the average income per person for every country in the world. Data source: http://www.gapminder.org/data/.

Briefly speaking, precipitation is the result of the coalescence of the condensed water vapour in the clouds and the gravitational force that initiates the fall of the water droplets to the ground. Extensive reviews of the various processes leading to coalescence can be found in any textbook of meteorology [e.g. *Wallace and Hobbs*, 1977; *Salby*, 1996]. An interesting feature of the precipitation formation processes is the interaction between the land surface and the atmosphere. In the presence of complex orographic patterns, various mechanisms can have a strong effect on precipitation generation and subsequently on its spatial and temporal patterns. The effects are more prominent for relatively small spatial scales (e.g. meso-β and meso-γ scale or even less). In figure 1.2 an illustration of the mechanisms that impact the precipitating clouds over mountains are shown after *Houze* [2012]. An excellent review of orographic precipitation can be found in *Roe* [2005]. The various effects result in the local enhancement or decrease of precipitation. This thesis focuses on precipitation in an area of the European Alps where orography plays a major role, modifying the local structure of precipitation in a very complex manner [e.g. *Bougeault et al.*, 2001; *Houze et al.*, 2001; *Schiesser et al.*, 1995].

A review of the current knowledge of the physics of the precipitation formation process would be far beyond the scope of this thesis. Even thought precipitation is important for many fields, a complete physical understanding of its structure and dynamics is not yet available due to the high complexity of all the processes that are involved in precipitation formation [*Koutsoyiannis and Langousis*, 2011]. In principle, not even a universally acceptable framework for its statistical description exists due to the fascinating high variability of this phenomenon in space and time.

For this reason, rainfall modelling in any physically based numerical weather prediction model is still problematic. The gaps of understanding are even larger for complex terrains where the feedback from the

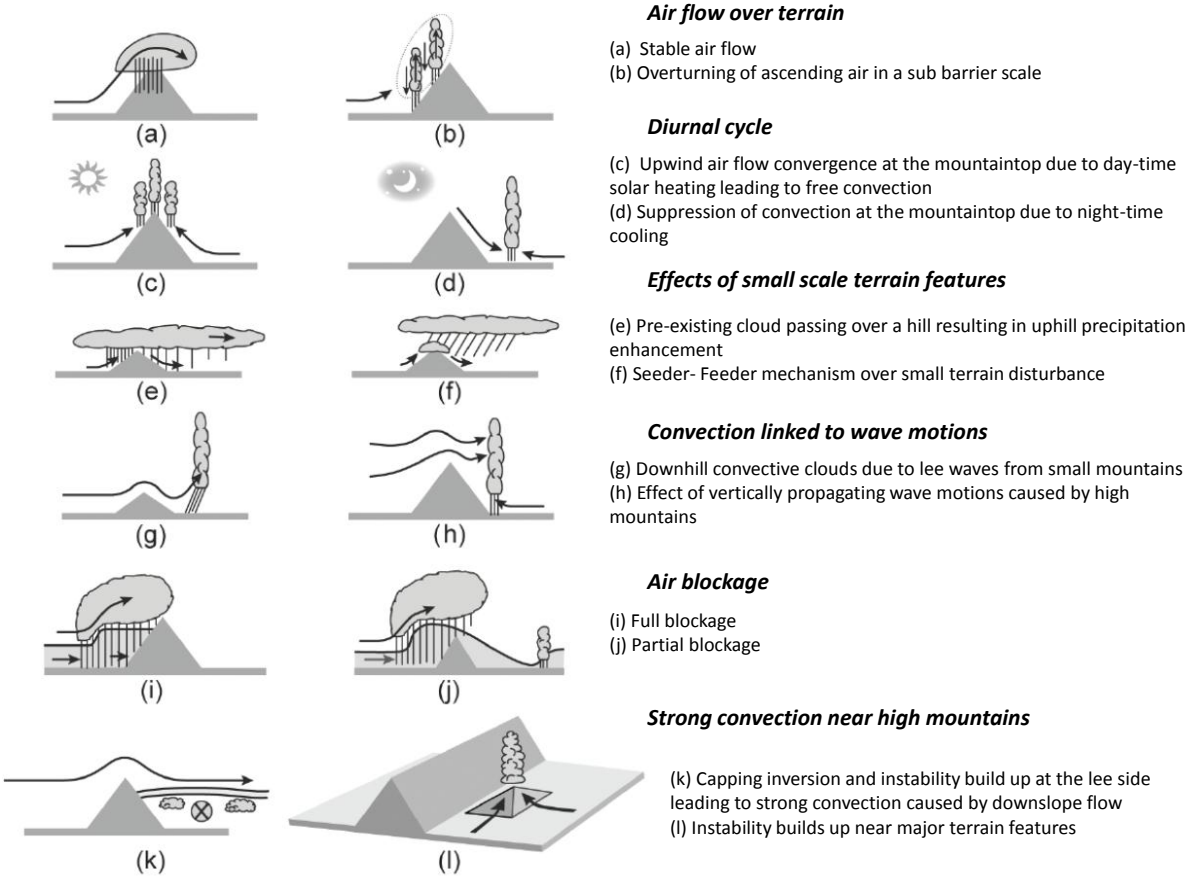


FIGURE 1.2: Orographic effects on the precipitation formation. Source: *Houze* [2012].

interactions between the atmospheric flow and topography play a major role. A clear recent illustration for these problems for the area of interest of this thesis can be found in *Bordoy and Burlando* [2012], who found that the detailed numerical models used for the fourth assessment report of IPCC [*IPCC*, 2007], which incorporate our state of the art knowledge concerning the thermodynamics of the atmosphere, cloud formation processes and land-atmosphere interactions, fail to capture even the seasonality of the precipitation for the Alpine complex terrain. The major problems are mainly due to our lack of complete knowledge of the precipitation formation, especially in the presence of convection.

Meteorological studies have mainly focused on the numerical modelling of various precipitation triggering mechanisms [*Houze*, 1993] in contrast to hydrological studies that have mainly aimed at developing new frameworks for the statistical description of the process.

Statistically speaking, some of the most fascinating features of the precipitation processes are its intermittent nature and diverse variability across an enormous range of scales, spanning in space from the order of magnitude of millimetres (e.g. cloud microphysics) to global scales and, in time from seconds to centuries [*Blöschl and Sivapalan*, 1995]. In fact one of the main differences between precipitation and most other geophysical process is its discontinuity (intermittency) in space and time, that makes its statistical description much more challenging. Moreover, the highly non-Gaussian distributions [e.g. *Papalexiou and Koutsoyiannis*, 2011] and correlation of precipitation for a very large range of spatio-temporal scales is also an important feature of this process.

In order to describe and understand the precipitation process, it is first necessary to monitor it accurately. The history of precipitation measurement dates back to ancient times [e.g. *Strangeways*, 2010]. However, systematic measurement of precipitation started in the 19th century. The measurements were obviously on the point scale (rain-gauges) with depth resolution typically of one day or coarser. This type of measurement continues until today and the data availability is large due to the dense gauging networks that have been developed. An improvement of the measuring techniques has been achieved through time due to the large technological advancements of computer science and the automatic measuring procedures at high temporal scales and large electronic storing capabilities. Probably the greatest innovation concerning the measurement of precipitation was achieved after the Second World War when the capabilities of radars in quantifying precipitation were identified [*Marshall et al.*, 1947; *Maynard*, 1945; *Berne and Krajewski*, 2012]. Finally, the achievements of aeronautics and space exploration allowed the possibility of measuring precipitation from space, on a global scale, [e.g. TRMM, *Kummerow et al.*, 1998].

Advanced knowledge of the precipitation process is possibly the most important prerequisite in hydrology, since it is the driving force of the entire hydrological cycle. A primary concern in engineering is the correct estimation of the river discharge. A key question therefore is how the spatio-temporal structure of precipitation can affect the response of basins and runoff production. This is still an open question with many issues that still to be addressed and will be raised towards the end of this thesis.

1.2 Modelling of Precipitation as a Stochastic Process

Operational hydrology and especially the task of flood risk and water resources management require long and reliable precipitation records. Due to the fact that such data are not always available it may be necessary to model them.

A first idea could be to create long records by exploiting our knowledge about the physical responses of the atmospheric system that lead to cloud formation and precipitation. For such a physically based modelling approach to be useful there are two essential prerequisites. The first one concerns our understanding of the complex nature of the physical system that subsequently leads to its mathematical representation, and the second one concerns our ability of numerical computations. Even though such a mechanistic approach, based on weather prediction models could be ideal since it links directly the physics of the process to the numerically simulated precipitation, none of the two prerequisites are adequately met. Our lack of understanding of the complex feedbacks of the natural system leads to systematic errors in the quantification of precipitation using "physically"-based numerical schemes of the atmosphere [*Knutti*, 2008; *Koutsoyiannis et al.*, 2008]. Moreover the computational requirements of such models are enormous.

An elegant alternative solution to this problem which inherently includes the concepts of variability and uncertainty is with the use of probability theory and stochastic processes. The principle idea is that random sampling of equiprobable precipitation series/fields could be a substitute for the lack of measured data. The ideas of solving complex problems through large random simulation exploiting the *theory of large numbers* is not new. The first ideas date back to the studies of *Metropolis et al.* [1953] and *Metropolis and Ulam* [1949]. Even though the computational requirements in stochastic simulation may be high, they are substantially lower than the ones needed for a physically based numerical simulation scheme.

In order to apply stochastic modelling simulation, a rigorous statistical framework for the description of precipitation as a stochastic process is needed. Such a framework should reflect the important statistical properties of precipitation as observed by data. In this sense, stochastic modelling of precipitation is only

a conceptual representation of the natural process. Throughout the years various statistical theories have inspired researchers. Ideas originating from very diverse scientific fields have influenced substantially the study of precipitation, spanning from mining engineering [e.g. geostatistics, *Matheron*, 1963] and astronomy [*Neyman and Scott*, 1952] to physics and the study of turbulence [*Kolmogorov*, 1940; *Benzi et al.*, 1984; *Frisch*, 1991].

The simulation approach became very popular for tackling hydrological problems. Probably the first effort of applying stochastic modelling in hydrology goes back to the study of *Thomas and Fiering* [1962] who focused on the simulation of discharge series. Stochastic modelling of precipitation though, in contrast to other geophysical processes, has probably been the most challenging problem in stochastic modelling in natural sciences due to its multi-scale characteristics. Several ideas influenced the "*art*" of stochastic modelling of precipitation, starting from temporal modelling of precipitation for coarse time scales (e.g. daily) and evolving in time to complete high resolution space-time models. Also practical solutions have been achieved with such methodologies, in real world applications, such as water resources management [*Koutsoyiannis and Economou*, 2003] and storm/flood nowcasting [e.g. *Berenguer et al.*, 2011; *Burlando et al.*, 1996].

Physically based and stochastic approaches in precipitation modelling need not be exclusive. In the last decades the potential of combining information from both physically based models and the capabilities of stochastic rainfall models has been recognized. This research direction is mainly oriented at the investigation of future risk under a changing climate and linking the scale gaps between the large physically based numerical modelling of the atmosphere (e.g. GCMs, RCMs) and the small scale data requirements of hydrology [e.g. *Burlando and Rosso*, 2002; *Burton et al.*, 2010a; *Fatichi et al.*, 2011; *Fowler et al.*, 2007].

1.3 Motivation of the Study

The motivation of the study in this thesis is threefold. The first goal is to achieve a better understanding of the statistical structure of the precipitation process in an orographically complex area of the European Alps. The second goal is more engineering-oriented and aims at the development of appropriate stochastic modelling tools, suitable for the examined area, that could also serve as a practical tool for operational hydrology. The third goal is to investigate the influence of the spatio-temporal structure of rainfall on the basin response.

In order to achieve these goals, state of the art methods concerning the statistical structure of precipitation will be employed for an extensive data analysis, exploiting a large precipitation data base both for temporal and spatial precipitation within the region of Switzerland where the main mass of the mountain range of the Alps is located. From this analysis, a better insight about the structure of precipitation in complex terrain, with strong orographic influences will be achieved.

A basis for this thesis is that the area of Switzerland is monitored very efficiently for precipitation. Long term and high quality precipitation measurements span several decades, and come from diverse measurement techniques (gauges, weather radars etc) covering adequately the orographically complex terrain. This has resulted in a vast amount of data. A major motivation thus of this thesis and also one of its major novelties is that it is one of few extensive studies with simultaneous precipitation records from rain-gauges and weather radars. Previous studies were mostly case and site specific. Here the aim is to confirm or reject hypotheses about the spatio-temporal structure of precipitation with the strong support of a high

quality database.

A better understanding of the general statistical description of the precipitation processes in space and time is a prerequisite to develop the appropriate stochastic modelling approaches. A *main scientific question* in this thesis is which stochastic modelling techniques are the optimal for precipitation simulation; how do existing models perform; and how can they be improved in order to capture best the statistical properties of precipitation in space and time. As mentioned before, the excellent available database provides the possibility of a comprehensive stochastic model intercomparison, which can guide their applicability. To my knowledge, such a comparison has not yet been reported in the literature. A strong motivation of this thesis is to fill this gap, and achieve generalized results concerning stochastic modelling of precipitation and provide methodological advancement, by improving the available simulation tools.

Finally, a crucial question that motivates this work is the effect of the spatio-temporal variability of the precipitation structure on basin response in mountainous catchments. The practical implications of this question are many. The huge economic and social impacts that disastrous flash floods can cause are of major importance. This is the reason why intense research has been devoted to the investigation of the precipitation processes in those areas, especially in terms of data collection and precipitation monitoring [e.g. *Houze et al.*, 2001; *Savina*, 2011; *Schneebeli and Berne*, 2012]. The approach in this thesis is to use the analysis and simulation of precipitation in space and time for an investigation of the impact of rainfall variability on the basin response. The motivation is to combine knowledge and tools coming from both stochastic and process hydrology. The premise is that exploiting the developed tools of this thesis for precipitation modelling with advanced tools of hydrological modelling can result in unique conclusions about the impact of the space-time structure of precipitation on catchment response.

1.4 Thesis Outline

The thesis is structured into eight chapters.

After this short introduction, a detailed description of the study area is given in chapter 2. The topography and precipitation climatology are presented, and the precipitation data availability and monitoring networks are explained.

In chapter 3 the investigation of the temporal structure of precipitation for the study area is illustrated. The chapter begins with a detailed description of the mathematical framework on which this study is based followed by the analysis of the ground point-scale precipitation measurements (rain-gauges). Various problems of the statistical estimation procedures are demonstrated, robust analysis procedures are established and potential connections between the temporal precipitation structure and various physical parameters are explored.

Following the analysis of the time-series from a large rain-gauge data base an extensive stochastic modelling analysis for temporal precipitation is conducted in chapter 4. The applicability of traditional, well established, modelling tools is assessed. Their strengths and deficiencies are shown and new modelling approaches are proposed. Finally a generalized model intercomparison is performed and the results are validated at several stations across the world.

Chapters 5 and 6 follow the same analysis methodology as chapters 3 and 4 by generalizing the results from temporal precipitation to the whole picture of the spatio-temporal structure of precipitation. In chapter 5 the analysis of spatial precipitation derived from a large radar data base is presented. Similar problems to those identified for the time-series data are found also in the spatial domain. The statistics of

the spatial structure of precipitation are linked to the mechanisms that trigger these process.

In chapter 6 a novel spatio-temporal stochastic model for precipitation is presented. Building on previous work, the model developed here yields very promising results for high resolution space-time precipitation, being able to mimic the essential features of the precipitation processes including its basic dynamics (e.g. growth-decay), structural characteristics (e.g. spatial clustering) and kinematics (e.g. advection). The model is one of the few attempts for simulating high resolution space-time precipitation in a continuous manner, and is shown to outperform other state of the art stochastic models.

In chapter 7, exploiting the capabilities of the new spatio-temporal model, a numerical investigation of the impact of the rainfall variability in space and time on basin response is conducted. By employing a fully distributed hydrological model, the impact of precipitation and soil saturation on runoff generation is explored by simulation.

Finally, chapter 8 concludes the thesis, summarising all the findings and discussing their generality and importance in environmental engineering. Also suggestions for further research are provided.

2

STUDY AREA AND DATA DESCRIPTION

This thesis focuses on the geographic region of the European Alps, with specific emphasis on Switzerland. In this chapter a short description of the topographical features of the area and its climatology are given and an overview of the data availability concerning precipitation measurement is provided.

2.1 Precipitation Climatology

Precipitation climatologies for the complex orographic terrain of the European Alps have been constructed in the past [e.g. *Legates and Willmott*, 1990; *Hulme et al.*, 1995]. Such a task is not straightforward due to the existence of various orographic effects on precipitation, such as precipitation enhancement, rain shadowing etc [*Houze*, 2012]. Developments of precipitation climatologies for complex terrains has also been restricted due to technical difficulties connected to the installation and maintenance of dense and high quality monitoring networks . Here the results obtained by *Frei and Schär* [1998] and *Schiemann et al.* [2010] are adopted (see figure 2.1) due to the extensive rain-gauge database from the records of the meteorological services of 7 countries.

The main orographic feature of Switzerland is the Alpine mountain range (figure 2.2) that has a distinct effect on the spatial and temporal distribution of precipitation. The two morphological characteristics of the Alpine mountain range that affect the precipitation generation mechanisms are its high altitudes (up to 4810 m.asl) and its bow shaped main divide (figure 2.2). Another area with complex terrain is the North-West part of the country, where the Jura mountain range lies. Altitude in this area is much lower in comparison to the Alps but yet has a strong influence on the precipitation distribution.

The Alpine mountain range roughly divides the country into four distinct climatic zones [*Frei and Schär*, 1998; *Molnar and Burlando*, 2008; *Paschalis et al.*, 2012] (figure 2.2). In the northern part, there is the pre-Alpine area of the Swiss plateau that has distinct features on the Eastern and Western part due to the different degree of influence of the Jura mountain range. The central part of the country consists of the main mass of the Alps. Elevation spans from 500 [m.asl] at the lowest valleys and reaches 4810 [m.asl] at the Mont-Blanc. Valleys are mainly oriented in the North-South direction but several are oriented West-East (e.g. Rhone Valley). The Southern part is the pre-Alpine area around the Lake Maggiore.

Each of the regions described above is affected to a different degree by topographic features and this is reflected on the spatial and also seasonal variations of precipitation structure.

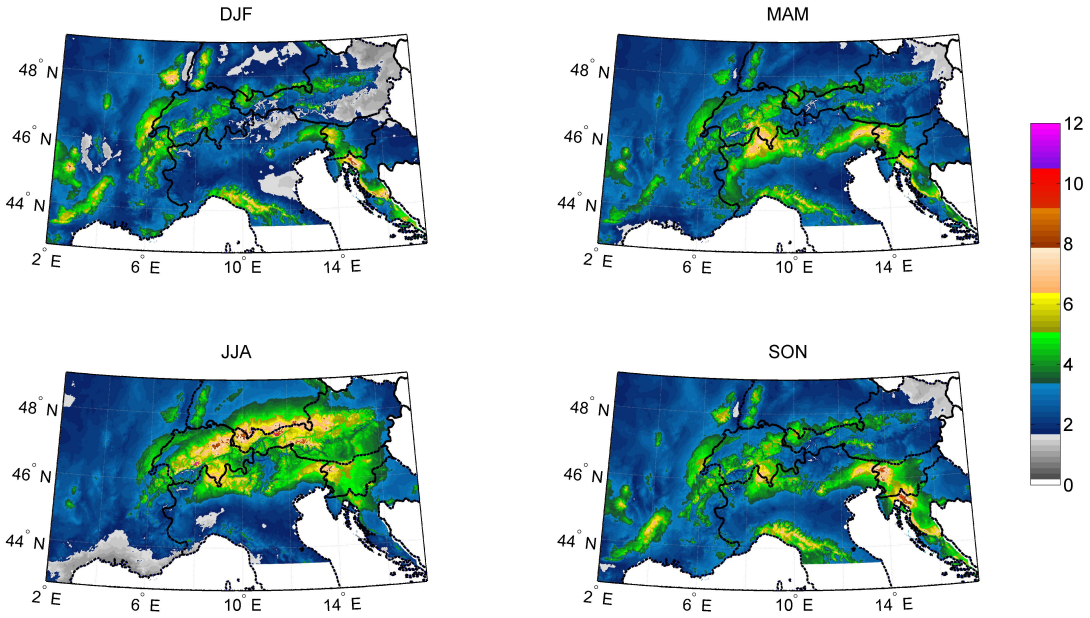


FIGURE 2.1: Precipitation Climatology of the European Alps per climatological season after *Frei and Schär* [1998].

In the Southern part, there is a strong wet anomaly located on the Lake Maggiore. The main source of moisture comes from the Mediterranean sea. In this area intense convective events with strong orographic enhancement occur during the warm seasons (spring-summer). These events lead to high precipitation accumulation depths in this area. Another feature is that rain events are rather scarce but intense. During summer the effect of the diurnal cycle is also present and intense storms usually take place in the afternoon. During winter storms are mostly stratiform, driven by frontal systems with low precipitation accumulations. Also in this area the intensity of the convective events is highly dependent on the air flow blockage mechanisms [*Houze et al.*, 2001], with high intensity events mainly triggered when the mean atmospheric flow can overcome the barrier of the high mountain range.

The Alpine area is the most complex one. Precipitation distribution is mainly affected by orographic enhancement and rain shadowing. Valleys that run eastwards and westward (e.g. Rhone valley) are very dry in comparison to the ones running north or south. A clear quantification of the effect of the various orographic-dependent mechanisms [*Roe*, 2005; *Houze*, 2012] that enhance precipitation is not easy though due to the complex terrain and the spatial sparsity of the data.

The Northern Swiss plateau has low spatial variability. The main characteristic is higher precipitation accumulations in the North-Western part, where the Jura mountains are located.

For the entire Switzerland the main rainy season is summer (JJA) and the driest is winter (DJF). The strength of the seasonality is however different between the various climatic areas [*Frei and Schär*, 1998] (figure 2.1). Seasonal differences are more intense on the Mediterranean side of the Alps.

2.2 Precipitation Monitoring Network

Precipitation data that are used throughout this thesis are mainly collected by the Swiss Federal Institute of Meteorology and Climatology (MeteoSwiss). The precipitation measuring network consists of both

point scale ground measurements (rain-gauges) and also weather radars.

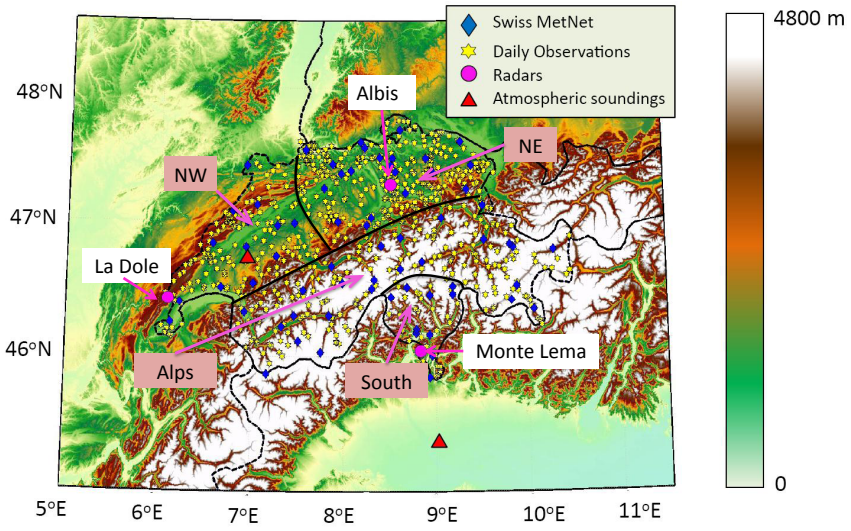


FIGURE 2.2: Precipitation measurement network of MeteoSwiss and the division of Switzerland into 4 precipitation climatological regions.

Point Measurements

The ground precipitation measurements that are used in this thesis can be divided into two different rain-gauge networks operating in different temporal resolutions. The high resolution network is named SwissMetNet [1] and consists of 68 rain-gauges. The recording mechanism is a tipping bucket and they are all heated in order to melt snow. The gauges are produced by Lambrecht (1518 H3 and 15188) [*Savina et al.*, 2012]. The sampling resolution is 10 minutes and the depth accuracy (tip volume) is 0.1 mm. The data availability covers 25 years on average. The spatial distribution of the gauges is rather uneven, since high altitude areas are under-sampled (figure 2.3). Data recorded from this network have already been extensively used in meteorological and hydrological studies [e.g. *Beuchat et al.*, 2011; *Wüest et al.*, 2010; *Schiemann et al.*, 2011, 2010; *Molnar and Burlando*, 2008; *Paschalis et al.*, 2012]. The coarse temporal resolution network consists of 679 rain-gauges in total that measure on a daily basis. The data availability is very variable and in some areas extends to more that 100 years (figure 2.3). Also in this case the high altitude under-sampling is present [*Frei and Schär*, 1998].

Measurement of precipitation using gauges is considered a robust methodology with relatively minor precision errors, especially when appropriate wind corrections are applied and the design of the network meets the recommendations of the World Meteorological Organisation (WMO). A quantification of the measurement errors for the Lambrecht tipping bucket rain-gauge can been found in *Savina et al.* [2012]. It was found that evaporation losses due to the melting of snow can be large in winter. The problem is also enhanced for low intensity events, when the lag of the tipping mechanism causes an additional delay that leads to higher evaporation and larger biases. This can be serious at the stations where snowfall is the main form of precipitation during winter. *Sevruk et al.* [2000] have also quantified the biases caused by wind effects for the Swiss high resolution gauges.

[1] http://www.meteoschweiz.admin.ch/web/en/climate/observation_systems/surface/swissmetnet.html

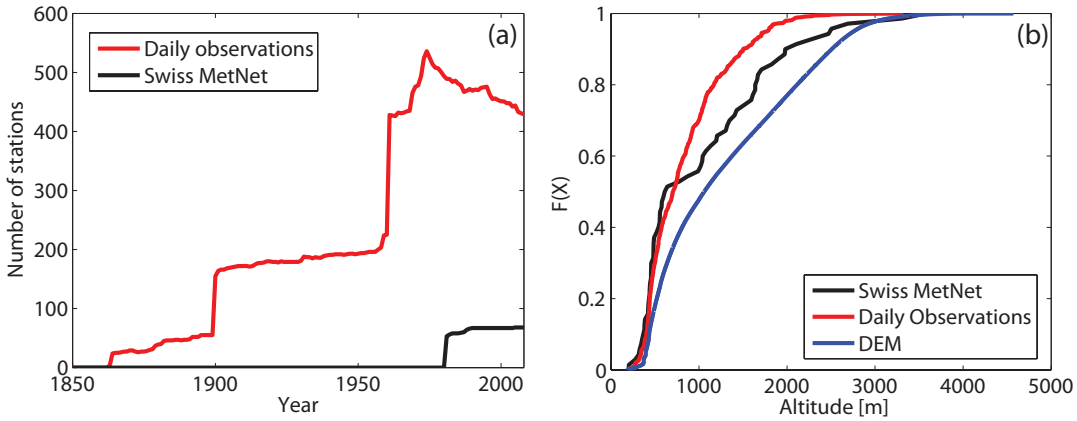


FIGURE 2.3: (a) Evolution of the number of rain-gauges in Switzerland.(b) Cumulative distribution function of the station altitude and elevation over all of Switzerland (DEM).

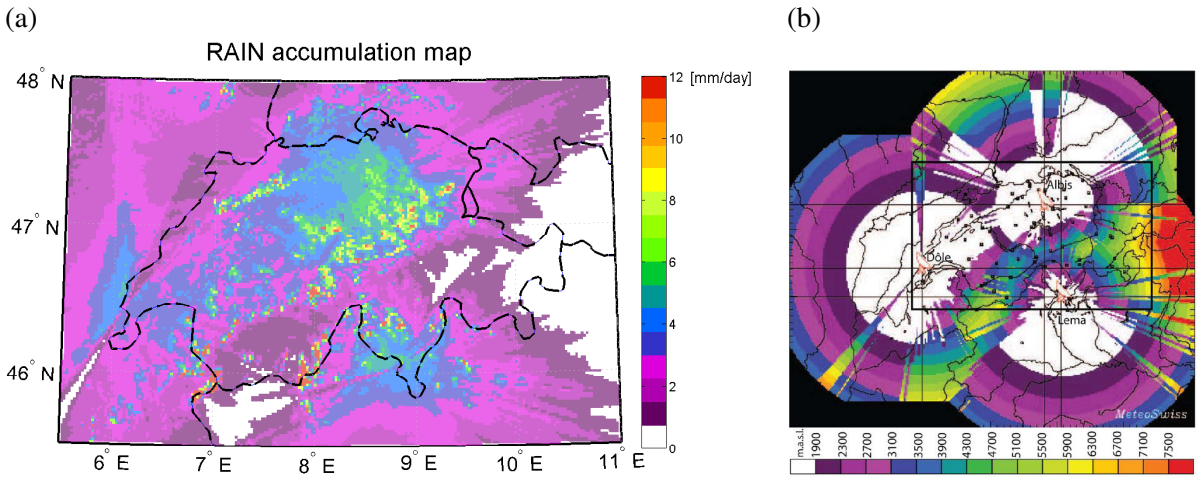
Radar Precipitation Measurements

The operational weather radar network of MeteoSwiss consists of 3 C-band single-polarization Doppler radars (5 cm wavelength) located at Albis close to Zurich, Monte-Lema in Tessin and La-Dôle close to Geneva (see figure 2.2). The final "best" ground precipitation estimate product is the composite of the measurements from all 3 radars that incorporates appropriate corrections of the Vertical Profiles of Reflectivity (VPR). It is named RAIN and is available on a 5 minute temporal resolution and 2×2 km^2 areal coverage. From April 2011, the spatial resolution of the product has increased to 1×1 km^2, but those data are not used in this thesis. Radar reflectivity factors are transformed into precipitation intensities according to a Marshall-Palmer equation [*Marshall and Palmer*, 1948] ($Z = 316R^{1.5}$, Z in [mm^{6}m^{-3}], R in [mmh^{-1}]) and are corrected for clutter contamination, signal attenuation and noise corruption. The data that are used in this thesis span from January 2004 to October 2010. Precipitation intensities are reported in a 16 value discrete scale, not uniformly distributed [*Savina*, 2011]. The data corrections that are implemented can be found in the MeteoSwiss report which is publicly available [2] and also in *Germann et al.* [2006]. Further developments on the correction algorithms in order to achieve better representation of the total precipitation amounts have been also implemented [*Gabella et al.*, 2005].

Radar measurements can serve as an excellent data-base for the investigation of the space-time structure of precipitation for high resolution scales. This becomes obvious if we compare the typical spatial density of the high resolution rain-gauge network ($\sim$ 100 km^2) to the radar estimated precipitation (4 km^2). On the other hand the uncertainty of the precipitation measurements using weather radars can be large, especially in orographically complex terrains [e.g. *Krajewski et al.*, 1996; *Germann et al.*, 2006; *Berne and Krajewski*, 2012]. A detailed description of the radar induced errors on precipitation measurement will be given in 5.2.1.

Due to the strong orography of the study area, radar coverage can be problematic due to partial beam shielding and poor visibility of the lower altitudes [*Foresti et al.*, 2012; *Germann et al.*, 2006]. As shown in figure 2.4(b) the areas of Valais (South west) and the central eastern part of the country are not captured well by the radars. Also, in figure 2.4(a) it becomes obvious that even though clutter correction algorithms are applied on the RAIN product, there are several areas that suffer from very strong clutter contamination mainly associated with atmospheric clutter caused by the anomalous propagation of the electromagnetic

[2] http://www.meteoswiss.admin.ch



Source: *Foresti et al.* [2012]

FIGURE 2.4: RAIN accumulation map and the lowest altitude of the radar sampling. On the left (a) the mean daily accumulation depths derived from the RAIN product for 2004-2010 is shown, on the right (b) the lowest altitude that is seen by the radars is shown after *Foresti et al.* [2012]

.

signal that leads to non systematic errors which in principle are very difficult to be eliminated. Another major issue with precipitation estimation with radars is that in principle they quantify the precipitable water content of the air volume above ground that does not necessarily reflect the precipitation at the ground. This problem is enhanced in areas with a complex orographic environment. The problem is illustrated in figure 2.4(a) since precipitation captured by radars cannot reproduce the distinct features of the area's climatology.

In order to tackle with the issues of not adequate representation of the precipitation in the Alpine region, several alternatives have been also explored. A recent effort to overcome shielding problems and increase the spatial resolution of precipitation measurements has been done by installing a small, cost-efficient, high frequency X-band radar in the area of Valais with very promising results *Savina* [2011], highlighting the potential of the installation of a dense X-band network across complex terrain. More MeteoSwiss has started the installation of 2 new C-band weather radars in order to solve the problems of visibility.

3

TEMPORAL STRUCTURE OF PRECIPITATION

This chapter focuses on the investigation of the temporal structure of precipitation. The main framework of the analysis is based on the concept of scale invariance. First the definitions and the basic mathematical formulations of scale invariance for stochastic processes are given. Second an investigation of the estimation procedures for assessing scale invariance is presented. Several Monte-Carlo numerical experiments are constructed in order to asses and quantify the robustness and reliability of each of the estimation procedures as a function of common rain-gauge measurement artefacts. Finally, an extensive data analysis of temporal precipitation is presented with the aim to quantify precipitation structure across scales and link it to various parameters that are known to have an effect on the generation of the precipitation processes (e.g. topography, seasonality etc.).

3.1 Introduction

The literature concerning the structure of precipitation in time is substantial since, on the point scale, relatively long and reliable records generally exist. The very first studies, were based on relatively sparse networks with coarse temporal resolutions. Their main goal was to study the basic statistical properties of temporal precipitation. A vast majority of these early studies was oriented towards the construction of global precipitation climatologies [e.g. see *Frei and Schär*, 1998, and references therein].

Probability theory and statistics, provided the statistical theories that describe the structural features of various natural processes. Some of these theories describe precipitation as point processes or Markov chains. However their applicability been limited due to their inability to capture the multi-scale statistical structure of the process. This major concern has led to the search for an appropriate theory that can describe precipitation structure for the entire range of temporal scales of hydrological interest.

Possibly the greatest boost was given through the theories of scale invariant processes that provided the formulations which link statistics of the involved processes across scales. The fact that these theories originally found great application in the study of turbulence also inspired further research on their applicability to natural processes related to atmospheric turbulence and of course precipitation. After the introduction of the concept of scale invariance and multifractals in hydrology, more and more studies have been identifying such scaling relationships in precipitation temporal records [e.g. *Molnar and Burlando*, 2008; *De Lima and Grasman*, 1999; *Svensson et al.*, 1996; *Venugopal et al.*, 2006; *Veneziano and Lepore*,

2012]. Generally the concept of scale invariance became common ground for the analysis of precipitation and this opened new horizons for several engineering related topics, such as the estimation of extremes [e.g. *Veneziano et al.*, 2009, 2006b; *Langousis and Veneziano*, 2007].

In addition, several studies investigated potential links of the scaling behaviour of precipitation and several climatic factors [*Purdy et al.*, 2001; *Harris et al.*, 1996; *Molnar and Burlando*, 2008]. This issue is still an open question, since research on this topic has yielded up to now, interesting yet sometimes controversial results.

On the other hand, there is also strong recent evidence that systematic deviations from scale invariance exist for temporal precipitation [e.g. *Veneziano et al.*, 2006a; *Cârsteanu and Foufoula-Georgiou*, 1996; *Rupp et al.*, 2009; *Serinaldi*, 2010a; *Fraedrich and Larnder*, 1993; *Paschalis et al.*, 2012]. The major problems of the applicability of the theory of scale invariance concern its limited range of temporal scales where the representation of the process as scale invariant holds [*Fabry*, 1996; *Fraedrich and Larnder*, 1993; *Marani*, 2003] and demonstrations of deviations from "perfect scaling" [*Veneziano et al.*, 2006a; *Paschalis et al.*, 2012; *Cârsteanu and Foufoula-Georgiou*, 1996; *Rupp et al.*, 2009]. This problem, even though important, has not been adequately addressed so far, and thus is one of the focus points of this chapter.

3.2 Scale Invariant Stochastic Processes

The main features of precipitation are its intermittent nature, its evolving structure across different aggregation scales and typically its heavy tailed and positively skewed distributions. The term intermittency describes the property of precipitation of containing zero values. The term ***intermittency*** has been commonly used in fluid mechanics as the property of sudden bursts of local velocities of fluids in fully developed turbulence [e.g. *Meneveau et al.*, 1990; *Benzi et al.*, 1984; *Schertzer and Lovejoy*, 1987; *Frisch*, 1991]. In order to avoid confusion, throughout this thesis intermittency reflects ***only*** the property of a non-negative process to have a probability mass at zero. This notation of intermittency is also the most commonly used in hydrology [*Molini et al.*, 2001; *Schleiss et al.*, 2011; *Kundu and Siddani*, 2011].

One of the most fruitful theories that attempt to describe in a unified and parsimonious way the properties of precipitation, is the theory of (stochastic) scale invariance. The concept of scale invariance originates from geometry, and loosely speaking, describes geometrical objects that are invariant under scale-change transformations. Those geometrical objects are named "fractals" due to the seminal work of *Mandelbrot* [1983]. The terms reflects their property of having fractional dimensions in contrast to Euclidean geometrical constructions that only have integer dimensions. Various terms for describing scale-invariance can be found in the literature, such as self-similarity, fractality, etc.

The deterministic framework of self-similarity has also been extended for stochastic processes. Generally speaking, this describes processes whose probability distributions remain invariant under scale change transforms. The first to define such a process was *Kolmogorov* [1940]. *Hurst* [1951] was also the first who identified self similar behaviour in geophysical signals while working on the design of the Aswan dam.

The self-similarity principle for stochastic processes has been extended to the generalized framework of stochastic self-similarity, or "multi-fractality" as it is widely known [e.g. *Schertzer and Lovejoy*, 1987; *Veneziano et al.*, 2009]. This principle has been proven to be applicable in many fields spanning from geophysical processes such as precipitation, river discharge, temperature etc. [e.g. *De Lima and Gras-*

man, 1999; *Deidda et al.*, 1999; *Molini et al.*, 2009; *Koutsoyiannis et al.*, 2011] to internet traffic [*Riedi et al.*, 1999], stock market prices [*Lo*, 1989], sunspot activity [*Movahed et al.*, 2006] and even heart beat dynamics [*Stanley et al.*, 1999].

3.2.1 Mathematical Formulation of Stochastic Self Similarity

The mathematical formulation of self similarity has followed various formalisms in the literature, which can lead to some confusion. Throughout this thesis we adopt the notation developed by *Over* [1995] and *Veneziano and Langousis* [2010]. The definitions will be given for generalized processes defined in one dimension. Generalizations for higher dimensions are straightforward and applicable for two or three dimensional processes as will be discussed in chapter 5 for the analysis of spatial precipitation.

A generalized process can de defined as [*Veneziano and Langousis*, 2010]:

$$X(h) = \int_{-\infty}^{\infty} h(t)X(t)dt, \tag{3.1}$$

where $X(t)$ is a stochastic process and $h(t)$ is a convolution kernel, or test function. The test function $h(t)$ is defined in a class Θ has to be closed under scaling, i.e. if

$$h(t) \in \Theta, \text{ then, } h_r(t) = rh(rt) \in \Theta, \text{ for } 0 < r < \infty. \tag{3.2}$$

A generalized process is said to be stochastic self-similar or multifractal under contraction if there exists a set of random variables A_r such that [*Veneziano and Lepore*, 2012]

$$X(h_r) \overset{d}{=} A_r X(h), \quad \text{for any } r \geq 1, h \in \Theta, \tag{3.3}$$

where $\overset{d}{=}$ stand for equality of the probability distributions. For $r < 1$ self similarity under dilation is defined.

Various different names have been used in the literature in order to describe stochastic self-similar processes such as multi-scaling [*Menabde et al.*, 1997a] or multi-affine processes [*Benzi et al.*, 1993], processes with anomalous scaling laws [*Deidda et al.*, 1999] and multifractals [*Verrier et al.*, 2010; *Svensson et al.*, 1996; *Olsson*, 1996], which even though it is the author's opinion not the most enlightening term, has been the most extensively used.

Veneziano and Langousis [2010] define also several sub-classes of the above formalism that have been extensively used and are briefly mentioned here. In the case where A_r are not random variables but a deterministic contraction factor (say α_r), it can be proven by successively applying equation 3.3 [*Waymire*, 1985; *Gupta and Waymire*, 1990] that the only possible form of α_r is a power law

$$\alpha_r = r^H. \tag{3.4}$$

This form of deterministic scaling is usually referred in the literature as simple scaling or self-similarity [e.g. *Koutsoyiannis*, 2009; *Klemeš*, 1974]. The scaling exponent H typically inherits the name from H.E. Hurst [*Hurst*, 1951].

Also another interesting class which can be defined as a subclass of the definition in 3.3 is when the convolution kernel $h(t)$ is a function whose j-th moments vanish (i.e $\mu_k = \int_{-\infty}^{\infty} t^k h(t)dt = 0,$ for $k = 0, 1, ..., j$). A class of functions that satisfy this property are the wavelets of order j. This has given rise to alternative multifractal formalisms dependent on the wavelet decomposition of the processes [e.g. *Hwang and Mallat*, 1993; *Davis et al.*, 1994; *Venugopal et al.*, 2006; *Arneodo et al.*, 1988].

It can be shown that for the general case, where A_r is a random variable, it can be expressed as [*Over*, 1995]

$$A_r = r^{\mu} \exp(Z_{-\log r}), \quad r \leq 1, \tag{3.5}$$

where μ is an arbitrary real number and $Z_{-\log r}$ is a process with stationary increments (for any scale r). A similar connection between A_r and a stochastic process with stationary increments can be found in *Veneziano* [1999] who generalized the work of *Lamperti* [1962], which is the first study that provides the basic linkages between self-similarity (simple scaling) and stationary and limit processes.

Stochastic self similar (multifractal) processes are often reported as multiplicative in contrast to the simple scaling processes that are described as additive. The reason is that multifractal processes can be defined as the limit product of independent and identically distributed (iid) random variables. This comes from the fact that since processes with stationary increments can be expressed as sums of identically distributed random variables, the exponentiation of equation 3.5 would yield [*Over*, 1995]:

$$X(h) = X(1) \prod_i W_i, \tag{3.6}$$

where W are iid random variables.

It follows that each multifractal process can be described from the distribution of the random variable A_r. The moments of A_r can be expressed as [*Veneziano and Furcolo*, 2009]

$$K(q) = \log_r E[A_r^q]. \tag{3.7}$$

The function $K(q)$ is non-linear, and as proven by *Gupta and Waymire* [1990] [1] has to be convex.

3.2.2 Multiplicative Random Cascades (MRC)

Definition of the MRC

As shown in 3.6, a multifractal process can be expressed as a limit product of iid random variables. One of the most commonly used discrete approximation of the construction of multifractal processes is the

[1] The notation in *Gupta and Waymire* [1990] is different. The notation of $K(q)$ has been introduced by *Schertzer and Lovejoy* [1987].

one of the discrete multiplicative random cascades (MRC) (see figure 3.1 for the schematic representation of MRC in one dimension).

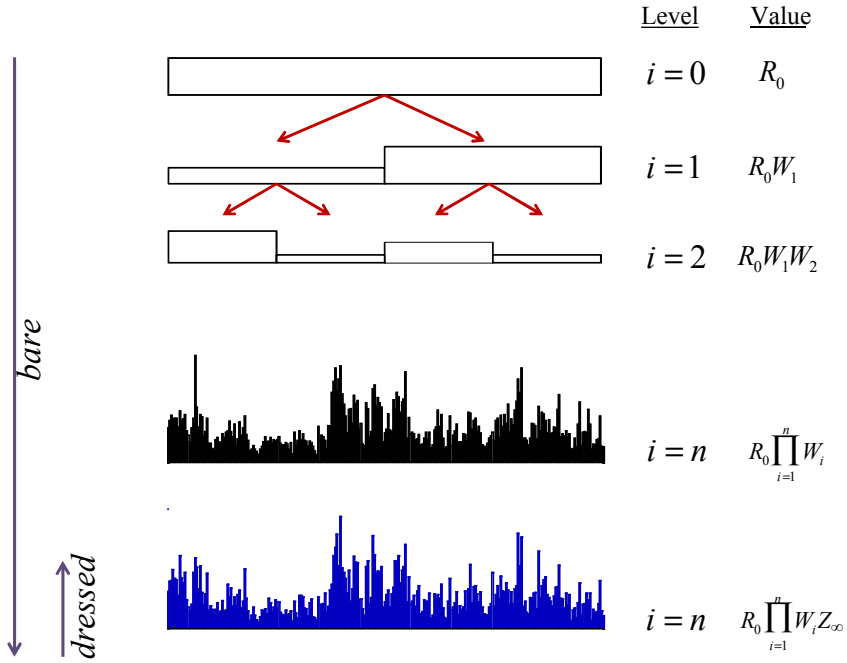


FIGURE 3.1: Schematic representation of the discrete multiplicative random cascade model for branching number $b = 2$.

The idea originates from the concept of energy cascades of the kinematic energy in fully developed turbulence. It reflects on the conceptual phenomenology of energy propagation across scales. A MRC construction can be expressed as a recursive multiplication of a measure defined in a previous scale with an iid random variable W which is called the cascade generator [e.g. *Over*, 1995; *Molnar and Burlando*, 2005]. In order to proceed to the finer scale, the measure is divided into b equally spaced segments in the D-dimensional space. $b \in \mathbb{N}$ is referred as the branching number of the MRC.

A multifractal measure, which is a special case of the generalized process described in equation 3.1 can be defined for a specific scale l_{max} as:

$$\mu_n(\Delta_n^i) = R_0 l_{max}^D b^{-n} \prod_{j=1}^{n} W_j, \tag{3.8}$$

where D is the Euclidean dimension on which the measure is embedded (e.g 1 for time series, 2 for planar fields etc.), R_0 the initial "mass" that is redistributed across scales, n is the cascade development level, b the branching number and W the cascade generator. In order to have mass conservation on average, W should have $E[W] = 1$, where E stands for the expected value ($E[X] = \int_{-\infty}^{\infty} x f(x)\, dx$). The product $l_{max}^D b^{-n}$ guarantees that the aggregated mass over any scale is conserved on average.

This case corresponds to the so-called "bare" construction of the MRC [*Schertzer and Lovejoy*, 1987; *Veneziano et al.*, 2006b; *Veneziano and Langousis*, 2010; *Veneziano et al.*, 2006a]. A very interesting property of the limit behaviour of a measure defined as a discrete MRC can be found in *Holley and*

Waymire [1992]

$$\mu_\infty(\Delta_n^i) = \mu_n(\Delta_n^i) Z_\infty(\Delta_n^i), \quad i = 1, 2, \cdots, b^{-n}. \tag{3.9}$$

The random variable $Z_\infty(\Delta_n^i)$ is independent of $\mu_n(\Delta_n^i)$ and represents the sub-grid variability of the process. This random variable is also known in the literature as the dressing factor of the cascade [*Langousis et al.*, 2009]. Analytical expressions of the dressing factors are generally not straightforward. For discrete MRC and specific distributions of the cascade generator, analytical expressions or numerical approximations of Z_∞ have been calculated by *Veneziano and Furcolo* [2003] and *Veneziano et al.* [2006b]. The property of the dressing factors are of major importance and should be taken into account in simulation since, if a process defined at a specific scale is assumed to arise from a MRC, then the simulation should take into account the distribution of the limit measure, and not just apply the bare MRC construction (see figure 3.1). Due to the non-existence of analytical solutions of the probability distribution of Z_∞, it is commonly approximated numerically by developing the discrete MRC for several more steps, and finally aggregating back to the desired scale. Furthermore, the distribution of Z_∞ can have also serious effects on the distribution of the dressed measure for each scale with specific emphasis on its extremes [*Langousis et al.*, 2009].

The classical notation of the MRC, has a serious disadvantage when dealing with precipitation. It completely neglects the intermittent nature of the process. To the best of my knowledge the first effort to inherently add a probability mass at zero at the construction of the MRC was by *Over and Gupta* [1994]. The idea is that the cascade generator W has an "atom" at zero, which means that $P(W = 0) > 0$. The distribution of W has thus the general form

$$P(W = 0) = 1 - b^{-\beta} \quad \text{and} \quad P(W > 0) = b^{-\beta}. \tag{3.10}$$

In this case in order to achieve mass conservation on average, the expected value of the positive part of the generator W^+ should be

$$E[W^+] = b^\beta. \tag{3.11}$$

The cascade generator can be more conveniently expressed as the product of two independent random variables

$$W = BY, \tag{3.12}$$

where B is a two state process with:

$$P(B = 0) = 1 - b^\beta, \quad \text{and} \quad P(B = b^\beta) = b^{-\beta}, \tag{3.13}$$

and Y is a strictly positive random variable following some probability distribution function and $E[Y] = 1$.

In principle this cascade construction is a multiplication of two separate MRCs, a binary cascade often called the beta model [*Schmitt et al.*, 1998], and a continuous MRC. There are also other approaches to describe the intermittency in precipitation, such as thresholding the output of a MRC generation and setting all values below a threshold to zero. This approach is inferior in the sense that no analytical results can be obtained for its statistical properties and thus throughout this thesis the formalisms developed by *Over* [1995] and *Over and Gupta* [1996] will be employed. More a unified description of the process statistical properties is generally more desirable that introducing subjective threshold decisions.

Adopting the methodology described above (MRC with atom at zero), intermittency is introduced to the simulation scheme and it can be shown that the probability of zero is a power law function of the scale. This contradicts the results of *Koutsoyiannis* [2006] that showed that entropy maximization under specific constraints, leads to a non-power law probability of zero across scales. This can be a rather significant problem of the MRC both as an analysis and simulation tool, if a non-power law relationship for the probability of rain is true. This is also the reason why the MRC cascades with an atom at zero have been found not to be very good in describing the distributions of dry and wet spell durations [*Schmitt et al.*, 1998]. It will be later though shown that for a limited scale range the assumption of the power law is in fact a very good approximation.

Discrete MRC cascades are not the only tool for the simulation of multifractal measures. Indeed, even though their simplicity is one of their main appealing characteristics, several problems have been identified. First of all, their scale invariance holds true only on their b-branch tree construction. This issue has been named as the "on-off" grid cascade issue [*Over*, 1995]. More, discrete MRC cascades produce non-stationary measures [*Lombardo et al.*, 2012; *Veneziano and Langousis*, 2010]. In order to overcome this issue, *Chainais et al.* [2007] developed continuous and quasi-continuous multifractal random measures. Also the most widely applicable methodology in order to densify the cascade construction, can be achieved in the frequency domain. This methodology has been proposed by *Schertzer and Lovejoy* [1987]. The MRC of the log-stable type in the frequency domain has been referred as "continuous universal" multifractals and the numerical simulation procedure can be found in *Pecknold et al.* [1993]. Briefly summarizing the steps:

1. A D-dimensional field of an iid alpha stable white noise is sampled $\sim S_\alpha(\sigma, -1, 0)$ [*Samorodnitsky and Taqqu*, 1994].

2. Then its Fourier transform is obtained and each component is multiplied by $|k|^{-D/\alpha'}$, where k is the wavenumber and $1/\alpha + 1/\alpha' = 1$.

3. The multifractal field is obtaining as the inverse Fourier transform of the previous field.

This construction can be further expanded to non-stationary (referred also as non conservative) cascade by a further fractional integration [e.g. *Parke*, 1999] of order H. Fractional integration of fields defined on a regular D-dimensional grid can be achieved also in the frequency domain.

Unfortunately those simulation methodologies are not able to generate signals with probability mass at zero values. When analysing or simulating precipitation, this problem is of major importance and that is the reason why in the main body of this thesis only discrete MRC are considered.

MRC Properties

One of the basic properties of the MRC has been already mentioned in equation 3.7 and concerns the moments of the MRC generator W. *Kahane and Peyriere* [1976] following *Mandelbrot* [1974] defined that, for the discrete cascades,

$$\chi_b(q) = \log_b E[W^q] - (q-1). \tag{3.14}$$

This function has been abbreviated as MKP (*Mandelbrot* - *Kahane* - *Peyriere*) from *Over* [1995]; *Over and Gupta* [1994]; *Gupta and Waymire* [1993]; *Molnar and Burlando* [2005]. Its analytical form depends on the probability distribution of the MRC generator W. The MKP function can also be found in the literature with several similar notations and names [e.g. see notations in *Chainais et al.*, 2005; *Ossiander and Waymire*, 2002]. In this thesis the name MKP function is used.

Mandelbrot [1974] defined three different classes of the MKP function dependent on the existence/divergence of the moments of Z_∞. The classes are namely, the regular, the irregular and the degenerate class. *Kahane and Peyriere* [1976] provide the following results for properties of the limit measure and Z_∞.

If $\chi_b(q)$ is the MKP function of a MRC with cascade generator W then [*Over*, 1995]:

1. If $\chi^{(1)}(1) < 0$, then $E[Z_\infty] > 0$ and $E[Z_\infty] = 1$.
2. For $q > 1$, $\mu_\infty(J)$ has finite moment of order q for $q < q_c$, where q_c is the critical order defined as $q_c = \inf\{q \geq 1; \chi_b(q) \geq 0\}$. [2]
3. If W is bounded by b and $P(W = b) < 1/b$ then $E[Z_\infty^q]$ exists for all $q > 0$.
4. If $E[Z_\infty \log Z_\infty] < \infty$ then $\mu_\infty(J)$ is almost surely supported by a Borel subset of J of Hausdorff dimension $D = -d\chi_b^{(1)}(1)$.

A definition of the Haussdorf dimension is beyond the scope of this thesis, but a complete definition can be found in *Falconer* [1990].

Another property of the random variable that serves as the cascade generator W is that it should belong to the class of the log-infinitely divisible distributions [*Veneziano and Furcolo*, 2009; *Waymire and Williams*, 1996]. Common infinitely divisible distributions are the normal, the Poisson and the general class of Levy alpha stable distributions [*Samorodnitsky and Taqqu*, 1994]. Those distributions have been used in the literature for multifractal process simulation [e.g. *Deidda*, 2000; *Pathirana et al.*, 2003b; *Over and Gupta*, 1996]. The Levy stable distributions especially have been proven an excellent multifractal generator and due to the study of *Schertzer and Lovejoy* [1987] usually referred to as "Universal Multifractals". In table 3.1 there is a brief summary of some of the most widely known multifractal generators and their respective MKP-functions

Another interesting property of multifractal stationary measures concerns its power spectrum. If a stationary process is a multifractal measure then its power spectral density has a power-law decay [*Veneziano*

[2] $\inf(S)$ refers to the Infimum (i.e. largest lower bound) of a set S

[3] *Samorodnitsky and Taqqu* [1994]

TABLE 3.1: MRC models with an atom at zero

Name	Distribution	MKP-function
Log-Poisson	$Y = b^{\gamma+\alpha X_\lambda}$ $\Pr(X_\lambda = k) = \frac{\lambda^k e^{-\lambda}}{k!}$	$\chi_b(q) = (\beta-1)(q-1) + \gamma q + \frac{\lambda(b^{\alpha q}-1)}{\log b}$
Log-Normal	$Y = b^{\gamma+\sigma X}$ $X_\lambda \sim N(0,1)$	$\chi_b(q) = (\beta-1)(q-1) + \frac{\sigma^2 \log b}{2}(q^2-q)$
Log-Levy	$Y = \exp[S_\alpha(C_{ls},-1,1)]$ $S_\alpha(C_{ls},-1,1)$, Levy-stable [3]	$\chi_b(q) = (\beta-1)(q-1) + \frac{C_{ls}}{\alpha-1}(q^\alpha - q)$

and Langousis, 2010].

$$S(k) \propto |k|^\nu = |k|^{-D+1-\chi_b(2)}, \tag{3.15}$$

where k corresponds to the wavenumber and D to the Euclidean dimension. This equation is valid only for non intermittent MRC (i.e. $\beta = 0$), but a power law was also found to be a good approximation for the ones that posses a probability mass at zero, especially in the high frequency regime (see 3.2).

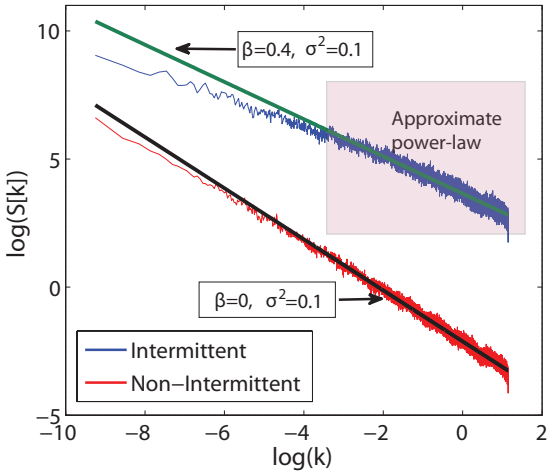


FIGURE 3.2: Power law decay of the power spectral densities for MRC. An exact power law is valid for the non intermittent MRC and an approximate power law, with a very good approximation for the intermittent MRC. The series have been generated with a discrete version of the beta-lognormal MRC and have a mean value of 1.

In fact, an analytical expression can be obtained for the multifractal processes with an atom at zero, if both the expressions of the spectral densities of the binary "beta-model" [*Over*, 1995; *Schmitt et al.*, 1998] and the non-intermittent multifractal [*Veneziano et al.*, 2006a] are known. Since the MRC with the atom at zero is a multiplication of such mutually independent processes, their spectrum can be found using equation 3.39. Since both the beta model and the lognormal MRC model have power law spectral decays [*Schmitt et al.*, 1998; *Veneziano et al.*, 2006a], it can be proven that the resulting beta-lognormal MRC

doesn't. In any case though, a power law approximation also for the intermittent cascades was adequate for the data analysis purposes of this study.

3.3 Scaling Parameter Estimation

One of the most crucial issues, either when dealing with modelling, or data analysis, is the parameter estimation of the MKP function. The most common procedure in the literature is by analysing the scaling behaviour of the moments.

First let us define the statistical moments of a process as

$$M_n(q) = \sum_{i=1}^{b^n} \mu_\infty^q(\Delta_n^i), \tag{3.16}$$

and then the scaling of the *ensemble* statistical moments as the small scale limit

$$\lim_{\lambda \to 0} \frac{\log EM_n(q)}{-\log \lambda} = \lim_{n \to \infty} \frac{\log_b EM_n(q)}{n/D}. \tag{3.17}$$

It can be shown [*Over*, 1995] that this limit converges to the MKP function.

$$\lim_{n \to \infty} \frac{\log_b EM_n(q)}{n/D} = D\chi_b(q). \tag{3.18}$$

In practice, an ensemble average of the moments is impossible since the signal for which parameter estimation is needed for the MKP function is a single realization of the process. For this case let us define the scaling of the moments as:

$$\tau(q) = \lim_{n \to \infty} \frac{\log_b M_n(q)}{n/D}. \tag{3.19}$$

The main question is under which conditions does the function $\tau(q)$ converges to the ensemble moment scaling (i.e. ergodicity holds true). *Holley and Waymire* [1992] provide those conditions. As stated by *Over* [1995], the theorem of *Holley and Waymire* [1992] has been developed for strongly bounded cascades, but a generalization is also possible since the boundedness does not pose any restriction to the proof of the theorem. Here only the final results are shown.

For a discrete MRC with cascade generator W, if

$$\frac{EW^{2q}}{E^2W^q} < b, \tag{3.20}$$

and

$$EZ_{\infty}^{2q} < \infty, \tag{3.21}$$

then

$$\tau(q) = D\chi_b(q). \tag{3.22}$$

Also

$$\lim_{n\to\infty} \frac{M_n(q)}{EM_n(q)} = Y(q), \tag{3.23}$$

where $Y(q)$ is a random variable for each order q. This means that the even though the moments do not converge to their expected values, but when the ensemble numbers goes to infinity, the scaling of the moments for the range of moments where the above conditions are valid, converges to the MKP function. Those results show then the limits of the procedure of the estimation of the parameters of the MKP function from a single realization. The maximum order for which $\tau(q)$ explains the MKP function will be stated hereafter as q^*.

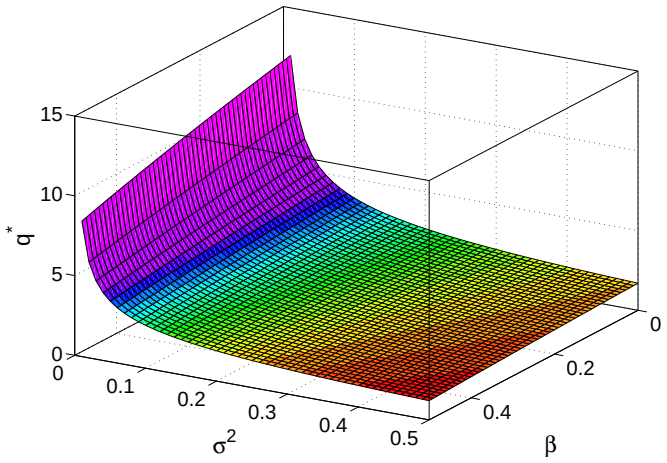


FIGURE 3.3: Maximum order for which the moment scaling function converges to the true MKP function. The example corresponds to the beta-lognormal model with branching number $b = 2$.

These convergence criteria have been commonly neglected in the literature, which can lead to highly biased results. The usual procedure in data analysis is that an upper threshold for the orders is chosen arbitrarily. For an illustrative example in figure 3.3 the maximum order of moments is plotted for the beta-lognormal model. It should be noted that data analysis studies commonly analyse moments up to the order of 5 or more. In figure 3.4 the behaviour of $\tau(q)$ is shown in comparison to the theoretical $\chi_b(q)$ function for several cases of the beta-lognormal model. $\tau(q)$ typically has an almost linear behaviour for the order beyond the upper level estimated by *Holley and Waymire* [1992] [e.g. *Veneziano and Furcolo*, 2009; *Lashermes et al.*, 2004; *Ossiander and Waymire*, 2002].

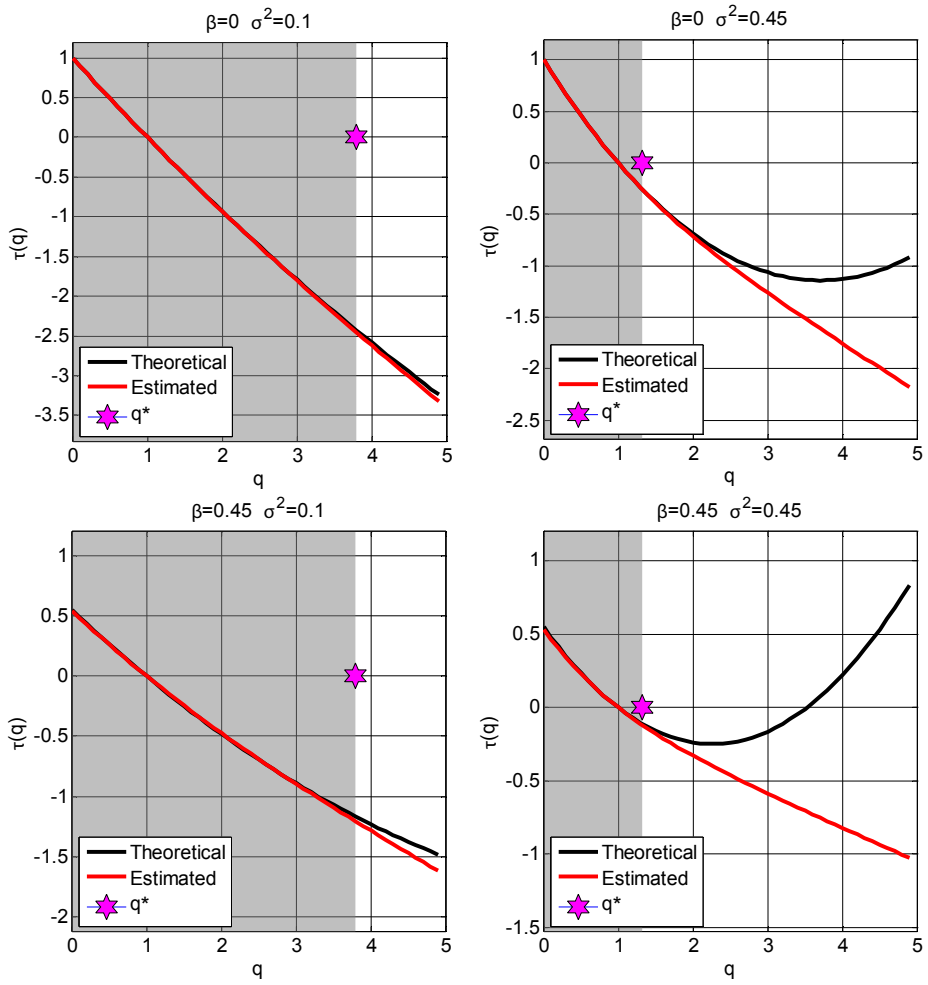


FIGURE 3.4: Examples of the $\tau(q)$ function and their respective order q^* for which $\tau(q)$ converges to the $\chi_b(q)$ function. Red lines correspond to the $\tau(q)$ function and black lines to the theoretical MKP function. The range of the convergence is marked with a grey shadow. The maximum order of q is shown with a star. The examples correspond to the beta-lognormal MRC model with a branching number $b = 2$.

3.3.1 Scale Invariance Estimators

As shown in the previous section, there are several ways to assess scale-invariance from geophysical signals. The most common procedure is from the analysis of the moment scaling as described in 3.2.2. One of the main concerns is the bias and robustness of the estimation procedure. In order to estimate the $\tau(q)$ function the moments M_n^q for several aggregation intervals (scales) in the range of orders q that $\tau(q)$ converges to $D\chi_b(q)$ are estimated. Afterwards, least square fitting on the logarithms of M_n^q versus the logarithms of aggregation scale is used to estimate the scaling exponents per order q. Since the logarithm is a concave function, it can be shown [*Veneziano and Furcolo*, 2009] that the estimation of the $\tau(q)$ function according to this procedure is biased.

The estimation of the moments $M_n(q)$ can be done either for local averages for each scale [e.g. *Molnar and Burlando*, 2008; *Harris et al.*, 1996; *Menabde et al.*, 1997b] or for wavelet coefficients for the respective scale [e.g. *Abry and Veitch*, 1998; *Turiel et al.*, 2006; *Venugopal et al.*, 2006]. In principle the first case

can be considered as a specific sub-case since local averages may be considered as zero-order wavelets [*Veneziano and Furcolo*, 2009]. One other procedure developed by *Lavallée et al.* [1991] and *Tessier et al.* [1993] is called the double trace moment technique and is a variant of the simple moment scaling analysis. This procedure has been proven to have some theoretical consistency issues [*Veneziano and Furcolo*, 1999], since it only applies to bare quantities, it is not appropriate for data analysis and thus not used here.

An analytical solution of the bias introduced due to the logarithmic conversion of the moments, does not exist. Here a numerical quantification is conducted. For this reason a numerical Monte-Carlo experiment is constructed. Several realizations are drawn from the beta-lognormal model and their parameters (β, σ^2) are estimated using classic moment scaling analysis. Then the expected value of the bias of each of the parameters is estimated, and the robustness of the procedure is quantified by the standard deviation of the estimates. The beta-lognormal model that will be extensively used in the entire thesis, has proven to be excellent for rainfall [*Veneziano and Lepore*, 2012] and since it is also the most parsimonious one, it is the one chosen in the examples in this thesis.

As shown in table 3.1, the MKP function of the beta lognormal model is

$$\chi_b(q) = (\beta - 1)(q - 1) + \frac{\sigma^2 \log b}{2}(q^2 - q). \tag{3.24}$$

For rainfall, the structure of intermittency can be quantified with only one parameter β and within-storm variability with the parameter σ^2 which controls the distribution of the positive part of the MRC generator. The quantification will be done only by estimating the moments $M_n(q)$ for local averages, since it is the most common procedure in data analysis. The range of moment orders for which the estimation of $\tau(q)$ is conducted is $[0\ q_{max}]$ where $q_{max} = \min(3, q^*)$. Negative order moments are not defined for MRC with an atom at zero.

For the beta-lognormal model we have

$$q^* = \frac{1 - \beta}{\sigma^2 \log b} \quad \text{, for } q_c \leq 2, \tag{3.25}$$

$$q^* = \sqrt{\frac{1 - \beta}{\sigma^2 \log b}} \quad \text{, for } q_c > 2, \tag{3.26}$$

where $q_c = 2(1 - \beta)/(\sigma^2 \log b)$ is the critical order of moment divergence.

The generated time-series from the MRC cascade are "partially" dressed in the sense that the limit measure Z_∞ is only approximated by generating the bare cascade for $n_{bare} = n + 5$ development steps and then aggregating back to the desired scale n. The quantification is done for $n = [8 \quad 14]$ for branching number $b = 2$ that corresponds to simulation length $l = [256 \quad 16384]$. Parameter estimation for the parameters β and σ^2 can be done in several ways. The one adopted here is as follows. The β parameter is estimated from the intersect of the $\tau(q)$ function at zero

$$\beta = 1 - \tau(0). \tag{3.27}$$

The estimation of σ^2 is done with ordinary least square fitting to the MKP function defined in 3.24. Several other methodologies exist, such as estimating the parameters from first order derivatives of $\chi_b(q)$ [*Molnar and Burlando*, 2008; *Over*, 1995]. However the differences between the methodologies are minor and thus only one methodology is shown here.

The figures (3.5 and 3.6), summarize the main results for β and σ^2. Only some of the cases are plotted for illustrative reasons.

First of all, since the estimation of the β parameter is based in principle only on the binary process $[0-1]$ (e.g. rain, no-rain), it is unbiased. It's robustness is quantified by the standard deviation of the estimated values of β. In figure 3.6 as expected, it is shown that the longer the time-series are, the most robust the estimation is. This result is trivial but here an exact numerical quantification is achieved. The distribution of the positive part of the generator Y has no effect, since the estimation of β is only affected by the occurrence process. Even for very long time-series, the standard deviation of $\hat{\beta}$ can reach 20% of the true value of the simulated β. This can cause serious robustness issues of the estimation of β especially when short series are used.

As far as the σ^2 parameter is concerned, when ordinary least square fitting is used on the logarithms of the moments $M_n(q)$ the estimation is generally negatively biased (figure 3.5) which means consistent underestimation of the parameter. The bias is independent of the sample size and enhances with increasing σ^2, which means that the bias issues are higher when the signals are "rougher". The average value of the bias can reach values up to roughly 20% of the simulated parameter. In order to illustrate the significance of such a bias for the extremes of the process, in figure 3.7 the exceedance probabilities of a non-intermittent lognormal cascade with $E[X] = 1$ are shown for 3 different values of σ^2. It can be clearly seen that the importance of a 20% underestimation of of the σ^2 parameter can lead to significant errors concerning the extremes of precipitation and thus yield serious problems in risk analysis.

An important result concerning the analysis conducted so far for the estimation of self similarity parameters can be derived by comparing figures 3.8 and 3.5. The results for figure 3.8 are obtained with the exact same methodology, but the upper order q_{max} has been arbitrarily fixed to 3. The bias of the estimation procedure increases dramatically, and if the selected upper order is larger, the results are even worse.

An other alarming result is the very low robustness of the estimation of σ^2 (see figure 3.6). The standard

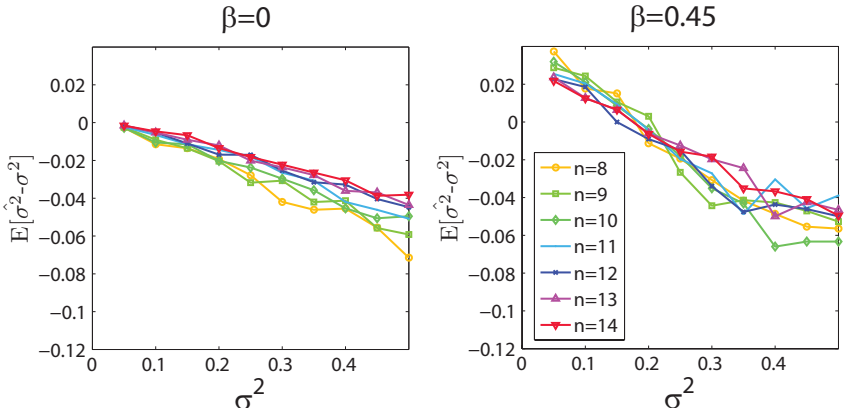


FIGURE 3.5: Mean value of the estimated bias $\overline{\hat{\sigma^2} - \sigma^2}$ due to the logarithmic conversion for the parameter estimation. Lines with different colours correspond to different cascade development steps n according to the legend. Left panel corresponds to non-intermittent series and right panel to highly intermittent ones.

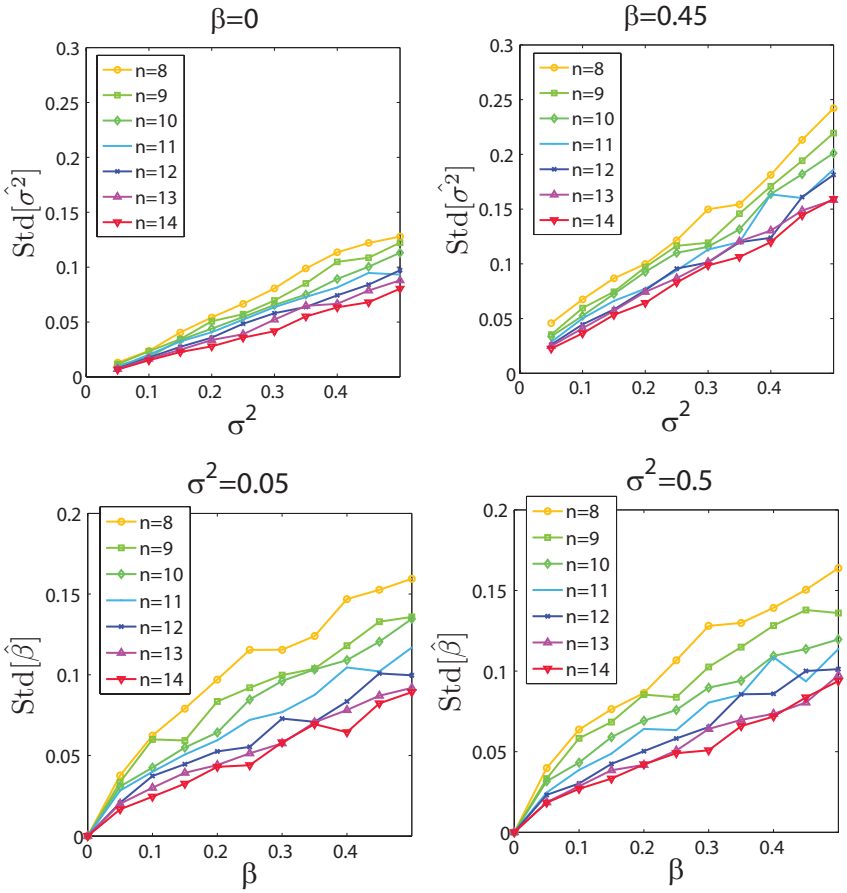


FIGURE 3.6: Standard deviation of the estimated $\hat{\sigma^2}$ and $\hat{\beta^2}$ parameters of the beta-lognormal MRC. Lines with different colours correspond to different cascade development steps n according to the legend. The upper panel corresponds to the σ^2 parameter and the lower one to the β parameter

deviation of the estimate increases with decreasing sample size, and also with increasing number of zeros in the record (see figure 3.6 right panel). In some cases, the uncertainty can reach up to 50% of the parameter value. This has to be taken seriously into account when dealing both with data analysis and simulation. This issue can be of major importance in the study of precipitation extremes, since a 50% deviation of the σ^2 can significantly impact extremes as has been shown.

Several studies have been focused on the way how to overcome the issue of low parameter estimation robustness. Most of them deal with the performance of wavelet based estimators [e.g. *Serrano and Figliola*, 2009; *Muzy et al.*, 1993; *Venugopal et al.*, 2006; *Kantelhardt et al.*, 2002]. Unfortunately these methods are not directly applicable to signals that contain a large amount of zeros and can be described of as arising from a MRC with an atom at zero. *Veneziano and Furcolo* [2009] developed a technique for improving the robustness of the estimation of the non-parametric $\tau(q)$ function. The idea is simple and is based on increasing the "realizations" of the process by splitting the data-series into non-overlapping subsamples. A similar idea is widely used in spectral analysis for smoothing the periodograms by windowing the

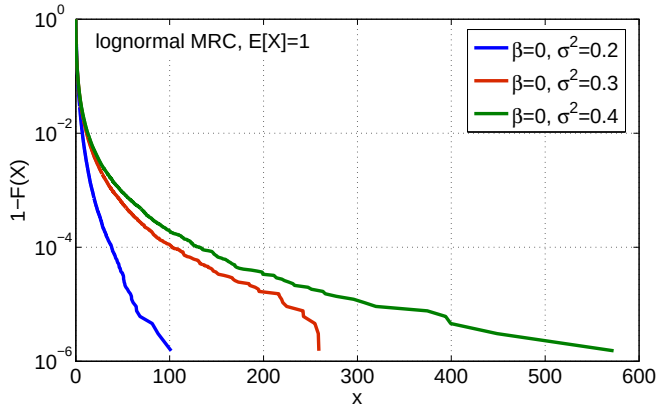


FIGURE 3.7: Exceedance probability distributions for a unit mean lognomal MRC with various values of σ^2 parameter. The results come from a realizations of a cascade with $n = 16$ and $b = 2$ dressed for 5 more levels.

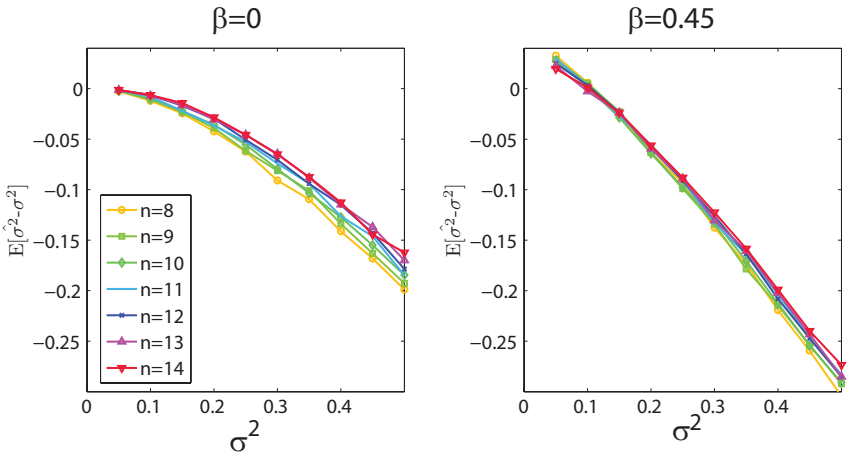


FIGURE 3.8: Mean value of the estimated bias $\overline{\hat{\sigma^2} - \sigma^2}$ for q_{max}=3. Lines with different colours correspond to different cascade development steps n according to the legend. Left panel corresponds to non-intermittent series and right panel to highly intermittent ones.

signals [*Welch*, 1967].

$$\overline{\tau(q)} = \frac{1}{k}\sum_{i=1}^{k}\hat{\tau}_i(q). \tag{3.28}$$

This procedure increases slightly the bias of the estimation with the benefit of higher robustness.

A numerical justification of the improvement of the robustness can be shown in figure 3.9. The figures correspond to a Monte-Carlo simulation similar to the ones reported before. In this case the sample length of the simulation was fixed to $2^{17} = 131072$ and the simulated time-series where split into 20 sub-samples. The number of realizations drawn per parameter set $\theta = \{\beta, \sigma^2\}$ was 100. Those parameters where not arbitrarily chosen but rather based on our rain-gauge database as will be explained in 3.4. An improvement of the robustness is achieved for both parameters expressed in terms of the standard deviation of the estimates.

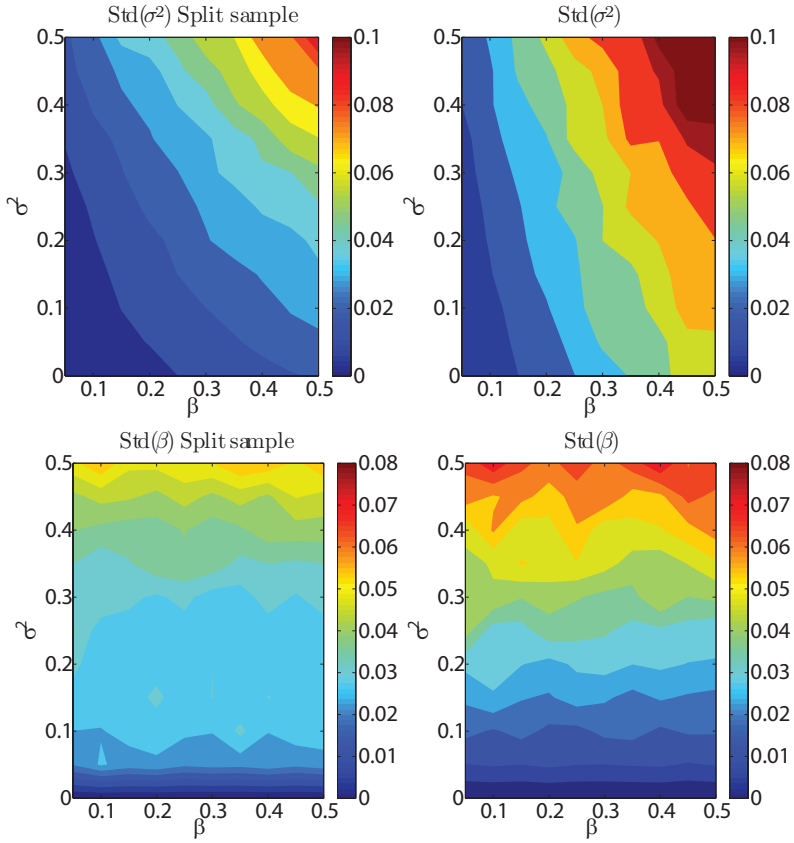


FIGURE 3.9: Standard deviation of the estimated σ^2 and β parameters. Left panel corresponds to the windowing technique and right panel to the classical moment scaling. Upper figures correspond to the standard deviation of the σ^2 parameter and lower figures to the β parameter.

3.3.2 Measurement Artefacts on Scaling Estimation

After having assessed in the previous part of the thesis the performance of various self-similarity estimation techniques for "perfect" data, here we quantify the behaviour of those techniques when data suffer from problems associated with various measuring techniques. Previous research has shown that measurement errors of various types can have a significant effect on the parameter estimation of self-similarity indices [e.g. *Harris et al.*, 1997; *Mandapaka et al.*, 2010]. In this section several measurement artefacts associated with gauge-based measurements are identified and the biases they impose on the estimated statistics are quantified.

The Zermatt Example

In order to demonstrate potential problems that the quality of the data can propagate to the various scaling estimators, a comparison of two different precipitation measurement techniques, that of weighing gauges and tipping bucket gauges is presented here. The study area is located in Zermatt, in South-Western Switzerland (see figure 3.10). The gauges are located 1638 m a.s.l. which means that winter precipitation is mostly snow, summer precipitation is mainly rain and spring/autumn have mixed conditions. The data have been previously analysed from *Savina et al.* [2012] and *Paschalis et al.* [2012]. The two measuring

instruments used here are the standard tipping bucket rain-gauge of the operational SwissMetNet and a high resolution weighing gauge (MPS [4]). The technical characteristics of the tipping bucket can be found in 2.2. The temporal resolution of the weighing gauge is one minute and its depth resolution is 0.001 [mm]. Due to the high accuracy of the second instrument, here it will be considered as the benchmark.

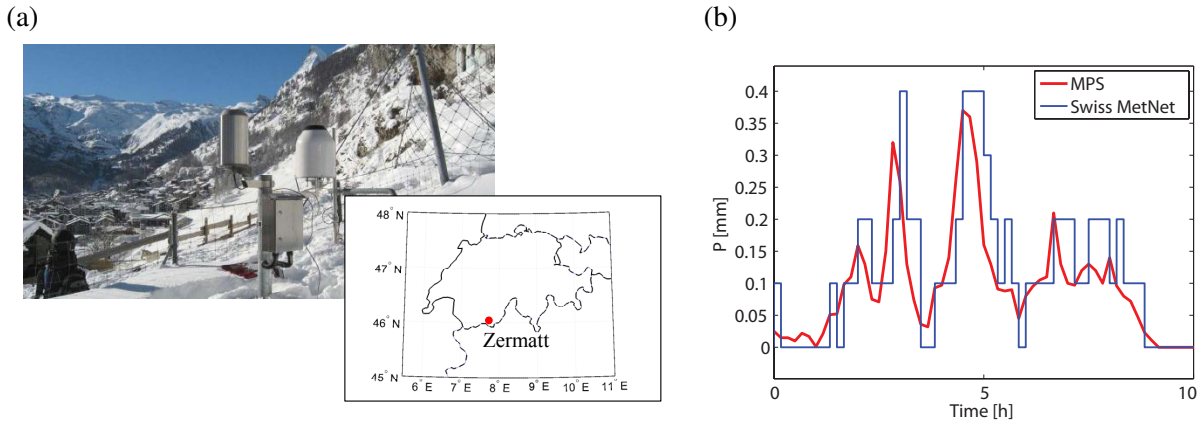


FIGURE 3.10: (a) Photo of the two gauges, weighing (right) and tipping bucket (left) for the Zermatt area, after *Savina et al.* [2012] (b) A 10-hour storm sampled using a weighing gauge (MPS) and a tipping-bucket rain-gauge (Swiss MetNet). Both records are plotted for 10 minute aggregations.

The two time periods for which both gauges were functioning simultaneously are from 22 January to 2 April 2009 and from 29 November 2009 to 10 March 2010. The main difference between the two precipitation records that was found by *Savina et al.* [2012] is that winter solid precipitation is highly underestimated from the tipping-bucket gauge due to the measurement lag caused by the tipping mechanism that subsequently leads to higher evaporation losses. An example of the two different measurements are demonstrated in figure 3.10 (b) where a 10-hour spring snowfall event is shown. The main discrepancies between the two are the different sampling depth resolution and their relevant time shift due to the delay of the melting of the snow and the lag of the tip.

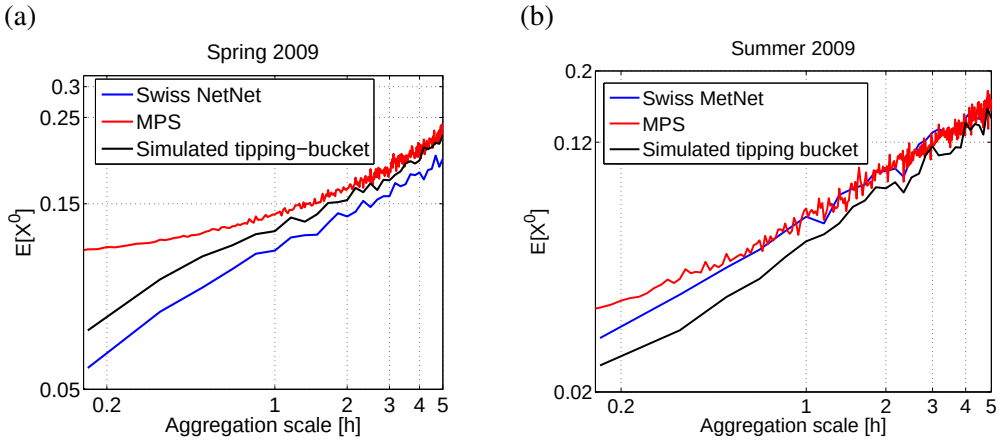


FIGURE 3.11: Scaling of the zero order moment for the Zermatt records during (a) Spring 2009 and (b) Summer 2009

[4] www.mps-system.sk

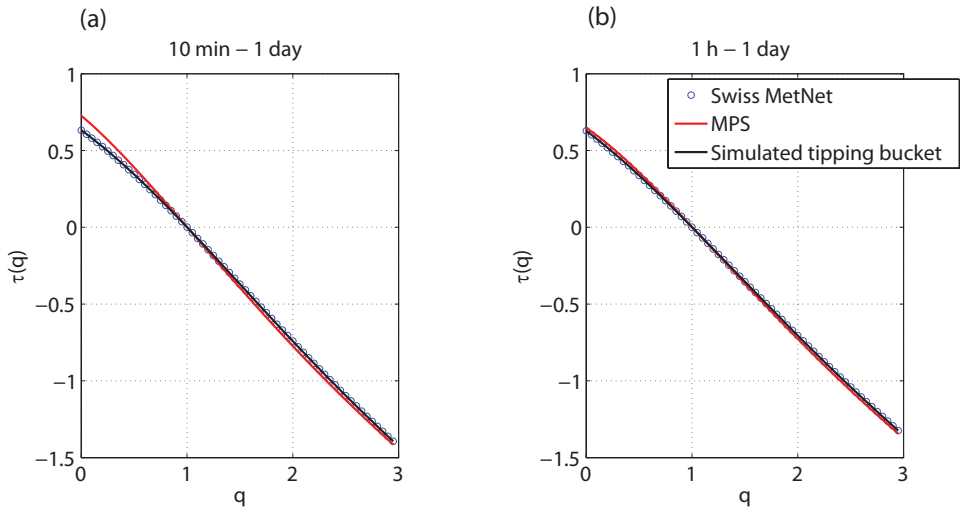


FIGURE 3.12: Moment scaling function $\tau(q)$ estimated for two different temporal regimes during spring 2009. The figure on the left (a) was estimated for scales spanning from 10 min to 1 day and on the right (b) was estimated for scales spanning from 1 hour to 1 day

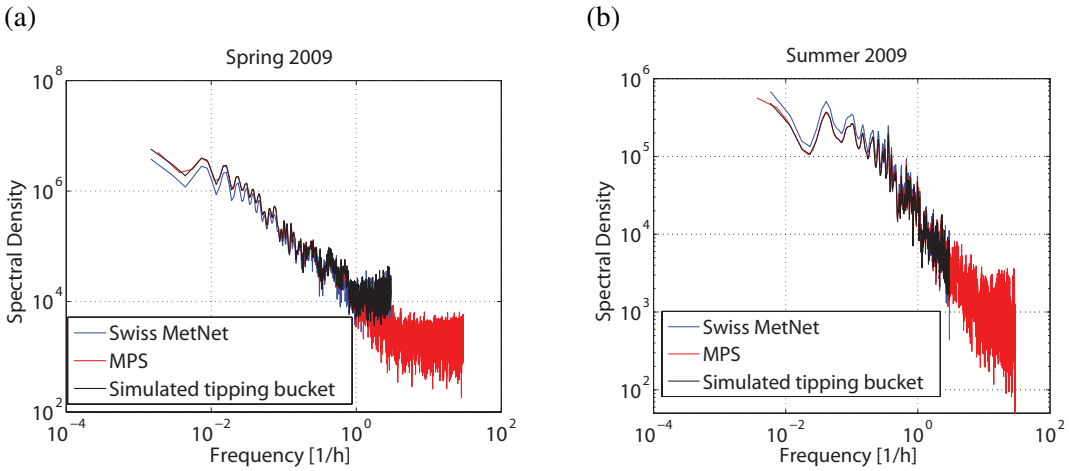


FIGURE 3.13: Periodograms of the precipitation records in Zermatt for (a) Spring 2009 (b) Summer 2009.

Here an investigation for two seasons is presented (spring and summer 2009). The study period is rather short and the number of recorded event small. Thus the results cannot be generalized, but they can serve as a first indication for the effect of measurement artefacts on the estimation of some of the most common scaling estimators. The two techniques that are examined are the classical moment scaling and the spectral analysis of the data series.

The periodograms for the series are calculated using the windowed spectral estimation technique developed by *Welch* [1967]. The time-series are split into sub-records with 50% of overlap and the total periodogram is estimated as the average of the ones of the sub-records. The estimate is biased but smoother.

In order to asses the role of the impact of the recording mechanism, equivalent "tipping bucket" records are developed from the weighing gauge measurements. The procedure is straightforward. The simulated tipping-bucket record is developed by quantizing the initial record according to a specific depth resolution equal to that of the tip volume, taking into account the lag of the tip (i.e. The recorded tip is reported only

when the bucket fills with the amount of water of the selected depth resolution). Here no evaporation losses are taken into account.

In figure 3.11(a) the differences between the two gauges for the scaling of the zeros order moment are shown. This property is crucial for the estimation of the β parameter of the MRC with an atom at zero. This estimation illustrated the probability of having zero values across scales. A big discrepancy occurs for scales less than one hour, where the tipping bucket records underestimate the probability of rainfall detection. This behaviour can be solely attributed to the tipping-bucket recording mechanisms, since the simulated record of the equivalent "tipping-bucket" records from the weighing gauge, show the same behaviour like the ones operated by the SwissMetNet. Those discrepancies lead to large differences of the estimation of the $\tau(q)$ function (figure 3.12(a)) for orders $q < 1$. This is expected since for moments $q < 1$ the behaviour of $\tau(q)$ depends on the small scale fluctuations that cannot be well represented when the sampling depth resolution is coarse. Scaling of the moments $q > 1$ is practically identical between the two different measurements. Another interesting result comes from the spectral analysis of the two signals (figure 3.13(a)). The inability of the recording mechanism to capture small scale variability, can be shown from the differences of the spectral components for scales less than one hour. More, an artificial scaling break is introduced for small scales , showing small scale decorrelation due to the flattening behaviour of the spectral density. Also in this case, the errors can be also attributed mainly to the tipping mechanism. It is clearly shown that all the estimation artefacts can be neglected for aggregation scales larger than 1 hour.

Interestingly, as shown in figures 3.11(b),3.13(b), these discrepancies disappear for the summer season. The main reason is that during summer, precipitation is liquid and concentrated in more intense events. This leads to shorter lag times, and also less quantized events. The faster response of the tipping bucket leads to a much better representation of the probability of zero detection across scales that ameliorates the effects of the estimation of the β parameter. Also, less quantized signals, can lead to much more consistent representation of the power spectrum for small scale fluctuation (figure 3.13(b)).

The most sticking result of this data analysis is that even though high temporal gauge resolution may exist (e.g. SwissMetNet), special attention should be given to various measurement artefacts which can lead to highly biased estimates of scaling behaviour.

Tipping Bucket Quantization

The previous example of the data intercomparison gave some insights concerning the biases that can be introduced due to low quality measurements. One of the most significant results of the previous analysis is that a large part of the discrepancies can be attributed to artefacts associated with the recording mechanism of the tipping bucket. For this reason, in this section a thorough numerical quantification of the effect of the tipping bucket mechanism is provided. A simple numerical Monte Carlo experiment is constructed in order to quantify those biases. The statistics that are explored are:

1. The $\left\{\beta, \sigma^2\right\}$ parameters of the beta-lognormal model.

2. The slope of the approximate power-law spectral decay

The parameters of the model are evaluated with the typical moment scaling analysis explained in 3.3.1. The power spectral densities were evaluated with two different techniques. First the classic periodogram

technique was calculated, using the Fourier decomposition of the signal, exploiting the Fast Fourier Transform (FFT) algorithm.

$$S(k) = |\mathcal{F}(k)|^2 , \tag{3.29}$$

where the Fourier transform of a function $f(x)$ in one dimension is defined as

$$\mathcal{F}(k) = \int_{-\infty}^{\infty} f(x) \exp(-2\pi i x k) dx. \tag{3.30}$$

Second, the estimation of the power spectrum using the wavelet decomposition of the series was used. The continuous wavelet transform of a function $f(x)$ is defined as [*Katul and Parlange*, 1995]

$$\mathcal{W}(b,a) = c_g^{-1/2} \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} \psi\left(\frac{x-b}{a}\right) f(x) dx, \tag{3.31}$$

where $\psi(x)$ the wavelet function, a is the scale dilatation and b is the position translation. $\psi(x)$ has to fulfil special conditions described in *Mallat* [1998]; c_g is defined as

$$c_g = \int_{-\infty}^{\infty} \frac{1}{|k|} |\psi^*(k)|^2 \, dk, \tag{3.32}$$

where $\psi^*(k)$ is the Fourier transform of $\psi(x)$ defined in 3.30.

Then the power spectrum can be defined as

$$S(k = 2\pi/a) = \int \mathcal{W}(a,b) db. \tag{3.33}$$

Wavelet decomposition is done with the discrete wavelet algorithm [*Mallat*, 1989]. Details of the algorithmic approach can be found in *Katul and Parlange* [1995]. The wavelet function that was chosen in this analysis (D4) belongs to the Daubechies family [*Daubechies*, 1988]. The two different methodologies for the estimation of the power spectrum have two strong differences. The first one is the method's sensitivity, and thus robustness as explained by *Katul and Parlange* [1995]. The second one is associated with discrete techniques that are applied for the estimation of the Fourier of wavelet transform. The difference is that the 2 methodologies (FFT, DWT) estimate the spectral densities for different frequencies, with the FFT algorithm giving more emphasis (i.e more frequencies) on the lower scales, if all of them are taken into account. So in principle the two different methodologies will be affected to a different degree from the small scale discrepancies.

The simulation parameter space is defined in table 3.2 and the number of simulations per parameter set is 100 mainly defined due to computational restrictions. The quantization levels are expressed as a percentage of the series average value (i.e $V_{tip}/E[R] * 100$ [%], where V_{tip} [mm] is the volume of the

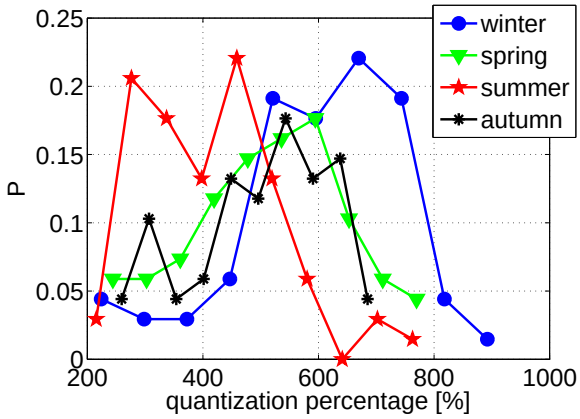


FIGURE 3.14: Histogram of the quantization values for the Swiss MetNet gauge network expressed as the percentage of the mean value of the record.

Parameter	min	max	spacing
β [-]	0	0.5	0.05
σ^2 [-]	0.05	0.5	0.05
quantization level [%]	0	1000	100
noise level [%]	10	100	10
cascade level [-]	7	14	1

TABLE 3.2: Parameter Space of the Monte Carlo Experiment for the quantification of the biases introduced by the tipping bucket mechanism.

tip and $E[R]$ [mm] is the average precipitation depth at the finest temporal resolution). The limits of the parameter space are determined by analysing the records of the Swiss MetNet. In figure 3.14 the histograms of the quantization values derived from the 68 10-min resolution gauges are shown dependent on the season. The quantization values are lower for summer due to the higher depth accumulations.

Similar efforts have been previously done [e.g. *Harris et al.*, 1997] illustrating the importance of the errors that can be introduced. The main novelty of the current investigation is that it expands the previous results on a much more extended parameter space. Previous efforts were illustrating the results only for specific cases, based on data analysis of a typically small number of records of various geophysical signals. The results reported in figures 3.15 and 3.16, correspond to the case of $n = 14$ cascade steps. The approximation of the dressing factor Z_∞ if done with 5 more cascade steps.

In figure 3.15, the mean value of the estimated biases are presented for the calculation of the parameters $\{\beta, \sigma^2\}$. Obviously, when the quantization level increases, the results become more biased concerning both parameters. For the exponent β that governs the intermittency of the simulated and quantized series, the bias is larger when the "true" simulated series contain a small number of zeros, and the σ^2 parameter, that describes the "spikiness" of the process, increases. The reasons why this happens are two. First, the simulated tipping-bucket procedure that quantizes the signal, enhances the number of zero records, since a "quasi" tip and thus a valid measurement occurs only when the sampling volume of the gauge has been filled. This effect of the measurement lag enhances when σ^2 is larger. This can be explained due to the fact that the marginal distribution of the simulated series is more skewed when σ^2 is larger. This leads to a higher proportion of low values in the simulated record that are affected to a larger degree by the lag of the tipping bucket mechanism.

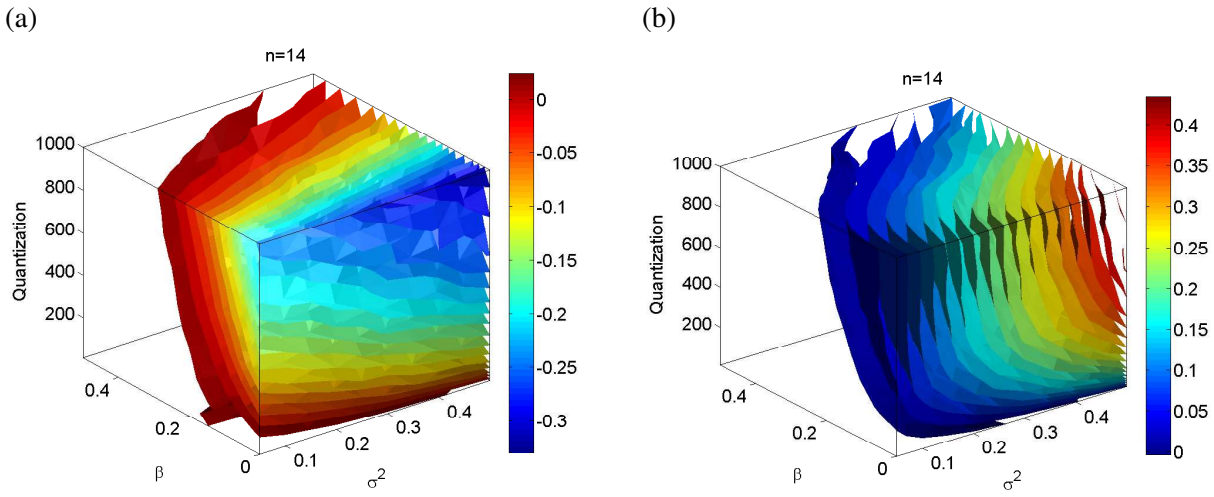


FIGURE 3.15: Estimated bias due to signal quantization for the multifractal beta-lognormal model. Left (a) the bias $\overline{\hat{\beta} - \beta}$ and right (b) $\overline{\hat{\sigma^2} - \sigma^2}$. The simulated series correspond to a MRC with cascade level generation $n = 14$

Also in figure 3.15(b) the bias for the parameter σ^2 is shown. The estimated σ^2 is always lower than the simulated one, The bias increases as the simulated σ^2 increases and also as β decreases. It should be noticed that the previous bias of the β estimate influences as well the estimation of σ^2 due to the calculation procedure mentioned in 3.2.2. The reason why for a specific value of simulated σ^2 the discrepancies increase with decreasing β is that the total measure is redistributed in a larger support, which give rise to lower simulated values which are much more affected by the quantization procedure.

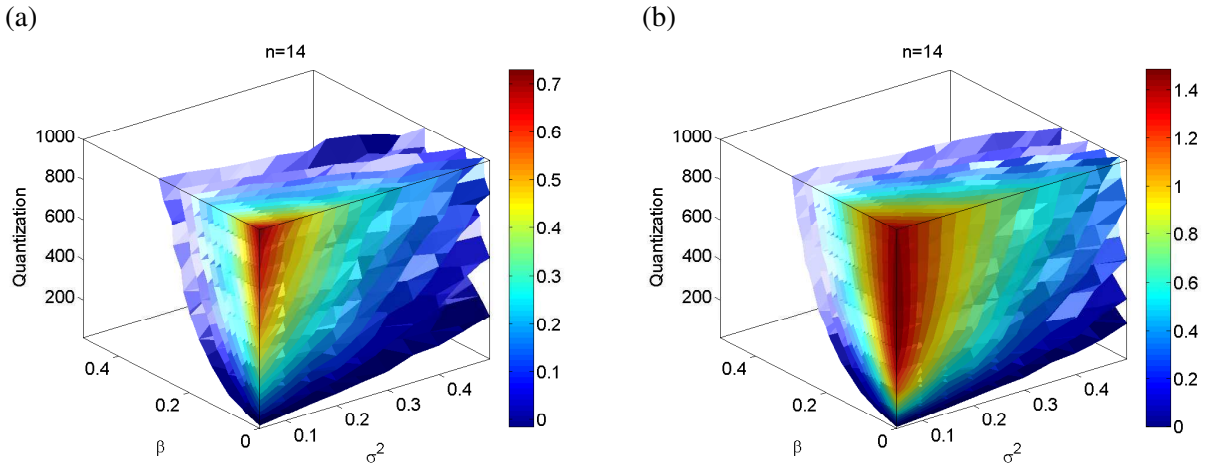


FIGURE 3.16: Estimated bias of the power-law spectral decay due to signal quantization for the multifractal beta-lognormal model. Left (a) the bias $\overline{\hat{\nu} - \nu}$ for the wavelet-based spectral estimation and right (b) the equivalent one for the Fourier based estimation. The simulated series correspond to a MRC with cascade level generation $n = 14$.

In figure 3.16, the estimated biases of the spectral slopes are presented. The estimation procedure is ordinary least square fitting of the logarithms of the frequencies against the logarithms of the spectral density. The bias pattern is, as expected, the same for both estimation techniques. The highest biases

are expected for highly quantized signals with small intermittency (e.g. small β) and generally smooth profiles (e.g. σ^2 small).

An interesting result is that the estimation using wavelet decomposition appears to be much more reliable. The fact why this is happening is that the wavelet coefficients computed using the discrete wavelet transform [*Mallat*, 1989] can be estimated only for scales that are integer powers of 2 which is as already mentioned very different from the FFT algorithm. Subsequently this gives less weight on the high frequency components in comparison to the periodogram calculated using the FFT decomposition and since quantization leads to higher errors on the high frequency components (see 3.13) this results the fact that spectral slopes are much better reproduced by the wavelet decomposition. The common pattern is that the spectral power law decay is always underestimated, and the underestimation is higher when both β and σ^2 parameters are low.

Taking into account that winter precipitation in the study area investigated here is mainly stratiform and in the form of snow, the highest biases on the data analysis of the records are expected in winter. The reason is that statiform precipitation is much more uniformly distributed than convective which leads to low σ^2 parameters, and considering the higher percentages of quantization levels (figure 3.14), higher biases on both β and ν are expected (see figures 3.15 and 3.16). Another interesting conclusion is that precipitation measurements for very high temporal scales can be highly misleading if the quantization levels are high. A striking example of this could be the precipitation records of NCDC[5] that adopt a depth resolution of 1/10 inch ($\sim$2.5 mm), which at a 10 minutes resolution can be of the order of 6000% of the mean precipitation value for a temperately humid climate (e.g. 6 mm/day).

Noise Corruption

The results above illustrate the problems associated with the recording principle of the tipping-bucket gauges which in some of the cases can be dominant. As shown by *Harris et al.* [1996] and *Veneziano et al.* [2006a] another main source of bias can be the data noise corruption. An excellent review concerning the gauge measuring errors can be found in [*Lanza and Stagi*, 2009; *Vuerich et al.*, 2009]. The study of *Harris et al.* [1996] focused on some illustrative examples of the problem and *Veneziano et al.* [2006a] analysed mainly the effects on the moment scaling function for negative orders, which in our case diverge since the analysis is conducted for MRC with atom at zero.

Here the previous two studies are expanded taking into account two different types of noise corruption.

1. Additive zero-mean Gaussian uncorrelated noise $\sim N(0, s_n)$.
2. Unit mean lognormally distributed uncorrelated multiplicative noise $\sim LN(m_{ln}, s_{ln})$ where $m_{ln} = -s_{ln}/2$.

The standard deviation of the noise (s_n for the additive noise and $\sqrt{(\exp(s_{ln}^2 - 1)\exp(2m_{ln} + s_{ln}^2))}$ for the multiplicative noise) can be expressed as a percentage of the standard deviation of the data exactly as in *Veneziano et al.* [2006a]. The noise corruption levels that are investigated here are shown in table 3.2. Typical values of the expected errors are not known for the gauge-network that will be analysed later, since a lot of parameters may play a major role, such as e.g. wind disturbances etc. [*Sevruk et al.*, 2000]. For this reason a rather large sample size of probable noise corruption levels is chosen.

[5] http://www.ncdc.noaa.gov/oa/ncdc.html

(a) Multiplicative noise (b) Additive noise

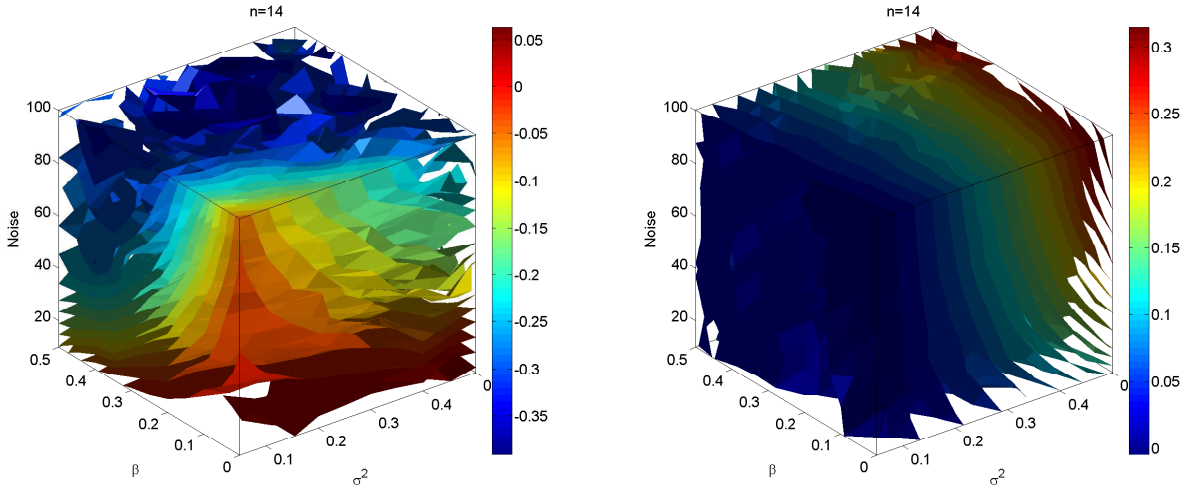


FIGURE 3.17: Bias of the estimation of the σ^2 parameter of a lognormal MRC due to noise corruption. In the left panel (a) the results for multiplicative noise corruption are shown for series developed using a beta-lognormal MRC with cascade level generation $n = 14$. In the right panel (b) the same results for the additive noise are shown.

First the bias for the estimation of the σ^2 parameter is shown in figure 3.17. It should be noted that the estimation of β is unaffected since the noise perturbations are imposed only on the positive values of the simulated series. For the case of additive noise, the small fraction of the negative values that were simulated, where replaced with zero, introducing a minor bias on β that was neglected in further analysis. A main difference is that the two different kinds of noise have an opposite effect on the estimated statistics. The lognormally distributed multiplicative noise leads to higher values of the estimated σ^2 for the perturbed signal. The effect, as expected, is higher as the noise corruption levels increase. This happens if we consider the right-skewed shape of the lognormal distribution, that can lead to values $>> 1$ that when they are multiplied with high values of the original signal, finally lead to very strong spikes at the perturbed signal. This results much higher values of σ^2. On the other hand, corruption with additive noise leads to lower values of σ^2. For the multiplicative noise, the errors enhance for lower values of the intermittency parameter β. On the contrary β appears to have no effect, when signals are corrupted with additive noise.

Taking into account that the power spectrum is the Fourier transform of the auto-covariance function, analytical expressions can be derived for the power spectrum of the noise contaminated signal. For the case of the additive noise we have that

$$X_P = X + \epsilon_g, \tag{3.34}$$

where X is the multifractal process and ϵ_g is the Gaussian zero mean additive noise. X and ϵ_g and independent. It can be shown that

$$\text{Cov}[X_P(t), X_P(t+\tau)] = \text{Cov}[X(t), X(t+\tau)] + \text{Cov}[\epsilon_g(t), \epsilon_g(t+\tau)], \tag{3.35}$$

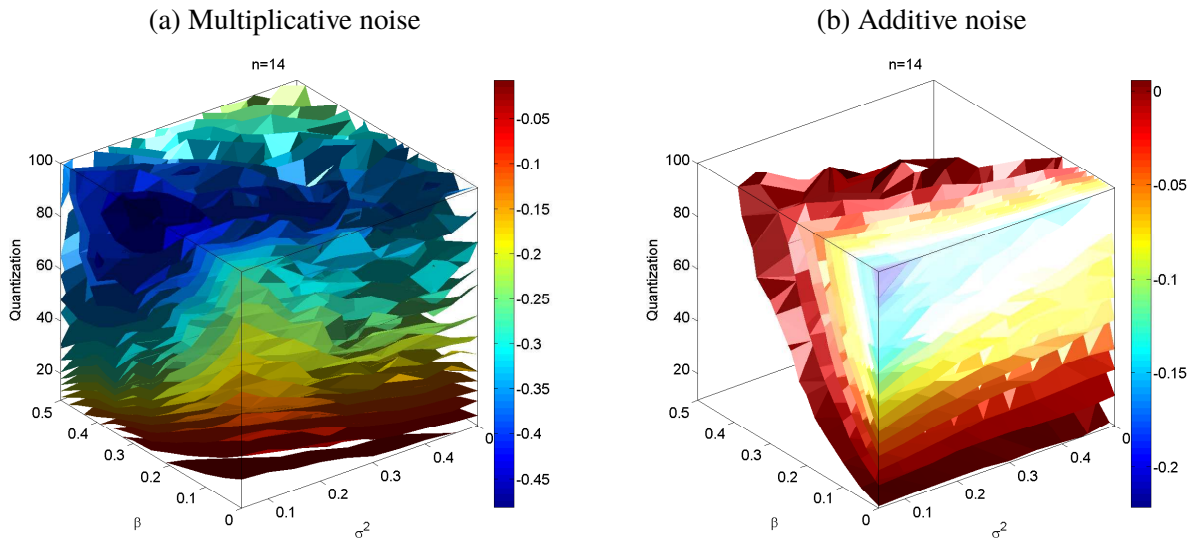


FIGURE 3.18: Bias of the estimation of the spectral slopes. The power spectral density is estimated according to a wavelet decomposition of the signal. Left panel (a) illustrates the result of multiplicative noise corruption, and the right panel (b) for the additive noise corruption. The series are simulated using a beta-lognormal MRC with cascade level generation $n = 14$

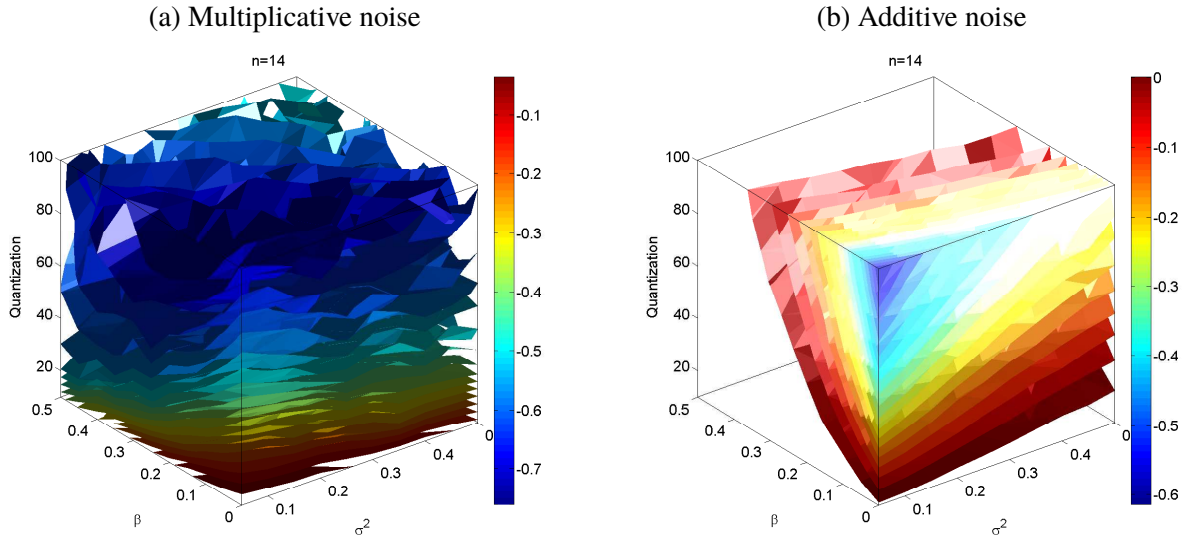


FIGURE 3.19: Bias of the estimation of the spectral slopes. The power spectral density is estimated according to a Fourier decomposition of the signal. Left panel (a) illustrates the result of multiplicative noise corruption, and the right panel (b) for the additive noise corruption. The series are simulated using a beta-lognormal MRC with cascade level generation $n = 14$.

where Cov[.] are the respective autocovariance functions if X_P, X, ϵ_g are all stationary. Taking the Fourier transforms it can be shown that

$$S_{X_P}(k) = S_X(k) + S_{\epsilon_g}(k). \tag{3.36}$$

On the other hand, for the case of multiplicative noise we have that if

$$X_P = X\epsilon_{LN}, \tag{3.37}$$

then

$$\begin{aligned}\text{Cov}[X_P(t), X_P(t+\tau)] &= \text{Cov}[X(t), X(t+\tau)]\text{Cov}[\epsilon_{LN}(t), \epsilon_{LN}(t+\tau)] \\ &+(E[\epsilon_{LN}])^2\text{Cov}[X(t), X(t+\tau)] + (E[X])^2\text{Cov}[\epsilon_{LN}(t), \epsilon_{LN}(t+\tau)],\end{aligned} \tag{3.38}$$

which leads to

$$S_{X_P}(k) = S_X(k) \star S_{\epsilon_{LN}}(k) + (E[\epsilon_{LN}])^2 S_X(k) + (E[X])^2 S_{\epsilon_{LN}}(k), \tag{3.39}$$

where $\star$ stands for the convolution operator.

In figure 3.20 the theoretical power spectra for additive and multiplicative noise corruption is presented for a case of a non-intermittent cascade with $\sigma^2 = 0.2$ and unit variance. The correlated noise fields are assumed to be stationary and have an exponential autocorrelation function

$$\rho(\tau) = \exp\left(-\frac{\tau}{l}\right), \tag{3.40}$$

which leads to a spectrum of the form

$$S(k) \propto \frac{1}{\pi} \ \frac{1/l}{k^2 + (1/l)^2}. \tag{3.41}$$

It is obvious that the high frequency components are the most affected in the case of uncorrelated noise and lead to the "whitening" of the signal for high frequencies fact that can be shown with the tendency of the spectrum to be flat in this regime. As previously shown this is also the typical appearance of the power spectrum of the high resolution recorded time series in Zermatt. On the other hand, correlated noise leads to different behaviour, mainly affecting the frequency components around its decorrelation length. For temporal precipitation records, correlated noise is not expected and thus not included in the extensive Monte Carlo experiment for the bias quantification. It will be later though analysed when spatial rainfall will be analysed.

In figures 3.18, 3.19 the effects on the estimation of the power law decay of the power spectral density is presented. In both cases, the spectral slope is underestimated. Since the power spectrum is the Fourier transform of the auto-covariance function, this leads to the obvious results that the autocorrelation of the corrupted signal, is lower, since the noise in uncorrelated. Moreover the biases that are introduced using the wavelet decomposition of the signal, are typically lower. For the additive noise, a clear pattern of the bias is observed. The bias increases for lower values of β and decreases for higher values of

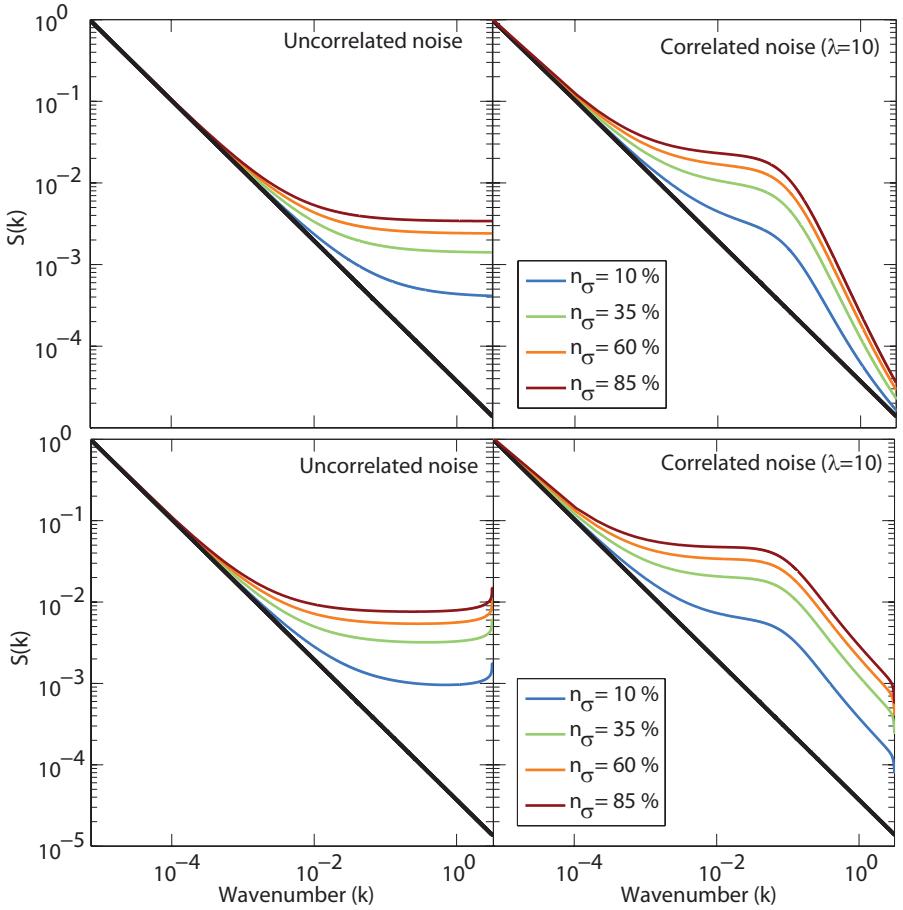


FIGURE 3.20: Theoretical power spectra for noise corrupted signals for a unit variance MRC with $\sigma^2 = 0.2$ and $\beta = 0$. n_σ corresponds to the percentage of noise corruption. Upper panel corresponds to additive zero-mean Gaussian noise and the lower panel to lognormally distributed multiplicative noise. In all the cases the power spectrum of the original signal is shown with a thick black line.

σ^2. For the multiplicative noise corruption, no clear pattern is observed. This is probably due to the fact that the number of the simulated series of the Monte-Carlo experiment was not large enough in order to have good convergence to the desired results. The only clear result is that the bias in this case increases when the series are more intermittent, which is the opposite behaviour to one when series are perturbed with additive noise. The maximum errors can reach up to 50%. In both cases, due to the fact that noise corruption has a major effect on the high frequency components, estimation of the power-law spectral decay is recommended for data analysis that the estimation should be done for higher scale (low frequency) components. This though would be only possible is scaling holds true for a wide range of temporal scales, which is something that will be later analysed in detail.

3.4 Time - Series Data Analysis

Taking into account all the previous results in terms of parameter estimation, a thorough investigation of the precipitation records from our rain-gauge database is presented in this section, having as a background

the theory of scale invariance.

3.4.1 Moment Scaling Analysis

The first statistic that is investigated is the scaling of the moments. As in our previous analysis, in order to have a parametric notation in the investigation of scale-invariance in the gauge-data sets, a specific case of a multi-scaling behaviour should be adopted. Also here, all the results are based on the beta-lognormal model (table 3.1). It is conceptually the simplest, most parsimonious and yet complete notation to investigate the temporal structure of precipitation across scales. Also, interesting conclusions can be drawn with the comparison of the results achieved here, and the study of *Molnar and Burlando* [2008] that used the same data-base (with shorter records), but used a different analysis procedure.

The analysis is conducted on a seasonal basis in order to take into account the different types of precipitation (rain/snow) that occur and also to roughly separate the different major precipitation triggering mechanisms like stratiform and convective precipitation.

In the previous sections it has been shown that one of the main difficulties of the parameter estimation of beta-lognormal model is that the maximum moment order the moment scaling function converges to the MKP function is not known a priori. In order to overcome this difficulty, the following algorithm presented schematically in figure 3.21 is constructed.

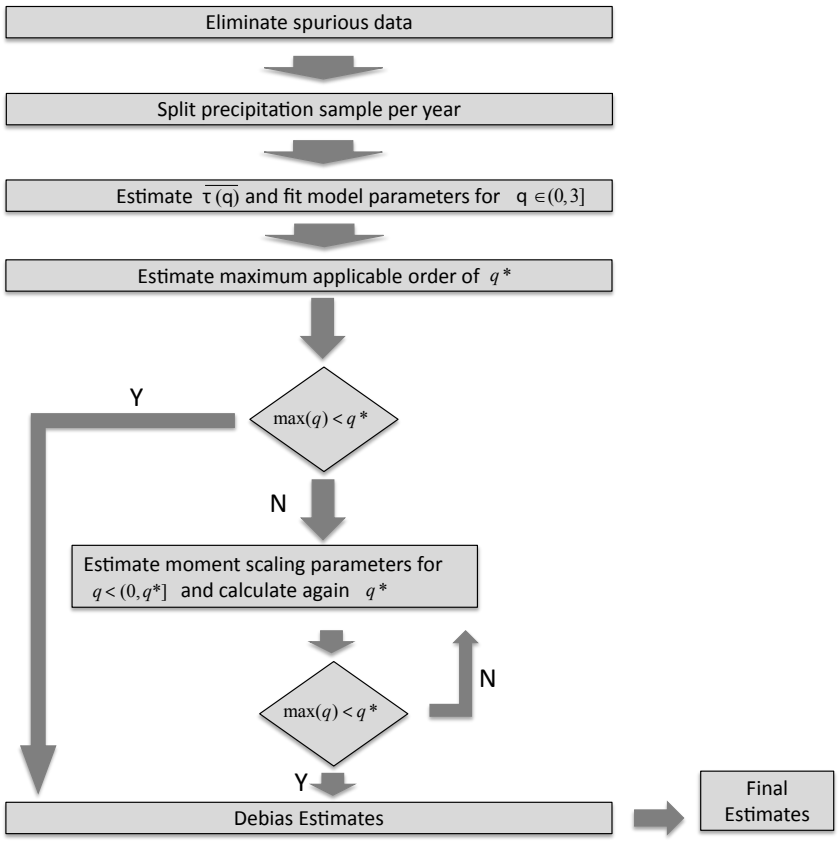


FIGURE 3.21: Algorithm for the estimation of moment scaling.

First, all spurious values are eliminated from the records. Those values can be the result of various errors such as digitization errors, corrupted data transmission, etc. Here a simple threshold criterion was used

in order to identify those values. All values that exceed an intensity of 50 mm/10min at the Swiss MetNet data and 300 mm/d at the daily observations, were flagged as errors. Those values, were either filled with simple linear interpolation, if the missing values were not connected, or neglected if a continuous gap of missing values existed. The simple linear interpolation is expected to have negligible errors in the case of high temporal correlations. The years for which the records contained a fraction of non-available data larger than 5% were considered as unreliable and excluded from the analysis. Taking into account that the data measuring network is well maintained, error occurrence or extensive data unavailability was scarce.

Second in order to increase the robustness of the estimation (see 3.3.1), the moment scaling function of each year was calculated and the non-parametric $\overline{\tau(q)}$ was derived as the average. In order to proceed to an analysis like that, the hypothesis of stationarity across the years was invoked. The assumption of stationarity may be considered as crude if the results of the climate change modelling have to be taken into account [e.g. *IPCC*, 2007].

Afterwards, an iterative procedure was applied. The MKP function of the beta-lognormal model was fitted for $q \in [0, 3)$ and the parameters $\{\widehat{\beta}, \widehat{\sigma^2}\}$ were estimated. A first estimate of the maximum order $\widehat{q^*}$ was the obtained according to equation, 3.25. In case the previous maximum order for which the MKP-function was fitted is higher than $\widehat{q^*}$, a new parameter set is obtained by least square fitting for $q \in [0, \widehat{q^*})$. The procedure is iterated until convergence. Numerical experiments were conducted with synthetically generated discrete MRC, to show that this procedure yields very good results with minor biases.

One major assumption however is that precipitation series are a priori assumed to be realizations of a discrete MRC of some branching number (here $b = 2$). This can yield some physical inconsistencies especially if the "on-off" grid cascade issue is taken into account [*Over*, 1995]. To the author's opinion this is a topic that still has to be further investigated. But if such an assumption is made, the estimation of the maximum order of q where ergodicity is valid has to be assessed. The main problem of the assumption of the discreteness of the cascade is that the maximum order q is dependent as well on the branching number. The problem could be partially solved if a continuous cascade formalism was adopted, but for consistency with the results of *Molnar and Burlando* [2008] and for comparative reasons, this is not further analysed. The orders for which ergodicity holds true in the continuous case can be found in *Veneziano and Furcolo* [2009], *Ossiander and Waymire* [2002, 2000].

For the Swiss MetNet data the MRC parameters were estimated for two scaling regimes. One spans from 40 minutes to roughly 1 day (21.8 hours) and the second one from $\sim$1 day (21.8 hours) up to $\sim$1 week. The reason for this is that this two scale ranges behave on a different manner and belong to distinct scaling regimes. More details about this are provided in 3.5.1. Sub-hourly scales were excluded due to the fact that, even though they could potentially add valuable information, the quantization and noise errors (explained in 3.3.2) can dominate the entire estimation procedure.

Physical Interpretation of the Results

For the low temporal scales, the σ^2 parameter is much higher (figure 3.22) during summer. This can be explained due to the fact that summer precipitation is much more intense and concentrated in time, fact that leads to more rough and "spiky" profiles. Winter precipitation, which is mainly composed by stratiform events, creates much smoother records, with lower σ^2 values. During winter, almost half of the stations show an almost simple scaling behaviour with $\sigma^2 = 0$. Autumn and Spring have similar behaviour that lies between the two extreme cases (winter, summer). For coarse temporal scales, the seasonal differences are slightly smaller but the absolute values of the σ^2 parameter are higher. Comparing the results with the

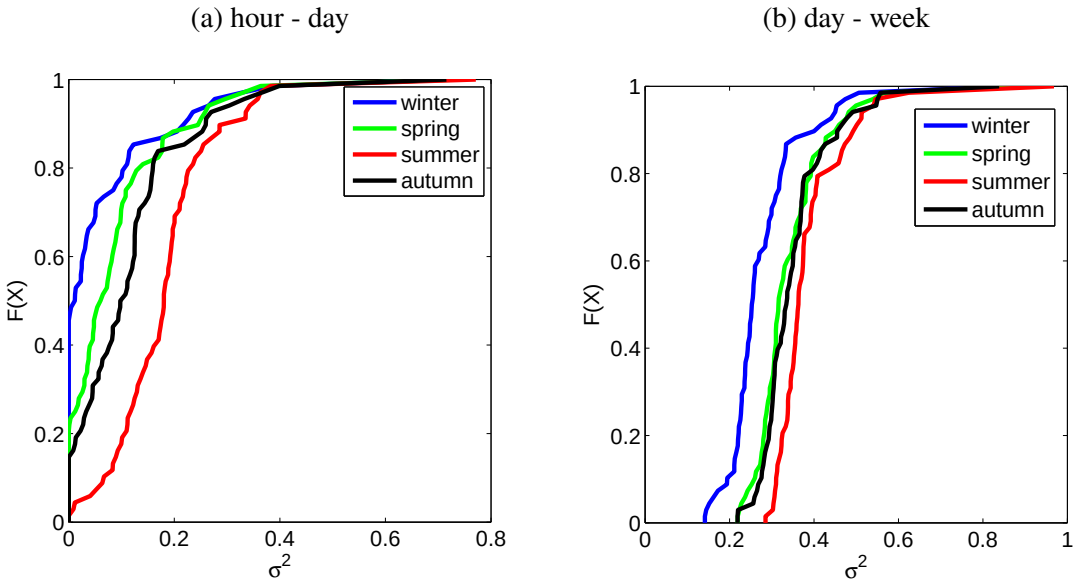


FIGURE 3.22: Cumulative distribution function of the σ^2 parameter derived by the Swiss MetNet stations. Left panel (a) is for scales from 1 hour up to 1 day and the right panel (b) for scales from 1 day up to 1 week.

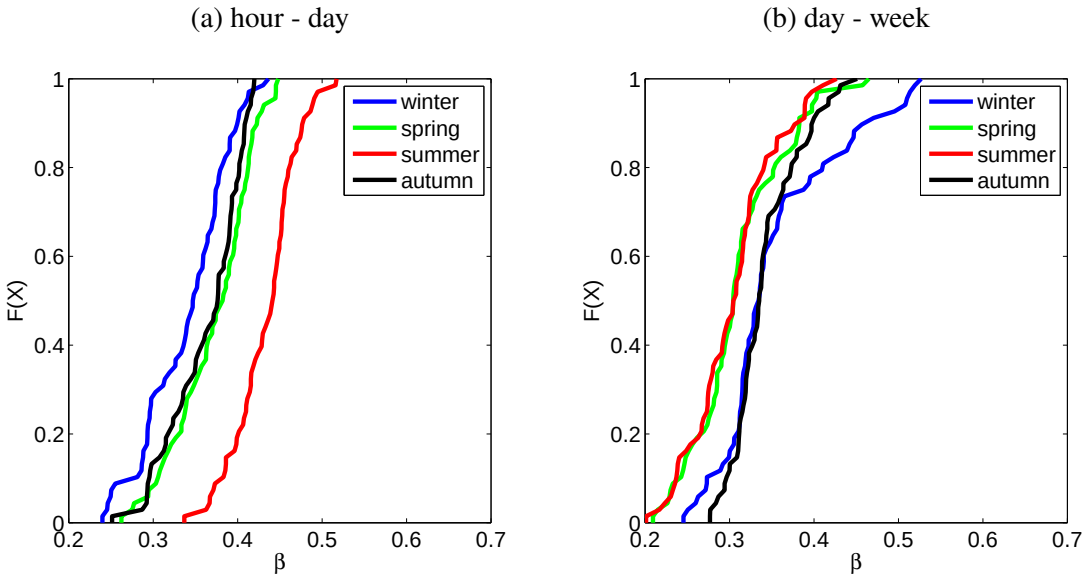


FIGURE 3.23: Cumulative distribution function of the β parameter derived by the Swiss MetNet stations. Left panel (a) is for scales from 1 hour up to 1 day and the right panel (b) for scales from 1 day up to 1 week.

ones obtained by *Molnar and Burlando* [2008], there are two major differences. First, the magnitude of the σ^2 parameter was always higher, and the uncertainty expressed in terms of the standard deviation of the estimates from all the stations is substantially higher. Even though the scope of the analysis is essentially the same, the estimation approach is different and this yielded some contradictions in the results. Those contradictions can be attributed to the biased estimations that were obtained by *Molnar and Burlando* [2008] due to the a priori selection of the orders for which the MKP function was fitted (0-4 in this case).

There are also some regional differences on the results. The North-East part of the Swiss plateau has

for both aggregation intervals higher values. This results though is not highly statistically significant, taking into account the large uncertainty between the stations (see figure 3.24(b),(d)). Generally the NW part of the country appears to have higher values of σ^2 that the alpine mountain range or the Southern Mediterranean side of the Alps. This is also in agreement with the findings of *Molnar and Burlando* [2008].

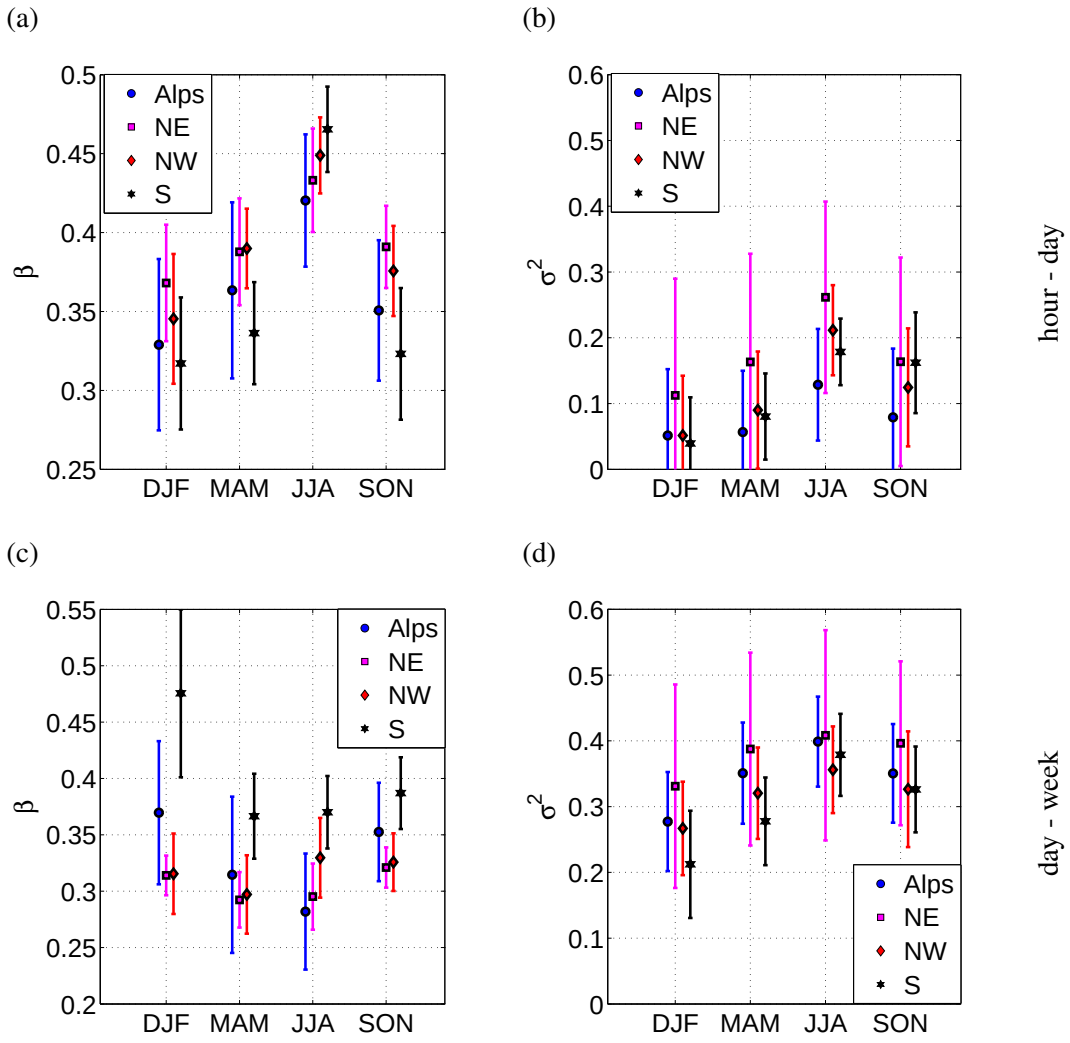


FIGURE 3.24: Seasonal and regional variation of the β and σ^2 of a lognormal MRC for the Swiss MetNet data.(a),(b) correspond to the scaling regime of $\sim$1 hour to $\sim$1 day and (c),(d) for the scaling regime of $\sim$1day-$\sim$1week. The error bars indicate the standard deviation from all the stations belonging to each region.

The β parameter that is an indication of the intermittency of the process and describes the probability of zero rainfall across scales is higher for summer and lower for winter. This can be explained by the fact that summer precipitation consists of small duration intense events in contrast to long lasting winter ones. Moreover the β parameter is highly connected with the mean value of the series. Intuitively the result demonstrates that higher precipitation accumulations occur by events that have a larger temporal extend. This result has been confirmed again for various types of precipitation data [e.g. *Pathirana and Herath*, 2002; *Molnar and Burlando*, 2005; *Over and Gupta*, 1996; *Veneziano et al.*, 2006a; *Serinaldi*, 2010a]. An interesting feature is also that there is a weak connection between the β parameter that governs the

intermittency of the process and the altitude of the station. There is a higher probability of low values of β with increasing altitude. Since this weak dependence appears for both summer and winter, the chance of it being an artefact due to the melting of the snow in the gauge, due to which the amount of collected precipitation is distributed to longer time periods, is probably small.

The values of β for the Southern part of Switzerland that are connected with the wet anomaly around the lake Maggiore are much higher for higher aggregation intervals, especially during winter (see figures 3.24, 3.27). Surprisingly the same does not hold true for low aggregation intervals (Figure 3.24). The same result is also valid for the dry valley of Vallais (South-West). This result cannot be easily illustrated with the data from the Swiss MetNet, due to their sparsity in space, but can become apparent by analysing the daily records. This can be considered as a strong indication that different precipitation triggering mechanisms lead to much different structure of the storm arrival process, that affects the probability of precipitation detection across scales and consequently to the estimation of the intermittency exponent β.

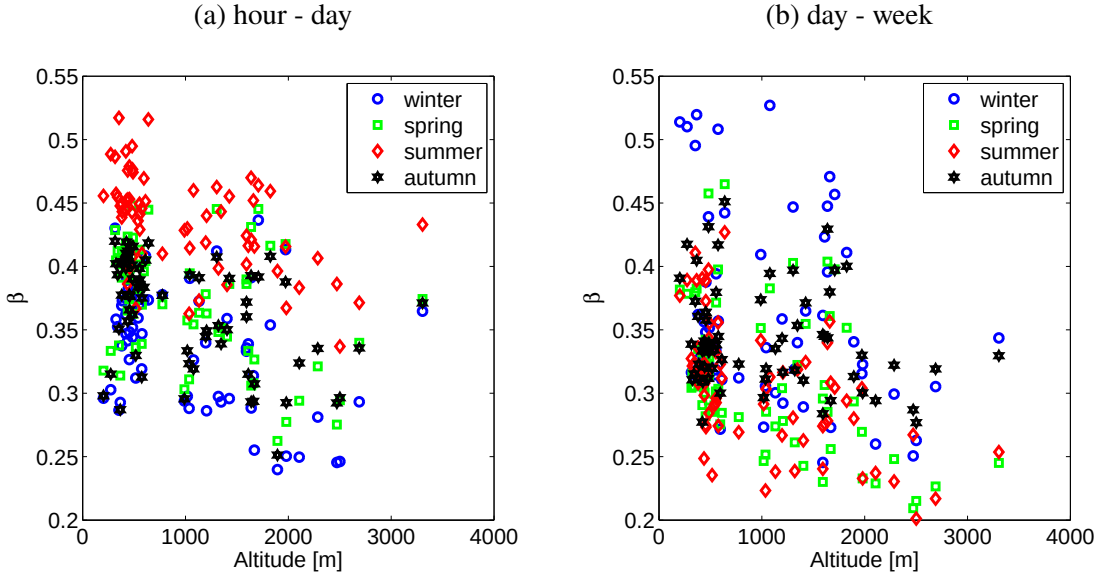


FIGURE 3.25: Relation between the β parameter and the altitude of the station. In the left panel the results for the temporal scales from $\sim$ 1 hour up to $\sim$ 1 day are shown. The right panel corresponds to temporal scales from $\sim$ 1 day up to $\sim$ 1 week.

Similar results concerning the spatial distribution of the σ^2 parameter could not be obtained, since a distinct influence of the terrain topography was not observed and thus not plotted here. The hypothesis of topographic influence cannot be confirmed or rejected due to the large degree of uncertainty on the estimation of σ^2.

One of the most significant features is that the seasonality of the σ^2 parameter is the same for both scaling regimes that are analysed here, in contrary to the β parameter that shows an inverse seasonal pattern for the two scaling regimes.

For the coarse temporal scales, the same analysis can be conducted exploiting the data recorded by the classic daily observation network. Some of the results coincide with the ones of the Swiss MetNet data, and some are different (see tables 3.3,3.4).

[6] ρ is the absolute value of the correlation coefficient

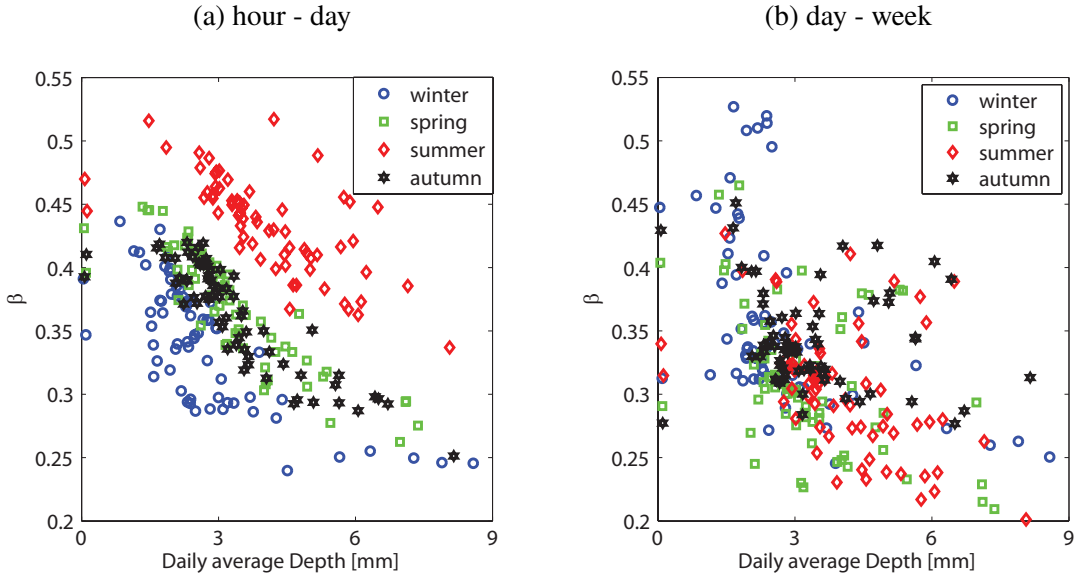


FIGURE 3.26: Relation between the β parameter and the average precipitation of the station. In the left panel the results for the temporal scales from $\sim$ 1 hour up to $\sim$ 1 day are shown. The right panel corresponds to temporal scales from $\sim$ 1 day up to $\sim$ 1 week.

TABLE 3.3: Summary statistics of the multifractal estimates for the Swiss MetNet data. The analysis is based on the beta-lognormal MRC model which parameters are estimated for two different scaling regimes

	Winter (DJF)	Spring (MAM)	Summer (JJA)	Autumn (SON)
mean depth $[mmd^{-1}]$	2.733	3.319	4.118	3.441
		40 min - 21.8 h (Swiss MetNet)		
σ^2	0.075	0.102	0.198	0.130
β	0.341	0.373	0.434	0.364
ρ [6] (σ^2-altitude)	0.063	0.002	0.117	0.023
ρ (β-altitude)	0.413	0.427	0.495	0.419
ρ (σ^2- mean depth)	0.132	0.204	0.149	0.134
ρ (β-mean depth)	0.753	0.896	0.727	0.891
		21.8 h - 1 week (Swiss MetNet)		
σ^2	0.319	0.392	0.442	0.408
β	0.355	0.311	0.305	0.342
ρ (σ^2-altitude)	0.203	0.274	0.202	0.171
ρ (β-altitude)	0.205	0.447	0.559	0.212
ρ (σ^2- mean depth)	0.170	0.200	0.188	0.046
ρ (β-mean depth)	0.503	0.467	0.495	0.196

The seasonality of the parameters β and σ is the same for both data sets. Also the connection of the β parameter to the average precipitation of the station is also the same. Surprisingly, for high aggregation temporal intervals (1 day-1 week), the correlation of β to the average precipitation depth is high for winter, spring and summer, but almost insignificant for autumn, a fact that can be confirmed by both data sets. This correlation is also much lower for high aggregation intervals in comparison to the fine scale ones. On the other hand the high correlation of the β parameter to the altitude station does not hold true for the results derived from the daily observation network. There are several possible reasons for this observation. Since the estimation of the β parameter is based of the probability of precipitation detection

TABLE 3.4: As in 3.3

	Winter (DJF)	Spring (MAM)	Summer (JJA)	Autumn (SON)
mean depth $[mmd^{-1}]$	2.752	3.283	4.342	3.364
		1 day - 1 week (daily observations)		
σ^2	0.209	0.268	0.306	0.244
β	0.310	0.274	0.274	0.319
ρ (σ^2-altitude)	0.106	0.072	0.153	0.123
ρ (β-altitude)	0.021	0.023	0.045	0.005
ρ (σ^2- mean depth)	0.020	0.079	0.084	0.066
ρ (β-mean depth)	0.406	0.386	0.580	0.079

across scales, the main discrepancies between the data sets is on the way precipitation is recorded. Daily records are obtained by an observer, which means that low precipitation intensities that may occur during a day probably cannot be measured the following day due to evaporation losses of the gauge. Also the spatial distribution of the two networks is very different. Unfortunately it is not straightforward to identify if the contradictions are due to the different degree of the spatial representation of precipitation from the two networks or due to the inherent problems of each of the measuring techniques.

The Role of Precipitation Clustering

As shown in 3.4.1, the major differences in the scaling structure of precipitation series can be attributed to seasonality, where distinct patterns can be identified. Recently *Molini et al.* [2009] attempted to identified the role of precipitation clustering to the differences in its scaling. According to previous analysis presented here, the β parameter could be one of the indicators of the effects of precipitation clustering

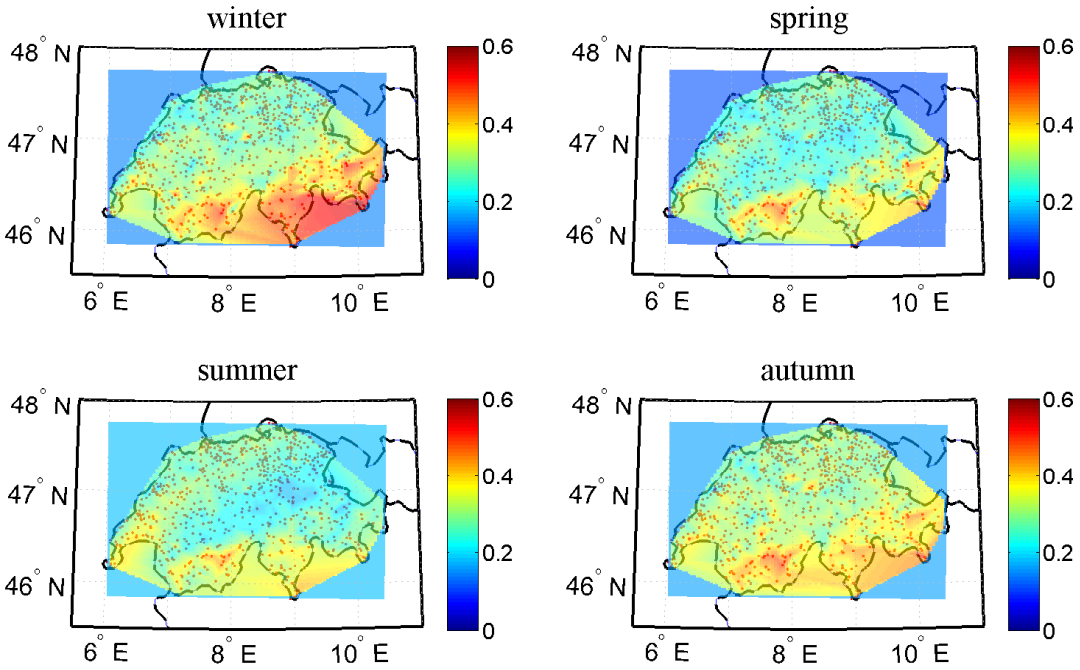


FIGURE 3.27: Spatial distribution of the β parameter derived from the daily observations network. The spatial interpolation is done with a simple linear interpolation technique.

in precipitation series, since it directly captures the structure of the process intermittency. *Molini et al.* [2009] following the ideas of *Sreenivasan and Bershadskii* [2006] adopted a different approach that was first applied to explore the clustering processes of velocity records for turbulent signals, in the sense of separating laminar and turbulent phases. The results of *Molini et al.* [2009] showed that clustering properties do not have an effect on the difference of scaling between the analysed precipitation time series that were used in their study. For this reason the same analysis is applied to the SwissMetNet precipitation records in order to assess the validity of the results in the study area of this thesis.

The first indicator about the effect of the process clustering is the scaling of the running density of the zero crossings.

Let $\theta(T)$ be the number of zeros crossings at the time intervals T and

$$\delta\theta(T) = \theta(T) - <\theta(T)> \tag{3.42}$$

its fluctuations. Then an assumption is that $\delta\theta(T)$ should be a power law function of T i.e.

$$\delta\theta(T) \sim T^{-\phi}. \tag{3.43}$$

A unique ϕ value would be the necessary condition to prove that the differences in the moment scaling function are not due to the clustering properties of precipitation. In principle, ϕ has a direct link to the parameter β.

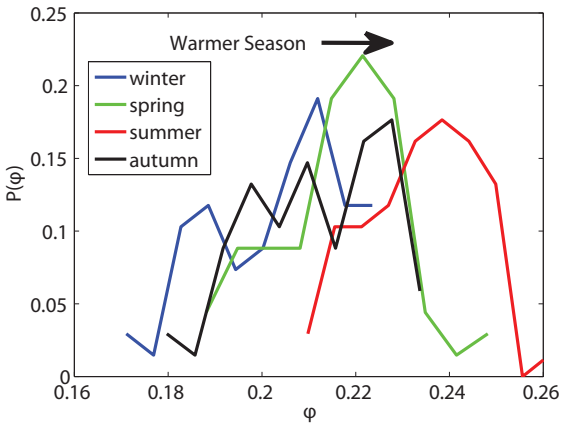


FIGURE 3.28: Histogram of ϕ parameter estimated for temporal scales spanning from 1 hour to 1 day.

As shown in figure 3.28, the estimated ϕ parameter for a range of temporal scales from 1 hour up to 1 day for the Swiss MetNet data records, has a clear connection with the season, a fact that contradicts the results of *Molini et al.* [2009] that precipitation scaling is not affected by the clustering of the process.

3.5 Deviations from Scaling in Data

To this point, the main hypothesis of the analysis was that precipitation data follow a "perfect" scaling behaviour as if arising as an ideal realization of a random multiplicative cascade in time. Recently this

hypothesis has been challenged. First *Fraedrich and Larnder* [1993] identified distinct scaling regimes for rainfall each one connected to a different atmospheric and climatic forcing. Those regimes have been obtained by analysis the power spectra of precipitation time-series. Those results have been confirmed in various studies [e.g. *Olsson and Burlando*, 2002; *De Lima and Grasman*, 1999]. *Marani* [2003] identified as well distinct precipitation scaling regimes, by investigating the scaling of the variance. He also proved that due to the intermittent nature of precipitation, a fine scale regime must exist where $\mathrm{Var}[X_T] \sim T^2$,(where T is the aggregation interval) for $T \to 0$. Similar scaling regimes have been identified as well for various hydrological processes [e.g. streamflow records *Thompson and Katul*, 2012] . In the context of the MRC model, there is a strong evidence that the weights of a MRC are not iid, which is needed in order to have a perfectly multi-scaling process. *Cârsteanu and Foufoula-Georgiou* [1996] and *Paschalis et al.* [2012] found clear correlation patterns on the weights of a MRC for various precipitation time series. *Veneziano et al.* [2006a], *Rupp et al.* [2009], *Serinaldi* [2010a], *Molnar and Burlando* [2005] identified as well clear connection to the distribution parameters of the MRC with precipitation intensities and scales. All those findings yield serious concerns for the validity of the self similarity framework as a universally accepted analysis tool for precipitation, an assumption that has been considered as common ground in hydrology the last decades. In the following sections, each of these deviations are discussed in detail and demonstrated on the MeteoSwiss precipitation data.

3.5.1 Scaling Regimes

First of all a concrete quantification of the issue of the scaling regimes, even though it has been clearly shown to be present for a vast amount of analysed time series, and the scaling breaks have been more or less universal since they are directly linked to the atmospheric processes such as the structure of frontal systems, the diurnal cycle, the seasonal periodicity [*Fraedrich and Larnder*, 1993], has not been provided in the literature. Here a quantification of those regimes is attempted exploiting the data-base of the Swiss MetNet, using spectral analysis.

The scaling regimes that are quantified here, correspond to low scale scaling regimes (<1 month) and are only linked with the storm generation mechanisms such as fronts, convection etc. However, as identified by *Fraedrich and Larnder* [1993] other scaling breaks also exist connected with the interannual variability of precipitation. The analysis is also done on a seasonal basis. The estimation of the spectral densities is achieved both with Fourier decomposition of the precipitation signals and also with wavelet decomposition. In order to smooth the spectral fluctuations, an approach conceptually similar to *Welch* [1967] is applied. The power spectrum of the series is calculated as the average of the normalized power spectra derived for each year. In order to increase also the number of scales estimated with the wavelet estimation of the power spectrum, the continuous power spectrum is used here.

The procedure that was followed in order to identify the scaling regimes is simple and straightforward. First the margins of temporal scales where a break is expected are defined [*Fraedrich and Larnder*, 1993]. For the high frequencies, the margins of expected scaling break are [10 min - 4 hours] and for the low frequencies [8 hours - 4 days]. Linear least square fits were performed on the logarithms of the frequencies and spectral densities for all the possible scaling break combinations and the one that was maximizing the goodness of fit R^2 was selected. An illustrative example can be shown for the Swiss MetNet station located in Zurich, for spring season in figure 3.29(a-b). For the power spectrum derived from the continuous wavelet transform, the lower scaling break was omitted due to very low information for those scales.

The scaling break for the high frequencies, as shown in 3.3.2 can be associated with the measurement

TABLE 3.5: Mean values of the spectral slopes and their respective scaling breaks as identified with Fourier (FFT) and wavelet (CWT) analysis

	Winter (DJF)	Spring (MAM)	Summer (JJA)	Autumn (SON)
FFT based estimates				
ν (FFT) low	-0.430	-0.824	-1.168	-0.808
ν (FFT) middle	-1.070	-0.946	-0.687	-0.962
ν (FFT) high	-0.745	-0.527	-0.474	-0.501
High freq. scaling break [h]	0.669	1.045	1.057	0.886
Low freq. scaling break [h]	18.648	29.227	17.412	31.556
CWT based estimates				
ν low (CWT)	-0.996	-0.811	-0.572	-0.853
ν high (CWT)	-0.512	-0.350	-0.250	-0.322
Scaling Break [h] (CWT)	26.493	37.056	41.176	38.971

precision errors and especially with the signal quantization of the tipping bucket procedure. This gives a direct indicator for which low aggregation scales a hypothesis of scale invariance can be supported without being seriously affected by measurement errors. This leads to the result (see table 3.5) that even though precipitation records exist for sub-hourly temporal scales, it is not a good choice to calculate scaling parameters for those scales. The season for which the high frequency discrepancies are expected to be highly affecting the results is winter (DJF) (see 3.3.2). The fact that for winter the high frequency scaling break is not highly variable and fluctuates around 40 min, is a strong evidence that sub-hourly records cannot be reliably used for scaling parameter estimation. This also justifies the previous choice of temporal scales for the estimation of the MRC parameters.

For both estimation procedures, the low frequency scaling regime is associated with the diurnal cycle. The spectral slopes,have a clear seasonal pattern for the intermediate range of scales ($\sim$1 hour - $\sim$1 day). Summer precipitation is much less correlated than the winter one and leads to lower values of the power law decay of the spectral density. The estimates of derived from the two different procedures yield very similar results (figure 3.30).

Also a regional pattern is as well identified for the scaling regime that corresponds to the sub-daily temporal scales (figure 3.29 (c)). For all the seasons, precipitation in the high altitude alpine areas display a steeper power law spectral decay, that leads to the result that in this area precipitation is highly correlated. Unfortunately, no clear indication of which exact precipitation mechanism lead to those higher correlations is possible, since the topography of the area is very complex, and a distinct quantification of the effect of orographic enhancement of precipitation cannot be done. Still this result can yield some connections with the precipitation mechanisms that affect precipitation and are rather different between each of the areas.

As far as the scaling breaks are concerned, no clear seasonal or regional pattern is observed. The low frequency scaling break is generally higher for the wavelet estimates but since the uncertainty of the estimation is rather high, any solid statement is premature.

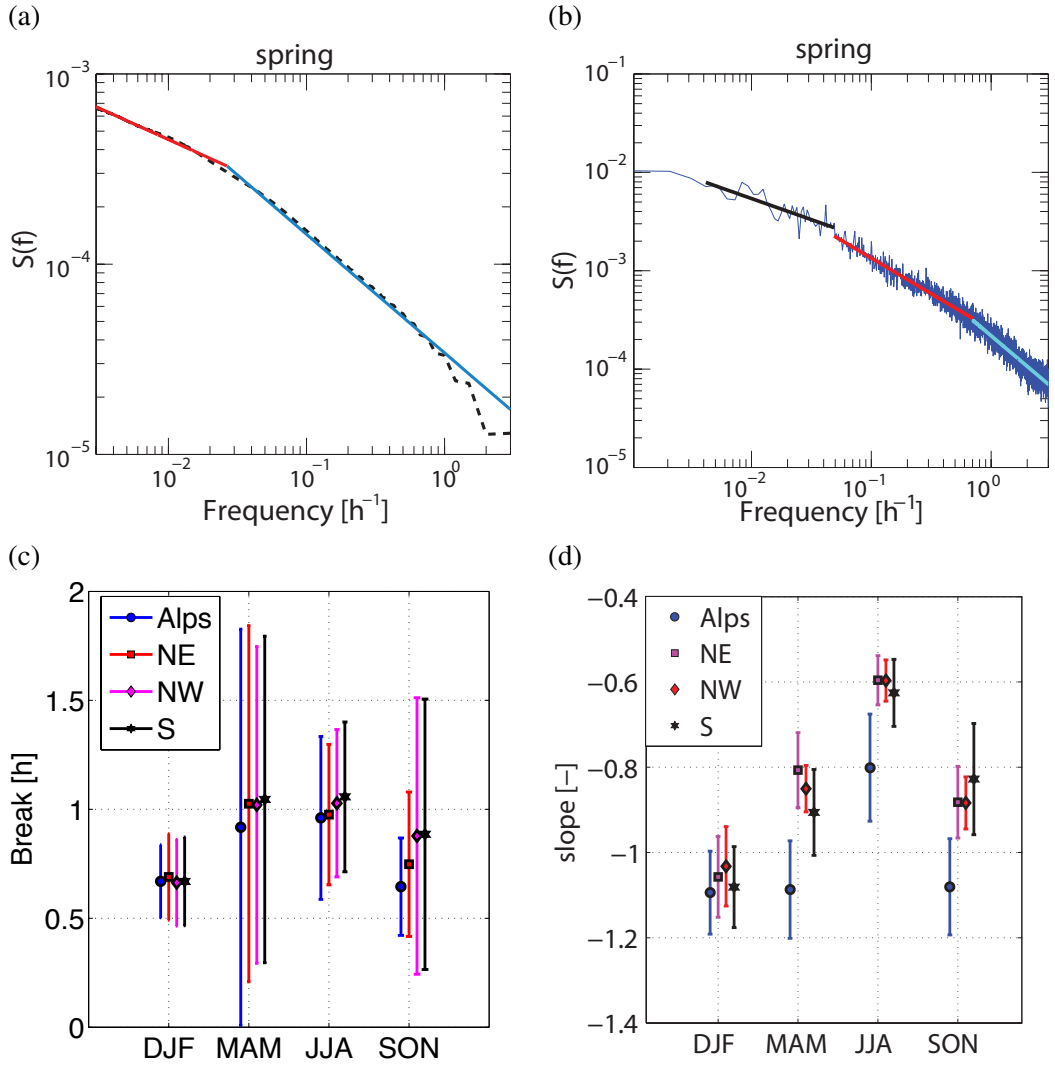


FIGURE 3.29: (a-b) Identification of the distinct scaling breaks for the spring season for the SMA station in Zurich. (a) corresponds to the FFT based spectral estimation and (b) to the wavelet based spectral estimation. The linear fits that maximize the R^2 of the linear fits are shown as well. (c) The low frequency scaling break per climatic region and season as estimated by the FFT spectrum (d) Spectral slopes as identified by the FFT spectrum of the time-series per climatological region. The slopes correspond to roughly the temporal scale regime of $\sim$1 hour - $\sim$1 day dependent on the identification of the scaling breaks.

3.5.2 Correlation of the MRC Weights

A quantification of the possible temporal correlation of the weights (e.g. the cascade generator) of a discrete MRC was first identified by *Cârsteanu and Foufoula-Georgiou* [1996]. Even though the results raised serious concerns about the applicability of the initial forms of MRC as a rainfall modelling tool, they have been somewhat ignored from the hydrological studies. Here an extension of those results is attempted, and the correlation patterns for the extensive record of the SwissMetNet are quantified. The main novelties in comparison to the study of *Cârsteanu and Foufoula-Georgiou* [1996] is that here a generalization of the results in terms of intermittent MRC is done. The results reported here can be found in [*Paschalis et al.*, 2012].

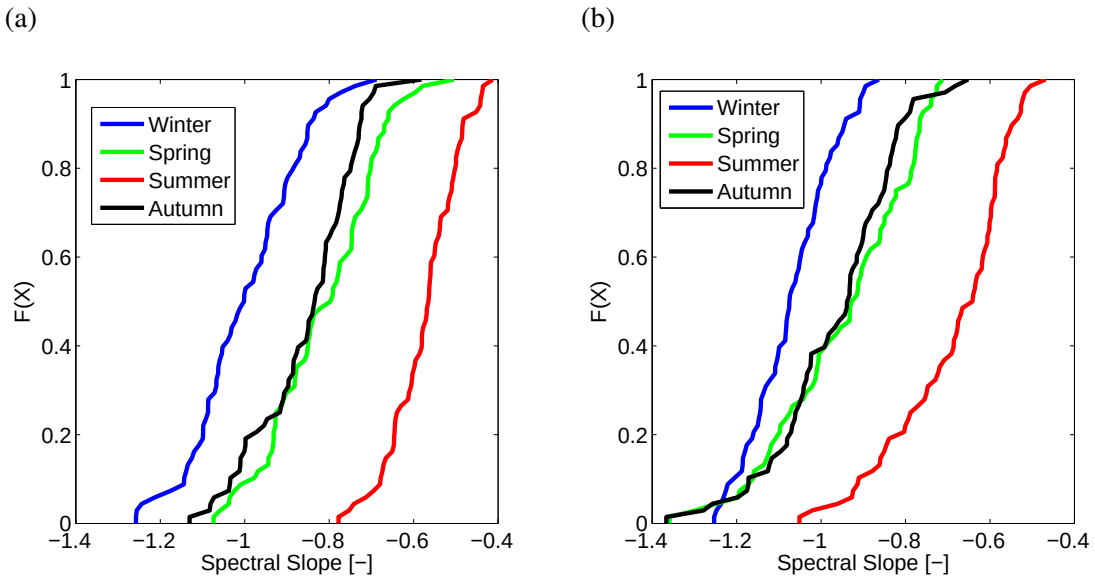


FIGURE 3.30: Cumulative distribution function of the spectral slopes derived from the Swiss MetNet for temporal scales (∼1 hour - ∼1 day). Left panel corresponds to the spectral estimates derived from Fourier analysis and the right panel to the estimates derived using wavelet analysis.

The approach taken is similar to *Cârsteanu and Foufoula-Georgiou* [1996] and consists of building a "toy-model" for the temporal correlation of weights in the MRC in order to asses the principle independency assumptions for the effects of the cascade generators of the MRC. Furthermore the whole dataset of the Swiss MetNet is used in order to asses the assumption of the independence of the MRC weights.

Methods

The temporal dependence structure in precipitation can be thought of as arising from two different processes: the sequence and extent of wet-dry spells and the intra-storm variability in precipitation intensity. To be able to separately quantify the effects of both types of dependence in a MRC a simple stochastic model to generate cascade weights with a known correlation in time due to both intermittency (two state zero-one Markov chain process) and intensity (positive first-order autoregressive process) is developed here. Here no dependence in the cascade weights between scales or any dependence of the distribution of the weights across scales [e.g., *Rupp et al.*, 2009] is taken into account and also the assumption that the processes that control the intermittency and the in-storm variability of the precipitation are independent. In principle, this is the basic assumption as well for the MRC with the atom at zero as described before. Making such an assumption it is possible, in a conceptual sense, to identify the potential impacts of the two separate processes on properties of the simulated time series. The model aims only to serve as an investigation tool (i.e toy-model) and by no means as a complete simulation tool which would require a joint treatment of intermittency and intra-storm variability. Furthermore, the main goal of the model is to reproduce the weight dependencies per scale and not the marginal distributions of precipitation events. More general models, that allow cross-scale dependencies and other general connections have been developed [e.g. *Serinaldi*, 2010a]. Such modelling approaches for temporal precipitation for a complete model construction will be discussed in depth in the next chapter 4.

In this study the canonical as opposed to the microcanonical MRC model framework in this study be-

cause it reproduces the intra-storm temporal variability in precipitation intensity and their extremes [e.g., *Molnar and Burlando*, 2005] and allows for a more convenient definition of independence in weights. Microcanonical models, that preserve the precipitation depth exactly at every branching step, have complementary weights. As a result the microcanonical cascade weight temporal lag one $\rho(1)$ and lag two $\rho(2)$ serial correlation coefficients are uniquely related, $\rho(1) = -(\rho(2) + 1)/2$, and the iid condition has to be defined on the basis of independence between pairs, i.e. $\rho(2) = 0$ and then $\rho(1) = -0.5$ [*Cârsteanu and Foufoula-Georgiou*, 1996]. In canonical models, cascade weights are not complementary but their distribution is restricted to preserve mass on the average in simulation. Temporal independence between weights can then be conveniently defined when $\rho(1) = 0$.

The notation of the oscillation coefficients defined by *Cârsteanu and Foufoula-Georgiou* [1996] is used in order to quantify the presence of dependence in cascade weights and in precipitation time series in general. Oscillation coefficients are a non-parametric measure of patterns of fluctuation in data. A comparison between the oscillation coefficients estimated from precipitation data with those simulated with the MRC model can result to the acceptance of rejection of the independence assumption. One of the main concerns in this study is that the validity of the results in the simulation framework can be affected by the inherent branching structure of the process, i.e. the estimation of the oscillation coefficients may also be affected by the discrete structure of the branching process and not only by the imposed correlation structure. However, the removal of the grid related biases (e.g. by aggregating the simulated series in intervals that does not fall on the branching tree of the cascade [*Over*, 1995, pp. 89]), will in turn change the imposed correlation properties of the weights in the simulated time series. Perhaps a more consistent way would be to build a model in continuous scale, but to my knowledge a model that can redistribute a measure in a multiplicative way across continuous scales and simultaneously produce non rainy spells by construction.

Stochastic Model for Cascade Weights

The simulation scheme that is used here belongs to the class of the discrete canonical random multiplicative cascades with an atom at zero and branching number $b = 2$ [e.g., *Over and Gupta*, 1994] as described in detail in 3.2.2.

Here the same recursive MRC construction procedure as in (3.2.2) is used, however the cascade weights W at each cascade level n are a realization of a composite first order autoregressive process similar to *Chebaane et al.* [1995]. The process is defined as the product of two independent autocorrelated processes. One generates serially correlated binary samples and is used in the cascading procedure to generate wet-dry sequences, and the other is a serially correlated positive process, following a defined marginal distribution which corresponds to within-storm variability in precipitation intensity. This model was selected for its simplicity which leads to few free parameters. A first order autoregressive process was chosen in order to have comparable results with the study of *Cârsteanu and Foufoula-Georgiou* [1996]. The composite process is defined at every level n in the cascade as

$$W_i = c \cdot B_i \cdot Y_i \quad \text{for} \quad i = 1, 2, ..., b^n, \tag{3.44}$$

where B_i is the binary process and Y_i is the continuous non-negative process. The step of the composite process is identical to the subdivision intervals at each scale in the cascade in (3.2.2). The value c is a constant introduced in order to achieve conservation of mass on the average.

The binary part B_i is a first order two state Markov chain model where $B = 0$ or $B = 1$ with the transition probability matrix [*Sponsler*, 1957],

$$(p_{i,j}) = \begin{bmatrix} 1-\alpha\gamma & \alpha\gamma \\ \alpha(1-\gamma) & 1-\alpha(1-\gamma) \end{bmatrix}. \tag{3.45}$$

The mean of the process is

$$E(B) = \gamma, \tag{3.46}$$

the variance is

$$\text{Var}(B) = \gamma(1-\gamma), \tag{3.47}$$

and the lag one autocorrelation coefficient is

$$\rho_B = 1 - \alpha, \tag{3.48}$$

with the restriction that

$$1 \geqslant \gamma + \rho_B(1-\gamma) \geqslant 0. \tag{3.49}$$

The effect of this restriction is, loosely speaking, that for a given mean there is a threshold of autocorrelation that is not accessible. The parameter γ is related to the intermittency parameter β of the binary *Over and Gupta* [1994] model as

$$\gamma = b^{-\beta}, \tag{3.50}$$

and because of mass conservation on the average for the process W_i,

$$c = b^{\beta} = 1/\gamma. \tag{3.51}$$

The continuous part Y_i is an AR(1) process with $E[Y] = 1$ that has a prescribed marginal distribution. Since the oscillation coefficients which are used to quantify correlation are distribution independent,the lognormal distribution was arbitrarily chosen here. The process Y_i is defined as

$$Y_i = \exp[Y_i^*], \tag{3.52}$$

where Y_i^* is a normally distributed AR(1) variable,

$$Y_i^* = \mu(y^*) + \phi_Y[Y_{i-1}^* - \mu(y^*)] + \epsilon_i, \tag{3.53}$$

in which $\mu(y^*)$ is the mean and ϕ_Y the autoregressive coefficient of the process, and ϵ_i is an iid random variable, normally distributed with zero mean and variance

$$\text{Var}(\epsilon) = \text{Var}(Y^*)(1 - \phi_Y^2). \tag{3.54}$$

The unit mean of Y_i implies that

$$E(Y^*) = -\frac{1}{2}\text{Var}(Y^*) \tag{3.55}$$

and the variance of Y_i is

$$\text{Var}(Y) = \exp[\text{Var}(Y^*)] - 1. \tag{3.56}$$

The lag one autocorrelation coefficient of the process Y_i is given by

$$\rho_Y = \frac{\exp[\phi_Y \text{Var}(Y^*)] - 1}{\exp[\text{Var}Y^*)] - 1}. \tag{3.57}$$

Due to the independence of the two processes in (3.44) we finally have for the composite process W_i:

$$E[Y] = cE[B]E[Y] = 1, \tag{3.58}$$

$$\text{Var}[W] = \frac{1}{\gamma}\exp[\text{Var}(Y^*)] - 1, \tag{3.59}$$

$$\rho_W = \frac{[\rho_B(1-\gamma) + \gamma] \cdot [\rho_Y(\exp[\text{Var}(Y^*)] - 1) + 1] - \gamma}{\exp[\text{Var}(Y^*)] - \gamma}. \tag{3.60}$$

In summary, the key parameters that capture the dependence in cascade weights in this model are the lag-one autocorrelation coefficients for the intermittent process ρ_B in (3.48) and for the continuous process ρ_Y in (3.57). The level of intermittency in the data is estimated by β in and variability by Var$[Y^*]$ of the AR(1) process in (3.55). It should be emphasized that all the properties of the compound Markov - AR(1) process hold only within a single scale n of the cascade procedure and not across scales.

Oscillation Coefficients and the iid Assumption

In order to be able to quantify the effect of temporal dependence in cascade weights on precipitation simulated with a MRC model and to compare the results with observations, it is convenient to use oscillation coefficients first introduced for that purpose by *Cârsteanu and Foufoula-Georgiou* [1996]. Oscillation coefficients are the sample probability of total n-tuples in a precipitation record to obey a specific "up-down" pattern. The pattern that provides a descriptor of the correlation structure is the one that two consecutive overlapping 2-tuples (pairs) have a different trend, i.e.

$$\begin{array}{lll} & r_i > r_{i+1} \quad \text{and} \quad r_{i+1} < r_{i+2} \\ \text{or} & r_i < r_{i+1} \quad \text{and} \quad r_{i+1} > r_{i+2}, \end{array} \tag{3.61}$$

where r_i is rainfall depth at time i. The frequency of this fluctuating pattern can be estimated by computing

$$a_i = (r_i - r_{i+1}) \cdot (r_{i+1} - r_{i+2}), \qquad i \in [1, m-2], \tag{3.62}$$

for the entire record length m and defining the oscillation coefficient as

$$C_{\updownarrow} = \frac{N(a<0)}{N(a \neq 0)}, \tag{3.63}$$

where N is the number of cases for the conditions in a. Note that for intermittent precipitation data a no rain pair $(a = 0)$ does not contribute to 3.63.

Cârsteanu and Foufoula-Georgiou [1996] showed that the oscillation coefficient $C_{\updownarrow}$ is scale invariant when applied to aggregated precipitation data (from seven analysed storms), it can be used to detect the presence of noise in data compared to the MRC, and is also independent of the distribution of the cascade weights. In a microcanonical MRC model, they showed by simulation that $C_{\updownarrow}$ is related to the lag two correlation coefficient of the weights and for independent weights $\rho(2) = 0$ the coefficient $C_{\updownarrow} = 0.6$. Here the oscillation coefficient is used for three purposes.

(1) Relations between parameters of the cascade weight model and $C_{\updownarrow}$ by Monte Carlo analysis are developed. The main focus is on the correlation coefficients $\rho_B = 1 - \alpha$ and ϕ_Y and the extent of dry spells expressed in terms of the intermittency coefficient β. Several samples with the MRC model with the key parameters belonging within plausible ranges which were identified from data, were simulated. The two cases of simulations were

$$\phi = 0 \quad and \quad \begin{cases} \text{Var}[Y^*] \in [0.05, 0.5] \\ \beta \in [0.05, 0.5] \\ \rho_B \in [-0.9, 0.9] \end{cases} \tag{3.64}$$

and

$$\rho_B = 0 \quad and \quad \begin{cases} \text{Var}[Y^*] \in [0.05, 0.5] \\ \beta \in [0.05, 0.5] \\ \phi_Y \in [-0.9, 0.9] \end{cases} \tag{3.65}$$

The number of steps in the cascade in each simulation was chosen to match the sample size corresponding to a single climatological season (3 months) of data at a given temporal resolution. For hourly data this was a sample size of 2048, i.e. $n = 11$ levels in the random cascade with $b = 2$. For each combination in the parameter space we simulated 1000 realizations and calculated the numerical approximation of the 90% confidence intervals. The 90% confidence intervals for each value of the intermittency coefficient β are estimated as follows. For every intermittency coefficient β, we numerically estimate the 90% intervals for an ensemble of the correlation parameters belonging in the parameter space defined in 3.64 and 3.65, depending on the kind of introduced correlation. The upper and the lower 90% confidence bound from the ensemble were as well estimated. The parameter space of the correlation coefficients is explored using a step of 0.05 for all coefficients.

Sub-grid variability was also taken into account by approximating the limit measure Z_∞ by "dressing" the cascade for 5 more cascade steps.

(2) The confidence bounds for $C_\updownarrow$ under the iid assumption of cascade weights were estimated as a function of the intermittency parameter β and compared with the oscillation coefficients estimated from the station data at the hourly temporal resolution. To be comparable with simulations, each season of every complete year of precipitation measurements is considered an independent sample. In addition to evaluating the iid assumption for the weights the confidence bounds for $C_\updownarrow$ for two limiting cases of correlation were estimated. The first case assumes that the temporal dependence structure in the weights is only the result of serial correlation in the wet-dry sequence, and the second case assumes that it is only dependent on the correlation strength in the continuous process. The reason to investigate only those limiting cases is that they distinguish between the processes of clustering and within-storm variability. Simultaneous correlation of the zero-one Markov process and the positive AR(1) process is possible but it would add another degree of freedom to the analysis, and was not investigated here.

(3) Finally the scale-invariance of $C_\updownarrow$ on precipitation data below the 1 hour resolution (down to the 10-min sampling resolution of the gauges) was quantified. The main goal here was to identify the effect of the gauge measurement mechanism by comparing a tipping-bucket and weighing gauge at high resolutions, and its possible relation to precipitation type (see 3.5.2).

Dependence Structure in Cascade Weights and $C_\updownarrow$

The oscillation coefficient $C_\updownarrow$ was found to be strongly related to the correlation structure in the cascade weights and intermittency. (Fig 3.31 a-b), shows that $C_\updownarrow$ generally increases with intermittency β and decreases with correlation in both the continuous (Fig 3.31 a) or the binary (Fig 3.31 b) part of the cascade weight model.

As an example, for the non-intermittent case $\beta = 0$ (Fig 3.31 c-d), $C_\updownarrow$ decreases with increasing correlation in the continuous process ϕ_Y with very low uncertainty. As expected, more persistence in the cascade weights leads to smoother simulated precipitation profiles, while anti-persistence increases roughness

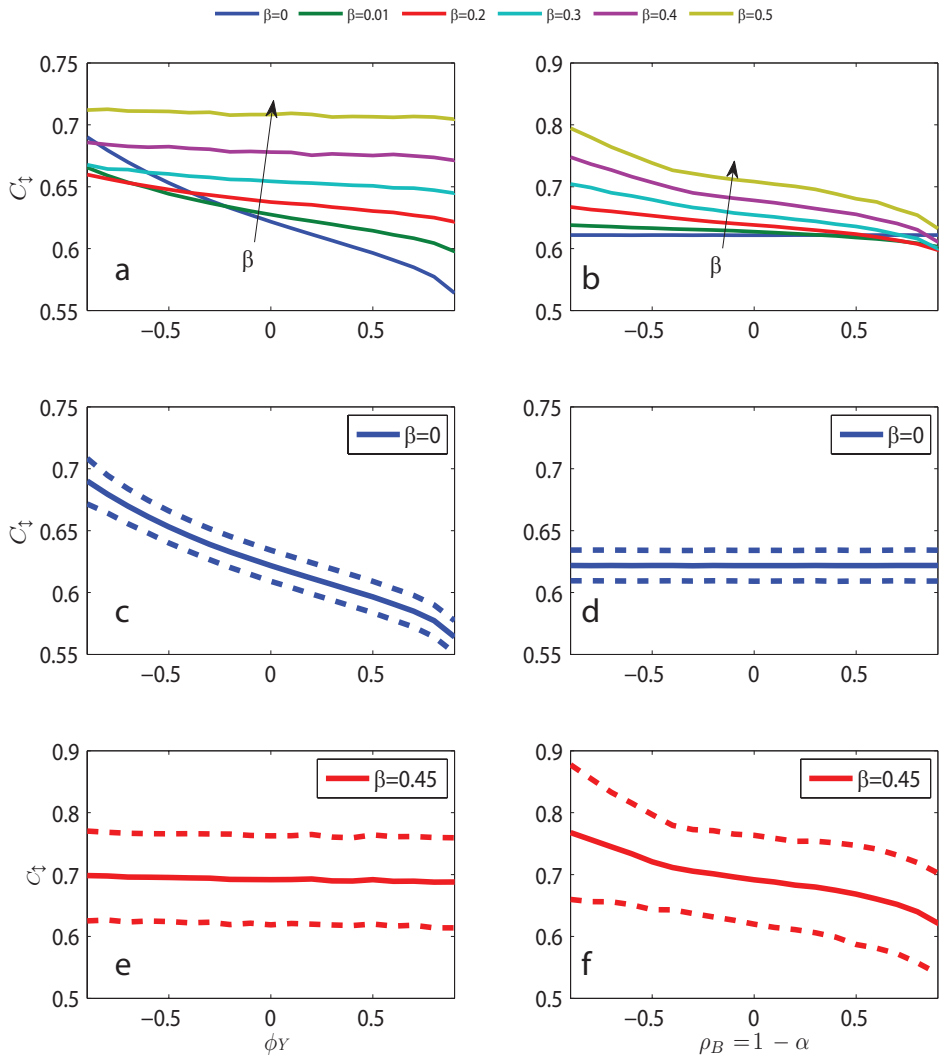


FIGURE 3.31: The relations between the oscillation coefficient and the correlation coefficients for the binary process ($\rho_B = 1 - \alpha$) and the logarithm of the positive process (ϕ), respectively. Panels (a) and (b) show the mean $C_{\updownarrow}$ as a function of a range of values for the intermittency coefficient β from 0 to 0.5. Panels (c) and (d) show the mean and 90% confidence intervals for $C_{\updownarrow}$ vs $1 - \alpha$ and ϕ_Y for the case of non intermittent data, $\beta = 0$. Panels (e) and (f) shows the same for the case of highly intermittent data, $\beta = 0.45$.

(fluctuations). Since $C_{\updownarrow}$ is a measure of fluctuations, smoother (rougher) fields lead to lower (higher) values of $C_{\updownarrow}$. For independent weights $\phi_Y = 0$, the mean oscillation coefficient $C_{\updownarrow}$ was found to be slightly larger than 0.6. This discrepancy with *Cârsteanu and Foufoula-Georgiou* [1996] for the microcanonical MRC model is probably caused by the the effect of subgrid variability which was not considered in their framework. Naturally, the correlation coefficient of the intermittent process ρ_B has no effect on $C_{\updownarrow}$ in the case when $\beta = 0$.

For the highly intermittent case $\beta = 0.45$ (Fig 3.31 e-f), $C_{\updownarrow}$ is generally higher than for the non-intermittent case and decreases with increasing correlation in both continuous and intermittent processes. Intermittency increases the roughness of the generated precipitation profiles, even for the case of independent weights, however uncertainty is also very large.

Comparison with Observations

In order to distinguish different dominant precipitation mechanisms and influences, the study domain was divided into the four climatic regions (see Fig 2.2) similarly to [*Molnar and Burlando*, 2008]. Furthermore, precipitation records were examined on a seasonal basis. Also for the time periods where temperature records are available at a given station, the difference between solid and liquid precipitation was explored by adding a temperature threshold criterion. Events during which the hourly temperature $T \leq 0°C$ are considered as snowfall events and in the opposite case ($T > 0°C$) as rainfall events.

In Figure 3.32 the estimated $C_{\updownarrow}$ computed from the data at the hourly time scale, plotted as a function of the intermittency coefficient β for the different climatic regions of Switzerland, together with the simulated mean iid case and 90% confidence intervals and the lower 90% confidence bounds for our model with correlation in the intermittent and continuous parts are shown for the alpine area. The time series ensemble was derived from the entire record considering each season of every station and year as an independent sample.

The first result is that for all climatic regions and seasons, with high probability, the random multiplicative cascade framework with iid weights fails to describe satisfactorily the observed natural behaviour in precipitation. Furthermore, for all regions the same pattern of dependence was observed, whereby $C_{\updownarrow}$ in data was generally lower than the mean iid case, meaning that natural precipitation was smoother than that simulated by the MRC iid model. A weak seasonal pattern was also observed. Warm seasons with possibility of convective precipitation show a weaker correlation structure and higher $C_{\updownarrow}$ closer to the iid case. The winter period in which a large fraction of precipitation falls in solid form, especially in the Alpine region, exhibits on average the strongest dependence, low $C_{\updownarrow}$ and greater deviations from the iid case (Table 3.6), and also the largest variation between the samples (Figure 3.33). However the differences due to seasonality are rather small and statistically not significant. Table 3.6 summarizes the statistics of $C_{\updownarrow}$ and its deviation from the iid case for all the regions and precipitation type. The deviation from the iid case was defined as the distance between the estimated $C_{\updownarrow}$ from the data and the one for the iid case for the same intermittency coefficient β. Overall precipitation type also does not have a very significant effect on $C_{\updownarrow}$, except that solid precipitation in winter is generally closer to the iid case than liquid precipitation, which could be attributed to generally lower intensities and the filtering effect of the gauge in measuring snowfall.

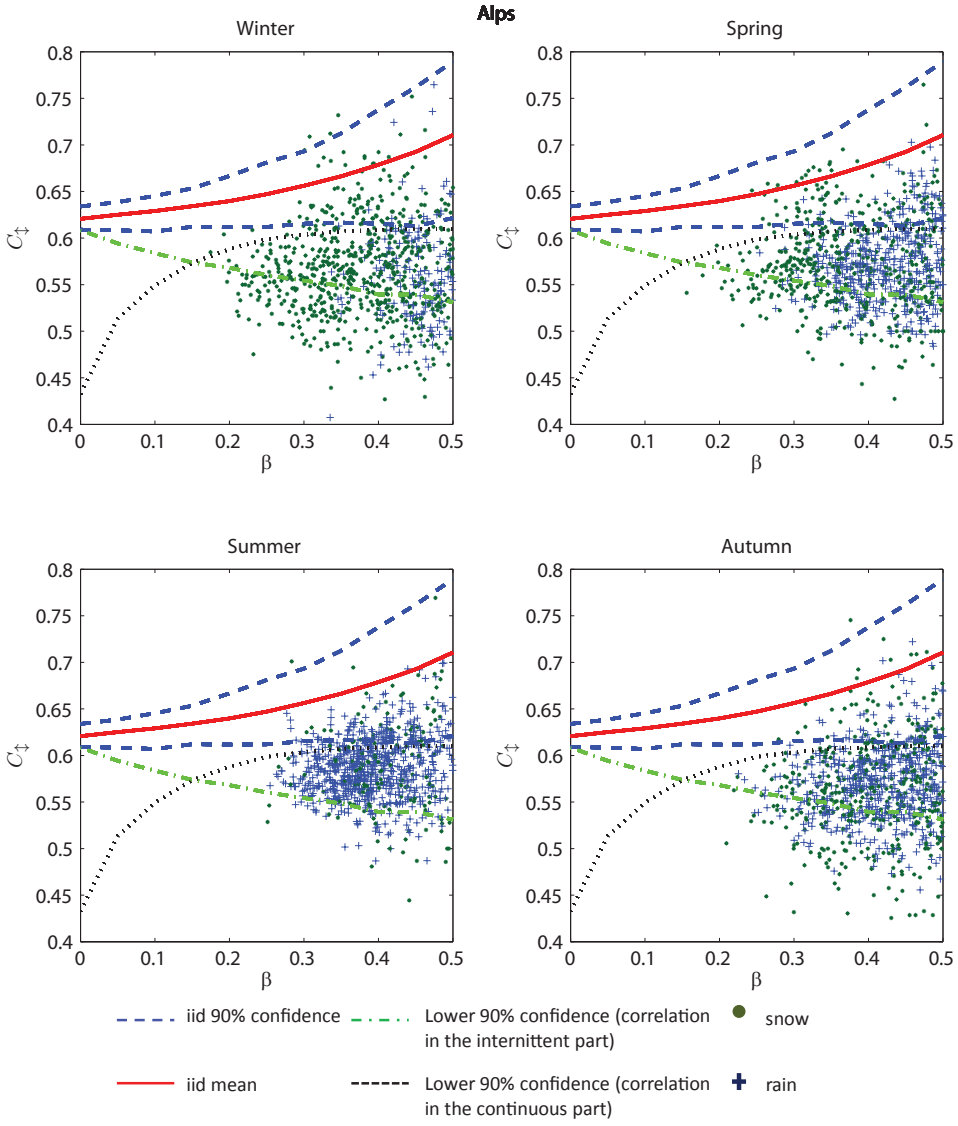


FIGURE 3.32: $C_\updownarrow$ versus the intermittency coefficient β for the Alpine region estimated for the hourly scale. The red solid line corresponds to the mean value for the iid case. The dashed blue lines correspond to the 90% confidence intervals of the iid case. The lower 90% confidence bounds for the correlated intermittent process (green dash-dot) and the continuous process (black dot) are also shown. As lower confidence bound we define the lowest numerically estimated 90% quantile for the respective intermittency coefficient, and the rest of the parameters varying according to (3.64–3.65). For precipitation observations β was estimated from the moment scaling spectrum $\tau(q)$ defined as the slope of the scaling of $q-$order statistical moments [*Over and Gupta*, 1994]. The moment scaling spectrum was calculated for scales spanning from one hour up to 2^7 hours. The value β was estimated from $\beta = 1 - \tau(0)$. Seasonal series for which the estimated value of β was greater than 0.5 were excluded from our analysis due to very low data availability since almost the entire record consisted of zero values. Blue crosses correspond to rainfall events, green dots correspond to snowfall events.

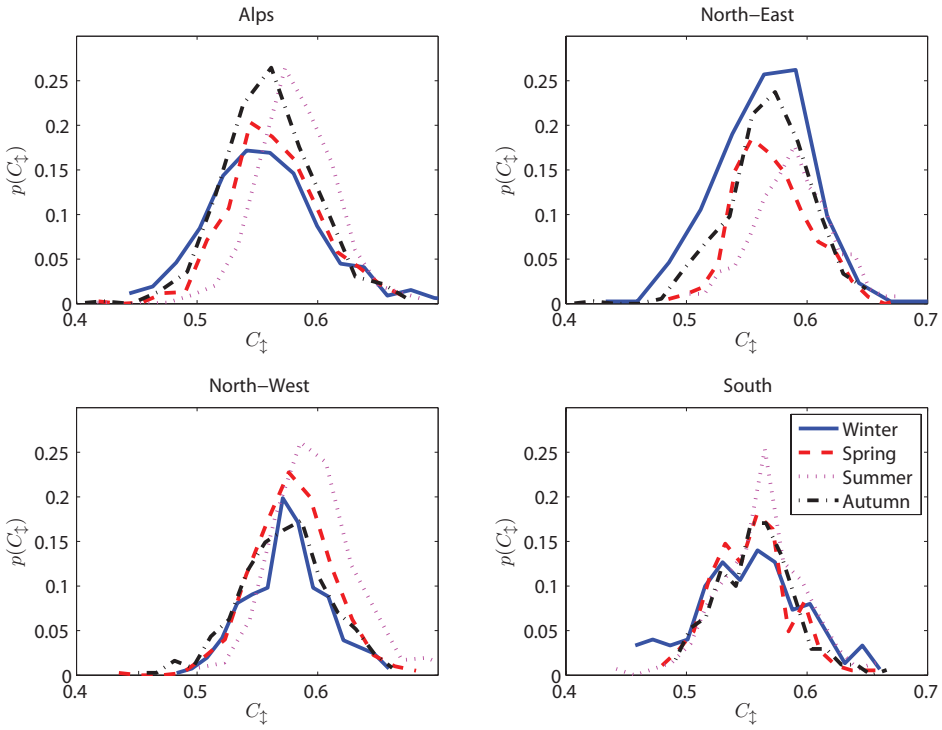


FIGURE 3.33: Experimental histogram of $C_\updownarrow$ for every climatic region, for each season.

The second result is that the adopted structure in the stochastic model for temporally correlated cascade weights cannot fully explain the data-derived oscillation coefficients. Although a proportion of data points lie above the 90% lower confidence bound which means they could be explained by clustering effects or within-storm correlation, there are also significant departures (e.g. Figure 3.32). Correlation in the continuous process explains more of the variability in $C_{\updownarrow}$ than correlation in the binary process. Compared to the model, some observations show very low values of $C_{\updownarrow}$, indicating very smooth precipitation profiles. There are several possible explanations for this. For example, clustering and within-storm correlation could act together, a case that was not explored here, or a different nonlinear form of the dependence structure could be more appropriate, for example including correlation between scales. Another possible explanation is that the measurement resolution of the tipping-bucket gauges used in this analysis produces a quantization of the precipitation which affects the results.

TABLE 3.6: Statistics of the oscillation coefficients and deviations from the iid case for each region and season based on hourly precipitation data. Non available data are marked with dash.

Oscillation coefficients

	Mean				Standard deviation			
	Winter	Spring	Summer	Autumn	Winter	Spring	Summer	Autumn
All precipitation								
Alps	0.556	0.564	0.582	0.559	0.047	0.038	0.032	0.039
NE	0.565	0.571	0.588	0.568	0.041	0.032	0.033	0.036
NW	0.575	0.579	0.597	0.573	0.032	0.033	0.036	0.036
S	0.553	0.554	0.562	0.557	0.045	0.033	0.036	0.032
Rainfall								
Alps	0.548	0.567	0.585	0.567	0.056	0.045	0.034	0.042
NE	0.570	0.580	0.589	0.574	0.041	0.034	0.033	0.038
NW	0.575	0.585	0.598	0.575	0.042	0.036	0.038	0.040
S	0.562	0.560	0.565	0.559	0.052	0.037	0.038	0.034
Snowfall								
Alps	0.569	0.584	0.604	0.574	0.053	0.054	0.064	0.067
NE	0.584	0.580	-	0.561	0.072	0.058	-	0.080
NW	0.598	0.605	-	0.589	0.050	0.078	-	0.080
S	0.560	0.573	-	0.562	0.071	0.077	-	0.053
Deviation of oscillation coefficients from average iid								
All precipitation								
Alps	-0.115	-0.106	-0.091	-0.115	0.051	0.043	0.034	0.041
NE	-0.104	-0.099	-0.088	-0.107	0.040	0.031	0.034	0.038
NW	-0.090	-0.091	-0.084	-0.099	0.034	0.036	0.037	0.036
S	-0.124	-0.117	-0.126	-0.115	0.049	0.034	0.037	0.033
Rainfall								
Alps	-0.147	-0.121	-0.095	-0.119	0.057	0.045	0.035	0.044
NE	-0.119	-0.100	-0.087	-0.106	0.043	0.034	0.035	0.042
NW	-0.110	-0.098	-0.085	-0.105	0.045	0.037	0.038	0.041
S	-0.129	-0.117	-0.125	-0.115	0.052	0.036	0.040	0.037
Snowfall								
Alps	-0.108	-0.091	-0.085	-0.108	0.055	0.056	0.063	0.066
NE	-0.099	-0.097	-	-0.123	0.069	0.050	-	0.077
NW	-0.082	-0.076	-	-0.097	0.050	0.078	-	0.080
S	-0.126	-0.112	-	-0.123	0.071	0.078	-	0.056

Measurement Effects on $C_{\updownarrow}$ and Scale-Invariance

Measurement artefacts are potentially very critical also for the correlation analysis of high resolution precipitation data. To illustrate this effect a comparison between the oscillation coefficient estimated for the two precipitation records used previously 3.3.2 was done. It has been shown that the tipping bucket measurement procedure leads to a bias in the intermittency of the recorded process [*Molini et al.*, 2001] and thus can have a significant effect on the estimation of $C_{\updownarrow}$. The comparison was done for a 2.5 month period in 2009 (22 January - 2 April). The data were aggregated in the range from 10–90 minutes and $C_{\updownarrow}$ at every aggregation scale was plotted in Figure 3.34.

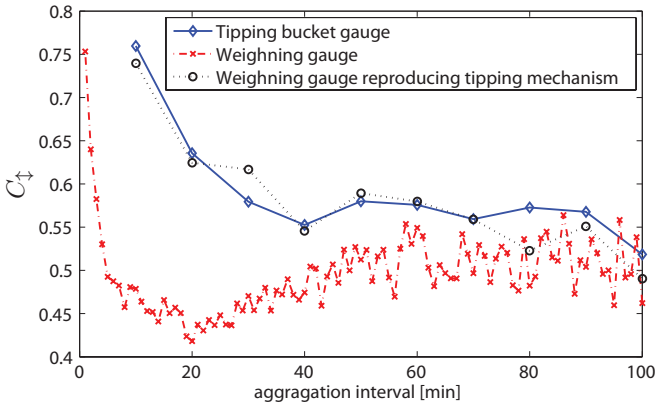


FIGURE 3.34: Comparison between the estimated $C_{\updownarrow}$ for the high precision electronic weighing gauge and the lower resolution tipping-bucket gauge which is installed at all SwissMetNet stations. The comparison was conducted at the Zermatt gauge for 2.5 months in 2009.

The oscillation coefficient for the electronic weighing gauge operating at 1-min resolution started at very high values, which are connected to known noise problems of the device at the 1 min scale, then had consistently lower values than the ones calculated for the tipping bucket and converged to a constant average value $C_{\updownarrow} \approx 0.5$ after about 50 minutes. On the other hand, the tipping-bucket gauge had very high oscillations at high resolutions and converged after about 40 minutes to $C_{\updownarrow} \approx 0.55$. By converting the weighing gauge data to equivalent tipping bucket data by computing the number of 0.1 mm tips in each 10 minute period, it can be shown that this significant discrepancy in high resolution oscillations was caused entirely by the tipping-bucket mechanism (Figure 3.34). This means that the tipping-bucket measurement procedure which integrates precipitation intensity into a multiple of the tip volume at the given time sampling resolution generates a very quantized and highly intermittent record at high resolutions. This is a major deficiency of this particular measurement device and a warning for the use of the 10-min tipping-bucket precipitation data for correlation analysis.

To further investigate the scale-dependence of $C_{\updownarrow}$ from the tipping-bucket data, the mean seasonal $C_{\updownarrow}$ for a range of aggregation levels from 10–100 minutes is shown in figure 3.35 for all stations in the SwissMetNet network. This plot clearly shows that the calculated $C_{\updownarrow}$ is scale-dependent in all seasons except summer for aggregation scales less than about one hour. Based on the results of *Savina et al.* [2012] probably the large departure from scale-invariance in cold seasons is due to the snow melting lag caused by the heating mechanism of the tipping bucket gauge and generally lower precipitation intensities. At aggregation times above 1 hour, the quantization effects are insignificant, and the oscillation coefficient can be considered scale-invariant. In another study of the variance of aggregated precipitation for scales

ranging from 10 minutes up to one day, it was found that precipitation is reasonably scale invariant at least in a specific range of scales above 1 hour which are commonly used in hydrological applications [*Marani*, 2003].

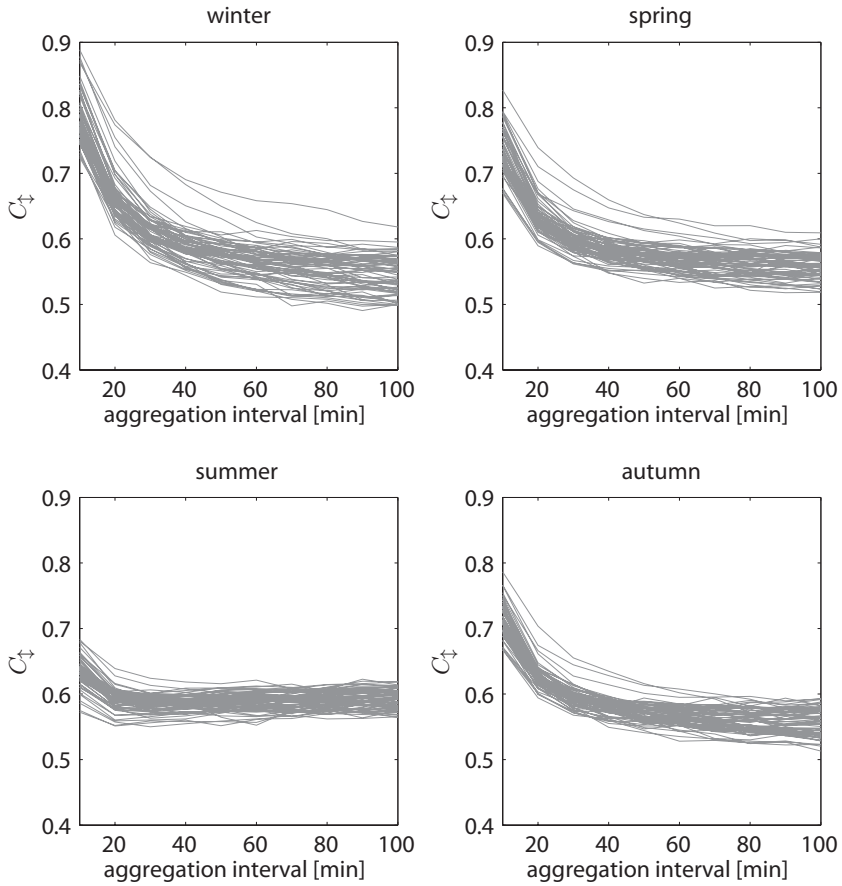


FIGURE 3.35: Effect of the low precision tipping-bucket mechanism on the estimation of $C_{\updownarrow}$ across different aggregation levels. Each line of the graph corresponds to the mean value of $C_{\updownarrow}$ estimated at a different aggregation interval over all years for all the stations.

3.5.3 Imperfect Scaling

Another kind of deviation from the classical MRC phenomenology is the dependency between the distribution of the MRC generator W and the scale or intensity of the series. Such connections have been well established in previous studies [e.g. *Molnar and Burlando*, 2005; *Over and Gupta*, 1994; *Rupp et al.*, 2009; *Veneziano et al.*, 2006a; *Pathirana and Herath*, 2002]. The most common dependence identified by all the studies is the direct link of the β parameter to the average precipitation intensity. Conceptually this fact is easy to understand since, as expected, large precipitation depth accumulations occur when precipitation events are more frequent. On the other hand the parameter(s) that controls the within-storm variability of precipitation, does not always have a clear connection with precipitation intensity. First *Veneziano et al.* [2006a] identified that distinct patters of dependencies between the generator W parameters and both aggregation scale and intensity can be identified for several precipitation datasets. The methodology that is followed is in principle identical to *Veneziano et al.* [2006a] with the main difference that here the beta-lognormal MRC instead of the beta log-Levy model is used. The aim is to illustrate the existence of such relationships between the weights of the MRC and the scale an intensity. This property will be as well later used for the construction of precipitation stochastic models for temporal precipitation in chapter 4.

The MKP function of a MRC can be defined as in equation 3.14. If the distribution W is allowed to be dependent on the measure density of the previous aggregation scale and also the scale itself, a new MKP function can be defined as

$$\chi_b(q|n, X_{n+1}) = \log_b EW(n, X_{n+1})^q - (q-1), \tag{3.66}$$

then an estimator of $\chi_b(q|n, X_{n+1})$ can be defined as

$$\hat{\tau}(q|n, X_{n+1}) = \log_b(\hat{E}[X_n|X_{n+1}]) - \log_b(X_{n+1}) - (q-1), \tag{3.67}$$

taking into account that

$$W = X_n/X_{n-1}, \tag{3.68}$$

where X_n is the measure density defined at scale n (i.e. the precipitation intensity and not depth). From $\hat{\tau}(q|n, X_{n+1})$, the estimation of the parameters $\beta(n, X_{n+1})$ and $\sigma^2(n, X_{n+1})$ is straightforward and is achieved as previously explained.

As noted by *Veneziano et al.* [2006a] the ratio defined in equation 3.68 reflects on the bare densities, and the estimation of $\hat{\tau}(q)$ will be a biased estimation of $\chi_b(q)$. In any case, here the biased $\hat{\tau}(q)$ is reported which is a clear indicator about the dependencies of the distribution of W. The unbiased estimation could be achieved with an iterative procedure that can yield serious computational restrictions due to the amount of data, and would be useful in case the true parameters would be needed in order to be used in a simulation framework. Here the goal is to identify the general tendencies and the biased estimates are sufficient.

Two main problems of the estimation of the parameter per intensity class and scale have to be taken into account. First for low aggregation scales and low intensities, due to the finite sampling volume of the

recording mechanism (tipping bucket), the estimates of W as the ratio between the intensities of two consecutive classes can be highly affected. This is due to the small number of possible values that the ratio can take. Second, for high intensity classes, due to the scarcity of high intensity events, the sample is rather small.

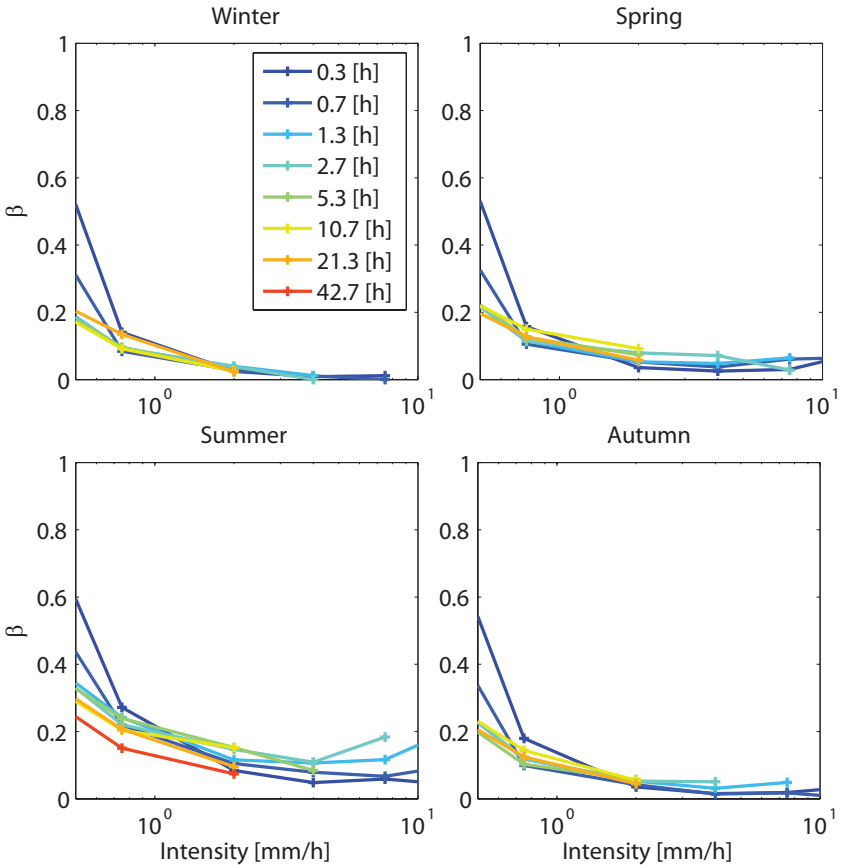


FIGURE 3.36: Dependence of the β parameter of the beta-lognormal MRC model on the scale of aggregation and the intensity of the previous scale for the SMA station in a seasonal basis

In figures 3.36 and 3.37 an illustrative example is shown for one station located in Zurich (Switzerland). The results for all the stations yield similar dependence patterns and thus only one station is reported here. An interesting result is that the β parameter is highly dependent on the intensity but rather independent of the scale. This result agrees with the finding of *Veneziano et al.* [2006a]. No significant seasonal differences are as well identified.

On the other hand the σ^2 parameter depends both on intensity and aggregation scale. This issue can partially give an explanation why mainly studies that looked connections between the distribution of the MRC generator and only intensity did not yield consistent results, since they were neglecting an additional degree of freedom that should taken into account. The σ^2 parameter generally increases with the aggregation scale. The identified dependence pattern on the intensity is not always unique and certainly not trivial, but is consistent between the temporal scales. Further assessment especially of σ^2 would be statistically insignificant due to the small sample size per aggregation scale and intensity class, and also taking into account the low robustness of the estimation procedure, as has already been investigated.

Another major problem that can affect the validity of these results is that equation 3.67 holds true only

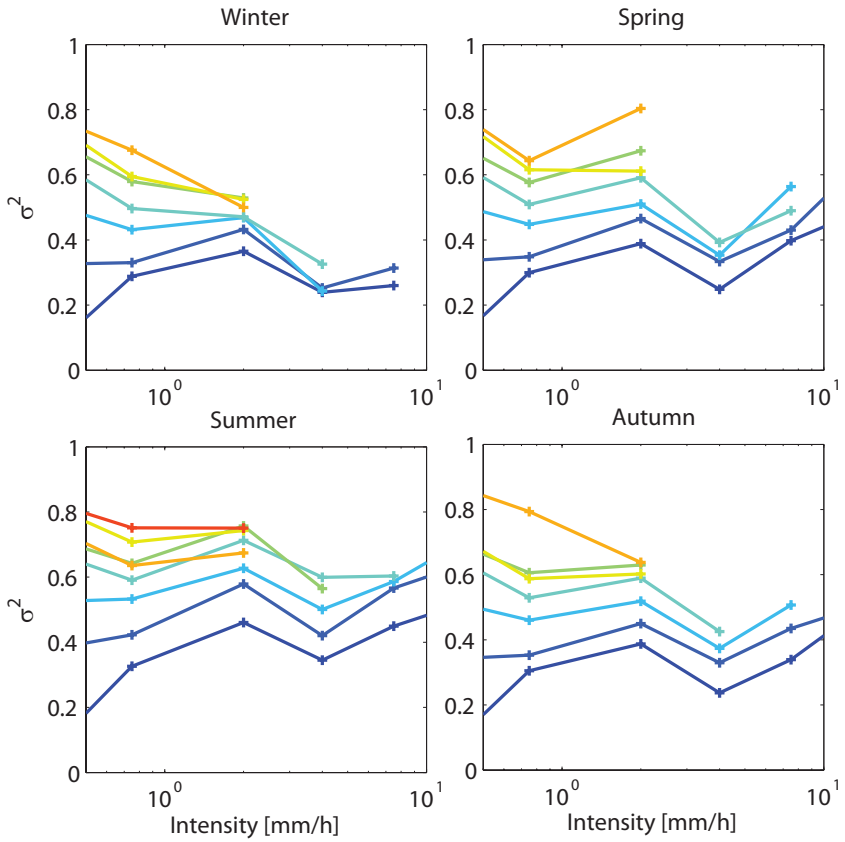


FIGURE 3.37: Dependence of the σ^2 parameter of the beta-lognormal MRC model on the scale of aggregation and the intensity of the previous scale for the Zurich station on a seasonal basis. Colour legend is the same as in figure 3.36

for very fine discretization of the intensity classes. This leads either to a poor approximation of the scale and intensity dependent MKP function is coarse discretisation or to results of low statistical significance if fine scale intensity classes are used that lead to very small sample sizes per intensity bin.

In any case, it appears that strong dependencies between the probability distribution of W and both scale and intensity hold true. In the following chapter, such parametrizations of such connections will be used for the construction of efficient stochastic modelling tools of temporal precipitation.

4

MODELLING THE TEMPORAL VARIABILITY OF PRECIPITATION

The lack of long and reliable precipitation time series has been always a problem in hydrology and the most common way to overcome this issue is by stochastic simulation. In this chapter, following an extensive data analysis of temporal precipitation, an investigation of several stochastic modelling approaches for temporal precipitation is presented. The models consist of both traditional models of rainfall and new developed ones. A thorough model inter-comparison is conducted and useful results concerning their applicability and potential universality are derived. The ultimate goal is to identify where there are still opportunities on improving stochastic models of precipitation and provide new modelling techniques. Following up chapter 3, special attention is given in this chapter on the abilities of the various stochastic models to simulate the statistical statistics of precipitation across temporal scales.

4.1 Introduction

Stochastic modelling of precipitation has always been one of the most challenging topics in hydrology. The reason is that despite the fact that stochastic modelling approaches had been already developed and applied in various scientific fields, a description of precipitation as a stochastic process is not straightforward. In contrast to other geophysical signals, precipitation has some distinct features that have to be all taken into account, without overparametrizing the structure of the models, in order to be generic. Among those features, the most important ones are the process's intermittent nature, its highly skewed probability distributions, its strong correlations and possibly long memory. Moreover, the correct description of those features across the entire range of temporal scales of hydrological interest is of major importance. Taking all these into account, it becomes clear that developing models that can simultaneously capture all these properties and at the same time are parsimonious in their structure is a very challenging topic.

The history of stochastic modelling of precipitation spans over 4 decades. The first approaches mainly focused on the simulation at the daily time scale. The reason for that is the lack of availability of high resolution reliable and long precipitation records and the also that early hydrological models were usually built to operate at the daily time scale. Among the most famous processes adopted in hydrology for stochastic simulation of precipitation were based on Markov chains [*Todorovic and Yevjevich*, 1969; *Todorovic and Woolhiser*, 1975; *Chin*, 1977; *Katz*, 1977; *Foufoula-Georgiou and Lettenmaier*, 1987] and

the alternating renewal process [*Schmitt et al.*, 1998; *Ivanov et al.*, 2007]. Those processes were used for precipitation simulation in some of the most widely used weather generators such as the LARS-WG and WGEN [e.g. *Richardson*, 1981; *Racsko et al.*, 1991]. However, these models were incapable of reproducing high order precipitation statistics and their simulation was usually not adequate for a wide range of precipitation scales. In principle, they were not designed for temporal scales finer than daily despite some exceptions [e.g. *Katz and Parlange*, 1995].

In order to overcome the deficiencies of these modelling techniques, and especially their inability of simulating sub-daily precipitation, new modelling approaches have been introduced. Among them, point processes based on the early work of *LeCam* [1961] and *Cox and Isham* [1988] have been proven very promising. Those processes appeared as a very good candidate for precipitation modelling, since they reproduce a structure of storm arrivals and persistence that mimic, in a conceptual sense, the physics of the processes. Extensions on the initial ideas led to the development of the Poisson cluster models [e.g. *Kavvas and Delleur*, 1981; *Ramirez and Bras*, 1985; *Rodriguez-Iturbe et al.*, 1987, 1988] that introduced the in-storm clustering properties of precipitation as well. This approach became a very popular modelling technique [e.g. *Entekhabi et al.*, 1989; *Cowpertwait*, 1991, 1994; *Cowpertwait et al.*, 1996; *Cowpertwait*, 1998; *Onof et al.*, 2000; *Burton et al.*, 2008, among others]. These models have recently been introduced into detailed, fine temporal scale (hourly), weather generators [e.g. *Fatichi et al.*, 2011]. Their wide applicability can be justified by the fact that they have been also extensively used in real world applications [e.g. *Onof and Wheater*, 1993; *Burlando and Rosso*, 2002]. Lately, non-stationary versions, with time dependent parameters, have been developed for the Poisson cluster models in order to account for climate change [*Evin and Favre*, 2012].

After the recognition that precipitation processes exhibit self similar behaviour (see chapter 3) and can be described as simple/multi-scaling processes, a large effort has been devoted to exploit those properties for stochastic modelling of precipitation [e.g., *Koutsoyiannis et al.*, 2011; *Lombardo et al.*, 2012; *Rebora et al.*, 2006; *Ferraris et al.*, 2003; *Over and Gupta*, 1994]. One of the most significant achievements of the application of the theories of scale invariance for temporal modelling, has been the very good representation of the fine scale properties of rainfall (e.g. 5-10 min). For this reason the most common application of the use of scaling processes, and especially multifractals with the use of discrete MRC, has been rainfall downscaling and disaggregation to fine temporal scales [*Molnar and Burlando*, 2005; *Serinaldi*, 2010b; *Rupp et al.*, 2009; *Güntner et al.*, 2001]. The use of simple scaling processes for rainfall modelling has attracted less attention, since the multiscaling nature of the precipitation process has been widely accepted. Nevertheless, there have been efforts that show the great potential of simple scaling modelling approaches in reproducing a rich variety of precipitation patterns, at least on the event scale [*Papalexiou et al.*, 2011].

Despite the large variety of precipitation modelling approaches, none of them can be characterized as free of problems or universal. As already mentioned, models based on Markov chains and alternating renewal processes can only be applied for temporal scales coarser that the daily. Poisson cluster models have been generally applicable for temporal scales spanning from 1 hour to several days, but are incapable of reproducing statistics of fine scale precipitation [*Cowpertwait et al.*, 2011] or its inter-annual variability [*Katz and Parlange*, 1998]. Also their calibration procedure has as well proven cumbersome and generally involves high numerical efforts and advanced optimization techniques [e.g. *Vanhaute et al.*, 2011]. The problems of the event based models are obvious, since they have to be nested with different occurrence models for continuous rainfall simulation. Finally the problems associated with the MRC's can be of major importance as has been shown and quantified in 3.5.3 where inconsistencies of the MRC's due to several scaling regimes of temporal rainfall and various correlations and dependencies of the MRC

weights have been identified.

In order to cope with the problems mentioned above, several modifications of the original models above have been proposed. For example, several extensions for the Poisson cluster models have been developed, mainly in order to improve its small scale statistical properties [*Cowpertwait et al.*, 2011; *Onof and Wheater*, 1993; *Gyasi-Agyei and Willgoose*, 1997, 1999; *Cowpertwait et al.*, 2007]. Many additional parametrizations have been proposed for the MRC in order to take into account the various dependencies of the weights that have been identified [*Molnar and Burlando*, 2005; *Veneziano et al.*, 2006a; *Molnar and Burlando*, 2008; *Rupp et al.*, 2009; *Serinaldi*, 2010b; *Paschalis et al.*, 2012].

Even though such model improvements have been found to be reasonable, a generalized assessment of their applicability is not yet available. The scope of this chapter is to fill this gap by providing a general overview of the capabilities of the "state of the art" stochastic models for temporal rainfall and also provide new modelling approaches.

4.2 Data and Study Locations

Since the aim of this study is to asses the general applicability of each of the modelling approaches investigated here, several precipitation records across the world are used. Each of the data sets used here is of high temporal resolution (max 10 minutes) and belong to a different climatological region (figure 4.1, table 4.1).

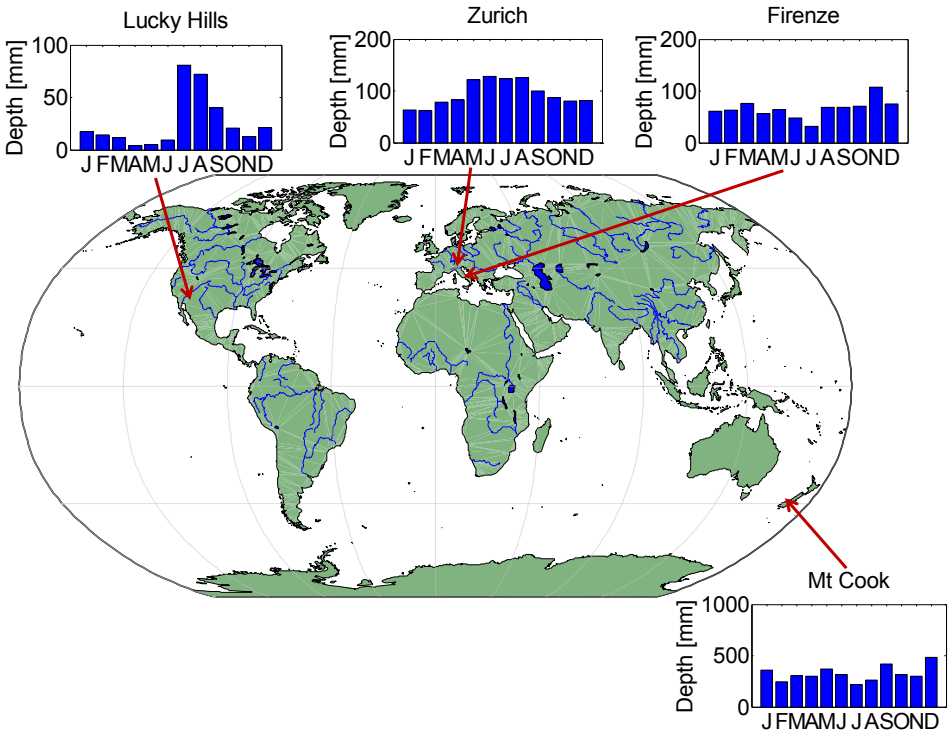


FIGURE 4.1: Study locations for the temporal stochastic models.

In total 4 stations are used in this study. The first station, Zurich (Switzerland), is representative of a temperate continental climate with distinct seasonality and a mean annual precipitation of about 1130 mm. The station belongs to the SwissMetNet that has already been analysed in the previous chapters.

A detailed description of the climatology of the area can be found in chapter 2. The second station, Firenze Ximeniano (Italy), is representative of a Mediterranean climate with a more temperate seasonality, consisting of dry summers and wet winters, and mean annual precipitation of about 800 mm. The precipitation recording mechanism is a tipping bucket and has a temporal resolution of 5 minutes. Data from this station have been analysed extensively in the past [*Olsson and Burlando*, 2002; *Cowpertwait et al.*, 2002; *Molini et al.*, 2009; *Veneziano*, 2002; *Burlando and Rosso*, 1991; *Becchi et al.*, 1994].

Outside of Europe two stations with contrasting precipitation regimes have been selected. The station Lucky Hills (Arizona, USA) is representative of the semiarid climate of southwestern USA. It is located in the Agricultural Research Service (ARS) Walnut Gulch Experimental Watershed [*Goodrich et al.*, 2008]. Precipitation has a pronounced seasonality with a wet monsoon season during summer and rare stratiform events occurring during the dry season. Mean annual precipitation is 340 mm. The station is equipped with a weighing gauge with a temporal resolution of 1 minute. The last studied station is Mount Cook (New Zealand), which is representative of the oceanic climate of southern New Zealand with an approximately uniform distribution of precipitation throughout the year. Its mean annual precipitation is 3900 mm. This high precipitation total is mainly generated by the strong orographic enhancement imposed by the southern Alps (max elevation 3754 m) that affects the southwestern area of New Zealand. The gauge is a tipping bucket with a temporal resolution of 10 minutes. Data were obtained from the National Climatic Database operated by NIWA [1].

TABLE 4.1: Properties of the weather stations analysed.

Station	Climate	Altitude [m]	P [mm/yr]	Δt [min]	Data
Zurich (Switzerland)	Temperate continental	556	1130	10	1981-2009
Ximeniano Firenze (Italy)	Mediterranean	50	800	5	1962-1986
Lucky Hills (Arizona, USA)	Semi-arid	1372	340	1	1962-2012
Mt Cook (New Zealand)	Oceanic	765	3900	10	2000-2012

For all the stations, there is also a distinct diurnal pattern for their respective summer season, when rainfall events are mainly convective (figure 4.2). The pattern is similar for all the stations, with afternoon depth accumulations being larger. This pattern is much more evident for the Lucky-Hills station than all the others.

In order to take into account the observed seasonal differences, the modelling analysis presented in this chapter is conducted on a seasonal basis. The diurnal cycle is neglected since none of the models that will be analysed later in this chapter can take this into account. Further development for the cyclostationary versions of these models incorporating a daily cycle is still needed.

4.3 Modelling Approaches for Temporal Precipitation

Since the objective is to comprehensively compare the efficiency of precipitation modelling approaches, several of the most common procedures are evaluated, and new approaches are developed. A summary of all the models and their respective parameters can be found in Table 4.2.

[1] http://cliflo.niwa.co.nz/

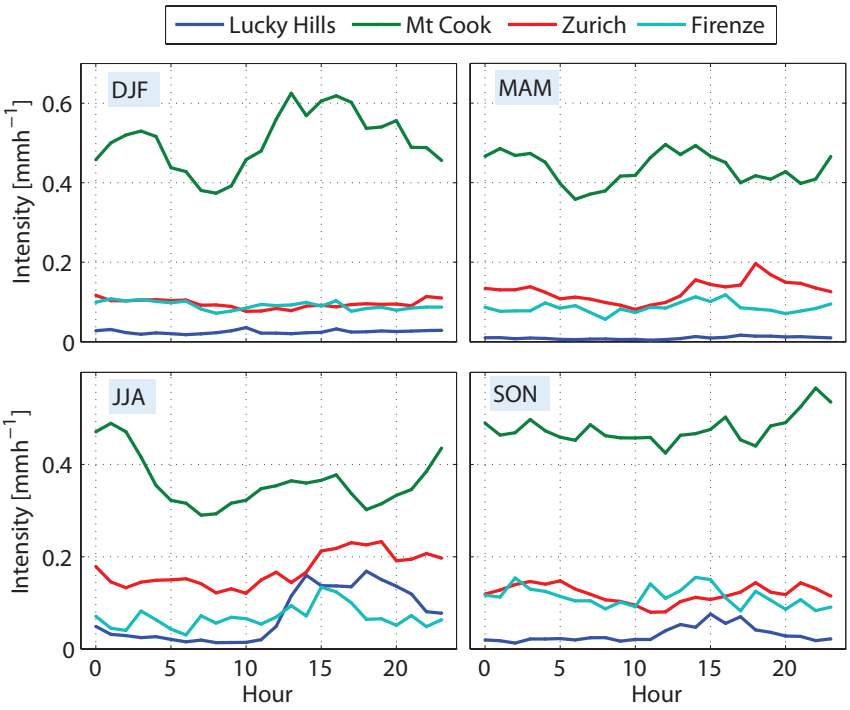


FIGURE 4.2: Variation of the mean hourly accumulation depths on a seasonal basis.

4.3.1 Poisson Cluster Models

Precipitation stochastic models based on point processes have been common ground in stochastic modelling of precipitation in the last 3 decades. One of their main attractive features is their parsimonious structure that allows tractable analytical derivations of their statistical properties. Among the point processes that have found great applicability in rainfall modelling are the Poisson cluster models and especially those that are structured on rectangular pulses. The two most widely applied variants of the Poisson cluster models are the Neyman-Scott rectangular pulses model (NSRP) and the Bartlett-Lewis rectangular pulses model (BLRP). Both of these models have the common structural feature that intensity in time is modelled as the superposition of rain cells that have uniform intensity across their lifetime. A schematic representation of these two models is shown in figure 4.3.

The Neymann-Scott Rectangular Pulses Model

The NSRP model inherits its name from the work of *Neyman and Scott* [1952] who investigated the spatial distribution of galaxies. The structure of the model can be summarized as follows:

- Storm origins arrive in time as a Poisson process with rate λ [h^{-1}].
- Each storm generates a random number of cells $C(C \geq 1)$ with mean value μ_c [-]. Here the geometric distribution is used similar to *Fatichi et al.* [2011]. Any positive and discrete distribution can be a good candidate, such as the Poisson distribution, etc.
- Cell origins of a storm are located after a time period that follows an exponential distribution with parameter β_c [h^{-1}].

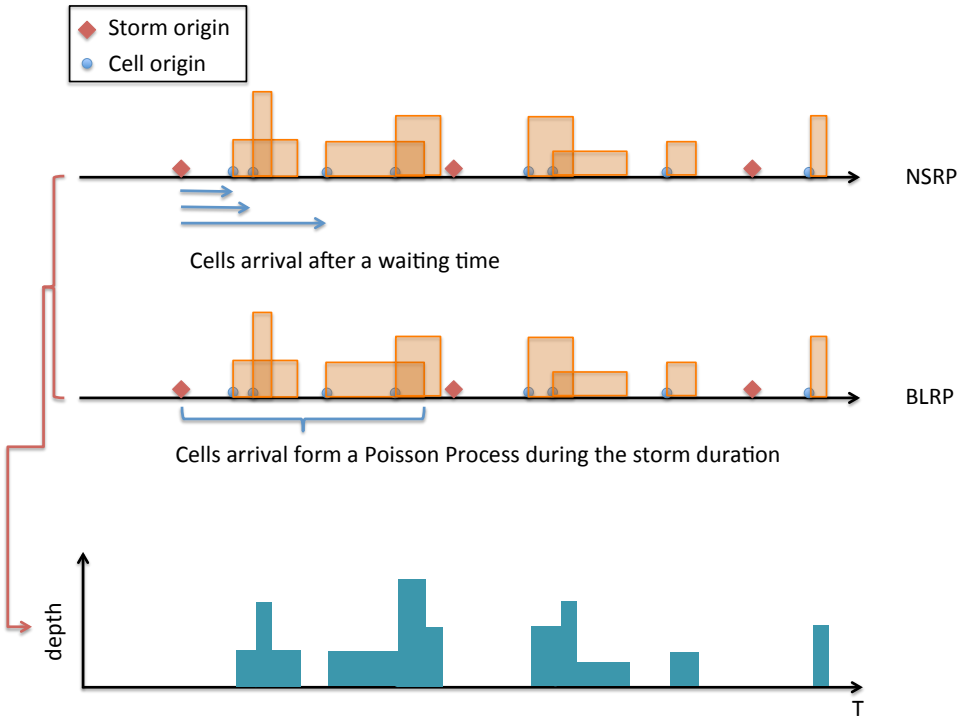


FIGURE 4.3: Schematic representation of the NSRP and BLRP models.

- Cell durations follow an exponential distribution with parameter η [h^{-1}].
- Cell intensities follow some distribution defined on the positive real line. Popular distributions include the the exponential, the Weibull and the gamma. Here the two parameter gamma $\Gamma(\alpha_c, \theta_c)$ distribution is used.

The model has been successfully applied in several regions of the world with reasonable results [*Burlando and Rosso*, 2002; *Beuchat et al.*, 2011; *Onof and Wheater*, 1993]. The statistical properties of the model, including its first three moments and the probability of zero precipitation across scales have been developed [*Rodriguez-Iturbe and Eagleson*, 1987; *Rodriguez-Iturbe et al.*, 1988, 1987; *Cowpertwait*, 1994; *Cowpertwait et al.*, 1996] and can be found in the Appendix B.1.

One of the major difficulties concerning the applicability of the model is its calibration. Since an analytical derivation of the model's likelihood function would be cumbersome, if not impossible, the calibration procedure is based on a generalized method of moments framework. An objective function quantifying the differences between the various statistical properties of the model and the data is defined and its numerical minimization yields the model's parameter estimates. The minimization problem that yields the models parameters, is the one defined by *Cowpertwait* [2006]:

$$\hat{\boldsymbol{\theta}_{\boldsymbol{p}}} = \arg\min_{\boldsymbol{\theta_p}} \sum_{h=1}^{n_h} \sum_{i=1}^{n} w_i \left[\left(1 - \frac{\hat{g}_i(h)}{g_i(\boldsymbol{\theta_p}, h)} \right)^2 + \left(1 - \frac{g_i(\boldsymbol{\theta_p}, h)}{\hat{g}_i(h)} \right)^2 \right], \tag{4.1}$$

where $g_i(\boldsymbol{\theta_p}, h)$ are the theoretical statical properties of the NSRP model for the parameter values $\boldsymbol{\theta_p} = [\lambda, \beta_c, \mu_c, \eta, \theta_c, \alpha_c]$ for various temporal scales h; and $\hat{g}_i(h)$ are the estimated ones from the data. The

function defined in equation 4.1 is highly non-linear and the numerical minimization procedure is not trivial. Several techniques have been reported in the literature spanning from traditional steepest descent algorithms to global optimization techniques [*Vanhaute et al.*, 2011]. The approach that is followed here is the one that *Burton et al.* [2008] found to be an adequate compromise between computational time and sufficient convergence of the algorithm. The algorithm is a constrained multi-start downhill simplex [*Nelder and Mead*, 1965]. Several starting points are taken into account in order to reduce the possibility of converging to local minima and the minimization is constrained in order to avoid unrealistic values of the model parameters, that conceptualize the true precipitation structure. For simplicity, the weights w_i are considered 1 for all the statistics. The statistics that are taken into account in the minimization procedure are the mean, the coefficient of variation, the probability of no precipitation and the transition probabilities ψ_{ww}, ψ_{dd} (see Appendix B.1). The temporal scales that are taken into account are [1-6-12-24-72] hours.

The Bartlett-Lewis rectangular pulses model

The structure of the BLRP model (the name of which originates from the work of *Bartlett* [1963] and *Lewis* [1964]) is very similar to the NSRP (see figure 4.3). The only difference is the cell arrival process. The model's structure can be summarized as follows:

- Storm origins arrive in time as a Poisson process with rate λ [h^{-1}].
- Each storm generates a random number of cells $C(C \geq 1)$ the follow a second Poisson process with rate β_c [h^{-1}].
- The storm duration for which the cell origins are generated follows an exponential distribution with parameter η [h^{-1}].
- Cell durations follow an exponential distribution with parameter γ [h^{-1}].
- Cell intensities follow a gamma distribution $\Gamma(\alpha_c, \theta_c)$.

Due to the very strong similarities of the BLRP model and the NSRP model, their efficiency is expected to be very similar. For this reason the original BLRP model is not further investigated. The version of the model that is taken into account here, introduces one more degree of freedom by considering that the parameter η is not any more constant but a random variable that follows a gamma distribution $\eta \sim \Gamma(\alpha_\eta, \nu_\eta)$. The model is abbreviated hereafter as the modified Bartlett-Lewis rectangular pulses model (MBLRP).

The modified version was first introduced by *Rodriguez-Iturbe et al.* [1988] in order to fix the problem of the original model that was not capable of reproducing well the probability of zero precipitation across scales. It has been subsequently widely used [e.g. *Onof et al.*, 2000; *Onof and Wheater*, 1994; *Islam and Entekhabi*, 1990]. The statistical parameters including the first two moments and the probability of zero precipitation can be found in the Appendix B.2. Even though several studies have highlighted the potential of the model to overcome the deficiencies of the original BLRP formulation, there are also others that challenge the model's applicability [*Burlando and Rosso*, 1991], reporting that the traditional versions of the model perform equivalently well.

The calibration procedure that was followed is identical to the one for the NSRP model. The parameter space for the minimization algorithm was considered the same as in *Vanhaute et al.* [2011].

4.3.2 Improved Representation of the Poisson Cluster Models

Several problems concerning both the NSRP and BLRP models have been identified. The most important one was their general inability to capture small scale variability. Precipitation statistics below one hour cannot be well reproduced due to the simplification of the rectangular pulses.

In order to overcome this issue several modifications/improvements of the original models have been developed. For example recently in order to capture fine scale variability and extremes, *Cowpertwait et al.* [2011] introduced a third Poisson process to the traditional Poisson cluster models that is associated with "rain pulses" that arrive as a Poisson process for each rain cell. The most widely used modification of the original model is the introduction of a *jitter* to the original NSRP, BLRP, MBLRP models [e.g. *Onof et al.*, 2000; *Onof and Wheater*, 1994; *Gyasi-Agyei and Willgoose*, 1997]. The jitter is nothing more than a multiplicative "noise" term that has been added to the initial formulation of the models in order to overcome the smooth profiles generated by the rectangular pulses. A schematic representation of the model is shown in figure 4.4.

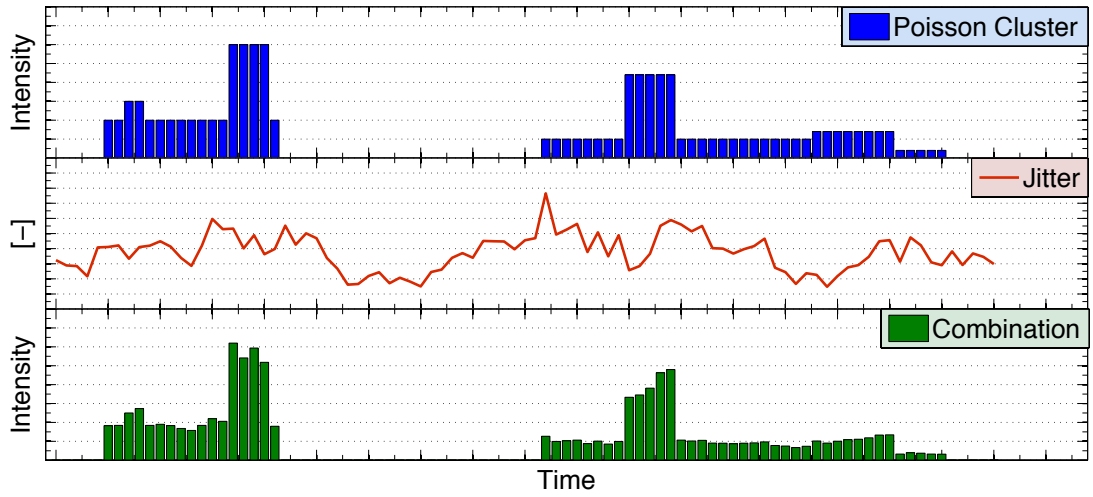


FIGURE 4.4: Schematic representation of the introduction of the jitter process for the Poisson cluster models.

The model with jitter is defined as:

$$Y(t) = Y_{PC}(t)Z_{jit}(t) \quad , t \in \Re, \tag{4.2}$$

where $Y_{PC}(t)$ is the Poisson cluster model and $Z_{jit}(t)$ a positive stochastic process independent of $Y_{PC}(t)$. The multiplication of the processes is done on the finest temporal scale of the simulation. Since the two processes are independent, the following properties for the first two moments and the autocovariance hold true for the scale at which the processes have been multiplied [*Gyasi-Agyei and Willgoose*, 1997]:

$$E[Y(t)] = E[Y_{PC}(t)]E[Z_{jit}(t)], \tag{4.3}$$

$$\begin{aligned}\mathrm{Var}[Y(t)] &= \mathrm{Var}[Y_{PC}(t)]\,\mathrm{Var}[Z_{jit}(t)] \\ &\quad + E[Y_{PC}(t)]^2\,\mathrm{Var}[Z_{jit}(t)] + \mathrm{Var}[Y_{PC}(t)]E[Z_{jit}(t)]^2,\end{aligned} \tag{4.4}$$

$$\begin{aligned}\mathrm{Cov}\left[Y(t), Y(t+\tau)\right] &= \{\mathrm{Cov}\left[Y_{PC}(t), Y_{PC}(t+\tau)\right] + E[Y_{PC}(t)]^2\} \\ &\quad \times\{\mathrm{Cov}[Z_{jit}(t), Z_{jit}(t+\tau)] + E[Z_{jit}(t)]^2\} - E[Y(t)]^2.\end{aligned} \tag{4.5}$$

The jitter was selected here as an exponential transform of a Gaussian autoregressive process of order one AR(1) $A_G(t) \sim N(0,1)$

$$Z_{jit}(t) = \exp\left[\sigma_G A_G(t) + \mu_G\right], \tag{4.6}$$

where

$$A_G(t) = \phi_G[A_G(t-1)] + \epsilon_G(t), \quad \epsilon_G \sim N\left(0, \sqrt{1-\phi_G^2}\right). \tag{4.7}$$

The exponentiation is needed in order to avoid negative values. To conserve the mean of the original process, the jitter process $Z_{jit}(t)$ has a unit mean. This means that $\mu_G = -\sigma_G/2$.

Even though approximations of the statistical properties of the process defined in 4.2 are available for any aggregation interval [*Onof and Wheater*, 1994], those are not used directly for the calibration of the model in this study. A slightly different approach is adopted. First the Poisson cluster models are calibrated exactly as defined in 4.3.1. Afterwards, the discrepancies of the small scale statistical properties of rainfall are fixed/improved by estimating the parameters of the jitter process for the finest scale adopting the equations 4.3- 4.5. The jitter process has two free parameters (ϕ_G, σ_G). The two equations that are used in order to define a system in order to estimate them are the variance and the lag one autocorrelation. This approach, is more suitable since the properties of the jitter are only calibrated for the finest temporal scale, the properties of which are intending to improve.

4.3.3 Composite Model Approaches

Another possible approach to improve performance across scales, is to combine stochastic models. Two different combination procedures are explored here. The first one nests different models across the range of temporal scales where they have been found to perform best. The second one is a multiplicative combination of different models, one responsible for the precipitation occurrence process and the other for the in-storm variability.

Nesting Models Across Temporal Scales

The first technique that is explored here, is the nesting of different stochastic models for the temporal scales for which they have been found to perform best. Very few analogues of such an approach can be found in the literature. For example, *Menabde and Sivapalan* [2000] following a similar modelling approach to *Schmitt et al.* [1998] successfully applied a combination of an alternating renewal process with a bounded MRC in order to reproduce rainfall extremes at fine (sub-hourly) temporal scales. *Fatichi et al.* [2011] combined an autoregressive model with a Poisson cluster one in order to introduce into the latter the inter-annual variability that it failed to capture. *Veneziano* [2002] obtained encouraging results combining an external model of conventional renewal type with an internal model based on an iterated random pulse process. Finally, *Koutsoyiannis* [2001] developed a model-free theoretical framework in order to couple different stochastic models.

In this section the stochastic models which are used as building blocks of the composite model are presented (Figure 4.5). The composite model consists of an external model which captures the processes connected with storm arrivals (timing and intensity) and inter-storm periods. The lower time scales reached by the external model are on the order of hours–days. The nested internal model is then used to capture the processes which describe rainfall variability within a storm by disaggregating the output of the external model to finer time scales on the order of minutes.

Although the different combinations of external and internal models are calibrated independently at their appropriate time scales, the performance of the composite model is tested jointly across all simulated temporal scales. This means that the composite model is used to simulate precipitation at the highest time scale of interest and then simulated precipitation series are aggregated back to any other chosen temporal scale, where relevant statistics are computed and compared.

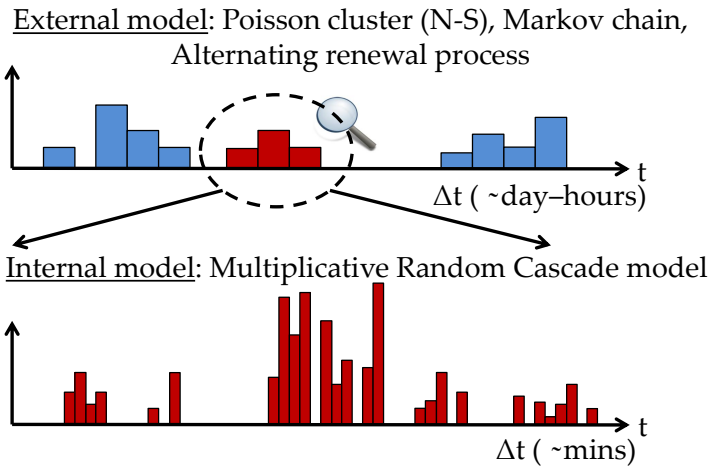


FIGURE 4.5: Schematic representation of nesting models across temporal scales

External Models:

For the coarse scale simulation, three different models, with various degrees of complexity are used. The first is based on alternative renewal processes, the second on Markov chains and finally the NSRP model, that has previously been explained in detail (see 4.3.1).

(a) The Alternating Renewal Process (ARP) describes the precipitation process as a sequence of wet and dry runs [*Roldan and Woolhiser*, 1982; *Buishand*, 1978; *Bernardara et al.*, 2007]. The durations of the runs are assumed to follow a prescribed probability distribution. The main assumption is that the wet and dry runs are, identically distributed (iid) and mutually independent. A constant intensity from a prescribed probability distribution is assigned to each day. These uniform intensities are also iid and independent of the duration of the wet run. Several distributions have been used in the literature and the choice is always subjective and dependent on the analysed dataset [e.g. *Srikanthan and McMahon*, 2001; *Roldan and Woolhiser*, 1982; *Buishand*, 1978]. Here, an exponential distribution for the duration of the dry and wet runs [*Ivanov et al.*, 2007] was chosen and the two parameter gamma distribution for the precipitation intensity. The two parameter gamma distribution is a specific case of the generalized gamma distribution which has been found to yield very good results for daily precipitation depths worldwide [*Papalexiou and Koutsoyiannis*, 2011]. However, it should be acknowledged that daily precipitation series often have heavy tails, a behaviour that the two parameter gamma distribution cannot capture. Potential candidates for the daily precipitation depths may include the Weibull distribution, the Pareto etc, specifically if a good simulation of daily extremes is essential.

The ARP model has in total 4 parameters to be estimated (Table 4.2). The estimation is based on the maximum likelihood method. Dry and wet spells and their respective daily depths are extracted from the rainfall records and their probability distributions are fitted by maximizing their likelihood functions.

(b) The Markov Chain (MC) describes the precipitation occurrence process in time through transition probabilities between wet and dry states. The two-state first order MC simulates a binary [0-1] process which corresponds to the dry (d) and wet (w) states with the probability transition matrix,

$$M_{i,j} = \begin{pmatrix} p_{w,w} & p_{w,d} \\ p_{d,w} & p_{d,d} \end{pmatrix}, \tag{4.8}$$

where $p_{i,j}$ is the probability of transitioning from state i to state j in the successive time step. A precipitation depth is assigned for each time step that the MC stays in the wet state. Precipitation depth is assumed iid and has a prescribed probability distribution, which we take to be a two parameter gamma distribution identical to the ARP.

The probability transition matrix $M_{i,j}$ is dependent on the duration of the selected time step of the MC model and has to be estimated from data at the same time scale. In this study the selected time step was approximately 1 day (1280 minutes). Equally, the distribution of rainfall depth is related to this time step duration and has to be estimated at the same time resolution. The fit of the rainfall depth distribution was conducted by the maximum likelihood method [*Papoulis and Unnikrishna*, 2002]. Altogether the MC model has 4 parameters (Table 4.2), two parameters of the gamma distribution for rainfall depths and two parameters of the probability transition matrix $p_{w,w}$ and $p_{d,d}$, which are related to the remaining probabilities $p_{w,w} + p_{w,d} = 1$ and $p_{d,w} + p_{d,d} = 1$.

Stochastic models based on the Markov chain process have been widely used in hydrology, especially for daily precipitation, due to their simplicity and parsimonious parametrization yielding very good results [e.g. *Haan et al.*, 1976; *Katz*, 1977; *Foufoula-Georgiou and Lettenmaier*, 1987; *Srikanthan and Pegram*, 2009].

Here the temporal scale of the external models was selected to be roughly the daily scale (1280 min). The reason of this choice is that ARP and Markov chains are not adequate simulation tools for sub-daily time scales.

Internal Models:

The disaggregation procedure of the daily output of the external models is based on the notion of the MRC. MRCs have been shown to be a very useful disaggregation/downscaling tool, especially for time scales ranging from daily to few minutes [*Molnar and Burlando*, 2005; *Gaume et al.*, 2007; *Lombardo et al.*, 2012; *Paulson and Baxter*, 2007; *Güntner et al.*, 2001]. However other approaches could also be considered for precipitation disaggregation [*Koutsoyiannis*, 1992; *Pui et al.*, 2012].

As shown in section 3.2.2, the behaviour of the MRCs is based on the distribution of their cascade generator W. For all the variants of the MRCs used here, the cascade generator has a probability mass at zero, in order to generate non precipitation intervals. In this study, for the internal model two variants of the micro-canonical MRC model parametrized by *Rupp et al.* [2009] and the classic canonical beta-lognormal model [*Over and Gupta*, 1996; *Paschalis et al.*, 2012] were used.

The difference of the micro-canonical MRC in comparison to the ones discussed in section 3.2.2 is that they are "mass conservative" in the sense that the finer scales preserve the exact precipitation depth of the coarser scales, and not only on average. This can be formulated into a restriction that the weights in every subdivision have to be complementary i.e.

$$\sum_{k=1}^{b} W_n(b(i-1)+k) = 1 \text{ for } i = 1, 2..., b^{n-1}. \tag{4.9}$$

The probability of the cascade to have an atom at zero can be introduced as:

$$P(W_n(j) = 0 \text{ or } W_n(j+1) = 0) = p_0, \tag{4.10}$$

where $j, j+1$ correspond to an arbitrary pair of complementary weights on the cascade development with branching number (here $b = 2$).

The positive part of the generator W^+, in order to achieve mass conservation, has to be bounded in $[0, 1]$. Here, the single parameter (symmetric) Beta distribution was chosen for W^+ with probability distribution function (pdf),

$$f(W^+) = \frac{1}{B(\alpha_W)} W^{+\alpha_W - 1}(1 - W^+)^{\alpha_W - 1}, \tag{4.11}$$

where $B(\alpha_W)$ is the Beta function and α_W is a parameter related to the variance of W [e.g. *Molnar and Burlando*, 2005].

In summary the microcanonical models are defined as:

1. MRC model A: The probability of a zero weight p_0 is dependent on the temporal aggregation scale and on the average intensity of precipitation at the next (coarser) aggregation scale. The parameter α_W of the Beta distribution of W^+ is dependent only on the temporal aggregation scale.
2. MRC model B: Both p_0 and the parameter α_W of the Beta distribution are dependent on temporal aggregation scale and average intensity of precipitation at the next (coarser) aggregation scale.

It was found that micro-canonical MRC models of lower complexity [*Molnar and Burlando*, 2005; *Rupp et al.*, 2009] give rather unsatisfactory results and are thus not further discussed. Similar results concerning the applicability of parsimonious versions of the micro-canonical MRCs were recently also obtained by *Pui et al.* [2012].

For the MRC model A, the probability of a non-zero weight at the cascade development is parametrized as

$$P_x(I,\tau) = 1 - p_0(I,\tau) = \frac{1}{2}\left(1 + \text{erf}\left[\frac{\log(I) - \mu_{MRC}}{\sqrt{2}\sigma_{MRC}}\right]\right), \tag{4.12}$$

where τ is the temporal scale of the cascade, I is the precipitation intensity of the coarser scale, and erf is the error function defined as $\text{erf}(x) = \frac{2}{\sqrt{\pi}}\int_0^x e^{-t^2}$. The parameters μ_{MRC} and σ_{MRC} depend linearly on the logarithm of the temporal scale of the cascade,

$$\mu_{MRC} = \alpha_\mu \log(\tau) + \beta_\mu, \tag{4.13}$$

$$\sigma_{MRC} = \alpha_\sigma \log(\tau) + \beta_\sigma, \tag{4.14}$$

The parameter of the Beta distribution is only dependent on the temporal scale, τ,

$$\alpha_W(\tau) = \alpha_{0W}\tau_W^{H_W}. \tag{4.15}$$

For the MRC model B, the parametrization of the probability of non-zero weights in the cascade is as in Equation (4.12). The α_W parameter is conditioned both on temporal scale and precipitation intensity at the coarser scale,

$$\log[\alpha_W^*(I)] = c_0 + c_1 \log(I) + c_2[\log(I)]^2, \tag{4.16}$$

and

$$\alpha_W(I,\tau) = \alpha_W^*(I)\alpha_{0W}\tau^{H_W}. \tag{4.17}$$

The estimation of the polynomial coefficients c_i for lower intensities (typically < 1 mm h^{-1}) can be problematic due to precision artefacts introduced by the tipping mechanism of the gauges. For this reason, the fitting procedure of Equation (4.16) was restricted to temporal scales coarser than 20 minutes. In some cases the second order polynomial introduced in Equation (4.17) could be simplified to a first order polynomial or even to a constant, however the above parametrization is kept for consistency with *Rupp et al.* [2009].

In all of the cases, the estimation of the α_W parameter of the Beta distribution was carried out using the method of moments [e.g. *Molnar and Burlando*, 2005],

$$\alpha_W = \frac{1}{8\text{Var}(W^+)} - 0.5\,, \tag{4.18}$$

where W^+ values were estimated from the data as the ratio between precipitation depths at two successive (embedded) temporal scales,

$$W^+ = \frac{R_{n+1}(\Delta^{bi}_{n+1})}{R_n(\Delta^i_n)} \quad \text{for} \quad i = 1, 2..., b^n\,, \tag{4.19}$$

if both precipitation depths are positive and not equal.

The canonical beta lognormal model has already been thoroughly analysed in section 3.2.2. Its parameters are estimated with the typical moment scaling analysis. Moreover, the fact that it has been found (3.5.1) that the temporal scales from minutes up to 1 day is a distinct scaling regime giver further rise to the applicability of such an approach.

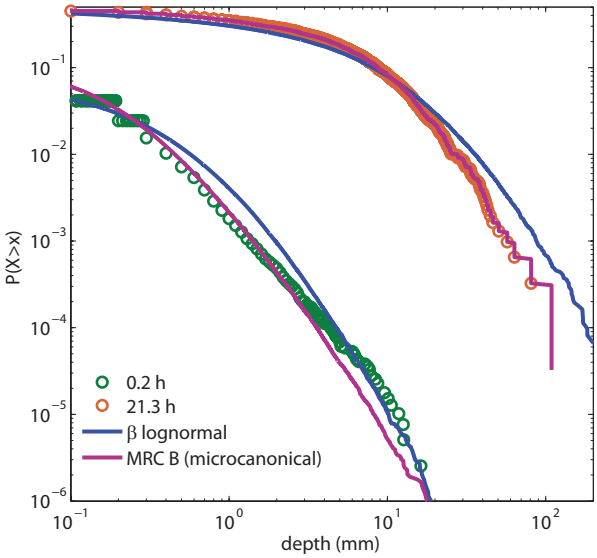


FIGURE 4.6: Generated exceedance probabilities for the disaggregation/downscaling of the daily time series for the Zurich station during spring (MAM), according to a canonical (β-lognormal) and a microcanonical (MRC B) model for the 10 min and the 1280 min aggregation interval.

An important issue concerning the applicability of the canonical MRC model is shown in figure 4.6. A disaggregation (downscaling) of the "daily" (1280 minutes) precipitation time series for the Zurich station for the Spring season (MAM) is performed with a microcanonical MRC model and a canonical MRC model and more specifically the β-lognormal model [*Over and Gupta*, 1996; *Molnar and Burlando*, 2005; *Paschalis et al.*, 2012]. Although the canonical beta-lognormal model is capable of reproducing well the heavy tails of the fine scale (10 min) precipitation, aggregating the time series back to the original temporal scale (1280 min) distorts the distribution of the precipitation depth. This is happening because the aggregated series at the 1280 min temporal scale correspond to the so-called dressed quantities [*Veneziano*

and Langousis, 2010], and the distribution is distorted due to the distribution of the dressing factor (see also section 3.2.2). In a hydrological perspective, this behaviour can be a serious disadvantage, since it is crucial that the original data that have to be downscaled, preserve not only their mean values but their exact distribution. In any case though, this possibility is also investigated here for the sake of completeness of the study.

Multiplicative Model Combination

Another possibility of model nesting is similar to the jitter process for the Possion cluster models and the introduction of the atom at zero for the intermittent MRC. In summary this approach is a simple multiplication of the realizations of two different and mutually independent stochastic processes at the same temporal time scale. The possibility that is investigated here, is a combination of the Poisson cluster model as a binary (0-1) occurrence model and a model describing the in-storm rainfall variability. A schematic representation of such a process is shown in figure 4.7. The general description of the model can be defined as:

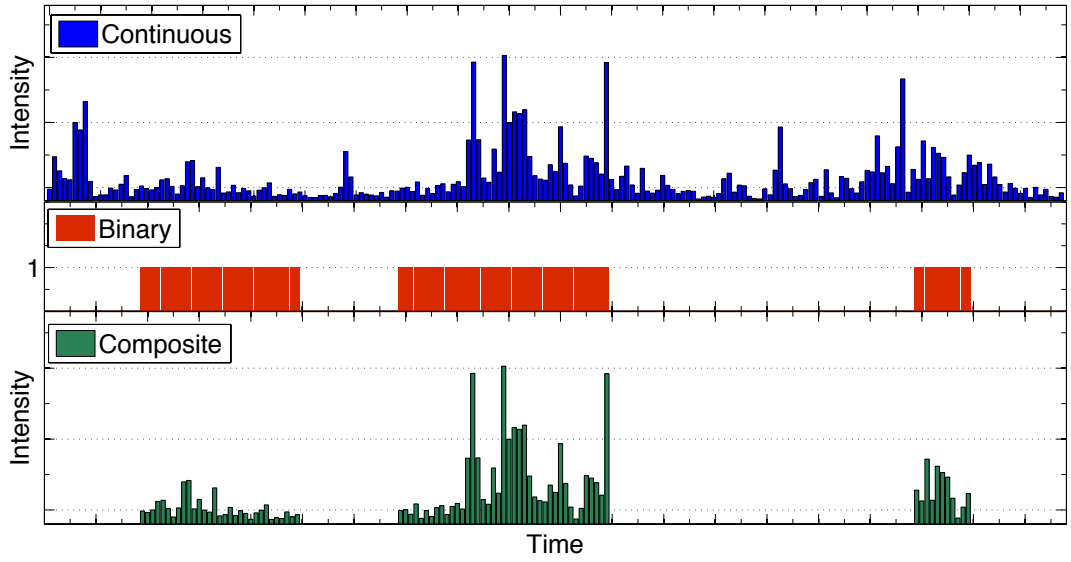


FIGURE 4.7: Schematic representation of the model combination through multiplication.

$$Y(t) = Y_{bin}(t)Y_{con}(t), \tag{4.20}$$

where $Y_{bin}(t)$ is the binary occurrence process and $Y_{con}(t)$ is the model that generated the in-storm variability. To my knowledge, even though this approach is very flexible, few applications can be found in the literature [*Gyasi-Agyei and Willgoose*, 1999].

As occurrence models, both the NSRP and the MBLRP model are used. In this case the models are structured with a reduced number of parameters, since the ones that describe the probability distribution of the rain cell intensities vanish.

For the continuous process $Y_{con}(t)$ several models could be selected, spanning from common linear models e.g. ARMA(p,q) [*Box and Jekins*, 1970] to ARFIMA models [*Hosking*, 1981; *Baillie*, 1996] etc. One

example is the study of *Gyasi-Agyei and Willgoose* [1999] where the authors investigated the applicability of the combination of the BLRP model and binary Markov chains with autoregressive models. Most of the continuous models usually adopt a Gaussian distribution, so a probability transformation should be adopted as well. For temporal rainfall, generally the distribution that should be selected is positively skewed and has a heavy tail. Several distributions can be appropriate candidates, such as the lognormal, the generalized Pareto etc. For the data that have been analysed for this study, it was found that even well-known and generally applied distributions fail to capture well both the body and tail of the positive part of the distribution (figure 4.8). For this reason a very flexible 4 parameter distribution was chosen. The distribution has been introduced by *Papalexiou and Koutsoyiannis* [2009] and abbreviated by the authors as the "JH-distribution".

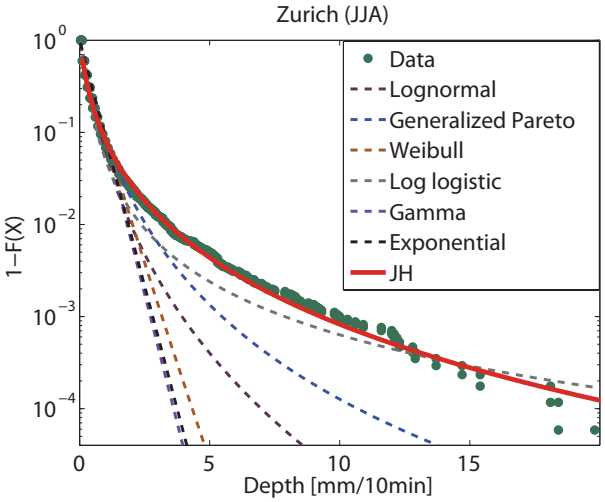


FIGURE 4.8: Fitting of various probability distributions to 10-min values of precipitation of the SMA Zurich station during summer (JJA).

Its probability distribution function is defined as

$$f_{JH}(x) = \frac{d}{cB(a,b)} \left(\frac{x}{c}\right)^{ad-1} \left[\left(\frac{x}{c}\right)^{d} + 1\right]^{-(a+b)}, \tag{4.21}$$

where $B(a,b)$ is the beta function. Its cumulative distribution function is:

$$F_{JH}(x) = I_{\frac{1}{1+(x/c)^{-d}}}(a,b), \tag{4.22}$$

where $I_z(\alpha,\beta) = \int_0^z t^{\alpha-1}\,(1-t)^{\beta-1}\,dt$, is the incomplete beta function.

The q-order raw moments of the distribution are defined as:

$$m_q = \begin{cases} \dfrac{c^q}{B(a,b)} B\left(a + q/d, b - q/d\right) & \text{if } q < bd \\ \infty & \text{if } q \geq bd \end{cases}. \tag{4.23}$$

Even though the distribution is not the most parsimonious choice, it was found to outperform all the widely known, less parametrized distributions for all the stations that have been examined (see figure 4.8).

In addition, the model that describes the in-storm variability has to be parsimonious. It has also been widely recognized that precipitation records exhibit long term correlations [e.g. *Montanari et al.*, 1996; *Kantelhardt et al.*, 2006]. For these reasons the model that is adopted here is similar to *Papalexiou and Koutsoyiannis* [2011] and consist of a non-linear transform of a fractional Gaussian noise. The model has been found to be able to generate rich patterns of high resolution precipitation on the event scale and this makes it a very good candidate for modelling the in-storm variability of precipitation. Moreover the model can deal with the long term correlations of the rainfall process and at the same time be parsimonious, since the correlation is determined only by one parameter H.

The model for the continuous part of the precipitation process can thus be expressed as

$$Y_{con}(t) = F_{JH}^{-1} \circ F_N \left([Y_{FGN}(t)] \right) = F_{JH}^{-1} \left(F_N \left[Y_{FGN}(t) \right] \right), \tag{4.24}$$

where $F_{JH}^{-1}(x)$ is the inverse cumulative distribution function of the fitted JH distribution and $F_N(x)$ the cumulative distribution function of the standard Gaussian distribution $N(0,1)$.

The Fractional Gaussian zero mean noise, is a simple scaling process defined as [*Mandelbrot and Van Ness*, 1968; *Koutsoyiannis*, 2003]:

$$Y_{FGN}^{(k)}(t) = k^{H-1} Y_{FGN}^{(1)}(t), \tag{4.25}$$

where k is the scale of aggregation and H the so-called Hurst exponent that defines the self similar properties of the process. Its autocorrelation function is the same for any aggregation interval and is defined as:

$$\rho_j^{(k)} = \frac{1}{2} \left[(j+1)^{2H} + (j-1)^{2H} \right] - j^{2H}. \tag{4.26}$$

Several methodologies have been defined for the simulation of fractional Gaussian Noises [*Mandelbrot*, 1971; *Koutsoyiannis*, 2000]. Here in order to increase the computational efficiency of the simulation scheme, the stochastic signal sampling was conducted in the frequency domain, exploiting Fast Fourier Transform [*Ignaccolo et al.*, 2009] [2]. More information about the simulation algorithm can be found in section 6.2.5.

The calibration of the model consists of the parameter estimation for the Poisson cluster models and fitting of the distribution and the estimation of the Hurst exponent. The parameters of the Poisson cluster models are estimated by the numerical minimization of equation 4.1 for aggregation scales spanning from 1 hour to 3 days. The only statistic that is taken into account in the calibration procedure is the probability of zero precipitation across scales (see Appendices B.1, B.2).

[2] The Matlab code used for the simulation can be found online in http://www.mathworks.com/matlabcentral/fileexchange/19797-simulation-of-fractional-gaussian-noise-exact

The fitting of the distribution is conducted with the method of moments for the positive values of the precipitation records. Specifically the moments of the positive values where estimated as

$$\hat{m}_q = \frac{\sum_{i=1}^{n} R_i^q}{n[1-\hat{\psi}(h)]}, \tag{4.27}$$

where n is the sample size of the precipitation record R_i at the finest temporal scale h. $\hat{\psi}(h)$ is the theoretical value of zero precipitation according to the estimated parameters of the binary Poisson cluster model at the scale h. The order of moments that are taken into account are (0.5-1-1.5-2-2.5), with the restriction that they are all defined and do not diverge.

The estimation of the Hurst exponent is not straightforward. The well known methods, such as the aggregated variance, or the "R/S" method [*Taqqu et al.*, 1995] cannot be directly applied since in the model structure the FGN process is used in a "hidden" sense. For this purpose a different approach is used here. From equation 4.3, it becomes clear that if the autocovariance function of the binary process is known, then the respective one of the continuous process can be estimated respectively.

The autocovariance of the binary process can be expressed as [*Gyasi-Agyei and Willgoose*, 1999]:

$$\begin{aligned}\mathrm{Cov}[Y_{bin}^{(k)}(\tau), Y_{bin}^{(k)}(\tau+l)] = [1-\psi(k)]^2 \{P_{dd}(l) - [P_{dw}(l)+P_{wd}(l)] \\ \times \left[\frac{\psi(k)}{1-\psi(k)}\right] + P_{ww}(l)\left[\frac{\psi(k)}{1-\psi(k)}\right]^2\Bigg\},\end{aligned} \tag{4.28}$$

where $\psi(k)$ is the probability of zero precipitation at the temporal scale k, and $P_{dd}(l), P_{dw}(l), P_{ww}(l)P_{wd}(l)$ are the lag l joint probabilities. For the lag-1 covariance at the temporal scale k, it can be shown [*Gyasi-Agyei*, 1999] that

$$P_{dd}(1) = \psi_{dd}(k)\psi(k), \tag{4.29}$$

$$P_{dw}(1) = \psi_{dw}(k)\psi(k),$$

$$P_{ww}(1) = \psi_{ww}(k)[1-\psi(k)],$$

$$P_{wd}(1) = \psi_{wd}(k)[1-\psi(k)],$$

and the analytical expression of the conditional transition probabilities $\psi_{ij}(k)$ can be found in the Appendices B.2,B.1.

It is also needed to express the autocovariance of the FGN process in terms of its non-linear distribution transform. *Guillot* [1999] defines this connection as follows:

$$\mathrm{Cov}[Y_{con}(t), Y_{con}(t+\tau)] = \sum_{k=1}^{\infty} \frac{\xi_k^2}{k!} \mathrm{Cov}^k[Y_{FGN}(t), Y_{FGN}(t+\tau)], \tag{4.30}$$

where

$$\xi_k = \int_{-\infty}^{\infty} f_{tr}(x) H_k(x) f_N(x) \,\mathrm{d}\, x. \tag{4.31}$$

$H_k(x) = (-1)^k e^{x^2} \frac{d^k}{dx^k} e^{-x^2}$ are the Hermite polynomials and $f_N(x)$ is the probability density distribution function of the standard Gaussian distribution $N(0,1)$. $f_{tr}(x) = F_{JH}^{-1} \circ F_N(x)$ is the transformation of the probability distributions from Gaussian $N(0,1)$ to the JH distribution. The estimation can only be achieved numerically. For the numerical implementation, it was found that an order $k = 3$ is enough to approximate the autocovariance of the transformed process. For the estimation of H only the lag-1 autocorrelation coefficient is used. An alternative approach would by to estimate H from the power-law decay of the autocorrelation function for large lags. This approach was not applied here because, due to the relatively small sample, the significance of the autocorrelation at large lags is low. The significance of the estimation is also highly dependent of the H itself [*Koutsoyiannis*, 2003].

One possible artefact of this choice of the calibration procedure is that the a priori choice of the parameters of the Poisson cluster occurrence model can restrict the values that the lag 1 autocorrelation coefficient can take (Equation 4.3, Figure 4.9). In the case that the estimated autocorrelation coefficient lies outside the range of values that are plausible, this can affect the model's efficiency. For the stations analysed here, this problem only occurred when very high, short term correlations are present in the data (e.g. Mt. Cook station).

4.4 Results

In order to assess the performance of the models, each of them was applied for all four stations on a seasonal basis. The assumption of the seasonal stationarity is perhaps crude, and that is the reason why most of the models are commonly calibrated on a monthly basis, but for the purpose of a comparison between models, this assumption is considered adequate. It should be noted though that if any of the above models has to be used in any application (e.g. flood risk management etc.), it is strongly suggested to be calibrated on a monthly basis. Furthermore all the modelling approaches that are adopted here do not include the option of cyclo-stationarity, which as was shown before (see figure 4.2) may need to be included in some cases.

For illustrative purposes, the results for only one station and one season will be presented here in order to highlight the pros and cons of each of the models. A complete summary of the statistics can be found

TABLE 4.2: Temporal Model Parameters

Model	Parameters	Description
Modified Bartlett Lewis	λ, μ_c, β_c, α_η, β_η, α_c,θ_c	Model parameters as described in *Onof and Wheater* [1994]
Neyman Scott	λ, μ_c, β_c, η, α_c,θ_c	Model parameters as described in *Cowpertwait* [1998]
Markov chain	$M_{i,j}$	Probability transition matrix
	a_{MC}, b_{MC}	Gamma distribution parameters for the depth accumulations
Alternating renwal process	λ_w, λ_d	Parameters of the exponential distribution of the dry and wet spell durations
	a_{ARP}, b_{ARP}	Gamma distribution parameters for the depth accumulations
Beta - Lognormal Canonical MRC	β, σ^2	Model Parameters of the beta lognormal MRC model as in Table 3.1
MRC model A	α_μ, β_μ, α_σ, β_σ	Model parameters controlling the dependence of the probability of zero of the MRC generator on the temporal scale and intensity
	α_{0W}, H_W	Model parameters controlling the dependence of the α_W parameter of the Beta distribution on the temporal scale
MRC model B	α_μ, β_μ, α_σ, β_σ	Model parameters controlling the dependence of the probability of zero of the MRC generator on the temporal scale and intensity
	α_{0W}, H_W, c_0, c_1, c_2	Model parameters controlling the dependence of the α_W parameter of the Beta distribution on the temporal scale and intensity
Jitter model	ϕ_G, σ_G	The parameters of the AR(1) unit mean jitter process. σ is the parameter of the lognormal distribution that controls the variance and ϕ the regression coefficient
FGN model	H, a, b, c, d	H The Hurst exponent of the FGN, and a, b, c, d the parameters of the JH distribution

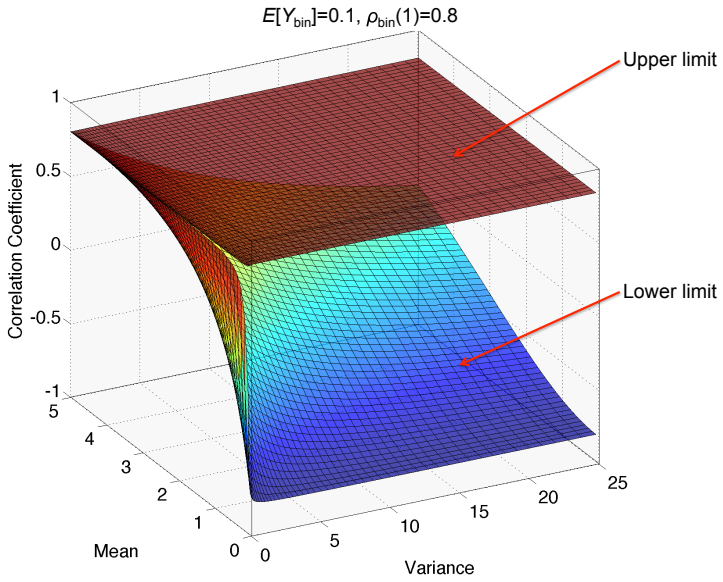


FIGURE 4.9: Envelope of the plausible values that the lag 1 autocorrelation coefficient can have in the case of the multiplicative model combination for the case $\rho(1)_{bin} = 0.8$ and probability of rainfall occurrence 0.1. The range is a function of the mean and the variance of the composite process.

in the tables 4.3 and 4.4 where the mean relative errors expressed as

$$\text{MRE} = \frac{1}{4n_h} \sum_{s=1}^{4} \sum_{h=1}^{n_h} \frac{g_s(h,s) - g_d(h,s)}{g_d(h,s)}, \tag{4.32}$$

are presented, where $g_s(h,s),\ g_d(h,s)$ stand for the values of the simulated and observed statistic g for the aggregation scale h for the season s.

4.4.1 Poisson Cluster Models

The Poisson cluster models have been found to be a very good approach for temporal precipitation, and that is the reason why for precipitation simulation they are the most common option. Due to the fact that the NSRP and BLRP model are very similar on their structure, their results should be very similar. Here an illustration of the capabilities and problems of the NSRP is reported.

In Figure 4.10 the performance of the NSRP model is summarized for the Zurich station. In general the model efficiency in reproducing precipitation statistics for temporal scales coarser than hourly is excellent, however significant departures are found for the sub-hourly scales. Due to the simplified representation of the precipitation process as a superposition of rectangular rain cells, small-scale variability cannot be simulated. The finest temporal scale of fluctuations is associated with the distribution of the rain cell lifetime that is typically not short enough in the model to properly characterize the true precipitation process at this scale.

The simulated precipitation time series severely underestimate the occurrence of 10-min rainfall depths greater than 1 mm (Figure 4.10a). Moreover, the simulated series are very smooth at this temporal scale,

leading to very high autocorrelations (Figure 4.10b). Another typical problem that is associated with the NSRP model is its inability to successfully reproduce the probability of zero precipitation [e.g. *Entekhabi et al.*, 1989] for large aggregation scales. In the examples analyzed here, this problem was not significant (Figure 4.10c), mainly because this statistic was explicitly considered in the calibration procedure.

Intermittency is connected to wet and dry spell durations through the clustering of the precipitation process. A wet period is defined as the duration for which precipitation accumulated at the selected temporal resolution stays continuously above zero is defined here. This notation is somewhat different from the definition of storm duration [e.g. *Vandenberghe et al.*, 2010]. The distribution of wet spell durations at the studied temporal aggregation scales is shown in Figure 4.10d. Similarly to the other statistics, the performance for time scales above 1 hour is very good, however for the 10-min scale there is a significant overestimation of short wet spell durations by the model. The NSRP model is not able to adequately generate gaps within storms and short rainfall pulses for structural reasons, despite the fact that the transition probabilities from dry to wet states were also taken into account in the calibration procedure. In order to compensate for this overestimation in the probability of precipitation occurrence, the durations of dry spells are as well overestimated.

The correlation structure across temporal scales was also examined with spectral analysis. Figure 4.12b) shows the observed and simulated power spectra for the Zurich station. The strong autocorrelation at high temporal resolutions in the model results in a steeper spectral decay for higher frequencies. The divergence between observed and simulated series starts roughly at temporal scales of about 5 hours. In summary the NSRP model alone performs very well for scales above 1 hr, but not for scales below 1 hr.

As mentioned before, a solution that has been proposed in order to fix the problem of the simulation of zero precipitation across scales was to include another degree of freedom in the model by constructing the modified versions of the NSRP and BLRP models respectively [*Entekhabi et al.*, 1989; *Onof and Wheater*, 1994, 1993; *Rodriguez-Iturbe et al.*, 1988]. For this reason, also the Modified BLRP model has been examined here. Since as already explained, the performance of the original NSRP model is very good for the probability of precipitation occurrence, for all the stations examined here, the differences between the NSRP and MBLRP model concerning this statistic are negligible (see table 4.3). It should be also noted that both models overestimate the durations of the wet spells (see figure 4.11). This means that both models, even though calibrated to reproduce the probability of zero precipitation across scales and their respective transition probabilities, can generate unreasonably high durations, as it is shown in figure 4.11. This behaviour has been previously identified by *Verhoest et al.* [2010] as well and can be very harmful if it is not taken into account. For example, if the model is to be used for flood risk estimation, it should be noted that the long durations can lead to unreasonably high event depth accumulations, and subsequently to very high flood estimates. The behaviour of the MBLRP model in comparison to the NSRP model was slightly superior. This overestimation can be attributed to the simplified way that precipitation is generated, as a superposition of raincells. Probably a different choice of the distribution of the durations would result to better results.

Taking into account the very diverse climatological conditions for the stations that have been used here, it can be said that the enhancement of the complexity of the model structure does not necessarily result in better performance. A major difference of the current study and the ones that pointed out the inefficiency of the NSRP model in reproducing the probability of zero precipitation, is the calibration procedure of the model which is not a trivial task. In this case it is not possible to conclude whether the problem arises as a structural deficiency of the NSRP model or its calibration procedure. From numerical experiments that were done in this study, it was found that the parameter estimation was very sensitive to the selection

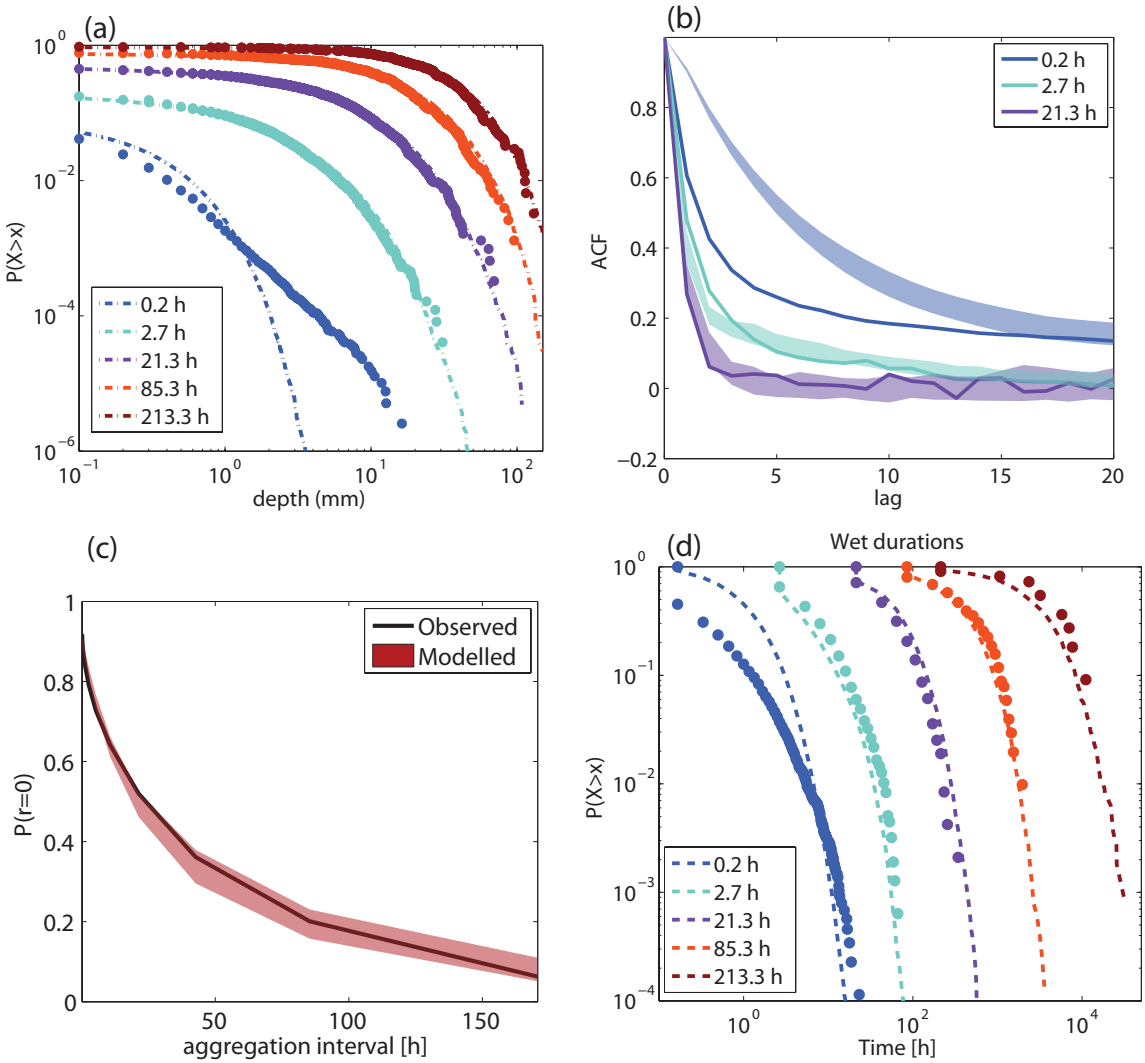


FIGURE 4.10: Simulated statistics for the Neyman Scott model for spring season (MAM) for the Zurich (SMA) station. (a) The exceedance probabilities of precipitation depth accumulated for different temporal scales. Dots represent observed values and lines simulated. (b) The autocorrelation functions. The lines show the data autocorrelation functions and the shaded areas correspond to the 90% confidence intervals for the ensemble of the simulations. (c) Probabilities of zero precipitation. Shaded red areas correspond to the 90% confidence intervals of the simulations (d) Exceedance probabilities of the wet spell durations for different aggregation intervals. Dots represent observed values and lines simulated.

of the objective function that was chosen for minimization as well as the selection of the temporal scales for which the model is calibrated.

As was clearly shown, both MBLRP and NSRP models are not able to reproduce well the sub-hourly statistics (see table 4.4). To resolve this issue, the jitter process that was previously explained was introduced. The first problem is the substantial underestimation of the tails of the precipitation depth accumulations for the fine temporal scales. As can be seen in figure 4.12(a) the introduction of the jitter improves the ability of the model to capture better the tails of the precipitation depth accumulations. At the same time, the gain in correct representation of extreme precipitation for fine temporal scale is still small (figure 4.18).

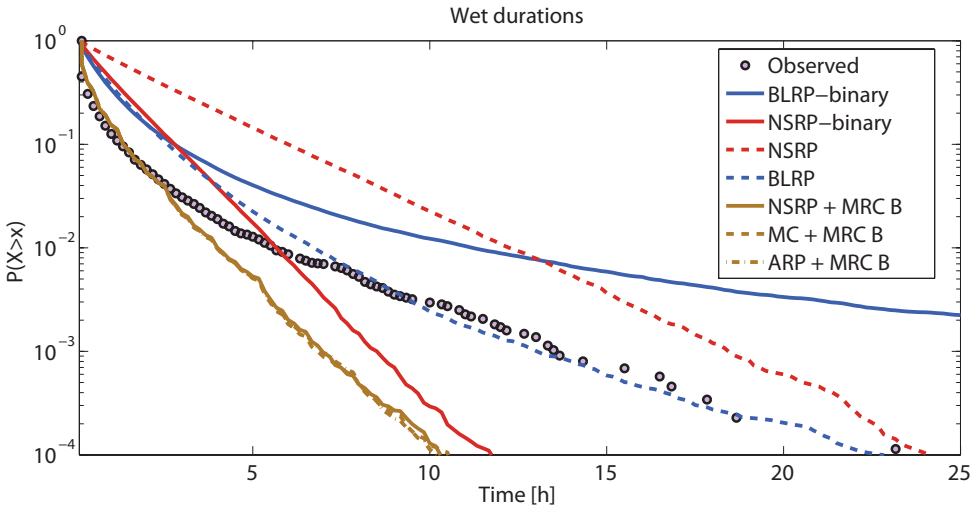


FIGURE 4.11: Comparison between temporal stochastic models of the exceedance probability for the wet spell durations as estimated from the 10 min time scale. The results are for the Zurich station (SMA) for the spring season (MAM).

The second problem is the overestimation of the small scale autocorrelation. One way to illustrate the overestimation of the autocorrelation structure using the Poisson cluster models can be shown using their power spectra. Figure 4.12(b) shows that, generally the Poisson cluster models lead to a "steeper" decay of the spectral densities for the high frequencies (typically for temporal scales < ∼5 hours). The introduction of the jitter, leads to the exact opposite result. As already explained, a multiplicative noise term such as the jitter leads to a typical "whitening" (e.g. flat spectrum) of the power spectral densities for the very high temporal scales. This can be a serious artefact which to the best of my knowledge, has not been discussed in the literature so far. Generally though as it can be shown in the summary of the statistics in the table 4.4, the inclusion of the jitter, improves the small scale issues of the classic Poisson cluster models.

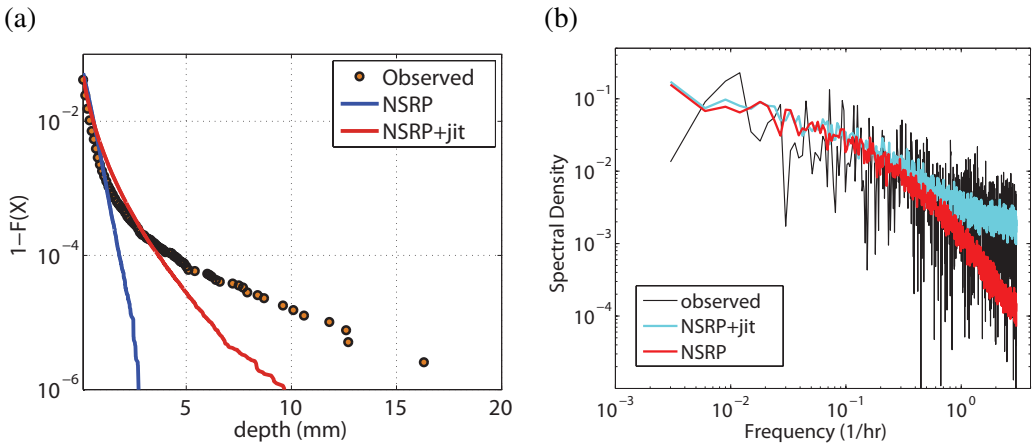


FIGURE 4.12: Differences between the NSRP model and its equivalent with the introduction of a jitter. (a) The exceedence probability of depth accumulation for the 10 min temporal scale and (b) the power spectral densities. The data are for the Zurich station during spring (MAM).

4.4.2 Model Combinations

The temporal nesting of a coarse scale parsimonious stochastic model and a fine-scale disaggregation model is based on the premise that nesting parsimonious stochastic models for the temporal scales where they perform best can enhance the range of temporal scales of applicability of the composite models in terms of the most significant statistical properties of the precipitation process.

The general assessment for the Zurich station can be shown in figures 4.14, 4.19.

As a first indication of the potential of the model, a direct comparison with the NSRP model alone, that has been traditionally used for precipitation simulation across temporal scales, can be used.

One of the most important improvements, and common to all the composite models, is a substantial enhancement in the reproduction of the probability distribution for the high resolution precipitation depths. The probability distributions of the wet spell durations are significantly better represented at small scales even though, in all of the cases, there is a systematic underestimation of the tails for the 10-min aggregation interval (Figure 4.11). This means that the composite models fall short in capturing long wet spells observed at this resolution.The behaviour of all the composite modelling approaches was almost identical.

A first result for the disaggregation procedures that have been chosen here is that the most parsimonious one, the beta lognormal canonical MRC gives systematically the worse results, overestimating severely the higher order moments (variance and skewness). In general, for the microcanonical versions of the MRC, the more complex the model is, the higher the performance concerning the high temporal scale distribution is. As shown in Figures 4.13 the reproduction of the tail of the distribution is much better, even though not perfect, when the more complex MRC model B is used. This also reflects in a more efficient reproduction of the extremes (Figure 4.13(b)). For all the other statistics, the improvement is comparable for both versions of the MRC model.

The probability of zero precipitation is reasonably well reproduced by all model combinations and no clear patterns could be identified.

In terms of the distributions of the wet spells, all the nested models perform well in reproducing the body of the distribution (at least for the 10 min scale) but underestimate the tails, leading to shorter rainy durations (figure 4.11). The same is true as well for the durations of the dry spells.

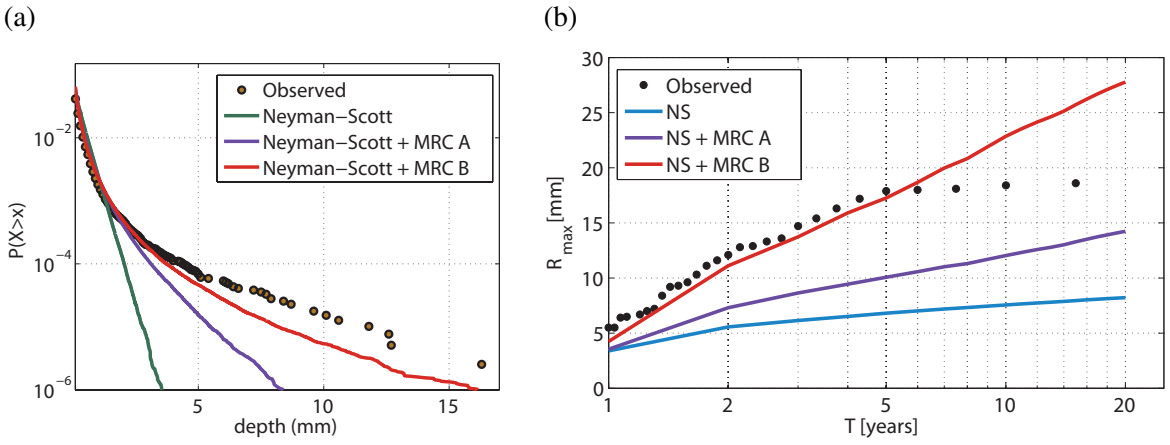


FIGURE 4.13: Generated exceedance probabilities (a) and annual maxima (b) for the 10 min aggregation interval according to the Neyman-Scott model and its combination with the MRC model A and MRC model B.

In table 4.4 only the lag-one correlation coefficient is reported. At temporal scales coarser than the daily (1280 mins), where the performance of only the external models can be validated, the NSRP model outperforms the other two models. The Markov chain and ARP external model systematically underestimate the autocorrelation function due to the assumption of the iid daily rainfall depths. This underestimation of the correlation can result in a relatively poor representation of the tails of the distribution at temporal scales coarser that 1 day (see Figure 4.14). At the intermediate scale (2.7 hours) all the model combinations underestimate the lag-one correlation coefficient. The underestimation is more dependent on the station and season rather than on the model but appears to be a structural deficiency of the MRC model itself [e.g., *Olsson*, 1998; *Lombardo et al.*, 2012; *Pui et al.*, 2012]. For the highest resolution (10-mins) the models do reasonably well, with a tendency to a slight underestimation which is also station dependent. The underlying reason can be found in the parametrizations of the microcanonical MRCs that are not always sufficient to capture the entire variability/correlation of rainfall. In this case, heavier parametrizations such as those in *Serinaldi* [2010b] could improve the situation, but would increase the model complexity.

4.4.3 Multiplicative Model Combination

The alternative approach for the combination of different models that was explored here was multiplying two independent processes, one modelling the occurrence of rainfall and the other its in-storm variability. This approach is quite parsimonious in terms of model parameters, and essentially introduces intermittency in a similar way, as the the atom at zero is included in the MRCs. In figure 4.16 the results for the Zurich station are shown. For this example, it can be shown that the model performance is excellent for both the reproduction of the probability distributions across scales, the probability of zero precipitation and the correlation structure.

The major problem of this modelling approach is that since the process that controls the in-storm variability is calibrated only for the finest temporal scale, it cannot be assured a priori that the model efficiency can be adequate for any temporal scale. This would happen if the parametric choices of all the model components are sufficient and provide a good fit to the data. As clearly illustrated in tables 4.3 and 4.4, the model efficiency is excellent for the finest aggregation interval but somewhat problematic for the average multi-scale efficiency (table 4.3). The problems mainly concern the higher order moments and the autocorrelation. The errors exhibit a systematic pattern. Whenever the autocorrelation function for large lags is overestimated at the finest scale, the higher order moments are overestimated as well for higher aggregation intervals and vice versa. This result is trivial, but reveals the potential problem of the choice of the FGN as the driving force of the in-storm variability. The reason for this is that, loosely speaking, the higher the correlation at the original scale, the higher the tails of the distributions are for increasing temporal aggregations. An example is shown in figure 4.15, where the exceedance probabilities of an exponentiated FGN for several aggregation scales is shown. It is clear that for higher values of the H exponent (i.e. higher correlations), the tails of the distributions are heavier (i.e. larger higher order moments), indicating a slower convergence to a normal (or Lévy alpha-stable) distribution.

The major reason behind those discrepancies is the selection of the FGN model as the Gaussian process that drives the process of the in-storm variability. Probably a more flexible parametric model for the autocovariance would be more appropriate. An alternative good candidate for modelling the in-storm variability can be the class of the ARFIMA (autoregressive fractionally integrated moving average) models, since it can cope well for both short and long term correlation properties. This approach was not investigated further due to the cumbersome and not robust parameter estimation [*Montanari et al.*, 1997;

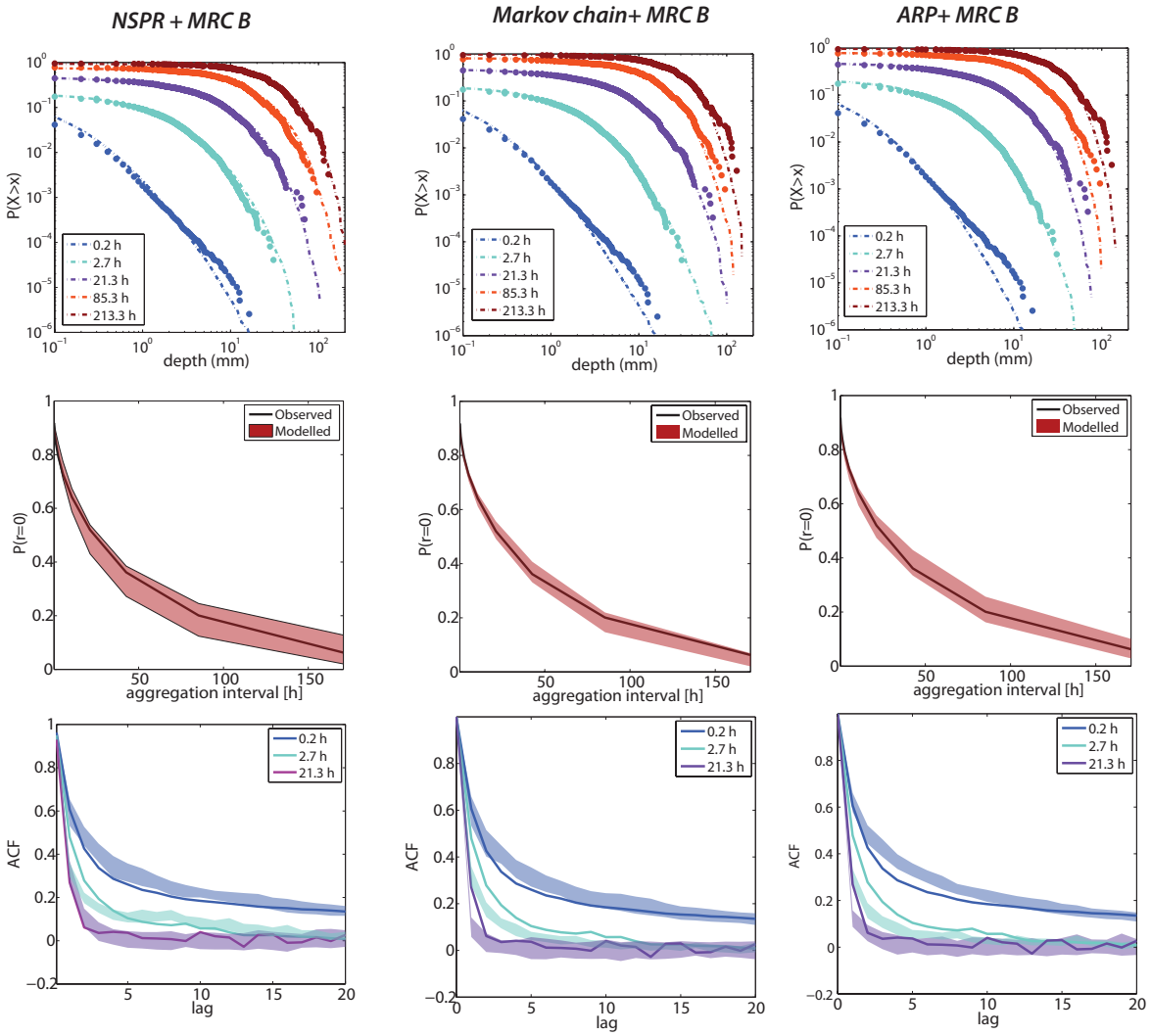


FIGURE 4.14: Comparison of the efficiency of the nesting model approaches for temporal precipitation. The left column corresponds to the combination of the NSRP model and the Model B of the microcanonical MRC. The middle corresponds to the combination of a Markov chain and MRC model A and the right to the combination of an Alternating renewal process and the MRC model A. On the first row The probability of zero precipitation is shown (dots-observed, lines-simulated), in the second the probability of precipitation occurrence (black line-observed, red area 95% confidence interval of the simulated) and in the third the respective autocorrelation functions (lines-observed, shaded areas, 95% confidence interval of the simulated series).The data are for the Zurich station during spring (MAM).

Taqqu and Teverovsky, 1998; *Fatichi et al.*, 2009]. Another possible candidate could be the very well known linear ARMA models [*Gyasi-Agyei and Willgoose*, 1999], that however have problems, if long term correlations are of major importance [*Papalexiou et al.*, 2011]. Overall the choice of an FGN model is good compromise between the parsimony, robustness of the calibration procedure and the ability of the model to capture long term correlations and statistics across scales.

In addition, as mentioned before, for this way of combining stochastic models there is a restricted range of values that the autocorrelation can take. It can be thus illustrated that for stations that exhibit very strong short term correlations, such as the one located in Mt Cook (figure 4.17(a)), the underestimation of the

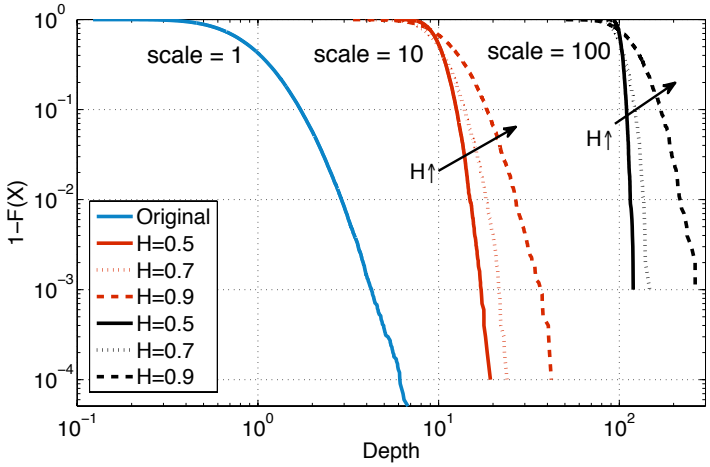


FIGURE 4.15: Exceedance probability distributions for an exponentiated Fractional Gaussian Noise for three different aggregation intervals. The three cases plotted here correspond to Hurst exponents 0.5, 0.7, 0.9. The data at scale 1 follow a lognormal distribution with parameters $\mu = -0.1$ and $\sigma^2 = 0.5$.

autocorrelation structure of the model can be considerable. Another reason contributing to this problem, could be the calibration procedure of the FGN model that is based on the lag 1 autocorrelation coefficient alone. Possibly, if a least square fitting of the autocorrelation function for a large number of lags was used, the results could be better. The robustness of the estimation procedure chosen here is a topic that needs further investigation. It is also true that since the two processes that are multiplied are mutually independent, this can as well lead to some non-realistic features in the simulated series, such as abrupt high intensities in the beginning or end of a storm event.

Another result, as can be illustrated in figure 4.17(b) is that in general, both the binary versions of the NSRP and MBLRP models perform very well in reproducing the probability of zero precipitation as rainfall occurrence models. It should be noted that the MBLRP model when calibrated on its probability of zero occurrence alone, is much worse in reproducing the distribution of the wet spells (figure 4.11) in comparison to the binary version of the NSRP which gives systematically better results. It is difficult to asses whether this behaviour is a general artefact of the MBLRP model, or its calibration procedure, but it is an issue that should be taken into account in applications.

Model Intercomparison

Summing up all the previous findings concerning the applicability of a wide range of stochastic modelling techniques, the general result is that the best approach depends on the requirements of each application. Indeed, no model was found to be the optimal choice, none of them was free of problems. The main topics that have to be taken into account are the ability of each model to simulate satisfactorily well the temporal precipitation structure across scales and at the same time being generally applicable and flexible in order to be considered as a universal model.

A first insight for this can be shown in the figures 4.18, 4.19, and the tables 4.4, 4.3, where the overall model intercomparison results are shown. In figure 4.18 the efficiency of all the modelling approaches in capturing precipitation extremes across temporal scales is shown. Generalized results for the Zurich station SMA can be found in figure 4.19 where the relative errors of the main statistical descriptors of the

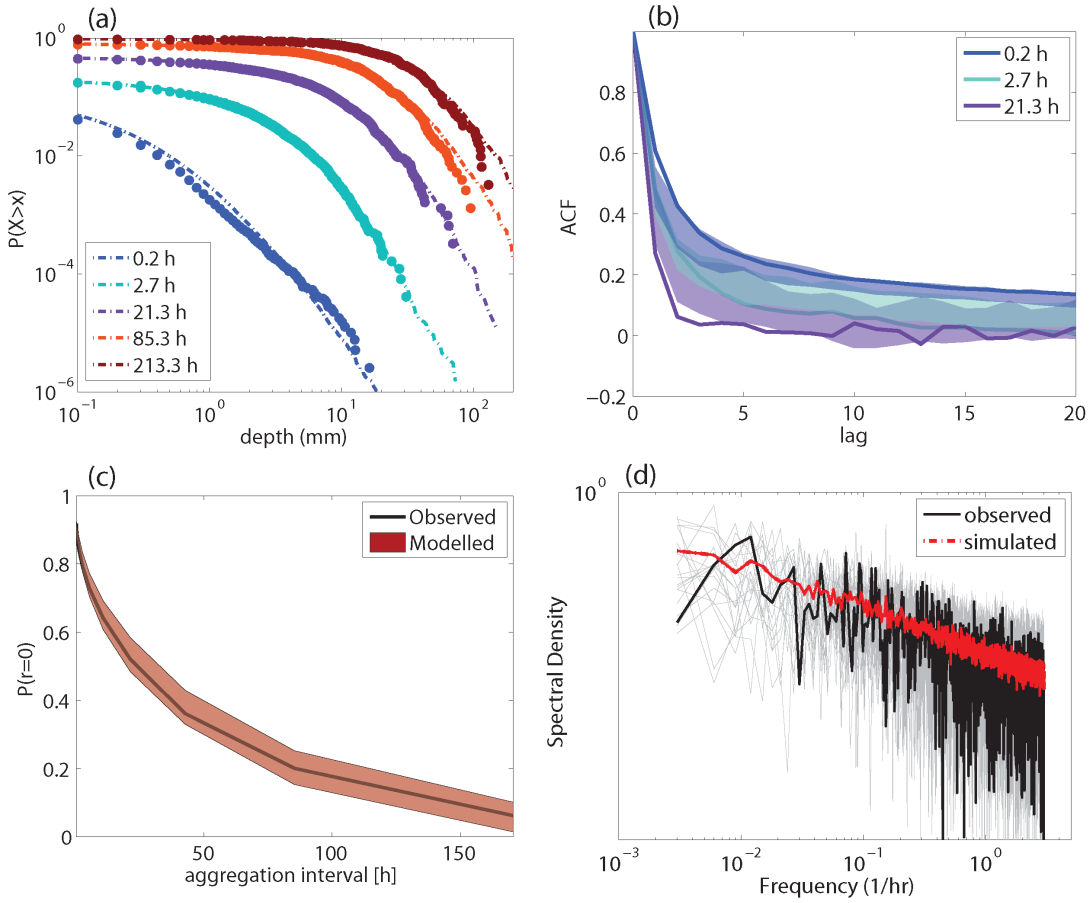


FIGURE 4.16: Efficiency of the hybrid model MBLRP-FGN for precipitation.(a) The exceedance probabilities, (b) The probability of zero precipitation, (c) the autocorrelation function and (d) the power spectrum. The data correspond to the Zurich station during spring (MAM).

process across scales are shown, revealing the multi-scale performance of all the models.

Figure 4.19 illustrates the difficulty of the model comparison attempted here, since model efficiency is highly dependent both on the modelling approach and the temporal scale for. Moreover the validation reported here is the composite effect of both the efficiency of the model structure and the efficiency of its calibration procedure. Separation between those two is not straightforward, and thus absolute statements could be uncertain. However, despite the large uncertainty of the results, some clear patterns that lead to a concrete model intercomparison could be identified.

One major result is that the simpler the structure of the model is, the most predictable its efficiency is. This can be easily illustrated by the fact that the simple Poisson cluster models have approximately the same efficiency for all the stations, and at the same time, for all the stations the same problems occur. This means that this kind of model can be used as a general application tool, but at the same time the modeller should be aware of its structural deficiencies that cannot be resolved. Summarizing their major problems are their incapability of reproducing small scale variability (less that 1 hour aggregation intervals) and also, their inability to reproduce well the distributions of the durations of the dry and wet spells.

The same result holds true also for the various disaggregation/downscaling techniques that have been

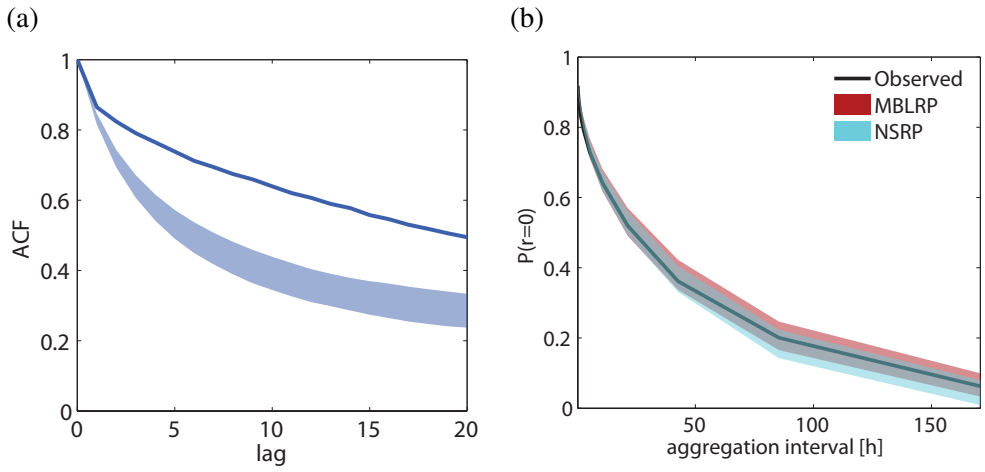


FIGURE 4.17: (a) Typical underestimation of the high lag autocorrelation function of the composite MBLRP-FGN model. The data are for the Mt Cook station during winter (JJA) for the 10 min aggregation scale(b) Comparison of the probability of zero precipitation for the binary NSRP and MBLRP models, across several temporal scales for the Zurich station during spring (MAM).

studied here. The more parsimonious the choice is (e.g. the beta lognormal model), the more predictable their behaviour is. It can be shown from the table 4.3 that the same errors, both in terms of magnitude and sign, occur for all the stations. In this case the main problems were the substantial overestimation of the higher order moments and also the underestimation of the autocorrelation functions. As far as the the microcanonical models are concerned, for all the cases the more complex MRC Model B outperformed the simpler MRC model A, clearly indicating that a dependence structure between the distribution of the cascade generator and the temporal scale and intensity exists.

Another general result of this study is that the model combinations used here, can significantly improve the efficiency of the traditional Poisson cluster models for the fine temporal scales (table 4.4; figure 4.18), and especially the reproduction of the tails of the distributions and their extreme values. The overall multi-scale gain on the performance though is not very significant, mainly because the Poisson cluster models are excellent for reproducing statistic above one hour. It should be noted though that the results reported in the table 4.3 are favouring the Poisson cluster models since they have been explicitly calibrated on the exact same statistics.

The general result that can be extracted from this analysis is that simple parsimonious modelling procedures can be robust for precipitation modelling, but typically not able to reproduce well all the properties of the the observed data structure, and more complex models that are generally heavily parametrized, can be very good for the data that have been developed, but their ability is not universal, in the sense that a direct extrapolation can lead to erroneous discrepancies.

This study thus indicates that none of the modelling techniques that have been applied here are free of problems, and the model that has to be chosen, should reflect the most significant aspects of the problem, e.g. the best reproduction of extremes, if flood risk analysis is concerned, the best reproduction of the dry-wet probabilities across scales if the model has to be used for drought analysis and water resources management etc. Each time the nature of the problem should guide to the best choice since there seems to be no universal model that can deal with all the aspects of the precipitation process across the entire range of temporal scales that are of hydrological interest.

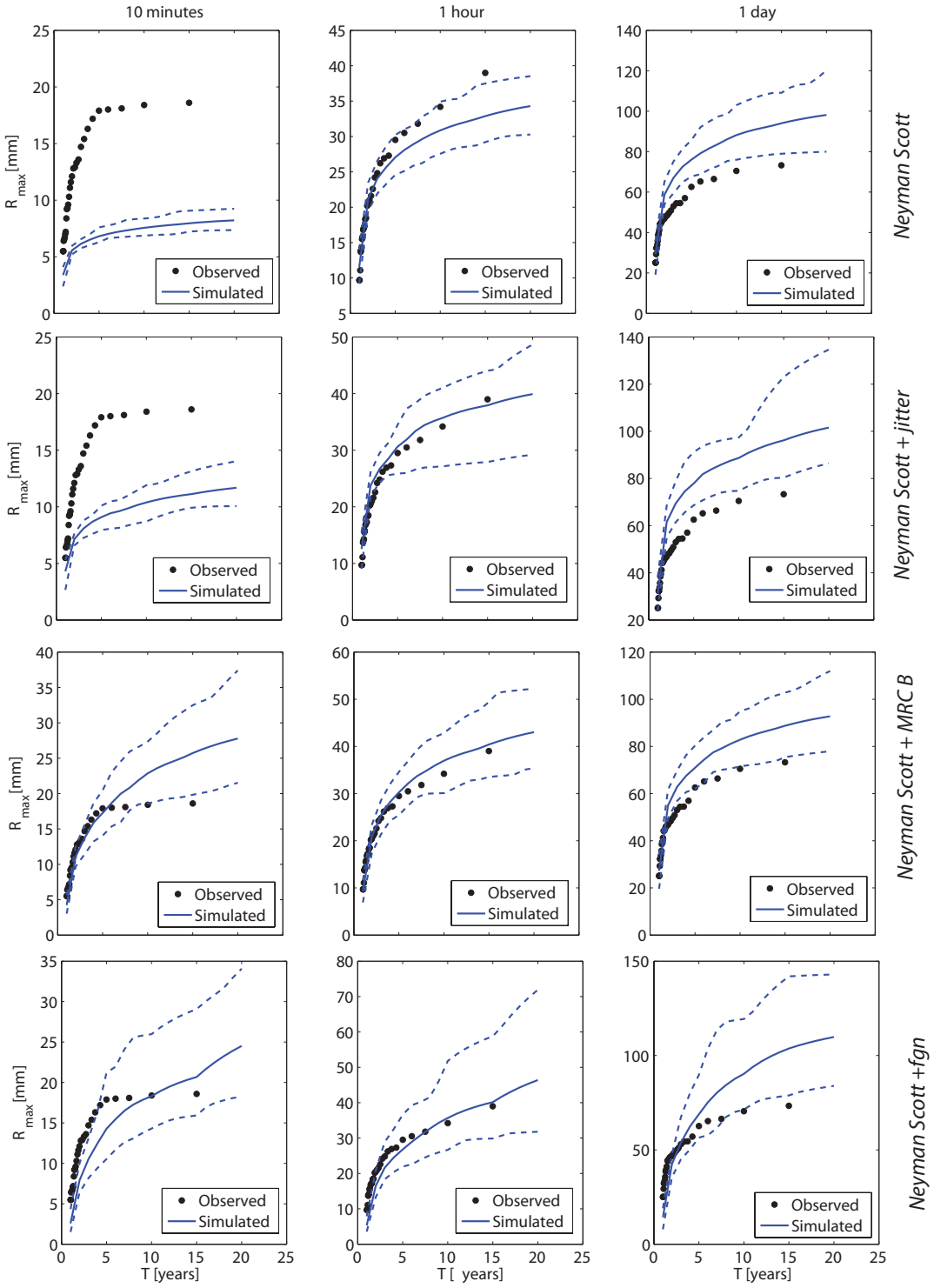


FIGURE 4.18: Comparison of the models for the reproduction of the extremes for the aggregations intervals 10min-1 hour- 1day.(a) The Neyman Sott model (b) The improved NSRP model with the jitter(c) The nested NSRP model with MRC B and (d) The combination of the binary NSRP and the FGN. The results correspond to the SMA station during summer (JJA).

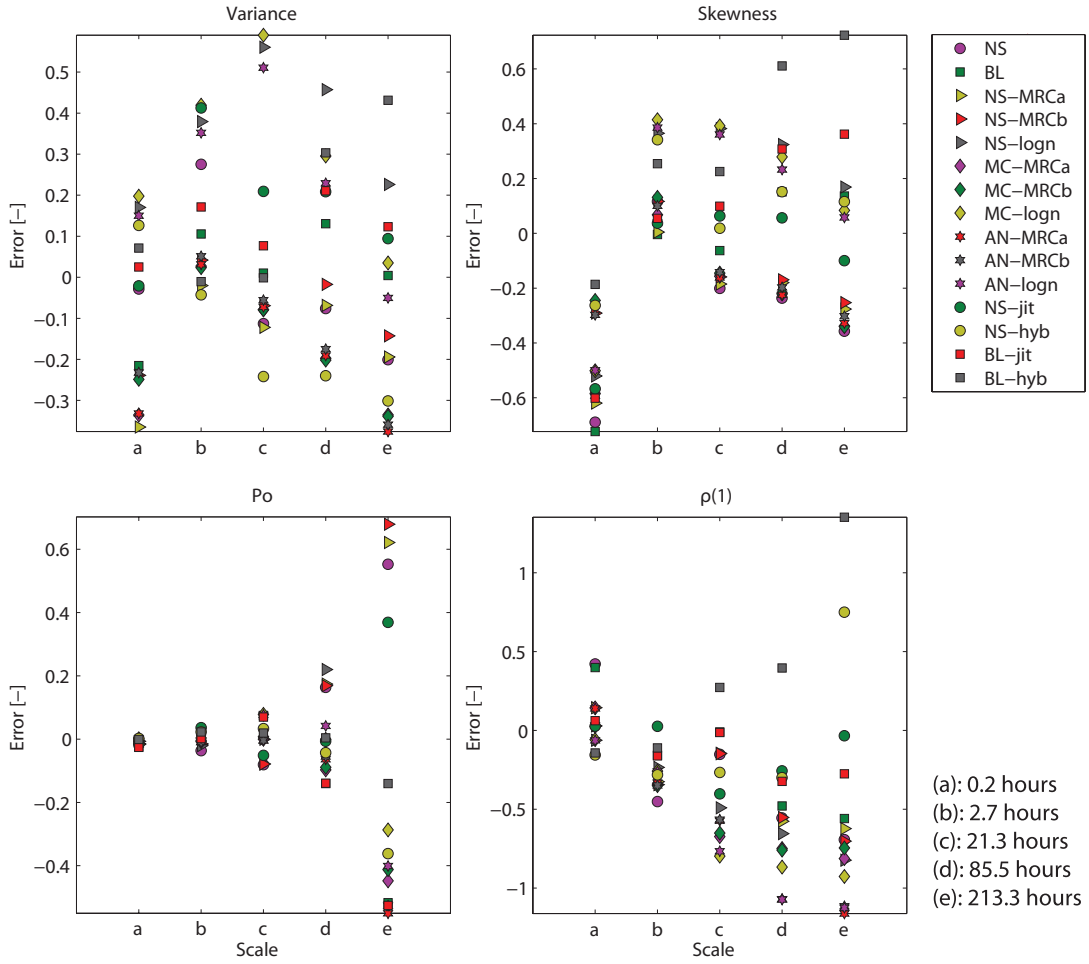


FIGURE 4.19: Comparison of the relative error values for the investigated stochastic models. The results correspond to the Zurich station (SMA) for the spring season (MAM).Relative errors of the (i) simulated variance (ii) the coefficient of skewness (iii) the probability of zero precipitation and (iv) the lag-1 autocorrelation coefficient.

This extensive inter-comparison thus, that is one of the very few reported in the literature, indicates that further research is needed in the topic of precipitation stochastic modelling. The overall conclusion of this extensive model intercomparison can be that the development of a generic stochastic model for rainfall that can be adequate for all the temporal scales is still an open and challenging question, even after a thorough investigation in the hydrological community of more that 4 decades. My belief is that starting with a generalized model intercomparison, such as the one reported in this chapter can lead to a clear illustration of the problems we deal with, pose questions and reveal new open questions for further research in this topic. Thus such a study can serve as a first step towards the the development of new ideas of stochastic modelling of precipitation that could eventually lead to a universally accepted generic and parsimonious model.

TABLE 4.3: Mean relative errors of the statistics simulated by the temporal stochastic models for all the stations. The values correspond to the average values of the relative errors for all seasons and for temporal scales [10min-2.7 hr-21.3 hr].

	'NS'	BL'	'NS-MRCa'	'NS-MRCb'	'NS-logn'	'MC-MRCa'	'MC-MRCb'	'MC-logn'	'AN-MRCa'	'AN-MRCb'	'AN-logn'	'NS-jit'	'NS-hyb'	'BL-jit'	'BL-hyb'
	Mt Cook														
Variance	-0.022	0.000	0.227	0.005	0.332	0.311	0.043	0.414	0.298	0.044	0.434	0.020	-0.257	0.008	-0.155
Skewness	-0.098	-0.114	0.434	-0.044	0.194	0.727	0.094	0.396	0.675	0.082	0.395	-0.080	-0.190	-0.107	-0.071
P0	0.035	0.063	0.015	0.016	0.042	-0.009	-0.009	0.007	-0.013	-0.011	0.001	0.042	0.003	0.061	-0.026
ρ(1)	0.009	-0.029	-0.248	-0.178	-0.307	-0.433	-0.356	-0.450	-0.413	-0.337	-0.438	-0.040	-0.166	-0.018	0.267
	Firenze														
Variance	-0.042	0.031	-0.120	-0.098	0.136	-0.034	0.030	0.365	0.053	0.126	0.477	0.021	0.145	0.071	0.280
Skewness	-0.210	-0.195	-0.122	-0.107	0.193	-0.001	0.043	0.433	-0.034	0.004	0.390	-0.182	0.100	-0.179	0.257
P0	0.013	0.043	0.000	0.001	0.009	0.005	0.005	0.013	-0.011	-0.010	-0.001	0.024	0.028	0.043	-0.015
ρ(1)	-0.058	0.043	-0.235	-0.278	-0.232	-0.228	-0.278	-0.249	-0.164	-0.190	-0.220	-0.065	0.062	-0.020	0.602
	Lucky Hills														
Variance	-0.041	0.008	-0.129	-0.082	0.359	-0.131	-0.070	0.356	0.038	0.114	0.727	0.057	0.266	0.055	0.303
Skewness	-0.138	-0.109	-0.085	-0.017	0.382	-0.079	0.010	0.336	-0.173	-0.094	0.282	-0.055	0.192	-0.065	0.279
P0	0.004	0.007	-0.002	-0.002	0.010	-0.001	-0.001	0.011	-0.013	-0.013	0.000	0.008	0.017	0.007	0.003
ρ(1)	-0.125	-0.088	-0.234	-0.283	-0.247	-0.287	-0.338	-0.287	-0.174	-0.231	-0.218	-0.191	-0.025	-0.068	0.222
	SMA														
Variance	-0.070	-0.036	-0.177	-0.136	0.210	-0.042	0.002	0.412	-0.029	0.007	0.413	0.058	-0.076	0.041	0.068
Skewness	-0.242	-0.207	-0.173	-0.077	0.067	-0.001	0.133	0.261	-0.016	0.111	0.245	-0.148	0.007	-0.135	0.172
P0	-0.015	0.027	-0.021	-0.022	0.012	-0.005	-0.005	0.027	-0.010	-0.010	0.023	0.014	0.013	0.026	-0.006
ρ(1)	-0.038	0.075	-0.154	-0.190	-0.245	-0.302	-0.342	-0.348	-0.273	-0.312	-0.330	-0.085	-0.150	-0.015	0.211

Table Abbreviations
NS: Neyman-Scott
BL: Modified Bartlett-Lewis
NS-MRCa: Nesting of NS and type A of the MRC
NS-MRCb: Nesting of NS and type B of the MRC
NS-logn: Nesting of NS and the canonical beta lognormal MRC
MC-MRCa: Nesting of a Markov chain and type A of the MRC
MC-MRCb: Nesting of a Markov chain and type B of the MRC
MC-logn: Nesting of a Markov chain and the canonical beta lognormal MRC
AN-MRCa: Nesting of an alternating renewal process and type A of the MRC
AN-MRCb: Nesting of an alternating renewal process and type B of the MRC
AN-logn: Nesting of an alternating renewal process and the canonical beta lognormal MRC
NS-jitter: NS with jitter
NS-hyb: Multiplicative combination of NS and FGN
BL-jitter: BL with jitter
NS-hyb: Multiplicative combination of BL and FGN;

TABLE 4.4: Mean relative errors of the statistics simulated by the temporal stochastic models for all the stations for the finest temporal scale (10 min)

	'NS'	BL'	'NS-MRCa'	'NS-MRCb'	'NS-logn'	'MC-MRCa'	'MC-MRCb'	'MC-logn'	'AN-MRCa'	'AN-MRCb'	'AN-logn'	'NS-jit'	'NS-hyb'	'BL-jit'	'BL-hyb'
	Mt Cook														
Variance	-0.076	0.116	0.500	0.030	0.637	0.591	0.050	0.739	0.577	0.054	0.763	0.024	-0.015	0.126	-0.029
Skewness	-0.168	-0.096	0.995	-0.011	0.250	1.390	0.079	0.471	1.320	0.073	0.468	-0.063	-0.024	-0.089	0.017
P0	-0.027	-0.005	-0.033	-0.028	0.010	-0.031	-0.027	-0.001	-0.032	-0.028	-0.003	-0.023	-0.002	-0.007	-0.069
ρ(1)	0.102	0.057	-0.112	-0.016	-0.229	-0.111	-0.010	-0.227	-0.111	-0.011	-0.228	0.005	-0.153	0.050	-0.101
	Firenze														
Variance	-0.062	-0.038	-0.152	-0.080	-0.074	-0.099	0.025	0.105	-0.016	0.122	0.204	-0.028	0.023	0.069	-0.028
Skewness	-0.476	-0.518	-0.200	-0.159	-0.125	-0.119	-0.025	0.076	-0.134	-0.061	0.060	-0.404	-0.098	-0.426	0.011
P0	0.028	0.035	0.010	0.011	0.009	0.012	0.013	0.010	0.008	0.009	0.007	0.031	0.012	0.035	-0.039
ρ(1)	0.239	0.195	-0.196	-0.231	-0.148	-0.179	-0.216	-0.146	-0.183	-0.216	-0.155	0.152	-0.047	0.035	-0.055
	Lucky Hills														
Variance	-0.075	-0.035	-0.182	-0.044	0.093	-0.189	-0.033	0.119	-0.026	0.157	0.411	0.032	-0.016	0.018	-0.006
Skewness	-0.433	-0.480	-0.178	0.000	0.019	-0.168	0.059	-0.001	-0.249	-0.056	-0.062	-0.214	-0.062	-0.441	-0.062
P0	0.007	0.009	0.000	0.000	0.003	0.001	0.001	0.004	-0.002	-0.002	0.002	0.005	0.004	0.009	0.001
ρ(1)	0.207	0.121	-0.300	-0.385	-0.244	-0.293	-0.384	-0.242	-0.298	-0.380	-0.236	0.009	0.028	0.060	-0.048
	SMA														
Variance	-0.119	-0.159	-0.275	-0.190	0.058	-0.176	-0.067	0.226	-0.164	-0.063	0.227	-0.025	0.042	0.011	-0.007
Skewness	-0.545	-0.532	-0.327	-0.106	-0.352	-0.181	0.128	-0.228	-0.194	0.106	-0.247	-0.380	-0.108	-0.393	-0.089
P0	-0.001	-0.019	-0.015	-0.016	-0.002	-0.012	-0.012	0.001	-0.013	-0.014	0.000	-0.003	0.002	-0.020	-0.034
ρ(1)	0.341	0.298	0.012	-0.079	-0.081	0.024	-0.087	-0.083	0.023	-0.085	-0.082	0.081	-0.100	0.060	-0.056

5

PRECIPITATION STRUCTURE IN SPACE

One of the principle concerns in hydrology but also other fields such as meteorology is the identification of the spatial structure of precipitation. Knowledge of the organization of precipitation fields in space is crucial for the correct estimation of runoff generation and concentration and therefore flood discharges.

In this chapter, the analysis of the spatial structure of precipitation fields for the Alpine area of Switzerland is attempted and possible connections with the physical precipitation triggering mechanisms are investigated.

5.1 Introduction

Traditionally the analysis of the spatial distribution of precipitation was based on sparse rain-gauge networks with very coarse temporal resolutions and this is the main reason why they were mainly focused on deriving precipitation climatologies of large areas [*Frei and Schär*, 1998, and references therein]. The problems of data unavailability are much more prominent in orographicaly complex areas, where the installation and maintenance of precipitation monitoring networks is particularly challenging.

Remote sensing techniques have provided significant improvement in the estimation of precipitation in fine space-time resolution of the orders of few square kilometres in space and at sub-hourly temporal scales. Among the remote sensing techniques the most widely used ones were satellites and weather radars, both ground and space-born or air-born [*Klazura and Imy*, 1993; *Kummerow et al.*, 1998; *Webster and Lukas*, 1992]. Especially precipitation estimates derived from ground-based weather radars have been considered reliable enough to be used in hydrological modelling [*Ivanov et al.*, 2004; *Carpenter et al.*, 2001].

The vast majority of the literature can be roughly divided into two categories. The first category consists of studies that are mainly conducted by meteorologists and are oriented on the identification of event specific precipitation structures and reveal the linkages with the physics of the precipitation process and its generation mechanisms (convection, etc) [*Houze*, 1993; *Rotunno and Houze*, 2007]. Most of these studies are based on specific events during experiments, e.g. MAP observation period [*Houze et al.*, 2001]. The second category is mainly hydrologically oriented and aims at the general statistical description of the spatial features of precipitation, commonly neglecting the physical processes involved. This kind of investigation has driven new developments in the stochastic modelling of rainfall in space. Studies in the

second category are usually based on the infrastructure of the operational meteorological networks [e.g. *Gebremichael et al.*, 2008].

Several statistical descriptors have been employed in order to describe in a quantitative manner the spatial structure of the precipitation such their correlation functions [e.g. *Ciach and Krajewski*, 2006; *Bacchi and Kottegoda*, 1995; *Habib et al.*, 2001; *Crane*, 1990], the distribution of the sizes of rain cells [e.g. *Goldhirsh and Musiani*, 1986] etc. Even though the traditional structure descriptors gave a deep insight about precipitation structure, a generalized statistical theory had yet to be established.

The theory of scale invariance for spatial precipitation, originally inspired by the analysis of temporal rainfall provided the excellent mathematical background for such a generic theory of the precipitation structure. During the last three decades, this theory has been widely applied to spatial precipitation data and has become one of the main analysis tools, since it can describe the precipitation spatial structure across scales with a unified and parsimonious way. The literature concerning the description of spatial precipitation as a fractal process is vast [e.g. *Lovejoy*, 2008; *Deidda*, 1999; *Gires et al.*, 2012; *Verrier et al.*, 2010, among others]. Precipitation in space has been characterised both as a simple scaling [e.g. *Koutsoyiannis et al.*, 2011] and a multi-scaling process [e.g. *Over and Gupta*, 1994].

An ultimate goal of most of the studies was not just to quantify the statistical structure of spatial rainfall, but also to link its statistical properties with various descriptors of the natural process, trying to identify relationships that could explain the statistical properties as a function of the underlying physics of the process. Such descriptors span from simple statistical properties (e.g. mean intensities, etc) to the description of the thermodynamics of the atmosphere. A better understanding of the processes and the identification of such relationships can provide strong prognostic power for applications of stochastic modelling of precipitation as well.

In one of the earliest studies linking the scaling properties of precipitation to such descriptors, *Over and Gupta* [1996] found a clear connection between the scaling of the intermittency and the "large scale forcing" (i.e. mean areal precipitation). Similar relationships between large scale precipitation intensities and the scaling parameters of the fields have been also found on other studies [e.g *Deidda et al.*, 2006; *Pathirana et al.*, 2003a; *Jothityangkoon et al.*, 2000]. *Perica and Foufoula-Georgiou* [1996] found a clear dependence of the scaling of the standardized fluctuations of the precipitation fields across scales with the pre-storm convective available potential energy (CAPE). Also *Nykanen and Harris* [2003] and *Nykanen* [2008], by analysing radar derived precipitation images for an orographically complex area in the US, concluded that some of the parameters describing the multiscaling behaviour of the precipitation fields are dependent both on the storm direction (i.e. if the storm is on the windward or leeward side of the mountain range) and the mean altitude above which heavy precipitation occurs. *Ebtehaj et al.* [2010] recently also found a dependency of the correlation structure of the precipitation fields on the orography of the area they analysed. Overall all these studies find point out that orography, convection and large scale circulation combine to produce patterns in spatial rainfall with distinct statistical properties.

5.2 Measurement Artefacts on Precipitation Structure Estimation

One major problem associated with the characterization of the spatial structure of precipitation arises from the measurement errors that occur. Since the analysis of the spatial fields in this study will be conducted using weather radar data, the measurement errors connected with this remote sensing technique will be explored similarly to the analysis conducted in 3.3.2 for the rain-gauge induced errors.

Although several studies have analysed the structure of precipitation fields in space using radar data [e.g. *Kang and Ramírez*, 2010; *Verrier et al.*, 2010; *Rodriguez-Iturbe et al.*, 1998], very few have focused on the possible artefacts associated with well known measurement errors. One of the first studies on this topic was conducted by *Krajewski et al.* [1996]. They investigated the effects of various measurement errors associated with weather radars on the estimation of the mean, variance, autocovariance function and the parameters of a MRC model. They concluded that the biases of the inferred statistics can be highly significant. *Villarini et al.* [2007] and *Mandapaka et al.* [2010] analysed in a Monte-Carlo framework the problems that autocorrelated noise corruption on radar-derived precipitation fields can have on the estimation of the moment scaling functions, autocorrelation functions and power spectra. Their analysis was restricted to a small data set, derived from a small sample of storms and thus difficult to generalise.

The scope of the analysis presented here is to expand and generalize the effects of measurement induced errors on the quantification of the precipitation structure in space and give further insight on bias and uncertainty in various common methods used to on describe the precipitation structure in space.

5.2.1 Radar Measurement Errors and Uncertainties

Even though the estimation of precipitation using weather radars is very useful in order to capture the process details in very high temporal and spatial resolutions, the estimation procedure suffers from various inherent problems of the recording techniques. An excellent review concerning the use of radars in hydrology can be found in *Berne and Krajewski* [2012]. Here a short review of the most serious errors is presented.

The first source of measurement errors depends on the interaction between the transmitted electromagnetic wave and the ground. Three major errors depend on this interaction. First, when the transmitted electromagnetic waves hit the ground surface, they are backscattered and this results in very strong recorded echoes. This behaviour is known as ground clutter. This type or error can be divided into two components, the systematic and random errors. Systematic ground clutter can be defined as the echoes that are backscattered to the radar receiver when no precipitation occurs. This type of error is solely dependent on the topography of the area and is easy to identify and eliminate. The second kind of clutter, occurs due to the same interaction between the transmitted radiation and the ground, but does not always appear at the same place. The reason for that is the anomalous propagation of the electromagnetic wave. Various image processing techniques have been successfully applied to eliminate this source of error but still their efficiency can be low especially in very complex terrain [e.g. *Rico-Ramirez and Cluckie*, 2008; *Steiner and Smith*, 2002]. Moreover the capabilities of modern radars to estimate Doppler velocities, helped also to identify and eliminate of atmospheric clutter. Finally the third error is the beam blockage due to topography, which is the inability of the weather radar to observe precipitation behind mountains.

The second source is the signal attenuation of the transmitted radio-waves. The main reason for attenuation is the absorption and scattering of the transmitted waves from the hydrometeors [*Berne and Krajewski*, 2012]. Scattering and absorption is much more prominent for high frequency signals and in principal affects most high frequency C-band and X-band weather radars. The common behaviour attributed to this kind of error is that raincells that occur in large distances are not well captured. For this case several algorithms for the quantification of signal attenuation have been developed. Another way to overcome this issue is to estimate precipitation fields as composites from several radars. This is the main reason why dense radar networks have significant overlap (i.e. NEXRAD).

The third source of measurement errors depends on the way the recorded echoes are transformed into

precipitation intensities. The received backscattered reflectivity factors are transformed into precipitation rates according to a power law function $Z = aR^d$, Z in [dBZ] and R in [mmh^{-1}], as first identified by *Marshall and Palmer* [1948]. This transform only applies in a statistical sense (i.e. long term averages), a fact that increases the uncertainty of the instantaneous precipitation rate. Also the coefficients of this power law transform are not universal but depend on the drop size distribution. They have to be calibrated on the local meteorological conditions and are highly dependent on the form of precipitation (rain or snow). Moreover, the discrepancies of this power-law function can be enhanced due to the difficulties in linking the recorded echoes from a sample volume in the atmosphere to ground precipitation.

Finally, a major limitation concerning the hardware and data management capabilities is that in order to store radar data efficiently, the recorded intensities are binned into a limited number of intensity classes which are commonly not uniformly spaced (e.g. linear binning of the reflectivity factors Z [dBZ] leads to lognormal binning of the intensities R [mmh^{-1}]).

The detailed description of the data post-processing procedures that are applied to the reflectivity fields derived from the three C-band radars operated by MeteoSwiss in order to correct for clutter contamination and beam shielding can be found in *Germann et al.* [2006] and *Joss and Lee* [1995].

5.2.2 A Storm-Based Precipitation Radar Product Comparison

A comparison between different precipitation products can provide some first evidence on the possible artefacts of precipitation measurement by radar on the estimation of several descriptors of the spatial structure of precipitation. Here such a comparison is conducted by exploiting 2 different radar-based precipitation radar sets provided by MeteoSwiss. Such a comparison serves as a motivation for which are the main expected measurement artefacts.

The first one is the RAIN data set as described in detail in section 2.2 and the second is called NASS [*Savina*, 2011]. The infrastructure behind the two products is the same, the three C-band operational weather radars as described in section 2.2. The NASS product is an improvement of RAIN in terms of the algorithmic details of precipitation quantification.Visually, the most significant differences between the two products is the better elimination of the clutter and noise. Moreover, the spatial resolution of the two products is different. In order to have a fair comparison, the NASS product that has a finer spatial resolution was first re-sampled to a 2×2 km^2 resolution, so that it was identical to the RAIN product.

The storm that was selected for comparison took place in 18-23 August 2005 and caused disastrous flash floods in Switzerland. More information concerning this extreme event can be found in *Beniston* [2006].

The study area for this storm-based example, spans the entire spatial domain of Switzerland. In figure 5.1 the mean areal intensities as recorded in the two radar products are shown. It is clear that in terms of mean areal intensities, the differences between the products are minor.

In terms of precipitation structure, only some of the statistical and scaling descriptors are used here, since the main purpose is an illustrative indication of probable artefacts, rather than a complete quantification, which will be conducted in the following sections.

In figure 5.2 the comparison of the two parameters of the beta-lognormal MRC model estimated for both RAIN and NASS products are shown. The estimates were obtained by the commonly used moment scaling procedure, similar to the one described in 3.3.1, but applied in the two dimensions. The estimates were obtained for the absolute gradients of the precipitation fields, for reasons that will be explained in 5.3.3. It is clear that there are significant differences in the estimates between the two products despite

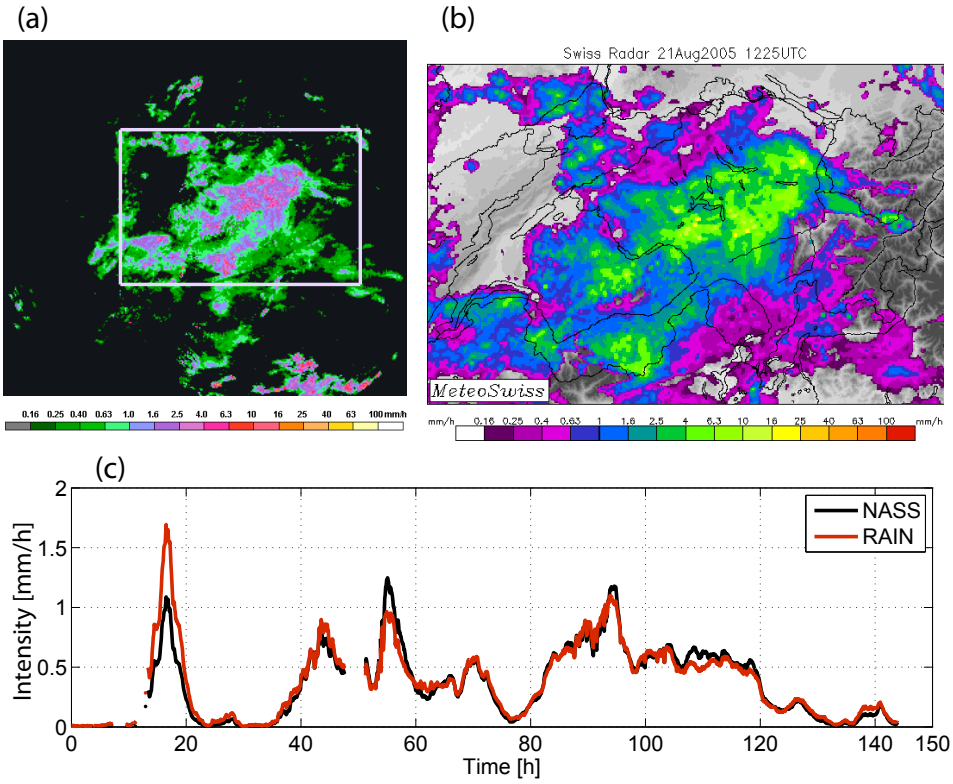


FIGURE 5.1: Radar precipitation from (a) RAIN (b) from NASS and (c) the mean areal intensity derived from the NASS and RAIN product. The discontinuity of the lines indicated data unavailability.

the fact that mean intensity is practically the same. As far as the intermittency exponent β is concerned (figure 5.2(b), there is a systematic bias but the general pattern is consistent between the two products. The estimated values of β for the RAIN product are higher. This can be attributed to the fact that RAIN underestimates the fraction of the area covered with precipitation in comparison to the NASS product. On the other hand, the spread of the estimates is small around the 1-1 line, which indicates the robustness of the estimation procedure, since it is based only on the occurrence of precipitation and not its intensity. The same does not hold true for the estimation of the σ^2 parameter (figure 5.2(a). The deviations from the one to one line are very large. Also the spread of the estimation is significant, a fact that indicated the lack of robustness of the estimation procedure for σ^2. Interestingly, both the bias and the spread are larger for images where the spatial coverage of precipitation is low i.e. images with generally low mean areal intensities. This behaviour could be caused by the remaining clutter in the RAIN product .

Another descriptor of the structure of the precipitation fields in space is their power spectrum. The power spectrum for the case of two dimensional random fields is direction dependent. For rainfall the assumption is that the precipitation fields at the mesoscale are approximately isotropic (see 5.3.2) and thus the radially averaged spectrum can be estimated. In figure 5.3, the analysis of the precipitation fields from the RAIN and NASS products is shown. It is of major importance that the large scale features described from the low frequency components have the exact same behaviour in both data sets. The main differences concern the low scale (high frequency) features. The RAIN product has a flatter behaviour for high frequencies, a fact that indicates lower correlation. This effect has been previously found to be caused by the noise corruption of the radar images [*Mandapaka et al.*, 2010] and since the only major noise difference between the two

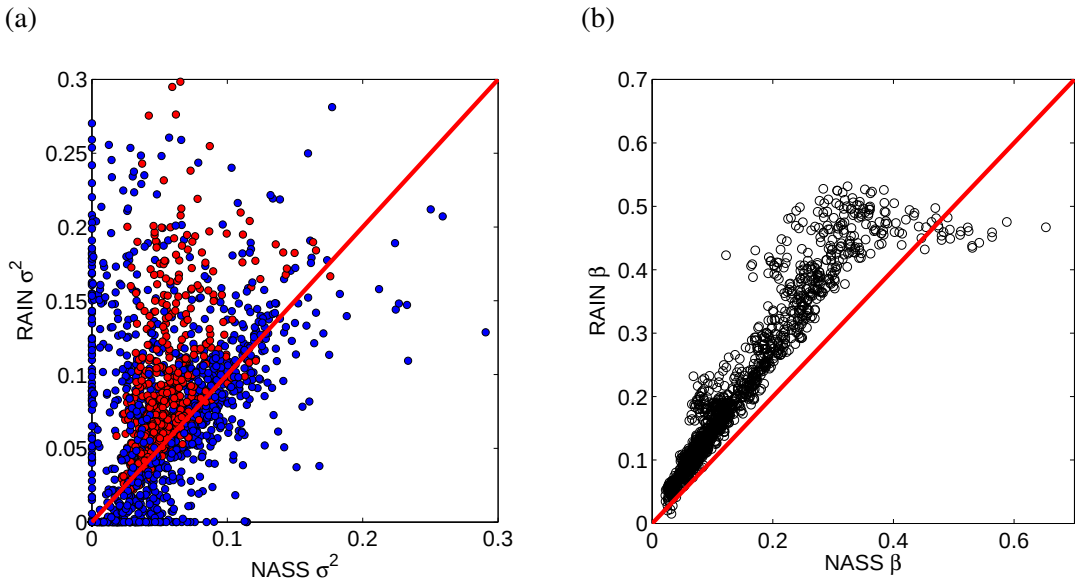


FIGURE 5.2: Parameters of the lognormal MRC cascade for the NASS and RAIN products. On the left panel, red dots correspond to the images where the spatial coverage of precipitation exceeds 30% of the area and the blue dots to the remained images.

products depends on the clutter elimination algorithms, this behaviour can be attributed to that.

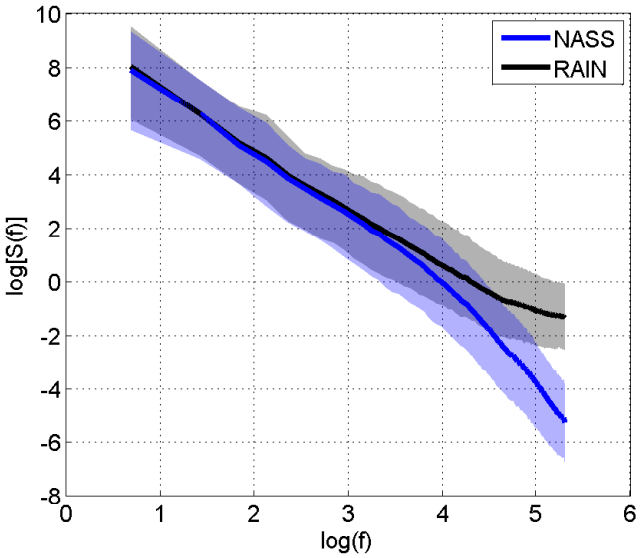


FIGURE 5.3: Radially averaged power spectral densities for the NASS and RAIN product. Lines illustrate the mean values and shaded areas the 90% confidence bounds derived out of all the radar images for the August 2005 storm.

Since the power spectrum is directly connected with the correlation of the precipitation fields, significant differences are expected to be found as well in their autocorrelation functions.

5.2.3 Evaluation Statistics

Similarly to the study conducted for the time series, the investigation of the impact of the various sources of errors on the estimation of the structure descriptors for spatial precipitation, was mainly based on Monte Carlo experiments. An illustrative example is shown in Figure 5.4. The quantification concept was identical to the time series analysis, i.e. simulated precipitation fields were perturbed with a chosen measurement artefact and then the quantification of each descriptor was done numerically through a large number of simulations.

Three basic precipitation structure descriptors were investigated :

- The multifractal scaling properties of the field
- The power-law spectral decay
- The autocorrelation function
- The marginal probability distribution functions of rain intensities

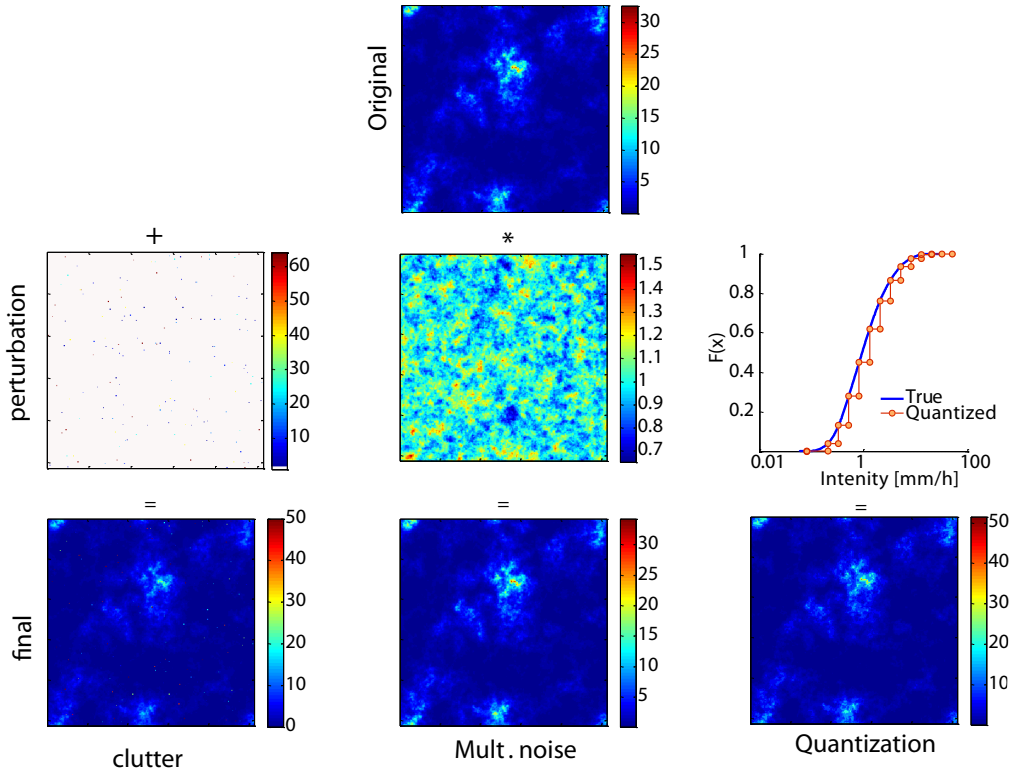


FIGURE 5.4: An example of the several radar related errors for rainfall.

In order to evaluate the statistics, appropriate models that have the desired properties should be used.

2D Multiplicative Random Cascades

In order to simulate two dimensional fields with a multi-scaling behaviour, a 2D MRC was used (see figure 5.5(a)). As it the temporal analysis, the lognormal model with branching number $b = 4$ was used. In

order to reduce the free parameters of the model, only non-intermittent fields were simulated (i.e. $\beta = 0$). In figure 5.5(a) a schematic representation of the 2D MRC model is shown.

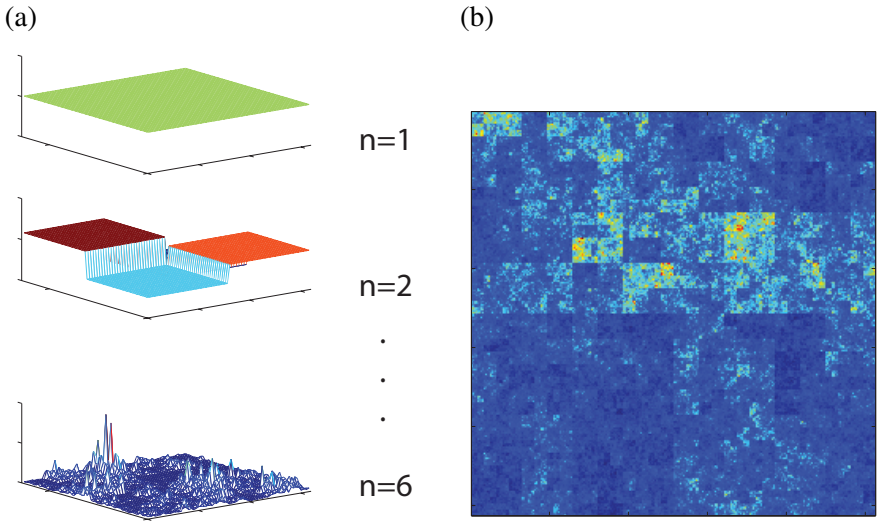


FIGURE 5.5: Schematic representation of the 2D MRC generation.(a) The generation procedure; (b) A random realization with branching number $b = 4$ and $\sigma^2 = 0.1$.

The 2D MRC model that was employed here has some serious restrictions to be used as a valuable precipitation model itself. First of all, the assumption of the description of precipitation as stationary multifractals has been previously challenged [e.g. *Menabde et al.*, 1999]. This the main reason why so-called bounded cascades have been developed [*Menabde et al.*, 1997b]. Another solution to this problem is by assuming that the precipitation fields are arising from a fractionally integrated MRC [*Schertzer and Lovejoy*, 1987]. In that case a MRC model is valid for the fractional derivatives or the increments of the 2D precipitation field [*Nykanen and Harris*, 2003; *Veneziano and Langousis*, 2010]. One other restriction is that the realizations of the discrete version of the spatial MRC model has a rather non-realistic "blocky" appearance (figure 5.5(b)).

In the case for the lognormal model of a MRC without an atom at zero, the only free parameters are the density of the measure on the initial scale (i.e. mean areal intensity or total precipitation depth over a radar image) and the σ^2 parameter of the lognormal distribution of the cascade weights.

Exponentiated 2D FBM

Another common characterization of the precipitation fields is through their radially averaged power spectral densities. A common assumption is that they follow a power law. Such power laws are connected with scale invariant models such as fractional Gaussian noise (FGN) [*Koutsoyiannis*, 2011], fractional Brownian Motion (FBM) [*Heneghan et al.*, 1996; *McGaughey and Aitken*, 2002] or Multiplicative random cascades. As mentioned before, MRC models have an exponent of power-law decay lower than the embedding dimension of the field (2 in the case of radar images).

The model that is used here in order to simulate 2D random fields with an (approximate) power law spectral decay, is an exponentiated fractional Brownian motion. *Veneziano and Langousis* [2010] have shown that such a process do not scale but only on the small scale limit. As shown in figure 5.6 an

approximate assumption of a power law spectral decay is adequate for the purposes of the numerical experiment conducted here.

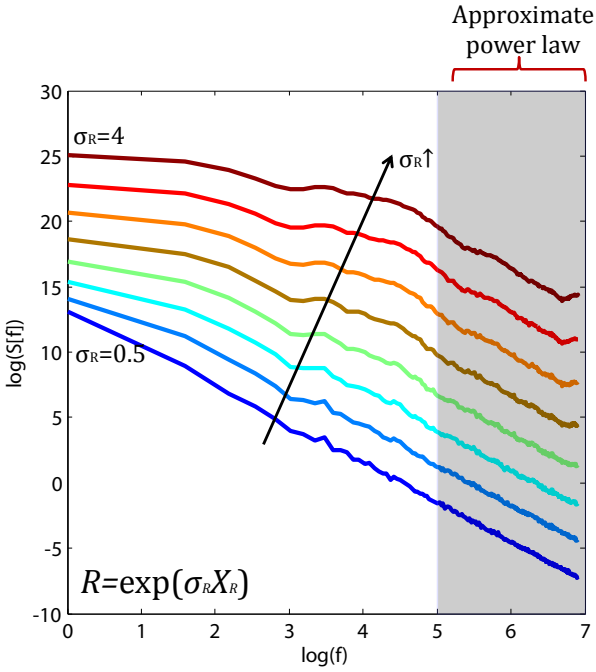


FIGURE 5.6: Power spectral densities of several exponentiated fractional Brownian motions.

The model that is used here for precipitation modelling can be formulated as follows:

$$R(x,y) = \exp\left[\mu_R + \sigma_R X_G(x,y)\right], \tag{5.1}$$

where $X_G(x,y)$ is a two dimensional FBM $\sim N(0,1)$. The simulation of $X_G(x,y)$ is done in the frequency domain using the computational efficiency of the Fast Fourier Transform (FFT) [*Pardo-Iguzquiza and Chica-Olmo*, 1993; *Lang and Potthoff*, 2011].

$$X_G(x,y) = (\mathcal{F}^{-1}|k|^{-\nu/2}\mathcal{F}Z)(x,y). \tag{5.2}$$

Generally speaking the simulation procedure can be summarized as follows:

- Sample a 2D Gaussian white noise field Z
- Obtain its Fourier transform using the FFT
- Multiply the Fourier transform with $|k|^{-\nu/2}$, where $|k| = \sqrt{k_x^2 + k_y^2}$ is the absolute value of the wavenumber.
- Finally $X_G(x,y)$ is its inverse Fourier transform

Detailed information concerning the simulation of random fields in space can be found in section 6.2.5.

Random Fields with Exponential Covariance Function

The two previous models (discrete MRC and exponentiated FBM) were both non-stationary [*Lombardo et al.*, 2012; *Pegram and Clothier*, 2001b] and thus an autocovariance function dependent only on the spatial distance cannot be defined. On the other hand, one of the most common assumptions for spatial rainfall is the stationarity of the field. Several parametric forms for the spatial autocovariance of precipitation have been used in various studies. Among them, the most commonly used, is to assume that precipitation in space is stationary and has an exponential autocorrelation [e.g. *Bell*, 1987; *Kundu and Bell*, 2003; *Sigrist et al.*, 2012a]. More flexible parametric models have also been proposed [*Habib et al.*, 2001; *Ciach and Krajewski*, 2006; *Villarini and Krajewski*, 2009a] for rainfall, but for this study, the most parsimonious exponential autocovariance was used. To test the effects of the various radar image errors, an appropriate model is needed.

In order to take also into account the positively skewed distribution of the precipitation fields the basic assumption was made that precipitation in space can be well described as an exponentiated Gaussian random. This means that precipitation intensities follow a lognormal distribution, which is commonly accepted as a good candidate for spatial fields of precipitation.

Matalas [1967] and *Mejia and Rodriguez-Iturbe* [1974] have shown that such a field

$$R(x,y) = \exp[\sigma_R X_G(x,y) + \mu_R],$$

with $X_G(x,y) \sim N(0,1)$ and $C_X(|r|) = \exp[-|r|/a_g]$ has an autocorrelation function:

$$\rho_R(|r|) = \frac{\exp[\sigma_R^2 \exp(-|r|/a_g)] - 1}{\exp(\sigma_R^2) - 1}. \tag{5.3}$$

More details on this, including the simulation scheme of those fields in the frequency domain will be given in section 6.2.4.

5.2.4 Clutter Contamination

As previously shown, the contamination of the radar images by atmospheric clutter is one of the major measurement induced errors which affects the estimation of spatial statistics for precipitation.To quantify this effect, a Monte Carlo numerical experiment is conducted. Two dimensional random fields with a known structure are drawn from appropriate models and then perturbed by atmospheric clutter. The bias in the estimation of several precipitation structure descriptors is then quantified.

First some reasonable assumptions about the atmospheric clutter contamination should be made in order to construct an appropriate model for it. The main assumption made is that atmospheric clutter can be described as a two dimensional "shot-noise". The probability of atmospheric clutter occurrence is homogeneous in the two dimensional domain. This latter assumption is not entirely true since the probability of atmospheric clutter occurrence is also dependent on the topography (see figure 2.2). For the numerical experiment conducted here it is more important to capture the occurrence rate and intensity of the clutter than its position. For this reason the spatial homogeneity is an adequate assumption. Also the

number of points with atmospheric clutter is dependent on the intensity class. So, the model for clutter contamination on a 2D regular grid is defined as:

$$I(x,y) = \sum_{i=1}^{n_c} \sum_{j=1}^{n(i)} r_c(i)\delta(x - m(j), y - l(i)), \quad x = 1,..,N_x \quad \text{and} \quad y = 1,..,N_y, \tag{5.4}$$

where n_c the number of intensity classes, $n(i)$ the number of clutter occurrences dependent on the intensity class, $r_c(i)$ the intensity value of each class, (N_x, N_y) is the grid size, $[m(j), l(j)]$ is the point of clutter occurrence for the j-th intensity class and

$$\delta(x,y) = \begin{cases} 1 & \text{if } x = 0 \text{ and } y = 0, \\ 0 & \text{if } x \neq 0 \text{ or } y \neq 0. \end{cases} \tag{5.5}$$

Then the perturbed field is defined as

$$R_p(x,y) = R(x,y) + I(x,y), \tag{5.6}$$

where $R(x,y)$ is the initial field. Since the assumption that clutter occurrence is homogeneous in space is adopted, $m(j)$ and $l(j)$ are integer values uniformly distributed in $[1, N_x]$ and $[1, N_y]$ respectively.

The only parameters left to estimate are the number of clutter occurrences per intensity class. The percentage of the area that is affected by clutter contamination is modelled here as a random variable following a lognormal distribution. In figure 5.7 the exceedance probabilities of the clutter percentage of the area around Monte-Lema are shown. The estimates were derived from the radar images when no precipitation occurred.

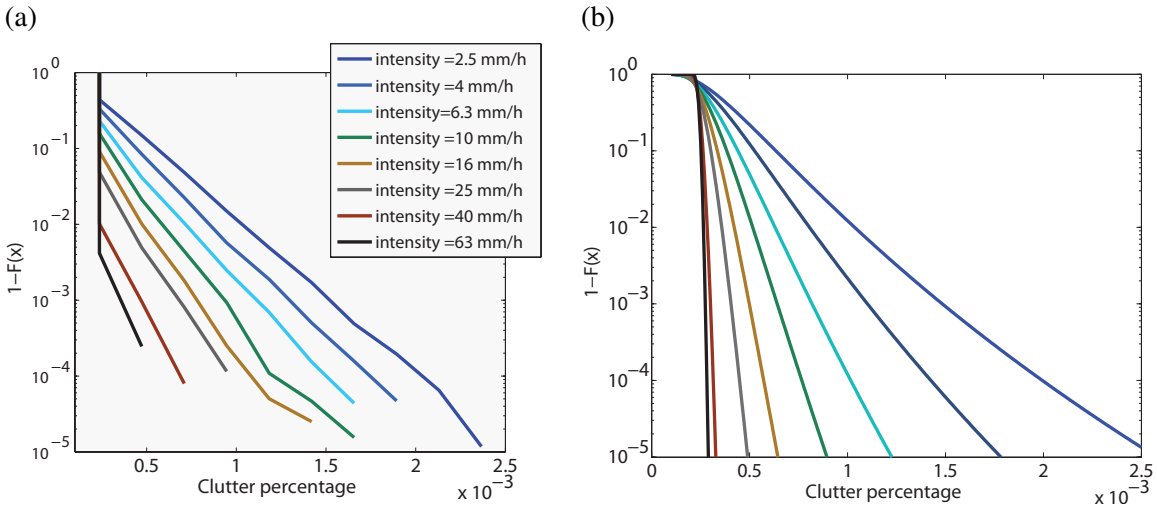


FIGURE 5.7: Exceedance probabilities of the percentages of clutter contamination for different intensity classes. The left panel (a) corresponds to the area located around the Monte-Lema area and the right one (b) to the fitting of lognormal distributions for each intensity class.

The results shown in figure 5.7 correspond to a rectangular area 128×128 km^2 around the radar of Monte-Lema (see figure 2.2). Similar results were obtained also for other regions of the radar domain (e.g. around the Albis radar), and thus only the results from the Monte-Lema region will be used used here.

Clutter Effect on 2D Multiplicative Random Cascade Parameters

The evaluation of the bias introduced by clutter contamination for the parameter of the 2D MRC was performed numerically using a Monte Carlo simulation. The parameter space of the simulation was for the mean areal intensity $[0.05 - 4.5]$ mmh^{-1} and for the σ^2 parameter $[0.05 - 0.3]$. The parameter space was defined from an extensive data analysis of the radar data set (see section 5.3). The sample size of the simulation was 500 for each point of the parameter set. The grid size of the simulated MRC is 256×256 pixels which corresponds to a cascade development of 8 steps with branching number $b = 4$. In order to approximate the dressing factor Z_∞ two more cascade steps were simulated and then aggregated to the desired scale.

The estimation procedure for the parameter σ^2 is the same as described in 3.3.1 and is based on the scaling of the moments of various orders.

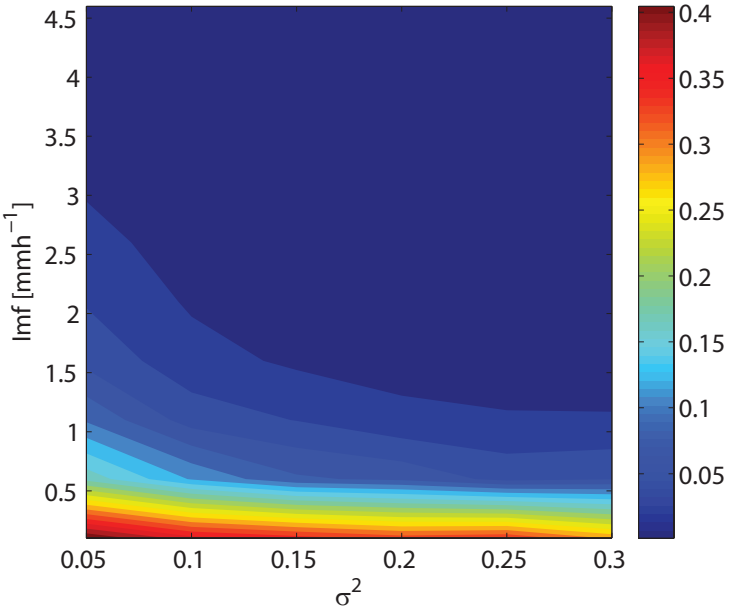


FIGURE 5.8: Bias of the lognormal MRC σ^2 parameter for radar image clutter corruptions.

The results of the bias analysis are shown in figure 5.8. The most striking result is that the bias for the low intensity simulated fields is much higher. The reason is that when low intensity fields occur, the clutter intensities are much stronger than the true field intensities and are dominating the moment scaling estimation procedure, especially for moment orders greater than 1. The bias is stronger as the simulated σ^2 parameter is low. The reason for this is that when σ^2 is low, then the simulated intensities are less positively skewed (i.e. the simulated field has weaker spikes), and the possibility of having simulated intensities that are of comparable magnitude to the clutter intensities is lower. The main result from this numerical experiment is that the MRC estimates are much better for higher intensity fields, an issue that has to be taken into account in data analysis studies. Estimation of the MRC parameters for precipitation fields with mean intensities less that 0.5 mmh^{-1} is almost impossible if a better clutter correction algorithm

is not applied. Since most of the studies are focusing on intense rainfall events, and thus the expected magnitude of the errors in the MRC parameters is small.

Clutter Effect on Power Law Spectral Decay

The biases that are quantified here are the differences between the spectral slopes at the high frequency limit between the simulated and perturbed random fields. The parameter space of the simulation is for the mean areal intensity $[0.05 - 4.5]$ mmh^{-1} and for the spectral slope of $X_G(x, y)$, ν $[-2 - 3]$. The simulations were carried out for two different values of the field's coefficient of variation (0.5 and 3). The estimation of the spectral densities can be done using either the Fourier decomposition of the image [e.g. *Ebtehaj et al.*, 2010] or the wavelet decomposition of the image [*Nicolis et al.*, 2011]. Here both techniques are used. Moreover the analysis is conducted both for the perturbed lognormally distributed fields $R_p(x, y)$ that follow only an approximate power law spectral decay and for their logarithms, which for the case of zero clutter contamination have the same slope as $X_G(x, y)$.

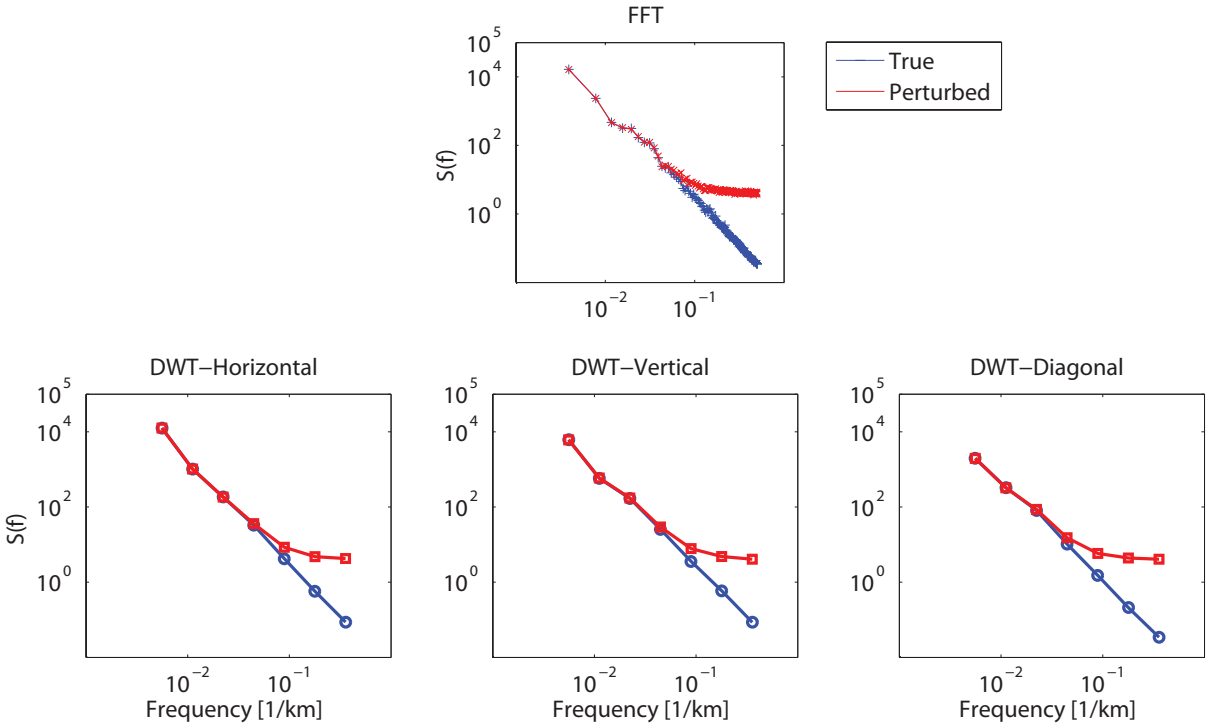


FIGURE 5.9: An example of the effect of clutter on the power spectral estimates. The upper panel shows the spectral estimated based on the FFT and the lower the 3 components of the wavelet based spectral estimates using the DWT. The true field is lognormally distributed with mean value 2 mmh^{-1} and standard deviation 2 mmh^{-1}

In figure 5.9 an illustration of the clutter effect is shown and in figure 5.10 the bias in the spectral slope estimation using the FFT algorithm is presented. The first common feature between all the simulations is that the perturbed fields systematically underestimate the absolute value of the spectral slopes. This can be attributed to the fact that for small spatial scales the clutter contamination "breaks" the correlation structure of the initial fields and this leads to a less steep spectral decay for high frequencies. Moreover, the bias is higher when the mean value and the coefficient of variation of the simulated fields are low, (i.e. lower variance). The theoretical reason for this behaviour is the same as explained for the time series

in 3.3.2. Since in principle the clutter corruption as defined here is an additive white noise, independent of the precipitation field, the power spectrum of the perturbed field will be the sum of the power spectra of $R(x, y)$ and $I(x, y)$. The power spectrum of $R(x, y)$, is an approximate power law for high wave numbers $S_R(|k|) \approx c_1|k|^{-\nu}$ and the power spectral densities of $I(x, y)$ are the same for all wavenumbers $S_I(|k|) = c_2$. Since the constants c_1 and c_2 are directly connected to the variance of the fields $R(x, y)$ and $I(x, y)$ respectively, it is straightforward to show that the effect of $I(x, y)$ is higher when the variance of $R(x, y)$ is low.

It can also be shown that the effect of an additive white noise has a strong effect mainly on the higher frequency scales (see Figure 5.9). The results thus shown in figures 5.10 and 5.11 are highly dependent on the choice of the range of frequencies for which the power-laws for the spectral decay are fitted. Since in most of the studies so far the range where an approximate power-law decay of the spectral density is assumed to hold true on the high frequency regime [e.g. *Pegram and Clothier*, 2001a; *Rebora et al.*, 2006; *Metta et al.*, 2009] the estimation of the bias is done for this range as well here. So no high frequency components of the estimated spectra are neglected. The estimation procedure itself is followed is the most common one in most of the studies, which is an ordinary least square fit at the logarithms of the frequencies against the logarithms of the estimated spectral densities. Another interesting feature is that the analysis conducted on the logarithms of the fields yielded the same bias patterns, but their magnitude is much smaller.

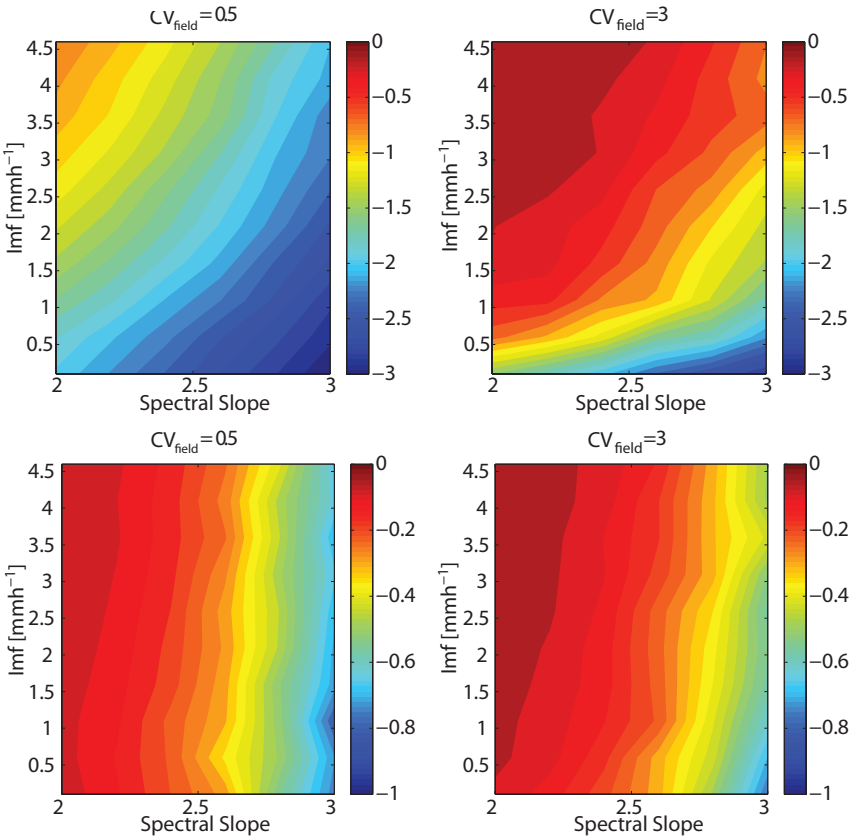


FIGURE 5.10: Bias of the spectral slope decay due to clutter contamination using the FFT based spectral estimates. In the upper two panels the results for the lognormal fields is shown for the two values of coefficient of variation. In the two lower panels the results for their logarithms in shown.

The differences between the estimation of the power spectrum using the Fourier or the wavelet transform are in principle the same as mentioned in the analysis of the time series (section 3.3.2). The estimation of the power spectral densities based on the Fourier decomposition of the image put higher emphasis on the small scale (high frequency) components in comparison to the discrete wavelet image decomposition that is evaluated in integer scales of 2. Due to that fact, the influence of the small scale features (high frequencies) on the least square fitting procedure is much smaller when the wavelet decomposition is used. The effect of that, as shown in figure 5.11 is that the bias of the estimates using the wavelet transform are lower.

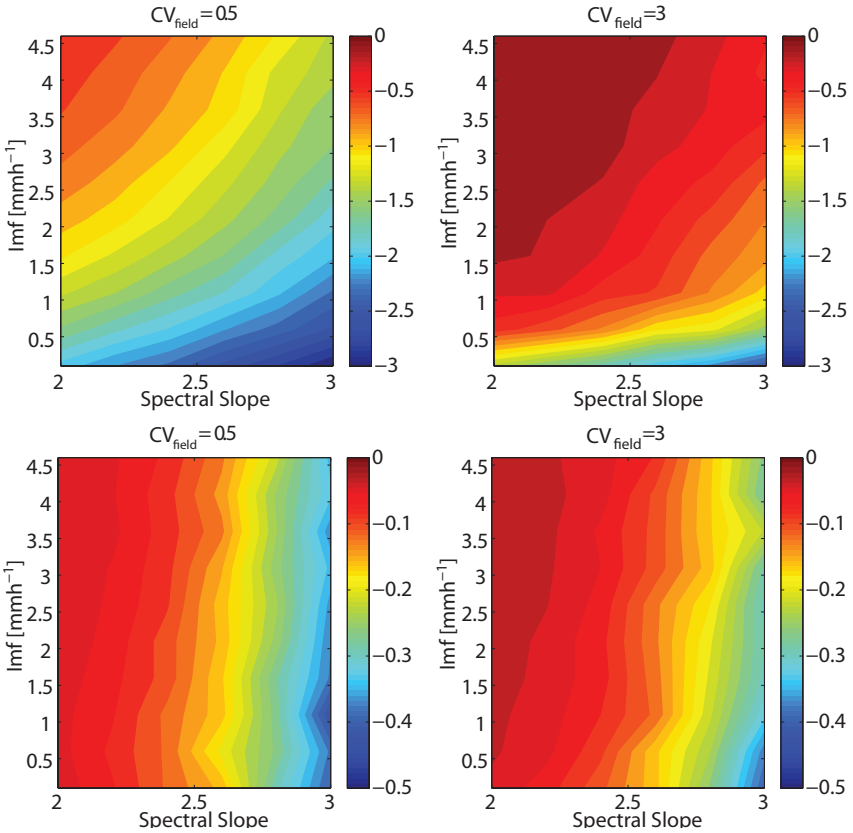


FIGURE 5.11: Bias of the spectral slope decay due to clutter contamination using the wavelet based spectral estimates. In the upper two panels the results for the lognormal fields is shown for the two values of coefficient of variation. In the two lower panels the results for their logarithms in shown.

So as it can be shown in general, the biases on the estimation of the spectral slopes due to clutter contamination of the radar images can be of major importance especially for low variance images. The way to overcome this problem thus could be either a better clutter elimination procedure or to neglect high frequency spectral components. The latter though could yield some more problems connected to data analysis, first due to the fact that a the sample size of high scale components can be low due to restricted data availability and thus the results would not be robust and also if there exists a distinct scaling break, similarly to the time series analysis [*Marani*, 2003; *Fraedrich and Larnder*, 1993], this would not be identified.

Clutter Effect on Correlation Structure

The effect of the clutter contamination on the correct estimation of the autocorrelation function is presented analytically here. The clutter contamination of the radar images is what is commonly referred to geostatistical analysis as the nugget effect [*Webster and Oliver*, 2001; *Kitanidis*, 1997]. The nugget effect describes the small scale variability that can be attributed to an additive white noise. Since as was shown in section 3.3.2, the autocovariance function of the clutter corrupted radar image will be

$$C_{R_p}(|r|) = C_R(|r|) + C_I(|r|), \tag{5.7}$$

where $C_R(|r|)$ is the autocovariance function of the true precipitation field and $C_I(|r|)$ the autocovariance function of the additive noise field. For the case of a white noise representing clutter corruption,

$$C_I(|r|) = \begin{cases} \sigma_I^2 & \text{if } |r| = 0, \\ 0 & \text{if } |r| > 0, \end{cases} \tag{5.8}$$

where σ_I^2 is the variance of the clutter field $I(x, y)$. Then the autocovariance of the perturbed fields will be

$$C_{R_p}(|r|) = \begin{cases} \sigma_I^2 + \sigma_R^2 & \text{if } |r| = 0, \\ C_R|r| & \text{if } |r| > 0. \end{cases} \tag{5.9}$$

It is obvious then that the variance of the clutter field together with the variance of the rainfall field is what controls the possible errors for the estimation of any parametric model for the autocorrelation function. For an illustration of the bias introduced in the estimation procedure if the nugget effect is neglected, the parametric model as described in section 5.2.3 is selected.

In figure 5.12 the bias of the estimation of the a_g parameter, which is the only one governing the correlation structure of the fields is shown. The bias is computed as the error between the theoretical values, and the estimated ones using ordinary least square fitting on the correlation function given in equation 5.9 for the theoretical model (equation 5.3) that does not take into account the nugget effect.

As mentioned earlier, the larger the variance of the precipitation field, the smaller the bias. For very low intensity and variance the underestimation of the correlation parameter a_g can be on the order of 100% of its value (i.e. uncorrelated fields). This is a useful result when data analysis studies are concerned. In this case only intense precipitation fields should be taken into account otherwise serious bias can affect the estimation procedure.

Clutter Effect on Probability Distributions

Finally, one of the most trivial but fundamental statistic is the probability density function of the precipitation intensities. As already mentioned, it is widely accepted that the precipitation intensities in space are approximately lognormally distributed [*Bell*, 1987; *Pegram and Clothier*, 2001a]. The true probability

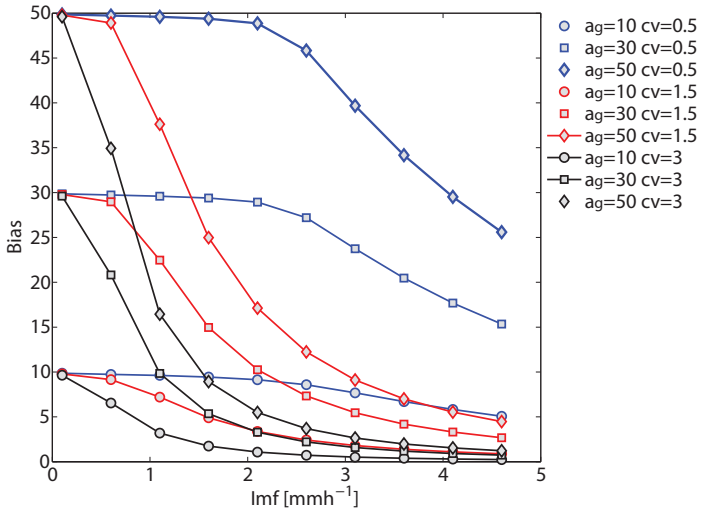


FIGURE 5.12: Bias of the estimated a_g correlation parameter for radar clutter corruption.

distribution function of the clutter-perturbed radar images is defined then as the convolution of the pdf's of the spatial precipitation and the clutter fields, if the assumption of independence holds true,

$$f_{R_p}(x) = \int_{-\infty}^{\infty} f_R(y) f_I(x-y)\, dy = \left(f_R * f_I\right)(x). \tag{5.10}$$

Distribution fitting can be achieved with several methodologies. Here, assuming that spatial precipitation is lognormally distributed and clutter is a process described by equation 5.4, a quantification of the errors that are incorporated in the three most common distribution fitting procedures is conducted. The three methodologies are the method of moments (MOM), the maximum likelihood method (MLE) and the L-moments method (L-MOM) [*Papoulis and Unnikrishna*, 2002; *Hosking*, 1990].

In Figure 5.13 the estimated mean square errors of the distribution fitting are shown for the two parameters of the lognormal distributions respectively for the three methods. The mean square error is defined as

$$MSE = \frac{1}{N}\sum_{i=1}^{N}\left\{F_{ln}^{-1}(U(i),\hat{\mu},\hat{\sigma}) - F_{ln}^{-1}(U(i),\mu,\sigma)\right\}^2, \tag{5.11}$$

where F_{ln}^{-1} is the inverse cdf of the lognormal distribution for non-exceedance probabilities $0 < U(i) < 1$. The values of $U(i)$ that were chosen here cover a large proportion of the entire distribution. They cover the cumulative probabilities $[10^{-5} \sim 1]$ with grid spacing $\delta U = 10^{-5}$. $(\hat{\mu},\hat{\sigma})$ are the estimated distribution parameters and (μ,σ) the theoretical ones.

As it can be seen the maximum likelihood method yields the most reliable results. The second more efficient methodology is based on L-moments and the worst performing method was the method of moments which is very sensitive to outliers. An interesting result is that each of the three estimation methods produces highest errors for different areas of the simulated parameter space. The method of moments is sensitive to errors for low intensity and highly variable fields, the method of L moments is producing the highest errors for low intensity and low variability fields, and the MLE is prone to strong biases for high intensity and high variability fields.

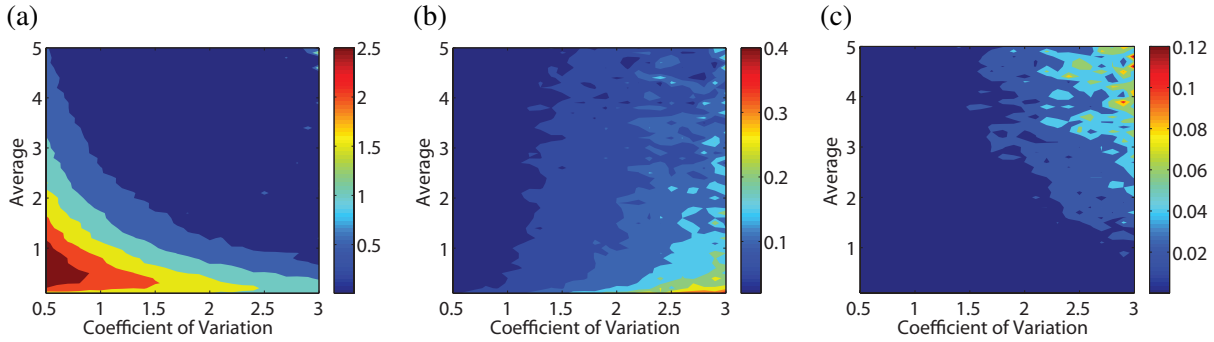


FIGURE 5.13: Mean square error of the inverse of the lognormal distribution function of the estimated and theoretical distributions in the presence of clutter corruption. The three panels from left to right correspond to the estimation based on (a) the method of moments,(b) L-moments and (c) maximum likelihood.

5.2.5 Noise Corruption

One of the most significant sources of error for the rainfall intensity estimation using weather radars is their measurement bias, usually termed as noise corruption. The errors are range dependent and also spatially and temporally correlated [*Berenguer and Zawadzki*, 2008]. In order to model this kind of noise the bias term is considered to be a unit mean multiplicative noise term spatially correlated [e.g. *Germann et al.*, 2009; *Villarini and Krajewski*, 2009b; *Villarini et al.*, 2007]. This approximation neglects the range dependent errors caused by the differences of the sampled volume and the altitude at which rainfall is measured. In the case study examined here, these artefacts are expected to be small since the final radar product is the composite of the recorded reflectivity factors from 3 radars. The general model that described the noise perturbed spatial fields can be formulated as:

$$R_p(x,y) = R(x,y)\epsilon_n(x,y) \quad x,y \in \Re, \tag{5.12}$$

where $R(x,y)$ is the true precipitation field and $\epsilon_n(x,y)$ is a two dimensional random field. The main assumption adopted here is that the two fields $R(x,y)$ and $\epsilon_n(x,y)$ are mutually independent. For the examples that are provided below a parametric form of the noise random field had to be selected. In favour of parsimony, the simplest possible form of the autocorrelation function and marginal distribution is selected. The noise fields are lognormally distributed with a covariance function as in equation 5.3. The noise fields should have a unit mean in order to be mass conservative.

The estimation of the structure of the noise corruption is not easy and requires extensive gauge-radar measurements [*Berenguer and Zawadzki*, 2008]. For the case of Switzerland a quantification of those errors can be found in the study of *Germann et al.* [2009] where the authors also quantified the importance of the error propagation to flood estimation, highlighting the caveats of the use of radar-estimated precipitation for operational hydrology.

Noise Corruption Effect on 2D Multiplicative Random Cascade Parameters

The first study to investigate the effects of correlated multiplicative noise on the estimation of the scale invariant structure of precipitation was conducted by *Villarini et al.* [2007] and subsequent studies [*Mandapaka et al.*, 2010; *Villarini et al.*, 2009]. The results of those studies where highlighting the general

effect of noise corruption, but they were restrictive in the sense that their analysis was case-dependent and a small number of rainfall events was analysed. Here a generalization of those results is made by a Monte Carlo simulation similar to the one conducted for the case of clutter contamination (see figure 5.4).

The parameters that influence the results are the "magnitude" of the noise field expressed by its variance as a percentage of the variance of the true precipitation field, its correlation structure and the statistics of the initial "true" precipitation field. Also the estimation procedure will influence the results. The estimation procedure that was used here is based on the moment scaling. It is expected that estimation based on wavelet decomposition of the image or the use of different variants of the moment scaling procedure [*Veneziano and Furcolo*, 2009] would result in better estimates, but since the classic moment scaling procedure is the most commonly used, it is the one analysed here.

In figure 5.14 the mean values if the introduced bias on the estimation of σ^2 parameter are shown.

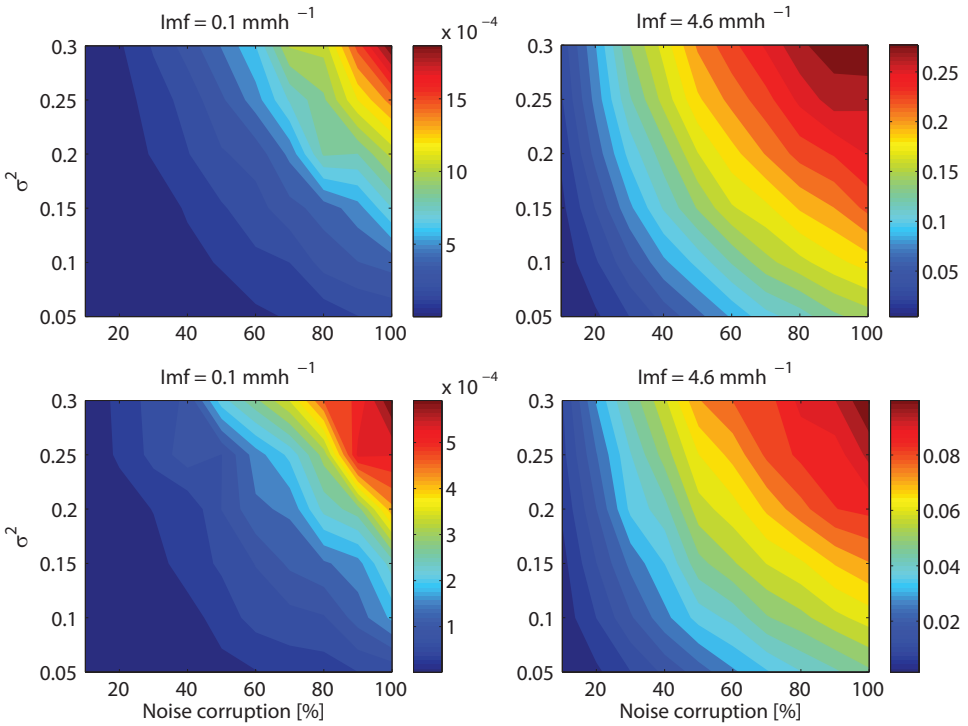


FIGURE 5.14: Bias in the estimation of the σ^2 parameter of the 2D lognormal MRC in the presence of correlated lognormally distributed noise. The example shown here correspond to noise with correlation length 5 cells (upper 2 panels) and 50 cells (lower 2 panels).

Due to the relatively small sample of realizations for the parameter set, the results are rather noisy, but the overall picture is clear. The first result is that for all the cases the bias is always positive indicating that there is a systematic underestimation of the σ^2 parameter regardless of the strength and the correlation structure of the noise. The other result is that the errors enhance for high values of the precipitation field and for low noise correlation. This result can be of major importance since typically in data analysis studies based on radar data [e.g. *Nykanen*, 2008], the high intensity fields are examined, and if the possible effect of the multiplicative noise is not taken into account this can result in high bias. Nevertheless very large errors only occur for noise fields with very high variance which are not expected in well calibrated radars.

Noise Corruption Effects on Power Law Spectral Decay

The other commonly estimated statistical descriptor the spatial rainfall is its radially averaged power spectrum, adopting the assumption of isotropy for the rainfall fields. *Mandapaka et al.* [2010] identified that multiplicative noise corruption of radar images leads to strong discrepancies of the spectrum estimated at high frequencies. Here the bias on the estimation of the spectral slopes is quantified numerically for both Fourier and wavelet spectra. The results are shown in figures 5.16, 5.17. The exponent of the power law spectral decay was estimated with ordinary least square fitting for all the frequencies.

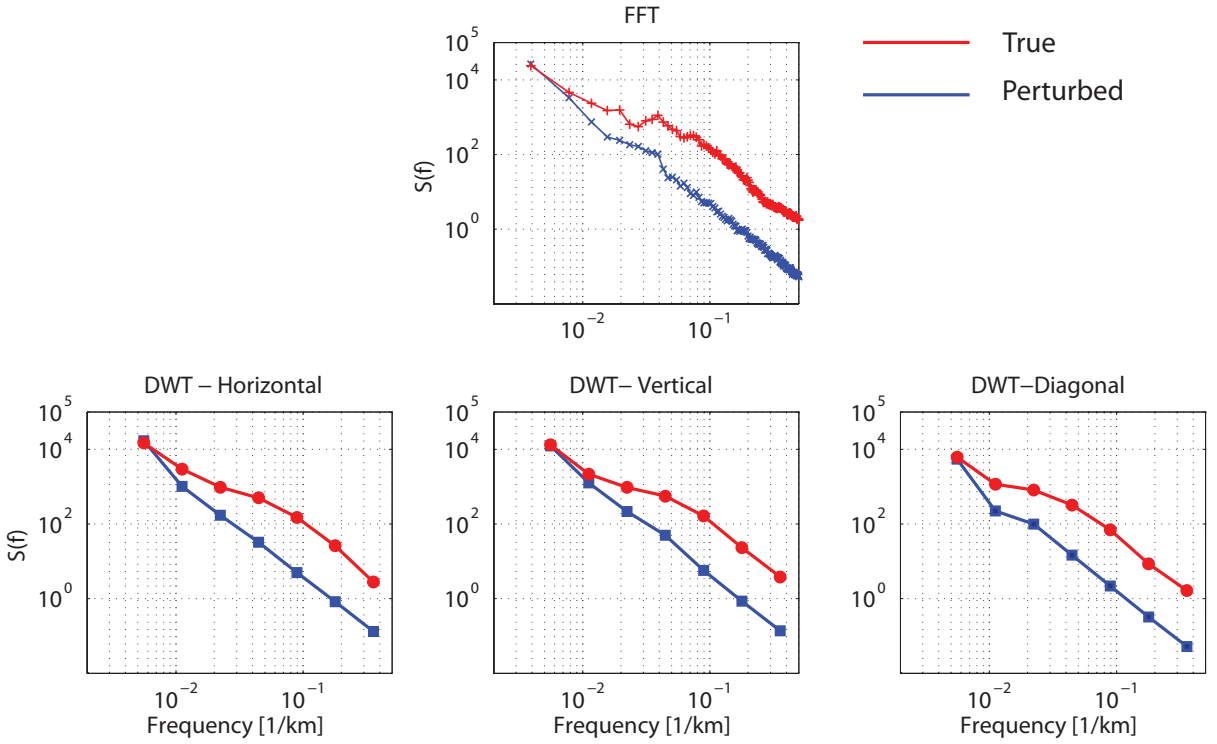


FIGURE 5.15: Example of the effect of multiplicative noise corruption on the spectral estimates. The rainfall field is lognormally distributed with mean value 2 mmh^{-1} and standard deviation 2 mmh^{-1}. The standard deviation of the multiplicative noise is 20% of the rainfall field and has a correlation length of 5 km.

In figure 5.15 an example of the estimated power spectral densities according to the two methodologies is shown. The first result is that due to the fact that the variance of the perturbed fields is higher than the variance of the original fields, the spectral densities are much higher. This is explained by the fact that variance is the integral of the power spectral density for all the frequencies. The perturbation affects the whole range of spectral densities and not only the high frequency components. The reason for this is the autocorrelation of the noise field. The spectral slopes for only the high frequency components remain almost the same.

The results of the bias estimates from the numerical experiment shown in figure 5.16 indicate that the bias is dependent on all of the free parameters included in the experiment. Due to the high number of free parameters for this numerical experiment (mean value, coefficient of variation, spectral slope, level of noise corruption, autocorrelation of the noise field), only the significant cases are presented.

A comparison between figures 5.16 and 5.17 shows that the bias pattern is the same between the two estimation procedures. The magnitude however is different. The estimates based on the wavelet spectra

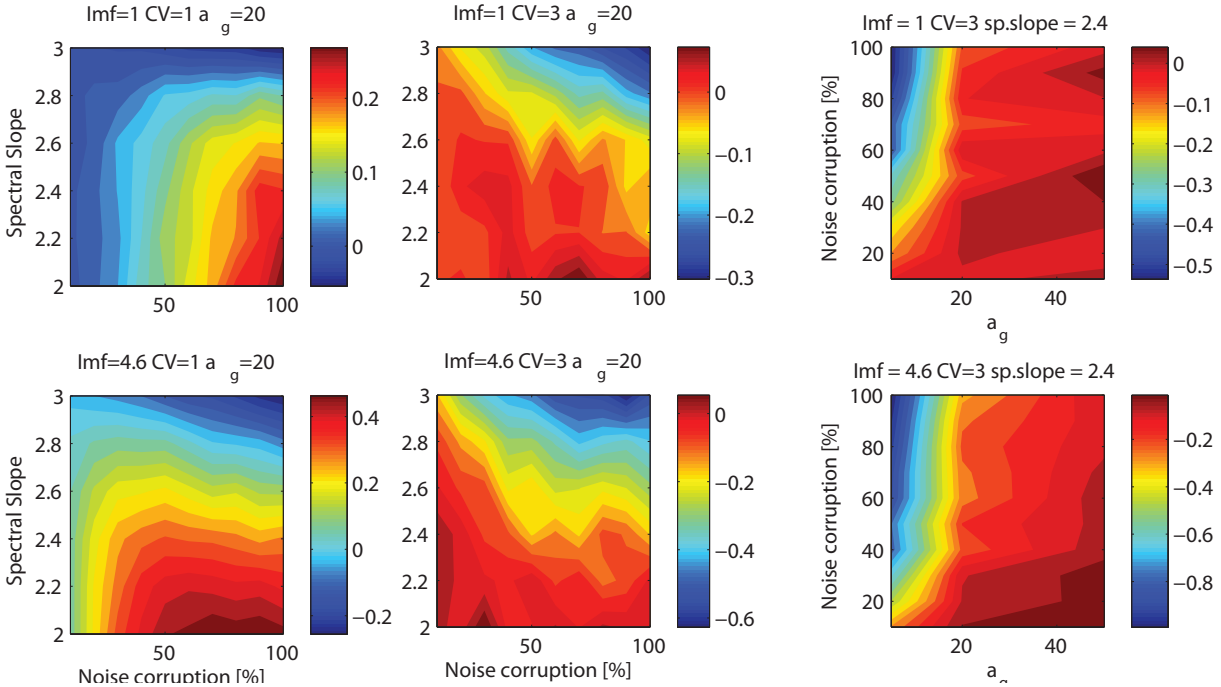


FIGURE 5.16: Estimated bias of the power law spectral decay due to multiplicative correlated noise corruption according to the FFT based estimated spectra. The upper panel corresponds to the case of mean areal intensity of 1 mmh^{-1} and the lower one of 4.6 mmh^{-1}. The first two columns illustrate the bias as a function of the true field spectral slope and the level of noise corruption. The third row shows the bias as a function of the correlation length of the noise term and the level of noise corruption.

are systematically more negatively biased than the FFT based estimates.

The bias is generally higher when the intensities are high. Obviously, the bias is also enhanced when the level of noise corruption is higher. Moreover the bias is stronger when the spectral slope of the original field is higher (i.e. more spatially correlated patterns.). Finally, the bias is substantially higher when the noise field is less correlated in space. This happens due to the fact that the lower the correlation, the higher is the image distortion in terms of spatial structure. It should be noted that the errors are substantial and can reach up to 20% of the simulated value of the spectral slopes.

Finally a comparison between the figures 5.15 and 5.9 shows that different sources of measurement errors result in discrepancies at different frequency ranges. This means that in the case of the simultaneous presence of clutter and a multiplicative error term, there will be no frequency range without problems. This excludes the possibility of choosing a restricted scale range for estimating the spectral slopes. It should also be noted that the presence of multiplicative noise is also difficult to identify and eliminate from the radar records, in contrast with the clutter corruption, which is more straightforward to correct. One solution to overcome this problem could be a real-time combination of the information of the precipitation amount from the rain-gauge network, and the spatial information from the radar derived estimates. Recent efforts have been aiming at this direction in the MeteoSwiss CombiPrecip project [1].

[1] http://meteoswiss.ch/web/en/research/current_projects/climate/nccr_iii/combiprecip/sii.html

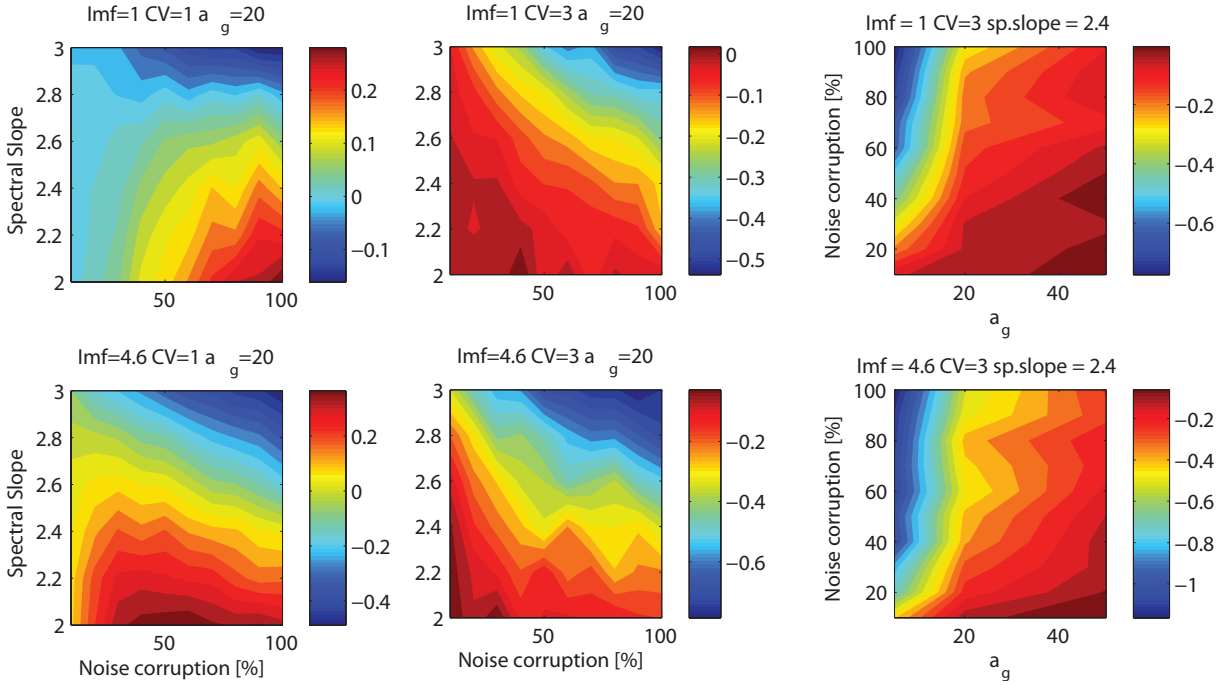


FIGURE 5.17: Estimated bias of the power law spectral decay due to multiplicative correlated noise corruption according to the wavelet based estimated spectra. The upper panel corresponds to the case of mean areal intensity of 1 mmh^{-1} and the lower one of 4.6 mmh^{-1}. The first two columns illustrate the bias as a function of the true field spectral slope and the level of noise corruption. The third row shows the bias as a function of the correlation length of the noise term and the level of noise corruption.

Noise Effects on Correlation Structure

If it is assumed that both the precipitation and noise fields are stationary in space and isotropic then the covariance of the product field will be

$$C_{R_p}(|h|) = C_{R_p}(|h|)C_{\epsilon_R}(|h|) + C_{R_p}(|h|)E[\epsilon_R]^2 + C_{\epsilon_R}(|h|)E[R_p]^2, \tag{5.13}$$

where $|h|$ is the Euclidean distance. This means that the autocorrelation functions, and the variances and mean values of the two fields will affect the final autocorrelation. *Mandapaka et al.* [2010] analysed these effects on a storm specific study. The analytical solution is shown in equation 5.13. For an illustration, figure 5.18 shows the bias that would be introduced in the estimation of the correlation length if the estimation would be conducted using ordinary least square fitting on the perturbed fields and their logarithms. The main hypothesis is that the true precipitation fields have a covariance function as in equation 5.3 and the noise fields a covariance function exponentially decaying.

The autocorrelation function of the logarithm of a lognormally distributed field is [*Mejia and Rodriguez-Iturbe*, 1974],

$$\rho_g(r) = \frac{\log(\rho_R(r)[\exp(\sigma_R^2) - 1] + 1)}{\sigma_R^2}. \tag{5.14}$$

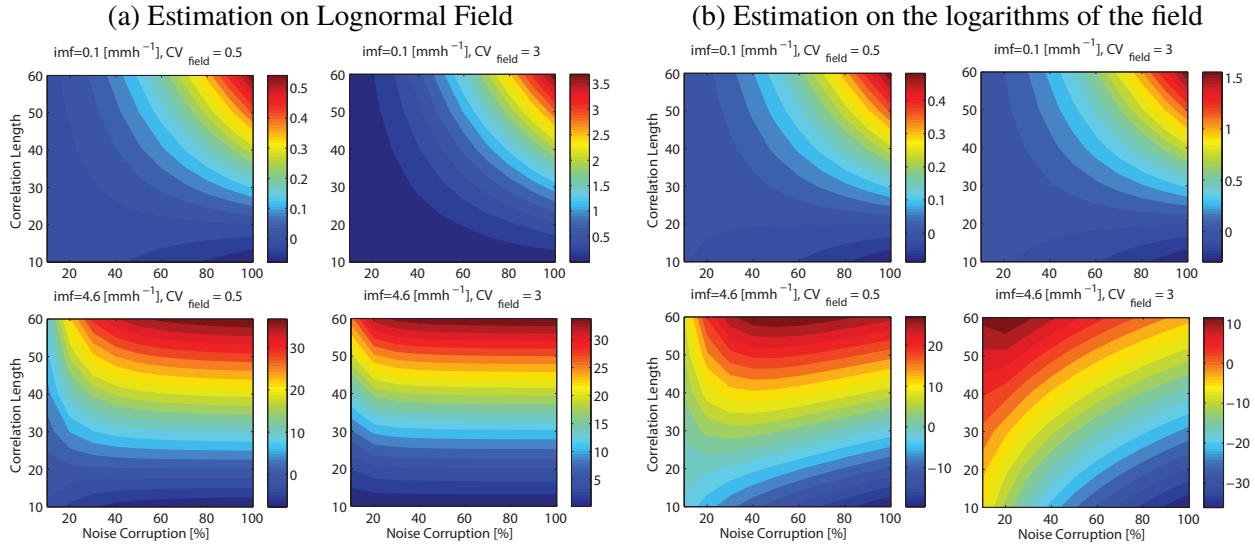


FIGURE 5.18: Bias of the estimated correlation length due to multiplicative noise for different combinations if mean intensity and coefficient of variation. The left panels (a) correspond to the estimation of the correlation of the lognormally distributed fields and the right panels (b) on their logarithms. In all the cases the correlation length of the noise term is 20 cells.

As a first result, it can be shown that in the case of the estimation of the model's correlation on the lognormally distributed fields, the bias is almost always positive, for the cases investigated here, which means that the correlation parameter is always underestimated. The same is not valid when the estimation of the correlation parameter is done on the logarithms of the field. In this case, also strong overestimation cases can be found. The patterns of the bias are the same and indicate that the largest values occur obviously as the noise corruption level increases and also when the unperturbed "true" precipitation field is highly correlated itself. Moreover, the biases are also affected by the coefficient of variation of the field, showing that the more variable the fields, the higher the bias is in absolute terms.

Noise Corruption Effect on Probability Distributions

The marginal probability density function of the perturbed fields will be different as well. In the case of multiplicative noise corruption, if the assumption of indecency between precipitation and noise holds true, the pdf of the noise corrupted field will be [*Springer and Thompson*, 1966],

$$f_{R_p(z)} = \int_{-\infty}^{\infty} \frac{1}{|x|} f_{R_p,\epsilon_R}\left(x, \frac{z}{x}\right) dx = \int_{-\infty}^{\infty} \frac{1}{|x|} f_{R_p}(x) f_{\epsilon_R}\left(\frac{z}{x}\right) dx. \tag{5.15}$$

Specifically, if both precipitation and noise are assumed to be lognormally distributed with parameter (μ_R, σ_R) and $(\mu_\epsilon, \sigma_\epsilon)$ respectively, then the perturbed fields will have as well a lognormal pdf with parameters $(\mu_R + \mu_\epsilon, \sqrt{\sigma_R^2 + \sigma_\epsilon^2})$. In this case there is no point comparing the various fitting methods as was done for the case of clutter contamination since all the methods would yield the same estimation parameters (at least for large enough samples) since the product distribution is as well lognormal.

5.2.6 Class Binning

As it was shown for the time series analysis, data quantization can potentially have a major effect on the estimation of various precipitation structure estimators. Here the same issue concerning the radar derived precipitation estimates in the fonm of binning into intensity classes is examined.

Most of the operational weather radar products are typically stored classified into intensity classes. Due to the power-law equation between the reflectivity factors [dBZ] that the radars record and precipitation intensity [mmh^{-1}], those classes are usually logarithmically distributed. In the case of the RAIN product used in this study, precipitation rain rates are binned into 16 classes [*Savina*, 2011]. The fact that the classes are not uniform increases the measurement uncertainty for high intensities. This can be of major importance in hydrological studies, since those large uncertainties propagate into the modelling of all the other hydrological processes, such as runoff generation.

In order to quantify the effects of this source of error on the various statistical estimators a Monte Carlo numerical experiment was conducted like in the previous sections. The approach to bin the various realizations from the appropriate models to discrete values. The values that were chosen here are the intensity classes of the operational RAIN radar product of MeteoSwiss. A more general bias quantification would investigate both the number of classes as well as their distribution (equally or non-equally spaced).

Class Binning Effect on 2D Multiplicative Random Cascade parameters

First, the quantification of the bias introduced in the single parameter of the non-intermittent lognormal cascade is shown in figure 5.19.

It should be noted that the quantization into classes always leads to an underestimation of the σ^2 value. The underestimation can reach up to 30% of the simulated σ^2 parameter. The bias is dependent both on the mean precipitation intensity of the field and the simulated σ^2 parameter. The bias is higher when the σ^2 parameter of the true field is higher and the mean areal intensity is high. The reason why this behaviour occurs is that high values of σ^2 and mean intensity lead to high values of pixel scale intensities, that are not well represented due to the signal quantization. For example all the intensities between 63-100 mmh^{-1} belong to the same intensity class, and since in the moment scaling analysis mainly the high intensity values affect the estimation of the high order moments $q > 1$, it becomes obvious that in the presence of high values, the estimation problems are enhanced.

Since the probability distribution of the simulated fields is dependent on the σ^2 parameter, with higher values of σ^2 leading to more heavily tailed distributions [e.g. *Langousis et al.*, 2009; *Lovejoy and Schertzer*, 1985; *Tessier et al.*, 1993; *Veneziano and Furcolo*, 2003; *Bernardara et al.*, 2008], a substantial underestimation of σ^2 can lead to serious errors in precipitation extremes.

Class Binning Effect on Power Law Spectral Decay

The numerical quantification of the estimation of the spectral slopes was conducted in the same way as described before.

A first result as shown in the example illustrated in figure 5.20 is that the signal quantization only affects the high frequency components. The spectral densities of the higher scales are almost identical to the original ones.

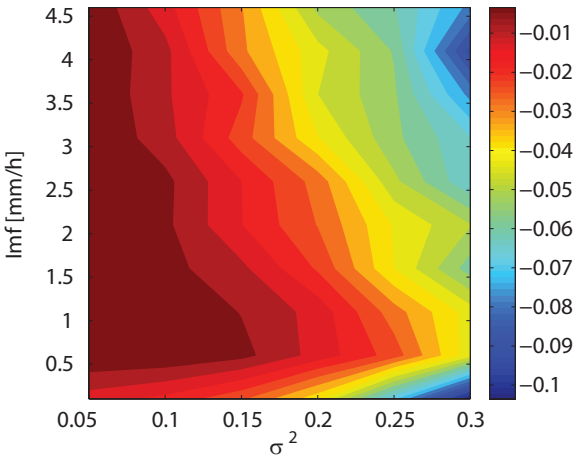


FIGURE 5.19: Bias of the lognormal MRC σ^2 parameter for radar image data quantization.

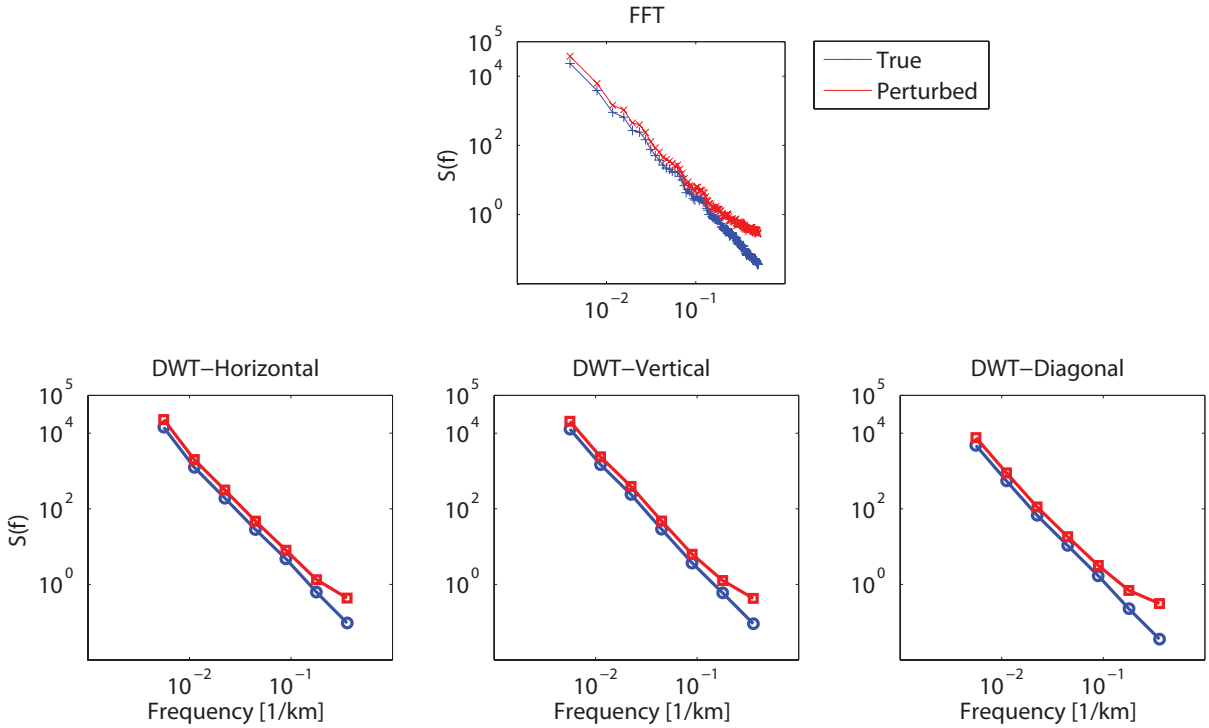


FIGURE 5.20: Example of the effect of the intensity class binning on the spectral estimates. The rainfall field is lognormally distributed with mean value 2 mmh^{-1} and standard deviation 2 mmh^{-1}. The standard deviation of the multiplicative noise is 20% of the rainfall field and has a correlation length of 5 km.

For all the cases, the errors that are introduced are mainly dependent on the initial spectral slope and the marginal distribution of the rainfall intensities, rather than the mean areal intensity. The slopes are always underestimated (negative bias), and the bias is stronger when the original "true" precipitation fields are highly correlated, i.e. have a steeper spectral decay and are less variable. Moreover, the biases are generally smaller for high intensity fields with higher spatial variation.

In addition since the discrepancies that are introduced mainly affect the small scale variations, which

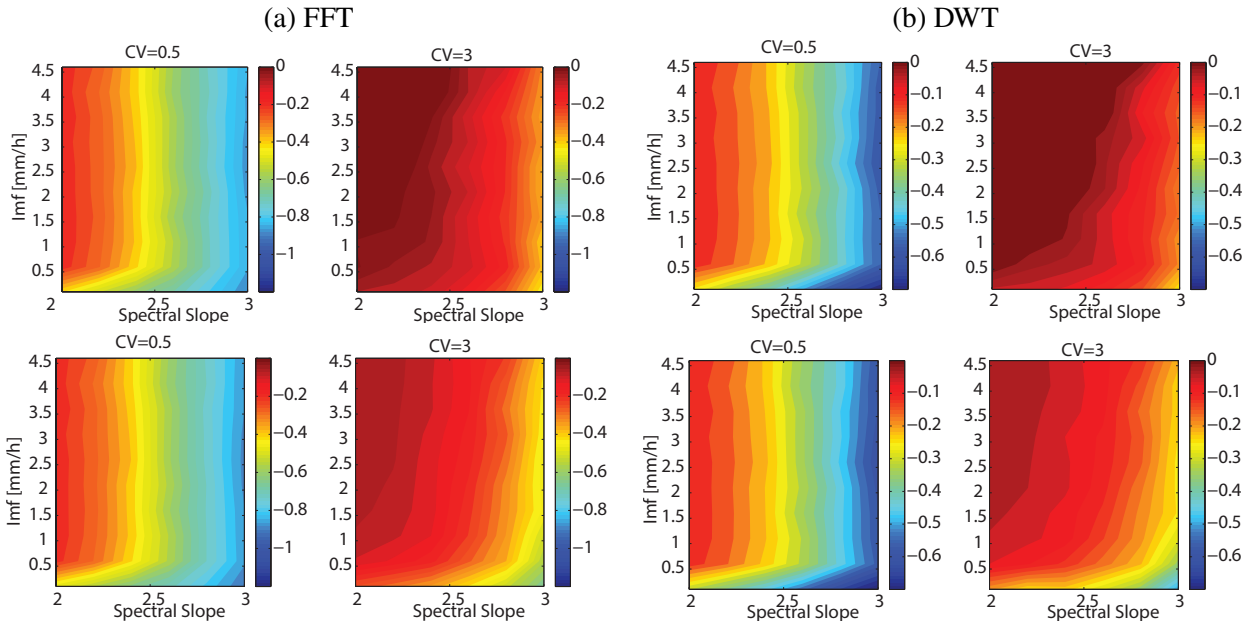


FIGURE 5.21: Bias of the spectral slope decay due to intensity quantization using the FFT and DWT based spectral estimates. In the upper two panels the results for the lognormal fields is shown for the two values of coefficient of variation. In the two lower panels the results for their logarithms in shown.

translates to deviations for the high frequency spectral components, the estimation based on the wavelet decomposition is more reliable, since it is less dependent on the high frequency components as was already mentioned (see Figure 5.20).

One encouraging result is that for the most intense precipitation fields (>3mmh^{-1}), that also typically have a high coefficient of variation (~ 3), the estimation of the power law spectral decays are almost unbiased. This means that stochastic models that are calibrated on this statistic [e.g. *Pegram and Clothier*, 2001b] may not seriously suffer from this kind of measurement error if an appropriate selection criterion for the radar images is selected. The estimation of the power spectrum can be highly problematic for low intensity fields with low spatial variation (e.g. winter stratiform precipitation). In this case the higher frequencies of the power spectral estimates should be discarded from the analysis.

Class Binning Effect on Correlation Structure

As far as the correlation structure of the precipitation fields is concerned, the estimation of the errors was done numerically for the same models as described in 5.2.4. Since an analytical expression of the correlation structure of the perturbed fields was not available the quantification is done using a Monte Carlo simulation experiment. The radially averaged autocorrelation of the perturbed field was estimated and the theoretical models using ordinary least square fitting. The numerical experiments were conducted for both lognormally distributed fields having an autocorrelation function as in equation 5.3 and to their logarithms that have an autocorrelation with exponential decay.

The results are shown in figure 5.22. An interesting result is that the correlation length can be both over and underestimated. Overestimation only occurs when the spatial coefficient of variation is high and the estimation procedure is conducted on the lognormally distributed field. In all the other cases the

correlation parameter a_g is underestimated.

For all the cases examined here, the errors are higher for low intensity precipitation values. In general, when the estimation of the correlation parameter is performed on the logarithms of the field, it is more reliable. For high intensity fields which are mainly of interest in hydrological studies, the results are almost unbiased. The effect of the spatial variability of the field is not entirely clear. When the estimation conducted on for the logarithms of the field, highly variable fields yield smaller errors.

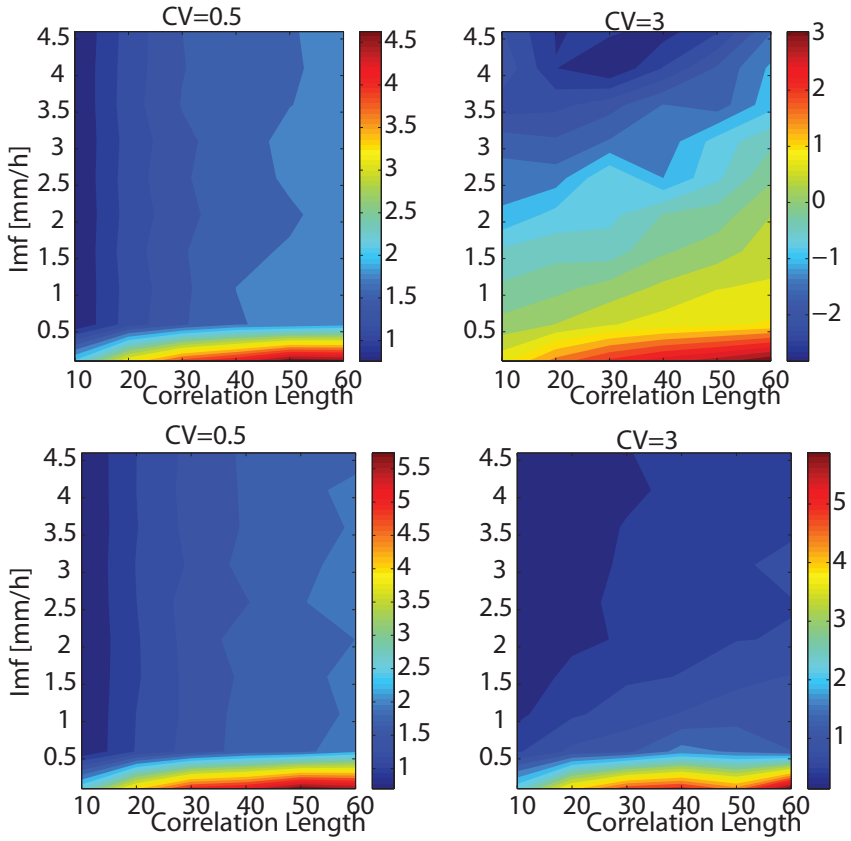


FIGURE 5.22: Bias of the estimated a_g correlation parameter for precipitation intensity quantization. In the upper two panels the results for the lognormal fields is shown for two values of the coefficient of variation. In the two lower panels the results for their logarithms in shown.

Also here an encouraging result as far as data analysis is concerned is that for the intense fields, that provide the maximum information about the precipitation fields, the introduced errors can be considered as negligible. Thus models that depend on the estimation of the autocorrelation function [e.g. *Bell*, 1987] can be reasonably well calibrated.

Class Binning Effect on Probability Distributions

Finally, the impact of the data quantization on the distribution fitting is presented . For the same reasons as explained in section 5.2.4 the lognormal distribution was chosen as an appropriate candidate for precipitation intensities. The three methodologies for distribution fitting are identical as in 5.2.4.

For the maximum likelihood method, for the case of grouped data, the log-likelihood function to be maximized is different from the case of continuous data. In this case it is defined as [*Giesbrecht and*

Kempthorne, 1976; *Pegram and Clothier*, 2001a]:

$$\mathcal{L} = n_0 N\left(\frac{y_1-\mu}{\sigma}\right) + \sum_{i=1}^{M-1} n_i \log\left\{N\left(\frac{y_{i+1}-\mu}{\sigma}\right) - N\left(\frac{y_i-\mu}{\sigma}\right)\right\}, \tag{5.16}$$

where $i = 0, 1, 2, ..., M$ are indicators of the intensity classes, n_i is the number of elements of the i-th class, y_i is the logarithm of the intensity of the class, N is the cdf of the standard normal distribution,and μ, σ are the parameters of the lognormal distribution. Obviously since this methodology is designed for grouped data, the estimation of the distribution parameters is exact (see figure 5.23(c)) .

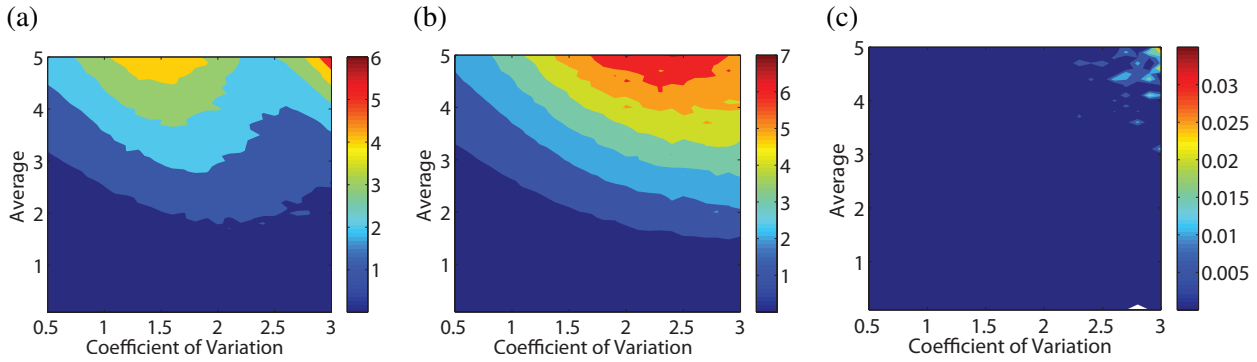


FIGURE 5.23: Mean square error of the inverse of the lognormal distribution function of the estimated and theoretical distributions for quantized data. The three panels from left to right correspond to the estimation based on the (a) method of moments,(b) L-moments and (c) maximum likelihood .

The comparison of the efficiency of the two other methods may also be found in figure 5.23.In terms of mean square error defined as in equation 5.11, the method of moments and the method of L-moments are comparable. Surprisingly though, the method of moments outperforms the one of L-moments. So as a general result, it can be concluded that in the case of grouped data, the MLE provide the most reliable results and is the recommended method. The non-exceedance probabilities that are used here for the estimation of the MSE are the same as in section 5.2.4

5.2.7 Summary of the effect of Measurement Errors on the Precipitation Structure Estimation

In table 5.1 an overview of result previously reported for all the measurement errors are shown. A crucial question is how can we deal with the various errors in order to proceed to a statistically sound data analysis of the available precipitation data.

A general result is that the highest discrepancies occur for low intensity fields. For this reason all the radar images reporting low average intensities sound be excluded for any kind of analysis. This results to a significant reduction of the sample size of the statistical analysis. However, low intensity events are of minor importance in hydrology and excluding them from the data analysis does not impose serious problems.

A second result is that atmospheric clutter can yield the most serious errors in terms of magnitude and has to be eliminated. Fortunately identification of the clutter is straightforward and simple image processing techniques as will be later described can resolve this issue.

A third result is that quantization of rainfall intensities into classes does not result into very significant errors (e.g. in comparison to the clutter). In the case of class binning the errors are mainly associated with low intensity rainfall images. However if low intensity radar images are eliminated from the analysis the overall impact of such an error are expected to be negligible. This is very important because a retrieval of the true intensity from the quantized ones is impossible.

Finally, the issue of the multiplicative noise can yield significant errors. Fortunately the radar network of MeteoSwiss is very well maintained and such errors are expected to be negligible. Moreover the algorithmic corrections implemented in 2004 for the estimation of precipitation [*Germann et al.*, 2006] had very good results. For this reason only radar data beyond 2004 will be used here. In the recent future when the new operational data product obtained through the CombiPrecip Project (Combination of radar and gauge measurements) with very high accuracy will be available, a validation of the results of the data analysis reported later in this chapter can be validated.

TABLE 5.1: Summary of the most significant errors on the statistical estimators for spatial rainfall due to measurement artefacts.

	Clutter	Multiplicative noise	Class Binning
MRC parameters	Significant Errors for Low Intensity fields	Larger biases for high intenity fields and low correlated noise	a) Not significant errors in general b) Larger errors for low intensity fields
Spectral Slope	a) Significant discrepancies for high frequencies b) Larger errors for low intensity fields c) Wavelet estimation less vulnerable	Larger biases for low correlated noise. Equivalent efficiency of Fourier and wavelet estimation	a)Larger bias for low intensities and steeper spectral decays b) Wavelet estimation less sensitive
Autocorrelation Function	Large underestimation for low intensities andlow spatial variation	a) Large underestimation for high intensities for estimation at the lognormally distributed fields b) Both under and overestimation for the correlation at the logarithms of the field	Significant errors only for low intensities
Probability Distribution	Maximum Likelihood estimates are least sensitive	Large errors for high intensity, higly variable fields	Significant errors both for the Method of moments and the method of L moments

5.3 Radar Data Analysis

Having analysed the impacts of the various sources of errors on the estimation of several indicators of spatial-structure indicators, an extensive analysis of the radar data derived as a composite of the 3 C-band weather radars operated by MeteoSwiss is conducted in this section.

Two areas of a spatial extend of 64×64 km^2 are analysed. The first one is the area located around the "Albis" radar close to Zurich, in the NE part of Switzerland (Figure 2.2). The area extend also into parts of SE Germany. The second one is located around the "Monte-Lema" area in Tessin located in the SE part of Switzerland. This area extends into the Northern part of Italy (Lombardia and Piemonte). Those two areas were selected because the radar efficiency there is considered as optimal, with negligible effects of beam shielding and ground clutter. The NE part of Switzerland, located around the third "La-Dôle" radar, suffers from serious beam blockage due to the Jura mountain range (see figure 2.2).

The two selected areas have very different climatological and orographic characteristics. Their main climatological features are determined by the various orographic effects of the Alpine mountain range. The detailed description can be found in section 2.1. Due to the relatively small spatial extent of those areas, they can be considered as climatologically uniform, with local orographic effects.

In order to take into account the strong seasonal differences that occur in both areas, the analysis is conducted on a seasonal basis.

5.3.1 Data pre-Processing

First of all, since the radar records are binned into 16 intensity classes, the simplified assumption that each class is represented by its central value is adopted. Without any knowledge about the true distribution of the precipitation intensities per intensity class, this assumption is adequate.

As shown in section5.2.4, the remaining ground clutter, due to anomalous propagation, that is present in the radar data set, can cause serious errors on the estimation of most of the spatial structure descriptors. This effect is also the strongest source of errors associated with radar measured precipitation. Here, a simple and efficient algorithm for eliminating those points is presented.

The points where ground clutter occurs report very high intensities localized in space. Subsequently, the spatial intensity gradient field is very high in those points. Here the selected measure for the identification of spurious points in the precipitation fields $R(\mathbf{x})$ is the absolute value of the discrete Laplacian defined as:

$$L = \left| \frac{\nabla^2 R}{4} \right| = \left| \frac{1}{4} \left(\frac{d^2 R}{dx^2} + \frac{d^2 R}{dy^2} \right) \right| \tag{5.17}$$

and approximated on a regular grid with the five point stencil finite difference:

$$L(i,j) \approx \left| \frac{R(i-1,j) + R(i+1,j) + R(i,j-1) + R(i,j+1) - 4R(i,j)}{h^2} \right|, \tag{5.18}$$

where h stands for the grid spacing. The points that are selected as spurious, were the ones where $L(i,j)$ exceeded the 99.5% quantile of the entire sample of $L(i,j)$. This corresponds to 20 spurious points on a 64×64 square grid. The selected threshold was selected subjectively since it yielded reasonable results. This procedure can however only eliminate strong ground clutter and not low intensity noise. This is not a serious problem, since as was shown in 5.2.5 low intensity noise has a minor effect on the analysed statistics.

The values that were identified as spurious were substituted with simple linear interpolation of their neighbouring points (see figure 5.24). Since there is the possibility of elimination of true intensities, the estimation of the magnitude of the errors associated with the linear interpolation is also needed.

Let $\hat{R}(x_i)$ be the estimate of the unknown intensity of a pixel on the radar domain defined as

$$\hat{R}(x_i) = \frac{1}{n} \sum_{k=1}^{n} R(x_k), \tag{5.19}$$

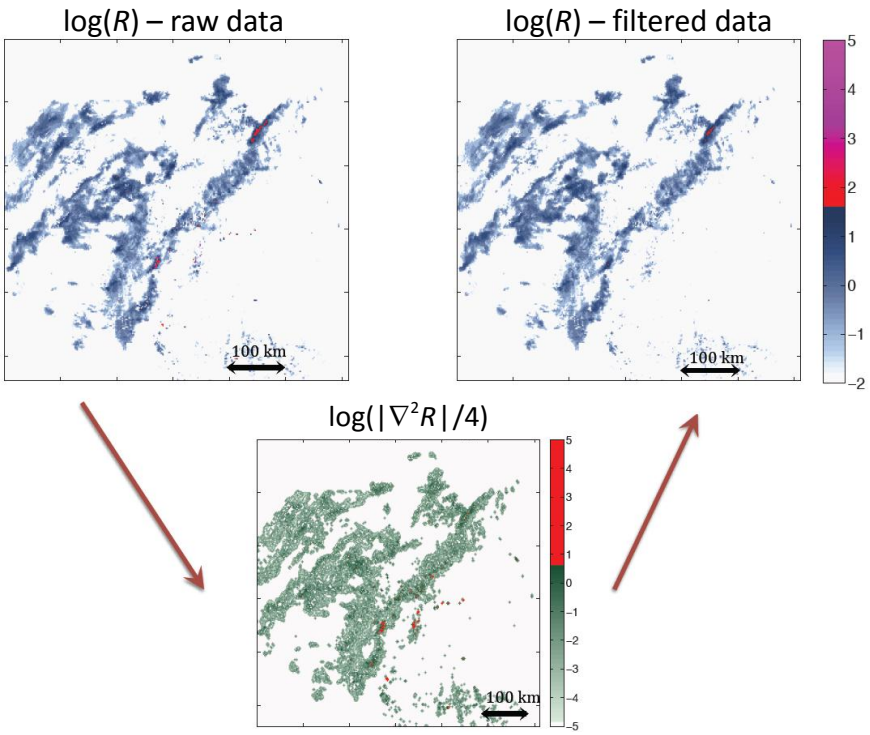


FIGURE 5.24: Schematic illustration of the elimination of spurious intensities. On the upper left panel the logarithm of the intensities of the raw data are shown. On the upper right panel the corrected intensities are shown and on the lower panel the logarithms of the absolute discrete Laplacians are shown. The spurious intensities are marked as red.

where $x_i, x_k \in \Re^2$ and x_k are the neighbouring points of x_i. If the precipitation random field is assumed to be wide sense stationary with mean value μ_R, standard deviation σ_R and spatial autocorrelation $\rho_R(h)$, then the expected value of the mean square error defined as

$$MSE = err_R^2 = E\left[\left(\hat{R}(x_i) - R(x_i)\right)^2\right], \tag{5.20}$$

is estimated as follows

$$err_R^2 = E\left[\left(R(x_i) - \frac{1}{n}\sum_{k=1}^{n} R(x_k)\right)^2\right] \tag{5.21}$$

$$= E\left[R(x_i)^2 - 2R(x_i)\frac{1}{n}\sum_{k=1}^{n} R(x_k) + \frac{1}{n^2}\left(\sum_{k=1}^{n} R(x_k)\right)^2\right] \tag{5.22}$$

and taking into account that

$$\text{Cov}\left[R(x_i), R(x_j)\right] = \sigma_R^2 \rho_R(h_{ij}) = E[R(x_i)R(x_j)] - E[R(x_i)]E[R(x_j)], \tag{5.23}$$

it can be shown that

$$E\left[R(x_i)^2\right] = \sigma_R^2 + \mu_R^2, \tag{5.24}$$

$$E\left[\frac{2}{n}R(x_i)\sum_{k=1}^{n}R(x_k)\right] = \frac{2}{n}\left(\sigma_R^2\sum_{k=1}^{2}\rho_R(h_{ik}) + n\mu_R^2\right), \tag{5.25}$$

$$E\left[\frac{1}{n^2}\left(\sum_{k=1}^{n}R(x_k)\right)^2\right] = \frac{1}{n^2}\left(n(\sigma_R^2+\mu_R^2) + 2\sum_{k=1}^{n-1}\sum_{l=k+1}^{n}(\sigma_R^2\rho_R(h_{kl}) + \mu_R^2)\right), \tag{5.26}$$

where k, l are indices referring to the surrounding points of $R(i, j)$. This means that the MSE is only dependent on the autocovariance of the field, i.e the autocorrelation and the variance. For highly correlated fields with small variance, the magnitude of MSE is negligible and is not expected to have a significant impact on the estimated statistics. Since this is usually the case for rainfall fields, the errors associated with the linear interpolation of the areas affected with clutter is not expected to influence the results.

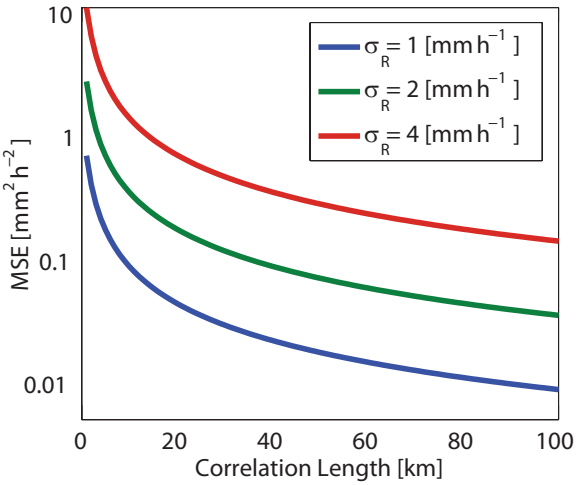


FIGURE 5.25: Expected values of MSE for the local linear interpolation in a 2D radar image. The assumption that the unknown intensity value of pixel on a regular grid can be estimated as an unweighed average of all its eight neighbours. The spatial covariance function is of the exponential type.

In figure 5.25 an illustration of the expected values of the MSE are shown for the exponential model, where the autocovariance function is defined as $\mathrm{Cov}(R(x_i), R(x_j)) = \sigma_R^2\left(\exp[-h/a]\right)$ for several values of correlation lengths a and variances. A further assumption for figure 5.25 is that the estimate of the unknown intensity of a point on a regular grid is defined as the (unweighed) average value of all eight surrounding point, which is in most of the cases true, since clutter points are localized and not clustered in space.

5.3.2 Anisotropy of the Precipitation Fields

One main issue concerning the structure of the precipitation fields is their anisotropy in space. The most common assumption, especially for stochastic modelling, is that precipitation fields at the meso-β

scale (typically 20-200 km) can be described as stationary and isotropic random fields [e.g. *Pegram and Clothier*, 2001a; *Over*, 1995; *Koutsoyiannis et al.*, 2011].

In terrain with complex topography this assumption may not hold true. *Ebtehaj et al.* [2010] found by analysing the two-dimensional power spectrum of an intense storm in Madison County (USA) that small and large scale features exhibit anisotropic structures dependent on the orientation of the mountain range. In particular, they found that high scale (low-frequency) features exhibit an elliptic anisotropic structure aligned parallel to the orientation of the mountain range and small scale features have the exact opposite behaviour.

Here a thorough analysis of the 7 years radar data-set is conducted in order to asses whether or not similar anisotropic structures occur, and what are the possible orographic effects.

The most common anisotropic behaviour is the so-called geometric-elliptic anisotropy. In random field analysis and especially in geostatistics the estimation of anisotropy is traditionally based on the two-dimensional variograms of the field. This methodology suffers from two major problems. First a parametric form of the covariance function, and thus the shape of the variogram, has to be selected a-priori. This can lead to serious problems if the fitting to the data is not good. Secondly, the estimation of the experimental variograms is computationally very expensive with complexity $O(N^2)$. The complexity can be reduced to $O(N \log N)$ for regular grids such as radar data, by exploiting the computational efficiency of the Fast Fourier Transform (FFT) [e.g. *Marcotte*, 1996], but the fitting algorithm of the theoretical covariance model remains computationally demanding.

In order to avoid both those problems, an alternative solution is applied here. The methodology has been developed by *Chorti and Hristopulos* [2008] and has the advantages of being non parametric, and thus no assumption of the covariance function is needed and also computationally very efficient with algorithmic complexity scaling as $O(N)$.

The methodology is based on the estimation of the gradients of the precipitation fields. Here only a brief description is given. The mathematical proofs can be found in *Chorti and Hristopulos* [2008] and *Swerling* [1962].

Let $R(\mathbf{r})$ be a wide sense stationary two dimensional random field with covariance function $C_R(\mathbf{r})$ with $\mathbf{r} \in \Re^2$. It can be shown that if

$$H_{ij} = -\frac{\partial^2 C_R(\mathbf{r})}{\partial r_i \partial r_j} \tag{5.27}$$

is the Covariance Hessian Matrix and

$$\begin{aligned} Q_{ij} &= \langle R_{ij} \rangle = \left\langle \nabla R(\mathbf{s}) \otimes \nabla^{\mathbf{T}} R(\mathbf{s}) \right\rangle & (5.28) \\ &= \langle \partial_i R(\mathbf{s}) \partial_j R(\mathbf{s}) \rangle \,, & (5.29) \end{aligned}$$

where $\otimes$ stands for the Kronecker product, then according to *Swerling* [1962],

$$\mathbf{Q} = \mathbf{H(r)}|_{r=0}. \tag{5.30}$$

This equation relates directly the spatial gradients of the random field R with the derivatives of the auto-covariance function.

Omitting all the related proofs that can be found in *Chorti and Hristopulos* [2008], here the final results for the two-dimensional stationary random fields are given. If the 2D random field $R(\mathbf{r})$ has correlation lengths ξ_1, ξ_2 along the principle axes of the covariance function that deviate anticlockwise with an angle θ from the coordinate system, and $\kappa_{2(1)} = \xi_1/\xi_2$ then

$$Q_{11} = \frac{\sigma_R^2 \zeta^2}{\xi_1^2} \left(\cos^2\theta + \kappa_{2(1)}^2 \sin^2\theta \right), \tag{5.31}$$

$$Q_{22} = \frac{\sigma_R^2 \zeta^2}{\xi_1^2} \left(\sin^2\theta + \kappa_{2(1)}^2 \cos^2\theta \right), \tag{5.32}$$

$$Q_{12} = Q_{21} = \frac{\sigma_R^2 \zeta^2}{\xi_1^2} \left[\sin\theta \cos\theta (1 - \kappa_{2(1)}^2) \right], \tag{5.33}$$

where σ_R^2 is the variance of the 2D field and the parameter ζ depends on the parametric form of the covariance function in the transformed isotropic coordinate system as

$$\zeta = \frac{1}{2} \Delta \widetilde{c}_R(0). \tag{5.34}$$

It is possible to overcome the dependence on the parametric form of the covariance function by dividing the equations 5.31, 5.32 ,5.33 and obtaining the result that

$$\theta = \frac{1}{2} \tan^{-1} \left(\frac{2 q_{off}}{1 - q_{diag}} \right), \tag{5.35}$$

$$\kappa_{2(1)} = \left(1 + \frac{1 - q_{diag}}{q_{diag} - (1 + q_{diag}) \cos^2\theta} \right)^{1/2}, \tag{5.36}$$

where

$$q_{diag} = \frac{Q_{22}}{Q_{11}} = \frac{\kappa_{2(1)}^2 + \tan^2\theta}{1 + \kappa_{2(1)}^2 \tan^2\theta} \tag{5.37}$$

and

$$q_{diag} = \frac{Q_{12}}{Q_{11}} = \frac{\tan\theta (1 - \kappa_{2(1)}^2)}{1 + \kappa_{2(1)}^2 \tan^2\theta}. \tag{5.38}$$

Even though this procedure is straightforward to apply, there are some restrictions that have to be taken into account. First of all the two dimensional random fields have to be wide sense stationary. For spatial precipitation, this assumption may not be valid. There are several examples of non-stationary random fields, especially in precipitation modelling. Probably the most widely known is the "universal multifractal" model in the case that the non-conservative fractional integration H is used [*Schertzer and Lovejoy*, 1987] or the 2D bounded multiplicative cascades *Menabde et al.* [1999]. Also another case of non-stationary random fields that have been proposed for precipitation simulation are the exponentiated 2D fractional Brownian motions [e.g. *Pegram and Clothier*, 2001a; *Rebora et al.*, 2006]. In the case of non-stationary fields, the analysis should be conducted on the spatial increments of the field.

The methodology applies for normally or lognormally distributed random fields [*Chorti and Hristopulos*, 2008]. The lognormal distribution for precipitation intensities as mentioned before is a very reasonable assumption and thus the approximation can be considered reliable. One problem that can probably arise is that the precipitation fields in space are intermittent. This means that they are not "perfectly" lognormally distributed but they rather follow a two state distribution with a probability mass at zero.

In order to asses the impact of this effect on the estimation of the anisotropy, a Monte Carlo experiment was designed. The numerical experiment was to threshold synthetic random fields at different levels and estimate the relative errors of the estimated anisotropy ratio. The random fields were generated as exponentiated Gaussian $\sim N(0,1)$ fields with a Whittle-Matern covariance function with $\nu = 2$. The details of the simulation procedure can be found in [*Chorti and Hristopulos*, 2008, Appendix B] [2]. In figure 5.26 it is shown that the introduced biases are negligible in comparison to ones associated with the method itself. Moreover, it is shown that the estimation procedure is more reliable for spatial fields with low correlations.

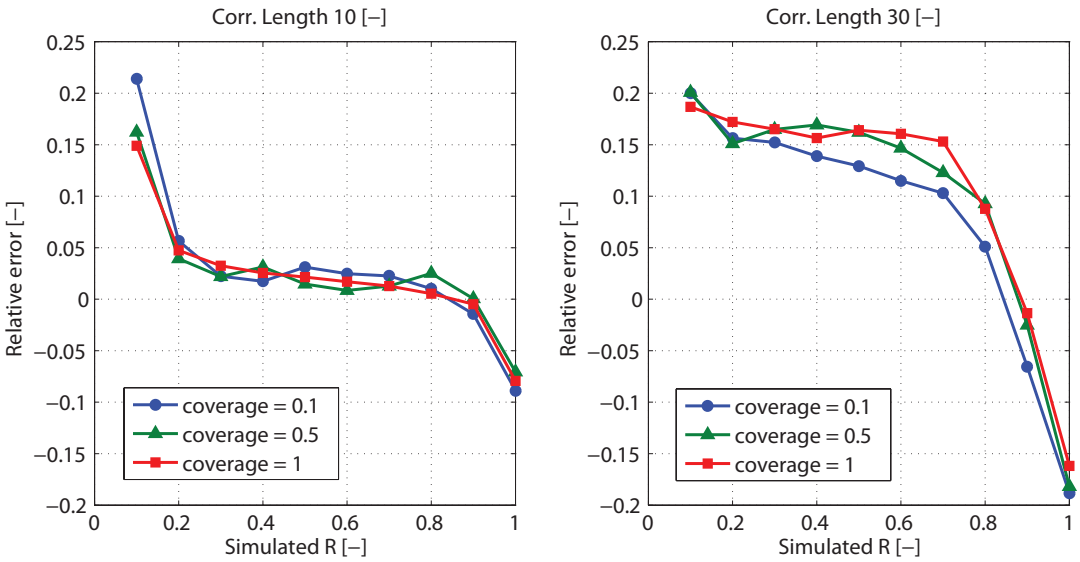


FIGURE 5.26: Relative error on the estimation of the anisotropy ratio due to thresholing. On the left panel, the results for a low correlation length 10 km are shown and on the right for a high correlation length 30 km.

Also the covariance function has to be continuous at the origin with continuous derivatives. The requirement of the continuity of the covariance function at the origin does not hold true, when the "nugget-effect"

[2] Matlab routines for the generation of the 2D random fields have been developed by *Chorti and Hristopulos* [2008] are publicly available in http://www.mred.tuc.gr/home/hristopoulos/dionisi.htm

[*Kitanidis*, 1997] is present. Sometimes the presence of stochastic small scale variability (nugget-effect) can be dominant in spatial precipitation fields [e.g. *Ciach and Krajewski*, 1999]. For the radar data of MeteoSwiss, the presence of such small scale variability is negligible. In addition, some of the most widely used models for the spatial covariance function (e.g. logistic, spherical) do not pose continuous derivatives at the origin. In data analysis though the issue of the parametric form of the covariance function is not a problems, as pointed out by *Chorti and Hristopulos* [2008], since the assumption that the measured fields are realizations of stochastic processes with covariance functions that poses a covariance function with continuous derivatives at the origin can be made. The choice of such parametric models is very large and flexible.

Taking into consideration all the above, the results for the two analysed areas are presented in figure 5.27.

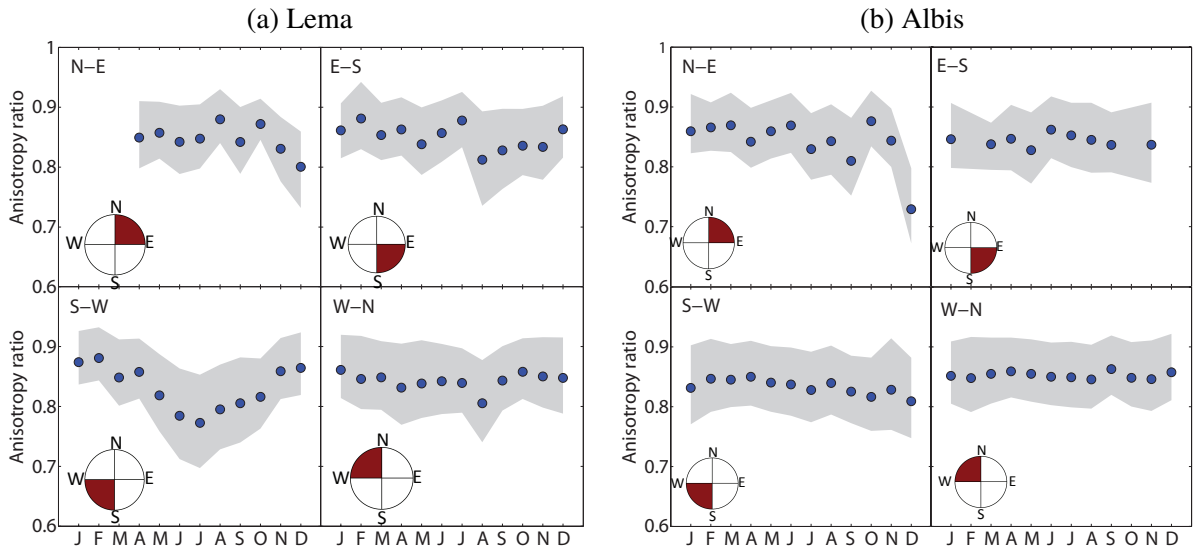


FIGURE 5.27: Monthly variation of the anisotropy ratio dependent on the wind direction for the two analysed areas (a) Monte Lema and (b) Albis. The dots represent the median value and the shaded area the 90% confidence intervals. The results are divided according to the 500 hPa wind direction of the upper air flow.

As is shown in the figure, the precipitation fields for all the cases are approximately isotropic since the ratio between the correlation lengths of the two principle axes (≤ 1) are close to 1. Ideally isotropic fields have a ratio equal to 1. For the area located in the Northern part of the Alps (Albis), it appears that there is not any significant monthly variation of the estimated values. Moreover there is no indication that the wind direction has an effect. Here the wind direction of the 500 [hPa] as measured from the closest atmospheric sounding location is used as an indirect measure of the mean driving force of the storms. This assumption was adopted since, as it will be later discussed in detail, the estimation of the storm velocities directly from radar data is not trivial and is prone to errors. One problem with this assumption is that balloon measurement are taken every 6 or 12 hours and the closest measurement in time to that of the recorded radar images was taken. The results for the NE direction are doubtful since the sample size was small. Indeed, precipitation generation linked to the anticyclonic atmospheric flow that generates NE upper level winds in this area is rare.

On the other hand, the results for the area lying on the Mediterranean part of the Alps (Lema) show a different picture. First of all, a weak seasonal pattern is identified when the 500 [hPa] wind direction is SW. This can be of major importance since in this area particularly all the large storms occur during

summer. Intense storms in this area are thus more anisotropic. One possible explanation is that storms arriving from the South are blocked by the Alpine mountain range and are organized in an anisotropic manner. Storm blocking in this area can be present due to the bow shape of the mountain range. *Houze et al.* [2001] and *Panziera and Germann* [2010] have shown that for this specific area, the wind flow pattern is a good descriptor of the precipitation accumulation with storms that are driven by intense wind conditions with high Froude numbers, showing significant differences from those that have lower ones. Storms with higher Froude numbers can overcome the mountain barrier in contrast to the ones that are driven by less intense wind conditions.

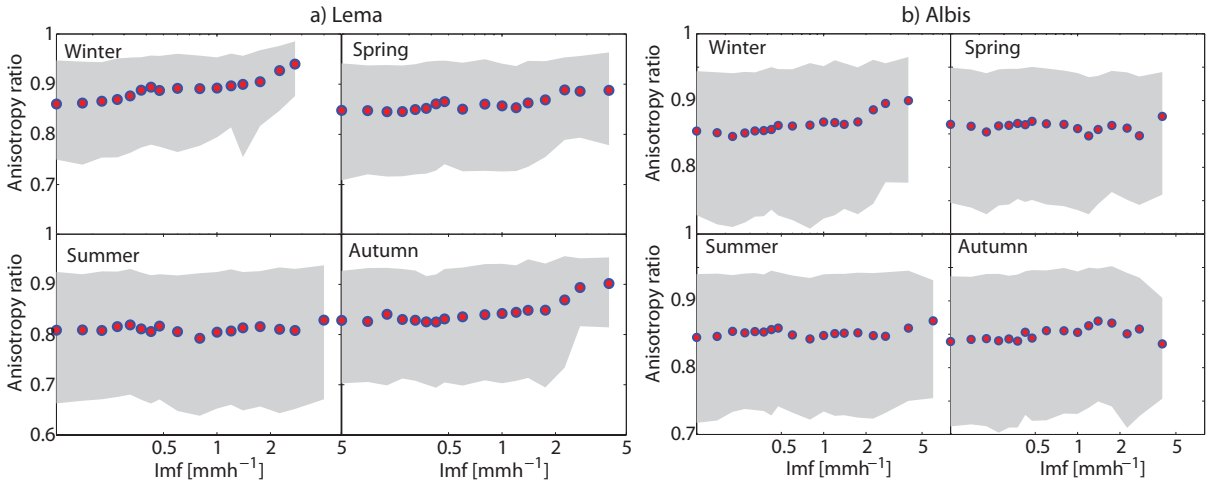


FIGURE 5.28: Dependence of the anisotropy ratio on the mean precipitation intensity on a seasonal basis. The left panel (a) shows the results for the area around the "Monte-Lema" radar and the right (b) for the area around "Albis" radar.

Another possible connection that can exist is the anisotropy ratio is related to the mean intensities of the precipitation fields. Figure 5.28 shows that these two variables are almost independent. There is a weak tendency of intense precipitation fields to be more isotropic, but due to the high variability of the results the statistical significance of such a connection is not strong.

In order to assess possible deviations from the pure geometric anisotropic behaviour as described before, the average two dimensional power spectrum was estimated for these areas [e.g. *Ebtehaj et al.*, 2010]. The power spectrum was estimated as the square of the absolute value of the 2D Fourier transform [*Bracewell*, 2000],

$$S(k_x, k_y) \quad = \quad \left|\mathcal{F}[R(x,y)]\right|^2 = \left| \int \int_{-\infty}^{\infty} R(x,y) e^{-i2\pi(k_x x + k_y y)} dx dy \right|^2 . \tag{5.39}$$

The Fourier transform was estimated using the FFT algorithm.

For the area of Albis the average 2D power spectrum showed a clear isotropic structure regardless of the season. For the area around the Lema radar (see figure 5.29), there is a clear anisotropic structure for the summer, the season during which all the intense convective storms occur. This is in agreement with the observations of figure 5.27(a). In contrast to *Ebtehaj et al.* [2010] the anisotropic structure is

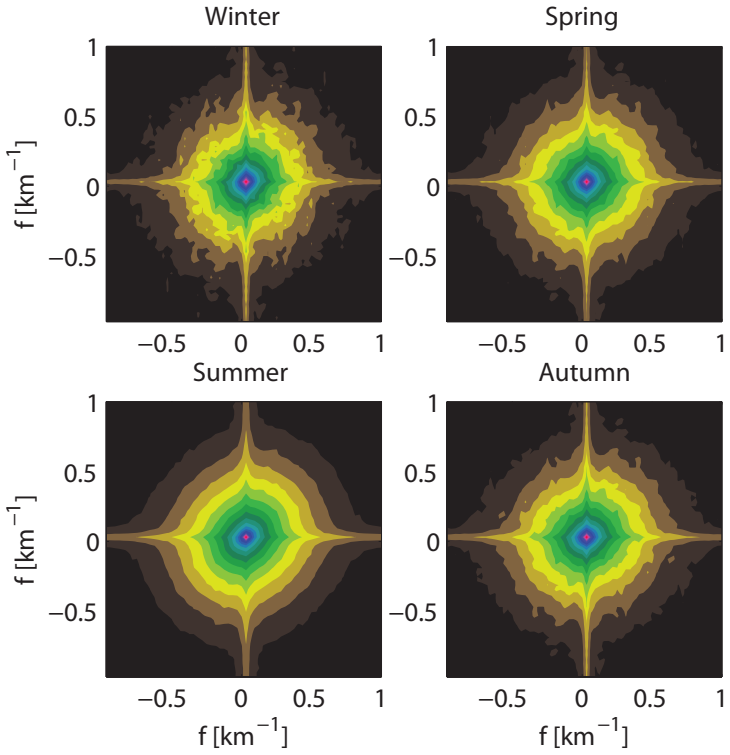


FIGURE 5.29: Logarithm of the average 2D power spectral density for the Monte Lema area on a seasonal basis.

the same, regardless of the wavelength, i.e. small and large scale features are anisotropically oriented in same direction. The axis of the highest directional correlation is oriented almost parallel to the direction of the mountain range. A physical explanation of this behaviour can be that the precipitation fields that are approaching the mountain range, are distorted and stretched parallel to the orientation of the mountain range when they are blocked and cannot overcome the mountain barrier.

5.3.3 Scaling of the Precipitation Fields in Space

As already mentioned before, the description of spatial rainfall as a scale invariance process has long attracted the interest of hydrology. In this section a detailed description of the radar derived precipitation as a scale invariant, multifractal process is provided. The focus is on seasonal and regional differences that can illustrate the effect of orography and the diffences between stratiform and convective precipitation and on the spatial structure of the precipitation fields.

Power Law Spectral Decay

One of the basic characteristics of the 2 dimensional scale invariant processes is that their power spectral densities are decaying as a power law with their associated wavenumber (or frequency). Here the spectral decays are estimated for the entire data set of 7 years of radar measurements for the two area sites.

The spectral slope was estimated using the Fourier and the Wavelet decomposition. The sample consisted of all the radar images that recorder precipitation for a fraction more than 10% of the area. Also the

very high frequency components (<4 km) were excluded from the fitting procedure in order to avoid the problems associated with data corruption due to clutter and noise, even though most of the problems have been eliminated with the data pre-processing filtering procedure.

First of all the monthly variations of the power law exponents were estimated. In figure 5.30 the slopes derived by the estimation of the radially averaged power spectra using the Fourier transform are shown. For both regions a clear seasonal pattern exist. Summer exhibits a steeper decay of the power spectrum in comparison to the other seasons. This shows that summer precipitation fields are more correlated in space rather than the winter ones. The differences are more apparent for the Mediterranean side of the Alps, where also the seasonal differences of depth accumulations are more pronounced (Figure 2.1). In all cases the spectral slopes are smaller than -2 which is an indication of nonstationarity in space. In summer, intense well organized convective cells in the mesoscale systems generate intense precipitation localized in space. On the contrary winter mesoscale systems are mainly driven by frontal activity with absence of areas of deep convection.

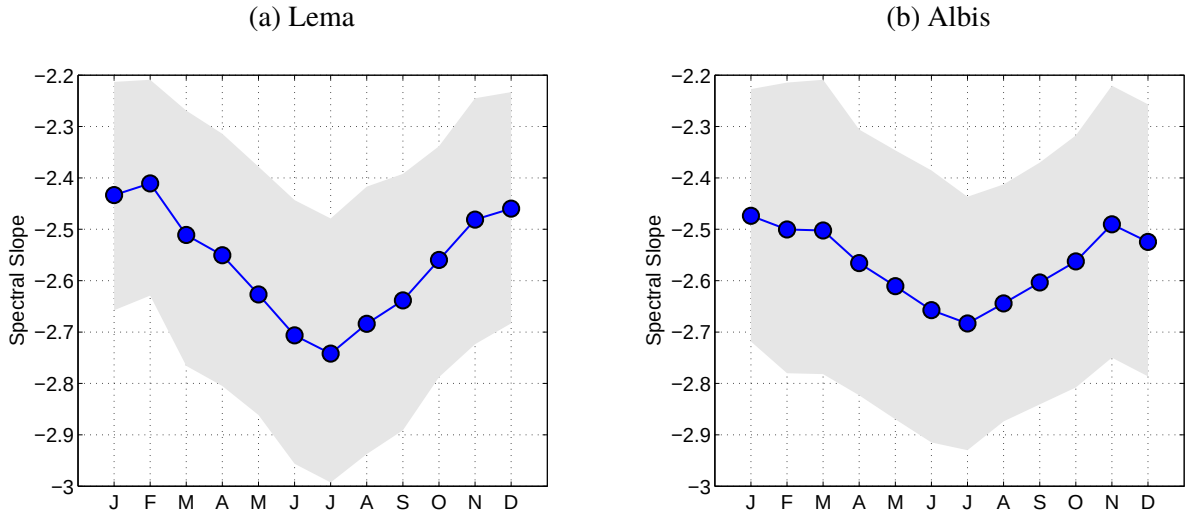


FIGURE 5.30: Monthly variation of the power law spectral slope using the FFT spectral estimation. The left panel (a) corresponds to the area around Lema and the right panel (b) for the area around Albis. The dots represent the median values and the shaded areas the 90% confidence limits.

The same behaviour concerning the seasonality of the spectral decay is found using wavelet analysis for the Mediterranean side of the Alps but not for the area around the Albis radar. The slopes that are reported in figure 5.31 correspond to the average value of the three diagonal components of the wavelet decomposition (horizontal, vertical and diagonal). The reason for this lack of seasonal signal can be that the spectral densities using the discrete wavelet decomposition of the signal, are evaluated only on scales that are integer powers of 2 (see Figure 5.32). So there is some incompatibility on the number of spatial scales that are used for the estimation of the power spectrum with the two methodologies. In the case of a less than perfect power law spectral decay, the two estimates can therefore be different. An example of this case can be illustrated in figure 5.32, where it is shown that the assumption of a power law spectrum is not perfect, and an indication of a transition regime between large and small scales is present.

Another interesting feature is that the main monthly differences of the power law coefficient of the spectral decay for the case of the Monte Lema area can be attributed to the power-law decay of the diagonal component of the field as shown in figure 5.33(a) [*Nicolis et al.*, 2011]. The differences in the direc-

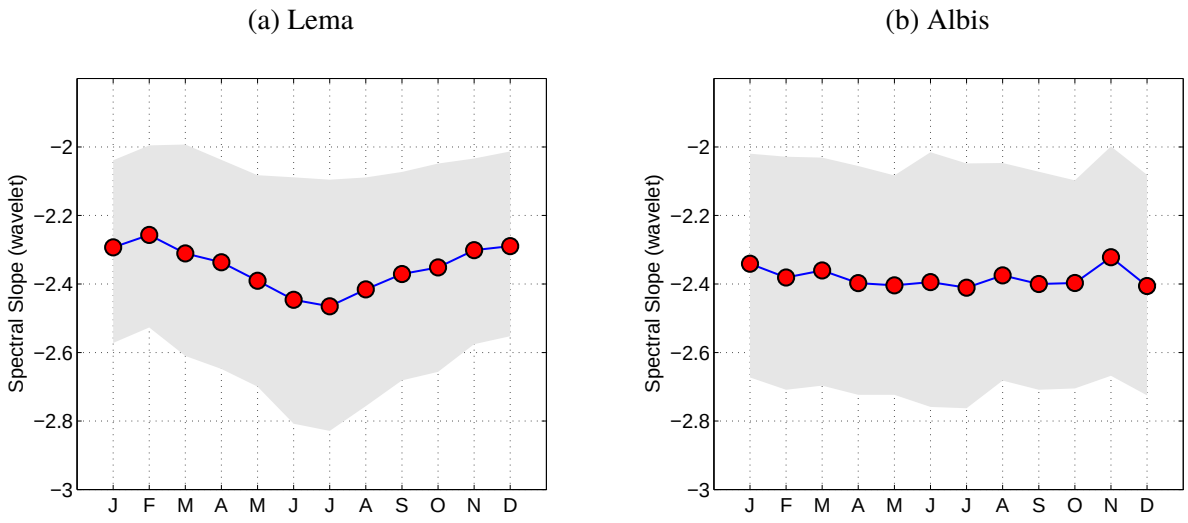


FIGURE 5.31: Monthly variation of the power law spectral slope using the wavelet decomposition of the radar image. The left panel corresponds to the area around Lema and the right panel for the area around Albis. The dots represent the median values and the shaded areas the 90% confidence bounds.

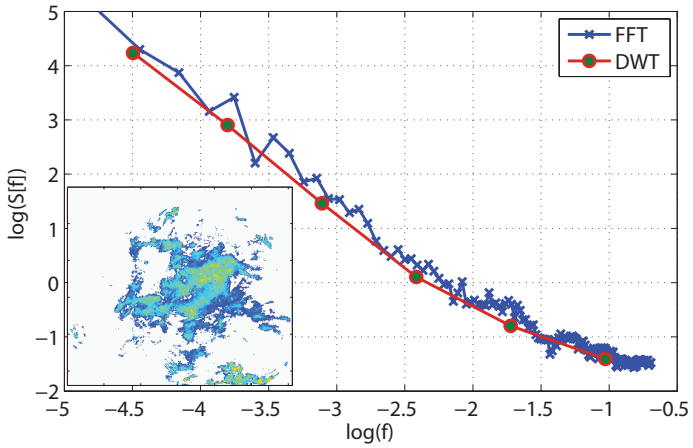


FIGURE 5.32: Comparison of the spectral densities estimated using the FFT and the discrete wavelet transform. The image correspond to the radar record for 21/8/2005(12:25 UTC). The selected wavelet is the db4 and the DWT spectrum corresponds to the average of its 3 components.

tional components of the power law spectral decay indicate the anisotropy of the field is dominant for the Mediterranean side of the Alps during summer (see also figure 5.29).

Having explored the seasonal patterns of the spectral decay, a possible connection to the mean intensity of the field is investigated next. As shown in figure 5.34, there are no strong dependencies between those two variables. Figure 5.34 corresponds to the spectral estimation using the Fourier image decomposition. For the case of the wavelet decomposition similar results were obtained and thus are not reported. The only possible connection is that for both areas there is a weak negative dependence of those two variables during the summer season. In general, precipitation fields with higher intensities lead to lower spectral slopes (i.e. higher correlation). The reason for that is that summer intense precipitation occurs when well structured convective systems that are highly spatially correlated are developed. On the other hand, the

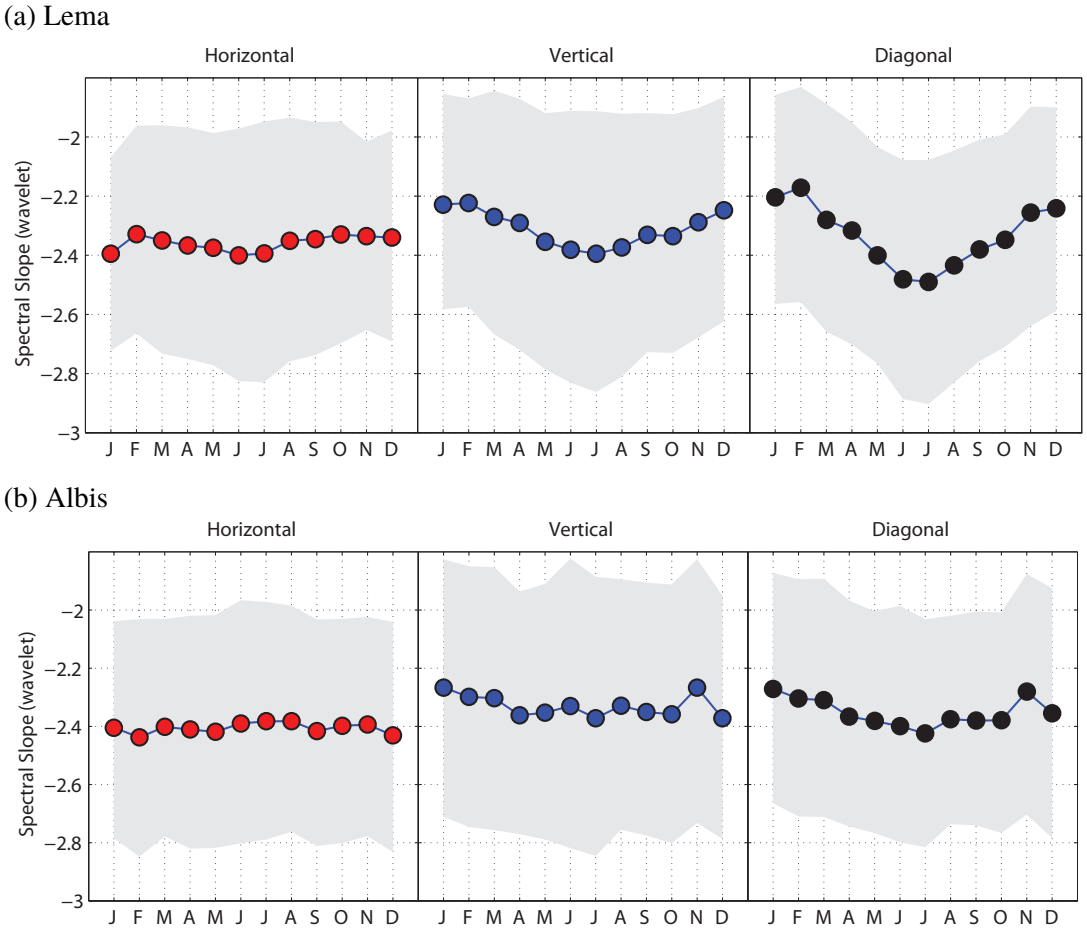


FIGURE 5.33: Monthly variation of the three (horizontal, vertical and diagonal) components of the power law spectral slope using the wavelet decomposition of the radar image. The upper panel corresponds to the area around Lema and the lower panel for the area around Albis. The dots represent the median values and the shaded areas the 90% confidence bounds.

correlation structure of winter stratiform precipitation is described by the general mesoscale features, the structure of which appears not to depend on the mean intensity of the precipitation fields.

The magnitude of the spectral slopes estimated here is in agreement with previous studies [e.g. *Pegram and Clothier*, 2001a; *Ebtehaj et al.*, 2010; *Mandapaka et al.*, 2010; *Clothier and Pegram*, 2001, among others] which suggest that they are rather general.

MRC Parameters

Precipitation structure in space has been described as the outcome of a two-dimensional MRC [e.g. *Over and Gupta*, 1994; *Kang and Ramírez*, 2010]. The most common procedure for the estimation of the parameters of a MRC in space is through the estimation of the scaling of the moments for some multi-resolution quantity. Usually this quantity is the local average of the spatial fields [e.g. *Over*, 1995; *Verrier et al.*, 2010; *Badas et al.*, 2005; *Deidda*, 1999] (elsewhere mentioned as the zero order wavelet coefficients [*Veneziano and Furcolo*, 2009]). One other class of multi-resolution quantities that have been also used for the estimation of MRC parameters are the wavelet coefficients of order larger than zero [e.g. *Ramírez-*

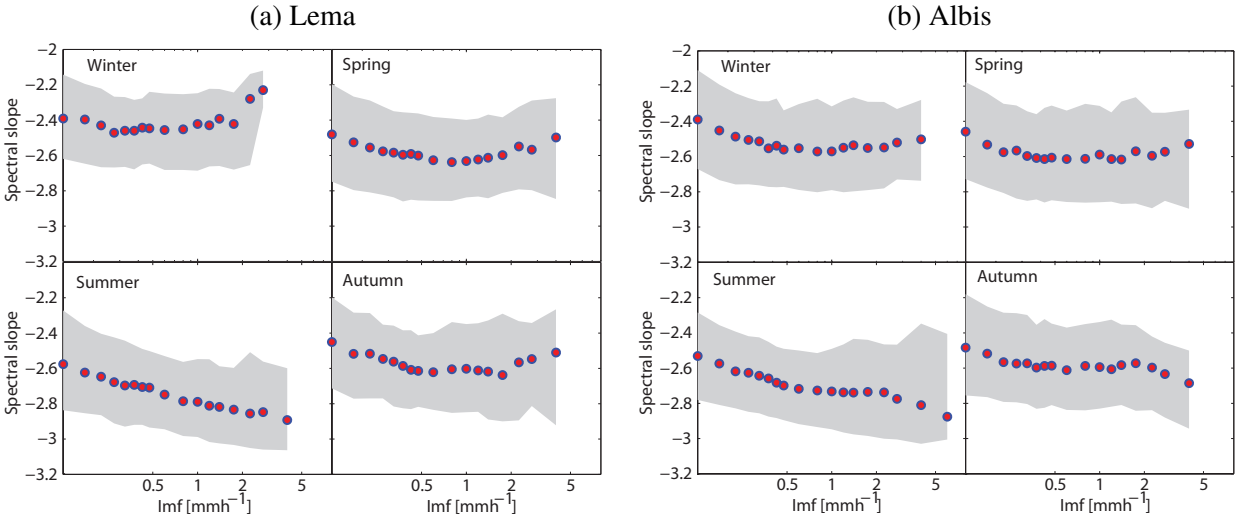


FIGURE 5.34: Dependence of the spectral slope on the mean areal intensity. The lefts panel illustrate the seasonal differences of the dependence of the mean areal intensity and σ^2 for the area around the Monte-Lema radar and the right one for the area around the Albis radar. The dots represent the median values and the shaded areas the 90% confidence bounds.

Cobo et al., 2011; *Audit et al.*, 2002; *Davis et al.*, 1994]. Several variants of these methods have been proposed with the aim to to increase their accuracy and robustness [e.g. *Serrano and Figliola*, 2009; *Decoster et al.*, 2000].

Here, the quantity that is used is the local average for several spatial resolutions. As shown in the previous section the power-law exponents of the spectral densities are less than -2 (Figure 5.30), this indicates that the precipitation field is not stationary in space [*Marshak et al.*, 1994]. In this case, the precipitation fields can be assumed as arising from a fractionally integrated realization of a MRC [*Tessier et al.*, 1993]. This is the main assumption also adopted in the simulation of the "Universal Multifractal" model [*Tessier et al.*, 1993; *Pecknold et al.*, 1993].

In this case, for the analysis of the precipitation fields, there are several options reported in the literature. The most straight-forward is to estimate the scaling of the moments of the absolute gradients [*Lovejoy et al.*, 1995] of the fields defined as

$$|\nabla R(x,y)| = \left| \sqrt{\left(\frac{\partial R}{\partial x}\right)^2 + \left(\frac{\partial R}{\partial y}\right)^2} \right|, \tag{5.40}$$

and estimated on a regular grid as

$$|\nabla R(i,j)| = \sqrt{\left(\frac{R(i+1,j) - R(i-1,j)}{2h}\right)^2 + \left(\frac{R(i,j+1) - R(i,j-1)}{2h}\right)^2}, \tag{5.41}$$

where h is the grid spacing.

Another methodology is to fractionally differentiate the precipitation field [*Nykanen*, 2008] or consider the field's spatial increments or generalized increments (e.g. using wavelet coefficients) [*Veneziano and Iacobellis*, 1999]. Even though *Veneziano and Iacobellis* [1999] have shown that the absolute values of the gradients do not scale and at small spatial scales this method can yield biased results, from numerical experiments conducted here, for typical MRC parameter values associated with precipitation the biases are negligible. Since this method is the most straightforward it is applied here.

The gradients of the precipitation fields are assumed to be a realization of a MRC. The parametric model that is chosen here is the beta-lognormal model, similarly to the analysis conducted for the time series. The maximum order where the moment scaling function can predict the MKP function is estimated with the exact same methodology as described in section 3.3. It should be noted that in this case, the intermittency that is described from the MRC corresponds to the gradient field and in terms of a physical interpretation is different from the intermittent cascades [*Over*, 1995; *Over and Gupta*, 1994], even though they are closely related. In previous studies, non-intermittent parametric forms of multiscaling fields have been used [*Nykanen*, 2008] for the analysis of precipitation fields. This approach however neglects the influence of the existence a probability mass at zero, and for this reason the parametric form of a MRC with an atom at zero has been used here. The propagation of the fraction of zeros across scales in the original, and in the gradient field, is very similar, and the interpretation of the β parameter directly refers to the intermittency properties of the rainfall fields. Figure 5.35 shows that the values of the β parameter for the gradient field is strongly correlated to the ones of the original fields. Thus an interpretation of the β parameter of the gradients as a descriptor of the scaling of the field's intermittency is reasonable.

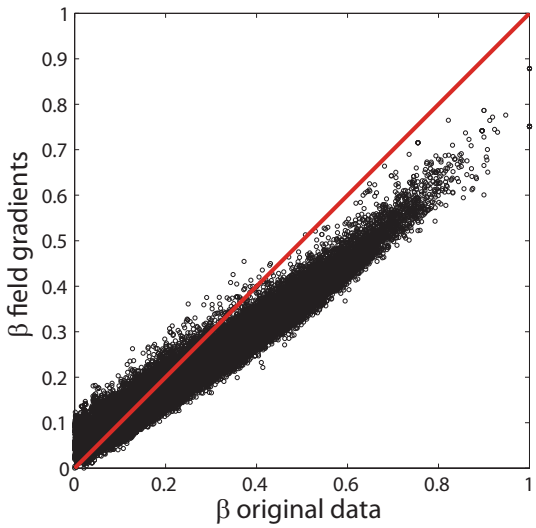


FIGURE 5.35: Scatter plot of the $beta$ parameter of a beta-lognormal MRC for the original radar fields and their gradients. The data correspond to the entire radar data set for the area around Monte Lema.

First of all, the patterns of the two parameters of the beta-lognormal model were estimated on a monthly basis (see figures 5.36, 5.37) for the two regions analysed here.

The monthly pattern of the intermittency parameter β is the same for both regions (figure 5.36). Summer season (JJA) has systematically higher values of β. This behaviour can be attributed to the convective nature of the summer precipitation that leads to very high intensity, spatially localised precipitation. Since β is directly connected to the proportion of the area that is covered with precipitation, this means that in summer precipitation fields are much more localized than in winter precipitation which is widely spread

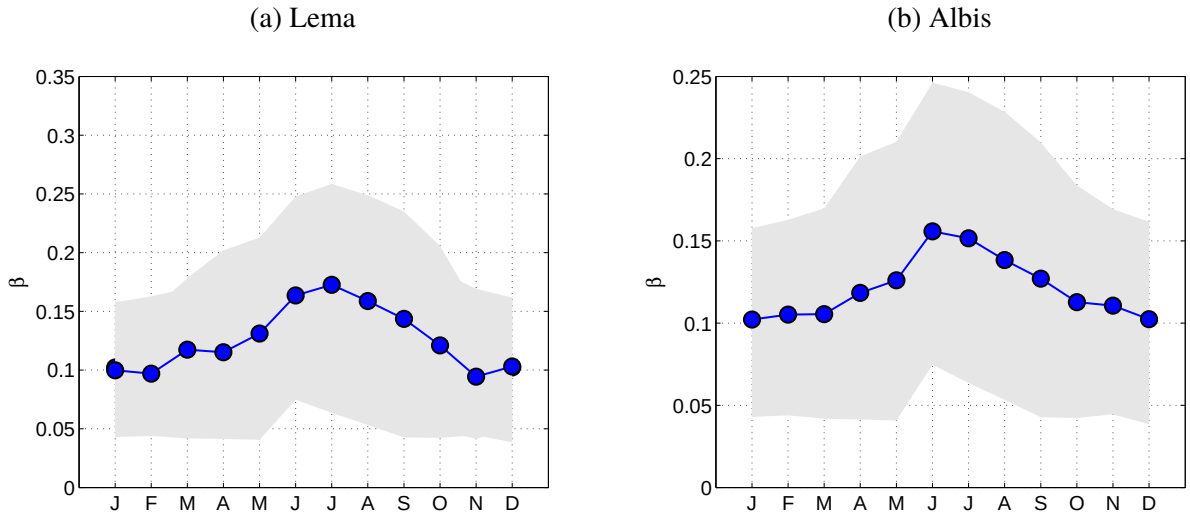


FIGURE 5.36: Monthly variation of the β parameter of the beta-lognormal MRC model. The left panel (a) corresponds to the area around Monte-Lema and the right panel (b) for the area around Albis. The dots represent the median values and the shaded areas the 90% confidence bounds.

(stratiform events).

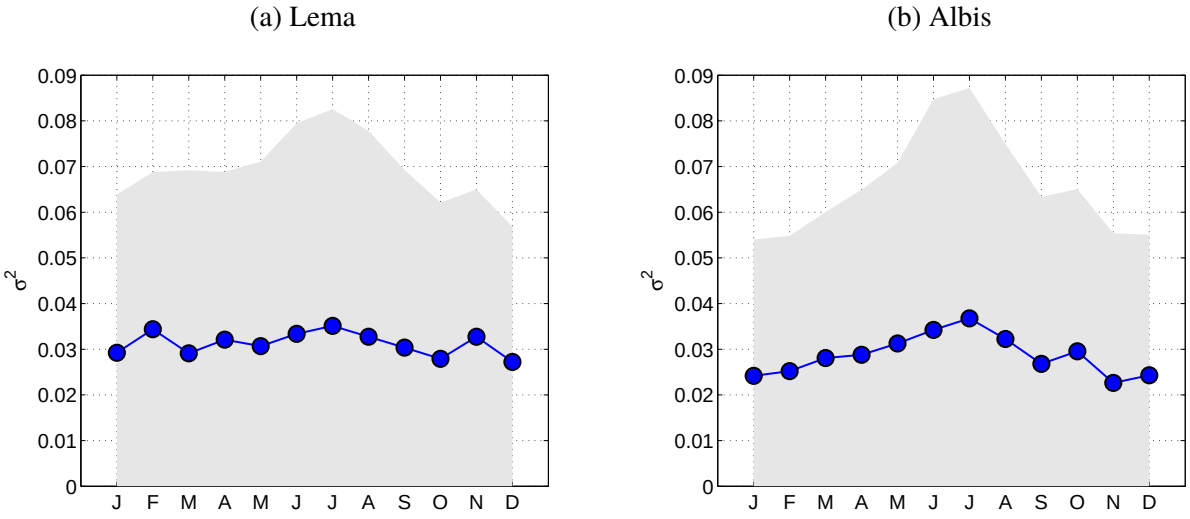


FIGURE 5.37: Monthly variation of the σ^2 parameter of the beta-lognormal MRC model. The left panel corresponds to the area around Lema and the right panel for the area around Albis. The dots represent the median values and the shaded areas the 90% confidence bounds.

On the other hand, the σ^2 parameter of the lognormal model does not show a distinct seasonal patter. For both regions the monthly differences of the median values of σ^2 are small. The main seasonal difference is that for the summer season the upper 90% quantile of the σ^2 parameter is higher. In other words the highest estimated values in the radar data-records are found during summer. The seasonal differences are more intense for the region located on the Northern part of the Alps.

One of the most common relationships between the MRC parameters and the nature of the precipitation that has been previously identified, is the connection of the mean areal precipitation intensity and the MRC parameters. *Over and Gupta* [1996] found that the intermittent parameter β for the precipitation fields was highly connected with the mean intensity of the fields. More, intense precipitation fields showed lower values of β. Intuitively, this behaviour is obvious since intense fields cover larger areas in space and thus the β parameter should be lower. They named this relationship "large scale forcing". In the same context, *Jothityangkoon et al.* [2000] identified similar connections of both σ^2, β parameters that they attributed to the spatial heterogeneity of the precipitation fields.

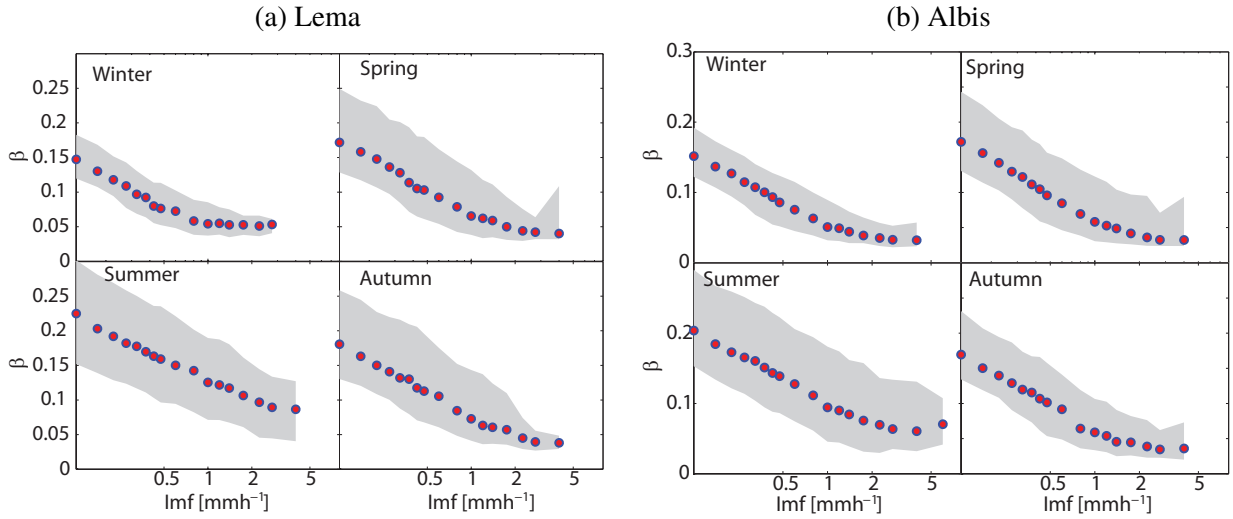


FIGURE 5.38: Dependence of the β parameter on the mean areal intensity. The left panels (a) illustrate the seasonal differences of the dependence of the mean areal intensity and β for the area around the Monte-Lema radar and the right (b) for the area around the Albis radar. The dots represent the median values and the shaded areas the 90% confidence bounds.

As shown in figure 5.38 the exact same behaviour as in *Over and Gupta* [1996] appears also in the data analysed here for both regions and for all the seasons. In order to obtain figure 5.38, the radar images were binned according to their mean areal intensities into classes and then for each one of them the median values and the 5-95% quantiles were estimated. One interesting feature is that during winter the spread is much smaller in comparison to the other seasons.

In figure 5.39 the dependence between the mean areal intensities and the σ^2 parameter is shown. The pattern that is observed is the same for all the seasons and areas that were studied. The pattern shows a positive correlation between those two parameters. High intensity precipitation fields give larger values of σ^2, i.e. higher multifractality of the field. The patterns are very similar to the ones obtained by *Jothityangkoon et al.* [2000] that analysed an extensive precipitation record in Australia. It must also be noted that the spread of the σ^2 values against the mean area intensity is very large. The large uncertainty, as shown by the spread of the estimates, limits the power of such a relationship as a prognostic tool.

Since both the β and σ^2 parameters are to some degree dependent on the mean intensity of the precipitation fields, they are also dependent between each other. The nature of this dependency is shown in figure 5.40. When the rainfall intensities are high, the β parameter is low and at the same time, the σ^2 is high. There is however also very high uncertainty in this relationship. Even though the overall tendencies are clear, the limitations of using such a relationship for stochastic simulation are apparent.

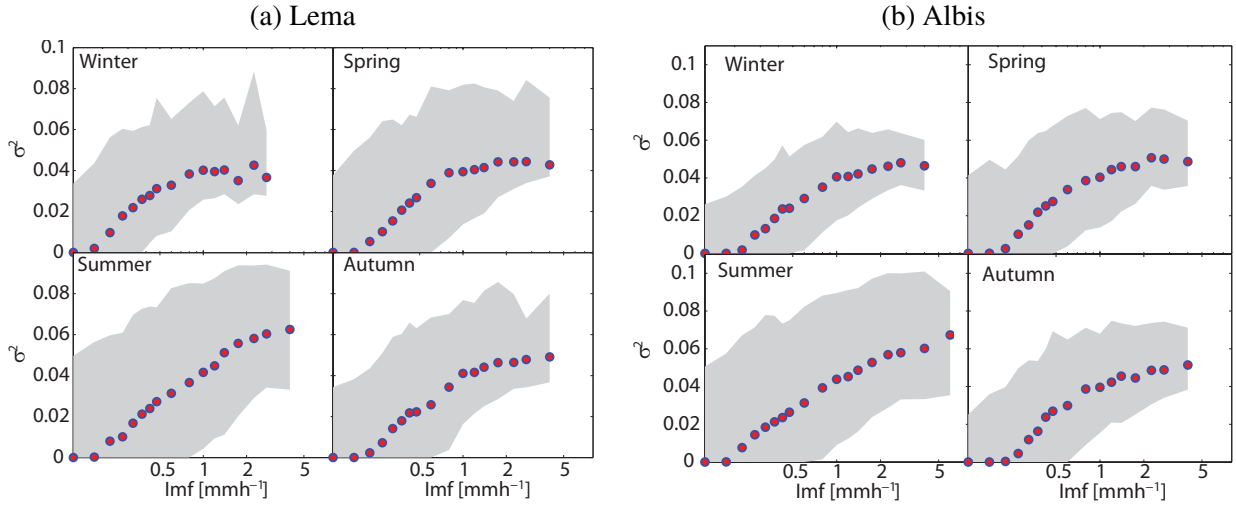


FIGURE 5.39: Dependence of the σ^2 parameter on the mean areal intensity. The left panels (a) illustrate the seasonal differences of the dependence of the mean areal intensity and σ^2 for the area around the Monte-Lema radar and the right (b) for the area around the Albis radar. The dots represent the median values and the shaded areas the 90% confidence bounds.

5.3.4 Linkages of Precipitation Structure to Atmospheric and Topographic Parameters

Establishing connections between the spatial structure of precipitation and various atmospheric and topographic descriptors, can shed light on the mechanisms generating precipitation. Such relationships can serve as predictive indicators for linking the statistical modelling of precipitation and atmospheric conditions. In addition, revealing such relationships can also advance our knowledge of the precipitaion formation process at small spatio-temporal scales, an issue that is still under investigation in meteorology and climatology.

Atmospheric Instability

One of the earliest attempts at linking precipitation structure and atmospheric thermodynamics can be found in *Perica and Foufoula-Georgiou* [1996]. The authors identified a strong relationship between the atmospheric instability and the scaling of the standardized precipitation fluctuations. This relationship was calculated for an extensive experimental set-up for a midlatitude convective system in central USA (Oklahoma-Kansas). In the analysed area orography did not influence the structure of precipitation.

One challenge in this section was to validate this observation and possibly expand it for other precipitation structure descriptors for a highly mountainous area. The challenge is twofold. First the dynamics of the precipitation generating mechanisms in mountain topography are much more complex [*Roe*, 2005; *Houze*, 2012] since the effect of orographic enhancement can be dominant. Second the quality of the data of operational weather monitoring networks is typically much lower. For instance the spatial density of the atmospheric sounding network, from which atmospheric instability is assessed is much smaller and the sampling resolution is much coarser (2 measurements per day). Taking into account those restrictions, if similar relationships are identified, this means that the strength of the predictive results of *Perica and Foufoula-Georgiou* [1996] is very high.

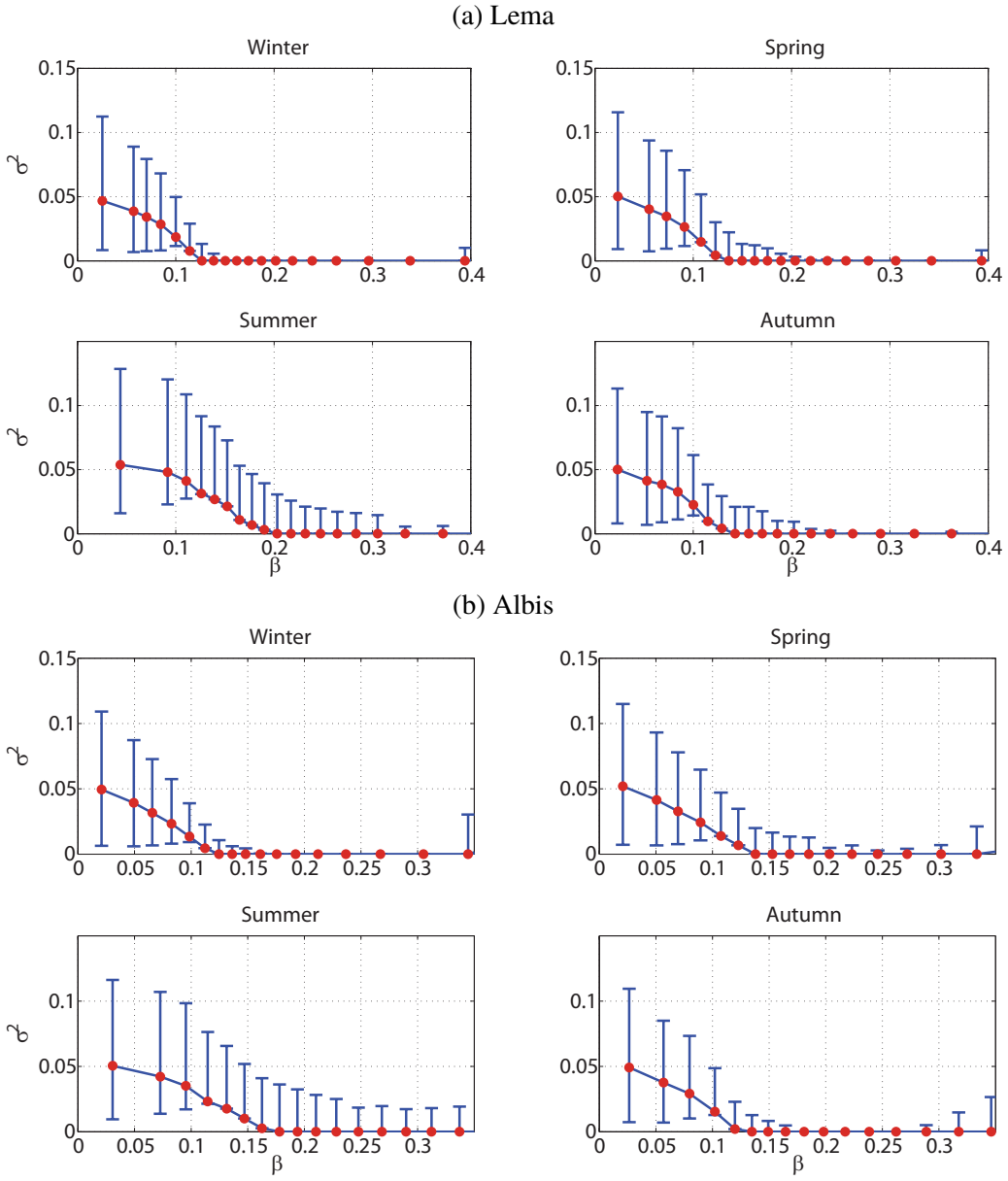


FIGURE 5.40: Dependence of the parameters of the beta-lognormal MRC for precipitation fields on a seasonal basis. The errorbars represent the 90% confidence bounds. The upper panel (a) shows corresponds to the area around the Monte-Lema redar and the lower (b) for the area around the Albis radar.

The indicator of atmospheric instability that will be used here is the convective available precipitation energy (CAPE). CAPE $[\mathrm{Jkg}^{-1}]$ is defined as,

$$\mathrm{CAPE} = g \int_{z_{LFC}}^{z_{LNB}} \left(\frac{T_p - T_e}{T_e} \right) dz, \tag{5.42}$$

where z_{LFC} is the altitude of the level of free convection and z_{LNB} the altitude of neutral vertical buoyancy. T_p is the virtual temperature of an air parcel adiabatically risen to a height z and T_e the virtual

temperature of its surrounding environment. In principle CAPE is the work per unit mass that the air parcel performs by rising from the level of free convection to its equilibrium level.

The integral of equation 5.42 can be numerically approximated from the readings of the balloon measurements. For the two areas that are analysed in this chapter, data from two different stations are taken into account. For the area around Monte-Lema, the data from the measurements from Milan are taken into consideration and for the area around the Albis the data from the Payèrne station are analysed (see figure 2.2). The two criteria that are used for the station selection, are spatial proximity and that the area under investigation and the balloon measurements lie on the same side of the main mass of the Alpine mountain range [3]. The second assumption is crucial since the meteorological conditions can be very different between the Northern and Southern sides of the Alpine mountain range.

Following the ideas of *Perica and Foufoula-Georgiou* [1996], the connections of the prestorm CAPE to the spatial statistics of precipitation are analysed. The statistics that are taken into account in this analysis are the mean intensity of the event and some spatial structure descriptors such as the mean storm power law spectral decay exponent, the mean storm σ^2 parameter of the beta-lognormal MRC, and the scaling exponent of the standardized fluctuations. The sample of radar images for every storm are the ones for which the ratio of wet area exceeds 30% of the area. This fraction of radar of images that are taken into account can be considered as a first approximation of the developed stage of a storm. As the value of the prestorm CAPE the closest measurement in time before the start of a convective event is taken. The reason why the prestorm measurement is taken into account is that this is a descriptor of the available energy that can be "released" during a convective event. During the storm evolution, CAPE is depleting and thus cannot serve as an indicator of the available energy for the storm development. There is an inherent problem of this definition, since there is not a necessity that the generation of the storm in the study area actually depends directly on the measured value of CAPE since the estimation of the energy could be several hours before the event initiation. According the study of [*Zhang and Foufoula-Georgiou*, 1997], since the temporal variation of the CAPE, as reproduced from a mesoscale model does not appear to be highly variable in time, the approximation of the definition used here as prestorm CAPE is considered reliable. Although a validation of such an assumption for an area with complex orography is not provided. In figures 5.41 and 5.42 the results for the Monte Lema area are shown. The results obtained for the area around the Albis area show the same behaviour and thus not reported here.

One of the very first indicators of how CAPE can affect the structure of the upcoming convective event is to identify potential connections between CAPE and the storm intensities. *Zawadzki and Ro* [1978] identified that the prestorm CAPE for an area in Canada was highly correlated to the maximum intensity of precipitation as derived both from radar and rain gauges. In Figure 5.41 a similar analysis is shown for the Monte Lema area. Since the areal precipitation statistics are of interest in this chapter, instead of analysing the point values of precipitation intensities, the analysis was conducted on the areal averaged rainfall intensities. As is shown, there is not a clear connection between CAPE and the mean storm intensity or the maximum storm intensity. Generally higher values of intensity are more likely to occur when the prestorm available energy is high, but a clear relationship cannot be established. This is a first indication of the complexity of the mesoscale system. A reason why a clear connection was not identified is that it is not clear if the analysed areas lie inside the area that is affected by the release of CAPE and the development of the convective system. In other word, there is a possibility that only part of the developed convective storm generates precipitation in the analysed area, that lead to the uncertainty of

[3] Atmospheric sounding data worldwide are publicly available in http://weather.uwyo.edu/upperair/sounding.html

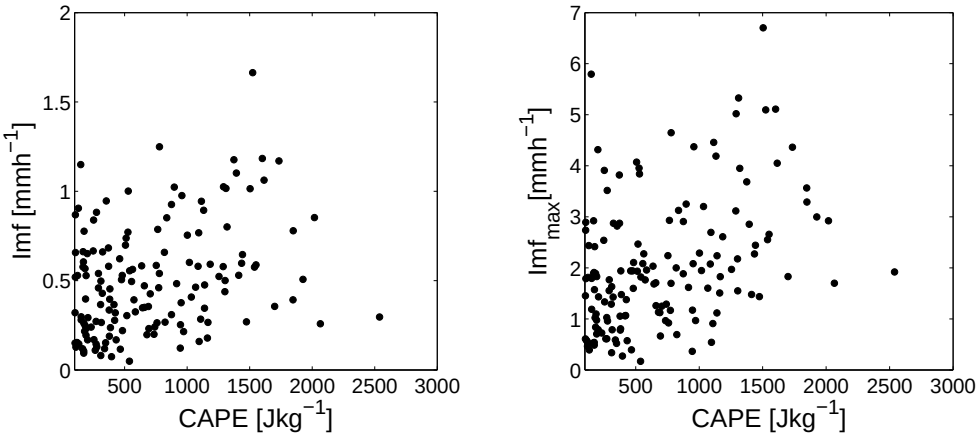


FIGURE 5.41: Prestorm CAPE versus the storm intensity. On the left, the mean spatially averaged storm intensity is shown and on the right the maximum spatially averaged storm intensity. The results correspond to the area around the Monte Lema.

the relationship between precipitation intensities and prestorm CAPE.

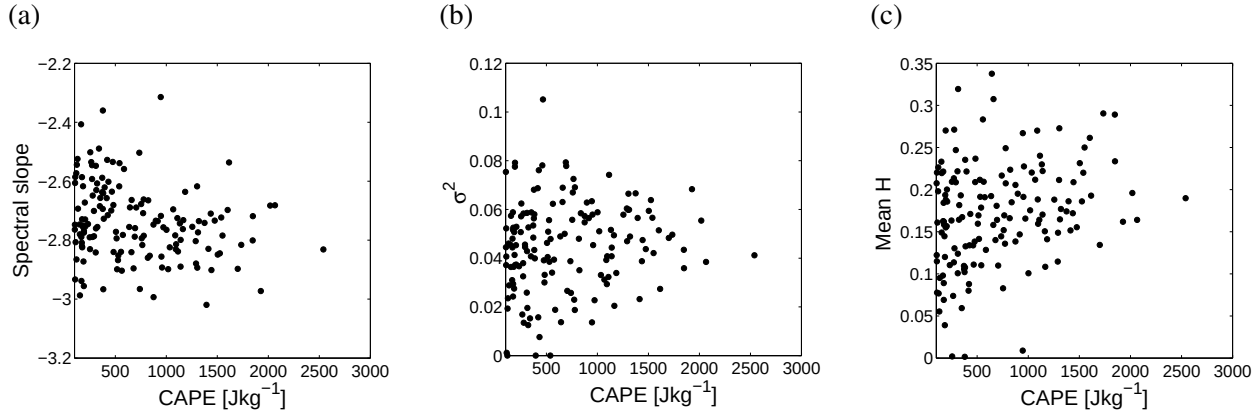


FIGURE 5.42: Prestorm CAPE versus scaling indicators. (a) CAPE vs the average value of the power law spectral decay. (b) CAPE versus the mean value of the σ^2 parameter of a beta lognormal MRC (c) CAPE versus the scaling exponent of the standardized fluctuations.

An extensive analysis expanding the one of *Perica and Foufoula-Georgiou* [1996] is shown in figure 5.42. The analysis searches for potential relationships between the various precipitation structure descriptors and the thermodynamic state of the atmosphere before the imitation of a convective storm. The first two descriptors are the power law decay of the radially averaged power spectral densities and the σ^2 parameter of the beta-lognormal MRC. The rational behind this investigation is to identify if CAPE can directly influence how well the precipitation fields are structured in space (spectral decays) and how locally intense ("spiky appearance") they can be. It is true that if the precipitation fields are realizations of a 2D scale invariant process, those two values are connected [*Veneziano and Langousis*, 2010]. The third descriptor is the one that was identified in *Perica and Foufoula-Georgiou* [1996] to be highly connected with the prestorm CAPE and is the scaling exponent of the standardized precipitation fluctuations. Roughly speaking the standardized fluctuations are defined in any spatial scale as the ratio between the wavelet (or detail) coefficients of the 2D wavelet image decomposition and their respective scaling (or approximation)

coefficients:

$$R'_{s,\lambda}(i) = \frac{R'_{\lambda}(i)}{\overline{R}_{\lambda}}, \quad i = 1, 2, 3 \quad , \tag{5.43}$$

where i stands for the three components of the 2D wavelet decomposition (horizontal, vertical, diagonal) and λ for the spatial scale of the wavelet decomposition. The spatial scale is defined here as $\lambda = 2^{m-1}$ where m is the level of wavelet decomposition. This quantity was found to follow a simple scaling relationship

$$R'_{s,\lambda}(i) \stackrel{d}{=} \lambda^{H_{st}(i)} R'_{s,1}(i). \tag{5.44}$$

The wavelet that was used here is the simplest one, the "Haar" wavelet, in order to be consistent with the study of *Perica and Foufoula-Georgiou* [1996]. The exponents $H_{st}(i)$ are estimated using linear relationships between the logarithms of the standard deviations of the standardized fluctuations and the respective logarithm of the scale [*Taqqu et al.*, 1995] and their average value is reported in figure 5.42. It can also be shown that there is not any clear relationship between those parameters and CAPE. One interesting feature though is that the power law exponent of the spectral decay is not highly variable during the storm, in contrast to the σ^2 parameter of the MRC and the scaling exponent of the standardized fluctuations that are (Figure 5.43), at least for the mature state of the storm that is analysed here. In particular the high variability for the scaling exponent of the standardized fluctuations, may reflect a major difference between this study and the one of *Perica and Foufoula-Georgiou* [1996] that reported the values of this exponent only for selected radar images during a storm and not their average value during the developed stage of each convective event. At the same time, the order of magnitude of the $H_{st}(i)$ coefficients is the same as the one reported in *Perica and Foufoula-Georgiou* [1996] and typically has a slight positive correlation to CAPE, even though not statistically significant.

As shown in figures 5.41 and 5.42, distinct relationships between atmospheric instability descriptors and precipitation structure do not appear to be valid. The main concern is whether this result is dependent on the data quality and restricted data availability, or the physics of the precipitation processes are so complex in the study area that do not allow such linkages to be identified. For example, one of the major atmospheric parameters that can potentially play a role in the study domain analysed here and not taken into account in this analysis is the strength of the wind flow, since it has been observed during the MAP experiment that precipitation processes are highly dependent on whether or not wind flows are blocked by the Alps [*Rotunno and Houze*, 2007; *Houze et al.*, 2001].

However an extensive data set was analysed in this study(7 years of radar data) and this is the main novelty in comparison to other studies that are in case specific. So concluding, in practical terms, the knowledge of the atmospheric instability from the available operational meteorological networks cannot provide a priori prognostic knowledge for the precipitation structure of the precipitation event on the long term. This statement does not necessarily contradict the results of *Perica and Foufoula-Georgiou* [1996] since it is not possible to identify whether the lack of such linkages between precipitation structure and atmospheric instability depend on the data quality or the physics of precipitation generation.

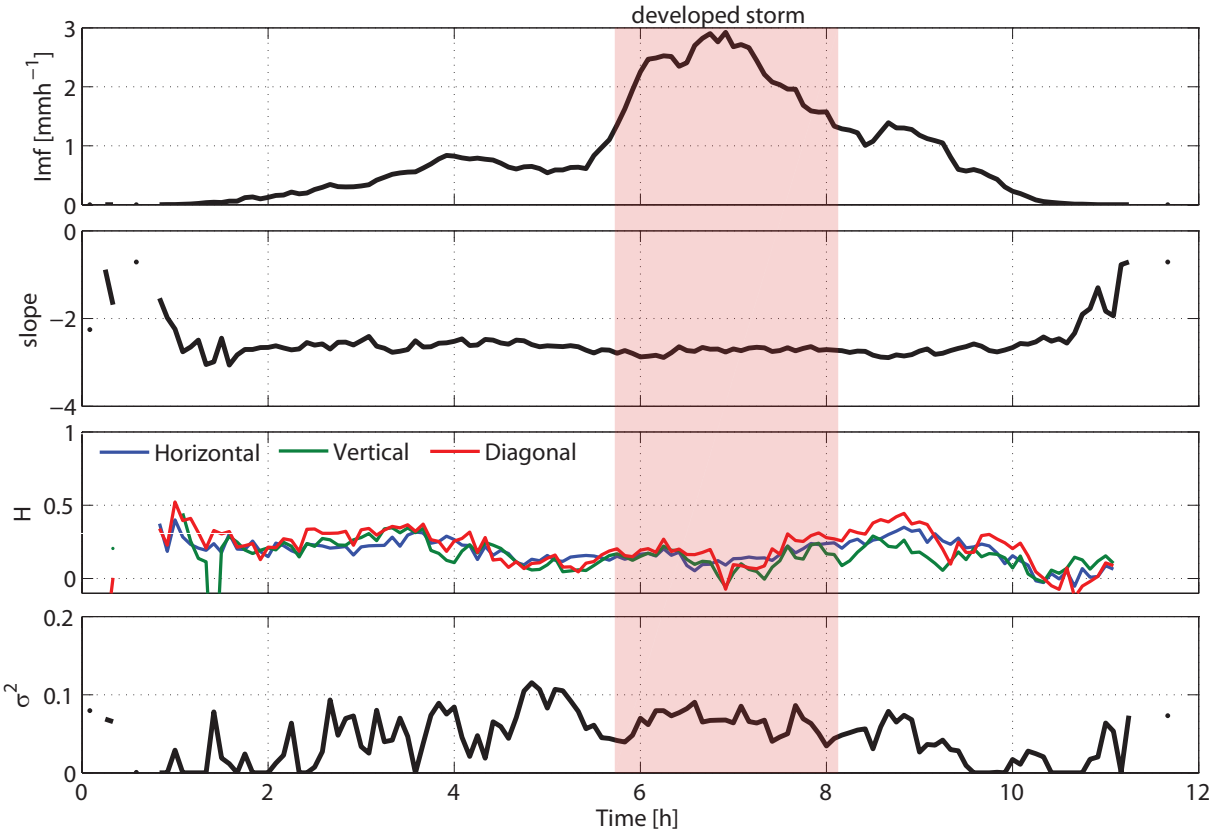


FIGURE 5.43: Temporal evolution of the various precipitation structure descriptors during a storm. From upper to lower panel the mean areal intensity, the spectral slope, the scaling exponent of the standardized wavelet coefficients and the σ^2 parameter of the lognormal MRC are shown. The illustrated example storm occurred in 6/8/2008 starting at 10:00 UTC.

Orography and Precipitation Enhancement

Since the areas that are analysed here have a highly complex orography, the effect of the topography on the spatial structure of precipitation has to be assessed as well. Previous studies show that topography can strongly influence the generation and development of precipitation fields. In qualitative terms, very informative reviews for orographic precipitation can be found in *Roe* [2005] and *Houze* [2012]. In contrast, numerical modelling and a precise quantification of those effects has been a very difficult task in meteorology, due to the various mechanisms that may be involved in the generation of precipitation processes over mountainous areas.

In statistical analysis, one of the first indicators of how topography modifies the precipitation processes, was the identification of precipitation gradients over mountain ranges [*Daly et al.*, 1994; *Johansson and Chen*, 2003; *Singh and Kumar*, 1997]. Those studies looked for empirical relationships between the long term accumulated precipitation depth and elevation for various areas in the world.

More recently, there have been several efforts to identify the connection of spatial precipitation structure features with orography. *Nykanen and Harris* [2003] and *Nykanen* [2008] tried to find relationships of the parameters of the multifractal log-Levy MRC model [*Schertzer and Lovejoy*, 1987] for radar derived precipitation intensities and mean storm elevations. The main result of those studies was that such a relationship exists for one of the model parameters and is dependent on the storm direction. Similarly to

this study, *Ebtehaj and Foufoula-Georgiou* [2010] identified that the power law exponent of the spatial spectral decay is also dependent on the mean topographic elevation of the storm, with higher altitudes having less steep decay (i.e. lower correlations). Generalizations of those results are not straightforward since previous studies are based on small amounts of data, typically a few storms. In this study the investigation uses a much larger sample size and is therefore statistically more robust.

A first question that is posed is whether or not the underlying topography of a precipitation field affects the precipitation occurrence process. In other words, the question is if precipitation fields are more prone to cover areas of higher or lower elevations. In order to asses this, a Monte Carlo experiment is constructed in order to identify the reference condition when the precipitation fields would be independent of the orography below them. The numerical experiment has the following steps:

- 2D random fields with equivalent size as the radar images and with selected correlation lengths are generated independent of the elevation
- A threshold is set to each of those fields according to a selected value of wet area ratio
- The mean altitude of the area where the random field exceeds the threshold is estimated

The procedure above is repeated 500 here for various correlation lengths and thresholds. Then the first three moments of the sample per parameter set are estimated. The autocorrelation function that was adopted here is the previously described exponential one, as it is a good descriptor of spatial precipitation processes [e.g. *Bell*, 1987; *Sigrist et al.*, 2012a]. In figure 5.44 the results for Monte-Lema and Albis are shown.

The lines that describe the independent case for the two areas are different due to the large differences of the elevation distribution in the two areas (figure 5.45). The area on the Mediterranean side of the Alps has a positive skewed distribution that has a wide plateau for a large range of elevations ($500 \sim 3000$ m) on the contrary to the Northern part of the Swiss Alps that has a also positive skewed distribution but without a plateau.

The first result from figure 5.44 is that the behaviour of precipitation occurrence in the two areas is substantially different. The Northern rim of the Swiss Alps (Albis) can be well described from the independent case, since all the data points lie close to the envelope that the independent simulations create. On the other hand this does not happen for the area around the Monte Lema, where it has been previously found that orographic enchantment can be strong [e.g. *Houze et al.*, 2001; *Rotunno and Houze*, 2007; *Rotunno and Ferretti*, 2001]. A weak seasonal pattern can be also identified in figure 5.44 (d). Summer precipitation seems to deviate more from the independent case than winter precipitation. This feature may have as well a physical explanation since the orographic effect in this area are expected to be much more intense during summer. The results shown in Figure 5.44 (d) suggest that if precipitation occurs with any wet area ratio, it is much more likely that will be located above areas of higher elevation in the Monte Lema area.

In addition it is worth investigating the dependency of some of the precipitation structure descriptors on the altitude of the precipitation field. This investigation is not always straightforward. Taking into account the dependencies that have already been identified, for example between the intensity and wet area ratio, or the intensity and the σ^2 parameter of the beta lognormal MRC, it is difficult to identify which is the direct effect of the precipitation field elevation on the various statistics and which is the result of interdependence between the parameters.

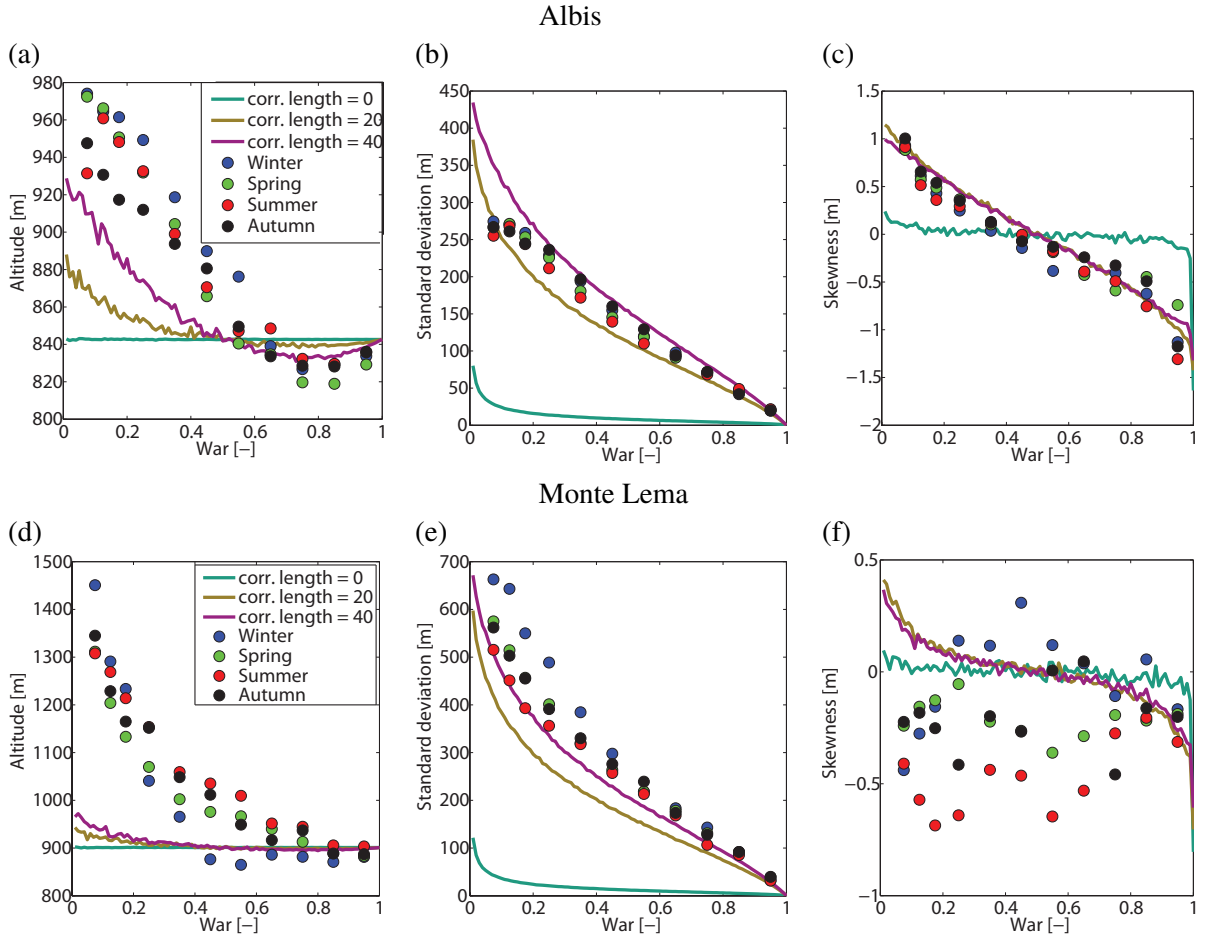


FIGURE 5.44: Dependence of the mean values of the first three statistical moment of the underlying altitude and fraction of wet area. The upper part corresponds to the area around the Albis radar and the lower to the area around Monte Lema. The dots correspond to observed data and lines to the simulated fields, independent of the underlying topography, with various correlation lengths, according to the respective legend.

The first connection that is examined is the one of the power-law decay of the power spectrum. As shown is figure 5.34 the relationship of the spectral slopes and the mean areal intensity is very weak, for all the seasons except summer. Moreover taking into account that the wet area ratio is very highly correlated to the mean areal intensity, it can be assumed that the altitude effects on the wet area ratio as shown in figure 5.44 will be negligible. In figure 5.46 the seasonal dependency of the spectral slopes on the altitude is illustrated. The mean value of the spectral slopes of precipitation fields that lie above distinct elevation ranges, is dependent on the altitude. The results for the cold seasons (Winter, Spring, Autumn) can be considered more reliable. First the seasonal pattern as previously described, shows that summer fields have steeper spectral decay. For the area located in the Northern part of the Swiss Alps (Albis), there is a clear positive correlation between the precipitation field's mean elevation and its corresponding spectral slope. In higher elevations, the precipitation fields have a lower absolute value of spectral slope indicating lower correlations. The same relationship though does not seem to hold for the Mediterranean site of the Alps (Monte Lema), perhaps only for mean altitudes > 1500m. The pattern is consistent for all the seasons which suggests that a probable connection between the uplifting of the air that leads to precipitation above a certain elevation areas and the statistical structure of precipitation exists. Moreover

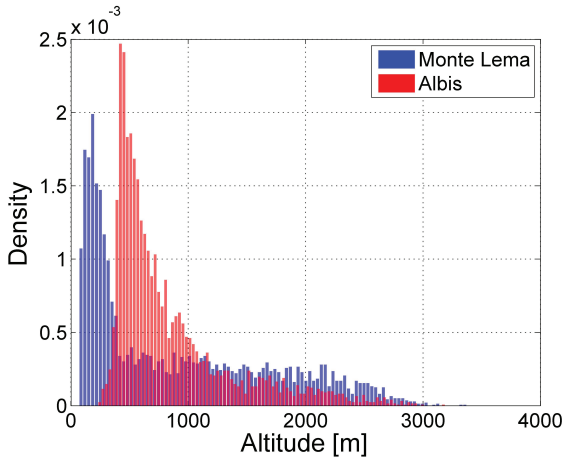


FIGURE 5.45: Distribution of the DEM for the areas "Mote Lema" and "Albis". The distributions are estimated in a resampled DEM at resolution 2×2 km^2 to match the radar sampling resolution.

this result confirms the results of *Ebtehaj and Foufoula-Georgiou* [2010], who found a similar pattern during one intense storm in Madison County (USA).

(a) Albis (b) Monte Lema

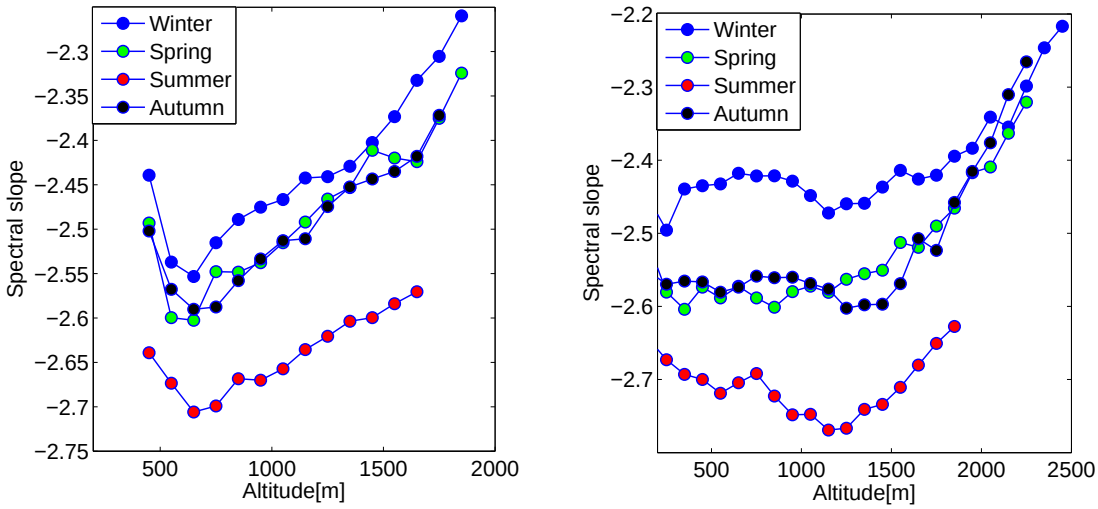


FIGURE 5.46: Mean elevation of the precipitation fields versus the mean value of the power law exponent of the radially averaged power spectrum.

The differences between the two regions are likely attributable to their very different distribution of elevations (figure 5.45),together with the very different blockage mechanisms of the airflow for the two areas. The Mediterranean site (Monte Lema) due to the bow-shaped mountain range of the Alps experiences much stronger air flow blockage conditions and as has been already found during the MAP observation period, the effects of the blockage can highly affect the convective precipitation effects [e.g. *Houze et al.*, 2001; *Rotunno and Ferretti*, 2001; *Panziera and Germann*, 2010; *Rotunno and Houze*, 2007]. TThe differences in precipitation orographic enhancement between the two areas are also strong [e.g. *Foresti and Pozdnoukhov*, 2011; *Foresti et al.*, 2012]

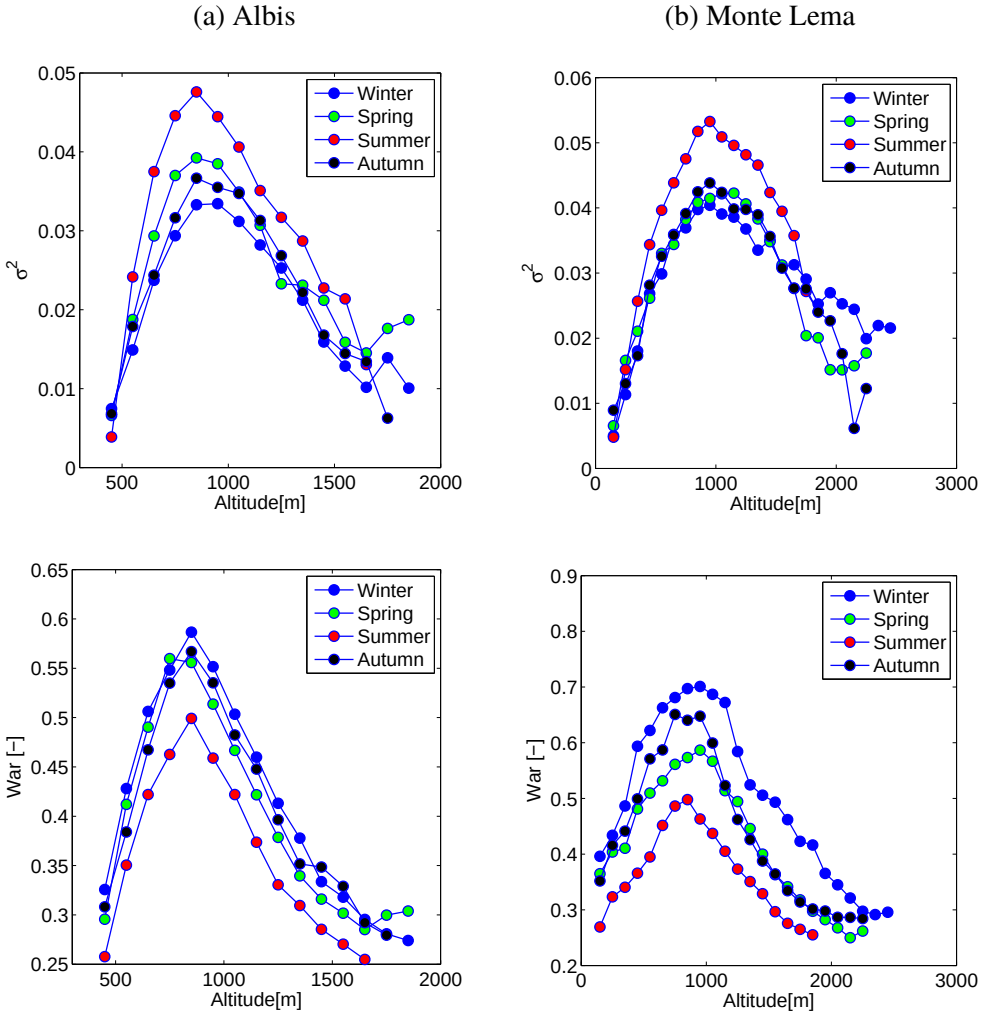


FIGURE 5.47: Dependence of the σ^2 parameter of the beta-lognormal MRC and the wet area ratio on elevation. In the upper panels, the mean value of σ^2 per elevation bin (100 m) is shown. In lower panels the same is shown for the mean value of the wet area ratio.

Another statistic that appears to have a connection to altitude is the σ^2 parameter of the beta-lognormal MRC. However since this parameter is positively correlated with the mean areal intensity (figure figure 5.39), which by itself is strongly dependent on the wet area ratio, a quantification of how much this statistic is dependent on the precipitation elevation is not straightforward. For example as shown in figure 5.47 the peak of the mean value of the σ^2 parameter coincides with the peak of the wet area ratio that in principle reflects the average altitude of the underline topography. But high values of the wet area ratio are connected with higher values of mean areal intensities, that are connected as well with high values of σ^2. Moreover, since figure 5.47 only shows the mean values of σ^2 per elevation bin, it must be noted that the samples $\{\sigma^2|(h_{min} < h < h_{max})\}$ have a very high variance that questions the statistical significance of the results. It was also investigated whether or not this relationship is different dependent on the storm direction as derived from the wind direction of the 500 hPa altitude from the most approximate atmospheric sounding, since *Nykanen* [2008] found that in a similar study this had a substantial effect. For both of the areas analysed here there was not any significant effect whether the storm was in the windward or leeward side of the mountain range.

For the other precipitation structure descriptors that have been analysed previously no direct dependence on the precipitation field mean elevation and are not reported here or any assessment would be highly biased by due to other dependencies (e.g. β parameter of the beta-lognormal MRC and wet area ratio).

The overall conclusion of this analysis is that in general there is not strong and highly statistical significant evidence that orography can modify precipitation in a distinct way. Since in this study a very large dataset was taken into account, a general conclusion can be that in a statistical sense no long term relationships can be identified. This is not however due to the lack of orographic influences that have been long established in meteorology [*Houze*, 2012], but perhaps mixing the various orographic effect of precipitation lead to those generally uncertain dependencies between storm structure and topography. This can have strong implication for the applicability of such identified relationships in stochastic modelling of spatial precipitation. On the other hand the few statistically significant dependencies that have been identified (e.g. Spectral slopes and topography) can be considered as robust and of fundamental importance.

6

STOCHASTIC MODELLING OF SPACE TIME PRECIPITATION

In the previous chapters an extensive analysis of the spatial and temporal structure of the precipitation process was conducted. In this chapter a novel stochastic modelling tool is developed in order to construct realizations of space-time precipitation fields. The overall goal of the model is to simulate space-time random fields that have the most significant structural properties of the precipitation process, capable of reproducing the statistics of the precipitation process as described in chapters 3, 5 for a wide range of spatial and temporal scales.

6.1 Introduction

Due to the increasing demand for high resolution spatial and temporal precipitation data, mainly driven by the requirements hydrological models, efforts have been devoted to the development of spatio-temporal stochastic rainfall models. The various precipitation measuring techniques spanning from classic point measurements (rain-gauges) to radar and even satellite measurements have given rise to diverse ideas for the simulation of precipitation.The most influential ideas are briefly reviewed here.

The various modelling approaches can be roughly divided into four separate classes. The first one simplifies the problem of spatiotemporal modelling to a multi-site temporal simulation framework [*Wilks*, 1998; *Brissette et al.*, 2007; *Wilks*, 1999; *Kleiber et al.*, 2012; *Bárdossy and Pegram*, 2009]. This approach is very appealing in the case when only point-scale measurements exist. Most of the models in this class operate on the daily time scale. This class of models exploits stochastic processes spanning from multivariate chain dependent processes [e.g. *Wilks*, 1998] to copula based approaches [*Bárdossy and Pegram*, 2009] and hidden Markov processes [*Hughes et al.*, 1999]. In general, since these models focus mainly on the stochastic simulation of precipitation in time at specific locations, their representation of the spatial distribution of rainfall is poor.

The second approach is based on the theory of point processes and is a generalization of the ideas that were first introduced for rainfall modelling in time. These models have become some of the most widely applicable and robust tools for precipitation simulation [*Burton et al.*, 2008, 2010b; *Leonard et al.*, 2008; *Cowpertwait et al.*, 2002; *Beuchat et al.*, 2011]. Several rainfall models based on point processes with

different degrees of complexity have been developed. All of the approaches simulate rainfall patterns in space as a superposition of two dimensional pulses (raincells). The simplest, yet the most widely applicable modelling approach is to assume the shape of the pulses as discs that have uniform intensity during their life time [*Cowpertwait et al.*, 2002; *Burton et al.*, 2008]. Under this assumption the analytical derivations of the model's statistical properties across scales is possible. This approach even though it has been found to yield reasonable results, several structural features of precipitation such as its enhanced spatial variability, its kinematic properties (e.g. advection) etc. More complex approaches include the well known WGR model [*Waymire et al.*, 1984] that is able to reproduce more realistic precipitation patterns including anisotropy, advection, etc. Unfortunately, the increase in the model complexity makes its calibration not an easy task. Other similar modelling techniques replace the uniforms raincells with more realistic constructions such as pulses with a Gaussian shape etc [e.g. *Féral et al.*, 2006]. Generally such models have been rather cumbersome to calibrate even if analytical expression for their statistical properties exist. Apart from the calibration, the major problem associated with these models has been their rather poor representation of high resolution precipitation properties both in time and space due to the simplifications that they adopt.

The third approach is based on the theory of random fields. Generally these models have been developed to fill the gap of the small scale precipitation simulation in space and time (typically $\sim$1 km^2, $\sim$5 min). The vast majority of these models simulate precipitation commonly as a non-linear transformation of two or three dimensional Gaussian random fields (mentioned as "latent" Gaussian processes as well) adopting some parametric form for their autocovariance [e.g. *Pegram and Clothier*, 2001b; *Bell*, 1987; *Durbán and Glasbey*, 2001; *Kumar and Bell*, 2006; *De Michele and Bernardara*, 2005; *Koutsoyiannis et al.*, 2011]. The most popular representations of the spatial structure of precipitation at the mesoscale, includes models with covariances of the exponential type [*Bell*, 1987; *Sigrist et al.*, 2012a], fractional Gaussian noises [*Koutsoyiannis*, 2011], and fractional Brownian motions [*Pegram and Clothier*, 2001a]. This class of models has been used more frequently for event based simulations and only few approaches have been directed at constructing a continuous simulation tool.

The final approach, is based on the theories of scale invariance and especially the use two or three dimensional multifractal processes for rainfall simulation. This approach due to its very attractive feature of linking statistics across spatio-temporal scales in a parsimonious manner has gained a lot of popularity [e.g. *Over and Gupta*, 1996; *Deidda*, 2000; *Gires et al.*, 2012; *Kang and Ramírez*, 2010; *Menabde et al.*, 1997b; *Pathirana and Herath*, 2002]. It has been shown in these studies that this simulation framework allows a satisfactory representation of several key aspects of precipitation, e.g. the power law distribution of extremes etc. One of the main constrains though of such models is their simulation procedure. The simulation tool for these processes has traditionally been the discrete version of the multiplicative random cascades, which simulate very unrealistic patters in space with a very characteristic "blocky" structure [*Kang and Ramírez*, 2010]. For this reason "continuous" simulations algorithms have been developed as well [*Pecknold et al.*, 1993], however with many fewer applications reported in the literature.

6.2 The STREAP Model

In this section the mathematical formulation of the building blocks of a new space-time stochastic model for rainfall are presented. The model is named STREAP as an abbreviation of **S**pace **T**ime **Re**alizations of **A**real **P**recipitation. The structure of the model is a substantial improvement of the previous works of *Bell* [1987], *Kundu and Bell* [2003] and mainly *Pegram and Clothier* [2001a]. The model adopts well

established ideas developed in the analysis of time-series and random fields spanning from point processes to multidimensional random fields. The aim of the model is to simulate space-time random fields that are consistent with the precipitation structure and especially reproduce the basic features and dynamics of the process for fine resolution scales (~ 5 min, ~ 1 km^2). The basic structural features of precipitation that are taken into account here are its intermittent nature, its correlation patterns in space and time, its positively skewed probability distribution and the dynamics are the growth and decay of storms as well as their movement (advection).

6.2.1 General Concept

The main idea of STREAP is that space-time precipitation can be considered as a realization of a stochastic process that is a composite of various sub-processes, each one describing some structural feature of precipitation. The structure of the STREAP model is based on a three-stage hierarchical basis each one described by a different stochastic process. The three stages are:

1. A storm arrival process
2. A process describing the within-storm temporal evolution of areal precipitation properties
3. A three dimensional process describing the space-time storm evolution

A schematic representation of the model is illustrated in figure 6.1.

This structure is key to the application of the model because it separates different processes in an explicit manner and draws clear connections with the natural phenomenology of precipitation at each stage.

TABLE 6.1: Parameters of the STREAP model

Parameters of the STREAP model	
1st stage	
$\kappa_{wet}, \sigma_{wet}, \theta_{wet}$	Parameters of the Gen. Pareto distribution of the wet runs
μ_{dry}, σ_{dry}	Parameters of the Lognormal distrubution of dry runs
$\alpha_{wind}, \beta_{wind}$	Parameters of the Gamma distribution of wind speed
2nd stage	
$\alpha_W, \alpha_I, \nu_W, \nu_I, \alpha_{WI}, \nu_{WI}, \rho_{WI}$	Parameters of the Whittle-Matern Covariance function
P_c, ν_c	Parameters of the T copula
$\alpha_\beta, \beta_\beta$	Parameters of the Beta distribution
$\alpha_\gamma, \beta_\gamma$	Parameters of the Gamma distribution
3rd stage	
α_g	Spatial correlation length
cv_r	Coefficient of variation of spatial precipitation
ϕ_i, θ_i	Parameters of the ARMA process

6.2.2 Storm Arrival Process

The storm arrival process is the simplest one possible and is an alternating renewal process describing a sequence of dry and wet periods [e.g. *Bernardara et al.*, 2007; *Ng and Panu*, 2010; *Veneziano*, 2002;

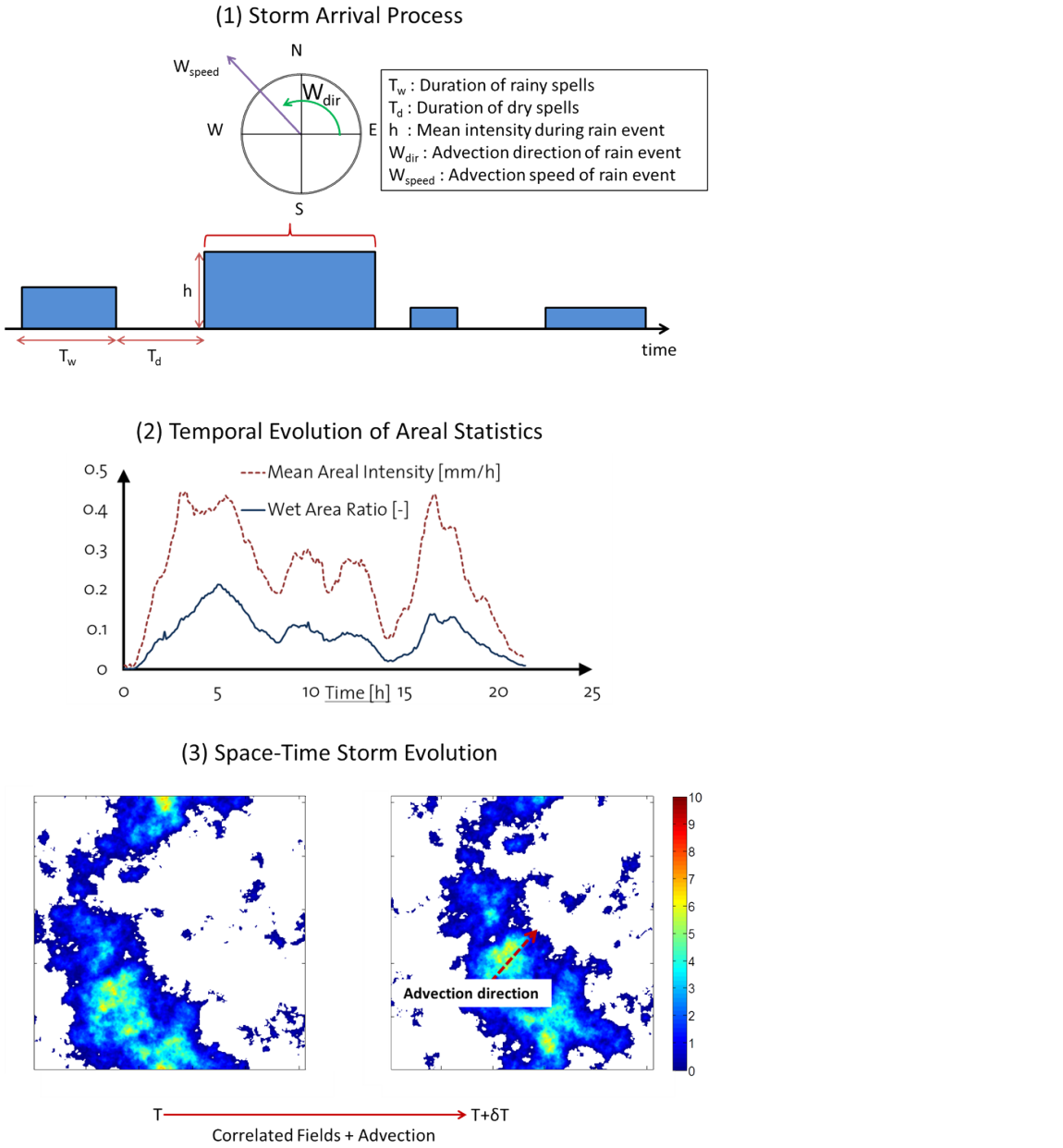


FIGURE 6.1: Schematic representation of the STREAP stochastic model

Roldan and Woolhiser, 1982; *Menabde and Sivapalan*, 2000]. This stage is almost identical to the String of Beads (SBM) model developed by *Pegram and Clothier* [2001b]. Specifically the dry and wet spell durations are sampled from appropriate probability distributions. The dry and wet lengths at each spell are iid and mutually independent.

The probability distribution of the wet durations is a generalized Pareto with a probability density function

[*Embrechts et al.*, 1997]:

$$f(x|\kappa_{wet},\sigma_{wet},\theta_{wet}) = \begin{cases} \left(\frac{1}{\sigma_{wet}}\right)\left(1+\kappa_{wet}\frac{x-\theta_{wet}}{\sigma_{wet}}\right)^{(-1-1/\kappa_{wet})} & \text{, if} \quad \kappa_{wet} \neq 0 \\ \\ \left(\frac{1}{\sigma_{wet}}\right)\exp\left(-\frac{x-\theta_{wet}}{\sigma_{wet}}\right) & \text{, if} \quad \kappa_{wet} = 0 \end{cases} \tag{6.1}$$

for $\theta_{wet} < x$, when $\kappa_{wet} > 0$, or for $\theta_{wet} < x < -\sigma_{wet}/\kappa_{wet}$ when $\kappa_{wet} < 0$.

and the dry durations follow a lognormal distribution [*Papoulis and Unnikrishna*, 2002]:

$$f(x|\mu_{dry},\sigma_{dry}) = \frac{1}{x\sqrt{2\pi\sigma_{dry}^2}}\exp\left(-\frac{(\log x-\mu_{dry})^2}{2\sigma_{dry}^2}\right). \tag{6.2}$$

For each wet period an average wind speed and direction are drawn. The distribution function of the wind speed follows a two parameter gamma distribution [*Papoulis and Unnikrishna*, 2002]:

$$f(x|\alpha_{wind},b_{wind}) = \frac{1}{b_{wind}^{\alpha_{wind}}\Gamma(\alpha_{wind})}x^{\alpha_{wind}-1}\exp\left(-\frac{x}{b_{wind}}\right) \tag{6.3}$$

and the wind direction follows the non-parametric pdf as identified from the data analysis, due to the fact that no well known parametric pdf fits well the bounded [0-360] distribution of the direction. Such a choice is considered generally very restrictive, especially for unbounded distributions, since no larger (or lower) values from the observed ones can be sampled. In the case of the wind direction though this is not a problem since it is obviously strictly bounded from both sides.

The selection of the probability distributions has no direct physical meaning and is simply based on a good fit to the data of this study. Several different distributions can be potentially used for various areas with different climatic characteristics.

6.2.3 Temporal Evolution of Mean Areal Statistics

In the second stage of the model, the simulation of the mean areal statistics of the field is performed. The two statistics that are required here are the fraction of the wet area and the mean areal precipitation. The two statistics are abbreviated hereafter as WAR and IMF following the notation of *Pegram and Clothier* [2001a, b]. The two series are auto/cross correlated and thus have to be simulated as a bi-variate stochastic process.

For each storm let us define the bivariate process $\text{WAR}_g \sim N(\mu_W,\sigma_W)$ and $\text{IMF}_g \sim N(\mu_I,\sigma_I)$ as Gaussian stochastic processes with covariance function belonging to the Whittle-Matérn class [e.g. *Gneiting*, 2010; *Hristopulos and Elogne*, 2009; *Storvik et al.*, 2002; *Christakos*, 1987; *Guttorp and Gneiting*, 2006], one of the most widely used parametric forms in geostatistics due to its generality and flexibility.

The auto-covariance function of each process is defined as

$$R_W(h) = \sigma_W^2 M(h|\nu_W, \alpha_W), \tag{6.4}$$

$$R_I(h) = \sigma_I^2 M(h|\nu_I, \alpha_I) \tag{6.5}$$

and the cross-covariance

$$R_{WI}(h) = R_{IW}(h) = \rho_{WI}\sigma_W\sigma_I M(h|\nu_{WI}, \alpha_{WI}), \tag{6.6}$$

where the subscripts W and I refer to the WAR_g and IMF_g processes respectively and h is the distance or time for temporal stochastic processes.

$M(h|\nu, \alpha)$ is defined as

$$M(h|\nu, \alpha) = \frac{2^{1-\nu}}{\Gamma(\nu)}(\alpha|h|)^\nu K_\nu(\alpha|h|), \tag{6.7}$$

where $K_\nu(\alpha|h|)$ is the modified Bessel function of the second kind which is defined as the solution of the modified Bessel differential equation

$$x^2\frac{d^2y}{dx^2} + x\frac{dy}{dx} - (x^2 + \nu^2)y = 0. \tag{6.8}$$

The class of the Whittle-Matérn covariance is a generalization of the widely used exponential model. *Gneiting* [2010] recently provided the restrictions for the bi-variate Whittle-Matérn covariance function to be positive definite. The model is valid if and only if

$$\begin{aligned}\rho_{WI} \quad \leq \quad & \frac{\Gamma(\nu_W + d/2)}{\Gamma(\nu_I)}\frac{\Gamma(\nu_I + d/2)}{\Gamma(\nu_W)}\frac{\Gamma(\nu_{WI})^2}{\Gamma(\nu_{WI} + d/2)^2}\frac{\alpha_W^{2\nu_W}\alpha_I^{2\nu_I}}{\alpha_{WI}^{4\nu_{WI}}} \\ & \times \inf_{t\geq 0}\frac{(\alpha_{WI}^2 + t^2)^{2\nu_{WI}+d}}{(\alpha_W^2 + t^2)^{\nu_W+(d/2)}(\alpha_I^2 + t^2)^{\nu_I+(d/2)}},\end{aligned} \tag{6.9}$$

where $d = 1$ for the case of one dimensional processes (e.g. time-series) and $\Gamma(.)$ is the gamma function. The bivariate process is stationary and thus the covariances are only dependent on the distance h.

The parameters $\mu_w, \mu_I, \sigma_W, \sigma_I, T_w$ where T_w is the storm length are interconnected. *Pegram and Clothier* [2001b] using a different covariance function (arising from a multivariate autoregressive model of 5th order) connected the above parameters by identifying linear relationships between all of them. Here a more general approximation of the relationships is sought by modelling the quantiles of the five parameters with a multidimensional copula. The marginals of $\mu_w, \mu_I, \sigma_W, \sigma_I$ are Gaussian and T_w follows a generalized Pareto distribution as previously explained.

A copula function for n random variables $X_1, X_2, \ldots, X_n$ can be defined as

$$C(u_1, u_2, \ldots, u_n) = P[U_1 \leq u_1, U_2 \leq u_2, \ldots, U_n \leq u_n], \tag{6.10}$$

where $U = F(X)$ is the cumulative distribution function. Here the parametric form of the multivariate T copula was chosen [*Demarta and McNeil*, 2007]:

$$C(\mathbf{u}) = \int_{-\infty}^{t_v^{-1}(u_1)} \cdots \int_{-\infty}^{t_v^{-1}(u_n)} \frac{\Gamma(\frac{\nu_c+n}{2})}{\Gamma(\nu/2)\sqrt{((\pi\nu_c)^n|P_c|)}} \left(1 + \frac{\mathbf{x'}P_c^{-1}\mathbf{x}}{\nu_c}\right)^{-\frac{\nu_c+n}{2}} d\mathbf{x}, \tag{6.11}$$

where $t_v^{-1}(x)$ is the quantile function of a standard univariate t_v t distribution. The choice was only based on the good fit to the data used in this study.

Finally WAR_g(t) and IMF_g(t) are transformed to WAR(t) and IMF(t) respectively according to a distribution anamorphosis scheme [*Schleiss et al.*, 2012] which is defined as

$$\text{WAR}(t) = F^{-1}[U(\text{WAR}_g(t))], \tag{6.12}$$

$$\text{IMF}(t) = F^{-1}[U(\text{IMF}_g(t))], \tag{6.13}$$

where U is the quantile function and F^{-1} is the inverse cumulative probability distribution function.

Finally the marginal distributions of WAR and IMF have to be chosen. WAR follows a beta distribution [*Papoulis and Unnikrishna*, 2002]

$$f(x|a_\beta, \beta_\beta) = \frac{x^{\alpha_\beta - 1}(1-x)^{\beta_\beta - 1}}{B(\alpha_\beta, \beta_\beta)} \tag{6.14}$$

and IMF a two parameter gamma distribution

$$f(x|\alpha_\gamma, b_\gamma) = \frac{1}{b_\gamma^{\alpha_\gamma}\Gamma(\alpha_\gamma)} x^{\alpha_\gamma - 1} \exp\left(-\frac{x}{b_\gamma}\right) \tag{6.15}$$

The parametric forms adopted both for the copula and the distribution of WAR(t) and IMF(t) were subjectively chosen due to the good fit they provided for the data in this study. There is no other physical consideration behind those choices.

6.2.4 Spatio-temporal Evolution of the Storm Structure

From the previous two simulation steps, for every temporal step, the fraction of the wet area and its respective intensity is known. In the third stage of the model this information is transformed into the space-time evolution of the precipitation fields.

Spatial Precipitation Fields

Precipitation fields in space are modelled as latent Gaussian fields [e.g. *Kleiber et al.*, 2012; *De Oliveira*, 2004; *Durbán and Glasbey*, 2001]. Let $G(x, y)$ be an isotropic, stationary two dimensional normally distributed ($\sim N(0, 1)$) random field with spatial autocorrelation function $\rho_g[(x, y), (x + s_x, y + s_y)]$. Since $G(x, y)$ is assumed to be wide sense stationary and isotropic, then the autocorrelation depends only on the distance $s = \sqrt{s_x^2 + s_y^2}$. The simplest 1 parameter exponential autocorrelation function is adopted here similarly to *Bell* [1987]. Its parametric form is:

$$\rho_g(s) = \exp\left(-\frac{s}{\alpha_g}\right). \tag{6.16}$$

Different assumptions concerning the spatial auto-covariance function have been adopted for spatial precipitation in previous studies. For example *Pegram and Clothier* [2001b] adopted a non-stationary two dimensional fractional Brownian motion (FBM) or [*Mandapaka et al.*, 2009] adopted a two-parameter exponential autocorrelation function. Here for the sake of parsimony the simplest possible form is used (Equation 6.16).

Then the intermittent precipitation fields can be expressed as:

$$R(x,y,t) = \begin{cases} 0 & \text{,for } U[G(x,y,t)] < 1 - \text{WAR}(t) \\ LN^{-1}\left(\dfrac{U[G(x,y,t)] - 1 + \text{WAR}(t)}{\text{WAR}(t)}, \mu_r, \sigma_r\right) & \text{,for } U[G(x,y,t)] \geq 1 - \text{WAR}(t) \end{cases}, \tag{6.17}$$

where LN^{-1} is the inverse cumulative distribution function of the lognormal distribution defined as:

$$LN^{-1}(x) = \frac{1}{\sigma_r\sqrt{2\pi}} \int_0^x \frac{\exp\left(\dfrac{-(\log t - \mu_r)^2}{2\sigma_r^2}\right)}{t} dt. \tag{6.18}$$

The parameters μ_r, σ_r can be expressed in terms of the coefficient of variation cv_r as

$$\mu_r = \log\left(\frac{\text{IMF}(t)}{\text{WAR}(t)\sqrt{\text{cv}_r^2 + 1}}\right) \tag{6.19}$$

and

$$\sigma_r = \sqrt{\log(\text{cv}_r^2 + 1)}, \tag{6.20}$$

or in terms of their L-coefficent of variation Lcv_r derived from the L-moments [*Hosking*, 1990], taking into account that for the lognormal distribution

$$\mathrm{Lcv}_r = \frac{\lambda_2}{\lambda_1} = \mathrm{erf}\left(\frac{\sqrt{\log(1+\mathrm{cv}_r^2)}}{2}\right). \tag{6.21}$$

The auto-covariance function of the lognormally distributed field for the case of $\mathrm{WAR}(t) = 1$ is as defined by [*Bell*, 1987; *Allard*, 2012; *Mejia and Rodriguez-Iturbe*, 1974]

$$R_r(s) = \mathrm{cv}_r * \mathrm{IMF}(t)\frac{\exp[\sigma_r^2 \exp(-s/\alpha_g)] - 1}{\exp(\sigma_r^2) - 1}. \tag{6.22}$$

The case of $\mathrm{WAR}(t) < 1$ is not trivial and an analytical solution does not exist. However a numerical estimation based on the Hermit polynomials of the transform function of the Gaussian field can be found [e.g. *Jeannin et al.*, 2012; *Guillot*, 1999]. A similar procedure was used in section 4.3.3 for the simulation of time series.

Temporal Evolution of the Precipitation Fields

Precipitation fields, in order to be realistic should also be correlated in time. This correlation is introduced in the STREAP model by assuming the following autoregressive moving average ARMA(p,q) representation for the Gaussian fields.

$$G(x,y,t) = \sum_{i=1}^{p} \phi_i G(x,y,t-i) + \sum_{j=1}^{q} \theta_j \epsilon(x,y,t-j) + \epsilon(x,y,t), \tag{6.23}$$

where noise term $\epsilon(x,y)$ is a Gaussian two dimensional random field with the same spatial autocorrelation as $G(x,y)$. The approach of *Pegram and Clothier* [2001b] was similar adopting an $AR(5)$ process. The variance of the noise term is estimated from the unit variance restriction of the field $G(x,y)$ [*Box and Jekins*, 1970] .

$$\sigma_\epsilon^2 = \frac{\gamma_0 - \sum_{i=1}^{p} \phi_i \gamma_i}{1 - \sum_{i=1}^{q} \psi_i \theta_i}, \tag{6.24}$$

where

$$\psi_j = \begin{cases} \sum_{i=1}^{p} \phi_i \psi_{i-1} - \theta_j & , j > 0 \\ 0 & , j = 0 \end{cases} \tag{6.25}$$

and

$$\gamma_k = \sum_{i=1}^{p} \gamma_{k-i}\phi_i + \sigma_\epsilon^2 \sum_{i=k}^{q} \theta_i \phi_{i-k}. \tag{6.26}$$

The temporal evolution of the field refers to the Lagrangian system of coordinates (i.e. following the field's motion). Here an ARMA(2,2) model was found to be sufficient, but this result cannot be generalized, and probably different orders could be eventually used for different case studies. Previously, an autoregressive model of order one AR(1), or its continuous representation, the Ornstein − Uhlenbeck process have been used [e.g. *Bell and Kundu*, 1996; *Kundu and Bell*, 2003; *Sigrist et al.*, 2012a] as well, with promising results. More *Sigrist et al.* [2012b] using an Ornstein–Uhlenbeck process, managed to express the Gaussian field evolution as the solution of a stochastic partial differential equation and derive closed form relationships for the non-separable space-time covariance function.

6.2.5 Modelling Random Fields in the Frequency Domain

One open question at this point is how to simulate Gaussian random fields with a known covariance function needed for stages 2 and 3 of the model. An exact method to simulate random fields is based on the Cholesky factorization of the covariance matrix [e.g. *Oliver*, 1995]. A major problem with this method especially when dealing with large domain simulations is its numerical complexity which is of the order $O(N^3)$. This can pose serious restrictions on the simulation time and also the computational resources that are needed.

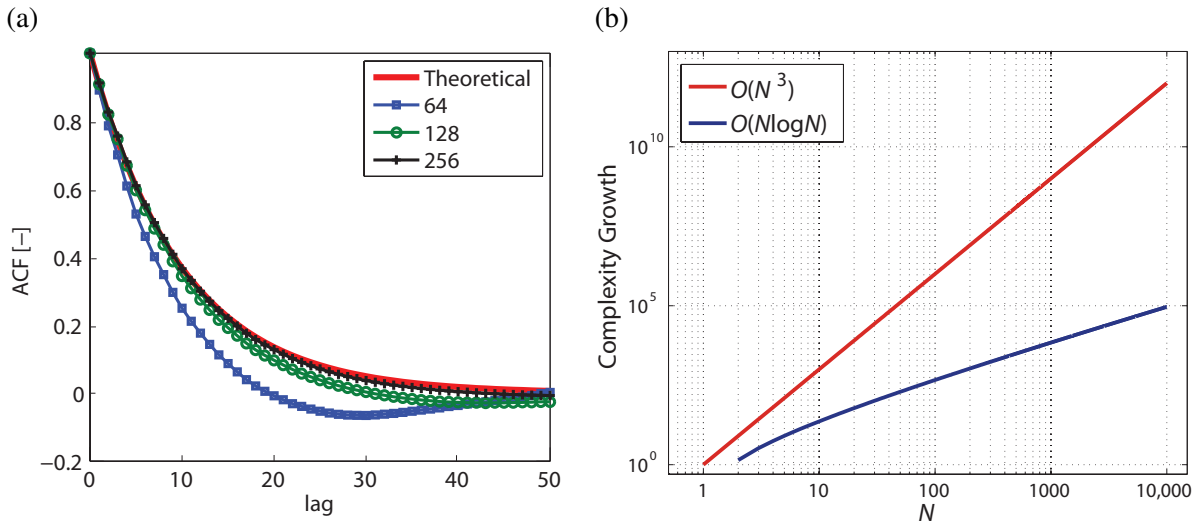


FIGURE 6.2: (a) Convergence to the theoretical auto-covariance function for 2D Gaussian random fields with correlation length $\lambda = 10$ for various sample sizes; and (b) the complexity growth comparison between the FFT method and the Cholesky factorization method

A significant reduction of the simulation time for random fields on a regular grid can be achieved exploiting the numerical efficiency of the Fast Fourier transform which is of the order $O(N \log N)$ [e.g. *Dietrich and Newsam*, 1993; *Ravalec et al.*, 2000; *Pardo-Iguzquiza and Chica-Olmo*, 1993] (see Figure 6.2(b)).

The main problem of those methods is that the covariance function of the simulated field converges only to the theoretical one for large sample sizes (see figure 6.2(a)). Another widely known method exploiting the spectral properties of the random field is the turning bands method *Mantoglou and Wilson* [1982]. Recently fast and exact methods for random field generation were introduced [*Dietrich and Newsam*, 1993; *Helgason et al.*, 2011] using the concept of the circulant embedding of the covariance matrix. In this thesis the simulation method based on the FFT is implemented. Details on the numerical implementation of the method can be found in *Chambers* [1995]; *Pardo-Iguzquiza and Chica-Olmo* [1993]; *Lang and Potthoff* [2011]. Here only the main results are reported.

Stage 2: Multivariate Series

The simulation of the second stage of the model requires a modelling scheme for simulating bi-variate time series. The one that is used here was developed by *Chambers* [1995]. Let $\boldsymbol{X}_i(t)$ be a q-variate time-series, with spectral density $\boldsymbol{S}(k)$, following a Gaussian distribution with $t = 0, 1 \ldots n$. Since the process $\boldsymbol{X}_i(t)$ is real valued, the spectral density is Hermitian. The eigen-decomposition of the spectral density matrix gives

$$\boldsymbol{S}(k) = \boldsymbol{U}(k)\boldsymbol{M}(k)\boldsymbol{U}^*(k). \tag{6.27}$$

Then the random variable $\boldsymbol{V}$ defined as

$$\begin{aligned} \boldsymbol{V}_k &= \boldsymbol{U}_k \boldsymbol{M}_k^{1/2} \boldsymbol{Z}_k \qquad , k =, \ldots, n/2, n \\ \boldsymbol{V}_{n-k} &= \overline{\boldsymbol{U}_k} \boldsymbol{M}_k^{1/2} \overline{\boldsymbol{Z}_k} \quad 1 \leq k \leq n/2 - 1 \end{aligned} \quad , \tag{6.28}$$

can be the Fourier coefficients for a multivariate series with spectrum $\boldsymbol{S}(k)$. The variable $\boldsymbol{Z}_k$ is complex and can be decomposed as

$$\boldsymbol{Z}_k = \text{Re}(\boldsymbol{Z}_k) + i\text{Im}(\boldsymbol{Z}_k), \tag{6.29}$$

with

$$\begin{bmatrix} \text{Re}(\boldsymbol{Z}_k) \\ \text{Im}(\boldsymbol{Z}_k) \end{bmatrix} \sim N \left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix} \right] \tag{6.30}$$

and

$$\boldsymbol{Z}_{n-k} = \boldsymbol{Z}_k, \quad \text{for} \quad 1 \leq k \leq n/2 - 1. \tag{6.31}$$

Over-bars denote the complex conjugate. I is the $q \times q$ identity matrix and $\boldsymbol{M}(k)$ is the $q \times q$ diagonal matrix with the real valued eigenvalues of $\boldsymbol{S}(k)$ and $\boldsymbol{U}(k)$ is the $q \times q$ matrix of the associate eigenvectors. The operator $[.]^*$ denotes complex conjugation and matrix transposition. $\boldsymbol{S}(k)$ is the matrix the

components of which are the spectral and cross-spectral densities. The simulated Gaussian series vector arises from the inverse Fourier transform of V_k.

So in order to simulate the bi-variate stochastic process for the second stage in the frequency domain, the estimation of the spectral densities is needed for the Whittle-Matérn covariance function. The spectral densities are the Fourier transforms of the covariance functions and thus:

$$S(k) = \mathcal{F}[cM(s|\nu,\alpha)] = c\frac{\Gamma(\nu+\frac{1}{2})\alpha^{2\nu}}{\Gamma(\nu)\pi^{\frac{1}{2}}}\frac{1}{(\alpha^2+k^2)^{\nu+\frac{1}{2}}}, \tag{6.32}$$

where $c \in \Re$ and $\mathcal{F}$ is the one dimensional Fourier transform.

The extension to multivariate fields of any arbitrary dimension can be found in *Ruan and McLaughlin* [1998].

Stage 3: Multidimensional Fields

For the simulation of the third stage, a method for simulating multidimensional, and in this case 2D, Gaussian random fields, is needed. A complete description can be found in [*Lang and Potthoff*, 2011], with the following main elements. Define G as:

$$G(x) = (\mathcal{F}^{-1}S^{1/2}\mathcal{F}Z)(x). \quad x \in \Re^d, \tag{6.33}$$

where S is the d-dimensional power spectrum, Z is a d-dimensional Gaussian noise and $\mathcal{F}^{-1}$ is the inverse Fourier transform.

So in order to simulate Gaussian random fields with an exponential decay of the autocorrelation, the 2D spectrum has to be calculated.

$$S(\mathbf{k}) = \mathcal{F}\left[\sigma^2 \exp\left(-\frac{|\mathbf{x}|}{\alpha_g}\right)\right] \quad , \mathbf{k}, \mathbf{x} \in \Re^2, \tag{6.34}$$

since the autocovariance function is symmetric the two dimensional Fourier transform of equation 6.34 can be reduced to the one dimensional Hankel transform [*Bracewell*, 2000, pp 336].

$$S(|k|) = 2\pi \int_0^\infty \sigma^2 \exp\left(-\frac{s}{\alpha_g}\right) J_0(2\pi|k|s)s \quad ds = \sigma^2 \frac{2\pi\frac{1}{\alpha_g}}{\left(4\pi^2|k|^2 + (\frac{1}{\alpha_g})^2\right)^{\frac{3}{2}}}, \tag{6.35}$$

where $|k| = \sqrt{k_x^2 + k_y^2}$ and J_0 is the Bessel function of the first kind of zero order.

Another interesting feature is that the ARMA representation of the Gaussian field as defined in equation 6.23 can be expressed with an equivalent ARMA representation of the Fourier coefficients $W_G = \mathcal{F}Z(x)$

as defined in equation 6.33.

$$W_G(k_x, k_y, t) = \sum_{i=1}^{p} \phi_i W_G(k_x, k_y, t-i) + \sum_{j=1}^{q} \theta_j \epsilon_W(k_x, k_y, t-j) + \epsilon_W(k_x, k_y, t), \tag{6.36}$$

where $\epsilon_W(k_x, k_y, t)$ is an uncorrelated hermitian white noise Gaussian complex random field. This can be easily shown by multiplying Equation 6.36 element by element with the square root of the power spectral density $S(k_x, k_y)^{1/2}$ and taking the 2D inverse Fourier transforms taking into account equation 6.33. This relationship gives the direct way of simulation the space-time Gaussian fields in Lagrangian coordinates.

Kundu and Bell [2003] and *Bell* [1987] stated that low frequency components should have higher correlation than the high frequency ones. Conceptually this means that large scale features such as e.g. fronts are much more correlated that small scale fluctuations e.g. rain-cells. Even though this feature would increase the model's realism, parameter estimation for the cases examined here was very difficult and thus not applicable.

The generated fields, due to the symmetries of the fast Fourier transform, can be folded (Fourier symmetry around the Nyquist frequency). This property can be exploited in order to simulate advection of the fields and results in the Eulerian coordinate system [*Pegram and Clothier*, 2001a] (see figure 6.3). Moreover in order to avoid any artefacts due to the field symmetries, the precipitation fields are simulated for a larger spatial extent (here 2 times more), and then only the central area of the simulations is used as the model output. The idea of simulating advection with this manner was used as well at the SBM model [*Pegram and Clothier*, 2001b].

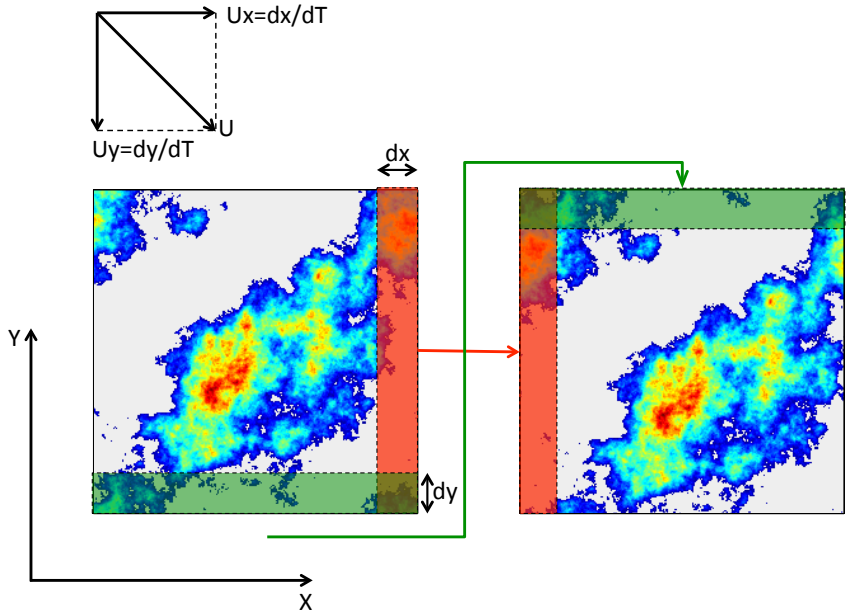


FIGURE 6.3: Schematic representation of the advection simulation using the symmetries of the FFT.

6.2.6 Model Calibration

In the previous section the description of each process was mathematically formulated. Here the model calibration procedure is given. One of the main disadvantages of such a complex model is that an analytical estimation of the statistical properties of the simulated fields can be mathematically untraceable. For example one of the major attractive features of the Poisson cluster models [e.g. *Burton et al.*, 2008; *Cowpertwait*, 2010; *Burton et al.*, 2010b] is that the relationships of their moments and covariance functions can be estimated and the numerical minimization of an objective function can yield a direct and straightforward parameter estimation. On the contrary, for the model developed here, each process has to be calibrated separately. In order to capture the seasonality of precipitation, most of the processes are calibrated on a seasonal basis and some are calibrated on a monthly basis.

Storm Arrival Process Calibration

This stage is the most straightforward to calibrate. From the data records, the durations of the dry and wet periods are extracted and the estimation of the distribution parameters can be achieved with several fitting procedures. The most popular among the distribution fitting procedures are the method of moments, L-moments, probability weighted moments and maximization of the likelihood. Here the latter one is adopted. Although the fitting procedure is clear, there is no commonly accepted way to separate precipitation events [e.g. *Veneziano and Lepore*, 2012; *Koutsoyiannis and Foufoula-Georgiou*, 1993; *Lakshmanan et al.*, 2003]. The method that is proposed here is simple and effective and has the following steps. First from the radar data the sequence of WAR(t) is extracted for the studied area. Then in order to avoid spurious behaviour and to smooth the signal, a moving average filter with a 1 hour extent is applied to the sequence. Finally the precipitation event is defined as the time when the filtered sequence of WAR(t) exceeds a low (2%) threshold. If the time between two storms is less than 2 hours then those events are considered as one. An illustration of the storm identification procedure is shown in figure 6.4, where a sample time-series of WAR and their identified storms are shown.

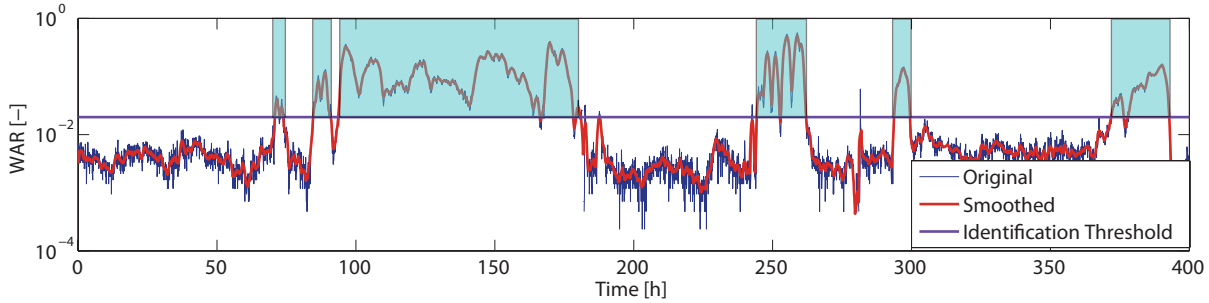


FIGURE 6.4: Illustration of the storm identification procedure. A sample time series of the WAR is shown, and the identified events are shown as cyan shaded areas.

WAR-IMF Calibration

The calibration of the WAR-IMF process has several components. First the parameters of the Whittle Matern function have to be estimated. The estimated sequences of war(t)-imf(t) [1] derived from the radar

[1] Random variables are noted as capital letters and their realisations with case letters.

data set were normalized. The normalization procedure can be formalized as

$$\text{war}_g(t) = N^{-1}(U(\text{war}(t)), 0, 1) \tag{6.37}$$
$$\text{imf}_g(t) = N^{-1}(U(\text{imf}(t)), 0, 1) \tag{6.38}$$

Then the sample auto and cross correlations were estimated, and finally the parameter estimation was done by constrained least square fitting of the theoretical parametric functions. Even though this procedure is sub-optimal [*Gneiting*, 2010] and maximum likelihood estimates exist, for the case studies analysed here it yielded reasonable results.

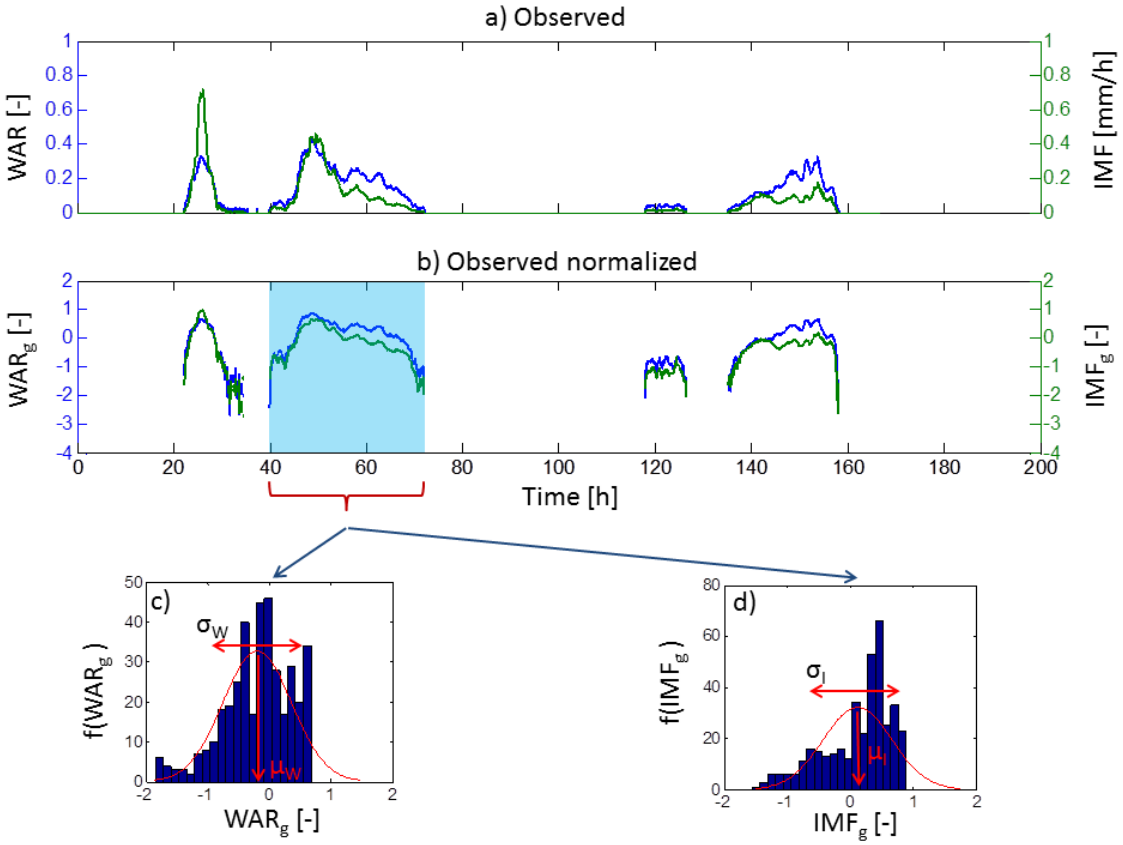


FIGURE 6.5: Illustration of the stage 2 calibration procedure.(a) The observed series of WAR and IMF. (b) The normalized series of WAR and IMF and (c-d) their corresponding probability density functions with a Gaussian assumption.

The estimation of the $\mu_W, \mu_I, \sigma_W, \sigma_I$ was done from the normalized $\text{war}_g(t), \text{imf}_g(t)$ as described in *Pegram and Clothier* [2001a] . An illustration of the procedure can be found in figure 6.5. The calibration of the T-copula was conducted using approximate maximum likelihood estimates and the fitting of the beta and gamma distributions for WAR and IMF respectively was also done by maximizing their likelihood functions.

Spatio-temporal Storm Evolution Calibration

The parameter estimation for the third stage of the model is not that straight-forward as for the previous two. The parameters that have to be estimated are the spatial correlation length, the coefficients of variation and the parameters of the ARMA process.

First as previously shown, a direct estimation of the spatial autocorrelation function can be highly problematic due to noise corruption and clutter contamination of the radar precipitation estimates. For this reason the parameter estimation of the autocorrelation function is conducted here only on the binary [0-1] rain-no rain process, which is expected to be well represented by the radar data and relatively free of errors.

Let $C_b(s)$ be the spatial autocovariance of the binary rain - no rain process (above a threshold h) of a latent Gaussian field with autocovariance $C_g(s)$. *Gutnisky and Josić* [2010] define that

$$C_b(s) = p(s) - r^2, \tag{6.39}$$

where

$$p(s) = \int_h^\infty \int_h^\infty \frac{1}{2\pi\sqrt{1 - C_g(s)^2}} \exp\left\{-\frac{a^2 + b^2 - 2C_g(s)ab}{2[1 - C_g(s)^s]}\right\} \mathrm{d}a\mathrm{d}b \tag{6.40}$$

and

$$r = \frac{1}{\sqrt{2\pi C_g(0)}} \int_h^\infty \exp\left(-\frac{y^2}{2C_g(0)}\right) \mathrm{d}y. \tag{6.41}$$

This means that the estimation of $C_g(s)$ is avoiding any possible problem of the measurement errors. α_g is the estimated with ordinary lest square fitting of the theoretical spatial covariance function. The estimation of $C_g(s)$ from $C_b(s)$ is achieved by a numerical solution of equation 6.40.

The most difficult and uncertain task, vulnerable to errors is the estimation of the parameters of the ARMA model for the temporal evolution of the precipitation fields. In their first efforts *Pegram and Clothier* [2001a] for the SBM model did not provide a calibration algorithm for this stage, but they rather used a "trial and error" procedure until they found a parameter set suitable for their data. Here the approximation that is adopted is to estimate the average autocorrelation function of the normalized precipitation fields in time in Lagrangian coordinates and use a constrained least square fitting procedure of the theoretical autocorrelation function of the ARMA process. The constraint arises from the stability criteria of the ARMA processes defined in *Box and Jekins* [1970].

The major difficulty is to identify the Lagrangian coordinate system from the data. The problem arises from the fact that observed precipitation fields are only approximately advected with a constant velocity but are rather prone to rotation and differential velocities (see 6.6(a) for an example). Here in order to identify storms that are moving approximately constantly a storm tracking algorithm is used. The storm tracking procedure adopted here is the COTREC algorithm *Li et al.* [1995] which is an improvement of the

TREC algorithm [*Rinehart and Garvey*, 1978]. The method is based on correlation matching imposing a mass conservation (continuity) restriction. Using such a tracking procedure the two dimensional velocity field can be estimated. Details concerning the numerical implementation of the method can be found in [*Li et al.*, 1995]. The storms that are selected to estimate the temporal autocorrelation function should fulfil some criteria. First in order to exclude events with strong differential movement, a threshold selection was implemented on the mean value of the standard deviation of the spatial velocity vectors estimated using COTREC. A similar criterion was applied on the standard deviation of the advection directions. Finally only storms that exceed a specific duration were selected. This procedure limits the sample size of storms that can be used for the estimation of the temporal autocorrelation of the Gaussian fields. Also, the issue of long lasting storms impose a serious restriction since generally only stratiform events could be taken into account, due to the fact that summer convective storms have very short duration.

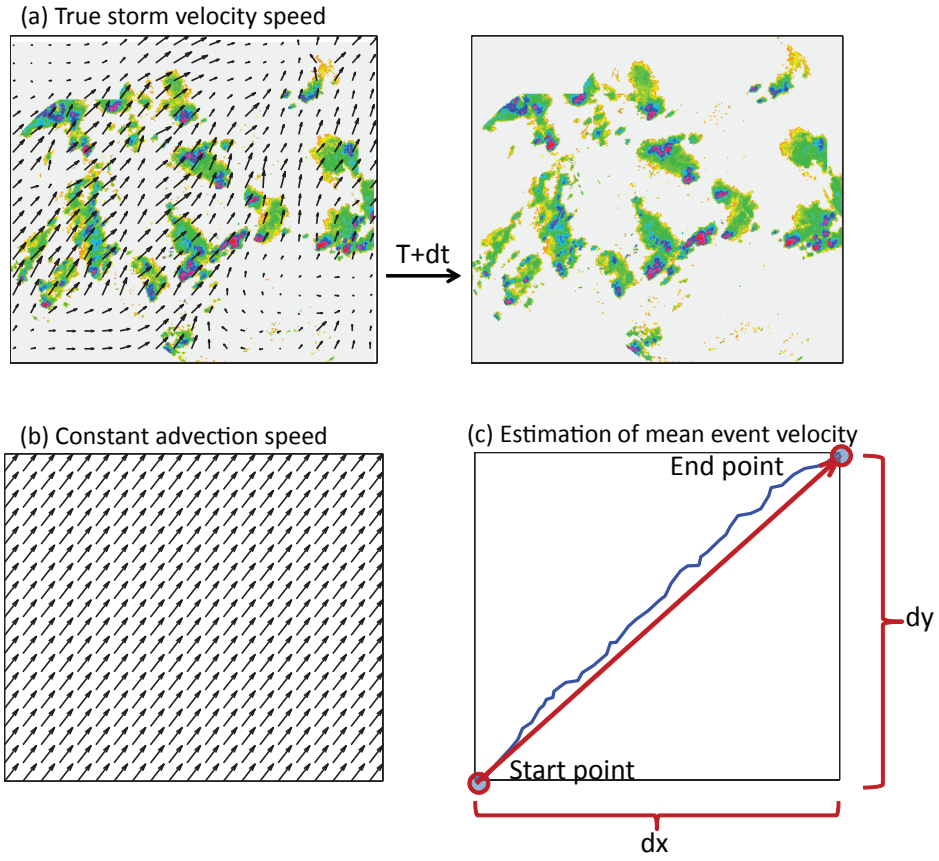


FIGURE 6.6: Schematic representation of the constant storm advection approximation and its estimation procedure. (a) The velocity vectors of the plotted radar sequences according to the COTREC algorithm. (b) The spatial uniformity assumption of the storm velocities. (c) The estimation of the mean storm velocity of an event.

Another difficulty that arises at this stage from the fact that precipitation data are quantized into coarse and not uniformly distributed intensity classes. The sample cross-correlation between consecutive images in the Lagrangian system of coordinates can be highly biased when only very few values of precipitation are available. For this reason another criterion that was posed on the storm selection was that the wet area ratio (WAR) should exceed a specific threshold, in order to decrease this bias.

In order to avoid model over-parametrisation, α_g, cv_r depend only on the month and are defined as the

average value from the data, and the parameters of the ARMA model are constant throughout the year. This assumption restricts the model variability but any further parametrization of each of the parameters as separate stochastic processes would reduce the parsimony of the model and the calibration procedure would be even more complex.

6.3 Model Validation

In order to asses the model's ability in reproducing precipitation statistics in time and space a model validation was performed on the Monte Lema area. The model STREAP is validated for its ability to reproduce both areal and temporal statistics for various aggregation scales. The space-time scales that are analysed here correspond to the most common ones that are of hydrological interest especially in catchment hydrology.

6.3.1 Study Area

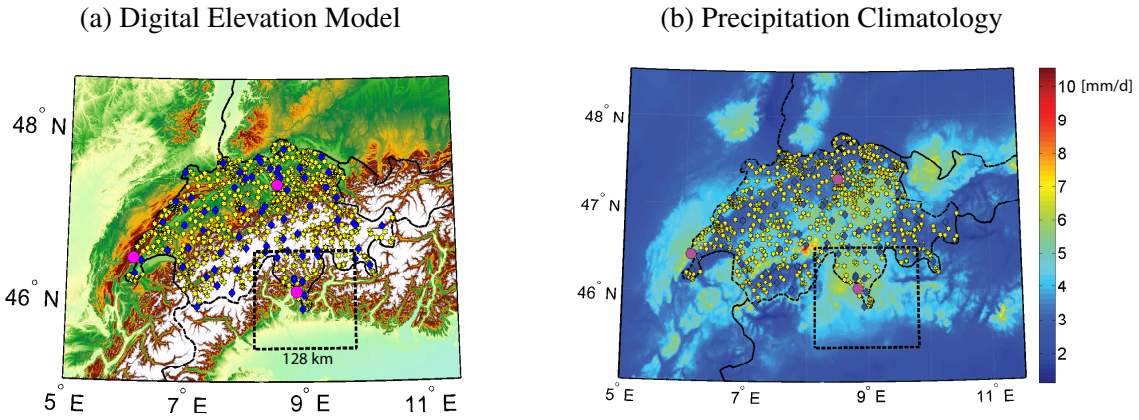


FIGURE 6.7: Study area of the application of the STREAP model. On the left (a) the digital elevation model of the area is shown and on the right (b) the mean daily accumulation depths in [mm] after *Frei and Schär* [1998].

The study area that was selected is shown in figure 6.7. First of all, as already mentioned, the radar visibility in this area is considered optimal, and precipitation estimates are very reliable [*Germann et al.*, 2006; *Foresti et al.*, 2012]. Moreover, since the STREAP model is homogeneous in space, this area is very convenient due to the fact that precipitation accumulations appear to be only slightly variable in space, with some orographic effects in the Northern part, and also the probabilities of precipitation concurrence are with a good approximation homogeneous.

In addition, as previously found (chapter 5) the spatio-temporal structure of precipitation in this area is of particular interest [e.g. *Germann et al.*, 2006; *Germann and Joss*, 2001]. This specific area was the case study of the MAP (**M**esoscale **A**lpine **P**rogramme) experiment [*Houze et al.*, 2001; *Bougeault et al.*, 2001] which was a collaborative research program that analysed the mesoscale precipitation structure in the Alpine area. The main climatological features of this area is a strong seasonality as shown in *Frei and Schär* [1998], with intense convective events during summer and weak and less frequent stratiform events

during winter and various patterns of orographic precipitation enhancement dependent on the wind speed and the wind blockage by the mountain range.

6.3.2 Areal Statistics

First of all, the model's capability in reproducing the mean areal storm characteristics is investigated. The areal statistics that are explored are:

1. The probability distribution of the fraction of wet areas
2. The probability distribution of the mean areal intensities
3. The auto/cross correlation of WAR, IMF
4. The probability distribution of the duration of dry spells
5. The probability distribution of the duration of wet spells

The temporal scales for which the model efficiency is evaluated are 10 min, 1 hour and 1 day. This selection reflects the major temporal scales of interest, both for urban and catchment hydrology. Even though the model calibration is performed on a seasonal basis, only the overall pattern throughout the year is presented here. It should be mentioned though that the model efficiency is comparable for all the seasons and thus the overall annual results are comparable with the efficiency on the seasonal basis.

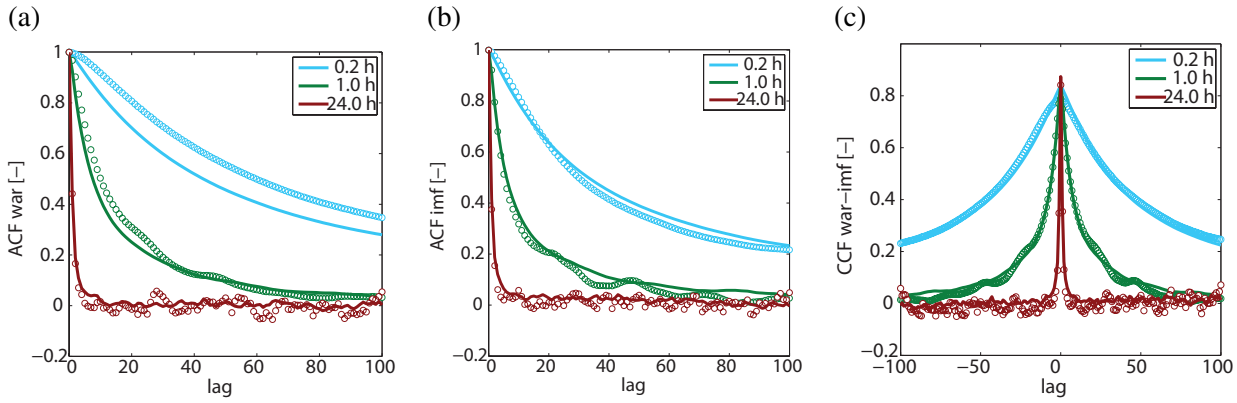


FIGURE 6.8: (a) The autocorrelation of the WAR process. (b) The autocorrelation of the IMF process. (c) Cross-correlation between WAR and IMF. Different colours correspond to various aggregation intervals as shown in the corresponding legends. Dots correspond to observations and lines to the mean value of an ensemble of simulated series.

In figure 6.8 it is clearly shown that the model is very good at reproducing both autocorrelations and crosscorrelations of the WAR and IMF processes. The only problem arises for the autocorrelation of the WAR process for the high frequency scale. The model fails to simulate realistically the start and end of the storm periods leads to abrupt jumps from positive values of WAR to zero and that has an impact to the autocorrelation function of the WAR. The problem is reduced for higher aggregation intervals and even on an hourly basis is not any more present. A solution to this problem could be that the stage 2 of the simulation of the WAR-IMF process per storms could be replaced with a conditional sampling procedure, fixing a priori the values of WAR-IMF at the start and end points of each storm.

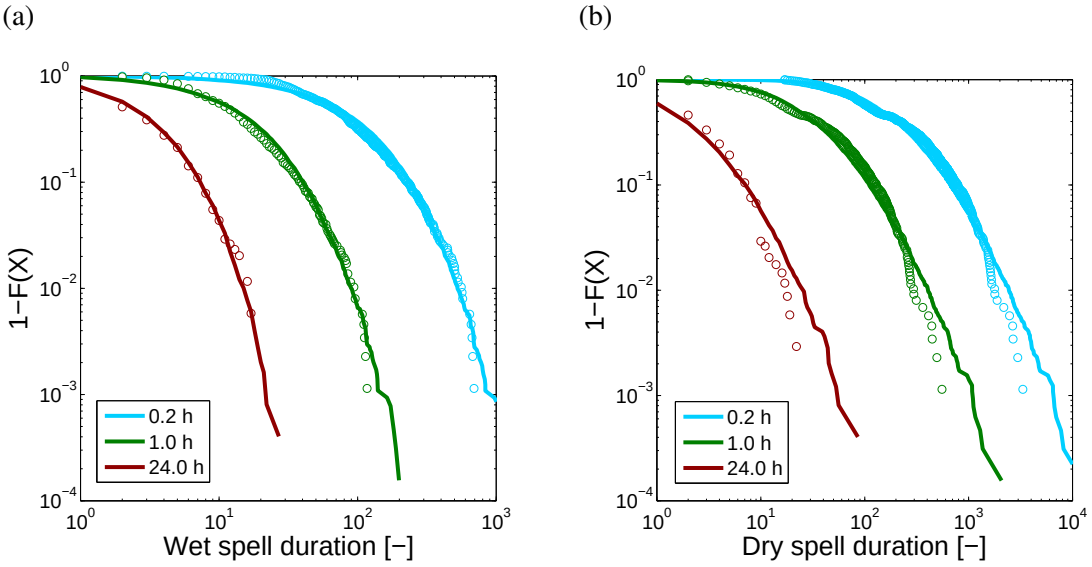


FIGURE 6.9: Exceedance probabilities of the dry and wet spells derived from the areal averaged process. (a) The duration of the wet spells for various aggregation intervals expressed in number of consecutive steps for each corresponding aggregation interval. (b) The same for the distribution of the dry durations.

The distributions of the storm lengths and their corresponding inter-arrival times is satisfactorily reproduced (figure 6.9). Only the tails of the dry spells are slightly overestimated. A way to overcome this issue would be a different choice of the probability distribution of the dry spells that would provide a better fit both for the body and the tail of the distribution. This was not further investigated here since the improvement would be marginal and it would increase the number of parameters of the model.

The probability distribution of the WAR and IMF processes are also very well reproduced as can be seen in figure 6.10. This leads to the conclusion that even a choice of a parsimoniously parametrized distribution for those processes (Beta and Gamma respectively) yields very reasonable results at least for the case study investigated here.

Finally another metric that was introduced by *Bell* [1987] which combines the effect of the simultaneous behaviour of the WAR process and the spatial autocorrelation is used. The metric corresponds to the probability distribution of the length of an arbitrary cross-section across a radar image that is covered with precipitation as shown schematically in figure 6.11. The behaviour of the model in reproducing the distribution of the length covered with precipitation is very good. This lead to the conclusion that the composite effect of the WAR process and the spatial correlation is very well represented in STREAP.

The very good results concerning the model efficiency in capturing the areal statistics of precipitation is not surprising, since it is calibrated directly on those properties. The most encouraging result is that even though the model is calibrated only for the finest temporal scales, its efficiency is very good also for coarser temporal aggregations.

6.3.3 Point Scale Statistics

For the model developed here to be adequate for applications it is crucial that it also reproduces well the point scale statistics. Unfortunately a direct comparison of the model output to the rain-gauges located the

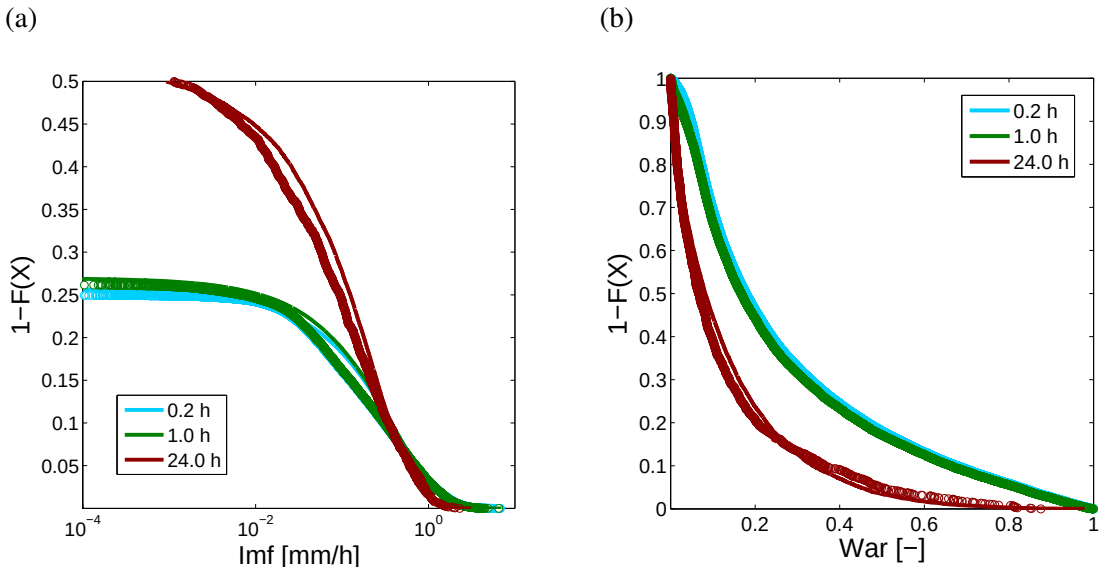


FIGURE 6.10: Survivor functions of the (a) mean areal intensities and (b) fraction of wet area for different aggregation intervals.

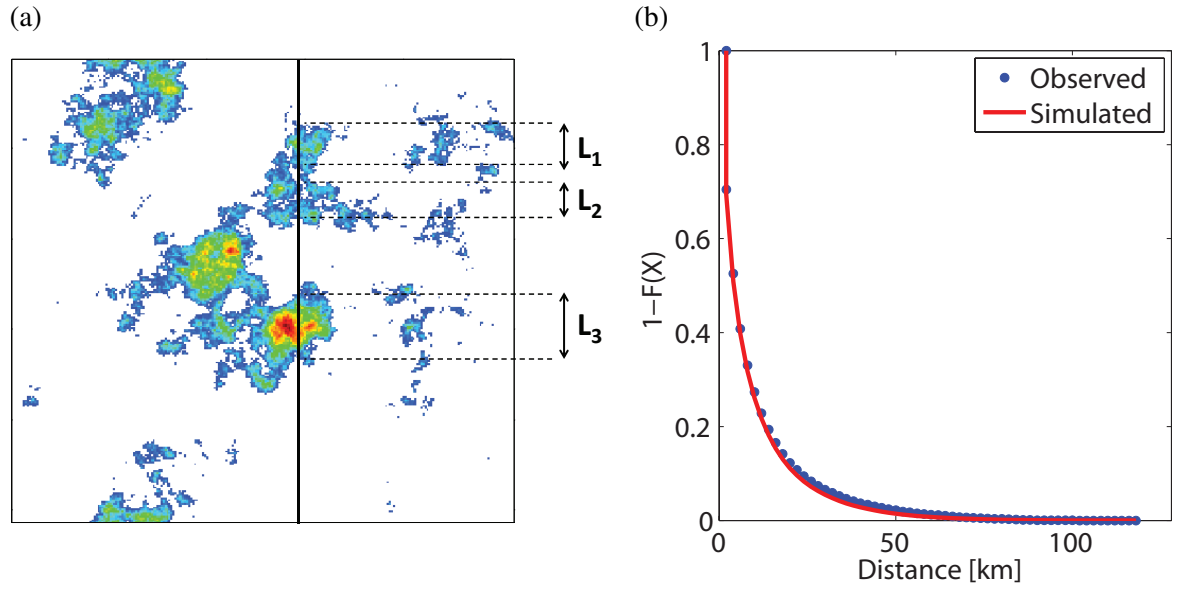


FIGURE 6.11: Probability distribution of the length of a cross-section covered with precipitation. (a) On the left a schematic of the calculation is shown; and (b) on the right the comparison between the exceedance probabilities of the observed and simulated values are given.

area studies here is not possible due to the large discrepancies between ground rain-gauge measurements and radar precipitation estimates. Also, the radar measurements suffer from the inherent disadvantage of the measuring technique which is the signal attenuation of the transmitted electromagnetic waves that leads to an underestimation of precipitation intensities with distance (see figure 2.2). This causes the mean spatial average of precipitation to be non-homogeneous in space, and since the model generates homogeneous random fields, any pixel (radar) to pixel (simulation) comparison would yield misleading results. In order to overcome this problem, the following procedure is followed. As data point-scale time

series the pixel values of the radar data set of several uniformly distributed random points across the study domain are chosen. What is defined here as data point scale statistics corresponds to the average values of the respective statistic derived from these radar time series. On the other hand the equivalent point-scale statistics of the simulated series are the same all over the simulated domain, since the model is by construction spatially homogeneous. The ensemble of the simulated time series is derived as an extraction of the simulated pixel intensities of 20 randomly selected 20 points across the simulation domain. The total length of the simulation was 30 years.

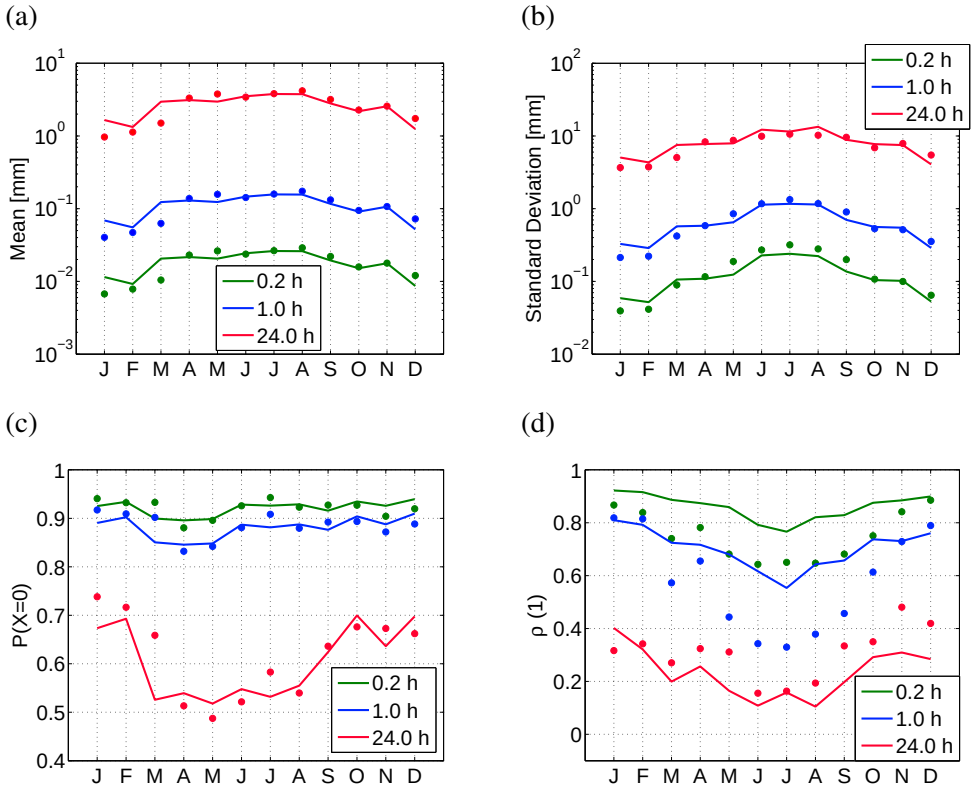


FIGURE 6.12: Validation of the STREAP model for the monthly point scale statistics. (a) Mean precipitation depth accumulation. (b) Precipitation standard deviation. (c) Probability of zero precipitation and; (d) Lag-1 correlation coefficient. Different colours correspond to different aggregation intervals. Dots stand for observed values and lined for simulated

In figure 6.12 it is shown how well some of the monthly point scale statistics are reproduced. It should be noted that the random selection of very few point on the entire domain increases the uncertainty. All the statistics except of the lag one correlation are very well reproduced. Estimation of higher order moments would be very uncertain for this small number of realizations and thus not reported. The major problem concerns the correct reproduction of the autocorrelation function that is systematically overestimated. This problem mainly arises from the problematic estimation of the parameters of the ARMA process due to the fact that most of the precipitation events on which the estimation is based are stratiform, highly correlated ones. This results in the overestimation observed here. The problem would be solved if longer radar records were available. Discrepancies in the mean intensity values on the order of 20% occur as well. The reason for this is that the IMF and WAR processes are calibrated on a seasonal basis and this reflects in slight differences on the average simulated precipitation. This issue could be solved if the calibration procedure was done on a monthly basis, but since the current study serves as a proof of concept, the

seasonal efficiency of the model is adequate. The complete reproduction of the probability distribution of precipitation depth is shown in figure 6.13 on a seasonal basis. In most of the cases, the main body of the distribution can be very well reproduced. Some problems arise for the correct simulation of the tails of the distributions. The reason for that is probably the choice of the lognormal distribution for the precipitation intensities in space and the assumption of a constant monthly value of the coefficient of variation. The overall behaviour is nevertheless satisfactory.

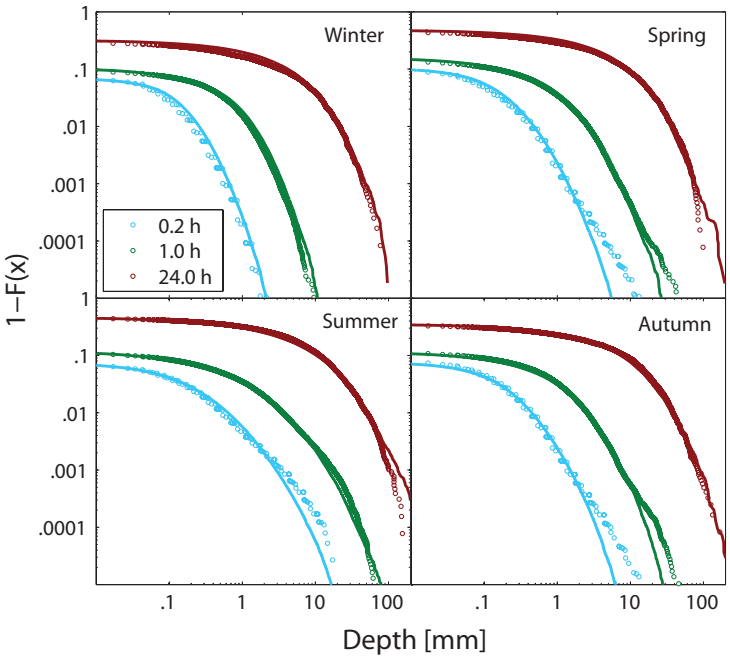


FIGURE 6.13: Exceedance probabilities for the point scale precipitation depth on a seasonal basis for three different aggregation intervals. The aggregation intervals are marked with different colours. Dots represent observed data and lines simulated series.

As already shown, the model is capable of reproducing well the storm durations and their inter-arrival times. This though does not assure that the same distributions are well reproduced at the point scale because the composite effect of the storm arrival process and the in-storm dry spells created by the WAR process should be taken into account. In figure 6.14 the comparison between observed and modelled distributions of the point scale wet durations is shown. Generally the model behaves well for large aggregation intervals (e.g. day) but there is a slight underestimation of the durations especially during winter and spring for the fine scale aggregation intervals. The same behaviour and discrepancies were also found for the distributions of the dry lengths.

Finally the point scale probability of zero precipitation is very well reproduced for a large range of temporal scales, for all the seasons as is illustrated in figure 6.15.

6.4 Model Benchmarking

In order to compare the model's ability in reproducing well both areal and point-scale statistics, here a comparison with one of the most widely used stochastic space-time models for rainfall is performed. The model is a spatiotemporal extension of the Poisson cluster model of the Neyman-Scott type NS(s-t)

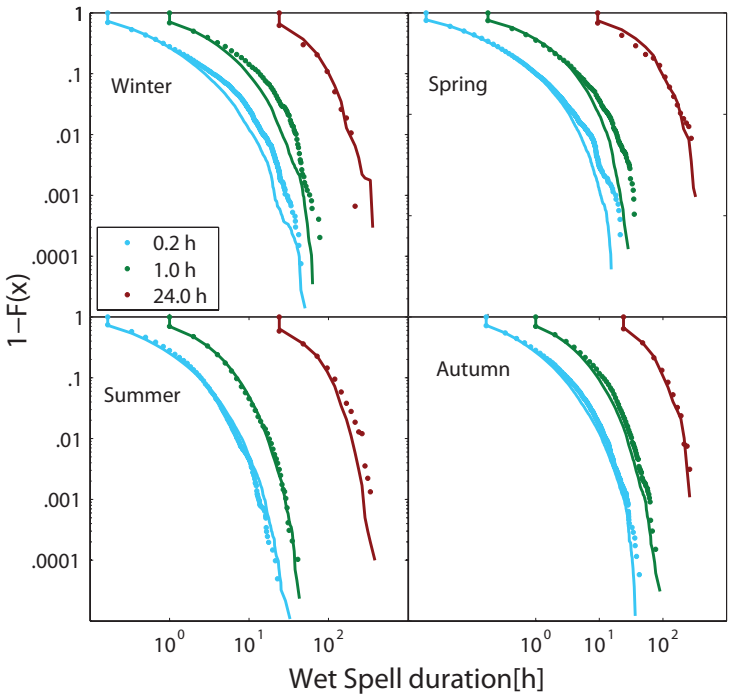


FIGURE 6.14: Exceedance probabilities for the point scale wet spell durations. Each panel corresponds to different seasons and each colour to different aggregation intervals as shown in the legends.

[*Burton et al.*, 2010b; *Cowpertwait*, 2006; *Cowpertwait et al.*, 2002]. This type of model has been found to yield very promising results and has been successfully applied to several types of climates such as oceanic [*Cowpertwait*, 2006], Mediterranean [*Cowpertwait et al.*, 2002] etc. It is true that the two models are of different complexity, but this can lead to a conclusion whether additional degrees of freedom should be added in order to reproduce well the precipitation statistics. It will be illustrated that models which explicitly describe all the main features of spatiotemporal precipitation are needed especially for high temporal resolutions (< 1 hour). Classic simplified Poisson-cluster models are very useful as multi-site stochastic simulation tools, but if the entire space-time structure of precipitation has to be taken into account they are inadequate.

6.4.1 The Space-Time Neyman-Scott Model

The model is essentially an expansion of its temporal formulation as described in 4.3.1. The formulation that is adopted here is the one developed by *Cowpertwait et al.* [2002]. The model structure is as follows (see figure 6.16):

1. Storms arrive as a Poisson process in time with rate λ [h^{-1}].
2. Each storm origin is followed by a generation of $N(N \geq 0)$ raincells that form a a two dimensional homogeneous Poisson process with density ϕ_c [km^{-2}].
3. The time between the storm origin and the cell arrival is exponentially distributed with parameter β_c [h^{-1}].
4. The life-time of a rain-cell follows an exponential distribution with parameter η [h^{-1}].

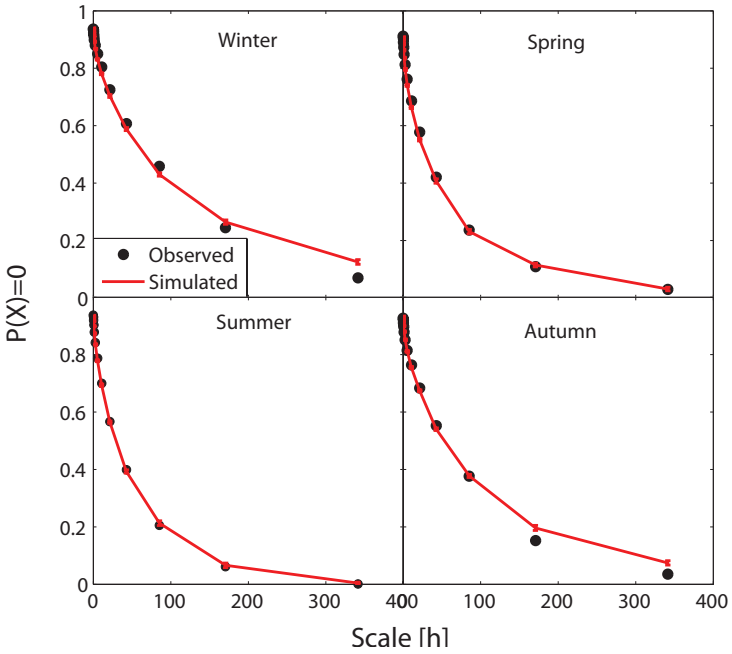


FIGURE 6.15: Probability of the point scale zero precipitation. Dots represent observed values and lines simulated values. the errorbars for each aggregation interval show the standard deviations of the simulated realizations.

5. The rain-cells have uniform intensity during their life time which is distributed according to a Weibull distribution with parameters α_c and θ_c.

6. The rain-cells are circular and their diameter is exponentially distributed with parameter φ [km^{-1}].

The simulated rainfall field is then the superposition of each rain-cell. In order to add spatial heterogeneity to the generated fields, *Burton et al.* [2008] proposed to weight the simulated output according to a weight function $\mathbf{\Phi(x)}$ that is defined as the ratio between the total mean precipitation and the average precipitation of each point that belongs to the simulation domain. This assumption is adopted here, which is somewhat different from the way that *Cowpertwait et al.* [2002] introduced heterogeneity on precipitation accumulation in their original version of the model.

One of the main drawbacks of the model is that it neglects the storm kinematics since all the generated storms are static. The representation of the areal storm development is rather simplistic but has a direct link to the physical process of precipitation, since it conceptualizes the rain-cell generation process. Further developments of the spatio-temporal version of the NS model have been recently introduced [*Cowpertwait*, 2010; *Burton et al.*, 2010b] in order to take into account spatial non-homogeneities in the storm arrival processes and also improve high temporal resolution statistics, but these were not studied here.

6.4.2 Model Calibration

One of the major advantages of the NS(s-t) model is that analytical expressions for the moments up to the third order and the spatial and temporal covariances exist for any aggregation interval. This gives the opportunity of calibrating the model parameters by applying a generalized method of moments technique

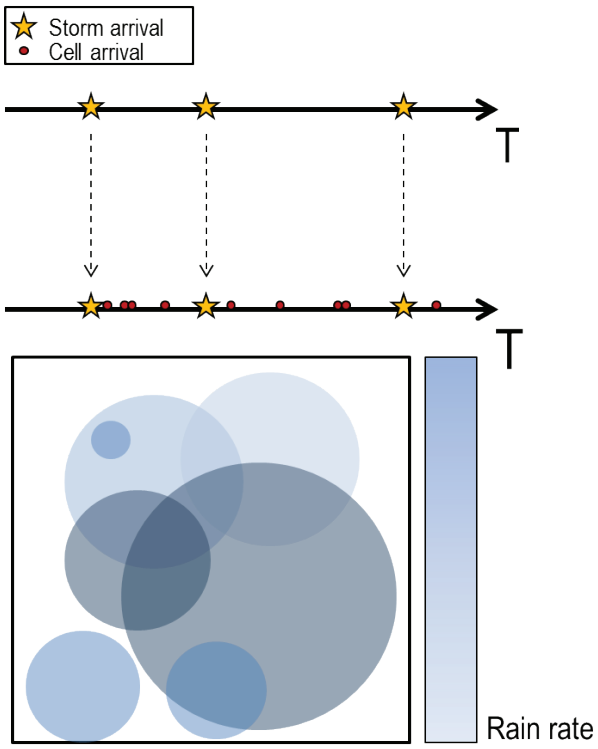


FIGURE 6.16: Schematic representation of the spatio-temporal Neyman Scott model.

[*Northrop*, 1998; *Evin and Favre*, 2012]. The minimization problem for the calibration of the model is the one used by *Burton et al.* [2010b]:

$$\hat{\boldsymbol{\theta}}_p = \arg\min_{\boldsymbol{\theta_p}} \sum_{h=1}^{n_h} \sum_{i=1}^{n} w_i \left(1 - \frac{g_i(\boldsymbol{\theta_p}, h)}{\hat{g}_i(h)}\right)^2, \tag{6.42}$$

where $\boldsymbol{\theta}_p$ is the parameter vector, g_i the statistics that are selected for the construction of the objective function and w_i their corresponding weights. h is the scale of aggregation where the statistics are evaluated. The statistics that were selected here are the mean, standard deviation, skewness, lag-one autocorrelation coefficient of precipitation intensity, the probabilities of zero rainfall, the conditional probabilities of state change ψ_{ww}, ψ_{dd} and the spatial autocorrelation for the hourly and daily scales. The analytical expressions of the statistics of the NS(s-t) model can be found in the Appendix B.1.

Similarly to the section 4.3.1 a multi-start downhill simplex algorithm [*Nelder and Mead*, 1965] was selected for the numerical minimization of the objective function. The fitting of the statistics was done using the data from the Swiss MetNet stations that lie in the study area. Precipitation time-series were first divided by the weight function $\mathbf{\Phi(x)}$ in order to be standardized. The minimization was carried out on a monthly basis.

In figure 6.17 the results of the calibration are given for the same area for which the space time model STREAP was validated. The results are very good for all the statistics except for the conditional probabilities that were very difficult to capture. One of the major possible problems is that the fitting procedure is applied to the average statistics (e.g. average standard deviation of all the stations) and if the assumption

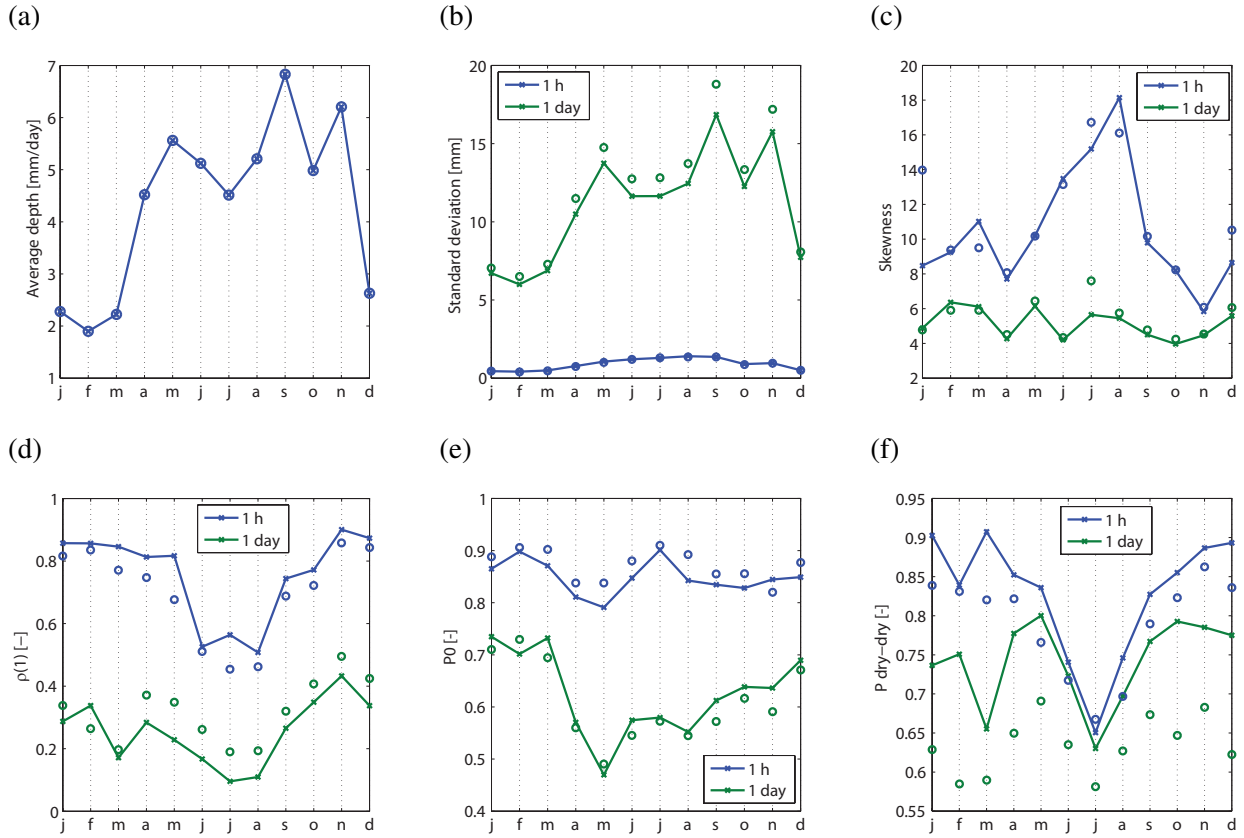


FIGURE 6.17: Comparison of the observed and fitted statistical parameters for the spatio-temporal Neyman-Scott model. Dots represent observed data and lines fitted. The data correspond to the area centred on the Monte Lema radar.

of homogeneous precipitation is not correct, then this can influence the model's efficiency. This problem can be of major importance in areas with complex topography, where local orographic effects affect the spatial distribution of precipitation.

6.4.3 Model Efficiency

In order to illustrate the model efficiency in reproducing the space-time statistics of the observed fields, results from only one station (Lugano) will be plotted. The Lugano station corresponds to a station with average performance. As it has already been previously shown, the Poisson cluster models cannot reproduce well high temporal resolution precipitation statistics.

This also holds true for the space time version of the model as can be seen in figure 6.18. The lines that represent the simulated exceedance probability distributions for depth accumulations, were derived from an ensemble of 50 simulations with temporal length of 30 years, approximately the same as the length of the SwissMetNet records. It can be clearly shown that the NS(s-t) model underestimates the tails of the precipitation depths for short accumulation times. On the other hand, the model behaves well for aggregation intervals larger than one hour. The two models (STREAP - NS(s-t)) appear to be equivalently good, with STREAP slightly outperforming the NS(s-t) for high resolution series (figures 6.13, 6.18).

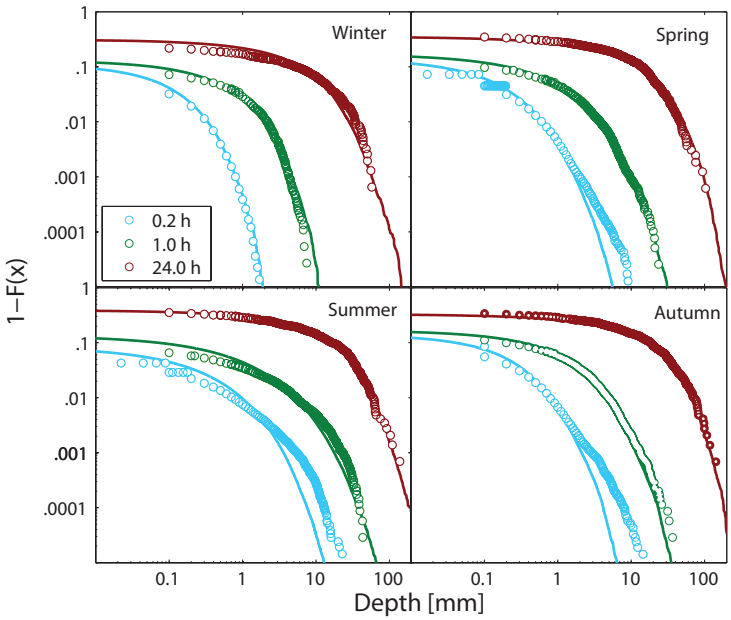


FIGURE 6.18: Exceedance probability distributions of the precipitation depth accumulation at the Lugano station on a seasonal basis for three different aggregation intervals. The aggregation intervals are marked with different colours. Dots represent observed data and lines simulated series.

Most of the studies that have applied the NS(s-t) model, mainly focus on the reproduction of the point scale statistics such as the three first moments and lag-one autocorrelation coefficients, on which the model is actually calibrated. Here a thorough investigation of both point scale and areal statistics is conducted in order to gain an overall idea concerning the model's behaviour. When point scale statistics are mentioned, the simulations were carried out taking into account the $\mathbf{\Phi(x)}$ as derived from the gauge data and when spatial statistics are investigated the simulations correspond to $\mathbf{\Phi(x)} = \mathbf{1}$ in order to avoid any probable issue of extrapolating the $\mathbf{\Phi(x)}$ for the entire 2D domain.

One of the most common problems concerning the Poisson cluster models [*Entekhabi et al.*, 1989], is the correct reproduction of the probability of no precipitation. This also the case here (figure 6.19). It should be noted though that another problem linked with this behaviour is associated with the calibration procedure. For the minimization of equation 6.42 only the average probability of precipitation occurrence is taken into account over all the stations. Spatial non-homogeneity for the occurrence process does exist in this area, since the probability of zero rainfall across scales is different for the various stations. Therefore, the problem shown in figure 6.19 is the composite effect of both the model structure and its calibration procedure. The latter one can be solved if a spatial inhomogeneous version of the model is used [*Burton et al.*, 2010b].

In comparison to figure 6.15 it can be clearly shown that STREAP achieves a much better reproduction of the probability of no precipitation, across all the investigated time-scales.

Another problem typically associated with the Poisson-cluster models is the reproduction of the probability distributions of the dry and wet lengths. Here the assumption is that wet-dry durations are defined as the consecutive time periods that the records stay above or equal to zero respectively. This assumption is adopted here in order to avoid the problem of event definition that is not straightforward [e.g. *Vandenberghe et al.*, 2010].

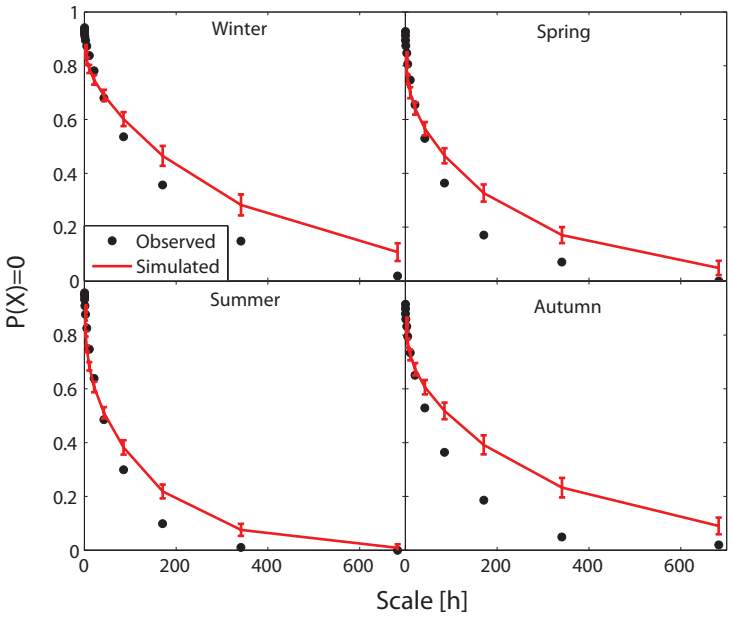


FIGURE 6.19: Probability of zero precipitation for the Lugano station. Dots represent observed values and lines simulated. The errorbars for each aggregation interval are the standard deviations for the model's realizations.

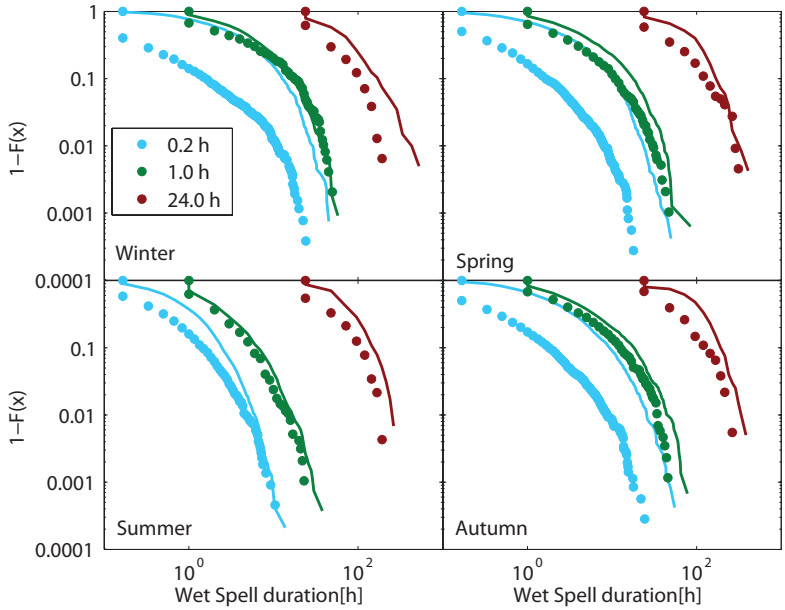


FIGURE 6.20: Exceedance probability for the wet spell durations at the Lugano station. Dots present observed data and lines simulated.

In figure 6.20 the duration of the wet lengths is shown. The spell durations are highly overestimated. For lower temporal scales the problem is enhanced significantly. In comparison, the STREAP model (figure 6.14) performs much better for this statistic. The main problem of the NS(s-t) model is the simplified way that the cell arrival process is simulated that leads to a strong persistence of the cells in time. The behaviour of the dry durations is similar. The overestimation of the dry durations can pose serious restriction of the

NS(s-t) model in case it used for draught simulation and risk analysis. Another problem associated with the strong over-estimation of the wet and dry durations is that the reproduction of the distributions of WAR and IMF is also problematic (figure 6.21), if we take into account that the total precipitation depth has to be preserved in the calibration procedure. In this case the NS(s-t) model leads to a compensation of several sources of errors in order to predict well the statistics that it is calibrated on.

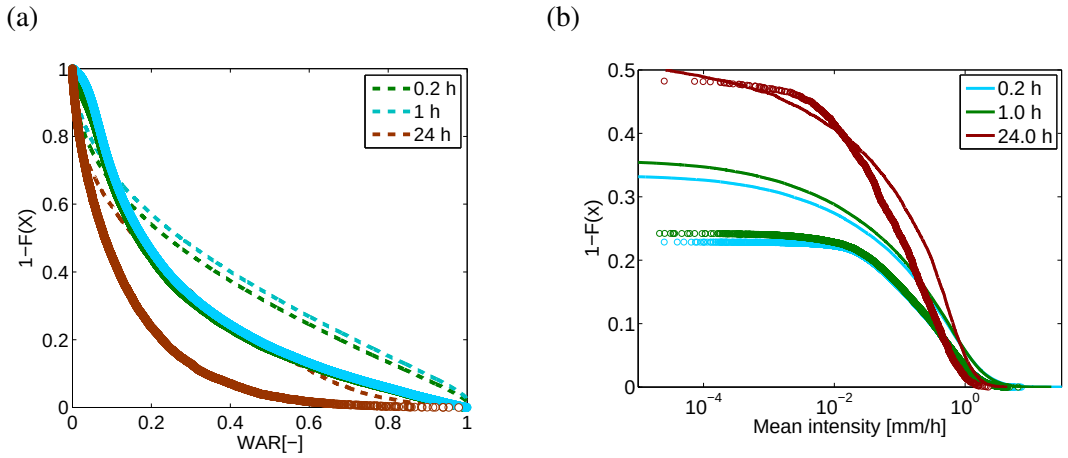


FIGURE 6.21: Comparison of the exceedance probabilities of IMF and WAR process for the NS-st model. (a) Left panel shows the exceedance probabilities of WAR; and (b) the right one of IMF. Different colours correspond their respective temporal scales according to the legend. Dots present observed data and lines simulated

Another problem is the correct reproduction of the autocorrelation function for a large number of lags. The model is calibrated to predict well the lag one autocorrelation coefficient but this in not adequate to capture its entire range as is shown in figure 6.22. For aggregation scales coarser than one hour the problems are reduced or even disappear, but the model is not able to reproduce small scale features concerning the temporal autocorrelation structure. Unfortunately the STREAP model also has deficiencies in reproducing well the temporal correlation structure due to calibration difficulties previously described.

The reproduction of the areal statistics is of major importance. One of the main difficulties of assessing the efficiency of the NS(s-t) model on reproducing areal statistics is that it is calibrated against gauge data. Since the number if gauges in the area is rather small (11), the estimation of the fractions of wet areas and mean areal intensities is impossible. In order to get a first, but not ultimately fair, impression of the model's efficiency, the statistics will be compared against the ones derived by the radar. Since the precipitation estimation from the gauge during summer is very good, the statistics that refer to summer liquid precipitation can be considered robust.

In figure 6.23 in comparison to 6.8, it is obvious that the STREAP model reproduces much better the auto(cross) correlation functions of WAR and IMF. This is expected since the second stage of STREAP is actually calibrated to reproduce these statistics. The problem of NS(s-t) model again is mainly observed on low temporal scales (10 min) where typically the autocorrelation of WAR is overestimated, the one of IMF is underestimated and their crosscorrelation is also underestimated.

Finally, the composite effect of the correct reproduction of the spatial autocorrelation and the fraction of wet areas can be show in figure 6.24. Due to the simplified way that the NS(s-t) model conceptualises the spatial structure of precipitation as uniform circular cells, it systematically overestimates the length of any crossection cover with precipitation. This is due to the fact that the cell sizes are frequently larger than

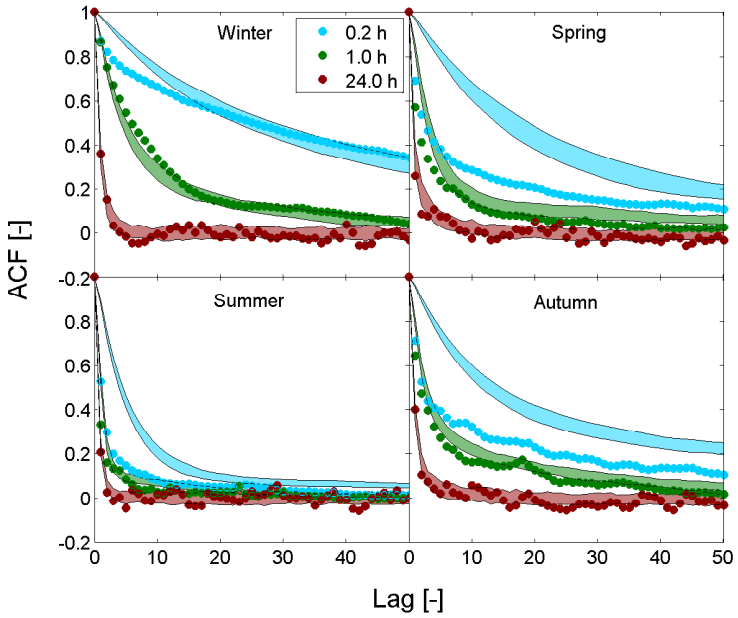


FIGURE 6.22: Autocorrelation function for the Lugano station.

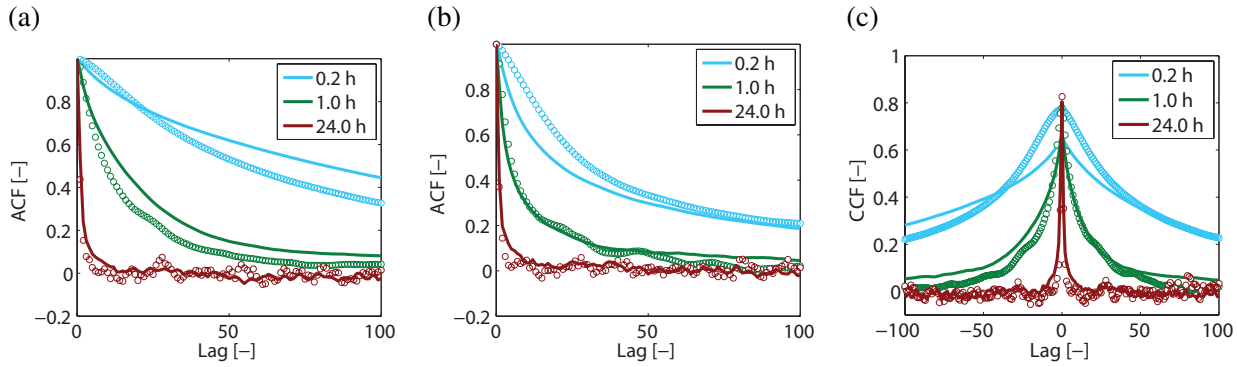


FIGURE 6.23: (a) On the left panel, the autocorrelation of the WAR process is shown. (b) In the middle the autocorrelation of the IMF process and (c) on the right their cross-correlation. Different colours correspond to various aggregation intervals as shown in their corresponding legends. Dots correspond to observations and lines to the mean value of an ensemble of simulated series.

the observed precipitation cells, especially during summer when very intense, localised small convective cells develop.

6.5 Further Developments of STREAP

The STREAP model is based on three main assumptions. The simulated precipitation fields are isotropic, stationary and the probability of precipitation occurrence is everywhere the same. Here some ways to overcome these limitations are proposed.

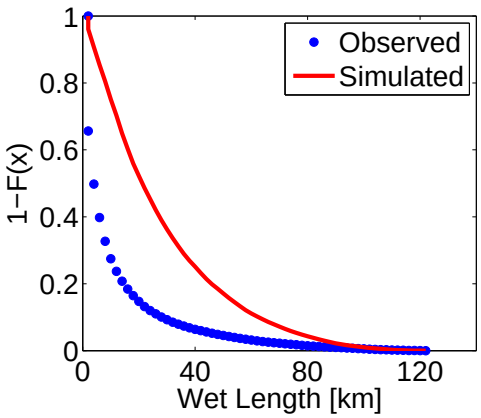


FIGURE 6.24: Comparison of the exceedence probabilities of the length of an arbitrary cross-section to be covered with precipitation between radar data and the output of the NS(s-t) model.

6.5.1 Anisotropy

It has been previously shown that isotropy is a valid assumption for precipitation at the $\beta-$ mesoscale ($20 \sim 200$ km). On the other hand in areas where orographic blockage is expected to play a major role for example, precipitation fields can be highly anisotropic [*Ebtehaj et al.*, 2010]. The most common type of anisotropy is the one of the elliptical type. A method for simulating random fields with anisotropic correlation in the frequency domain can be found in *Chorti and Hristopulos* [2008]. An illustration of the appearance of a rainfall field with geometric anisotropy is shown in figure 6.25. The method is a simple coordinate transform of the type

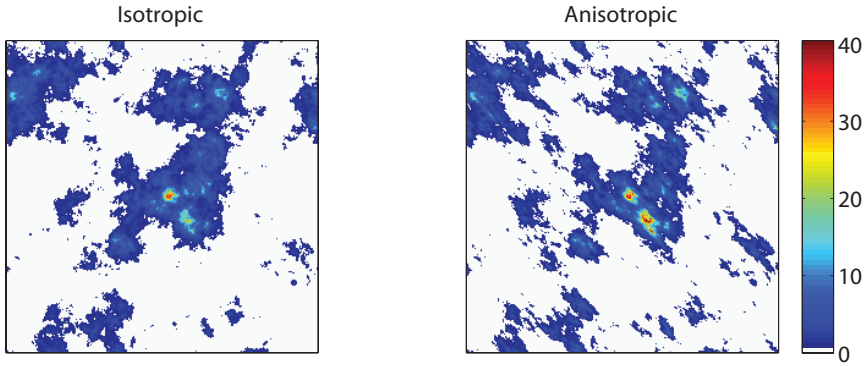


FIGURE 6.25: Illustration of a the simulation of spatial fields with geometric anisotropy. The isotropic and anisotropic fields were simulated with the same seed number of the random number generator. The anisotropic field has an anisotropy ratio of 0.5.

$$h_1 = \frac{\cos\theta}{\xi_1} r_1 + \frac{\sin\theta}{\alpha_1} r_2, \tag{6.43}$$

$$h_2 = -\frac{\sin\theta}{\xi_2} r_1 + \frac{\cos\theta}{\alpha_2} r_2, \tag{6.44}$$

where (r_1, r_2) are the transformed isotropic coordinates and (h_1, h_2) the initial coordinates of the anisotropic correlation function; α_1, α_2 correspond to the correlation lengths (maximum and minimum) along the principle axes (h_1, h_2) and R is their ratio $\zeta = \alpha_1/\alpha_2$.

The Euclidean distance $|h|$ can thus be expressed in terms of r_1, r_2 as

$$|h|^2 = A_1 r_1^2 + A_2 r_2^2 + A_{12} r_1 r_2, \tag{6.45}$$

where

$$A_1 = \frac{1}{\alpha_1^2}\left(\cos^2\theta + \frac{\sin^2\theta}{\zeta^2}\right), \tag{6.46}$$

$$A_2 = \frac{1}{\alpha_1^2}\left(\sin^2\theta + \frac{\cos^2\theta}{\zeta^2}\right), \tag{6.47}$$

$$A_{12} = -\frac{2}{\alpha_1^2}\sin\theta\cos\theta\left(\frac{1}{\zeta^2} - 1\right). \tag{6.48}$$

Finally the analytical expression of the power spectrum that is needed in order to simulate the Gaussian random fields is

$$S(k_1, k_2) = |\mathcal{J}_1| \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} C_g(h_1, h_2)\exp(-ih_1\tilde{k}_1 - ih_2\tilde{k}_2)\mathrm{d}h_1\mathrm{d}h_2, \tag{6.49}$$

where $|\mathcal{J}_1|$ is the Jacobian of the coordinate transform

$$|\mathcal{J}_1| = \left|\frac{\partial(r_1, r_2)}{\partial(h_1, h_2)}\right| = \zeta\alpha_1^2 \tag{6.50}$$

and

$$\tilde{k}_1 = \alpha_1(k_1\cos\theta + k_2\sin\theta), \tag{6.51}$$

$$\tilde{k}_2 = \alpha_1(-k_1\sin\theta + k_2\cos\theta). \tag{6.52}$$

For the case of the exponential correlation function that is used in the STREAP model the power spectrum is

$$S(k_1, k_2) = 2\pi\alpha_1^2 R\frac{1}{(1 + G_1k_1^2 + G_2k_2^2 + G_{12}k_1^2k_2^2)^{\frac{3}{2}}}, \tag{6.53}$$

where

$$G_1 = \alpha_1^2(\cos^2\theta + \zeta^2 \sin^2\theta), \tag{6.54}$$

$$G_2 = \alpha_1^2(\sin^2\theta + \zeta^2 \cos^2\theta), \tag{6.55}$$

$$G_1 = 2\alpha_1^2 \cos\theta \sin\theta(1 - \zeta^2). \tag{6.56}$$

Various other forms of anisotropy in precipitation may also exist. For example *Ebtehaj et al.* [2010] identified different anisotropic structures dependent on the spatial scale. In order to simulate such fields, a methodology similar to the simulation of fields according to the Generalized scale invariance concept can be used [e.g. *Lovejoy and Schertzer*, 1985]. Algorithmic details concerning its simulation can be found in *Pecknold et al.* [1993].

6.5.2 Non-Homogeneous Probability of Precipitation Occurrence

Another main restriction of the STREAP model is that the probability of precipitation occurrence on the simulated domain is spatially equal. That restriction can be crucial when areas with e.g. strong orographic enhancement are simulated. Generally the probability of precipitation occurrence increases with altitude and this has to be taken into account. Very few studies have focused on parametrizing this kind of nonstationarity of the field in stochastic models. One of the most recent examples is the development of the spatially non-stationary Poisson cluster model of the Neyman-Scott type [*Burton et al.*, 2010b].

Here a simple parametrization is proposed in order to include non-homogeneity in the STREAP model. The methodology is based on modifying appropriately the simulated Gaussian fields [e.g. *Kleiber et al.*, 2012].

The principle idea is to add a filter field on the simulated Gaussian fields. By doing that a non-homogeneous thresholding across the spatial domain occurs defining spatially non homogeneous probabilities of precipitation occurrence.

Let us define the modified $G_\omega(x, y)$ field as

$$G_\omega(x, y) = G(x, y) + a_\omega(x, y). \tag{6.57}$$

Moreover we define F_{G_ω} and $F_{G_\omega}^{-1}$ the respective cumulative and inverse cumulative distribution functions of the G_ω field. Also we define as $F_\omega(x, y)$ the cumulative distribution function of every pixel in the 2D domain. From equation 6.57 it is straightforward to show that

$$F_\omega(x, y) = F_N[a(x, y), 1] \tag{6.58}$$

where $F_N[a(x,y),1]$ is the cumulative distribution function of the normal distribution with mean value $a(x,y)$ and unit variance (i.e. $\sim N(a(x,y),1)$).

If these assumptions are made, then the probability of precipitation occurrence in the 2D domain can be estimated as

$$\overline{WAR}(x,y) = \frac{1}{T}\sum_{t=1}^{T} P(G_\omega(x,y,t) > \theta_\omega(t)), \tag{6.59}$$

where $\theta_\omega(t) = F_{G_\omega}^{-1}(1 - WAR(t))$ which implies that

$$\overline{WAR}(x,y) = \frac{1}{T}\sum_{t=1}^{T} 1 - F_N\left(F_{G_\omega}^{-1}\left(1 - WAR\left(t\right)\right), a(x,y), 1\right). \tag{6.60}$$

For simulation, the field $\overline{WAR}(x,y)$ and the series $war(t)$ can be estimated from the data and 6.60 can provide the estimates of the $a(x,y)$ field that are needed. Since $F_{G_\omega}^{-1}$ depends on the entire field of $a(x,y)$ the estimation procedure is not straightforward. A suggestion is to get the estimates of $a(x,y)$ from the nonlinear minimization problem defined as:

$$\hat{a}(x,y) = \arg\min_{a(x,y)} \sum_x \sum_y \left[\overline{WAR}(x,y) - 1 + \frac{1}{T}\sum_{t=1}^{T} F_N\left(F_{G_\omega}^{-1}\left(1 - war(t)\right), a(x,y), 1\right)\right]^2. \tag{6.61}$$

This problem is of high complexity since for e.g. a spatial domain of 64×64 pixels it evolves 4096 free parameters to be estimated. For this reason its calibration can be cumbersome and computationally very demanding.

It should be noted that the auto-covariance function of $G_\omega(x,y)$ is not equal to $G(x,y)$ and this has to be taken into account. In case if $a(x,y)$ is a stationary field, due to the independence of $G(x,y)$ and $a(x,y)$ then the auto-covariance of $G_\omega(x,y)$ equals the sum of covariances of $G(x,y)$ and $a(x,y)$.

Figure 6.26 shows an illustration of how the precipitation fields are modified in order to have a lower probability of precipitation occurrence in areas where rainfall events are more rare. It is clear that the proposed methodology modifies satisfactorily well the precipitation fields since the probability of zero precipitation enhances for dry areas without loss of the physical realism of the spatial precipitation fields.

A major problem in applying this methodology to precipitation radar data for mountainous areas occurs since there are usually areas that are affected by partial beam shielding. In order to apply such a filtering technique the correct estimation of $\overline{\mathrm{WAR}(x,y)}$ over the entire study area domain is needed, otherwise the adjustement proposed here could simply reproduce the beam shielding effect. Since, as shown in figure 2.4 this is not the case in the area studied here, this methodology cannot be used.

6.5.3 Non Homogeneous Spatial Precipitation Accumulation

Another form of non-stationarity that can be taken into account is the different spatial depth accumulations in space $\overline{R(x,y)}$. The most straightforward procedure is to multiply the simulation output with a field

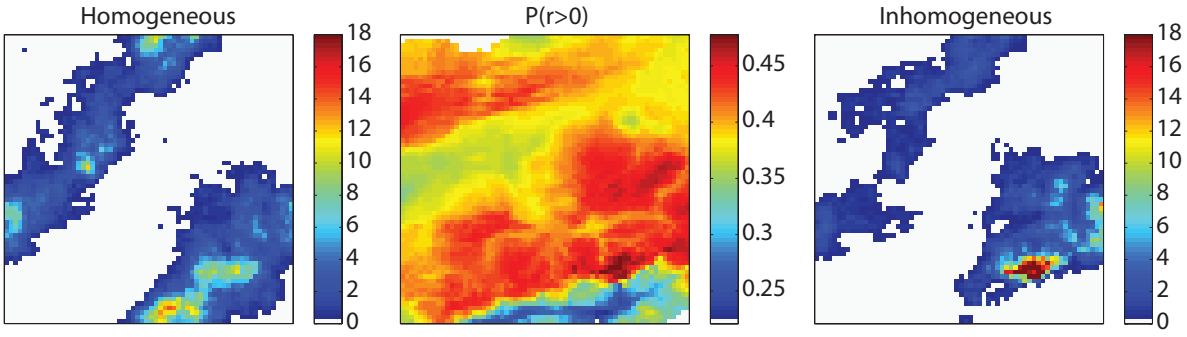


FIGURE 6.26: Illustration of the modified algorithm that takes into account inhomogeneous probabilities of precipitation occurrence in space. The illustration corresponds to the Alpine area of central Switzerland and has an extend of 128×128 km^2. The middle panel shows the probability of daily precipitation occurrence for this areas, according to which the latent Gaussian fields are modified.

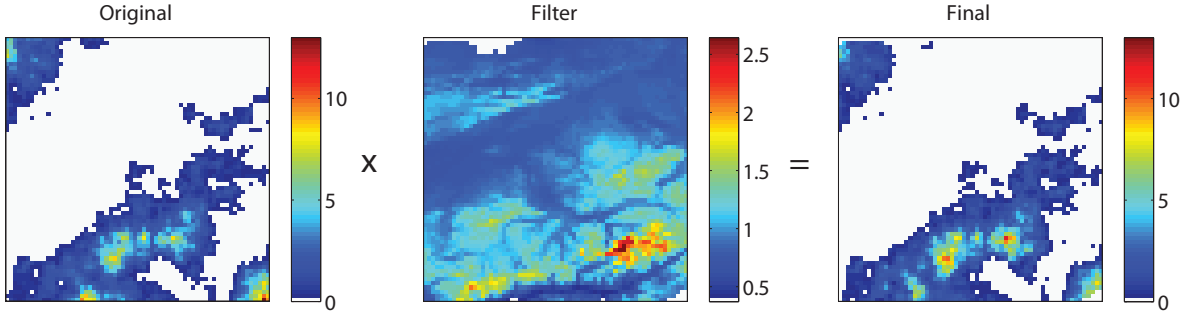


FIGURE 6.27: Schematic representation of the simulation output modification for the introduction of nonhomogenious spatial depth accumulations. The spatial precipitation distribution correspond to the Alpine area of central Switzerland and has an extent of 128×128 km^2.

of weights similar to *Burton et al.* [2008] modified in order to take into account the non-homogeneous precipitation occurrence. This multiplication factor can be defined as

$$\mathbf{\Phi}(x,y) = \frac{\iint\limits_{x,y} \overline{\mathrm{WAR}(x,y)}dxdy}{\iint\limits_{x,y} \overline{R(x,y)}dxdy} * \frac{\overline{R(x,y)}}{\overline{\mathrm{WAR}(x,y)}}. \tag{6.62}$$

A schematic representation of this modification is illustrated in figure 6.27. This solution however has an effect on the spatial auto-covariance of the simulated fields.

Summarizing, there are also ways to in include in a straightforward manner also orographic effects in the STREAP model, which can make it more efficient for orographically complex areas.

7

THE EFFECT OF SPACE-TIME RAINFALL VARIABILITY ON BASIN RESPONSE

One of the major open questions in hydrology is how the spatial and temporal structure of precipitation affects the response of river basins, i.e. the generated discharge. The impact of the spatio-temporal structure of precipitation on the catchment response can be crucial for the design of the precipitation monitoring network and flood protection infrastructure especially for areas prone to flash floods such as the mountainous steep basins of the European Alps. In this chapter a numerical quantification of the impact of the space-time variability of rainfall on basin response is presented exploiting the capabilities of the space-time stochastic model STREAP presented in the previous chapter.

7.1 Introduction

Various studies have focused on the topic of the effects of rainfall variability, sometimes yielding controversial results, highlighting the complex nature of the problem. Most of those studies were based on numerical experiments where precipitation input from data or stochastic generated ensembles, were used to force hydrological models.

Many studies are based on stochastic precipitation input into hydrological models to assess the sensitivity of the basin response to the precipitation spatio-temporal variability. One of the early studies by *Krajewski et al.* [1991] used the widely known WGR (Gupta-Rodrigues Iturbe-Waymire) stochastic rainfall model [*Waymire et al.*, 1984] to simulate continuous space-time rainfall sampled at different spatial and temporal resolutions and then used it as input into one fully distributed and one lumped hydrological model. For a small (7.5 km^2) rural catchment in Iowa, they found that the temporal sampling resolution is more important than the spatial one in terms of the sensitivity of peak discharge, time to peak and total volume, showing that the effect of temporal precipitation variability can be of major importance. The authors also addressed the fact that the results they obtained were model dependent with the distributed model giving a much better process understanding, as expected, due to its explicit physical representation of the hydrological processes. In a similar analysis *Shah et al.* [1996a, b] found that the sampling resolution is crucial for both space and time, using a different stochastic modelling approach for rainfall. They also highlighted the importance of the initial soil moisture conditions on the sensitivity of the runoff generation process, with dry initial conditions resulting in much more variable runoff responses, dependent on the

rainfall spatial and temporal sampling resolutions. Moreover, they addressed the issue of the space-time correlation of the precipitation fields. *Booij* [2002] following a similar modelling strategy identified as well that the effect of the spatial resolution of the hydrological model can significantly impact the results as well. In catchment hydrology, the discretisation of the system in space and time is crucial in order not to violate the process dynamics and their interactions. In addition, *Gabellani et al.* [2007] quantified the effects of the storm's spatial and temporal correlation on the peak of the generated discharge. In a recent study *Gires et al.* [2012] using a multifractal MRC showed as well that small scale spatio-temporal variability can be crucial also for urban catchments, where the concentration times are small. Finally, *Mandapaka et al.* [2009], identified that the spatio-temporal rainfall structure can have an effect on the statistical structure of peak flows and more specifically on the scaling relations between peak flows and contributing areas.

There have also been several data-based studies which investigated the topic of rainfall effects in floods [e.g. *Obled et al.*, 1994]. Those studies have the advantage that the rainfall input is observed and any artefact related to the stochastic model structure can be avoided. However these approaches suffer from the low accuracy of the data, especially if they are derived by weather radars, and short data records, that usually restrict the analysis to be event specific.

Nicotina et al. [2008] found that the spatial sampling resolution is of minor importance and does not affect significantly the prediction of the peak discharge. Moreover the authors explored the effect of catchment size, concluding that larger catchments are more prone to errors if the spatial variability is not well represented. They concluded that the key parameter for the correct estimation of the discharge is the total precipitable volume. To the author's opinion the simplified approach for modelling discharge using a Geomorphologic unit hydrograph (GUIH) approach [*Rinaldo and Rodriguez-Iturbe*, 1996] limits the validity and generality of their results due to the lack of a concrete physical background and due to the lack of an explicit preservation of spatial heterogeneity (retained only in probabilistic terms). Similarly to their approach, *Nikolopoulos et al.* [2011] used radar-derived precipitation fields aggregated to different scales and import it as input to the physically based distributed model TRIBS [*Ivanov et al.*, 2004; *Vivoni et al.*, 2005, 2009]. They mainly investigated the effects of different rainfall input resolutions dependent on the catchment sizes and initial soil moisture conditions. Their findings concerning the initial soil moisture are in complete agreement with *Shah et al.* [1996b] suggesting that the rainfall variability becomes significant when the catchment is drier. Another interesting finding is that the results are also basin size dependent with the sensitivity increasing with catchment size. Finally they showed that high sampling resolution is more important for low intensity events.

Another popular approach for the investigation of the effects of rainfall variability on generated runoff is based on the analytical framework first introduced by *Woods and Sivapalan* [1999]. Simplified approaches of the runoff generation mechanisms lead to analytical derivation of the moments (typically variance) of the discharges, travel times etc. This kind of method has been recently expanded by *Viglione et al.* [e.g 2010] and *Volpi et al.* [2012]. Even though the analytical derivations are appealing, the required simplifications of can be misleading and not representative of the natural processes. This limits the capabilities of such an analysis framework to capture the full effect of the spatio-temporal variability.

From this short review of the state of the art of the effect of space-time variability on basin response, it becomes clear that the complexity of the problem is very high and the composite effect of many variables has to be taken into account. In this chapter a quantification of the sensitivity of the basin response of a natural catchment, to the spatiotemporal structure of rainfall is conducted, in a similar manner as in *Krajewski et al.* [1991] and *Shah et al.* [1996b] by exploiting the simulation tool developed in chapter 6.

7.2 The TOPKAPI-ETH (TE) Model

During the last decades, when the computational resources have increased exponentially, several hydrological models have been developed with different degrees of complexity spanning from purely conceptual [e.g., *Kitanidis and Bras*, 1980] to semi-empirical and physically-based ones [e.g., *Kollet and Maxwell*, 2006, 2008; *Ivanov et al.*, 2004; *Fatichi et al.*, 2012a; *Rigon et al.*, 2006]. Dependent on the spatial organization of the computational elements of each model, they are commonly referred to as lumped (catchment scale), semi-distributed (large scale sub-catchments) or fully-distributed (small scale elements).

Through the years, the hydrological models have evolved in order to take into account with a higher detail the processes involved and especially the interactions between land and atmosphere etc. After the recognition of the importance of vegetation on the hydrologic response of the natural system [*Rodriguez-Iturbe*, 2000], several efforts have been done in order to take into account such information [*Ivanov et al.*, 2008; *Fatichi et al.*, 2012a, b]. Due to the complexity of the natural system though, a common practice is that each of the models specifically improved some of the processes involved. Models that are targeting on the correct reproduction of the generated response of a catchment are mainly oriented on the detailed description of water flow (surface and subsurface) [e.g. *Kim et al.*, 2012; *Kollet and Maxwell*, 2006]. Others that focus on the detailed description of the energetic components of the system are based on detailed descriptions on land-atmosphere interactions [e.g. *Ivanov et al.*, 2008; *Fatichi et al.*, 2012a, b].

The main processes that are important at the event scale, when discharge is the variable of interest, are the runoff generation mechanisms (infiltration-saturation excess), the surface water routing (overland and channel flow) and the subsurface water flow. In order to take all those processes into account, the fully distributed hydrological model TOPKAPI-ETH (TE) is used in this study. The model is an improvement of the original TOPKAPI (**Top**ographic **K**inematic **Ap**roximation and **I**ntegration) model [*Liu and Todini*, 2002; *Ciarapica and Todini*, 2002]. The main features of the model are:

1. Subsurface, Surface and channel water routing are approximated by the kinematic wave approach.
2. Infiltration is simulated according to an explicit expression of the Green-Ampt infiltration rate [*Salvucci and Entekhabi*, 1994].
3. Vertical subsurface flow is conceptualized as a cascade of two soil reservoirs and a deep percolation layer.

A detailed description of the model can be found in *Rimkus et al.* [2013]. The kinematic wave approach that is adopted here [*Chow et al.*, 1988; *Bates and De Roo*, 2000; *Dingman*, 1994] originates from the shallow water equations assuming that the energy slope is equal to the topographic one. This is generally accepted as an adequate approximation for catchment hydrology, and that is the reason why it is employed in the vast majority of hydrological models. It is also computationally efficient, in contrast to the numerically demanding dynamic wave approaches. The detailed description of the kinematic wave approximation and its solution is beyond the scope of this thesis and can be found in *Chow et al.* [e.g. 1988] or *Brutsaert* [2005].

The main simplification of the model is the conceptualization of the subsurface flow. Instead of solving the three dimensional Richards equations for variably saturated flow [*Hillel*, 1998], TE simplifies the problem introducing a cascade of non-linear subsurface reservoirs corresponding to a shallow and a deep soil layer [*Liu and Todini*, 2002]. Adopting such a formulation, the computational demand is decreased

significantly, since instead of solving the Richard's partial differential equation numerically, the approach used in TE that is non-iterative. This assumption is a good compromise between computational efficiency and physical realism of subsurface flow.

TE also models other key processes of the hydrological cycle such as evapotranspitation (ET), snow accumulation and melt, glacier melt etc. Investigating the details of those processes is not of major importance on the event scale and thus the process description and their numerical implementation are not further explained here. The detailed description of the model can be found in *Rimkus et al.* [2013].

7.3 Study Basin

The basin that was selected as the case study is the Kleine Emme river basin, a pre-alpine catchment located in central Switzerland (see figure 7.1(a)). The catchment size is 477 km^2 with a mean elevation of 1047 [m.asl.] and mean slope $\sim$ 16 degrees. The area is mainly pristine. Also there are no glaciers in this catchment. The area is prone to flooding, mainly due to summer intense convective events. For example this specific area suffered from serious flood related problems during the intense storms of August 2005. For this reason, the investigation of its response, dependent on the space-time variability of rainfall is crucial. The area is in a large extent covered by forest and grassland. A small fraction of the area ($\sim$4%) is used for agricultural purposes (figure 7.1(c), table 7.1). A detailed description of the soil composition and the landuse of the catchment is given in figure 7.1 and tables 7.1, 7.2. The data of the soil classification is provided by the Soil map (Bodeneignungskarte[1]) of Switzerland (Swiss Federal Statistical Office) whereas the landuse was obtained by the global cover product of the European Space Agency (ESA)[2].

The Kleine Emme river basin is a well monitored catchment in terms of both precipitation and river discharge. Three streamflow gauges are located in the catchment, 3 rain-gauges of the Swiss MetNet also lie inside the catchment and the area is relatively well captured from the weather radars with negligible beam shielding, minor clutter contamination and adequate representation of the sampling volume in the vertical direction. Also this catchment is representative in terms of size, and geomorphological features (steepness, soil structure etc.) of a typical Alpine catchment in Switzerland.

Discretisation of the Study Basin

The issue of the spatial and temporal resolution of the computational elements of a hydrological model is also of major importance [e.g. *Vivoni et al.*, 2005; *Kuo et al.*, 1999; *Fatichi et al.*, 2012a] and can have strong influences on the generated output. The computational elements of TE are structured on a regular square grid. In order to avoid excessive computational effort and at the same time not violate the hillslope dynamics, a spatial grid resolution of 100×100 m^2 was selected (Figure 7.2). This results to 47707 computational elements for the Kleine Emme river basin.

The vertical discretisation of the catchment soil layer is shown in figure 7.3. The soil is conceptualized as a cascade of 3 soil layers. The depth of the layers is assumed in this case dependent only on the soil class and not on the local topography of the area. This is a simplification, since it is known that the depth of the soil is dependent on the local topography of the area [e.g. *Heimsath et al.*, 1997, 1999; *Dietrich and*

[1] http://www.blw.admin.ch/dienstleistungen/00334/00337/index.html?lang=de

[2] http://www.esa.int/Our_Activities/Observing_the_Earth/ESA_global_land_cover_map_available_online

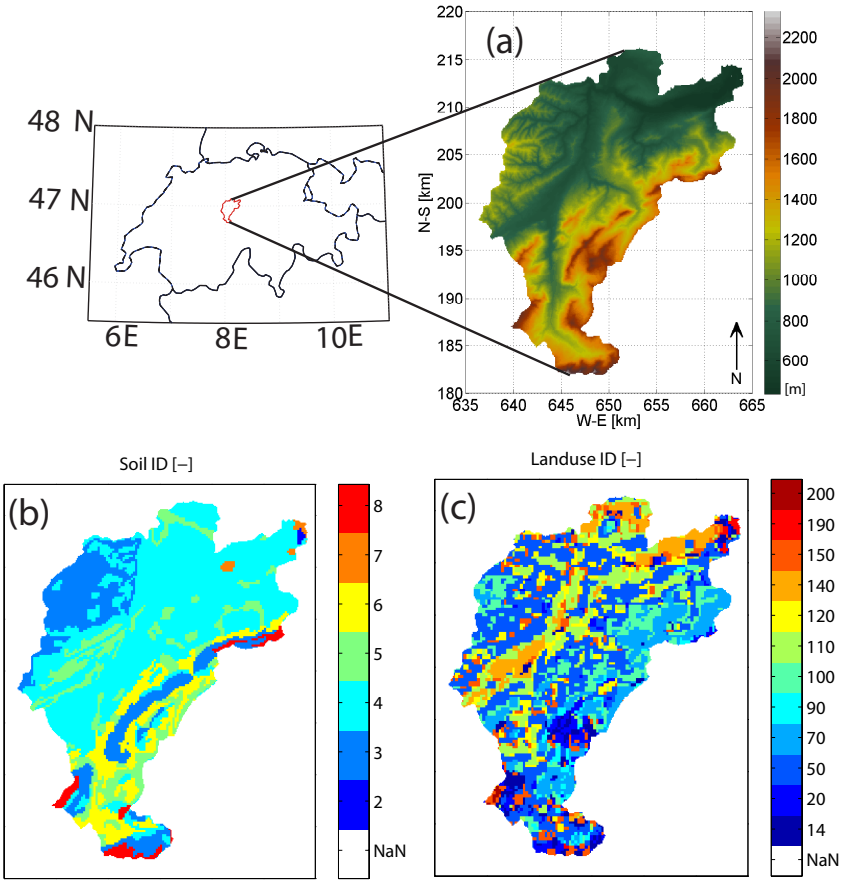


FIGURE 7.1: Description of the Kleine Emme river basin.(a) The DEM;(b) the soil map and (c) the landuse map. Properties for the various ID's can be found in tables 7.1 and 7.2

.

Reiss, 1995], due to soil generation and transport mechanisms in the catchment. However, since no data are available concerning the true spatial distribution of soil depth across the catchment, the simplified assumption of a constant depth is adopted here.

7.4 Model Calibration

The calibration of any hydrological model is crucial for its application. An extensive review of the calibration procedures is beyond the scope of this study. Since the model parameters have a physical meaning, a "manual" calibration procedure was chosen. Appropriate ranges for the model parameters, mainly concerning the soil, were first identified and subsequently, a random search was conducted in a defined parameter space, to find a good parameter set that is sufficient for this study. For a better, calibration though, an extensive numerical optimization algorithm should be employed, that is very expensive computationally, taking into account the distributed nature of the model.

As it can be shown in figures 7.4, 7.5 and table 7.3 TE performs very well in reproducing the correct discharge at the basin outlet. Especially in figure 7.5 it can be shown that the model can simulate well the most extreme flood event recorded in the Kleine Emme basin. This is of major importance here, since the

TABLE 7.1: Description of the Landuse map ID's

Landuse ID	Description
14	Rainfed croplands
20	Mosaic cropland
50	Decidous
70	Evergreen
90	Open (15-40%) needle leaved deciduous or evergreen forest (>5m)
100	Closed to open (>15%) mixed broadleaved and needle leaved forest (>5m)
110	Mosaic forest or shrubland (50-70%) / grassland (20-50%)
120	Mosaic grassland (50-70%) / forest or shrubland (20-50%)
140	Evergreen or deciduous, shrubland (<5m)
150	Closed to open (>15%) herbaceous vegetation (grassland)
190	Closed (>40%) broadleaved forest or shrubland permanently flooded - Saline or brackish water
200	Closed to open (>15%) grassland or woody vegetation on regularly flooded or waterlogged soil - Fresh

The Landuse ID's correspond to the ones from the global cover product of esa http://ionia1.esrin.esa.int/

TABLE 7.2: Description of the soil classes

Soil ID	2	3	4	5	6	7	8
Θ_s	0.46	0.46	0.48	0.52	0.5	0.475	0.99
Θ_r	0.05	0.14	0.06	0.27	0.3	0.09	0.01
Ks_h	5.58E-03	1.47E-03	1.76E-03	9.42E-04	8.27E-04	8.67E-04	1.00E-11
Ks_v	6.98E-05	1.83E-05	2.19E-05	1.18E-05	1.03E-05	1.08E-05	2.00E-04

Θ_s: Saturated water content [-]
Θ_r: Residual water content [-]
Ks_h: Horizontal saturated hydraulic conductivity [ms^{-1}]
Ks_v: Vertical saturated hydraulic conductivity [ms^{-1}]
The Soil IDs correspond exactly to the ones from the Bodeneignungskarte.

space-time variability of rainfall on extreme flooding is the main goal of the study.

TABLE 7.3: Efficiency of the TOPKAPI model for the Kleine Emme catchement across temporal scales

Scales	Correlation coefficient (r).	Nash-Sutcliff efficiency (R_{NSE}).
1 hour	0.8564	0.6627
1 day	0.9163	0.8017
1 month	0.9167	0.8189
1 year	0.9787	0.9102

7.5 Data Analysis

First, a thorough analysis of the available data was conducted in order to obtain some first insight for the behaviour of the natural system and to identify a reasonable parameter space for the Monte-Carlo experiment that will be formulated later.

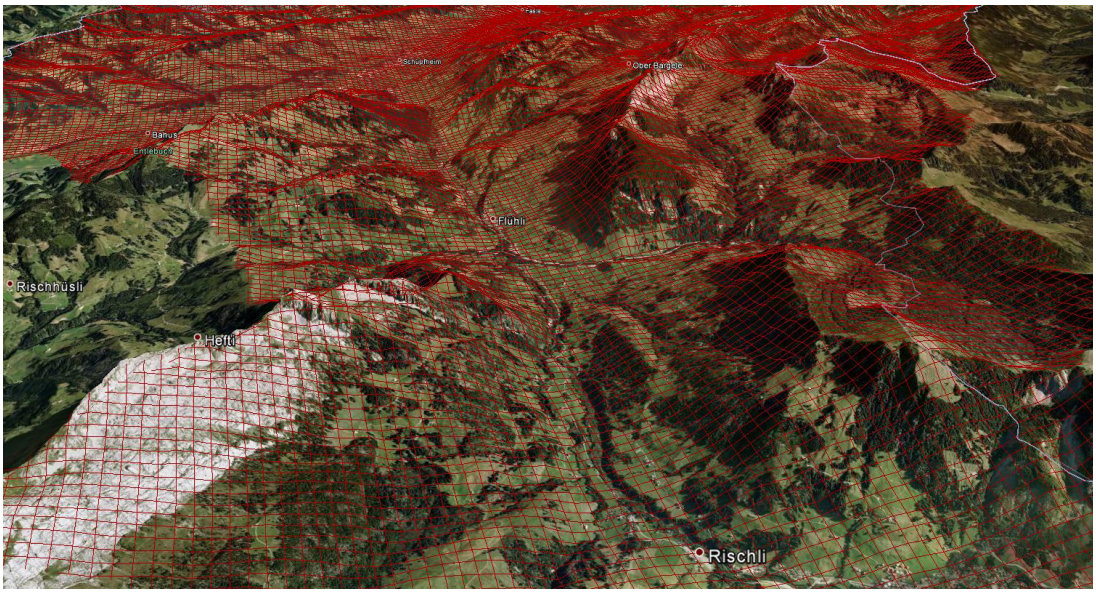


FIGURE 7.2: Spatial discretisation of the Kleine Emme river basin. Satellite imagery source (Google Earth).

The traditional extreme value analysis (Block maxima approach) is shown in figure 7.6 (a-c) for the three SwissMetNet stations laying in the catchment, in order to investigate the precipitation distribution in the examined area. Even though the extent of the area is small, it can be clearly seen that the local behaviour of the extremes can be very different, possibly due to localized orographic effects. The same analysis for the entire catchment area is shown in 7.6(d).

The estimation of the catchment hourly accumulation depth was done with the combination of two rain-gauge products. The first one (RhiresD) [*Frei and Schär*, 1998; *Schwarb*, 2000] is a gridded data set of daily accumulations (2×2 km^2), and the second one is hourly precipitation data from the Swiss MetNet stations located into the catchment. The hourly spatial distribution of precipitation across the catchment was defined as the temporal disaggregation of the daily depths according to the hourly rain-gauge data. In detail, the nearest station (Theissen polygons) was used for defining the hourly partition of the daily aggregated precipitation contained in RhiresD. The major advantage of such a methodology is that it can take into account explicitly the orographic effects of precipitation, as introduced in the RhiresD data set.

It is also important to derive the precipitation statistics on an event basis. The processes that are of major importance are the spatio-temporal structure of rainfall and storm motion. The high resolution radar data are used here for this purpose for the time period (Jan 2004 - Oct 2010). The precipitation events were identified with the procedure described in section 6.2.6. The cumulative probabilities of precipitation depth accumulations and durations are shown in figure 7.7

For the identified events, the storm velocities were also estimated according to a correlation matching algorithm identical to *Pegram and Clothier* [2001b]. Simulations using the TE model were conducted for the catchment for the 2000-2009 period, and the soil moisture conditions were derived from the model output, since no data were available. The precipitation forcing in this case was the combination of the RhiresD and the Swiss MetNet rain-gauge data, as previously described.

The correlations between hydrological variables that are expected to appear in such an analysis are between the initial soil moisture conditions, the precipitation depth during an event, and the generated discharge. In terms of discharge two variables are investigated. The magnitude of the peak and the time to

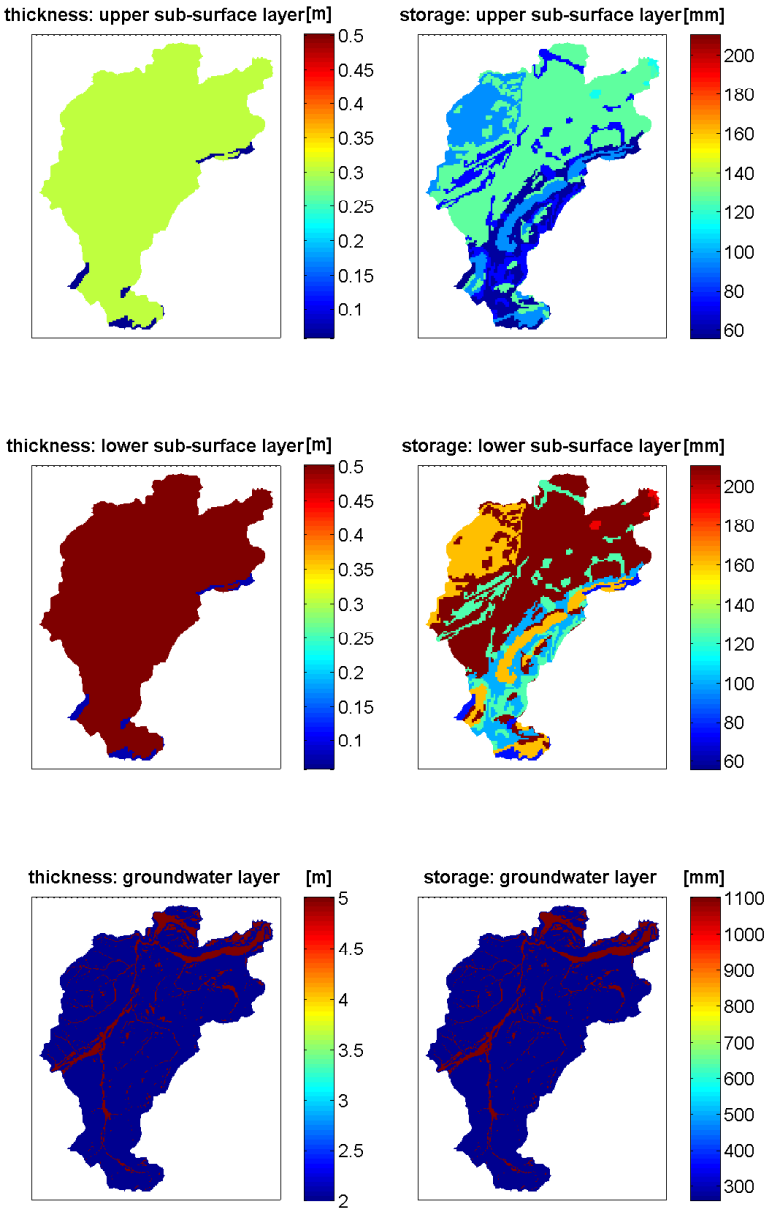


FIGURE 7.3: Vertical discretisation of the soil into a cascade of 3 vertical layers.

peak. It is not always straightforward to identify which part of the discharge is generated from a specific storm and thus a selection had to be done. The selection criteria were that 10 hours before the initiation and 10 hours after the end of the storm, no precipitation should occur. Taking into account that $\sim$10 hours correspond roughly to the concentration time of the Kleine Emme catchment, this assures that the contribution from previous rainfall events to the streamflow is negligible. Since all the major flooding events occur during warm seasons, only the months from May to September were analysed here. The peak discharge was defined as the maximum observed discharge at the basin outlet after the initiation of the storm and the time to peak was defined as the time between the center of mass of the storm hyetograph

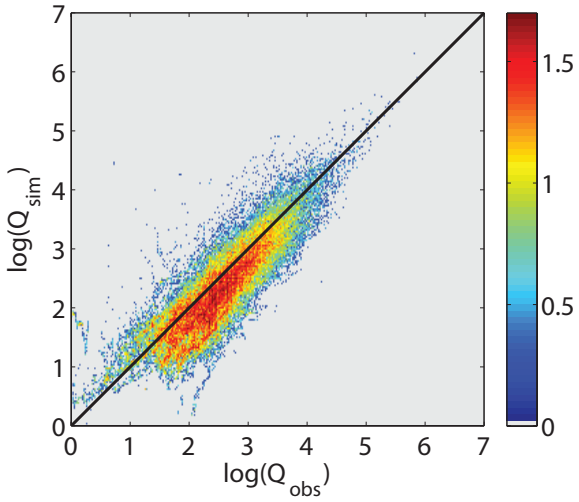


FIGURE 7.4: Scatter plot of the observed and simulated discharges at the outlet of the Kleine Emme catchment for the time period 2000-2009 and their corresponding bi-variate pdf.

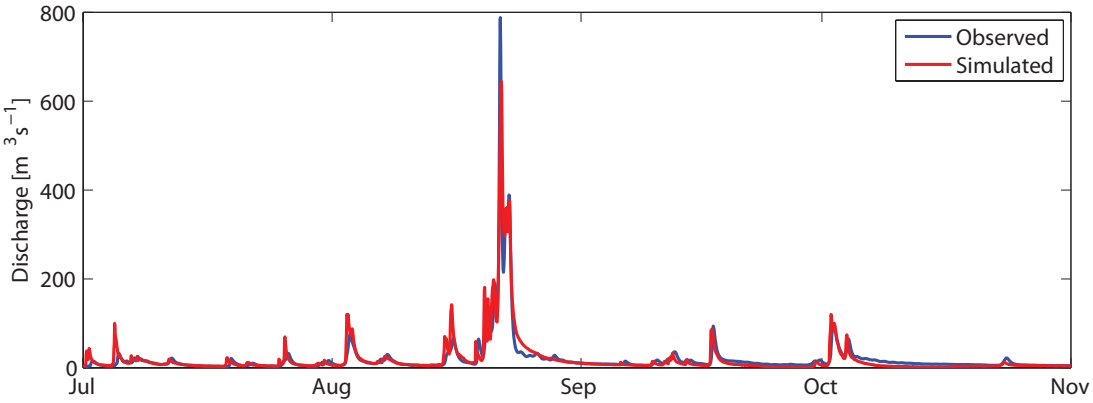


FIGURE 7.5: Simulated versus observed hydrographs at the outlet of the Kleine Emme catchment for the time period July-Nov 2005.

and the time of the peak. The center of mass of the event hyetograph is defined as

$$t_c = \frac{1}{T} \sum_{i \in T} P_i t_i, \tag{7.1}$$

where T the storm duration and P_i the mean areal precipitation depth for the corresponding time t_i.

The differences that can be identified concerning the depth accumulations shown in figures 7.7 and 7.8 can be attributed to the discrepancies that exist between the raig-gauges based data and the radar data. Generally the raw radar data underestimate precipitation, which is also the case here. However, the depth accumulations of gauge-based data are taken into account in figure 7.8 since they are expected to more and accurate.

From the analysis shown in figure 7.8 no clear conclusions could be drawn. Larger discharges in terms of peak response are associated with higher depth accumulations and with higher initial soil moisture con-

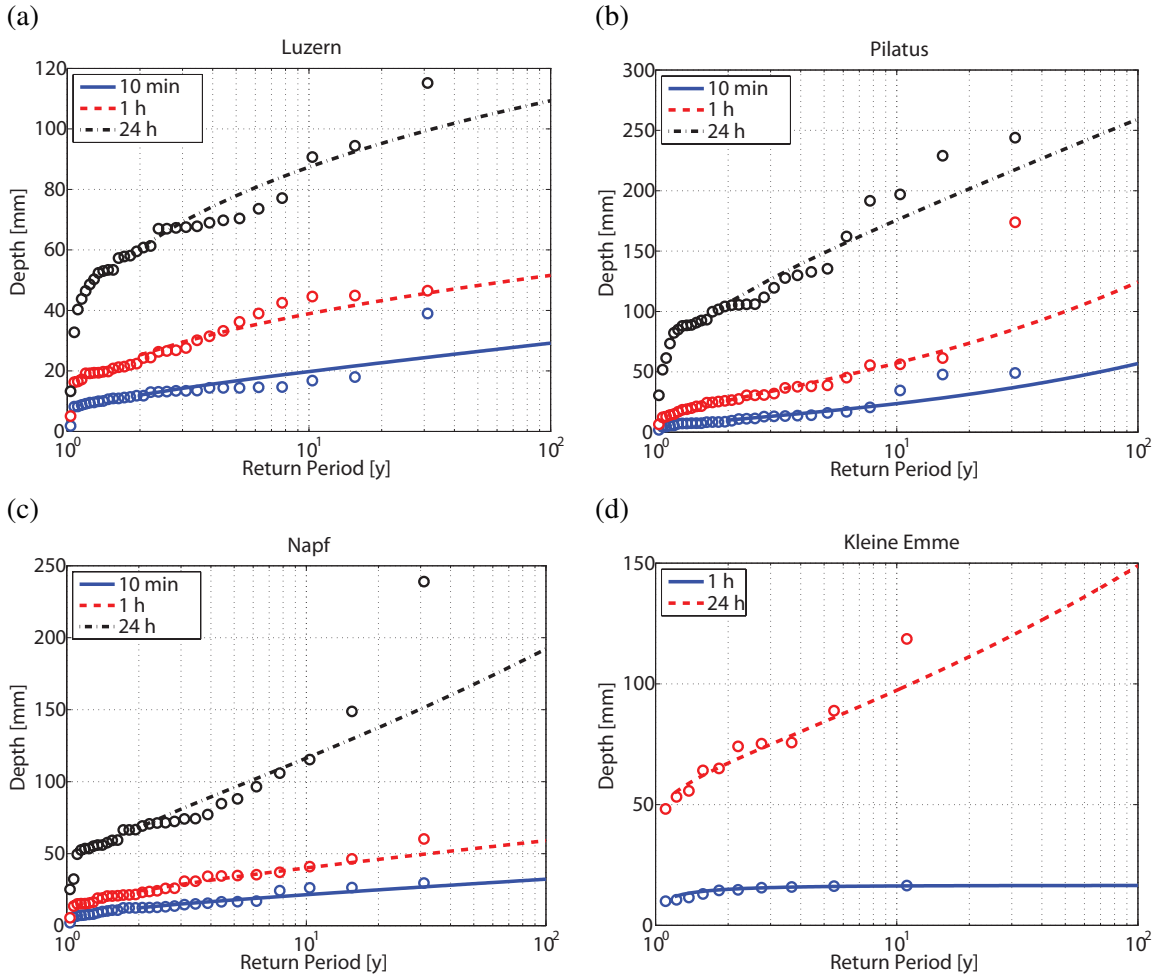


FIGURE 7.6: Analysis of the extremes for the 3 SwissMetNet (a-c) rain gauges in the Kleine Emme river basin; and (d) the overall catchment. The dots correspond to the observed values and the lines to the fitted Generalized Extreme Value distribution per aggregation interval.

ditions. There is also a tendency of shorter response times with increasing precipitation depth. However such a relationship is not statistically significant. The probable reason for this is the inherent non-linearity and feedbacks between the hydrological processes that play a role in runoff generation. Such feedbacks could be for example the spatial heterogeneity of ET, the shape of the hyetograph, the timing of snow melt, etc.

7.6 Numerical Experiment Set-up

A numerical Monte Carlo experiment is constructed to quantify the impact of the spatio-temporal structure of precipitation on the basin response in a controlled setting. A schematic description of the Monte Carlo experiment is shown in figure 7.9. The parameters that are investigated here are:

- The storm accumulation depth
- The spatial correlation of the precipitation fields

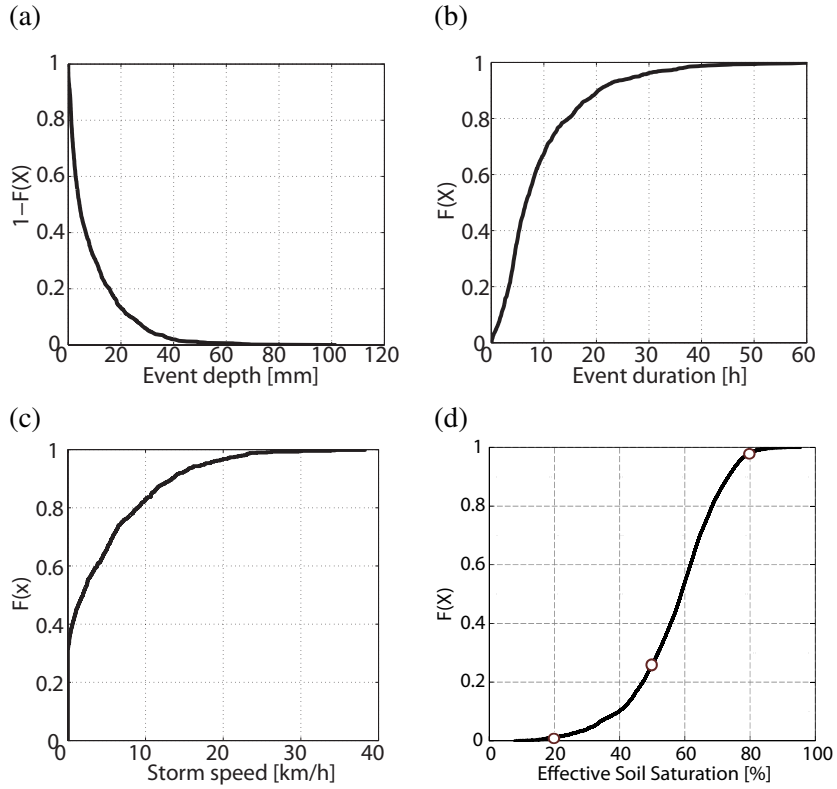


FIGURE 7.7: (a) Exceedance probability the event accumulation depth. (b) Non-exceedence probability event duration. (c) Non-exceedance probability for the storm speed. (d) Non-exccedence probability for the average upper layer effective water saturation.

- The spatial coefficient of variation of precipitation
- The storm advection
- The initial soil saturation conditions

Simulated rainfall fields by the STREAP model (see chapter 6) on an event basis are given as input to the TE model in order to asses their impact on the generated hydrographs. In all the cases the precipitation forcing is only rainfall. On a storm basis, ET is very small and is neglected here, and the catchment is assumed to be snow free in order to avoid any contribution from snow melt in river discharge.

The temporal resolution of the simulated rainfall fields is 5 minutes and the spatial resolution 1×1 km^2. Those resolutions were chosen since they correspond to the finest scale of information that is available in time and space as measured by the weather radars. In principle, the finest spatial scale of information derived by the RAIN product is 2×2 km^2, but from a preliminary comparison that was conducted with the new operational product of MeteoSwiss released on 2011 and delivered on a 1×1 km^2 grid, downsampling the simulations by a factor of 2 does not introduce significant errors.

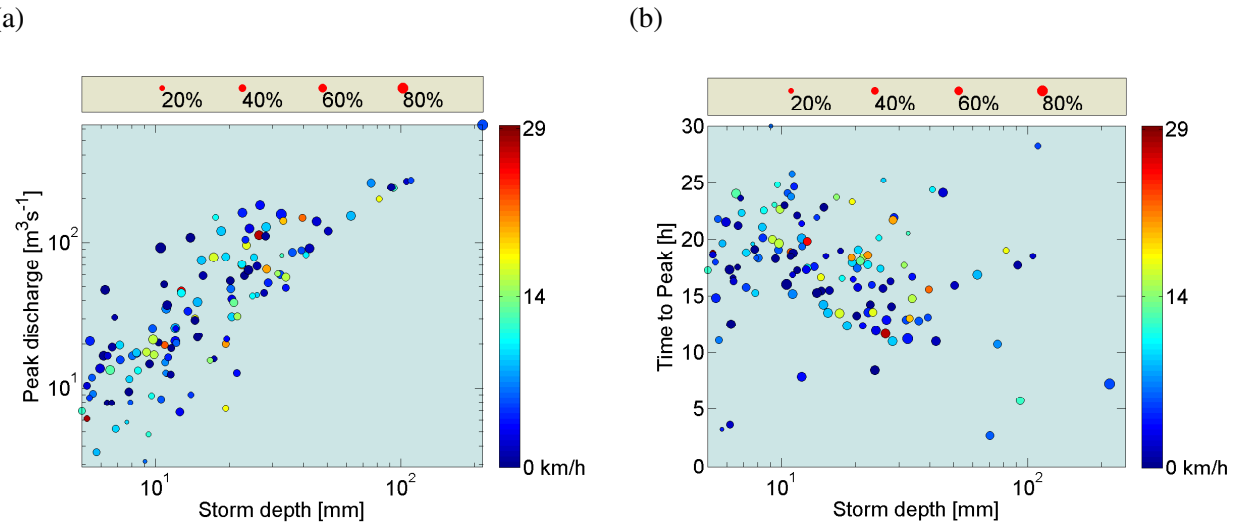


FIGURE 7.8: Peak discharge and time to peak as a function of the storm depth, soil saturation conditions and storm velocities for the Kleine Emme catchment. (a) The peak discharge; (b) the time to peak. Different colours correspond to different storm velocities according to the respective legend. The marker sites depend on the initial soil effective saturation of the upper soil layer according to the corresponding (upper) legend.

7.6.1 Storm Structure

The storm structure for the numerical experiment was selected in a way that corresponds to a realistic storm evolution. Given the relatively small area of the catchment relative to the extent of typical storms, the entire area is assumed to be covered by precipitation (i.e. WAR(t)=1). This simplification was chosen in order to restrict the degrees of freedom of the experiment, and also for intense storms.

The mean areal intensities (IMF) during a storm follow a Gaussian shape in order to have both a rising and a decaying part that is a good approximation of the typical storm evolution (see figure 7.9). Two different peak intensities were chosen, resulting in depth accumulations of 50 mm/d and 100 mm/d that approximately correspond to the annual and the 10 year return period (see figure 7.6). Low intensity storms generate very low values of discharge that cannot cause any serious flooding. The storm duration was chosen to be 1 day (90% percentile, figure 7.7(b)). An illustration of the storm structure and the corresponding generated discharges are shown in figure 7.10.

Spatial rainfall is assumed to follow a lognormal distribution. The parameters of the distribution are estimated once the mean value and the coefficient of variation are known. The mean value is given by the storm evolution and the coefficient of variation is assumed to be 1 and 3. These values correspond to two reasonable values of low and high variation respectively. This choice may also be seen as a separation between stratiform and intense convective rainfall, since during winter, precipitation fields have a much lower spatial variability in comparison to the intense summer events. This is as well supported by the radar data.

The spatial correlation of the fields is imposed by the correlation of the Gaussian random fields that are exponentiated in order to get the rainfall fields. Two different spatial correlation structures were chosen with correlation lengths of 10 km and 30 km respectively. This choice can also yield some physical connections between different types of precipitation (e.g. stratiform and convective). The choice of the correlation length came from the radar data analysis (chapter 6).

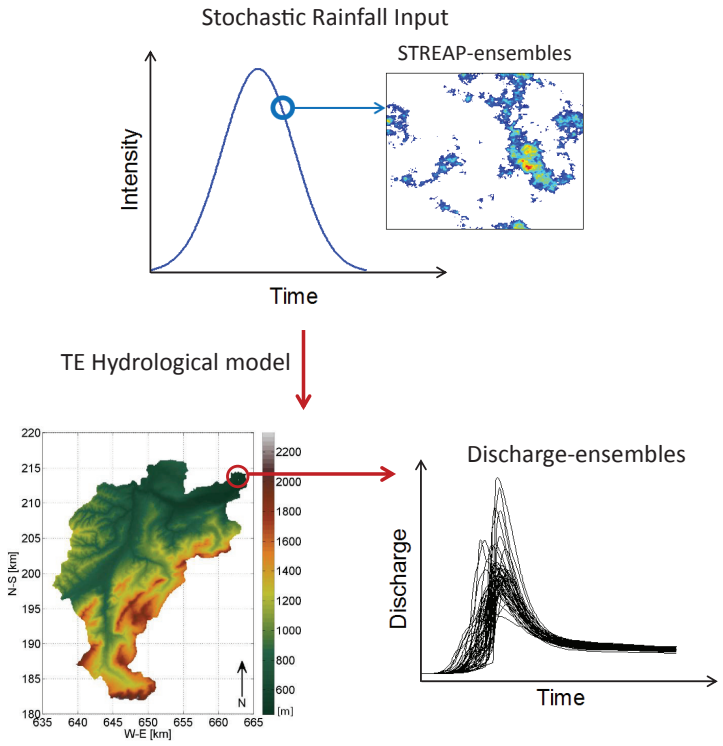


FIGURE 7.9: Schematic representation of the Monte Carlo numerical experiment for the quantification of rainfall variability on basin response.

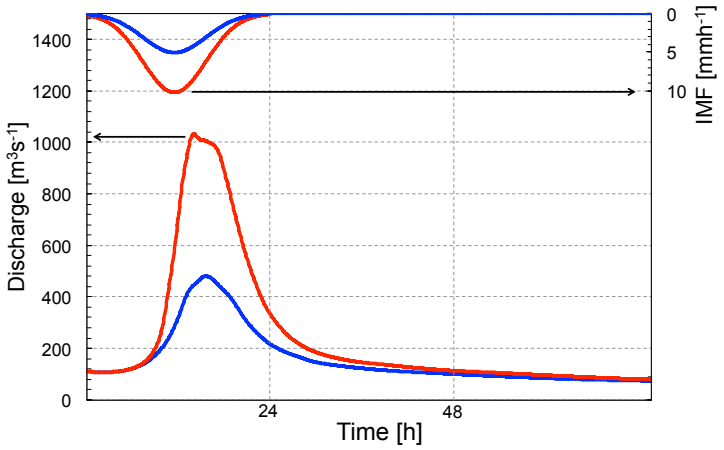


FIGURE 7.10: Illustration of the temporal evolution of the mean areal intensity and the respective generated discharges. The two cases reported here correspond to one realization of storms without advection, $cv = 1$, correlation length 10 km, and initial effective soil saturation $\theta_{ini} = 0.8$.

The rainfall fields were chosen to be either not moving (static) or moving with a constant advection velocity of 15 kmh^{-1}. Two sub-cases of moving fields were selected dependent on whether the storm direction is towards the outlet or in the opposite direction. The direction towards the outlet is defined as the one that links the geometric centre of the catchment and its outlet.

The temporal correlation of the fields in the Lagrange coordinate system was imposed by the ARMA

process of the Fourier coefficients (see equation 6.36). The coefficients were chosen identical as in the case study of the Monte-Lema area (see 6.3.1). The choice of only one correlation pattern in Lagrange coordinates does not pose the restriction that the generated precipitation fields have the same temporal correlations in the Eulerean coordinate system. The rainfall fields that are advected have significant lower temporal correlation in the Eulerean coordinate system. Therefore by imposing the field advection, also the temporal correlation is investigated without having to increase the degrees of freedom of the numerical experiment.

For each one of the cases described above, 50 storm realizations were simulated. The sample size was mainly chosen due to the computational requirements of the TE model. A summary of the parameters of the Monte Carlo experiment can be found in the table 7.4.

TABLE 7.4: Parameter space of the Monte Carlo numerical experiment

Parameter		Value	
Peak storm intensity [mmh^{-1}]	5	10	
Coefficient of variation of spatial rainfall [-]	1	3	
Correlation length [km]	10	30	
Storm velocity (towards the catchment outlet) [kmh^{-1}]	0	15	-15
Initial soil moisture of the upper soil layer [-]	0.2	0.5	0.8

7.6.2 Initial Soil Moisture Conditions

In order to get a realistic pattern of the spatial distribution of the initial soil moisture conditions in the basin, a methodology similar to *Noto et al.* [2008]; *Vivoni et al.* [2009]; *Moreno et al.* [2012] was followed. The basin is initially assumed to be completely saturated. Then it is allowed to drain without any precipitation forcing. The three states of the simulation that were chosen are when the mean value of the upper soil layer effective soil saturation was 80%, 50% and 20%. Those values correspond to two extreme cases and one regular case as estimated by the 10 year simulation described in section 7.5. The spatial distribution of the upper layer soil moisture is shown in figure 7.11.

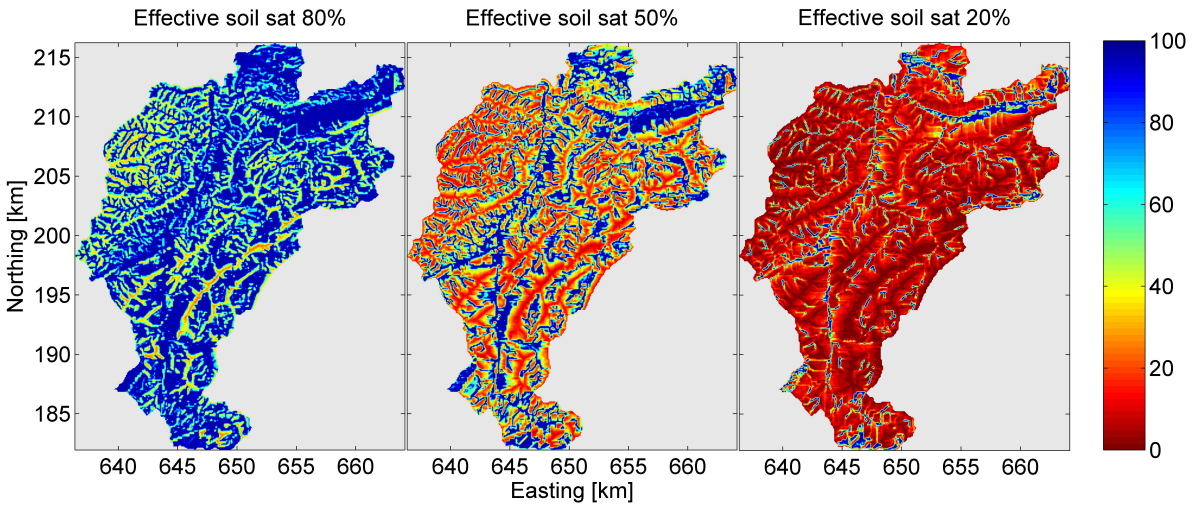


FIGURE 7.11: Spatial distribution of the initial upper layer effective soil saturation for the three initial soil moisture cases.

The spatial distribution of the soil moisture is mainly determined by the topography of the area since lateral subsurface flow is driven in the TE model only by gravity and ET is neglected. This simplification for obtaining the initial soil moisture conditions may yield some discrepancies between simulation and reality since processes such ET can modify the soil moisture spatial distribution, dependent on the topography and the vegetation cover of the catchment. However it is expected to give a reasonable approximation of the overall spatial distribution of soil moisture, for the specific scope of the analysis.

7.7 Results

The three key variables of flood discharges which are analysed here are:

- Peak discharge
- Time to peak
- Total discharged volume

The time to peak is defined as the time between the start of the storm and the time when peak discharge occurs. It is somewhat different from the notation introduced before, but since the centre of mass of the storm is always 0.5 days the two quantities are connected through a subtraction of the 0.5 days.

A first visual result for the effect of the spatio-temporal variability on river discharge is shown in figure 7.12 where the generated hydrographs at the basin outlet for various cases of rainfall input are shown. Moreover, the discharge of the basin assuming a uniform rainfall intensity over the entire catchment, which is a common assumption in catchment hydrology, is shown. It is clear that there exist cases where the small scale variability of rainfall can significantly influence the basin response and lead to strong differences. In what follows, each of the three key variables of the runoff are thoroughly discussed.

Effect of Spatio-temporal rainfall Variability in Peak Discharge

As shown in figure 7.13, the peak discharge, both in terms of magnitude and in terms of variability, are highly dependent on almost all the parameters analysed here. First the most trivial result is that higher storm intensities yield stronger peaks of the hydrograph for all the cases.

For lower initial saturation conditions, non-moving (static) storms generate much higher peaks on average and also much more variable basin responses. This behaviour can be attributed to the mechanisms that trigger runoff initiation. When the storm is not moving areas with local intense precipitation persist in time and space. This can lead to rapid saturation of smaller connected areas in the catchment and subsequently lead to higher discharges due to the locally higher runoff coefficients. The localization of the intense storm cells lying close or far from the basin outlet leads to high variability between realizations, since the travel times of the water in the river network can be highly variable. On the other hand, storms that are moving have lower temporal correlation and there is no time for a storm that is advected to saturate localized areas in the basin. This leads to a behaviour much closer to the uniformly distributed rainfall across the entire area (see figure 7.12).

The spatial correlation does not have a significant effect neither in terms of the magnitude of the peaks nor in their variability. A high coefficient of variation of the spatial rainfall increases both the mean value and the variance of the peak discharges for very low initial soil moisture conditions.

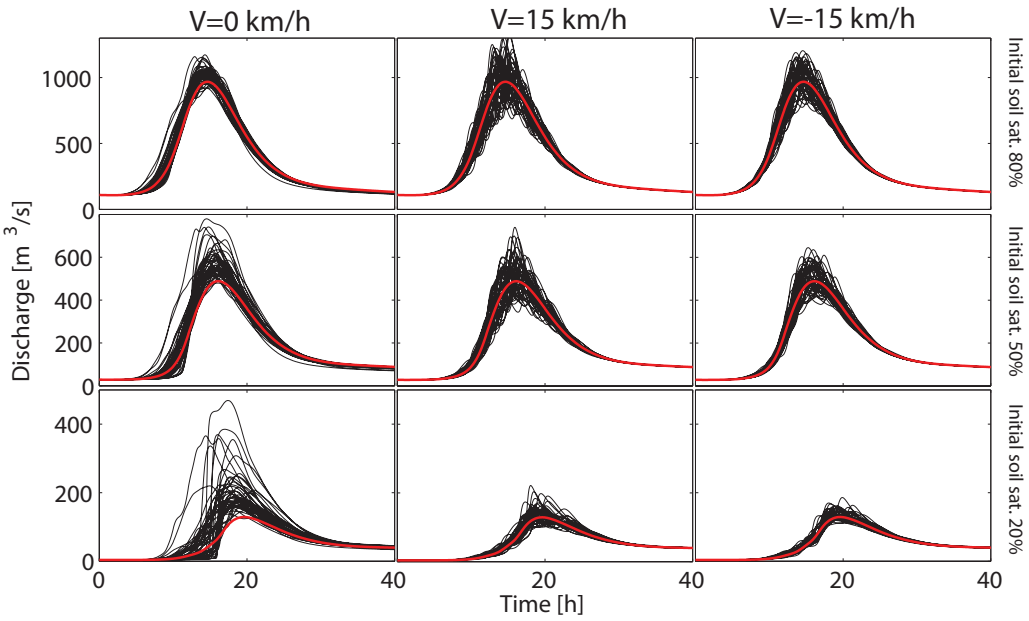


FIGURE 7.12: Simulated hydrographs dependent on the initial saturation ratio and the speed velocity. This case corresponds to peak storm intensity of 10 mm/h, spatial coefficient of variation 1 [-] and correlation length 30 km. Red lines illustrate the case where rainfall is uniform is space

As the initial soil moisture of the catchment increases, the differences between the generated hydrographs for all storm combinations are reduced. The variability between the generated responses is also reduced. Similar results were obtained by *Zehe and Blöschl* [2004] who found that larger rainfall intensities and higher initial saturation ratios, reduce the uncertainty of the generated hydrographs.

For the case of highly wet conditions, the storm direction seems to play a minor role, with storms moving toward the outlet leading to slightly higher and more variable peak discharges (figure 7.12, 7.13(c)). This result is not very significant probably due to the small number of storm realizations that were chosen.

These general results are also in line with the findings of both *Krajewski et al.* [1991] and *Shah et al.* [1996a] which strengthens the idea that temporal correlation along with the initial soil moisture conditions of the basin are the key variables for the correct estimation of the peak discharge.

Effect of Spatio-temporal rainfall Variability in Time to Peak Discharge

The time to peak is highly connected with the same variables as the peak discharge as shown in figure 7.14. Highly temporally correlated rainfall fields lead to shorter response times on average. This effect is stronger for low saturation levels. The reason for that is exactly the same as described above concerning peak runoff. Rainfall fields with strong temporal correlation initiate local overland flow that reaches the channel network much faster than the runoff generated by rainfall patterns that are weakly correlated in time. The higher variability originates from the fact that local saturated areas can exist on different parts of the catchment which affects the travel time to reach the channel network. The effect of enhanced variability is more obvious for low intensity events. The spatial coefficient of variation of rainfall has also the effect of slightly increasing the variability of the times to peak. In this case as well the spatial correlation of the rainfall fields does not affect the output significantly. The direction of the storm that could be potentially affect the time to peak does not have a significant impact.

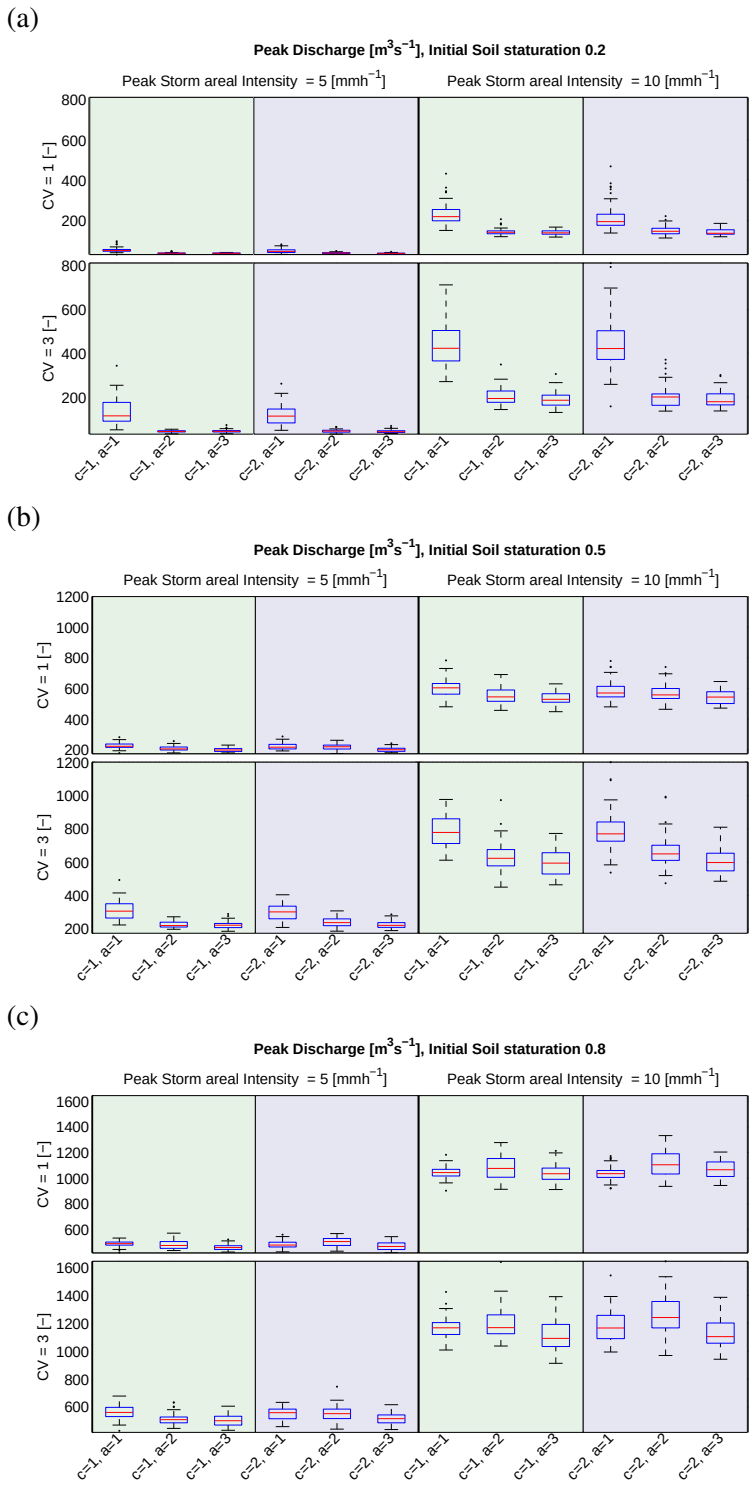


FIGURE 7.13: Boxplots of peak discharge. The three panels (a-c) correspond to different initial soil moisture levels as indicated on their titles. The two left columns of each plot correspond to peak intensity of 5 mm/h and the two right ones to 10 mm/h. Green plots correspond to coefficient of variation 1 and blue ones to 3. c corresponds to correlation length with 1 standing for 10 km and 2 for 30 km.a corresponds the speed advection towards the outlet with values 1, 2, 3 standing for 0,15,-15 km/h respectively.

Effect of Spatio-temporal rainfall Variability in Total Discharge Volume

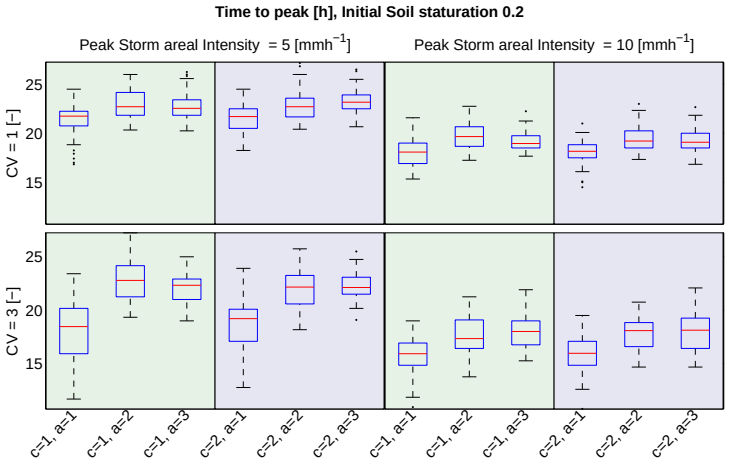
As far as the total discharge volume is concerned, also here the main key variables are the initial soil saturation conditions and the storm kinematics, closely related to the temporal variability of the storm. As shown in figure 7.15 non-moving storms lead to much higher total volumes. This behaviour is consistent regardless of the initial soil moisture conditions, but is more intense for dry conditions. Interestingly, the fact that for very high initial soil moisture conditions the peak discharges for all the storm combinations are almost the same, does not reflect in the total discharged volume.

Physical Explanation of the Results

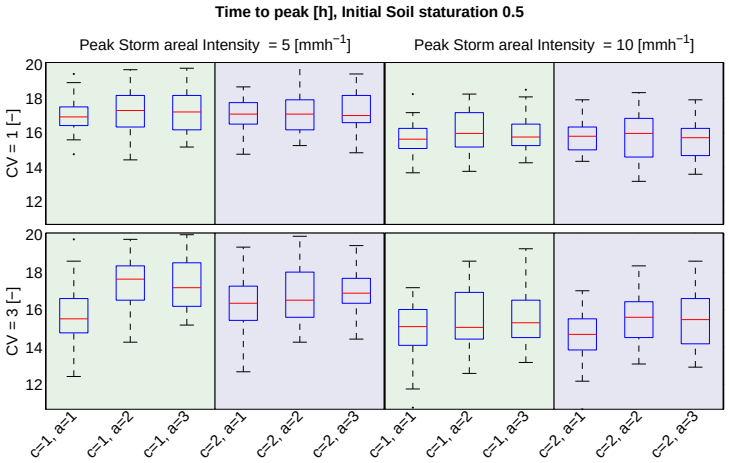
In order to illustrate and quantify in detail the reasons why the observed behaviour occurs, for one of the cases that was previously analysed, the spatial distribution of the effective soil saturation of the upper soil layer was analysed during the simulation period. The case corresponds to the storms with peak areal intensity of 10 mmh^{-1}, correlation length 30 km, coefficient of variation 1, for initial effective soil saturation of the upper layer of 20% and 50%.

Figure 7.16 shows the spatial distribution of the upper layer soil saturation ratio plotted for one realization of non-moving rainfall fields and the respective for a moving one with storm speed 15 kmh^{-1} towards the basin outlet. The figure corresponds to the soil moisture maps after 20 hours from the storm start with average initial soil saturation 20%. From figure 7.16 it becomes clear that due to the storm persistence in time, saturated areas tend to be localized in space, a fact that leads to higher and earlier peak discharges, exactly as shown in the simulations.

(a)



(b)



(c)

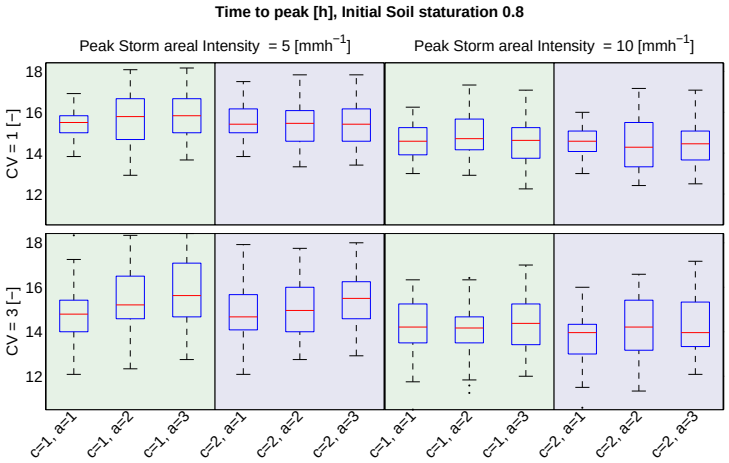
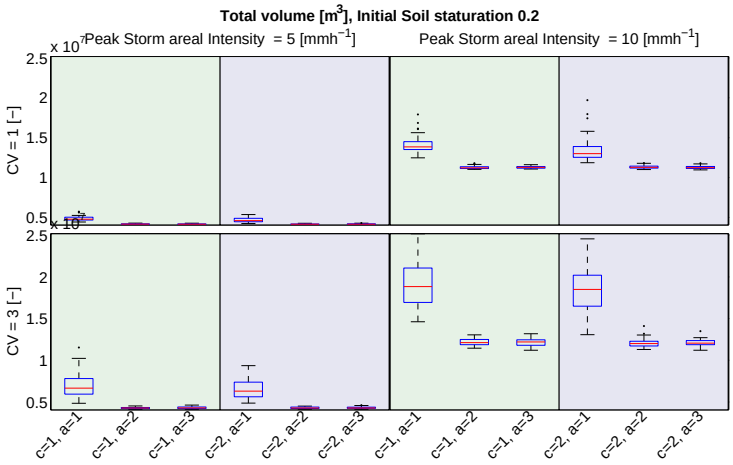
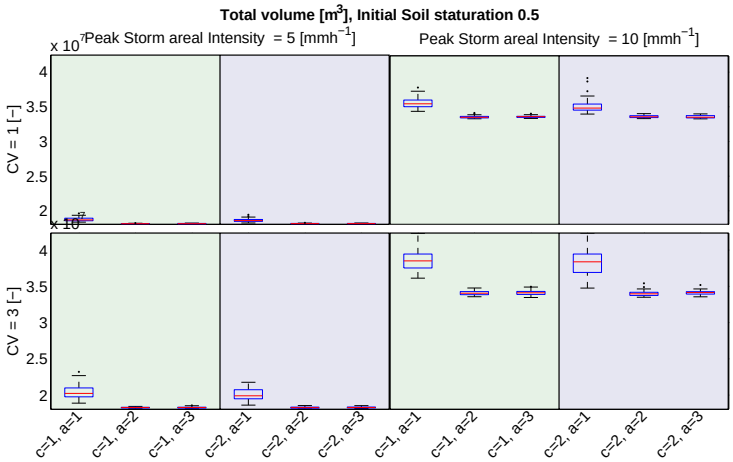


FIGURE 7.14: Boxplots of time to peak discharge. The plot description is the same as in figure 7.13

(a)



(b)



(c)

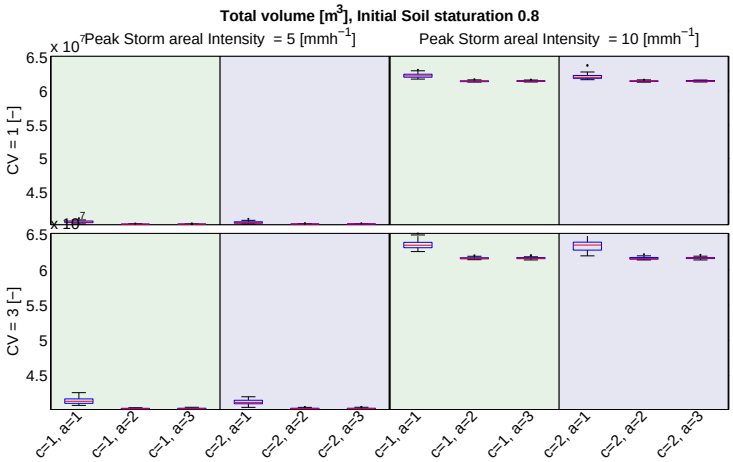


FIGURE 7.15: Boxplots of total volume. The plot description is the same as in figure 7.13

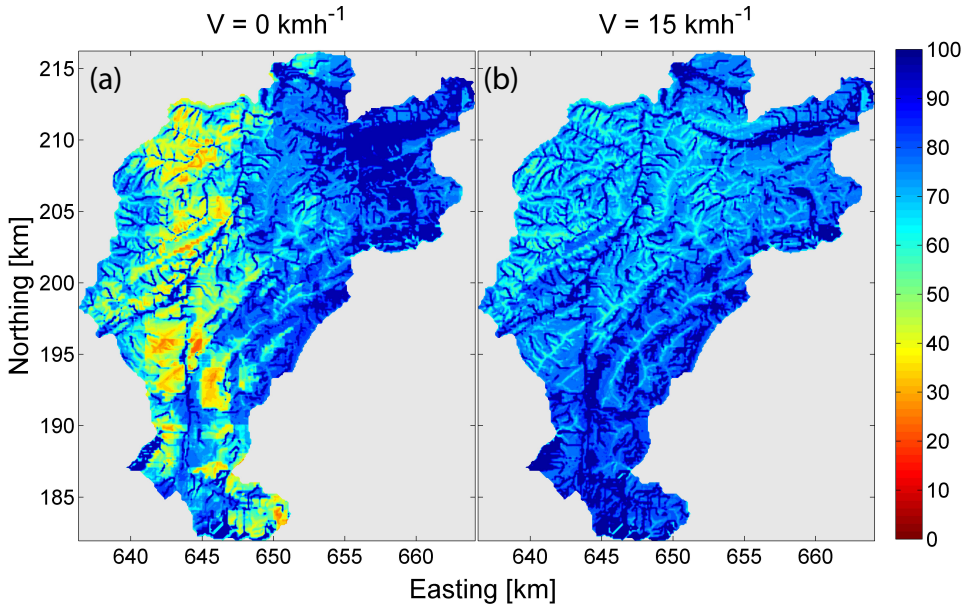


FIGURE 7.16: Effective soil saturation of the upper soil layer for initial soil saturation 20 % and 20 h after the storm start. The left panel (a) corresponds to the a non-moving storm and the right one (b) to a storm moving with 15 km/h towards the basin outlet.

In order to quantify the behaviour of saturation excess runoff due to higher temporal correlation for the case of static storms, the spatial coefficient of variation of the upper layer soil moisture is shown in figure 7.17 for the ensemble of the 50 generated storms. The common behaviour is that at the beginning of the event, the cv is high showing the spatial variability of the initial soil moisture field, the spatial distribution of which is dominated by drainage due to gravity. During the storm, the soil moisture of the catchment gradually increases and the differences in the spatial distribution are reduced, which leads to the consistent decrease of the cv. After the end of the storm, the spatial distribution of soil moisture tends to its initial condition due to gravity drainage. This soil moisture redistribution leads to an increase of the cv.

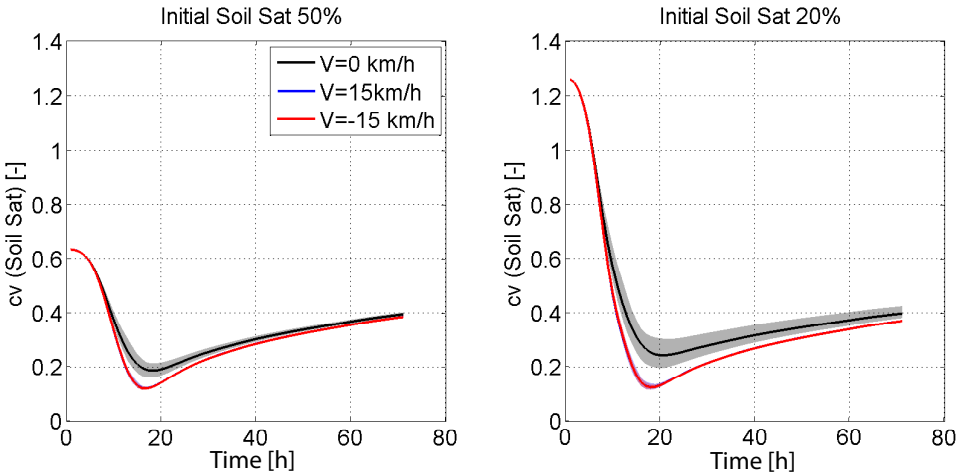


FIGURE 7.17: Spatial coefficient of variation of the upper soil layer. Different colours correspond to different storm kinematics. Shaded areas represent the 95% confidence intervals of the simulation ensemble and lines illustrate the mean value.

The localized behaviour of the storms that are non moving is identified by the fact that the value of cv is consistently higher during the storm period. This clearly shows that specific areas with high soil moisture are saturated and generate local high values of runoff, due to the persistence of intense precipitation in these areas. Clearly this behaviour is more intense for low initial saturation conditions due to the heterogeneous initial conditions. A similar pattern can be identified in figure 7.18 where the dependence of the spatial cv to the mean value of the effective saturation is shown for the cases examined here. In all the cases, the cv of the soil moisture was higher and much more variable for the cases of not-moving storms.

The behaviour of the moving fields is much closer to the one of the storms with uniform intensity across the entire catchment area. To illustrate this in figure 7.19 the standard deviation of the differences between the ensemble of the storms and the uniform intensity case are shown. The differences of the deviations are very strong both in terms of magnitude and variability. In principle, the storms that are moving with the moderate speed of 15 kmh^{-1} behave very similarly to the uniform intensity case. The differences between the storms with opposite directions are negligible (see figures 7.16, 7.19 and 7.17.).

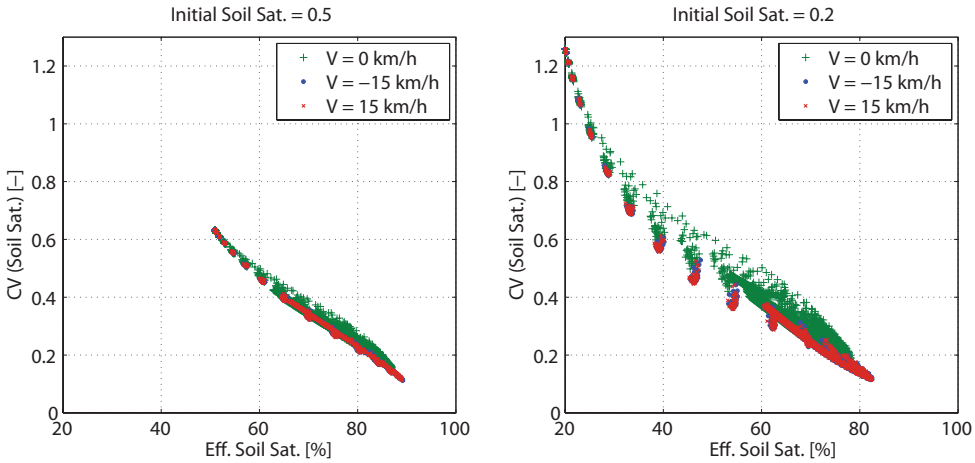


FIGURE 7.18: Dependence of the spatial cv on soil moisture to the average upper layer effective soil saturation.

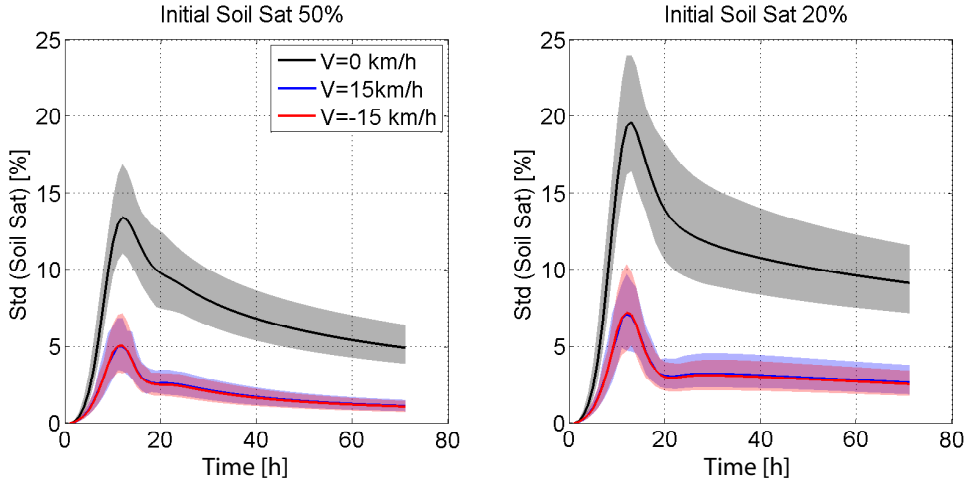


FIGURE 7.19: Spatial standard deviation of the soil moisture of the upper soil layer of the difference between the ensemble on the generated storms and the uniform intensity case. Different colours correspond to different storm kinematics. Shaded areas represent the 95% confidence intervals of the simulation ensemble and lines illustrate the mean value.

The fraction of the areas close to saturation (effective soil saturation >99%) in time and the time of saturation were estimated (see figures 7.20 and 7.21). It was found that the differences between all the cases analysed here are not significant. Thus taking this into account, it appears that the major reason why non-moving persistent storms generate higher discharge peaks and shorter response times, is the fact that locally connected areas are simultaneously saturated and generate locally higher values of runoff that go fast into the channel network. A similar result was also obtained by *Merz and Plate* [1997], *Merz and Bárdossy* [1998] and *Minet et al.* [2011], that found that when well structured soil moisture patterns exist, the discharge volumes and peak discharges are typically higher. On the other hand the generality of this result could be challenged since recently *Morbidelli et al.* [2012] found that the spatial variability of the soil moisture field does not have a high impact on the correct representation of the generated hydrographs.

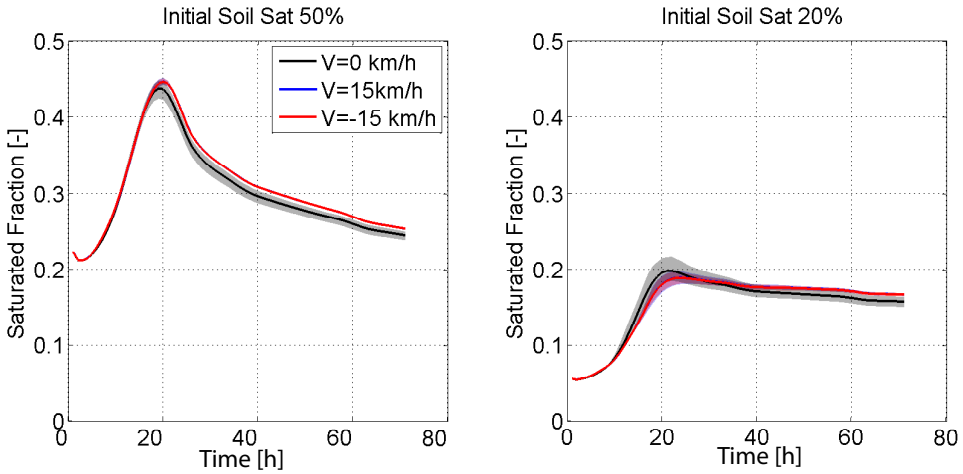


FIGURE 7.20: Fraction of the saturated upper soil layer. Different colours correspond to different storm kinematics. Shaded areas correspond to the 95% confidence intervals of the simulation ensemble and lines illustrate the mean value.

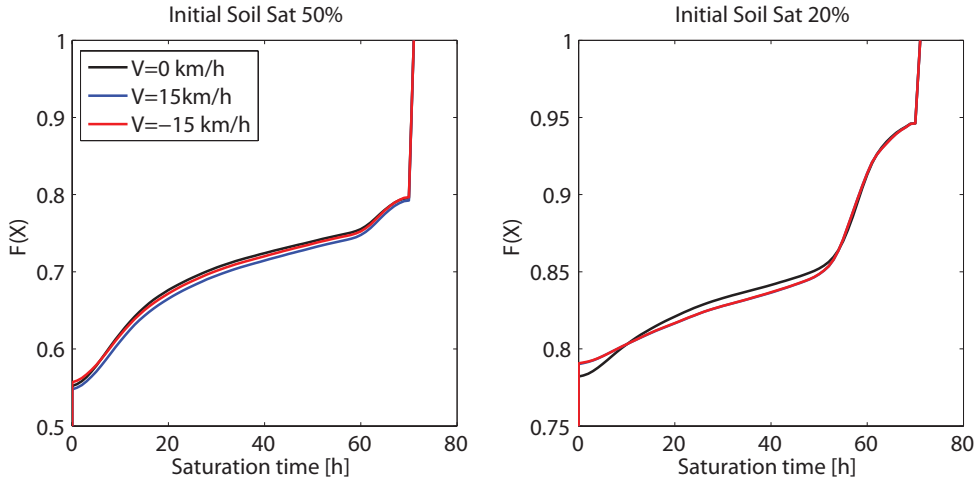


FIGURE 7.21: Cumulative distribution function of the time a pixel stays saturated.

In order to better understand the effect of locally connected saturated areas, a connectivity index for the saturated areas was defined. This index distinguishes between the effects of the spatial distribution of soil moisture due to topography and due to the rainfall structure. For a specific simulation, the areas close to saturation (>99%) were estimated together with the saturated areas of the same temporal step for the case of uniform rainfall. Then the two maps were subtracted. This subtraction leads essentially to the identification of the differences between the generated storms and the uniform case, where the spatial effective soil saturation patterns are dependent on the topography of the area. This can indirectly separate the effects of the catchment's topography and the ones dependent solely on the storm structure. From the subtracted maps, the areas of connected saturated area clusters are estimated. This gives a measure of how much more connected in space are the saturated areas for each simulation. From figure 7.22 it becomes clear that the differences between the moving and non-moving storms are large. Static storms generate larger connected saturated areas in locations different from those created by spatially uniform storms. This is a clear indication of the explanation that was previously provided concerning the reasons why the peak discharges are higher when a storm is static, and also why they occur faster on average. The spatial distribution of the localized areas, laying close or far from the basin outlet enhances the variability of the generated hydrographs as clearly shown in figure 7.12.

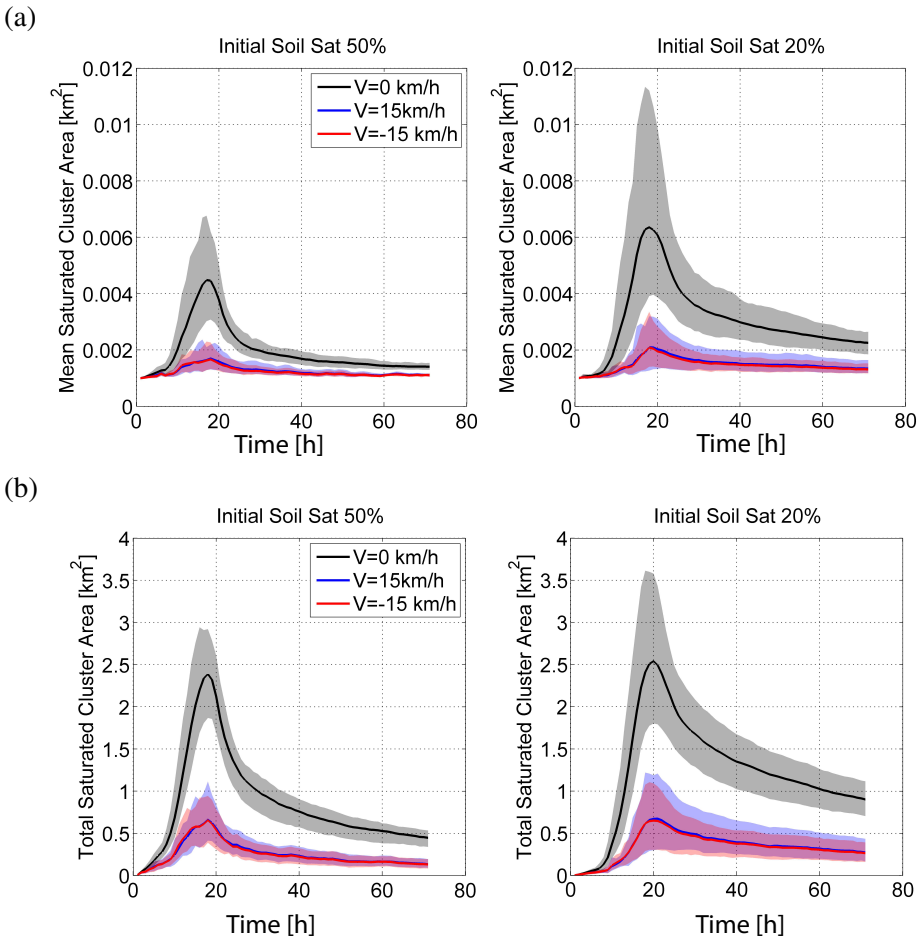


FIGURE 7.22: Temporal evolution of the sizes of the clustered saturated areas during generated storms. The areas correspond to the subtracted maps between the saturated areas of each stochastic realization and the equivalent one of the uniform spatial precipitation. The upper two panels show the mean size of the clusters and the lower two panels their total sizes.

The major conclusion from this numerical investigation is that the most crucial factor which influences the response of the basins is related to the storm velocity, which reflects the temporal correlation of precipitation. In order to illustrate this with a real world example, the storm that caused very strong floods in Switzerland during 21-25 August 2005 is used. In figure 7.23 the total depth accumulation for this period across the country is shown, and also the time that the local precipitation exceeded several intensity thresholds are shown. The data product for this analysis is again the combination of weather radars and rain-gauges [*Wüest et al.*, 2010]. During this event the areas that mostly suffered from severe flooding were in central Switzerland. Not surprisingly, this coincides with the locations where intense precipitation was persistent in time and thus intense precipitation for long durations resulted in very large depth accumulations (figure 7.23). This example cannot serve as a proof of the numerical investigation results, but it is an encouraging illustration of the validity of the reported results.

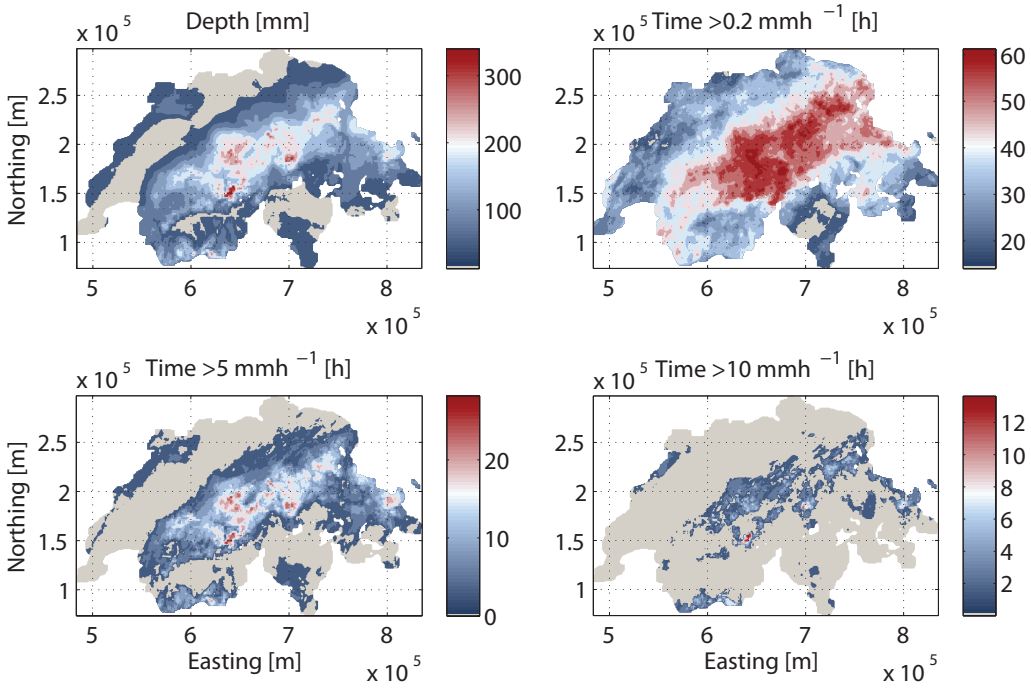


FIGURE 7.23: Precipitation depth accumulation in Switzerland during the period 21-25/8/2005 and the time for which precipitation exceeded several intensity thresholds.

7.8 Discussion and Open Questions

The study reported here, is comprehensive due to the large number of variables that were investigated. However it is also case specific and generalization of the results may be not straightforward.

First the effect of the catchment size was not explicitly taken into account. The reason why only one catchment was selected is the spatial resolution of the weather radars that are used to identify a reasonable parameter set for the STREAP model. If higher resolution data were available, the study could be potentially extended to catchments of smaller sizes that are common in the Alpine area and are prone to flash floods. Indeed the dampening effect of a catchment depending on its size has been thoroughly investigated before [e.g. *Nikolopoulos et al.*, 2011; *Nicotina et al.*, 2008; *Volpi et al.*, 2012] and can have a major impact. In this case, the reported results may be catchment dependent, but since the analysed area is representative of many Alpine basins that have approximately same sizes and similar morphological and geological features, the conclusions can be generalized to some degree for the area of the European Alps. The results are not directly transferable to areas, with substantially different structural characteristics (e.g. highly eroded flat basins with very different drainage density) with much larger or smaller sizes. Another feature that is also expected to have an influence on the effect of the spatio-temporal structure of precipitation on the basin response, is the topology of the channel network itself. The relative dimensions of the storm and the networks that route it to the catchment outlet may play a significant role for the dampening of the heterogeneities within the basin.

Second, the study is not only catchment dependent but also model dependent. It is true that the investigation conducted in this study is a sensitivity analysis of the TE model that is assumed to represent adequately the physical processes in the catchment. The only processes that play a major role here are the surface and subsurface water routing. Since water routing is done, as mentioned, with the kinematic wave approach, which is a simplification of the shallow water equations, and the examined basin is steep enough, the kinematic wave approximation will likely not create any major artefacts. The main problems arise from the approximation of the subsurface flow. However, the fact that the model is reproducing very well the discharge at the outlet for a 10 year calibration period (2000-2009) is an encouraging fact that supports the validity of the results.

Moreover, since the analysis is based on the event scale, neglects other hydrological processes such as ET, snow-melt etc. Those processes, as already mentioned, can have a strong impact on the initial state of the catchment during a flood event.

In addition, another parameter that was not extensively examined in this experiment is the storm duration that was assumed constant for all the cases. Strongly connected with that is also the choice of the temporal evolution of the mean areal intensities during the rainfall event. In general, the study that has been presented here is based on an "impulse response" concept similar to the notion of the most widely used tool in applied hydrology, the unit hydrograph. Relaxing the assumptions may provide further insights but will also result to more degrees of freedom in this experiment.

Another parameter, that was probably treated in a simplified manner is the storm velocity. It indeed appeared to be one of the most influential parameters for the variability of the basin response. The assumption of using a single, yet reasonable, storm speed velocity may be restricting the range of results. For example *Singh* [1997] found that the relative speed of the storm and the water routed through the channels can have a significant impact on the generated hydrographs. Future analysis of this topic can investigate a broader spectrum of storm speed velocities in order to quantify in more detail the effect of storm speed.

Finally the results are also dependent on the structure of the precipitation forcing, i.e. the STREAP model. Since in section 6.3 was shown that this model can capture well almost all the substantial structural features of the space-time storm evolution, the possible artefacts connected with this model selection are considered negligible. This is indeed one of the major novelties of this work in comparison to previous studies that have been using rainfall input from simpler space-time models that cannot capture several important features of the storm structure.

8

CONCLUSIONS

This study had three main goals. The first goal was to review statistical approaches to describe precipitation structure in space and time and to apply these to a large dataset of high resolution time series and radar data in Switzerland. A comprehensive assessment of commonalities. connected to the topographic complexity of the country is given. The second goal was to quantify limitations of stochastic models of precipitation, and on this basis develop new stochastic modelling tools which satisfy requirements of parsimony and at the same time reproduce the key features of precipitation across spatial and temporal scales. The final goal was to use the developed stochastic models for rainfall to quantify the effect of the space-time precipitation structure on basin response in a typical Alpine basin. The main results are summarized in this chapter

8.1 Precipitation Structure in Space and Time

The study showed that within some constraint precipitation can be analysed with scaling and correlation based methods, both in time and space. However the statistical descriptors and the ranges of their applicability were variable among the stations and radar sites in Switzerland. Moreover it was found that the measurement errors associated with basic precipitation measurement techniques can propagate into the estimation of various statistical estimators.

It was found that under specific circumstances the measurement errors can be dominant and lead to large discrepancies leading to erroneous interpretations of the structure of rainfall. For temporal precipitation recorded with tipping bucket rain gauges, the measurement errors caused by the lag of the tipping mechanism were found to be the crucial parameter that causes most of the discrepancies in the estimation of almost all the statistical descriptors. It was found also that the problems typically appear for temporal scales below one hour and are generally affecting low intensity precipitation events.

Following a similar concept, for the spatial precipitation derived from weather radars, it was also found that measurement errors can harm most of the descriptors of the statistical structure of spatial precipitation. For this case the most severe type of error was found to be the atmospheric clutter (e.g. large intensity values that appear generally as a result of the anomalous propagation of the electromagnetic waves that the weather radars transmit).

After quantifying possible errors a robust estimation framework was developed both for time-series and

radar measured precipitation and the extensive precipitation data base of Switzerland was analysed.

A main focus of the data analysis was oriented towards the identification of potential links between the statistical structure of the precipitation processes both in time and space and their respective driving mechanisms, with specific emphasis given to the orographic signatures of the European Alps.

As far as the *temporal precipitation* is concerned, several interesting results concerning the description of precipitation based on the theory of scale invariance were obtained. First it was found that the scaling parameters of precipitation are dependent both on the season and the location of the station. This reveals the influence of the various precipitation generating mechanisms and the orographic signature of the underlying topography on the statistical properties of precipitation. This complements earlier findings of *Molnar and Burlando* [2008]. The major advancement of the current study was the way parameter estimation was conducted, avoiding any potential influences of measurement errors.

One of the main novelties of this work is that several types of deviations of the self similarity assumption in time were identified. Distinct temporal scaling regimes were found, associated with the typical time scales of storm structure organizations ($\sim$1 hour) and the diurnal cycle ($\sim$ 1 day) confirming previous results [*Fraedrich and Larnder*, 1993]. Dependencies of the weight of the multiplicative random cascades were identified, putting limits on the universal applicability of the theory of multifractals in rainfall analysis. The weights were found to be dependent on the temporal scale, the intensity of the precipitation itself (elsewhere mentioned as large scale forcing [*Over*, 1995]) and also they were found to be correlated in time [*Cârsteanu and Foufoula-Georgiou*, 1996; *Paschalis et al.*, 2012].

Also for the *spatial precipitation* several interesting results concerning its structure were obtained by analysing 7 years of high spatial (2×2 km^2) and temporal resolution (5 min) radar data for two areas located on the Northern and Southern part of the the European Alps.

First of all, it was found that in general the spatial structure of precipitation patterns at the β-mesoscale can be characterized as isotropic, with some exception in the Mediterranean side of the Alpine mountain range, where preferential wind flow patterns cause air flow blocking and thus lead to anisotropic patterns. The scaling properties of precipitation spatial fields were found to be seasonally dependent and also dependent on their mean areal intensities. This result can be also potentially exploited for their stochastic simulation.

Some orographic influences on the statistical structure of spatial precipitation fields were identified. The relatioships to altitude was found to be very different for the two different sides (Northern and Mediterranean) side of the Alps. In the Mediterranean part the orographic signatures are substantially higher. The underlying topography was found to have an effect on the spatial organization of the precipitation fields. Generally, storms located on higher altitudes were found to have less steep spectral decays, which illustrates their lower correlation structure in space. This result applied to all the seasons.

In addition, several possible connections that have been previously reported in the literature concerning the scaling properties of precipitation and the thermodynamic instability of the atmosphere [*Perica and Foufoula-Georgiou*, 1996] were investigated. It was not able to confirm these results in this study possibly because of the different climatological conditions and topography of the area.

8.2 Stochastic Modelling of Precipitation

The study also focused on the stochastic modelling approaches for the simulation of precipitation. The aim was to compare existing methods and propose improvements.

For *temporal precipitation*, an extended stochastic model inter-comparison was conducted, with the goal to quantify the capabilities and problems of various stochastic modelling approaches for temporal precipitation. The traditional models based on Poisson-cluster properties were reviewed and novel strategies dependent on the combination of various stochastic models were developed. The performance of these combined models was found to be very good for diverse climates across the world, especially for the small temporal scales (e.g. 10 minutes).

The main conclusion was that the most parsimonious and easy to calibrate models were most robust, in the sense that they show a similar degree of applicability for all the stations. On the other hand more complex model structures dependent on subjective parametrizations can yield excellent results for the data that they have been developed but typically cannot be easily generalized.

For *space-time precipitation* a novel model named STREAP was developed in order to simulate the spatio-temporal variability of precipitation as captured from high resolution weather radars. The model is of higher complexity in comparison to traditional spatio-temporal modelling approaches but to the author's knowledge, it is one of the few models that can carry out simulations in continuous time, and not on an event basis, and can as well mimic most of the precipitation structural and dynamical characteristics. STREAP has been found to give good results for one of the most challenging areas in Europe, the Mediterranean site of the Alpine mountain range, which is prone to intense precipitation orographic enhancement. A comparison of STREAP and one of the most widely known space-time models based on Poisson processes concluded that STREAP performed substantially better. This comparison served to illustrate that even though Poisson cluster models for spatio-temporal precipitation can be an excellent tool for multi-site point precipitation tools, they fail to capture the full characteristics of the spatio-temporal statistical structure of high resolution precipitation. This can be a major deficiency if those models are used as forcing in distributed hydrological model for impact studies.

8.3 The Effect of Spatio-Temporal Precipitation Variability on Basin Response

Finally the effect of the space-time rainfall variability on the runoff generation was investigated through an extensive numerical experiment. Stochastic rainfall input with prescribed statistical properties simulated by the STREAP model were used as forcing to a detailed fully distributed hydrological model (TE).

The major conclusion was that the sensitivity of the basin response is highly connected to the statistical properties of precipitation and also to the initial soil moisture conditions of the catchment. For very high initial soil saturation levels the key factor that characterizes the basin response is the total water volume and the internal structure of the storm has generally negligible effects.

The situations where the internal spatio-temporal structure of rainfall was found to play a major role, is when the catchment has an initial state of low soil moisture conditions. In this case, the key parameter that enhances the variability of the basin response was found to be the correlation of rainfall in time. As shown in chapter 7 the temporal correlation of rainfall is also connected with the storm motion, with static storms giving rise to a higher temporal correlation. Highly correlated rainfall fields in time can lead to strong peak discharges, faster responses and much more variable hydrographs, in comparison to a spatially uniform rainfall event with the same total volume, that is a common assumption traditionally in hydrology. Taking this into account, if a reliable flood protection system has to be designed, deep knowledge of the high resolution space time structure of precipitation is crucial.

This study was restricted to a single catchment. This means that sensitivity of the response of a basin dependent on its size could not be investigated. Further investigation on this topic will reveal more about the filtering/dampening mechanisms that influence the runoff processes in a basin and more general results can be extracted.

8.4 Outlook for Future Research

This study has tackled various issues related to the description and simulation of the statistical structure of precipitation and its impact on flood generation and, to the author's opinion, contributed to the advancement on knowledge of this topic. Yet this study had several limitations, which did not allow generally valid conclusions to be made in some cases. However the lessons learned lead to the following outlook.

A major concern is for the *data availability*. Even though new monitoring technologies and computer storage capabilities have inflated the last decades the amount of data concerning precipitation, a precise knowledge of precipitation intensities for small temporal and spatial scales is far from been achieved. It seems that remote sensing techniques have a strong advantage in comparison to the traditional measuring practices, however a worldwide high quality precipitation network is still not existing and would be a major prerequisite in order to strengthen our knowledge both in meteorology and hydrology. The radar and satellite era have still a lot to contribute in this perspective.

A *generalized statistical theory* of describing the precipitation process across space and time despite good partial efforts still does not exist. Several caveats were demonstrate in this thesis for most of the theories that have been used for precipitation description. In my opinion no theory can be yet considered as universal and new ones have to be developed in the future. Furthermore, no such theory can serve as a successful research tool, if no clear connections between the physics of the precipitation process and its observed statistical structure can be demonstrated. This is still an open and challenging question that in order to be tackled will reply on data availability. A clear illustration of this is that major achievements concerning our understanding of the precipitation process were always connected with strong investments in monitoring networks (e.g. NEXRAD) or research specific missions (e.g. TRMM). If such knowledge can be achieved, then new universal, robust and precise models, both deterministic and stochastic, will be developed.

REFERENCES

Abry, P., and D. Veitch, Wavelet analysis of long range dependent traffic, *IEEE Transactions on Information Theory*, *44*(1), 2–15, doi:10.1109/18.650984, 1998.

Allard, D., Advances and Challenges in Space-time Modelling of Natural Events, *Statistics*, *207*, 141–164, doi:10.1007/978-3-642-17086-7, 2012.

Arneodo, A., G. Grasseau, and M. Holschneider, Wavelet transform of multifractals, *Physical review letters*, *61*(20), 2281–2286, doi:10.1103/PhysRevLett.61.2281, 1988.

Audit, B., E. Barcy, J.-F. Muzy, and A. Arneodo, Wavelet based estimators of scaling behavior, *IEEE Transactions on Information Theory*, *48*(11), 2938–2954, doi:10.1109/TIT.2002.802631, 2002.

Bacchi, B., and N. Kottegoda, Identification and calibration of spatial correlation patterns of rainfall, *Journal of Hydrology*, *165*(1-4), 311–348, 1995.

Badas, M. G., R. Deidda, and E. Piga, Orographic influences in rainfall downscaling, *Advances in Geosciences*, *2*, 285–292, 2005.

Baillie, R., Long memory processes and fractional integration in econometrics, *Journal of econometrics*, *73*(1), 5–59, doi:10.1016/0304-4076(95)01732-1, 1996.

Bárdossy, A., and G. Pegram, Copula based multisite model for daily precipitation simulation, *Hydrology and Earth System Sciences*, *13*, 2299–2314, doi:10.5194/hess-13-2299-2009, 2009.

Bartlett, M., The spectral analysis of point processes, *Journal of the Royal Statistical Society. Series B (Methodological)*, *25*(2), 264–296, 1963.

Bates, P., and A. De Roo, A simple raster-based model for flood inundation simulation, *Journal of Hydrology*, *236*(1-2), 54–77, doi:10.1016/S0022-1694(00)00278-X, 2000.

Becchi, I., E. Caporali, L. Castellani, and F. Castelli, Multiregressive analysis for the estimation of the spatial zero rainfall probability, in *Workshop on climate change and hydrogeological hazards in the Mediterranean area*, pp. 133–145, Colombella, Perugia, Italy, 1994.

Bell, T., A space-time stochastic model of rainfall for satellite remote-sensing studies, *Journal of Geophysical research*, *92*(D8), 9631–9643, doi:10.1029/JD092iD08p09631, 1987.

Bell, T., and P. K. Kundu, A study of the sampling error in satellite rainfall estimates using optimal averaging of data and a stochastic model, *Journal of climate*, *9*(6), 1251–1268, 1996.

Beniston, M., August 2005 intense rainfall event in Switzerland: Not necessarily an analog for strong convective events in a greenhouse climate, *Geophysical Research Letters*, *33*(5), 1–5, doi:10.1029/2005GL025573, 2006.

Benzi, R., G. Paladin, G. Parisi, and A. Vulpiani, On the multifractal nature of fully developed turbulence and chaotic systems, *Journal of Physics A: Mathematical and General*, *17*(18), 3521–3531, doi:10.1088/0305-4470/17/18/021, 1984.

Benzi, R., L. Biferale, A. Crisanti, G. Paladin, M. Vergassola, and A. Vulpiani, A random process for the construction of multiaffine fields, *Physica D: Nonlinear Phenomena*, *65*(4), 352–358, doi:10.1016/0167-2789(93)90060-E, 1993.

Berenguer, M., and I. Zawadzki, A Study of the Error Covariance Matrix of Radar Rainfall Estimates in Stratiform Rain, *Weather and Forecasting*, *23*(6), 1085–1101, doi:10.1175/2008WAF2222134.1, 2008.

Berenguer, M., D. Sempere-Torres, and G. G. Pegram, SBMcast - An ensemble nowcasting technique to assess the uncertainty in rainfall forecasts by Lagrangian extrapolation, *Journal of Hydrology*, *404*(3-4), 226–240, doi:10.1016/j.jhydrol.2011.04.033, 2011.

Bernardara, P., C. Demichele, and R. Rosso, A simple model of rain in time: An alternating renewal process of wet and dry states with a fractional (non-Gaussian) rain intensity, *Atmospheric Research*, *84*(4), 291–301, doi:10.1016/j.atmosres.2006.09.001, 2007.

Bernardara, P., D. Schertzer, E. Sauquet, I. Tchiguirinskaia, and M. Lang, The flood probability distribution tail: how heavy is it?, *Stochastic Environmental Research and Risk Assessment*, *22*(1), 107–122, doi:10.1007/s00477-006-0101-2, 2008.

Berne, A., and W. Krajewski, Radar for hydrology: Unfulfilled promise or unrecognized potential?, *Advances in Water Resources*, *51*, 357–366, doi:10.1016/j.advwatres.2012.05.005, 2012.

Beuchat, X., B. Schaefli, M. Soutter, and a. Mermoud, Toward a robust method for subdaily rainfall downscaling from daily data, *Water Resources Research*, *47*(9), 1–18, doi:10.1029/2010WR010342, 2011.

Blöschl, G., and M. Sivapalan, Scale issues in hydrological modelling: A review, *Hydrological processes*, *9*(3-4), 251–290, 1995.

Booij, M., Modelling the effects of spatial and temporal resolution of rainfall and basin model on extreme river discharge, *Hydrological sciences journal*, *47*(2), 307–320, 2002.

Bordoy, R., and P. Burlando, Bias correction of Regional Climate Model simulations in a region of complex orography, *Journal of Applied Meteorology and Climatology*, *In Press*, doi:10.1175/JAMC-D-11-0149.1, 2012.

Bougeault, P., et al., The MAP special observing period, *Bull. Amer. Meteor. Soc*, *82*(3), 433–462, 2001.

Box, G., and M. Jekins, *Time series analysis, Forecasting and control*, Holden Day, San Fransisco, 1970.

Bracewell, R., *The Fourier Transform and its Applications*, 3rd ed., McGraw Hill, San Fransisco, 2000.

Brissette, F., M. Khalili, and R. Leconte, Efficient stochastic generation of multi-site synthetic precipitation data, *Journal of Hydrology*, *345*(3-4), 121–133, doi:10.1016/j.jhydrol.2007.06.035, 2007.

Brutsaert, W., *Hydrology: An Introduction*, Cambridge University Press, Cambridge, 2005.

Buishand, T., Some remarks on the use of daily rainfall models, *Journal of Hydrology*, *36*(3-4), 295–308, 1978.

Burlando, P., and R. Rosso, Comment on " Parameter Estimation and Sensitivity Analysis for the Modified Bartlett-Lewis Rectangular Pulses Model of Rainfall " by S . Islam et al ., *Journal of Geophisical Research*, *96*(D5), 9391–9395, 1991.

Burlando, P., and R. Rosso, Effects of transient climate change on basin hydrology. 1. Precipitation scenarios for the Arno River, central Italy, *Hydrological processes*, *16*(6), 1151–1175, doi:10.1002/hyp.1055, 2002.

Burlando, P., A. Montanari, and R. Ranzi, Forecasting of storm rainfall by combined use of radar, rain gages and linear models, *Atmospheric research*, *42*(1-4), 199–216, doi:10.1016/0169-8095(95)00063-1, 1996.

Burton, A., C. Kilsby, H. Fowler, P. S. Cowpertwait, and P. O'Connell, RainSim: A spatial-temporal stochastic rainfall modelling system, *Environmental Modelling & Software*, *23*(12), 1356–1369, doi:10.1016/j.envsoft.2008.04.003, 2008.

Burton, A., H. Fowler, S. Blenkinsop, and C. Kilsby, Downscaling transient climate change using a Neyman–Scott Rectangular Pulses stochastic rainfall model, *Journal of Hydrology*, *381*(1-2), 18–32, doi:10.1016/j.jhydrol.2009.10.031, 2010a.

Burton, A., H. J. Fowler, C. G. Kilsby, and P. E. O'Connell, A stochastic model for the spatial-temporal simulation of nonhomogeneous rainfall occurrence and amounts, *Water Resources Research*, *46*(11), 1–19, doi:10.1029/2009WR008884, 2010b.

Carpenter, T. M., K. P. Georgakakos, and J. a. Sperfslagea, On the parametric and NEXRAD-radar sensitivities of a distributed hydrologic model suitable for operational use, *Journal of Hydrology*, *253*(1-4), 169–193, doi:10.1016/S0022-1694(01)00476-0, 2001.

Cârsteanu, A., and E. Foufoula-Georgiou, Assessing dependence among weights in a multiplicative cascade model of temporal rainfall, *Journal of Geophysical Research. D. Atmospheres*, *101*(D21), 26,363–26,370, doi:10.1029/96JD01657, 1996.

Chainais, P., R. H. Riedi, and P. Abry, On non-scale-invariant infinitely divisible cascades, *IEEE, Transactions on Information Theory*, *51*(3), 1063–1083, doi:10.1109/TIT.2004.842570, 2005.

Chainais, P., R. H. Riedi, and P. Abry, Infinitely divisible cascades, *Tech. Rep. 12*, 2007.

Chambers, M., The simulation of random vector time series with given spectrum, *Mathematical and computer modelling*, *22*(2), 1–6, 1995.

Chebaane, M., J. Salas, and D. Boes, Product periodic autoregressive processes for modeling intermittent monthly streamflows, *Water Resources Research*, *31*(6), 1513–1518, doi:10.1029/95WR00144, 1995.

Chin, E., Modeling daily precipitation occurrence process with Markov chain, *Water resources research*, *13*(6), 949–956, 1977.

Chorti, a., and D. Hristopulos, Nonparametric Identification of Anisotropic (Elliptic) Correlations in Spatially Distributed Data Sets, *IEEE Transactions on Signal Processing*, *56*(10), 4738–4751, doi: 10.1109/TSP.2008.924144, 2008.

Chow, V., D. Maidment, and L. Mays, *Applied Hydrology*, McGraw Hill, New York, 1988.

Christakos, G., Stochastic simulation of spatially correlated geo-processes, *Mathematical Geology*, *19*(8), 807–831, doi:10.1007/BF00893018, 1987.

Ciach, G. J., and W. F. Krajewski, On the estimation of radar rainfall error variance, *Advances in Water Resources*, *22*(6), 585–595, doi:10.1016/S0309-1708(98)00043-8, 1999.

Ciach, G. J., and W. F. Krajewski, Analysis and modeling of spatial correlation structure in small-scale rainfall in Central Oklahoma, *Advances in Water Resources*, *29*(10), 1450–1463, doi:10.1016/j.advwatres.2005.11.003, 2006.

Ciarapica, L., and E. Todini, TOPKAPI: a model for the representation of the rainfall-runoff process at different scales, *Hydrological Processes*, *16*(2), 207–229, doi:10.1002/hyp.342, 2002.

Clothier, A. N., and G. G. Pegram, Space-Time Modelling of Rainfall using the String of Beads Model : Integration of Radar and Raingauge Data Report to the Water Research Commission Space-Time Modelling of Rainfall using the String of Beads Model : Integration of Radar and Raingauge Data E, *Tech. rep.*, 2001.

Cowpertwait, P. S., A generalized point process model for rainfall, *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences*, *447*(1929), 23–37, doi:10.1098/rspa.1994.0126, 1994.

Cowpertwait, P. S., A Poisson-cluster model of rainfall: some high-order moments and extreme values, *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *454*(1971), 885–898, doi:10.1098/rspa.1998.0191, 1998.

Cowpertwait, P. S., A spatial–temporal point process model of rainfall for the Thames catchment, UK, *Journal of Hydrology*, *330*(3-4), 586–595, doi:10.1016/j.jhydrol.2006.04.043, 2006.

Cowpertwait, P. S., A spatial-temporal point process model with a continuous distribution of storm types, *Water Resources Research*, *46*(12), 1–12, doi:10.1029/2010WR009728, 2010.

Cowpertwait, P. S., P. O'Connell, A. Metcalfe, and J. Mawdsley, Stochastic point process modelling of rainfall. I. Single-site fitting and validation, *Journal of Hydrology*, *175*(1-4), 17–46, doi:10.1016/S0022-1694(96)80004-7, 1996.

Cowpertwait, P. S., C. Kilsby, and P. O'Connell, A space-time Neyman-Scott model of rainfall: Empirical analysis of extremes, *Water resources research*, *38*(8), 1131, doi:10.1029/2001WR000709, 2002.

Cowpertwait, P. S., V. Isham, and C. Onof, Point process models of rainfall: developments for fine-scale structure, *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *463*(2086), 2569–2587, doi:10.1098/rspa.2007.1889, 2007.

Cowpertwait, P. S., G. Xie, V. Isham, C. Onof, and D. C. I. Walsh, A fine-scale point process model of rainfall with dependent pulse depths within cells, *Hydrological Sciences Journal*, *56*(7), 1110–1117, doi:10.1080/02626667.2011.604033, 2011.

Cowpertwait, P. S. P., Further developments of the Neyman-Scott clustered point process for modeling rainfall, *Water Resources Research*, *27*, 1431–1438, 1991.

Cox, D., and V. Isham, A simple spatial-temporal model of rainfall, *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences*, *415*(1849), 317–328, doi:10.1098/rspa.1988.0016, 1988.

Crane, R., Space-time structure of rain rate fields, *Journal of Geophysical Research*, *95*(D3), 2011–2020, doi:10.1029/JD095iD03p02011, 1990.

Daly, C., R. Neilson, and D. Phillips, A statistical-topographic model for mapping climatological precipitation over mountainous terrain, *Journal of Applied Meteorology*, *33*(2), 140–158, 1994.

Daubechies, I., Orthonormal bases of compactly supported wavelets, *Communications on Pure and Applied Mathematics*, *41*(7), 909–996, doi:10.1002/cpa.3160410705, 1988.

Davis, A., A. Marshak, and W. Wiscombe, Wavelet-based multifractal analysis of non-stationary and/or intermittent geophysical signals, in *Wavelets in Geophysics*, edited by E. Foufoula-Georgiou and P. Kumar, August, pp. 249 – 298, Academic Press, San Diego,CA, 1994.

De Lima, M., and J. Grasman, Multifractal analysis of 15-min and daily rainfall from a semi-arid region in Portugal, *Journal of Hydrology*, *220*(1-2), 1–11, doi:10.1016/S0022-1694(99)00053-0, 1999.

De Michele, C., and P. Bernardara, Spectral analysis and modeling of space-time rainfall fields, *Atmospheric Research*, *77*(1-4), 124 – 136, doi:10.1016/j.atmosres.2004.10.031, 2005.

De Oliveira, V., A simple model for spatial rainfall fields, *Stochastic Environmental Research and Risk Assessment*, *18*(2), 131–140, doi:10.1007/s00477-003-0146-4, 2004.

Decoster, N., S. Roux, and A. Arneodo, A wavelet-based method for multifractal image analysis. II. Applications to synthetic multifractal rough surfaces, *The European Physical Journal B-Condensed Matter and Complex Systems*, *15*(4), 739–764, doi:10.1007/s100510051179, 2000.

Deidda, R., Multifractal analysis and simulation of rainfall fields in space, *Physics and Chemistry of the Earth, Part B: Hydrology, Oceans and Atmosphere*, *24*(1-2), 73–78, doi:10.1016/S1464-1909(98)00014-8, 1999.

Deidda, R., Rainfall downscaling in a space-time multifractal framework, *Water Resources Research*, *36*(7), 1779–1794, doi:10.1029/2000WR900038, 2000.

Deidda, R., R. Benzi, and F. Siccardi, Multifractal modeling of anomalous scaling laws in rainfall, *Water resources research*, *35*(6), 1853–1867, doi:10.1029/1999WR900036, 1999.

Deidda, R., M. G. Badas, and E. Piga, Space-time multifractality of remotely sensed rainfall fields, *Journal of Hydrology*, *322*(1-4), 2–13, doi:10.1016/j.jhydrol.2005.02.036, 2006.

Demarta, S., and A. J. McNeil, The T Copula and Related Copulas, *International Statistical Review*, *73*(1), 111–129, doi:10.1111/j.1751-5823.2005.tb00254.x, 2007.

Dietrich, C., and G. N. Newsam, A fast and exact method for multidimensional Gaussian stochastic simulations, *Water Resources Research*, *29*(8), 2861–2869, 1993.

Dietrich, W., and R. Reiss, A process [U+0090]based model for colluvial soil depth and shallow landsliding using digital elevation data, *Hydrological Processes*, *9*, 383–400, 1995.

Dingman, S., *Physical Hydrology*, Macmillan, New York, 1994.

Doswell, C., H. Brooks, and R. Maddox, Flash flood forecasting: An ingredients-based methodology, *Weather and Forecasting*, *11*(4), 560–581, 1996.

Durbán, M., and C. Glasbey, Weather modelling using a multivariate latent Gaussian model, *Agricultural and Forest Meteorology*, *109*(3), 187–201, 2001.

Ebtehaj, M., and E. Foufoula-Georgiou, Orographic signature on multiscale statistics of extreme rainfall: A storm-scale study, *Journal of Geophysical Research*, *115*(D23), 1–14, doi:10.1029/2010JD014093, 2010.

Ebtehaj, M., H. Moradkhani, and H. V. Gupta, Improving robustness of hydrologic parameter estimation by the use of moving block bootstrap resampling, *Water Resources Research*, *46*(7), 1–14, doi:10.1029/2009WR007981, 2010.

Embrechts, P., C. Klüppelberg, and T. Mikosch, *Modelling Extremal Events: for Insurance and Finance*, Springer, New York, 1997.

Entekhabi, D., I. Rodriguez-iturbe, and P. S. Eagleson, Probabilistic representation of the temporal rainfall process by a modified Neyman-Scott Rectangular Pulse model: Parameter Estimation and Validation, *Water Resources Research*, *25*(2), 295–302, 1989.

Evin, G., and A.-C. Favre, Further developments of a transient Poisson-cluster model for rainfall, *Stochastic Environmental Research and Risk Assessment*, doi:10.1007/s00477-012-0612-y, 2012.

Fabry, F., On the determination of scale ranges for precipitation fields, *Journal of geophysical research*, *101*(D8), 819–826, doi:10.1029/96JD00718, 1996.

Falconer, K., *Fractal Geometry, Mathematical Foundations and Applications*, John Wiley & Sons, Chichester, England, 1990.

Fatichi, S., S. M. Barbosa, E. Caporali, and M. E. Silva, Deterministic versus stochastic trends: Detection and challenges, *Journal of Geophysical Research*, *114*(D18), 1–11, doi:10.1029/2009JD011960, 2009.

Fatichi, S., V. Y. Ivanov, and E. Caporali, Simulation of future climate scenarios with a weather generator, *Advances in Water Resources*, *34*(4), 448–467, doi:10.1016/j.advwatres.2010.12.013, 2011.

Fatichi, S., V. Y. Ivanov, and E. Caporali, A mechanistic ecohydrological model to investigate complex interactions in cold and warm water-controlled environments: 1. Theoretical framework and plot-scale analysis, *Journal of Advances in Modeling Earth Systems*, *4*(2), 1–31, doi:10.1029/2011MS000086, 2012a.

Fatichi, S., V. Y. Ivanov, and E. Caporali, A mechanistic ecohydrological model to investigate complex interactions in cold and warm water-controlled environments: 2. Spatiotemporal analyses, *Journal of Advances in Modeling Earth Systems*, *4*(2), 1–22, doi:10.1029/2011MS000087, 2012b.

Féral, L., H. Sauvageot, L. Castanet, J. Lemorton, F. Cornet, and K. Leconte, Large-scale modeling of rain fields from a rain cell deterministic model, *Radio Science*, *41*(2), 1–21, doi:10.1029/2005RS003312, 2006.

Ferraris, L., S. Gabellani, and N. Rebora, A comparison of stochastic models for spatial rainfall downscaling, *Water Resources Research*, *39*(12), doi:10.1029/2003WR002504, 2003.

Foresti, L., and A. Pozdnoukhov, Exploration of alpine orographic precipitation patterns with radar image processing and clustering techniques, *Meteorological Applications*, *19*(4), 407–419, doi:10.1002/met.272, 2011.

Foresti, L., M. Kanevski, and A. Pozdnoukhov, Kernel-Based Mapping of Orographic Rainfall Enhancement in the Swiss Alps as Detected by Weather Radar, *IEEE Transactions on Geoscience and Remote Sensing*, *50*(8), 2954–2967, doi:10.1109/TGRS.2011.2179550, 2012.

Foufoula-Georgiou, E., and D. Lettenmaier, A Markov renewal model for rainfall occurrences, *Water Resources Research*, *23*(5), 875–884, 1987.

Fowler, H., S. Blenlinsop, and C. Tebaldi, Linking climate change modelling to impacts studies: recent advances in downscaling techniques for hydrological modelling, *International Journal of Climatology*, *27*(12), 1547–1578, doi:10.1002/joc, 2007.

Fraedrich, K., and C. Larnder, Scaling regimes of composite rainfall time series, *Tellus A*, *45*(4), 289–298, doi:10.1034/j.1600-0870.1993.t01-3-00004.x, 1993.

Frei, C., and C. Schär, A precipitation climatology of the Alps from high-resolution rain-gauge observations, *International Journal of Climatology*, *18*, 873–900, doi:10.1002/(SICI)1097-0088(19980630)18:8<873::AID-JOC255>3.0.CO;2-9, 1998.

Frisch, U., From global scaling, a la Kolmogorov, to local multifractal scaling in fully developed turbulence, *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences*, *434*(1890), 89–99, doi:10.1098/rspa.1991.0082, 1991.

Gabella, M., M. Bolliger, U. Germann, and G. Perona, Large sample evaluation of cumulative rainfall amounts in the Alps using a network of three radars, *Atmospheric Research*, *77*(1-4), 256–268, doi:10.1016/j.atmosres.2004.10.014, 2005.

Gabellani, S., G. Boni, L. Ferraris, J. von Hardenberg, and a. Provenzale, Propagation of uncertainty from rainfall to runoff: A case study with a stochastic rainfall generator, *Advances in Water Resources*, *30*(10), 2061–2071, doi:10.1016/j.advwatres.2006.11.015, 2007.

Gaume, E., N. Mouhous, and H. Andrieu, Rainfall stochastic disaggregation models: Calibration and validation of a multiplicative cascade model, *Advances in Water Resources*, *30*(5), 1301–1319, doi:10.1016/j.advwatres.2006.11.007, 2007.

Gebremichael, M., W. F. Krajewski, T. M. Over, Y. Takayabu, P. Arkin, and M. Katayama, Scaling of tropical rainfall as observed by TRMM precipitation radar, *Atmospheric Research*, *88*(3-4), 337–354, doi:10.1016/j.atmosres.2007.11.028, 2008.

Germann, U., and J. Joss, Variograms of radar reflectivity to describe the spatial continuity of Alpine precipitation, *Journal of Applied Meteorology*, *40*(6), 1042–1059, doi:10.1175/1520-0450(2001)040<1042:VORRTD>2.0.CO;2, 2001.

Germann, U., G. Galli, M. Boscacci, and M. Bolliger, Radar precipitation measurement in a mountainous region, *Quarterly Journal of the Royal Meteorological Society*, *132*(618), 1669–1692, doi:10.1256/qj.05.190, 2006.

Germann, U., M. Berenguer, D. Sempere-torres, and M. Zappa, REAL-Ensemble radar precipitation estimation for hydrology in a mountainous region, *Quarterly Journal of the Royal Meteorological Society*, *135*(639), 445–456, doi:10.1002/qj.375, 2009.

Giesbrecht, A. F., and O. Kempthorne, Maximum Likelihood Estimation in the Three-Parameter Lognormal Distribution, *Journal of the Royal Statistical Society. Series B (Methodological)*, *38*(3), 257–264, 1976.

Gires, A., C. Onof, C. Maksimovic, D. Schertzer, I. Tchiguirinskaia, and N. Simoes, Quantifying the impact of small scale unmeasured rainfall variability on urban runoff through multifractal downscaling: a case study, *Journal of Hydrology*, *442*, 117–128, doi:10.1016/j.jhydrol.2012.04.005, 2012.

Gneiting, T., Matérn Cross-Covariance Functions for Multivariate Random Fields, *Journal of the American Statistical Association*, *105*(491), 1167–1177, doi:10.1198/jasa.2010.tm09420, 2010.

Goldhirsh, J., and B. Musiani, Rain cell size statistics derived from radar observations at Wallops Island, Virginia, *IEEE Transactions on Geoscience and Remote Sensing*, *24*(6), 947–954, 1986.

Goodrich, D. C., T. O. Keefer, C. L. Unkrich, M. H. Nichols, H. B. Osborn, J. J. Stone, and J. R. Smith, Long-term precipitation database, walnut gulch experimental watershed, arizona, united states, *Water Resources Research*, *44*(W05S04), doi:10.1029/2006WR005782, 2008.

Guillot, G., Approximation of Sahelian rainfall fields with meta-Gaussian random functions, *Stochastic Environmental Research and Risk Assessment (SERRA)*, *13*(1-2), 100–112, doi:10.1007/s004770050034, 1999.

Güntner, 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. K., and E. Waymire, Multiscaling properties of spatial rainfall and river flow distributions, *Journal of Geophysical Research*, *95*(D3), 1999–2009, doi:10.1029/JD095iD03p01999, 1990.

Gupta, V. K., and E. Waymire, A statistical analysis of mesoscale rainfall as a random cascade, *Journal of Applied Meteorology*, *32*(2), 251–267, doi:10.1175/1520-0450(1993)032<0251:ASAOMR>2.0.CO;2, 1993.

Gutnisky, D. a., and K. Josić, Generation of spatiotemporally correlated spike trains and local field potentials using a multivariate autoregressive process., *Journal of Neurophysiology*, *103*(5), 2912–30, doi:10.1152/jn.00518.2009, 2010.

Guttorp, P., and T. Gneiting, Studies in the history of probability and statistics XLIX On correlation family On the Mate, *Biometrika*, *93*(4), 989–995, 2006.

Gyasi-Agyei, Y., Identification of regional parameters of a stochastic model for rainfall disaggregation, *Journal of Hydrology*, *223*(3-4), 148–163, doi:10.1016/S0022-1694(99)00114-6, 1999.

Gyasi-Agyei, Y., and G. Willgoose, A hybrid model for point rainfall modeling, *Water resources research*, *33*(7), 1699–1706, 1997.

Gyasi-Agyei, Y., and G. Willgoose, Generalisation of a hybrid model for point rainfall, *Journal of Hydrology*, *219*(3-4), 218–224, doi:10.1016/S0022-1694(99)00054-2, 1999.

Haan, C. T., D. M. Allen, and J. O. Street, A Markov Chain Model of daily rainfall, *Water Resources Research*, *12*(3), 443, doi:10.1029/WR012i003p00443, 1976.

Habib, E., W. F. Krajewski, and G. J. Ciach, Estimation of rainfall interstation correlation, *Journal of Hydrometeorology*, *2*(6), 621–629, 2001.

Harris, D., M. Menabde, A. Seed, and G. Austin, Multifractal characterization of rain fields with a strong orographic influence, *Journal of Geophisical Research*, *101*(D21), 405–426, doi:10.1029/96JD01656, 1996.

Harris, D., A. Seed, M. Menabde, and G. Austin, Factors affecting multiscaling analysis of rainfall time series, *Nonlinear Processes in Geophysics*, *4*(3), 137–156, 1997.

Heimsath, A., W. Dietrich, K. Nishiizumi, and R. Finkel, The soil production function and landscape equilibrium, *Nature*, *388*(July), 21–24, 1997.

Heimsath, A. M., W. E. Dietrich, K. Nishiizumi, and R. C. Finkel, Cosmogenic nuclides, topography, and the spatial variation of soil depth, *Geomorphology*, *27*(1-2), 151–172, doi:10.1016/S0169-555X(98)00095-6, 1999.

Helgason, H., V. Pipiras, and P. Abry, Fast and exact synthesis of stationary multivariate Gaussian time series using circulant embedding, *Signal Processing*, *91*(5), 1123–1133, doi:10.1016/j.sigpro.2010.10.014, 2011.

Heneghan, C., S. Lowen, and M. Teich, Two-dimensional fractional Brownian motion: wavelet analysis and synthesis, in *Image Analysis and Interpretation, 1996., Proceedings of the IEEE Southwest Symposium on*, pp. 213–217, IEEE, 1996.

Hillel, D., *Environmental Soil Physics*, Academic Press, San Diego,CA, 1998.

Holley, R., and E. Waymire, Multifractal dimensions and scaling exponents for strongly bounded random cascades, *The Annals of Applied Probability*, *2*(4), 819–845, 1992.

Hosking, J. R. M., Fractional differencing, *Biometrika*, *68*(1), 165–176, 1981.

Hosking, J. R. M., L-moments: analysis and estimation of distributions using linear combinations of order statistics, *Journal of the Royal Statistical Society. Series B (Methodological)*, *52*(1), 105–124, 1990.

Houze, R. A., C. James, and S. Medina, Radar observations of precipitation and airflow on the Mediterranean side of the Alps: Autumn 1998 and 1999, *Quarterly Journal of the Royal Meteorological Society*, *127*(578), 2537–2558, doi:10.1256/smsqj.57803, 2001.

Houze, R. J., *Cloud Dynamics*, Academic Press, San Diego,CA, 1993.

Houze, R. J., Orographic effects on precipitating clouds, *Reviews of Geophysics*, *50*(RG1001), 1–47, doi:10.1029/2011RG000365, 2012.

Hristopulos, D., and S. Elogne, Computationally efficient spatial interpolators based on Spartan spatial random fields, *IEEE, Transactions on Signal Processing*, *57*(9), 3475–3487, doi:10.1109/TSP.2009.2021450, 2009.

Hughes, J. P., P. Guttorp, and S. P. Charles, A non-homogeneous hidden Markov model for precipitation occurrence, *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, *48*(1), 15–30, doi: 10.1111/1467-9876.00136, 1999.

Hulme, M., D. Conway, P. Jones, J. T, E. Barrow, and C. Turney, Construction of a 1961-1990 European Climatology for climate change modelling and impact applications, *International Journal of Climatology*, *15*(12), 1333–1363, 1995.

Hurst, H., Long-term storage capacity of reservoirs, *Transactions of the Amererical Society of Civil Engineers*, *116*, 770–808, 1951.

Hwang, W., and S. Mallat, Characterization of self-similar multifractals with wavelet maxima, *Tech. rep.*, 1993.

Ignaccolo, M., C. De Michele, and S. Bianco, The Droplike Nature of Rain and Its Invariant Statistical Properties, *Journal of Hydrometeorology*, *10*(1), 79–95, doi:10.1175/2008JHM975.1, 2009.

IPCC, *Climate Change 2007: The physical science basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, Cambridge University Press, 2007.

Islam, S., and D. Entekhabi, Parameter estimation and sensitivity analysis for the modified Bartlett-Lewis rectangular pulses model of rainfall, *Journal of Geophisical Research*, *95*(2), 2093–2100, 1990.

Ivanov, V. Y., E. R. Vivoni, R. L. Bras, and D. Entekhabi, Preserving high-resolution surface and rainfall data in operational-scale basin hydrology: a fully-distributed physically-based approach, *Journal of Hydrology*, *298*(1-4), 80–111, doi:10.1016/j.jhydrol.2004.03.041, 2004.

Ivanov, V. Y., R. L. Bras, and D. C. Curtis, A weather generator for hydrological, ecological, and agricultural applications, *Water Resources Research*, *43*(10), doi:10.1029/2006WR005364, 2007.

Ivanov, V. Y., R. L. Bras, and E. R. Vivoni, Vegetation-hydrology dynamics in complex terrain of semiarid areas: 1. A mechanistic approach to modeling dynamic feedbacks, *Water Resources Research*, *44*(3), W03,429, doi:10.1029/2006WR005588, 2008.

Jeannin, N., L. Féral, H. Sauvageot, L. Castanet, and F. Lacoste, A Large-Scale Space-Time Stochastic Simulation Tool of Rain Attenuation for the Design and Optimization of Adaptive Satellite Communication Systems Operating between 10 and 50 GHz, *International Journal of Antennas and Propagation*, *2012*, 1–16, doi:10.1155/2012/749829, 2012.

Johansson, B., and D. Chen, The influence of wind and topography on precipitation distribution in Sweden: statistical analysis and modelling, *International Journal of Climatology*, *23*(12), 1523–1535, doi: 10.1002/joc.951, 2003.

Joss, J., and R. Lee, The application of radar-gauge comparisons to operational precipitation profile corrections, *Journal of Applied Meteorology*, *34*(12), 2612–2630, 1995.

Jothityangkoon, C., M. Sivapalan, and N. Viney, Tests of a space-time model of daily rainfall in southwestern Australia based on nonhomogeneous random cascades, *Water resources research*, *36*(1), 267–284, doi:10.1029/1999WR900253, 2000.

Kahane, J., and J. Peyriere, Sur certaines martingales de Benoit Mandelbrot, *Advances in mathematics*, *22*(2), 131–145, 1976.

Kang, B., and J. a. Ramírez, A coupled stochastic space-time intermittent random cascade model for rainfall downscaling, *Water Resources Research*, *46*(10), 1–17, doi:10.1029/2008WR007692, 2010.

Kantelhardt, J., S. Zschiegner, E. Koscielny-Bunde, S. Havlin, A. Bunde, and H. Stanley, Multifractal detrended fluctuation analysis of nonstationary time series, *Physica A: Statistical Mechanics and its Applications*, *316*(1-4), 87–114, doi:10.1016/S0378-4371(02)01383-3, 2002.

Kantelhardt, J. W., E. Koscielny-Bunde, D. Rybski, P. Braun, A. Bunde, and S. Havlin, Long-term persistence and multifractality of precipitation and river runoff records, *Journal of Geophysical Research*, *111*(D1), 1–13, doi:10.1029/2005JD005881, 2006.

Katul, G. G., and M. Parlange, Analysis of land surface heat fluxes using the orthonormal wavelet approach, *Water Resources Research*, *31*(11), 2743–2749, doi:10.1029/95WR00003, 1995.

Katz, R., Precipitation as a chain-dependent process, *Journal of Applied Meteorology*, *16*(7), 671–676, 1977.

Katz, R., and M. Parlange, Generalizations of chain-dependent processes: Application to hourly precipitation, *Water Resources Research*, *31*(5), 1331–1341, 1995.

Katz, R., and M. Parlange, Overdispersion phenomenon in stochastic modeling of precipitation, *Journal of Climate*, *11*(4), 591–601, 1998.

Kavvas, M., and J. Delleur, A stochastic cluster model of daily rainfall sequences, *Water Resources Research*, *17*(4), 1151–1160, 1981.

Kim, J., A. Warnock, V. Y. Ivanov, and N. D. Katopodes, Coupled modeling of hydrologic and hydrodynamic processes including overland and channel flow, *Advances in Water Resources*, *37*, 104–126, doi:10.1016/j.advwatres.2011.11.009, 2012.

Kitanidis, P., *Introduction to Geostatistics*, Cambridge University Press, 1997.

Kitanidis, P., and R. Bras, Real-time forecasting with a conceptual hydrologic model 2. applications and results, *Water Resources Research*, *16*(6), 1034–1044, 1980.

Klazura, G., and D. Imy, A description of the initial set of analysis products availiable from the NEXRAD WSD-88D system, *Bulletin of the Americal Meteorological Society*, *74*(7), 1293–1311, 1993.

Kleiber, W., R. W. Katz, and B. Rajagopalan, Daily spatiotemporal precipitation simulation using latent and transformed Gaussian processes, *Water Resources Research*, *48*(1), 1–17, doi:10.1029/2011WR011105, 2012.

Klemeš, V., The Hurst phenomenon: a puzzle, *Water Resources Research*, *10*(4), 675–688, doi:10.1029/WR010i004p00675, 1974.

Knutti, R., Should we believe model predictions of future climate change?, *Philosophical transactions. Series A, Mathematical, physical, and engineering sciences*, *366*(1885), 4647–64, doi:10.1098/rsta.2008.0169, 2008.

Kollet, S. J., and R. M. Maxwell, Integrated surface–groundwater flow modeling: A free-surface overland flow boundary condition in a parallel groundwater flow model, *Advances in Water Resources*, *29*(7), 945–958, doi:10.1016/j.advwatres.2005.08.006, 2006.

Kollet, S. J., and R. M. Maxwell, Capturing the influence of groundwater dynamics on land surface processes using an integrated, distributed watershed model, *Water Resources Research*, *44*(2), 1–18, doi:10.1029/2007WR006004, 2008.

Kolmogorov, A., Wienersche Spiralen und einige andere interessante Kurven im Hilbertschen Raum, *Acad. Sci. URSS*, *26*, 115–118, 1940.

Koutsoyiannis, D., A nonlinear disaggregation method with a reduced parameter set for simulation of hydrologic series, *Water Resources Research*, *28*(12), 3175–3191, 1992.

Koutsoyiannis, D., A generalized mathematical framework for stochastic simulation and forecast of hydrologic time series, *Water Resources Research*, *36*(6), 1519–1533, doi:10.1029/2000WR900044, 2000.

Koutsoyiannis, D., Coupling stochastic models of different timescales, *Water Resources Research*, *37*(2), 379–391, doi:10.1029/2000WR900200, 2001.

Koutsoyiannis, D., Climate change, the Hurst phenomenon, and hydrological statistics, *Hydrological Sciences Journal*, *48*(1), 3–24, doi:10.1623/hysj.48.1.3.43481, 2003.

Koutsoyiannis, D., An entropic-stochastic representation of rainfall intermittency: The origin of clustering and persistence, *Water Resources Research*, *42*(1), 1–20, doi:10.1029/2005WR004175, 2006.

Koutsoyiannis, D., HESS Opinions "A random walk on water", *Hydrology and Earth System Sciences Discussions*, *14*(3), 585–601, 2009.

Koutsoyiannis, D., Scale of water resources development and sustainability: small is beautiful, large is great, *Hydrological Sciences Journal*, *56*(4), 553–575, doi:10.1080/02626667.2011.579076, 2011.

Koutsoyiannis, D., and A. Economou, Evaluation of the parameterization-simulation-optimization approach for the control of reservoir systems, *Water Resources Research*, *39*(6), 1170, doi:10.1029/2003WR002148, 2003.

Koutsoyiannis, D., and E. Foufoula-Georgiou, A Scaling Model of a Storm Hyetograph, *Water Resources*, *29*(7), 2345–2361, doi:10.1029/93WR00395, 1993.

Koutsoyiannis, D., and A. Langousis, Treatise on Water Science Chapter 27 : Precipitation 1 Introduction, in *Treatise on Water Science*, edited by P. Wilderer, Elsevier, 2011.

Koutsoyiannis, D., A. Efstratiadis, N. Mamassis, and A. Christofides, On the credibility of climate predictions, *Hydrological Sciences Journal*, *53*(4), 671–684, 2008.

Koutsoyiannis, D., A. Paschalis, and N. Theodoratos, Two-dimensional Hurst-Kolmogorov process and its application to rainfall fields, *Journal of Hydrology*, *398*(1-2), 91–100, doi:10.1016/j.jhydrol.2010.12.012, 2011.

Krajewski, W. F., V. Lakshmi, K. P. Georgakakos, and S. C. Jain, A Monte Carlo Study of Rainfall Sampling Effect on a Distributed Catchment Model, *Water Resources Research*, *27*(1), 119–128, doi:10.1029/90WR01977, 1991.

Krajewski, W. F., E. N. Anagnostou, and G. J. Ciach, Effects of the radar observation process on the inferred rainfall statistics, *Journal Geophysical Research*, *101*(D21), 26,493–26,502, doi:10.1029/96JD01616, 1996.

Kumar, P., and T. Bell, Space-time scaling behavior of rain statistics in a stochastic fractional diffusion model, *Journal of Hydrology*, *322*(1-4), 49–58, doi:10.1016/j.jhydrol.2005.02.031, 2006.

Kummerow, C., W. Barnes, T. Kozu, J. Shiue, and J. Simpson, The Tropical Rainfall Measuring Mission (TRMM) Sensor Package, *Journal of Atmospheric and Oceanic Technology*, *15*(3), 809–817, doi:10.1175/1520-0426(1998)015<0809:TTRMMT>2.0.CO;2, 1998.

Kundu, P. K., and T. Bell, A stochastic model of space-time variability of mesoscale rainfall: Statistics of spatial averages, *Water Resources Research*, *39*(12), doi:10.1029/2002WR001802, 2003.

Kundu, P. K., and R. K. Siddani, Scale dependence of spatiotemporal intermittence of rain, *Water Resources Research*, *47*(8), 1–17, doi:10.1029/2010WR010070, 2011.

Kuo, W.-l., T. S. Steenhuis, C. E. Mcculloch, C. L. Mohler, D. A. Weinstein, D. Degloria, and D. P. Swaney, Effects of grid size on runoff and soil noisture for a variable-source-area hydrology model, *Water Resources Research*, *35*(11), 3419–3428, 1999.

Lakshmanan, V., R. Rabin, and V. DeBrunner, Multiscale storm identification and forecast, *Atmospheric Research*, *67-68*, 367–380, doi:10.1016/S0169-8095(03)00068-1, 2003.

Lamperti, J., Semi-stable stochastic processes, *Transactions of the American Mathematical Society*, *104*(1), 62 –78, 1962.

Lang, A., and J. Potthoff, Fast simulation of Gaussian random fields, *Monte Carlo Methods and Applications*, *17*(3), 195–214, 2011.

Langousis, A., and D. Veneziano, Intensity-duration-frequency curves from scaling representations of rainfall, *Water resources research*, *43*(2), 2422, doi:10.1029/2006WR005245, 2007.

Langousis, A., D. Veneziano, P. Furcolo, and C. Lepore, Multifractal rainfall extremes: Theoretical analysis and practical estimation, *Chaos, Solitons & Fractals*, *39*(3), 1182–1194, doi:10.1016/j.chaos.2007.06.004, 2009.

Lanza, L. G., and L. Stagi, High resolution performance of catching type rain gauges from the laboratory phase of the WMO Field Intercomparison of Rain Intensity Gauges, *Atmospheric Research*, *94*(4), 555–563, doi:10.1016/j.atmosres.2009.04.012, 2009.

Lashermes, B., P. Abry, and P. Chainais, New insights into the estimation of scaling exponents, *International Journal of Wavelets*, *2*(4), 497–523, doi:10.1142/S0219691304000597, 2004.

Lavallée, D., D. Schertzer, and S. Lovejoy, On the determination of the codimension function, in *Nonlinear Variability in Geophysics*, pp. 99–110, Kluwer, 1991.

LeCam, L., A stochastic description of precipitation, in *Fourth Berkeley Symposium on Mathematical Statistics and Probability, vol. 3*, pp. 165–186, J. Neyman ed. Berkeley, University of California., Berkeley, 1961.

Legates, D., and C. Willmott, Mean seasonal and spatial variability in gauge corrected, global precipitation, *International Journal of Climatology*, *10*(2), 111–127, 1990.

Leonard, M., M. Lambert, A. Metcalfe, and P. S. Cowpertwait, A space-time Neyman–Scott rainfall model with defined storm extent, *Water Resources Research*, *44*(9), doi:10.1029/2007WR006110, 2008.

Lewis, P., A branching Poisson process model for the analysis of computer failure patterns, *Journal of the Royal Statistical Society. Series B (Methodological)*, *26*(3), 398–456, 1964.

Li, L., W. Schmid, and J. Joss, Nowcasting of motion and growth of precipitation with radar over a complex orography, *Journal of Applied Meteorology*, *34*(6), 1286–1300, doi:10.1175/1520-0450(1995)034<1286:NOMAGO>2.0.CO;2, 1995.

Liu, Z., and E. Todini, Towards a comprehensive physically-based rainfall-runoff model, *Hydrology and Earth System*, *6*(5), 859–881, 2002.

Lo, A., Long-term memory in stock market prices, *Econometrica*, *59*(5), 1279–1313, doi:10.2307/2938368, 1989.

Lombardo, F., E. Volpi, and D. Koutsoyiannis, Rainfall downscaling in time: theoretical and empirical comparison between multifractal and Hurst-Kolmogorov discrete random cascades, *Hydrological Sciences Journal*, *57*(6), 1052–1066, 2012.

Lovejoy, S., The remarkable wide range spatial scaling of TRMM precipitation, *Atmospheric Research*, *90*(1), 10–32, doi:10.1016/j.atmosres.2008.02.016, 2008.

Lovejoy, S., and D. Schertzer, Generalized Scale Invariance in the Atmosphere, *Water Resources Research*, *21*(8), 1233–1250, doi:10.1029/WR021i008p01233, 1985.

Lovejoy, S., D. Lavallée, D. Schertzer, and P. Ladoy, the l1/2 law and multifractal topography: theory and analysis, *Nonlinear Processes in Geophysics*, *2*(1), 16–22, 1995.

Mallat, S., A theory for multiresolution signal decomposition: The wavelet representation, *IEEE transactions on pattern analysis and machine intelligence*, *11*(7), 674–693, 1989.

Mallat, S., *A wavelet tour of signal processing*, Academic Press, San Diego,CA, 1998.

Mandapaka, P. V., W. F. Krajewski, R. Mantilla, and V. K. Gupta, Dissecting the effect of rainfall variability on the statistical structure of peak flows, *Advances in Water Resources*, *32*(10), 1508–1525, doi:10.1016/j.advwatres.2009.07.005, 2009.

Mandapaka, P. V., G. Villarini, B.-C. Seo, and W. F. Krajewski, Effect of radar-rainfall uncertainties on the spatial characterization of rainfall events, *Journal of Geophysical Research*, *115*(D17), 1–16, doi:10.1029/2009JD013366, 2010.

Mandelbrot, B. B., A fast fractional Gaussian noise generator, *Water Resources Research*, *7*(3), 543–553, doi:10.1029/WR007i003p00543., 1971.

Mandelbrot, B. B., Intermittent turbulence in self-similar cascades: divergence of high moments and dimension of the carrier, *Journal of Fluid Mechanics*, *62*(02), 331–358, doi:10.1017/S0022112074000711, 1974.

Mandelbrot, B. B., *The Fractal Geometry of Nature*, Freeman, Wh., San Fransisco, 1983.

Mandelbrot, B. B., and J. W. Van Ness, Fractional Brownian Motions, Fractional Noises and Applications, *SIAM Review*, *10*(4), 422–437, doi:10.1137/1010093, 1968.

Mantoglou, A., and J. Wilson, The turning bands method for simulation of random fields using line generation by a spectral method, *Water Resources Research*, *18*(5), 1379–1394, 1982.

Marani, M., On the correlation structure of continuous and discrete point rainfall, *Water Resources Research*, *39*(5), 1128, doi:10.1029/2002WR001456, 2003.

Marcotte, D., Fast variogram computation with FFT, *Computers & Geosciences*, *22*(10), 1175–1186, doi:10.1016/S0098-3004(96)00026-X, 1996.

Marshak, A., A. Davis, R. Cahalan, and W. Wiscombe, Bounded cascade models as nonstationary multifractals, *Physical Review E*, *49*(1), 55–69, doi:10.1103/PhysRevE.49.55, 1994.

Marshall, J., and W. Palmer, The distribution of raindrops with size, *Journal of Meteorology*, *5*(4), 165–166, 1948.

Marshall, J., R. Langille, and W. Palmer, Measurement of rainfall by radar., *Journal of Meteorology*, *4*, 186–192, 1947.

Matalas, N., Mathematical assessment of synthetic hydrology, *Water Resources Research*, *3*(4), 937–945, 1967.

Matheron, G., Principles of geostatistics, *Economic geology*, *58*, 1246–1266, 1963.

Maynard, R., Radar and weather, *Journal of Atmospheric Sciences*, *2*(4), 214–226, 1945.

McGaughey, D., and G. Aitken, Generating two-dimensional fractional Brownian motion using the fractional Gaussian process (FGp) algorithm, *Physica A: Statistical Mechanics and its Applications*, *311*(3-4), 369–380, doi:10.1016/S0378-4371(02)00778-1, 2002.

Mejia, J., and I. Rodriguez-Iturbe, Correlation links between normal and log normal processes, *Water Resources Research*, *10*(4), 689–690, 1974.

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, doi:10.1029/2000WR900197, 2000.

Menabde, M., D. Harris, A. Seed, G. Austin, and D, Multiscaling properties of rainfall and bounded random cascades, *Water Resources Research*, *33*(12), 2823–2830, doi:10.1029/97WR02006, 1997a.

Menabde, M., A. Seed, D. Harris, and G. Austin, Self-similar random fields and rainfall simulation, *Journal of Geophysical Research*, *102*(D12), 509–515, doi:10.1029/97JD00915, 1997b.

Menabde, M., A. Seed, D. Harris, and G. Austin, Multiaffine random field model of rainfall, *Water Resources Research*, *35*(2), 509–514, doi:10.1029/1998WR900020, 1999.

Meneveau, C., K. Sreenivasan, P. Kailasnath, and M. Fan, Joint multifractal measures: Theory and applications to turbulence., *Physical review. A*, *41*(2), 894–913, doi:10.1103/PhysRevA.41.894, 1990.

Merz, B., and A. Bárdossy, Effects of spatial variability on the rainfall runoff process in a small loess catchment, *Journal of Hydrology*, *213*(1-4), 304–317, 1998.

Merz, B., and E. Plate, An analysis of the effects of spatial variability of soil and soil moisture on runoff, *Water Resources Research*, *33*(12), 2909–2922, 1997.

Metropolis, N., and S. Ulam, The monte carlo method, *Journal of the American Statistical Association*, *44*(247), 335–341, 1949.

Metropolis, N., A. Rosenbluth, M. Rosenbluth, A. Teller, E. Teller, and Others, Equation of state calculations by fast computing machines, *The journal of chemical physics*, *21*(6), 1087, 1953.

Metta, S., J. von Hardenberg, L. Ferraris, N. Rebora, and A. Provenzale, Precipitation Nowcasting by a Spectral-Based Nonlinear Stochastic Model, *Journal of Hydrometeorology*, *10*(5), 1285–1297, doi: 10.1175/2009JHM1120.1, 2009.

Minet, J., E. Laloy, S. Lambot, and M. Vanclooster, Effect of high-resolution spatial soil moisture variability on simulated runoff response using a distributed hydrologic model, *Hydrology and Earth System Sciences*, *15*(4), 1323–1338, doi:10.5194/hess-15-1323-2011, 2011.

Molini, A., P. La Barbera, L. Lanza, and L. Stagi, Rainfall intermittency and the sampling error of tipping-bucket rain gauges, *Physics and Chemistry of the Earth, Part C: Solar, Terrestrial & Planetary Science*, *26*(10-12), 737–742, doi:10.1016/S1464-1917(01)95018-4, 2001.

Molini, A., G. G. Katul, and A. Porporato, Revisiting rainfall clustering and intermittency across different climatic regimes, *Water Resources Research*, *45*(11), 1–13, doi:10.1029/2008WR007352, 2009.

Molnar, P., and P. Burlando, Preservation of rainfall properties in stochastic disaggregation by a simple random cascade model, *Atmospheric Research*, *77*, 137–151, doi:10.1016/j.atmosres.2004.10.024, 2005.

Molnar, P., and P. Burlando, Variability in the scale properties of high-resolution precipitation data in the Alpine climate of Switzerland, *Water Resources Research*, *44*(10), doi:10.1029/2007WR006142, 2008.

Montanari, A., R. Rosso, and M. Taqqu, Some long-run properties of rainfall records in Italy, *Journal of geophysical research*, *101*(D23), 29,431–29,438, 1996.

Montanari, A., R. Rosso, and M. Taqqu, Fractionally differenced ARIMA models applied to hydrologic time series: Identification, estimation, and simulation, *Water Resources Research*, *33*(5), 1035–1044, doi:10.1029/97WR00043, 1997.

Morbidelli, R., C. Corradini, C. Saltalippi, and L. Brocca, Initial Soil Water Content as Input to Field-Scale Infiltration and Surface Runoff Models, *Water Resources Management*, *26*(7), 1793–1807, doi: 10.1007/s11269-012-9986-3, 2012.

Moreno, H. A., E. R. Vivoni, and D. J. Gochis, Utility of Quantitative Precipitation Estimates for high resolution hydrologic forecasts in mountain watersheds of the Colorado Front Range, *Journal of Hydrology*, *438-439*, 66–83, doi:10.1016/j.jhydrol.2012.03.019, 2012.

Movahed, M. S., G. R. Jafari, F. Ghasemi, S. Rahvar, and M. R. R. Tabar, Multifractal detrended fluctuation analysis of sunspot time series, *Journal of Statistical Mechanics: Theory and Experiment*, *2006*(02), P02,003–P02,003, doi:10.1088/1742-5468/2006/02/P02003, 2006.

Muzy, J.-F., 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*(2), 875–884, doi:10.1103/PhysRevE.47.875, 1993.

Nelder, J., and R. Mead, A simplex method for function minimization, *The computer journal*, *7*(4), 305–313, doi:10.1093/comjnl/7.4.308, 1965.

Neyman, J., and E. Scott, A theory of the spatial distribution of galaxies, *The Astrophysical Journal*, *116*(1), 144–163, doi:10.1086/145599, 1952.

Ng, W., and U. Panu, Comparisons of traditional and novel stochastic models for the generation of daily precipitation occurrences, *Journal of Hydrology*, *380*, 222–236, doi:10.1016/j.jhydrol.2009.11.002, 2010.

Nicolis, O., P. Ramírez-Cobo, and B. Vidakovic, 2D wavelet-based spectra with applications, *Computational Statistics & Data Analysis*, *55*(1), 738–751, doi:10.1016/j.csda.2010.06.020, 2011.

Nicotina, L., E. A. Celegon, A. Rinaldo, and M. Marani, On the impact of rainfall patterns on the hydrologic response, *Water Resources Research*, *44*(12), W12,401, doi:10.1029/2007WR006654, 2008.

Nikolopoulos, E. I., E. N. Anagnostou, M. Borga, E. R. Vivoni, and A. Papadopoulos, Sensitivity of a mountain basin flash flood to initial wetness condition and rainfall variability, *Journal of Hydrology*, *402*(3-4), 165–178, doi:10.1016/j.jhydrol.2010.12.020, 2011.

Northrop, P., A clustered spatial-temporal model of rainfall, *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences*, *454*(1975), 1875–1888, doi:10.1098/rspa.1998.0238, 1998.

Noto, L. V., V. Y. Ivanov, R. L. Bras, and E. R. Vivoni, Effects of initialization on response of a fully-distributed hydrologic model, *Journal of Hydrology*, *352*(1-2), 107–125, doi:10.1016/j.jhydrol.2007.12.031, 2008.

Nykanen, D., and D. Harris, Orographic influences on the multiscale statistical properties of precipitation, *Journal of geophysical research*, *108*(D8), 8381, doi:10.1029/2001JD001518, 2003.

Nykanen, D. K., Linkages between Orographic Forcing and the Scaling Properties of Convective Rainfall in Mountainous Regions, *Journal of Hydrometeorology*, *9*(3), 327–347, doi:10.1175/2007JHM839.1, 2008.

Obled, C., J. Wendling, and K. Beven, The sensitivity of hydrological models to spatial rainfall patterns: an evaluation using observed data, *Journal of Hydrology*, *159*(1-4), 305–333, 1994.

Oliver, D., Moving averages for Gaussian simulation in two and three dimensions, *Mathematical Geology*, *27*(8), 939–960, 1995.

Olsson, J., Validity and applicability of a scale-independent, multifractal relationship for rainfall, *Atmospheric research*, *42*(1-4), 53–65, doi:10.1016/0169-8095(95)00052-6, 1996.

Olsson, J., Evaluation of a scaling cascade model for temporal rain-fall disaggregation, *Hydrology and Earth System Sciences*, *2*(1), 19–30, 1998.

Olsson, J., and P. Burlando, Reproduction of temporal scaling by a rectangular pulses rainfall model, *Hydrological Processes*, *16*(3), 611–630, doi:10.1002/hyp.307, 2002.

Onof, C., and H. Wheater, Improvements to the modelling of British rainfall using a modified random parameter Bartlett-Lewis rectangular pulse model, *Journal of Hydrology*, *157*, 177–195, 1994.

Onof, C., and H. S. Wheater, Modelling of British rainfall using a random parameter Bartlett-Lewis rectangular pulse model, *Journal of Hydrology*, *149*(1-4), 67–95, doi:10.1016/0022-1694(93)90100-N, 1993.

Onof, C., R. Chandler, A. Kakou, P. Northrop, H. Wheater, and V. Isham, Rainfall modelling using Poisson-cluster processes: a review of developments, *Stochastic Environmental Research and Risk Assessment*, *14*(6), 384–411, 2000.

Ossiander, M., and E. Waymire, Statistical estimation for multiplicative cascades, *Annals of statistics*, *28*(6), 1533–1560, 2000.

Ossiander, M., and E. Waymire, On estimation theory for multiplicative cascades, *Sankhyà: The Indian Journal of Statistics, Series A*, *64*(2), 323–343, 2002.

Over, T. M., Modeling space-time rainfall at the mesoscale using random cascades, Ph.D. thesis, University of Colorado, 1995.

Over, T. M., and V. K. Gupta, Statistical Analysis of Mesoscale Rainfall: Dependence of a Random Cascade Generator on Large-Scale Forcing, *Journal of Applied Meteorology*, *33*(12), 1526–1542, doi:10.1175/1520-0450(1994)033<1526:SAOMRD>2.0.CO;2, 1994.

Over, T. M., and V. K. Gupta, A space-time theory of mesoscale rainfall using random cascades, *Journal of geophysical research*, *101*(D21), 26,319–26,331, doi:10.1029/96JD02033, 1996.

Panziera, L., and U. Germann, The relation between airflow and orographic precipitation on the southern side of the Alps as revealed by weather radar, *Quarterly Journal of the Royal Meteorological Society*, *136*(646), 222–238, doi:10.1002/qj.544, 2010.

Papalexiou, S., and D. Koutsoyiannis, Entropy based derivation of probability distributions: A case study to daily rainfall, *Advances in Water Resources*, *45*, 51–57, doi:10.1016/j.advwatres.2011.11.007, 2011.

Papalexiou, S. M., and D. Koutsoyiannis, An all timescales rainfall probability distribution, in *European Geosciences Union, General Assebly*, April, Vienna, Austria, 2009.

Papalexiou, S. M., D. Koutsoyiannis, and A. Montanari, Can a simple stochastic model generate rich patterns of rainfall events?, *Journal of Hydrology*, *411*(3-4), 279–289, doi:10.1016/j.jhydrol.2011.10.008, 2011.

Papoulis, A., and S. Unnikrishna, *Probability, Random Variables and Stochastic Processes*, 4 ed., McGraw Hill, 2002.

Pardo-Iguzquiza, E., and M. Chica-Olmo, The Fourier integral method: an efficient spectral method for simulation of random fields, *Mathematical Geology*, *25*(2), 177–217, 1993.

Parke, W., What is fractional integration?, *Review of Economics and Statistics*, *81*(4), 632–638, 1999.

Paschalis, A., P. Molnar, and P. Burlando, Temporal dependence structure in weights in a multiplicative cascade model for precipitation, *Water Resources Research*, *48*(1), 1–14, doi:10.1029/2011WR010679, 2012.

Pathirana, A., and S. Herath, Multifractal modelling and simulation of rain fields exhibiting spatial heterogeneity, *Hydrology and Earth System Sciences*, *6*(4), 659–708, 2002.

Pathirana, A., S. Herath, and T. Yamada, On the modelling of temporal correlations in spatial-cascade rainfall downscaling, in *Symposium on Weather Radar Information and Distributed Hydrological Modelling held at the 23rd IUGG General Assembly*, 282, pp. 74–84, 2003a.

Pathirana, A., S. Herath, and T. Yamada, Estimating rainfall distributions at high temporal resolutions using a multifractal model, *Hydrology and Earth System Sciences*, *7*(5), 668–679, 2003b.

Paulson, K. S., and P. D. Baxter, Downscaling of rain gauge time series by multiplicative beta cascade, *Journal of Geophysical Research*, *112*(D9), doi:10.1029/2006JD007333, 2007.

Pecknold, S., S. Lovejoy, D. Schertzer, C. Hooge, and J. Malouin, The simulation of universal multifractals, in *Cellular Automata*, vol. 1, pp. 228–267, 1993.

Pegram, G. G., and A. N. Clothier, High resolution space-time modelling of rainfall: the "String of Beads" model, *Journal of Hydrology*, *241*, 26–41, doi:10.1016/S0022-1694(00)00373-5, 2001a.

Pegram, G. G., and A. N. Clothier, Downscaling rainfields in space and time, using the String of Beads model in time series mode, *Hydrology and Earth System Sciences*, *5*(2), 175–186, doi:10.5194/hess-5-175-2001, 2001b.

Perica, S., and E. Foufoula-Georgiou, Linkage of scaling and thermodynamic parameters of rainfall: Results from midlatitude mesoscale convective systems, *Journal of Geophysical Research*, *101*(D3), 7431–7448, doi:10.1029/95JD02372, 1996.

Pui, A., A. Sharma, R. Mehrotra, B. Sivakumar, and E. Jeremiah, A comparison of alternatives for daily to sub-daily rainfall disaggregation, *Journal of Hydrology*, *In Press*, doi:10.1016/j.jhydrol.2012.08.041, 2012.

Purdy, J. C., D. Harris, G. Austin, A. Seed, and W. Gray, A case study of orographic rainfall processes incorporating multiscale characterization techniques, *Journal of Geophysical Research*, *106*(D8), 7837–7845, doi:10.1029/2000JD900622, 2001.

Racsko, P., L. Szeidl, and M. Semenov, A serial approach to local stochastic weather models, *Ecological modelling*, *57*, 27–41, 1991.

Ramirez, J., and R. Bras, Conditional distribution of Neyman-Scott models for storm arrivals and their use in irrigation scheduling, *Water Resources Research*, *21*(3), 317–330, 1985.

Ramírez-Cobo, P., K. S. Lee, A. Molini, A. Porporato, G. G. Katul, and B. Vidakovic, A wavelet-based spectral method for extracting self-similarity measures in time-varying two-dimensional rainfall maps, *Journal of Time Series Analysis*, *32*(4), 351–363, doi:10.1111/j.1467-9892.2011.00731.x, 2011.

Ravalec, M., B. Noetinger, and L. Hu, The FFT moving average (FFT-MA) generator: An efficient numerical method for generating and conditioning Gaussian simulations, *Mathematical Geology*, *32*(6), 701–723, 2000.

Rebora, N., L. Ferraris, J. von Hardenberg, and A. Provenzale, RainFARM: rainfall downscaling by a filtered autoregressive model, *Journal of Hydrometeorology*, *7*(4), 724–738, doi:10.1175/JHM517.1, 2006.

Richardson, C., Stochastic simulation of daily precipitation, temperature, and solar radiation, *Water Resources Research*, *17*(1), 182–190, 1981.

Rico-Ramirez, M., and I. Cluckie, Classification of Ground Clutter and Anomalous Propagation Using Dual-Polarization Weather Radar, *IEEE Transactions on Geoscience and Remote Sensing*, *46*(7), 1892–1904, doi:10.1109/TGRS.2008.916979, 2008.

Riedi, R. H., M. Crouse, V. Ribeiro, and R. Baraniuk, A multifractal wavelet model with application to network traffic, *IEEE Transactions on Information Theory*, *45*(3), 992–1018, doi:10.1109/18.761337, 1999.

Rigon, R., G. Bertoldi, and T. M. Over, GEOtop: A Distributed Hydrological Model with Coupled Water and Energy Budgets, *Journal of Hydrometeorology*, *7*(3), 371–388, doi:10.1175/JHM497.1, 2006.

Rimkus, S., S. Fatichi, and P. Burlando, TOPKAPI-ETH, a flexible, fast, distributed hydrological model for natural and managed river basins, *In Prepara*, 2013.

Rinaldo, A., and I. Rodriguez-Iturbe, Geomorphological theory of the hydrological response, *Hydrological processes*, *10*(6), 803–829, doi:10.1002/(SICI)1099-1085(199606)10:6<803::AID-HYP373>3.0.CO;2-N, 1996.

Rinehart, R., and E. Garvey, Three-dimensional storm motion detection by conventional weather radar, *Nature*, *273*, 287 – 289, doi:10.1038/273287a0, 1978.

Rodriguez-Iturbe, I., Ecohydrology: A hydrologic perspective of climate-soil-vegetation dynamics, *Water Resources Research*, *36*(1), 3–9, 2000.

Rodriguez-Iturbe, I., and P. Eagleson, Mathematical models of rainstorm events in space and time, *Water Resources Research*, *23*(1), 181–190, doi:10.1029/WR023i001p00181, 1987.

Rodriguez-Iturbe, I., D. Cox, and V. Isham, Some models for rainfall based on stochastic point processes, *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences*, *410*(1839), 269–288, doi:10.1098/rspa.1987.0039, 1987.

Rodriguez-Iturbe, I., D. R. Cox, and V. Isham, A Point Process Model for Rainfall: Further Developments, *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *417*(1853), 283–298, doi:10.1098/rspa.1988.0061, 1988.

Rodriguez-Iturbe, I., M. Marani, P. D'Odorico, and A. Rinaldo, On space-time scaling of cumulated rainfall fields, *Water resources research*, *34*(12), 3461–3469, doi:10.1029/98WR02701, 1998.

Roe, G., Orographic precipitation, *Annual review of earth and planetary sciences*, *33*, 645–671, doi:10.1146/annurev.earth.33.092203.122541, 2005.

Roldan, J., and D. A. Woolhiser, Stochastic Daily Precipitation Models 1. A Comparison of Occurence processes, *Water Resources Research*, *18*(5), 1451–1459, 1982.

Rotunno, R., and R. Ferretti, Mechanisms of intense Alpine rainfall, *Journal of the atmospheric sciences*, *58*(13), 1732–1749, doi:10.1175/1520-0469(2001)058<1732:MOIAR>2.0.CO;2, 2001.

Rotunno, R., and R. Houze, Lessons on orographic precipitation from the Mesoscale Alpine Programme, *Quarterly Journal of the Royal Meteorological Society*, *133*(625), 811–830, doi:10.1002/qj, 2007.

Ruan, F., and D. McLaughlin, An efficient multivariate random field generator using the fast Fourier transform, *Advances in Water Resources*, *21*(5), 385–399, doi:10.1016/S0309-1708(96)00064-4, 1998.

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*(W07409), doi:10.1029/2008WR007321, 2009.

Salby, M., *Fundumentals of Atmoshperic Physics*, Academic Press, San Diego,CA, 1996.

Salvucci, G. D., and D. Entekhabi, Explicit expressions for Green-Ampt (delta function diffucivity) infiltration rate and cumulative storage, *Water Resources Research*, *30*(9), 2661–2663, 1994.

Samorodnitsky, G., and M. Taqqu, *Stable non Gaussian Random Processes*, Chapman & Hall, New York, 1994.

Savina, M., The use of a cost-effective X-band weather radar in Alpine region, Ph.D. thesis, ETH Zurich, 2011.

Savina, M., B. Schäppi, P. Molnar, P. Burlando, and B. Sevruk, Comparison of a tipping-bucket and electronic weighing precipitation gage for snowfall, *Atmospheric Research*, *103*, 45–51, doi:10.1016/j.atmosres.2011.06.010, 2012.

Schertzer, D., and S. Lovejoy, Physical modeling and analysis of rain and clouds by anisotropic scaling multiplicative processes, *Journal of Geophysical Research*, *92*(D8), 9693–9714, doi:10.1029/JD092iD08p09693, 1987.

Schiemann, R., M. a. Liniger, and C. Frei, Reduced space optimal interpolation of daily rain gauge precipitation in Switzerland, *Journal of Geophysical Research*, *115*(D14), 1–18, doi:10.1029/2009JD013047, 2010.

Schiemann, R., R. Erdin, M. Willi, C. Frei, M. Berenguer, and D. Sempere-Torres, Geostatistical radar-raingauge combination with nonparametric correlograms: methodological considerations and application in Switzerland, *Hydrology and Earth System Sciences*, *15*(5), 1515–1536, doi:10.5194/hess-15-1515-2011, 2011.

Schiesser, H. H., R. A. Houze, and H. Huntrieser, The mesoscale structure of severe precipiation systems in Switzerland, *Monthly weather review*, *123*(7), 2070–2097, doi:10.1175/1520-0493(1995)123<2070:TMSOSP>2.0.CO;2, 1995.

Schleiss, M., J. Jaffrain, and A. Berne, Statistical analysis of rainfall intermittency at small spatial and temporal scales, *Geophysical Research Letters*, *38*, 1–5, doi:10.1029/2011GL049000, 2011.

Schleiss, M., J. Jaffrain, and A. Berne, Stochastic Simulation of Intermittent DSD Fields in Time, *Journal of Hydrometeorology*, *13*(2), 621–637, doi:10.1175/JHM-D-11-018.1, 2012.

Schmitt, F., S. Vannitsem, and A. Barbosa, Modeling of rainfall time series using two-state renewal processes and multifractals, *Journal of geophysical research*, *103*(D18), 23,181–23, doi:10.1029/98JD02071, 1998.

Schneebeli, M., and A. Berne, An Extended Kalman Filter Framework for Polarimetric X-Band Weather Radar Data Processing, *Journal of Atmospheric and Oceanic Technology*, *29*(5), 711–730, doi:10.1175/JTECH-D-10-05053.1, 2012.

Schwarb, M., The Alpine precipitation climate: Evaluation of a high-resolution analysis scheme using comprehensive rain-gauge data, Ph.D. thesis, ETH Zurich, Diss. No. 13911, 2000.

Serinaldi, F., Use and misuse of some Hurst parameter estimators applied to stationary and non-stationary financial time series, *Physica A: Statistical Mechanics and its Applications*, *389*(14), 2770–2781, doi:10.1016/j.physa.2010.02.044, 2010a.

Serinaldi, F., Multifractality, imperfect scaling and hydrological properties of rainfall time series simulated by continuous universal multifractal and discrete random cascade models, *Nonlinear Processes in Geophysics*, *17*(6), 697–714, doi:10.5194/npg-17-697-2010, 2010b.

Serrano, E., and A. Figliola, Wavelet Leaders : A new method to estimate the multifractal singularity spectra, *Physica A*, *388*(14), 2793–2805, doi:10.1016/j.physa.2009.03.043, 2009.

Sevruk, B., Y. Roulet, and N. V, Corrections of the Wind Induced Error of Tipping-Bucket Precipitation Gauges in Switzerland Using Numerical Simulation, *Tech. rep.*, Observing Methods Report, No. 74 (TECO-2000 Beijing, China, WMO/TD No. 1028, World Meteorological Organization), Geneva, 2000.

Shah, S., P. O'Connell, and J. R. M. Hosking, Modelling the effects of spatial variability in rainfall on catchment response. 1. Formulation and calibration of a stochastic rainfall field model, *Journal of Hydrology*, *175*(1-4), 67–88, doi:10.1016/S0022-1694(96)80006-0, 1996a.

Shah, S., P. O'Connell, and J. R. M. Hosking, Modelling the effects of spatial variability in rainfall on catchment response . 2 . Experiments with distributed and lumped models, *Journal of Hydrology*, *175*(1-4), 89–111, doi:10.1016/S0022-1694(96)80007-2, 1996b.

Sigrist, F., R. K. Hans, and W. A. Stahel, A Dynamic Non-Stationary Spatio-Temporal Model for Short Term Prediction of Precipitation, 2012a.

Sigrist, F., R. K. Hans, and W. A. Stahel, An SPDE Based Spatio-temporal Model for Large Data Sets with an Application to Postprocessing Precipitation Forecasts, 2012b.

Singh, P., and N. Kumar, Effect of orography on precipitation in the western Himalayan region, *Journal of Hydrology*, *199*(1-2), 183–206, 1997.

Singh, V., Effect of spatial and temporal variability in rainfall and watershed characteristics on stream flow hydrograph, *Hydrological processes*, *11*(12), 1649–1669, 1997.

Sponsler, G., First-Order Markov Process Representation of Binary Radar Data Sequences, *IRE Transactions on Information Theory*, *3*(1), 56–64, doi:10.1109/TIT.1957.1057395, 1957.

Springer, M., and W. Thompson, The distribution of products of independent random variables, *SIAM Journal on Applied Mathematics*, *14*(3), 511–526, 1966.

Sreenivasan, K. R., and a. Bershadskii, Clustering Properties in Turbulent Signals, *Journal of Statistical Physics*, *125*(5-6), 1141–1153, doi:10.1007/s10955-006-9112-0, 2006.

Srikanthan, R., and T. A. McMahon, Stochastic generation of annual, monthly and daily climate data: a review, *Hydrology and Earth System Science*, *5*, 653–670, 2001.

Srikanthan, R., and G. G. Pegram, A nested multisite daily rainfall stochastic generation model, *Journal of Hydrology*, *371*(1-4), 142–153, doi:10.1016/j.jhydrol.2009.03.025, 2009.

Stanley, H. E., L. a. Amaral, a. L. Goldberger, S. Havlin, Ivanov PCh, and C. K. Peng, Statistical physics and physiology: monofractal and multifractal approaches., *Physica A*, *270*(1-2), 309–24, 1999.

Steiner, M., and J. a. Smith, Use of three-dimensional reflectivity structure for automated detection and removal of nonprecipitating echoes in radar data, *Journal of Atmospheric and Oceanic Technology*, *19*(5), 673–686, 2002.

Storvik, G., A. Frigessi, and D. Hirst, Stationary space–time Gaussian fields and their time autoregressive representation, *Statistical Modelling*, *2*(2), 139–161, doi:10.1191/1471082x02st029oa, 2002.

Strangeways, I., A history of rain gauges, *Weather*, *65*(5), 133–138, 2010.

Svensson, C., J. Olsson, and R. Berndtsson, Multifractal properties of daily rainfall in two different climates, *Water Resources Research*, *32*(8), 2463–2472, doi:10.1029/96WR01099, 1996.

Swerling, P., Statistical Properties of the Contours of Random Surfaces, *IRE Transactions on Information Theory*, *8*, 230–233, 1962.

Taqqu, M., and V. Teverovsky, On estimating the intensity of long-range dependence in finite and infinite variance time series, *A practical guide to heavy tails: statistical techniques and applications*, pp. 177–218, 1998.

Taqqu, M., V. Teverovsky, and W. Willinger, Estimators for long-range dependence: an empirical study, *Fractals*, *3*(4), 785–798, 1995.

Tessier, Y., S. Lovejoy, and D. Schertzer, Universal multrifractals: Theory and observations for rain and clouds, *Journal of Applied Meteorology*, *32*(2), 223–250, doi:10.1175/1520-0450(1993)032<0223:UMTAOF>2.0.CO;2, 1993.

Thomas, H., and M. Fiering, Mathematical synthesis of streamflow sequences for the analysis of river basins by simulation, in *Design of water resources systems*, edited by M. A, Harvard University Press, Cambridge, 1962.

Thompson, S. E., and G. G. Katul, Multiple mechanisms generate Lorentzian and 1/fα power spectra in daily stream-flow time series, *Advances in Water Resources*, *37*, 94–103, doi:10.1016/j.advwatres.2011.10.010, 2012.

Todorovic, P., and D. A. Woolhiser, A stochastic model of n-day precipitation, *Journal of Applied Meteorology*, *14*, 17–24, 1975.

Todorovic, P., and V. Yevjevich, Stochastic process of precipitation, *Tech. rep.*, Colo. State Univ. Fort Collins, Fort Collins, Colorado, 1969.

Turiel, A., C. Pérez-Vicente, and J. Grazzini, Numerical methods for the estimation of multifractal singularity spectra on sampled data: a comparative study, *Journal of Computational Physics*, *216*(1), 362–390, doi:10.1016/j.jcp.2005.12.004, 2006.

Vandenberghe, S., N. E. C. Verhoest, and B. De Baets, Fitting bivariate copulas to the dependence structure between storm characteristics: A detailed analysis based on 105 year 10 min rainfall, *Water Resources Research*, *46*(1), 1–17, doi:10.1029/2009WR007857, 2010.

Vanhaute, W. J., S. Vandenberghe, K. Scheerlinck, B. De Baets, and N. E. C. Verhoest, Calibration of the modified Bartlett-Lewis model using global optimization techniques and alternative objective functions, *Hydrology and Earth System Sciences Discussions*, *8*(6), 9707–9756, doi:10.5194/hessd-8-9707-2011, 2011.

Veneziano, D., Basic properties and characterization of stochastic self-similar processes in Rd, *Fractals*, *7*(1), 59–78, 1999.

Veneziano, D., Iterated random pulse processes and their spectral properties, *Fractals-an Interdisciplinary Journal on the*, *10*(1), 1–11, doi:10.1142/S0218348X02000884, 2002.

Veneziano, D., and P. Furcolo, A modified double trace moment method of multifractal analysis, *Fractals*, *7*(2), 181–195, 1999.

Veneziano, D., and P. Furcolo, Marginal distribution of stationary multifractal measures and their Haar wavelet coefficients, *Fractals-an Interdisciplinary Journal on*, *11*(3), 253–270, doi:10.1142/S0218348X03002051, 2003.

Veneziano, D., and P. Furcolo, Improved moment scaling estimation for multifractal signals, *Nonlinear Processes in Geophysics*, *16*, 641–653, 2009.

Veneziano, D., and V. Iacobellis, Self-similarity and multifractality of topographic surfaces at basin and subbasin scales, *Journal of geophysical research*, *104*(B6), 797–812, doi:10.1029/1999JB900083, 1999.

Veneziano, D., and A. Langousis, Scaling and fractals in hydrology, in *Advances in Data-based Approaches for Hydrologic Modeling and Forecasting*, 2010.

Veneziano, D., and C. Lepore, The scaling of temporal rainfall, *Water Resources Research*, *48*(8), 1–16, doi:10.1029/2012WR012105, 2012.

Veneziano, D., P. Furcolo, and V. Iacobellis, Imperfect scaling of time and space-time rainfall, *Journal of Hydrology*, *322*(1-4), 105–119, doi:10.1016/j.jhydrol.2005.02.044, 2006a.

Veneziano, D., A. Langousis, and P. Furcolo, Multifractality and rainfall extremes: A review, *Water Resources Research*, *42*(6), doi:10.1029/2005WR004716, 2006b.

Veneziano, D., A. Langousis, and C. Lepore, New asymptotic and preasymptotic results on rainfall maxima from multifractal theory, *Water Resources Research*, *45*(11), W11,421, doi:10.1029/2009WR008257, 2009.

Venugopal, V., S. G. Roux, E. Foufoula-Georgiou, and A. Arneodo, Revisiting multifractality of high-resolution temporal rainfall using a wavelet-based formalism, *Water Resources Research*, *42*(6), W06D14, doi:10.1029/2005WR004489, 2006.

Verhoest, N. E. C., S. Vandenberghe, P. Cabus, C. Onof, T. Meca-Figueras, and S. Jameleddine, Are stochastic point rainfall models able to preserve extreme flood statistics?, *Hydrological Processes*, *24*(23), 3439–3445, doi:10.1002/hyp.7867, 2010.

Verrier, S., L. de Montera, L. Barthès, and C. Mallet, Multifractal analysis of African monsoon rain fields, taking into account the zero rain-rate problem, *Journal of Hydrology*, *389*(1-2), 111–120, doi:10.1016/j.jhydrol.2010.05.035, 2010.

Viglione, A., G. B. Chirico, R. Woods, and G. Blöschl, Generalised synthesis of space–time variability in flood response: An analytical framework, *Journal of Hydrology*, *394*(1-2), 198–212, doi:10.1016/j.jhydrol.2010.05.047, 2010.

Villarini, G., and W. Krajewski, Inference of Spatial Scaling Properties of Rainfall: Impact of Radar Rainfall Estimation Uncertainties, *IEEE Geoscience and Remote Sensing Letters*, *6*(4), 812–815, doi:10.1109/LGRS.2009.2025891, 2009a.

Villarini, G., and W. F. Krajewski, Review of the Different Sources of Uncertainty in Single Polarization Radar-Based Estimates of Rainfall, *Surveys in Geophysics*, *31*(1), 107–129, doi:10.1007/s10712-009-9079-x, 2009b.

Villarini, G., G. Ciach, W. F. Krajewski, K. Nordstrom, and V. K. Gupta, Effects of systematic and random errors on the spatial scaling properties in radar-estimated rainfall, *Nonlinear dynamics in geosciences*, pp. 37–51, 2007.

Villarini, G., W. F. Krajewski, G. J. Ciach, and D. L. Zimmerman, Product-error-driven generator of probable rainfall conditioned on WSR-88D precipitation estimates, *Water Resources Research*, *45*(1), 1–11, doi:10.1029/2008WR006946, 2009.

Vivoni, E. R., V. Y. Ivanov, R. L. Bras, and D. Entekhabi, On the effects of triangulated terrain resolution on distributed hydrologic model response, *Hydrological Processes*, *19*(11), 2101–2122, doi:10.1002/hyp.5671, 2005.

Vivoni, E. R., K. Tai, and D. J. Gochis, Effects of Initial Soil Moisture on Rainfall Generation and Subsequent Hydrologic Response during the North American Monsoon, *Journal of Hydrometeorology*, *10*(3), 644–664, doi:10.1175/2008JHM1069.1, 2009.

Volpi, E., M. Di Lazzaro, and A. Fiori, A simplified framework for assessing the impact of rainfall spatial variability on the hydrologic response, *Advances in Water Resources*, *46*, 1–10, doi:10.1016/j.advwatres.2012.04.011, 2012.

Vuerich, E., C. Monesi, L. Lanza, L. Stagi, and E. Lanzinger, Instruments and Observing methods. Report No 99. WMO Field intercomparison of rainfall intensity gauges, *Tech. Rep. 99*, 2009.

Wallace, J., and P. Hobbs, *Atmospheric Science: An Introduction Survey*, Academic Press, 1977.

Waymire, E., Scaling limits and self-similarity in precipitation fields, *Water Resources Research*, *21*(8), 1271–1281, doi:10.1029/WR021i008p01271, 1985.

Waymire, E., V. K. Gupta, and I. Rodriguez-Iturbe, A spectral theory for rainfall intensity at meso-beta scale, *Water Resources Research*, *20*(10), 1453–1465, doi:10.1029/WR020i010p01453, 1984.

Waymire, E. C., and S. Williams, A cascade Decomposition Theory with Applications to Markov and Exchengeable Cascades, *Transactions of the Amererical Mathematical Society*, *348*(2), 585–632, 1996.

Webster, P., and R. Lukas, TOGA COARE: The coupled ocean-altmosphere response experiment, *Bulletin of the Americal Meteorological Society*, *73*(9), 1377–1416, 1992.

Webster, R., and M. Oliver, *Geostatistics for Environmental scientists*, John Wiley & Sons, Chichester, UK, 2001.

Welch, P., The use of fast Fourier transform for the estimation of power spectra: a method based on time averaging over short, modified periodograms, *IEEE Transactions on Audio and Electroacoustics*, *15*(2), 70–73, 1967.

Wilks, D., Multisite generalization of a daily stochastic precipitation generation model, *Journal of Hydrology*, *210*(1-4), 178–191, doi:10.1016/S0022-1694(98)00186-3, 1998.

Wilks, D., Simultaneous stochastic simulation of daily precipitation, temperature and solar radiation at multiple sites in complex terrain, *Agricultural and Forest Meteorology*, *96*(1-3), 85–101, doi:10.1016/S0168-1923(99)00037-4, 1999.

Woods, R., and M. Sivapalan, A synthesis of space-time variability in storm response: Rainfall, runoff generation, and routing, *Water Resources Research*, *35*(8), 2469–2485, 1999.

Wüest, M., C. Frei, A. Altenhoff, M. Hagen, M. Litschi, and C. Schär, A gridded hourly precipitation dataset for Switzerland using rain-gauge analysis and radar-based disaggregation, *International Journal of Climatology*, *30*(12), 1764–1775, doi:10.1002/joc.2025, 2010.

Zawadzki, I., and C. Ro, Correlation between maximum rate of precipitation and mesoscale parameters, *Journal of Applied Meteorology*, *17*, 1327–1334, 1978.

Zehe, E., and G. Blöschl, Predictability of hydrologic response at the plot and catchment scales: Role of initial conditions, *Water Resources Research*, *40*(10), 1–21, doi:10.1029/2003WR002869, 2004.

Zhang, S., and E. Foufoula-Georgiou, Subgrid-scale rainfall variability and its effects on atmospheric and surface variable predictions, *Journal of Geophysical Research*, *102*(D16), 19,559–19,573, doi:10.1029/97JD01392, 1997.

A

RAINFALL-RUNOFF SIMULATION STATISTICS

TABLE A.1: Statistics of the Peak discharge $[m^3s^{-1}]$. cv corresponds to the coefficient of variation $[-]$, cor to the correlation length [km], adv to the advection speed and in state the initial soil saturation ratio

				Imf = 5 mmh^{-1}		Imf = 10 mmh^{-1}	
cv [-]	cor [km]	adv [kmh^{-1}]	In state [-]	Mean	Std	Mean	Std
1	10	0	0.2	486.49	24.88	1043.52	48.93
1	10	0	0.5	225.61	19.28	603.34	56.44
1	10	0	0.8	54.23	12.67	232.36	56.94
1	10	15	0.2	479.57	32.74	1083.22	99.43
1	10	15	0.5	207.90	14.86	552.35	49.37
1	10	15	0.8	37.69	3.04	144.58	16.18
1	10	-15	0.2	460.80	23.33	1040.35	65.10
1	10	-15	0.5	199.36	11.94	533.50	41.98
1	10	-15	0.8	37.05	2.02	140.76	11.53
1	30	0	0.2	481.69	28.06	1035.60	57.03
1	30	0	0.5	221.77	21.24	587.85	63.04
1	30	0	0.8	49.22	11.03	217.97	70.37
1	30	15	0.2	500.11	37.64	1115.54	99.29
1	30	15	0.5	217.16	18.12	569.87	58.62
1	30	15	0.8	38.55	3.15	149.87	22.57
1	30	-15	0.2	471.01	33.36	1066.96	70.52
1	30	-15	0.5	201.21	15.20	542.45	44.71
1	30	-15	0.8	36.96	2.45	141.62	14.29
3	10	0	0.2	562.90	52.33	1167.22	87.49
3	10	0	0.5	312.25	59.12	790.91	100.82
3	10	0	0.8	131.63	59.55	446.18	110.81
3	10	15	0.2	515.51	43.43	1202.54	122.85
3	10	15	0.5	226.20	19.73	637.99	92.85
3	10	15	0.8	43.32	4.39	204.40	40.61
3	10	-15	0.2	503.50	42.59	1116.28	113.65
3	10	-15	0.5	224.38	22.74	600.27	80.95
3	10	-15	0.8	44.80	7.21	188.17	36.41
3	30	0	0.2	552.02	47.18	1180.55	114.13
3	30	0	0.5	300.31	49.71	787.46	128.37
3	30	0	0.8	118.71	47.42	441.31	132.44
3	30	15	0.2	548.55	55.55	1260.63	134.65
3	30	15	0.5	239.60	27.43	666.42	99.44
3	30	15	0.8	44.88	6.88	204.70	54.05
3	30	-15	0.2	512.91	41.86	1128.02	110.92
3	30	-15	0.5	225.59	22.14	612.49	84.47
3	30	-15	0.8	43.66	7.08	194.17	41.57

TABLE A.2: Statistics of the time to peak discharge [h]. cv corresponds to the coefficient of variation [−], cor to the correlation length [km], adv to the advection speed and in state the initial soil saturation ratio

				Imf = 5 mmh^{-1}		Imf = 10 mmh^{-1}	
cv [-]	cor [km]	adv [kmh^{-1}]	In state [-]	Mean	Std	Mean	Std
1	10	0	0.2	15.33	0.67	14.57	0.73
1	10	0	0.5	17.01	0.94	15.69	0.87
1	10	0	0.8	21.35	1.72	18.01	1.34
1	10	15	0.2	15.69	1.32	14.95	1.16
1	10	15	0.5	17.27	1.31	16.14	1.18
1	10	15	0.8	22.94	1.52	19.70	1.32
1	10	-15	0.2	15.80	1.12	14.51	0.99
1	10	-15	0.5	17.25	1.13	15.92	0.97
1	10	-15	0.8	22.82	1.48	19.22	0.99
1	30	0	0.2	15.49	0.78	14.58	0.65
1	30	0	0.5	17.07	0.91	15.75	0.83
1	30	0	0.8	21.39	1.47	18.09	1.25
1	30	15	0.2	15.40	1.05	14.45	1.21
1	30	15	0.5	17.18	1.21	15.86	1.35
1	30	15	0.8	22.76	1.51	19.40	1.28
1	30	-15	0.2	15.48	1.19	14.44	1.07
1	30	-15	0.5	17.21	1.10	15.66	1.13
1	30	-15	0.8	23.17	1.24	19.29	1.21
3	10	0	0.2	14.91	1.24	14.25	1.22
3	10	0	0.5	15.76	1.53	14.88	1.39
3	10	0	0.8	17.96	3.06	15.80	1.78
3	10	15	0.2	15.36	1.36	14.08	1.30
3	10	15	0.5	17.39	1.48	15.43	1.59
3	10	15	0.8	22.78	1.76	17.67	1.68
3	10	-15	0.2	15.78	1.36	14.32	1.16
3	10	-15	0.5	17.28	1.32	15.50	1.46
3	10	-15	0.8	22.08	1.36	18.01	1.55
3	30	0	0.2	14.78	1.16	13.72	1.11
3	30	0	0.5	16.26	1.57	14.63	1.34
3	30	0	0.8	18.73	2.55	15.94	1.84
3	30	15	0.2	15.07	1.36	14.18	1.43
3	30	15	0.5	16.77	1.47	15.58	1.33
3	30	15	0.8	21.99	1.82	17.89	1.43
3	30	-15	0.2	15.37	1.26	14.28	1.36
3	30	-15	0.5	16.94	1.11	15.47	1.52
3	30	-15	0.8	22.26	1.26	17.95	1.79

TABLE A.3: Statistics of the total discharge volume $[\mathrm{m}^3]$. cv corresponds to the coefficient of variation $[-]$, cor to the correlation length [km], adv to the advection speed and in state the initial soil saturation ratio

				Imf = 5 mmh^{-1}		Imf = 10 mmh^{-1}	
cv [-]	cor [km]	adv [kmh^{-1}]	In state [-]	Mean	Std	Mean	Std
1	10	0	0.2	40582381.20	162254.01	62331061.86	318319.50
1	10	0	0.5	18753386.63	268586.33	35532297.79	753821.49
1	10	0	0.8	4827641.78	291389.44	14100567.91	1110895.94
1	10	15	0.2	40240766.23	29081.00	61479283.82	87303.98
1	10	15	0.5	18146274.31	26938.71	33534583.60	180467.25
1	10	15	0.8	4163136.36	48750.06	11270835.28	179109.69
1	10	-15	0.2	40241582.98	27828.89	61499793.09	67321.65
1	10	-15	0.5	18149352.49	37079.95	33572760.43	134661.79
1	10	-15	0.8	4161306.14	50533.67	11290370.93	148960.62
1	30	0	0.2	40509266.59	161485.75	62187469.25	439379.29
1	30	0	0.5	18629146.39	247758.80	35150849.26	1054443.68
1	30	0	0.8	4666972.47	253974.13	13521306.59	1586742.42
1	30	15	0.2	40252475.54	36542.77	61509103.08	89668.23
1	30	15	0.5	18154296.07	32256.50	33597530.79	198595.40
1	30	15	0.8	4153462.02	50919.65	11322899.44	214601.36
1	30	-15	0.2	40254164.02	33484.73	61500358.18	82627.38
1	30	-15	0.5	18155686.10	34814.17	33572130.41	185323.40
1	30	-15	0.8	4164788.02	56438.49	11288112.99	197775.22
3	10	0	0.2	41384103.69	418756.76	63589329.74	603016.25
3	10	0	0.5	20403665.84	923840.61	38565406.29	1491195.03
3	10	0	0.8	6980283.92	1337933.08	19047200.68	2489831.76
3	10	15	0.2	40296374.74	58720.00	61667026.92	133873.17
3	10	15	0.5	18276803.39	77976.28	34099411.18	308523.31
3	10	15	0.8	4284175.25	93945.48	12203894.86	431914.17
3	10	-15	0.2	40306152.12	71613.81	61677823.93	133328.21
3	10	-15	0.5	18288739.02	101924.27	34100948.82	314777.54
3	10	-15	0.8	4307960.39	131190.20	12143979.46	445557.87
3	30	0	0.2	41248746.84	374669.38	63491339.95	650030.38
3	30	0	0.5	20131466.77	818115.28	38329399.10	1567223.69
3	30	0	0.8	6622192.70	1155917.80	18525889.09	2431275.09
3	30	15	0.2	40306397.14	78927.73	61650952.74	165664.80
3	30	15	0.5	18292422.77	108361.64	34018395.62	384880.57
3	30	15	0.8	4297799.63	104202.28	12042977.43	519782.57
3	30	-15	0.2	40320815.46	77637.13	61701361.88	123847.84
3	30	-15	0.5	18307871.30	100369.78	34116027.23	282053.16
3	30	-15	0.8	4311155.08	111210.10	12110990.07	387943.75

B

STATISTICS OF THE POISSON CLUSTER MODELS

B.1 The Neyman Scott model

In this Appendix the statistics properties of both the temporal and spatiotemporal version of the Neyman Scott models are given. The equations are derived from [*Cowpertwait et al.*, 1996, 2002].

Let Y_h be the depth accumulation for an aggregation interval h for some location $\mathbf{x} \in \Re^2$ defined as the superposition of raincells X. X in the temporal version are one dimensional rectangular pulses. In the spatiotemporal version are two dimensional circular discs the centres of which form a homogeneous two dimensional Poisson process with rate φ [km^{-2}]. The radius of the raincells follow an exponential distribution with rate ϕ [km^{-1}].

The mean is:

$$\mu_h = E[Y_h] = \lambda \mu_c E[X] h/\eta, \tag{B.1}$$

where μ_c is expected value of the number of cells in a storm.

In the case of the spatiotemporal version of the model μ_c is defined as

$$\mu_c = \frac{2\pi\varphi}{\phi^2}. \tag{B.2}$$

$E[X]$ is the expected value of the precipitation intensity of each raincell, and is dependent on the choice of its probability distribution. Popular distributions are the exponential, the Gamma, the Weibull and the Gumbel among others. The distribution has to be defined only for non negative values.

The temporal covariance for the one dimensional version of the model is defined as

$$\begin{aligned} \gamma_{h,l} &= \text{Cov}[Y_h^i, Y_h^{i+l}] = \lambda\eta^{-3} A(h,l)[2\mu_c E[X^2] + [E[X]]^2 \beta_c^2 E\{C^2 - C\}/(\beta_c^2 - \eta^2)] \\ &\quad -\lambda [E[X]]^2 B(h,l) E[C^2 - C]/[\beta_c(\beta_c^2 - \eta^2)] \end{aligned} \tag{B.3}$$

and the generalization of the space-time covariance structure of the spatiotemporal model is

$$\gamma_{x,y,h,l} = \mathrm{Cov}[Y_h^i(\mathbf{x}), Y_h^{i+l}(\mathbf{y})] \quad \text{(B.4)}$$
$$= \gamma_{h,l} - \lambda\{1 - P(\phi, d)\}\mu_c E[X^2] A(h,l)/\eta^3, \quad \text{(B.5)}$$

where h is the time aggregation and $l \geq 0$ is a integer lag of the autocorrelation and $A(h,l)$, $B(h,l)$, $P(\phi, d)$ are defined as

$$\begin{aligned} A(h,l) &= h\eta + e^{-\eta h} - 1\,, \quad \text{if } l = 0\,, \\ A(h,l) &= 0.5\left(1 - e^{-\eta h}\right)^2 e^{-\eta h(l-1)}\,, \quad \text{if } l > 0\,, \end{aligned} \quad \text{(B.6)}$$

$$\begin{aligned} B(h,l) &= h\beta_c + e^{-\beta_c h} - 1\,, \quad \text{if } l = 0\,, \\ B(h,l) &= 0.5\left(1 - e^{-\beta_c h}\right)^2 e^{-\beta_c h(l-1)}\,, \quad \text{if } l > 0\,, \end{aligned} \quad \text{(B.7)}$$

$$P(\phi, d) = \frac{2}{\pi}\int_0^{\pi/2}\left(\frac{\phi d}{2\cos y} + 1\right)\exp\left(\frac{-\phi d}{2\cos y}\right), \quad \text{where} \quad d = \|\mathbf{x} - \mathbf{y}\| \quad \text{(B.8)}$$

The third moment, $\xi_h = E[([Y_h - E[Y_h])^3]$,

$$\begin{aligned} \xi_h &= E[([Y_h - E[Y_h])^3] = 6\lambda\mu_c\, E[X^3](\eta h - 2 + \eta h e^{-\eta h} + 2e^{-\eta h})/\eta^4 \\ &\quad +3\,\lambda E[X]E[X^2]E[C(C-1)]f(\eta, \beta_c, h)/[2\eta^4\beta_c(\beta_c^2 - \eta^2)^2] \\ &\quad +\lambda E[X]^3 E[(C^2 - C)(C-2)]g(\eta, \beta_c, h) \\ &\quad /[2\eta^4\beta_c(\eta^2 - \beta_c^2)(\eta - \beta_c)(2\beta_c + \eta)(\beta_c + 2\eta)], \end{aligned} \quad \text{(B.9)}$$

where the function $f(\eta, \beta_c, h)$ and $g(\eta, \beta_c, h)$ are defined as:

$$\begin{aligned} f(\eta, \beta_c, h) &= -2\eta^3\beta_c^2 e^{-\eta h} - 2\eta^3\beta_c^2 e^{-\beta_c h} + \eta^2\beta_c^3 e^{-2\eta h} + 2\eta^4\beta_c e^{-\eta h} \\ &\quad +2\eta^4\beta_c e^{-\beta_c h} + 2\eta^3\beta_c^2 e^{-(\eta+\beta_c)h} - 2\eta^4\beta_c e^{-(\eta+\beta_c)h} - 8\eta^3\beta_c^3 h + 11\eta^2\beta_c^3 - 2\eta^4\beta_c \\ &\quad +2\eta^3\beta_c^2 + 4\eta\beta_c^5 h + 4\eta^5\beta_c h - 7\beta_c^5 - 4\eta^5 + 8\beta_c^5 e^{-\eta h} - \beta_c^5 e^{-2\eta h} \\ &\quad -2h\eta^3\beta_c^3 e^{-\eta h} - 12\eta^2\beta_c^3 e^{-\eta h} + 2h\eta\beta_c^5 e^{-\eta h} + 4\eta^5 e^{-\beta_c h} \end{aligned} \quad \text{(B.10)}$$

$$\begin{aligned} g(\eta, \beta_c, h) &= 12\eta^5\beta_c e^{-\beta_c h} + 9\eta^4\beta_c^2 + 12\eta\beta_c^5 e^{-\eta h} + 9\eta^2\beta_c^4 + 12\eta^3\beta_c^3 e^{-(\eta+\beta_c)h} \\ &\quad -\eta^2\beta_c^4 e^{-2\eta h} - 12\eta^3\beta_c^3 e^{-\beta_c h} - 9\eta^5\beta_c - 9\eta\beta_c^5 - 3\eta\beta_c^5 e^{-2\eta h} \\ &\quad -\eta^4\beta_c^2 e^{-2\beta_c h} - 12\eta^3\beta_c^3 e^{-\eta h} + 6\eta^5\beta_c^2 h - 10\eta^3\beta_c^4 h + 6\eta^2\beta_c^5 h \\ &\quad -10\eta^4\beta_c^3 h + 4\eta\beta_c^6 h - 8\beta_c^2\eta^4 e^{-\beta_c h} + 4\beta_c\eta^6 h + 12\beta_c^3\eta^3 \\ &\quad -8\beta_c^4\eta^2 e^{-\eta h} - 6\eta^6 - 6\beta_c^6 - 2\eta^6 e^{-2\beta_c h} - 2\beta_c^6 e^{-2\eta h} \\ &\quad +8\eta^6 e^{-\beta_c h} + 8\beta_c^6 e^{-\eta h} - 3\beta_c\eta^5 e^{-2\beta_c h}. \end{aligned} \quad \text{(B.11)}$$

The probability that an arbitrary interval of length h is dry, $\psi(h) = P(Y_h = 0)$, was derived from *Cowpertwait et al.* [1996].

$$\begin{aligned}\psi(h) &= \exp\Big(-\lambda\, h + \lambda\, \beta_c^{-1}\mu_c^{-1}[1 - e^{(-\mu_c + \mu_c e^{-\beta_c h})}] \\ &\quad -\lambda \int_0^\infty \big[1 - p_h(t)\big]\, dt\Big) \end{aligned} \tag{B.12}$$

where $p_h(t)$ is function of h, β_c, η and μ_c:

$$\begin{aligned}p_h(t) &= \Big[e^{-\beta_c(t+h)} + 1 - (\eta e^{-\beta_c t} - \beta_c e^{-\eta t})/(\eta - \beta_c)\Big] \cdot \\ &\quad \exp\Big[-\mu_c\beta_c(e^{-\beta_c t} - e^{-\eta t})/(\eta - \beta_c) - \mu_c e^{-\beta_c\, t} + \mu_c e^{-\beta_c(t+h)}\Big]. \end{aligned} \tag{B.13}$$

The conditional transition probabilities $Pdry - dry$ and $Pwet - wet$ are

$$P\left\{Y_{i+1}^{(h)} = 0 | Y_i^{(h)} = 0\right\} = \psi_{DD} = \psi(2h)/\psi(h) \tag{B.14}$$

and

$$P\left\{Y_{i+1}^{(h)} > 0 | Y_i^{(h)} > 0\right\} = \psi_{WW}(h) = \{1 - 2\psi(h) + \psi(2h)\} \,/\, \{1 - \psi(h)\}. \tag{B.15}$$

B.2 The Modified Bartlett Lewis model

The derivation of the statistics of the temporal Bartlett Lewis model are similar to the Neymann-Scott one [*Rodriguez-Iturbe et al.*, 1987]. The notation that is presented here follows *Islam and Entekhabi* [1990]; *Onof and Wheater* [1993].

First let us define

$$\kappa = \beta_c/\eta, \tag{B.16}$$

$$\phi = \gamma/\eta, \tag{B.17}$$

as two non-dimensional parameters and η be gamma distributed as $\sim \Gamma(\alpha_\eta, \nu_\eta)$. Then the mean is:

$$\mu_h = E[Y_h] = \frac{\lambda\mu_c E[X]h}{\alpha_\eta - 1}, \quad \mu_c = 1 + \kappa/\phi. \tag{B.18}$$

The variance is

$$\begin{aligned}\mathrm{Var}[Y_h] &= 2A1[(\alpha_\eta-3)h{\nu_\eta}^{2-\alpha_\eta}-{\nu_\eta}^{3-\alpha_\eta}+(\nu_\eta+h)^{3-\alpha_\eta}] \\ &\quad -2A2[\phi(\alpha_\eta-3)h{\nu_\eta}^{2-\alpha_\eta}-{\nu_\eta}^{3-\alpha_\eta}+(\nu_\eta+\phi h)^{3-\alpha_\eta}],\end{aligned} \tag{B.19}$$

where

$$A1 = \frac{\lambda\mu_c{\nu_\eta}^{\alpha_\eta}}{(\alpha_\eta-1)(\alpha_\eta-2)(\alpha_\eta-3)}\left[E(X^2)+\frac{\kappa\phi E(X)^2}{\phi^2-1}\right], \tag{B.20}$$

$$A2 = \frac{\lambda\mu_c{\nu_\eta}^{\alpha_\eta}\kappa E(X)^2}{\phi^2(\phi^2-1)(\alpha_\eta-1)(\alpha_\eta-2)(\alpha_\eta-3)}. \tag{B.21}$$

The covariance is defined as

$$\begin{aligned}\mathrm{Cov}[Y_h(i),Y_h(i+k)] &= A1\{[\nu_\eta+(k+1)h]^{3-\alpha_\eta}-2(\nu_\eta+kh)^{3-\alpha_\eta}+[\nu_\eta+(k-1)h]^{3-a}\} \\ &\quad -A2\{[\nu_\eta+(k+1)\phi h]^{3-\alpha_\eta}-2(\nu_\eta+k\phi h)^{3-\alpha_\eta}+[\nu_\eta+(k-1)\phi h]^{3-a}\}\end{aligned} \tag{B.22}$$

and the probability of zero precipitation can be approximated as

$$\begin{aligned}\psi(h) \approx \exp\Bigg(&-\lambda h-\frac{\lambda\phi^{-1}\nu_\eta}{\alpha_\eta-1}\left[1+\phi(\kappa+\phi/2)-\frac{1}{4}\phi(5\phi\kappa+\kappa^2+2\phi^2)\right. \\ &\left.+\frac{1}{72}\phi(4\kappa^3+31\kappa^2\phi+99\kappa\phi^2+36\phi^3)\right]+\frac{\lambda\phi^{-1}\nu_\eta}{\alpha_\eta-1} \\ &\times\left(1-\kappa-\phi+\frac{3}{2}\kappa\phi+\phi^2+\frac{1}{2}\kappa^2\right)\left\{\frac{\phi}{\phi+\kappa}+\frac{\kappa}{\phi+\kappa}\left[\frac{\nu_\eta}{\nu_\eta+(\nu_\eta+\phi)h}\right]^{\alpha_\eta-1}\right\}\Bigg).\end{aligned} \tag{B.23}$$

The mean value of the cell intensity for the gamma distribution $X\sim\Gamma(\alpha_c,\theta_c)$ is

$$E[X]=\alpha_c\theta_c \tag{B.24}$$

and

$$E[X^2]=\alpha_c\theta_c^2(1+\theta_c). \tag{B.25}$$

If a Weibull distribution is assumed for the cell intensities $X\sim\mathrm{Wbl}(\alpha_c,\theta_c)$

$$E[X]=\alpha_c\Gamma(1+1/\theta_c) \tag{B.26}$$

and

$$E[X^2] = \alpha_c^2 \Gamma(1 + 2/\theta_c). \tag{B.27}$$