Statistical Models for Extreme Events in Atmospheric Processes

by

Enrico Zorzetto

Division of Earth and Ocean Sciences

Duke University

Date: ______________________________

Approved:

____

Marco Marani, Advisor

____

A. Brad Murray

____

Ana P. Barros

____

Gabriel G. Katul

Dissertation submitted in partial fulfillment of the
requirements for the degree of Doctor of Philosophy
in the Division of Earth and Ocean Sciences
in the Graduate School of Duke University

2020

<u>ABSTRACT</u>

Statistical Models for Extreme Events in Atmospheric Processes

by

Enrico Zorzetto

Division of Earth and Ocean Sciences

Duke University

Date: ______________________________

Approved:

____

Marco Marani, Advisor

____

A. Brad Murray

____

Ana P. Barros

____

Gabriel G. Katul

An abstract of a dissertation submitted in partial fulfillment of the
requirements for the degree of Doctor of Philosophy
in the Division of Earth and Ocean Sciences
in the Graduate School of Duke University

2020

Abstract

Extreme events - that is, intense events so rare to be poorly represented in historical observational records - play a fundamental role in atmospheric processes, and can have far reaching consequences ranging from impacts on society, economy, the environment, as well as on the global water and energy budgets. However, characterizing the statistical properties of such events is a challenging task, as (i) by definition extremes are poorly sampled, and thus studying them often requires extrapolation beyond the range of available observations, and (ii) extremes are often the result of nonlinear and intermittent processes, which determine significant difficulties both in predicting them and in studying their frequency of occurrence beyond the range of observations. This dissertation focuses on new statistical methods specifically aimed at characterizing extremes events in rainfall and boundary layer turbulence, including contributions along three main lines of inquiry:

1. Developing extreme value models able to reduce estimation uncertainty in the case of short rainfall time series. To this end, a non-asymptotic approach is developed which, deviating from traditional extreme value theory, models the entire distribution of rainfall magnitudes and frequency of occurrence. Using compound distributions and the structure of latent-level Bayesian models, this

framework accounts for the effects of low-frequency variability of rainfall statistics on the tail decay of their probability distribution.

2. Characterizing the frequency of extreme values from remotely-sensed rainfall estimates. This objective is approached by developing a downscaling technique that allows comparing rainfall statistics across different spatial averaging scales, and by constructing a model of the error so as to permit their validation over poorly gauged locations. The framework developed here now allows for the production of large-scale estimates of the frequency of extreme rainfall based on satellite-derived rainfall datasets and their validation even in data-scarce regions.

3. Investigating the dynamics of scalar quantities transported in the atmospheric boundary layer, with a focus on fluxes of sensible heat and methane. In the case of sensible heat, I studied to what extent the extreme values properties of temperature fluctuations retain information on the turbulence generation mechanism. In the case of methane, I focused instead on an inverse problem: given the observed statistical properties of methane concentration fluctuations, is it possible to infer the spatial intermittency of its source at the ground? In both cases, I found that statistical properties of the scalar, including its extreme

value statistics, can be used to improve characterization of the turbulent flow and of its boundary conditions.

to my parents, Mario and Ivana

Acknowledgements

This achievement would not have been possible without the support of my family, and of the many longtime friends, mentors and colleagues which supported me during these years. In particular, I wish to thank my advisor Marco Marani, for his support and constant guidance during my Ph.D. at Duke. Thank you for giving me the opportunity of working together, and for passing on to me the fascination for extreme events! A special thank you goes to Gaby Katul, who has been a wonderful mentor and a constant source of inspiration. I also wish to thank my committee members Ana P. Barros and A. Brad Murray for their advice and guidance in countless occasions during my time at Duke. I am also indebted to several other collaborators: Antonio Canale, Gianluca Botter, Andrew D. Bragg, Olli Peltola, and Tiia Grönholm, Francesco Marra, Efrat Morin, Emmanouil Anagnostou, and Efthymios Nikolopoulos. I thank all the amazing professors whose classes I had the fortune to attend back at the University of Padova and here at Duke. In particular, I want to mention here Luigi D'Alpaos, Andrea Rinaldo, Thomas Witelski, Robert Behringer, and Amilcare Porporato. A special thanks to the many friends and colleagues I met in Durham: Fateme Yousefi, Bijan Seyanasdrollah, Salvatore Calabrese, Norman Frank, John Gardner, Khaled Ghannam, Jun Yin, Sara Bonetti, Gabriele Manoli, Assaad Mrad, Mazen Nakad, Shuolin Li, Marcella Roner, Saverio Perri, Yanlan Liu, Masih Eghdami, Mark Bartlett, Samantha Hartzell, Emilie Lefevre, Yun Jian, Stefano Basso, Arianna Miniussi, and many others. A special thank you also goes to all my friends and colleagues in the EOS department, and to the friends in the *'ricreazione'* group in Padova, for the wonderful time spent in their Ph.D. lab, and for the countless cups of coffee shared. Finally, a special thank you to Rachel. It has been wonderful having you by my side during these years. I acknowledge funding from the National Aeronautics and Space Administration through the NASA Earth and Space Science Fellowship program (fellowship 17-EARTH17F-0270, 18-EARTH18R- 0089, and 19-EARTH19R-27), from the National Science Foundation (grant NSF-EAR-13-44703) and appreciate student support from the Division of Earth and Ocean Sciences, Duke University.

Contents

Biography **291**

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Background and Motivation

Extreme events are usually defined as those events so rare to be poorly represented in historical observational records [Col01, DHF07]. From a physical perspective, they can also be defined by individuating regions of the phase space of a system which are rarely visited by the system's trajectories [AK05]. Clearly, the definition of "*rarity*", and thus the same definition of extreme event, depends in general on the observer's timescale, and on its willingness to wait "long enough". In the context of atmospheric and hydrological processes, extreme events can have far reaching consequences ranging from impacts on society, economy, the environment, as well as on the global water and energy budgets. It is precisely the dynamics of these impacted physical systems which dictates the range of waiting times we are interested in when studying extreme values of a given process. Thus, the metric used to quantify the rarity of an event is conventionally the average recurrence interval, or *return time* of that event. In this dissertation I focus on the statistical modelling of two main phenomena: Rainfall on the one hand, and the transport of scalar quantities in the atmospheric boundary layer on the other. While both these phenomena are deeply rooted in the dynamics of atmospheric turbulent flows, they also exhibit substantial differences. Notably, the different range of temporal and spatial scales which are under investigation, and the different nature of the available observations, ranging from high-frequency in-situ measurements to large-scale remotely-sensed estimates. In both contexts, characterizing the frequency of such extreme events in atmospheric processes is a challenging task due to the nonlinear, multiscale and chaotic nature of the physical processes involved. For this reason, observations and statistical methods have been instrumental in characterizing the variability of these processes, and in guiding the development of phenomenological theories. As a representative example, recognizing the importance of extreme fluctuations in energy dissipation rates led to one of the major developments in phenomenological theory of turbulence [LL59, Kol62, FD96]. Similarly, drawing from its similarity with turbulence, the adoption of multifractal models for rainfall led to significant advanced in describing the space-time fluctuations of rainfall fields [SL87, OG94, Mar05, NB15]. However, estimating statistical properties of extremes remains in general a very challenging task, since data are '*scarce*' by definition. Additionally, the complex nonlinear nature of the processes studied here determines a basic difficulty in predicting the occurrence of future values, and in characterizing their frequency beyond the range of observations. The overarching goal of the work presented here is developing novel statistical technique to advance our ability of characterizing rare fluctuations in both rainfall and scalar turbulence, and in turn, use these predictive tools to obtain a deeper understanding of the physical processes involved.

1.2 Overview

Toward this broad objective, my graduate work includes three main contributions. First, a novel characterization of the probability distribution of daily rainfall extremes, aimed at optimizing the information contained in samples of limited length [ZBM16, ZCM20]. Second, the development of a statistical technique to spatially downscale, validate, and correct satellite-retrieved rainfall statistics, with a particular focus over poorly gauged areas [ZM19, ZM20]. Finally, I investigated the nature of turbulent fluxes in the atmospheric boundary layer (ABL), with a focus on the exchange of both sensible heat [ZBK18] and methane [ZPGK20] with the surface. Here I have studied the role of intermittency - either internal intermittency, as in the case of temperature, or intermittency that is externally imposed by the boundary conditions, as in the case of methane - in modulating the exchange of mass and energy with the surface.

Therefore, this dissertation is structured along three main lines of inquiry - respectively focusing on (a) the statistical modelling of daily rainfall extremes, (b) the estimation of extreme values from remotely-sensed rainfall fields, and (c) the investigation of scalar fluxes in boundary layer turbulence. Each of these components will be discussed in two of the following Chapters:

- Chapter 2: A Non-Asymptotic Approach to Model the Frequency of Daily Rainfall Extremes (Theme a)
- Chapter 3: Bayesian Hierarchical Modelling of Extreme Values of Environmental Time Series (Theme a)
- Chapter 4: Downscaling of Extremes Rainfall Statistics from Satellite Observations (Theme b)
- Chapter 5: Extreme Value Analysis of Remotely-Sensed Rainfall in Ungauged Areas: Spatial Downscaling and Error Modelling (Theme b)
- Chapter 6: Extremes, Intermittency, and Time Directionality of Atmospheric Turbulence at the Crossover from Production to Inertial Scales (Theme c)
- Chapter 7: Intermittent Surface Renewals and Methane Hotspots over Natural Wetlands (Theme c)

The dissertation closes with Chapter 8, where the main results are discussed and their intellectual merit is discussed with respect to future research directions.

1.2.1 Improving the Characterization of Daily Rainfall Extremes

Existing statistical models of extreme rainfall make use of only a limited fraction of the available observations, such as annual maximum values or exceedances over high thresholds. Discarding most of the available information is inefficient, and can hinder the inclusion of information about the physical process that generates the observed phenomena. With the objective of overcoming this limitation, I have worked on a non-asymptotic approach to rainfall extremes which accomplishes two main tasks: i) models the entire rainfall process as opposed to just the right tail of the distribution, leading to a more efficient use of the data when only small samples are available, and ii) meaningfully anchors the statistical inference to parameters which can be more reliable to estimate. My work can explicitly model the inter-annual variability in the distribution of rainfall magnitudes. I have shown that accounting for this variability tends to yield probability distributions with *'heavier tails'* than one would otherwise obtain, thus underlining the fundamental role of low-frequency climate variability in controlling the observed frequency of intense precipitation events [ZBM16]. Building upon this result, I have then developed a Bayesian formulation of this model for precipitation extremes [ZCM20]. This improvement, while retaining the fundamental structure of the model introduced in [1], improves (I) the quantification of uncertainty, and (II) the inclusion of prior climatic information in the inference process, as made possible by the Bayesian framework adopted. This approach has been extensively tested and applied using a network of rainfall records over the Conterminous United

States.

1.2.2 Downscaling and Validating Remote Sensing Rainfall Datasets over Poorly Gauged Locations

Remote sensing datasets have become an essential source of information to study the fluctuations of the water cycle at the global scale. However, in the case of intermittent and highly-variable processes such as rainfall, validation of these datasets requires extensive measurement campaigns at the ground. In the case of rainfall, this is a daunting task, given that large areas worldwide are poorly gauged, including, in particular, arid regions, complex terrain areas, and developing countries. The goal is to harness the potential of the Tropical Rainfall Measurement Mission (TRMM), and now the Global Precipitation Measurement (GPM) mission to study rainfall extremes over poorly instrumented areas worldwide. I have thus developed a framework for spatially downscaling rainfall statistics obtained from satellite multi-sensor precipitation estimates, and for comparing them with point gauge observations over extended poorly-gauged areas [ZM19]. The method relies on the theory of stochastic processes to describe the space-time variability of the precipitation process, and allows for the comparison of statistics obtained from gridded precipitation datasets with their counterparts obtained from a single measuring gauge at the ground, thus relaxing the need for extensive field campaigns with multiple gauges. I then combined this approach with a non-parametric model of the error aimed at inferring the performance of remotely sensed rainfall datasets over ungauged locations, thus allowing for predictions uniquely based on the local topography and climatic variables [ZM20]. Taken together, these results provide an innovative framework for estimating the frequency of extreme events at the global scale from remotely sensed observations, and for extending their validation over ungauged locations.

1.2.3 Investigating Turbulent Fluxes in the Atmospheric Boundary Layer

A significant component of my graduate work investigates the transport of scalar quantities such as temperature and gases in the turbulent atmospheric boundary layer. My first objective was inferring the signature of atmospheric stability (i.e., of the interplay between buoyancy and shear forces in generating turbulence) on the statistical properties of scalar quantities transported in the boundary layer. Interestingly, from a campaign of sonic-anemometer measurements performed at the Duke forest, I found that atmospheric stability has a clear signature both on the frequency of extreme scalar fluctuations and on the symmetry of their time-evolution [ZBK18]. This finding parallels flight-crash dynamics recently observed in numerical simulation of turbulent flows. I then turned my attention to studying the fluxes of methane (CH4) originating from boreal peatlands, using data from a long-term measurement campaign conducted in Finland [ZPGK20]. Natural fluxes of methane - a strong greenhouse gas - are particularly difficult to characterize due to their sporadic behavior, as they not only occur by diffusion through the water table, but also through plant transpiration and by the localized release of methane bubbles through the wetland surface. This peculiarity sets them apart from fluxes of other quantities such as carbon dioxide and water vapor, and makes it difficult to correctly quantify the magnitude of these fluxes and their seasonal variability with the standard eddy covariance technique. The dynamics of the ground source cast this again as an extreme value problem, where relatively few short events account for a considerable fraction of the overall fluxes. I have analyzed this problem by applying a partition scheme in the wavelet domain aimed at detecting the signature of ebullition in CH4 concentration time series. This information was then used to calibrate an intermittent surface renewal scheme, which provides novel information on how CH4 source intermittency impacts the gas transfer velocity.

Chapter 2

A Non-Asymptotic Approach to Model the Frequency of Daily Rainfall Extremes

Adapted from: Zorzetto, E., G. Botter, and M. Marani. "*On the emergence of rainfall extremes from ordinary events*." Geophysical Research Letters 43.15 (2016): 8076-8082.

2.1 Introduction

Extreme Value Theory (EVT) [FT28, Gne43, Gum58] is a fundamental tool in the study of many geophysical processes, such as the local and global hydrologic cycle [KPN02], wind velocities [CH04], earthquake magnitudes [PSSR14], ecological processes [KBP05], storm-surge marine levels [CT90], pollutant dynamics in the environment [ET09], and many others. In the classical EVT, extremes are defined as "block maxima", i.e. as the events with maximum magnitude x occurred over a period of fixed lengeth (often one year). The n events occurring in each block are assumed to be independent and their magnitude is assumed to follow the same parent cumulative distribution $F(x)$. Hence, block-maxima have cumulative distribution $H_n(x) = F(x)^n$. This expression is not directly applicable as n is the value assumed by a random variable N. To obtain a closed-form expression for $H_n(x)$, the classical EVT makes one of two possible assumptions. A first approach is to assume the number of events per block to be "large enough" (i.e. $n \to \infty$), such that the succession $H_n(x)$, upon proper renormalization, tends to an asymptotic distribution, $H(x)$, which takes the form of the Generalized Extreme Value (GEV) distribution [VM36]. It has been noted that in many applications the number of events from which the maximum value is selected is not nearly sufficient for this asymptotic hypothesis to be valid [CH04, Kou04]. A second approach, termed Peak Over Threshold (POT) method [BDH74, P$^+$75] i) fixes a high intensity threshold, q, ii) assumes a Poisson occurrence of events above the threshold, and iii) models the excess values over q (assumed to be independent of the occurrence process) using a Generalized Pareto Distribution (GPD) [DS90]. Also in this second approach, sometimes referred to as Partial Duration Series [Ste93] the resulting EV distribution is GEV. Both these classical EVT approaches lead to formulations which neglect a significant proportion of the observations, as they fit the block-maxima distribution, $H(x)$, using only the block-maxima themselves, or a relatively small number of exceedances over a high threshold. Effectively, both these approaches discard the information contained in the bulk of the parent distribution $F(x)$, along with most of the observations. Here I refine and apply a statistical approach based on the assumption that the extreme events are block maxima among a finite and stochastically variable number of ordinary events. These are defined as the values obtained by the repeated sampling, in each block, from an underlying and possibly time-varying distribution (e.g., all rainfall occurrences in a given year in the case of daily rainfall). This simple consideration allows me to use the entire observational set to infer the distribution of extremes, by means of a Metastatistical Extreme Value framework (MEV), with obvious statistical adavantages. This approach is here applied to the relevant case of daily rainfall events using a worldwide data set of long rainfall records and a Monte Carlo approach to comparatively assess MEV and GEV high-quantile estimation uncertainties.

2.2 Theoretical Framework

I propose the use of a Metastatistical Extreme Value (MEV) approach that relaxes the limiting assumptions of the classical EVT by considering as random variables the parameters defining the number of events and the probability distribution of event magnitudes [MI15]. This leads to a compound distribution [Dub68] or superstatistics [BC03, PVF06, BBIR13] for the distribution of the block maxima. In the MEV approach the variability of these parameters accounts for i) the random process of event occurrence, which generates a finite and varying number of events in each block, and ii) the possibly changing probability distribution of event magnitudes across different blocks. The MEV approach accomplishes this by recognizing the number of events in each block, n, and the values of the parameters, $\vec{\theta}$, of the parent distribution $F(x;\vec{\theta})$ to be realizations of stochastic variables (N and $\vec{\Theta}$). The probability distribution of block-maxima can now be defined, by use of the total probability theorem, by considering all possible values N and $\vec{\Theta}$, thereby yielding a MEV cumulative distribution function:

$$\zeta(x) = \sum_{n=1}^{\infty} \int_{\Omega_{\vec{\Theta}}} F(x;\vec{\theta})^n \, g(n,\vec{\theta}) d\vec{\theta} \tag{2.1}$$

where $g(n,\vec{\theta})$ is the joint probability distribution of N and $\vec{\Theta}$ (discrete in N and continuous in $\vec{\Theta}$), and $\Omega_{\vec{\Theta}}$ is the population of all possible values of the parameters. The probability distribution of the extremes thus arises from the full distribution of the ordinary events (not just from a predetermined part of the tail), which is sampled - each year - a variable number of times n. For this reason, the MEV approach exploits all the available observations defining the probability distributions of ordinary events in each block, rather than censor the dataset to only include values from the tail of $F(x)$. It is interesting to note that if one assumes i) x to be the excess over a high threshold q ii) $F(x;\vec{\theta})$ to be a Generalized Pareto distribution (with fixed, deterministic parameters), and iii) n to be generated by a Poisson distribution, then the GEV distribution is recovered as a particular case of the MEV distribution by means of the POT approach. Rather than specifying the joint probability density function (pdf) $g(n, \vec{\theta})$, I obtain here an approximate expression for $\zeta(x)$ by substituting the expectations in eq. (2.1) with sample averages. I illustrate this derivation with application to the relevant case of daily rainfall observed at a point. Following Wilson and Toumi [WT05] and Marani and Ignaccolo [MI15], I adopt the Weibull, or stretched exponential distribution [LS98], $F(x; C, w) = 1 - e^{-\left(\frac{x}{C}\right)^{w}}$ to model the non-zero daily rainfall amounts (C and w being, respectively, the Weibull scale and shape parameters). One can thus define the MEV-Weibull cumulative distribution function as

$$\zeta(x) = \sum_{n=1}^{\infty} \int_{C} \int_{w} g(n, C, w) \cdot \left[1 - e^{-\left(\frac{x}{C}\right)^{w}}\right]^{n} dC dw \tag{2.2}$$

The Weibull distribution, $F(x; C_j, w_j)$, is assumed to describe the observations in each year on record ($j = 1, 2, ..., M$). A sample of yearly maxima distributions, $H_{n_j}(x) = F(x; C_j, w_j)^{n_j}$ (where n_j is the number of wet days in year j), can thus be defined, and a sample-average approximation can be computed $\zeta(x) \cong \zeta_m(x) = 1/M \sum_j F(x; C_j, w_j)^{n_j}$ (Fig. 2.1). The discrete expression of the MEV-Weibull distribution thus reads:

$$\zeta_m(x) = \frac{1}{M} \sum_{j=1}^{M} \left[1 - e^{-\left(\frac{x}{C_j}\right)^{w_j}}\right]^{n_j} \tag{2.3}$$

Convergence of (2.3) to (2.2) is ensured provided that (C_j, w_j, n_j) are sampled $\forall j$ from their joint distribution $g(N, C, w)$. I fit the Weibull distribution to observations in each single year by means of the Probability Weighted Moments method (PWM), which, compared to other methods (e.g. Maximum Likelihood, ML), attributes a greater weight to the tail of the distribution. Moreover, the PWM method performs well for small samples and is not very sensitive to the presence of outliers [GLM79].

ML is, on the contrary, known to be a biased estimator of the Weibull parameters, especially the shape parameter, for small samples [Sor06]. In the following I will often consider the event magnitude, $\hat{x}$, corresponding to a given value of the return period of interest, which I obtain by numerically solving $\zeta_m(\hat{x}) = (Tr-1)/Tr$ ($\zeta_m(x)$ being given by eq. (2.3)). I describe below extensive comparisons of MEV high-quantile estimates with those obtained from the traditional Generalized Extreme Value distribution. I estimate GEV parameters using the most efficient and most commonly used techniques: Maximum Likelihood (ML) [Col01], L-Moments [Hos90], the Peak Over Threshold approach [DS90], and Mixed Methods [MS02]. The POT approach was applied by selecting threshold values such that an average of 5 excesses/year are used to fit the parameters. Overall, I find that, for this 37-station dataset, the POT and L-moment methods yield the best estimates of GEV parameters (see Appendix A for more details), whereas ML estimators exhibit a larger error standard deviation, especially for smaller samples.

2.3 Data Sets

I gathered data from 37 rainfall records distributed globally and spanning different rainfall regimes. Many of the records were extracted from NOAA's Global Historical Climatology Network (GHCN) (ftp://ftp.ncdc.noaa.gov/pub/data/ghcn/daily/). Particularly long time series were gathered independently (e.g. the Padova time series, the longest daily rainfall record worldwide, consisting of 272 years of observations [MZ15]). See Table A.2 in Appendix A for a complete description of the data included in the analysis. The stations span different climatic conditions, thus allowing to test the ability of the MEV approach to capture observed extremes across a wide variety of precipitation regimes. I restricted my analysis to time series with length exceeding

100 years (mean length in the dataset is 135 years). Furthermore, only years with less than 10% of missing daily data were considered, which implies that about 2.4% of the years in the global data set were excluded from the analysis. I tested the ability of the Weibull distribution to describe observed daily rainfall for all the stations considered using the Kolmogorov-Smirnov and Cramer Von Mises statistical tests. The positive outcomes of the statistical tests performed for the different stations are described in Appendix A (Figure A.3).

2.4 Monte Carlo Analysis of Model Performance

The possible presence of non-stationarity or of periodicities in observed rainfall records adds an additional, and difficult to control, source of uncertainty in the comparative evaluation of extreme value analyses [SK15]. Hence, I used a Monte Carlo approach which by construction removes possible non-stationarities in the observations, while preserving the distribution of the rainfall accumulation values and number of events/year present in the observed dataset. To this end, for each station in the data set, I randomly reshuffle the observed numbers of wet days/year (n_j's), thereby preserving their original distribution, but destroying any serial correlation that may be present. Subsequently, I construct a m-year synthetic sample by randomly drawing (without resubstitution) n_j rainfall accumulation values ($j = 1, ..., m$) from the original record. The resulting synthetic time series lacks any serial correlation and preserves the original frequency distribution of rainfall depths. From each rainfall sequence generated through the above procedure (with length, m years, equal to the original observed time series), I extract the first s years to be used as a sample to fit the EV distributions. The training sample size s is varied from 10 to 80 years with a 2-year step, to explore the range of commonly available sample sizes (see Fig. 2.2A).

The remainder of the time series ($m - s$ years) is then used to independently test the MEV and GEV models performances. The selection of observed time series with length exceeding 100 years allows me to use empirical frequencies as references for the exceedance probabilities inferred through the EV models. The sample frequency of an yearly maximum value, x_i, is computed using the Weibull plotting position formula as $F_i = i/(m - s + 1)$, and is assumed to be the best estimate of the actual exceedance probability $F(x_i)$. i is the rank of x_i in the list, sorted in descending order, of the $s - m$ yearly maxima available in the validation sub-set. I finally compare $\hat{x}_i = F^{-1}(F_i)$ (where $F^{-1}(\cdot)$ denotes the inverse of one of the EV distributions to be tested) and x_i to determine the estimation error for the 20 largest events in each Monte-Carlo generated datasets. I repeat this bootstrap/reshuffling procedure $n_r = 100$ times for each observed time series, in order to obtain a large number of realizations over which to average the root mean square error. The accuracy of the empirical frequency estimates of the underlying probability improves with the length of the time series and with the number of Monte Carlo realizations considered, and decreases as the return period examined increases. For this reason I focus my attention on return times in the range 10-150 years. For every bootstrap realization, for every sample size and return time (s, Tr), theoretical quantiles, $\hat{x}$, were estimated from the EV distributions being compared. Using the observational quantiles x_{obs} relative to the same Tr_i, the non-dimensional estimation error can be computed as $\epsilon = (\hat{x} - x_{obs})/x_{obs}$. The values of ϵ_i obtained from each reshuffled series are then averaged over all the Monte Carlo realizations ($n_r = 100$) to obtain a global performance metric:

$$\rho(s, T_r) = \left[\frac{1}{n_r} \sum_{k=1}^{n_r} \left(\frac{\hat{x}_k(s, T_r) - x_{obs,k}(s, T_r)}{x_{obs,k}(s, T_r)} \right)^2 \right]^{\frac{1}{2}} \tag{2.4}$$

Figure 2.2 plots the ratio of ρ_{MEV} to ρ_{GEV} as a function of s and Tr, in which data from all the stations have been pooled together. In order to obtain meaningful statistics, individual values of the ratio of the RMSE's computed from eq. (2.4) are averaged over tiles in the plane (s, Tr) of size 20 yrs x 10 yrs.

2.5 Results

Figure 2.2B shows that MEV on average outperforms GEV when used to obtain estimates for return periods exceeding the length of the sample used to fit the distribution. For the largest return periods, often of greatest practical interest, the average MEV error is of the order of 50-60% the average GEV estimation error. This result has broad implications, as most of the time series globally available only span a few decades (Fig. 2.2A), while return periods of common interest are greater than $Tr = 100$ years. The analysis of the ratio of the estimation errors as a function of the dimensionless number Tr/s (Fig. 2.4B) clarifies this notion. While some scatter exists, the average of ρ_{MEV}/ρ_{GEV} over bins of Tr/s values clearly indicates that the MEV error tends to be smaller than the GEV error when Tr is greater than the sample size (i.e., $Tr/s > 1$), attaining a 50% improvement for Tr/s indicatively larger than 5. In absolute terms the average Root Mean Square Error for MEV and $Tr/s = 5$ is roughly 20% (Fig. 2.4A). Figures A.1 and A.4 in Appendix A show similar results for the comparison with the POT and GEV-ML approaches. Additionally, the comparison of the full distributions of the estimation errors for a common return period and sample size confirms that the MEV approach leads to a significantly narrower error distribution with a mode close to zero, as show in Figure 2.3 (see also Table A.1 and Figure A.5 in Appendix A).

2.6 Discussion

The MEV approach presents significant conceptual advantages with respect to traditional methods rooted in the EVT. It removes any asymptotic hypothesis and hence does not require that a sufficiently large number of events/year takes place. The hypothesis of a Poisson occurrence of events is also removed in the MEV approach, the POT approach being retrieved as a special case. The use of a distribution with varying parameters to describe the ordinary event intensities embeds the inter-annual variability of the rainfall generation process, and paves the way to the natural incorporation of trends or multi-annual climatic oscillations. The MEV approach recognizes that annual maxima do not necessarily come just from the tail of the underlying parent distribution, a known limitation of the classical EVT [VLL15].

Classical EVT shows that extremes can only exhibit three types of tail behaviours (upper bounded, exponential, power-law tailed), which becomes manifested in the value of the GEV shape parameter [FT28, VM36, Kou04], a conceptually important implication of the classical EVT requiring further discussion. This fat (power-law) vs. thin (exponential) tail asymptotic dichotomy is conceptually suggestive and practically relevant, such that one wonders if it is negated by the MEV-Weibull approach, which seems to invariably yield a thin-tailed behaviour dictated by the exponential nature of the Weibull distribution. However, the Weibull distribution has been noted to exhibit a sub-exponential tail when $w < 1$, with a behaviour which is intermediate between an exponential ($w = 1$) and a power-law [LS98, Sor06]. Furthermore, I note that the combination of exponential distributions with different decay parameters in a metastatistical framework can lead to power-law tails [Dub68, BC03, PVF06, GMFG$^+$10]. For example, in the present MEV formulation one can show that (see Appendix A for the details), when only the scale parameter of the Weibull distribution varies stochastically, the MEV distribution can assume a power-law form, i.e., a heavier tail than the underlying Weibull distributions in eq. (2.3). Hence, I conclude that the MEV-Weibull formulation, even though it is based on stretched-exponential building blocks, can reproduce thin- as well as fat-tailed extreme value distributions. The adoption of a single Weibull distribution to describe all daily events within each year implies that seasonality and different rainfall generation mechanisms are not explicitly resolved. Recent work on flood frequency analysis [MS02, VS10] indeed suggests that power-law tails may artificially emerge from a mixture of probability distributions associated to different rainfall-generating mechanisms. However, this interpretation is not fully in contrast with my approach, which explains thick-tailed extremes by the metastatistical mixing of distributions of the same type, but with stochastic parameters.

2.7 Conclusions

Analysis of extremes in several ultra-centennial daily rainfall records shows that the MEV approach on average outperforms traditional GEV methods when the return period of interest is longer than the length of the observational time series available. The GEV distribution does provide accurate descriptions of the specific samples used to fit it, as shown by the high goodness of fit obtained when the performance is evaluated on the same data used for its calibration (see Figure A.2 in Appendix A), but, compared to the MEV approach, it fails to properly generalize and capture the underlying statistical properties of the population. The MEV approach, on the contrary, uses information from the bulk of the distributions of ordinary values, and is able to more effectively capture the characters of the population of extremes, such that the estimated high quantiles are less sensitive to the specific sample used for fitting. In conclusion, I argue that the MEV approach should be preferred to the GEV distribution, especially when small samples are available and high-quantile extremes are to be estimated.

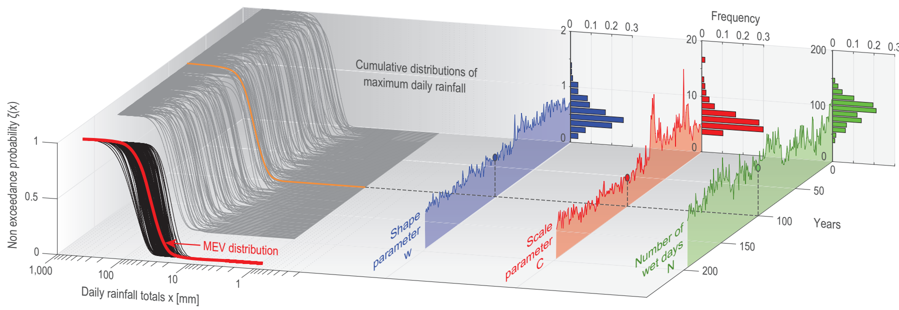


Figure 2.1: Conceptual representation of the MEV approach applied to daily rainfall EV analysis. Yearly values of the Weibull parameters (in blue and orange) and of the number of wet days (in green) define the cumulative distribution of maximum yearly rainfall as $H_n(x) = F(x; C, w)^n$ (grey distributions on the left, and their projections in black on the vertical xy plane in foreground). The MEV distribution (in red in the vertical xy plane in the foreground), accounting for the stochastic variability in C, w and n, is obtained by averaging over the empirical frequency distribution of the parameters.

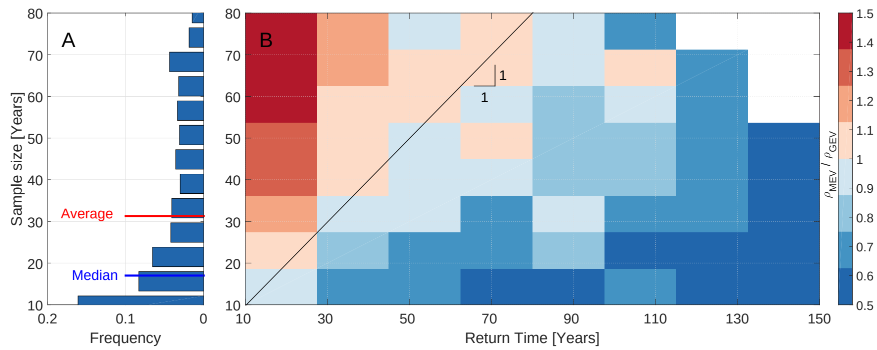


Figure 2.2: Comparative performance of MEV and GEV distributions. (A) Frequency distribution of sample sizes from the NOAA-NCDC global daily rainfall dataset. (B) Ratio ρ_{MEV}/ρ_{GEV}, of the Root Mean Square Errors of quantile estimates from the MEV-Weibull and GEV-LMOM approaches as a function of return period and size of the sample in our dataset. Individual ρ_{MEV}/ρ_{GEV} values from each site are pooled together and averaged over rectangular tiles on an uniform grid (with sides $\Delta Tr = 10$ years and $\Delta s = 20$ years). The $Tr/s = 1$ line is indicated as a reference. The MEV distribution outperforms the GEV distribution in the blue area. Areas in white contain no data.

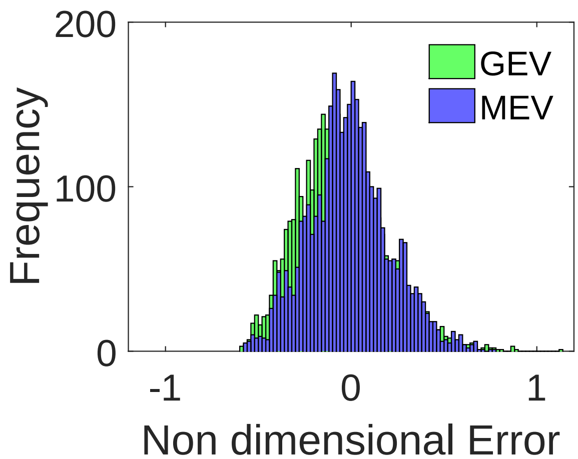


Figure 2.3: Distribution of the relative error $\epsilon = (\hat{x} - x_{obs})/x_{obs}$ for GEV-LMOM and MEV distributions. ϵ was computed over all the available stations and Monte Carlo realizations ($n_r = 100$). The return time is $T_r = 50$ years and the sample size $s = 30$ years (close to the mean length of the time series in the NOAA-NCDC global dataset). The mode of the MEV error is nearly zero, and the error distribution exhibits a smaller spread compared to the frequency distribution of the GEV error.

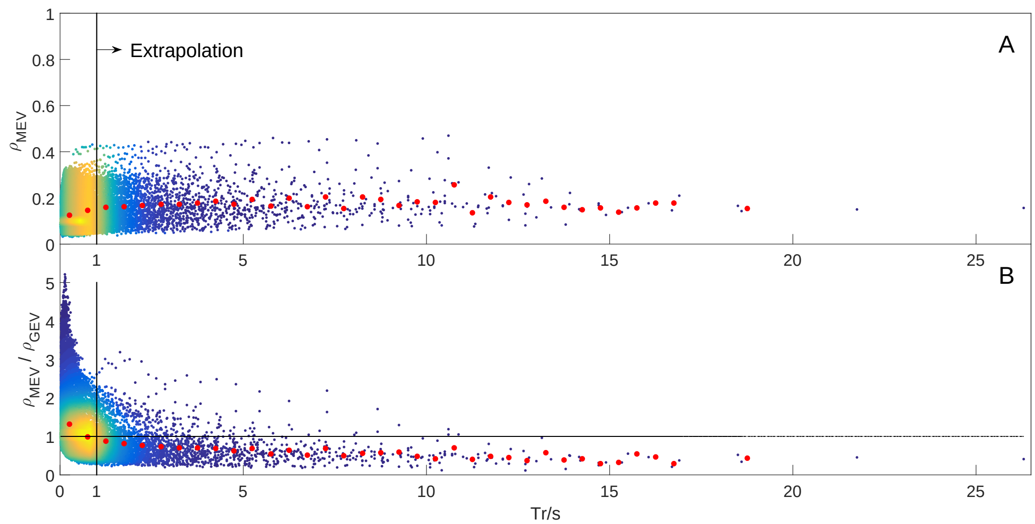


Figure 2.4: Performance of the MEV and GEV distributions as a function of the dimensionless parameter Tr/s. (A) Root Mean Square Error ρ_{MEV}, obtained from 100 Monte Carlo generations. Points denote values from single realizations, while red closed circles represent averages over bins of width 0.5 units. Colors denote the density (points/unit area of the plot, computed over circles of fixed radius) of the values falling in each area of the scatter plot (blue indicating the lowest density and yellow the highest one). (B) Ratio ρ_{MEV}/ρ_{GEV} of the Root Mean Square Errors obtained with the MEV and GEV approaches.

Chapter 3

Bayesian Hierarchical Modelling of Extreme Values of Environmental Time Series

Adapted from: Zorzetto, E., A. Canale, and M. Marani. "*Bayesian non-asymptotic extreme value models for environmental data* ." Manuscript submitted to Bayesian Analysis.

3.1 Introduction

The quantitative modelling of extreme events is of paramount importance in several disciplines, such as water science, geology, engineering, and finance, to name a few. In these contexts extremes are often defined as the maximum values observed in each year, or, more in general, as *block maxima* (BM). This approach avoids (by neglecting them) having to explicitly tackle issues related to seasonality, and introduces a unit of time to define the frequency of occurrence of extremes over time scales of applicative interest. This traditional approach has proven very fruitful and has generated a large theoretical body related to the max-stability property of the Generalized Extreme Value (GEV) distribution [FT28, Gne43, VM36, Col01]. An alternative modelling approach is based on defining extremes as exceedances over a high threshold, described through the theory developed by Balkema and De Haan [BDH74] and Pickands [P$^+$75]. Both approaches are asymptotic in nature.

In the Block Maxima approach, GEV is the non-degenerate distribution obtained for block maxima, after proper normalization, in the limit of an infinite number of independent and identically distributed (i.i.d) events in each block [FT28, Gne43], result later extended to the case of weak dependence structure [Lea83, LLR12]. Based on the value of its shape parameter, often denoted as $\xi \in \mathbb{R}$, the GEV family includes three possible limiting distributions for the block maxima: a double exponential (*Gumbel*, or *EV1*, for $\xi = 0$), a heavy-tailed (*Frećhet*, or *EV2*, for $\xi > 0$), and an upper bounded (*inverse Weibull* or *EV3*, for $\xi < 0$) distribution.

Conversely, in the Peaks Over Threshold (POT) framework, the Generalized Pareto Distribution (GPD) is derived as a model for excesses over threshold, in the limit of the threshold tending to the upper end point of the underlying random variables' support [Dav84, Smi84, DS90]. This approach was later also extended to the case of dependent sequences [Lea83, Smi92, BT98]. The GEV and GPD parametric models, respectively derived through the BM and POT frameworks, are deeply connected. In particular, by modelling the magnitude of threshold excesses with a GPD and their frequency of occurrence through a Poisson point process, again one obtains GEV as a model for the block maxima [DS90, Col01], with a parameter ξ equal to the corresponding GPD shape parameter. For a comprehensive introduction, see [Col01], [DHF07] and [EKM13].

Threshold models generally lead to a more efficient use of the data compared to the BM approach. However, the selection of the threshold is a relevant issue in this case, and a contrast exists between the desire of including as much data as possible in the EV model, while at the same time satisfying the asymptotic assumption, which would require the adoption of a high threshold. Therefore in general the optimal threshold selection requires a tradeoff between bias and variance of the resulting estimator [EKM13]. Several techniques have been developed for informing this decision [see

Dup99, Col01, EKM13, WT12].

The wide popularity enjoyed by approaches based on the GEV distribution led much of the extreme-value literature to focus on the block-maxima alone, or on few values above a high threshold, discarding and neglecting the 'ordinary values' from which these large events are extracted. In turn, this caused the widely accepted traditional Extreme Value Theory (EVT) 1) to be based on asymptotic results, to avoid the need of specifying details about the underlying distributions of the 'ordinary events', and 2) to focus only on few selected events, thereby 'wasting' most of the available information.

These issues have been receiving an increasing attention in recent times. Hydrological applications of EV models have shown that the number of yearly events is rarely sufficiently large for the asymptotic argument to hold [Kou04, MI15]. Moreover, for some parent distributions commonly used in a wide class of environmental applications, the actual extreme value distribution has been noticed to converge to its theoretical limiting form at a slow rate [CH04]. This is for example the case of the Weibull parent distribution, a parametric model widely adopted to describe several natural processes—such as wind speeds [HC14] and rainfall accumulations [WT05]—or in economics [LS98].

A more practical problem is related to the estimation of the GEV distribution shape parameter, ξ, which controls the nature of the tail of the distribution. When applied to precipitation data, maximum likelihood and L-moments estimates of ξ from block-maxima and POT techniques can be markedly biased depending on the size of available samples, and this can lead to an underestimation of the probability of large extremes in the case of small samples [Kou04, PK13, SK14]. This issue can be mitigated by use of sample statistics that are more efficient and robust than traditional ones [HW87], or, following a Bayesian approach, by penalizing the likelihood function with 'Geophysical Prior' distributions for ξ [MS00a, CPS03]. However, the limits, both conceptual and practical, of an approach that on the one hand heavily censors the data and, on the other, suffers by estimation bias and uncertainty, remain.

Another limitation of the traditional EVT which has been recently pointed out is related to the assumption of a single and invariable parent distribution [MI15]. In fact, many phenomena display changes in the event magnitude generation process that are imperfectly known and predictable due to the complexity of the system. In these circumstances the assumption of a time-independent form of the parent distribution can be questionable. Examples of this type of issues can be found in many Earth-system processes and variables, such as rainfall intensity [MI15, MNAM18], flood magnitudes [MMV20], wind speeds, and tropical storm intensities [HSM20].

Overall, though mitigated by advanced estimation approaches, the above limitations can have wide implications in the many applications requiring the accurate estimation of large quantiles, i.e. quantiles characterized by return times—average recurrence intervals—larger than the length of observed samples.

Recent contributions attempt to fill some of the gaps discussed above. Some of these contributions have focused on including the entire parent distribution of events in EV modelling, by using mixture of distributions [FHR02], by extending a GPD model to the entire range of observed values while retaining a Pareto tail [TAO06, PT13, NHRH16], by combining splines with an algebraic tail decay [HNZ19], or by use of a parametric family of distributions to model the entire range of ordinary values [MI15, JRM$^+$19].

The case of variable parent distribution has recently been tackled with the introduction of the Metastatistical Extreme Value Distribution (MEVD), a non-asymptotic extreme value approach in which a compound parametric distribution describes the entire range of ordinary values, with parameters varying across blocks [MI15, ZBM16, MNAM18, ZM20]. The main rationale behind the introduction of MEVD is describing the superposition of dynamics occurring over a wide range of time scales by use of compound distributions, i.e., by allowing the parameters of the distribution describing a *fast* dynamics to vary on a separate, much slower time scale.

Building upon the MEVD, here I introduce a Bayesian hierarchical model for extreme events which models the entire distribution of observed values, and explicitly incorporates the variability of their parent distribution across blocks. Latent variable models arise naturally in the Bayesian framework [GCS+13] and in the context of extremes have been widely used to develop spatial models [DPR+12, BHRM18] and to describe the temporal dependence of excesses over thresholds [BG14, BG16]. Here I harness the flexibility of Bayesian hierarchical modelling to account for the low-frequency variability in the underlying physical processes generating the data observed in different blocks, and to connect this variability with the tail properties of their extreme value statistics.

The use of Bayesian methods to model extremes of environmental data is quite general and successful [CT96, CPS03, FG18] and is particularly useful in the common case in which one has to rely on relatively short observational time series but has relevant and reliable expert prior information of the physical processes involved—as discussed in Section 3.2.3.

The Chapter is organized as follows: In Section 3.2 I introduce the general structure of the hierarchical model and subsequently specialize it to the analysis of rainfall data with a focus on informative prior specifications. In Section 3.3.2 the proposed formulation is empirically tested and compared to Bayesian implementations of standard extreme value models via a comprehensive simulation study. In Section 3.4 an application to a large data set related to daily rainfall measured over the United States is described. The Chapter ends with a a final discussion.

3.2 A Hierarchical Bayesian Extreme Value Model

3.2.1 Notation and general formulation

The proposed Bayesian Hierarchical Model for Extreme Values (HMEV) is formulated by denoting as n_j the number of events observed over the j-th *block* of time ($j = 1, \dots, J$, with J the number of blocks in the observed sample) and x_{ij} the magnitude of the i-th event within the j-th block ($i = 1, \dots, n_j$). The magnitudes of the n_j events occurring within a block are assumed to be realizations of independent and identically distributed (i.i.d.) random variables X_{ij}, with common parametric cdf $F(\cdot; \theta_j)$. $\theta_j \in \Theta$ is the possibly multivariate unknown parameter vector and $f(\cdot; \theta_j)$ the related probability density function. Under this framework, the block maxima $Y_j = \max_i \{X_{ij}\}$ have cdf

$$\zeta_j(y) = \Pr(Y_j \leq y) = F(y; \theta_j)^{n_j}. \tag{3.1}$$

In the following I define a generative hierarchical model for the data at hand. A graphical representation of its structure is illustrated in Figure 3.1. I let n_j be a realization of a random variable with probability mass function (pmf) $p(n; \lambda)$, where λ is an unknown vector of parameters. I further assume that latent θ_j's exist that are i.i.d. realizations of a random variable with probability density function $g(\cdot; \eta)$, where η is an unknown vector of parameters. With the convention that the symbol $\sim$ means "is a realization of a random variable having pdf/pmf," I can write the following hierarchical model,

$$n_j \mid \lambda \sim p(n_j; \lambda), \qquad \theta_j \mid \eta \sim g(\theta_j; \eta), \qquad x_{ij} \mid n_j, \theta_j \sim f(x_{ij}; \theta_j) \quad \text{for } i = 1, \dots, n_j. \tag{3.2}$$

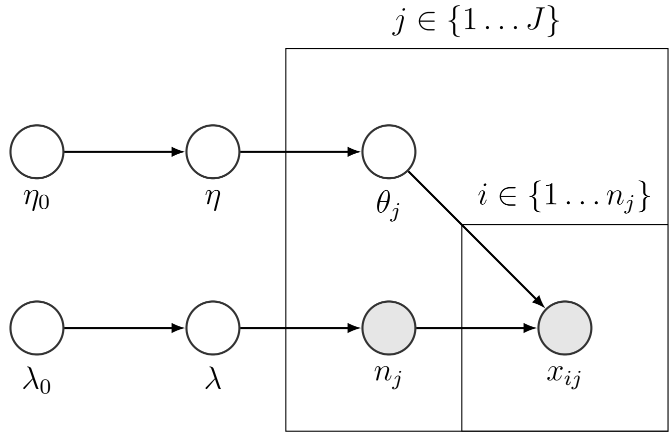


Figure 3.1: Hierarchical structure of the model described in equations (3.2)–(3.3). Grey dots represent observed variables.

Following a Bayesian approach, the hierarchical representation of the model is completed by eliciting suitable distributions, representing one's prior beliefs, for the unknown parameters λ and η,

$$\lambda \mid \lambda_0 \sim \pi_\lambda(\lambda; \lambda_0), \qquad \eta \mid \eta_0 \sim \pi_\eta(\eta; \eta_0). \tag{3.3}$$

In equation (3.3) λ_0 and η_0 represent suitable prior hyperparameters. Comments and suggestions about their elicitation are reported in Section 3.2.3. Denoting as $\boldsymbol{x}$ the collection of all x_{ij}'s and as $\mathbf{n}$ the collection of all the n_j's, I indicate with $\Pi(\eta, \lambda \mid \boldsymbol{x}, \mathbf{n}, \eta_0, \lambda_0)$ the posterior distributions of $(\eta, \lambda) \in \Omega$.

The main goal of extreme value analysis can be summarized in estimating the cdf in (3.1) or one of its functionals. This can be be done marginalizing out (3.1) with respect to the distributions of θ_j and n_j [MI15], obtaining the following expression (3.4), where h is function of the model's parameters λ and η:

$$h(y; \lambda, \eta) = \sum_{n=0}^{N_t} \int_\Theta F(y; \theta)^n g(\theta; \eta) p(n; \lambda) d\theta. \tag{3.4}$$

where N_t is the maximum number of events in a block (e.g. $N_t = 366$ days in the case of yearly blocks and daily observations of an environmental variable such as rainfall). A Bayesian estimator of (3.4) can then be obtained by integration over the posterior distribution of the model parameters λ and η:

$$\hat{\zeta}(y) = E[h(y; \lambda, \eta)|x_{ij}, n_j] = \int_{\Omega} h(y; \lambda, \eta)\Pi(\eta, \lambda \mid \boldsymbol{x}, \mathbf{n}, \eta_0, \lambda_0) d\lambda d\eta. \tag{3.5}$$

Other functionals of interest such as the variance, or the probability intervals corresponding to given quantiles, can be calculated accordingly. As customary in extreme value analysis, for an event of given intensity y I am interested in estimating the corresponding *return time* T_r, or its *average recurrence interval*, which is defined in terms of the cumulative distribution function as $\hat{T}_r(y) = \{1 - \hat{\zeta}(y)\}^{-1}$. Conversely, the return level $\hat{y}$ associated with a given non exceedance probability p_0, or return time $T_{r0} = 1/(1 - p_0)$, is obtained as $\hat{y} = \zeta^{-1}(1 - 1/T_{r0})$, where $\zeta^{-1}(\cdot)$ denotes the quantile function obtained by inverting the non exceedance probability function defined by eq. (3.5).

3.2.2 A specific formulation of HMEV for modelling daily rainfall

I now discuss how the model structure presented above can be applied to modelling extreme values of environmental time series. Here I provide a specification of the HMEV for modelling the frequency of annual maxima daily rainfall accumulations, based on the general hierarchical structure outlined in Section 3.2.1. To this end, I need to specify parametric models for event magnitudes and occurrence, and elicit suitable prior distributions for their unknowns parameters. In this process, I seek to harness information on the physical processes generating the data and include it in the Bayesian pipeline. Several parametric families have been employed to model rainfall accumulations, including the exponential [RICI87], gamma [SN14], Weibull [WT05], lognormal and Pareto [PKM13], or mixtures of Gaussian distributions [LL13]. Generally, the choice of the model for a particular application is merely based on some goodness of fit assessment, without seeking a physical justification for the choice of the distribution. However, physical arguments have been provided suggesting the body of the daily rainfall distribution should follow a gamma distribution [SN14, NSSB17, MVN19], and suggesting its right tail should decay as a stretched exponential (i.e., Weibull) distribution [WT05]. Since the focus of this work is on extreme values, I briefly review the latter argument and show how physical insight can be incorporated into the present Bayesian specification. Wilson and Toumi[WT05] noted that precipitation accumulations can be characterized as the product of three independent random variables, namely the average vertical air mass flux through a moist level, the air specific humidity, and the precipitation efficiency, i.e., the fraction of the vertical water vapor flux which is precipitated out as rainfall during each event. As these are all average quantities, it is assumed that, by the central limit theorem, their respective distributions can be approximated by Gaussians. By the theory of extreme deviations [Sor06, FS97], it can then be shown that, in the upper part of the distribution (i.e., for large enough rainfall accumulations), the product of a finite number K of standard normal random variables is approximately a stretched exponential or Weibull distribution with a shape parameter equal to $2/K$, where $K = 3$ is the number of variables in the present case. Therefore, not only this argument supports the choice of the stretched exponential distribution to model heavy rainfall accumulations, but additionally provides an indication on the value of its shape parameter. This argument provides valuable prior information to be exploited in our Bayesian hierarchical model.

Consistently with this argument, I model the magnitudes of daily rainfall accumulations x_{ij} in year j with a 2-parameter Weibull distribution with parameter vector $\theta_j = (\gamma_j, \delta_j)$ and pdf

$$f_w(x; \gamma_j, \delta_j) = \frac{\gamma_j}{\delta_j} \left(\frac{x}{\delta_j}\right)^{(\gamma_j - 1)} \exp\left\{-\left(\frac{x}{\delta_j}\right)^{\gamma_j}\right\} \tag{3.6}$$

where $\delta_j > 0$ and $\gamma_j > 0$ denote the scale and shape parameters respectively.

To allow for the inter-block variability discussed in Section 3.2.1, I assume that the latent variables $\delta_j \sim g_\delta(\delta_j; \mu_\delta, \sigma_\delta)$ and $\gamma_j \sim g_\gamma(\gamma_j; \mu_\gamma, \sigma_\gamma)$ are independent and have Gumbel pdfs, a flexible yet parsimonious 2-paramater model allowing for possible asymmetry.

$$g_\delta(\delta_j; \mu_\delta, \sigma_\delta) = \frac{1}{\sigma_\delta} \exp\left\{-\frac{\delta_j - \mu_\delta}{\sigma_\delta} - \exp\left(-\frac{\delta_j - \mu_\delta}{\sigma_\delta}\right)\right\}, \tag{3.7}$$

$$g_\gamma(\gamma_j; \mu_\gamma, \sigma_\gamma) = \frac{1}{\sigma_\gamma} \exp\left\{-\frac{\gamma_j - \mu_\gamma}{\sigma_\gamma} - \exp\left(-\frac{\gamma_j - \mu_\gamma}{\sigma_\gamma}\right)\right\} \tag{3.8}$$

Next, I need to specify $p(\cdot; \lambda)$ in equation (3.2). It is well known that the rainfall process often tends to be overdispersed at the interannual time scale [ET10]. This consideration would suggest a choice of $p(\cdot; \lambda)$ allowing a variance-to-mean ratio greater than one, to flexibly represent the possible presence of clustering. However, I show in the following that the distribution of n_j chiefly affects the probability distribution of extreme events, (3.4), through its mean value only. To show this, let us rewrite (3.4) in terms of the survival probability function $S(y; \theta) = 1 - F(y; \theta)$,

$$h(y; \lambda, \eta) = \sum_{n=0}^{N_t} \int_\Theta [1 - S(y; \theta)]^n g(\theta; \eta) p(n; \lambda) d\theta, \tag{3.9}$$

by expanding $[1 - S(y; \theta)]^n$ in a Taylor series around zero $[1 - S(y; \theta)]^n = 1 - nS(y; \theta) +$

$\mathcal{O}\left(S(y;\theta)\right)$, and by retaining only the linear term in the expansion—as justified for large values of n and extreme quantiles (i.e., for $S(y \mid \theta) \to 0$)—one finds:

$$\begin{aligned} h(y;\lambda,\eta) &\simeq \sum_{n=0}^{N_t} p(n;\lambda) \int_\Theta g(\theta;\eta)d\theta - \sum_{n=0}^{N_t} np(n;\lambda) \int_\Theta S(y;\theta)g(\theta;\eta)d\theta \\ &= 1 - E_\lambda[n] \int_\Theta S(y;\theta)g(\theta;\eta)d\theta. \end{aligned} \tag{3.10}$$

This expression depends on the distribution of n_j only through its expected value conditional to the sample of observed n_j. I therefore argue for the adoption of a minimalistic model, the binomial distribution, with a success probability $\lambda \in (0,1)$ and number of trials N_t equal to the block size (e.g., $N_t = 366$ in our application to annual maximum daily rainfall). This rationale is also supported by practical applications of Poisson processes of extremes [S$^+$89] and of MEVD, showing that the specific distribution adopted for the n_j's does not significantly affect the estimation of large extremes as long as the average is correctly reproduced [MZAM19, HSM20]

3.2.3 Prior elicitation

One of the main advantages of introducing a hierarchical model describing the entire distribution of daily rainfall accumulations is the possibility of eliciting priors directly on the underlying distribution of the observed "ordinary" events x_{ij} and on the distribution of n_j, rather than on the distribution of block maxima. By doing so, in particular, I avoid the difficulty of prescribing a prior directly on the shape parameter ξ of the annual maxima distribution, which is the main challenge in the inference on EV models, and to which it is difficult to attribute physical meaning. Studies at the global [PK13] and continental scale [PAFG18] showed that the shape parameters of extreme value models can vary significantly in space and is particularly difficult to estimate reliably [e.g. Col01], especially for small samples [SK14]. However, here I argue that using the entire distribution of daily rainfall provides inferential advantages, and allows for the inclusion of additional physical insight on the process at hand.

For what concerns the specific parametric family for $\pi_\eta(\cdot;\eta_0)$, with $\eta = \{\mu_\delta, \sigma_\delta, \mu_\gamma, \sigma_\gamma\}$, I opt for independent inverse gamma distributions, but other choices of 2-parameters distributions such as gamma lead to a similar model flexibility and to qualitatively similar results. What is crucial is the specification of the values of the parameters of the above distributions according to the physical understanding of the precipitation process. Prior belief on the typical intensity of the events, μ_δ, is not difficult to obtain empirically for a given location as the climatological mean. Furthermore, the physical argument outlined in Section 3.2.2 enables one to assume a priori that the inverse gamma prior distribution for μ_γ is centered around 2/3. Note that if additional physical insight is available on the types of storms characterizing the site of interest, or from similar sites, this prior elicitation could be further refined, e.g., based on studies of the value of γ over large geographic areas [PAFG18].

For the latent Gumbel scale parameters σ_δ and σ_γ, quantifying the variability of the Weibull parameters between blocks, I also choose informative distributions with expectations equal to 25% and 5% of the respective location parameters (μ_δ and μ_γ). This choice reflects the notion that I expect significant variability in the scale parameter across years—here quantified as 25% of its mean value —but, conversely, I do not expect the shape parameter to vary as much, as its expected value should be more strongly constrained by the general physical nature of precipitation processes. Of course different precipitation types can occur in different proportions in different years, and, since I do not model these components explicitly, I should include their effect in possible variations of the scale parameters. Guided by these considerations, I choose a latent scale parameter for the variability of the Weibull shape parameter equal to 5% of its prior expected value.

Sometimes, information can be available on the relative frequency of different precipitation mechanisms, for example as obtained through satellite or radar measurements. In this case, the prior location value of σ_δ could be for example increased in settings characterized by higher inter-annual variability of the relative frequency of different precipitation types, as suggested in [MZAM19].

An independent weakly informative beta prior for the binomial rate parameter for n_j concludes the prior elicitation. I found that eliciting an informative prior for n_j is not as important as for the other parameters is the model, as (i) inference on the single-parameter distribution for n_j is more robust than inference of the distribution of x_{ij} even for very small sample sizes, and (ii) the HMEV estimates are primarily affected by the expected value of the n_j's distribution rather than by its higher-order moments. The specific values of the prior parameters used in the in the remainder of the article are summarized in Table B.1 in Appendix B.

3.2.4 Posterior computation and posterior predictive checks

Given the complex structure of the models described in previous sections, it is clear that an analytical expression for the posterior distribution of the parameters or for $\hat{\zeta}(y)$ in (3.5) is not available and numerical procedures are needed. Here I chose to approximate the posterior distribution with Markov Chain Monte Carlo (MCMC) and specifically using a *Hamiltonian Monte Carlo* approach exploiting the flexibility of the Stan software [CGH$^+$17].

The implementation of the hierarchical model and related prior described in Section 3.2.2 is trivial under Stan and is provided as a standalone R package. In all the following examples, I run $n_c = 4$ parallel chains, with $n_g = 2000$ iterations in each chain. I discard the first half of each chain to account for the burn-in effect. The final sample on which I perform inference is therefore based on $B = n_c n_g/2 = 4{,}000$ draws.

Using MCMC I can make inference on any functional of the posterior distribution, calculating, at each iteration of the sampler, the current value of the functional of interests. For example, if the cumulative probability of block maxima approximating (3.4) is the target, one should compute at the generic iteration

$$\zeta^{(b)}(y) = \frac{1}{M_g} \sum_{j=1}^{M_g} F(y; \theta_j{}^{(b)})^{n_j^{(b)}} \tag{3.11}$$

where $\theta_j^{(b)}$ and $n_j^{(b)}$ for $j = 1, \cdots, M_g$ are drawn from the related posterior predictive distributions for each block, and M_g is a number of future blocks—$M_g = 50$ in our application. Therefore, the Monte Carlo approximation of the posterior expectation (3.5) is

$$\hat{\zeta}_{MC}(y) = \frac{1}{B} \sum_{b=1}^{B} \zeta^{(b)}(y). \tag{3.12}$$

Note that (3.11) approximates the functional $h(z; \lambda, \eta)$ where λ and η are the parameters describing the inner level of the hierarchical model and the averaging operation in (3.11) is performed on the values of θ_j and n_j. Conversely, (3.12) is obtained by averaging over the B draws from the posterior distribution thus accounting for the posterior uncertainty of the λ and η parameters.

To assess whether the parametric assumptions of the proposed HMEV provide a good fit to the observed data, it is important to perform posterior predictive checks [GCS+13] comparing relevant quantities—such that y_i, n_j, or x_{ij}—with their corresponding posterior predictive densities. Although the posterior predictive distributions are not analytically available, it is straightforward to simulate new data from them by leveraging the MCMC samples of the parameters and the hierarchical representation of the model reported in Figure 3.1. I recommend to focus on the distribution of block maxima, and, given the interest in consistent estimates of the probability of large extremes, particularly on its right tail.

3.3 Simulation Study

3.3.1 Description

To assess the empirical performance of the proposed HMEV model, and to compare it with standard alternative methods, I performed an extensive simulation study. Different synthetic data sets have been generated under four scenarios characterized by specific event magnitude distributions: Generalized Pareto (GP), Gamma (GAM), Weibull (WEI) with constant parameters in each block, and a dynamic Weibull model in which the variable scale and shape parameters in each block follow Gumbel distributions (WEI_G). While the latter specification reflects the structure of the proposed hierarchical model, the other 3 scenarios represent model misspecifications and will be used to assess the rubustness of the proposed formulation to the specific distribution of event magnitudes. Common to all scenarios, the number of events in each block is drawn from a beta-binomial distribution with mean $\mu_n = 100$ events/block, variance equal to $\sigma_n^2 = 150$, and $N_t = 366$ block size. This choice represents the case of overdispersion commonly observed in rainfall and other environmental time series [ET10]. Each of the $R_s = 100$ replicated data set consists of two independent time series of lengths M_{train} and M_{test} blocks, which are respectively used for training and testing the different EV models. Here I fix $M_{test} = 500$ yearly blocks, and train the different models focusing on sample size values of $M_{train} = 20$ and 50 years, representative of many geophysical datasets. Table B.2, reported in Appendix B, describes the specific values of the parameters used to generate the synthetic data.

The competing methods used to benchmark HMEV are Bayesian implementations of the classical generalized extreme valued distribution (GEV) and peak over threshold (POT) Poisson point process models, whose details, including prior specifications, are reported in Appendix B. In order to perform a fair comparison, also these competing models are estimated under a Bayesian approach, using informative priors. In particular, for both models the prior distribution for the shape parameter is centered around the value 0.114, determined from investigations of rainfall records at the global scale [PKM13], and has a standard deviation of 0.125, yielding a distribution close to the Geophysical prior suggested by Martins and Stedinger [MS00a].

To evaluate the predictive accuracy of the different competing methods in estimating the true distribution of block maxima, I use different criteria measuring both the global goodness of fit and the uncertainty in estimating the probability of extreme events. The log pointwise predictive density (lppd) [GCS+13] computed both for the in-sample data and for the out-of-sample data is often used as a measure of global performance of the models. An alternative measure is the logarithm of the pseudo-marginal likelihood (lpml), a convenient index that directly accounts—at no additional computational cost—for a leave-one-out cross validation measure [GD94]. Notably, since the lpml approximates the expected log pointwise predictive density, the difference between the in-sample lppd and the lpml represents the number of effective parameters of a model [see e.g., VGG17] and thus will be used to quantify overfitting.

Since the focus of this work is the right tail of the distribution of the block maxima, here I introduce an additional index that measures predictive performance for quantiles above a given non exceedance probability. To this end, I introduce the Fractional Square Error (FSE)

$$FSE = \frac{1}{m_T} \sum_{j=1}^{M_x} \mathbb{1}_{(\tilde{T},\infty)} \left(T_j\right) \sqrt{\frac{1}{B} \sum_{b=1}^{B} \left(\frac{\zeta^{(b)^{-1}} \left(p_j\right) - y_j}{y_j} \right)^2}, \tag{3.13}$$

where $\zeta^{(b)^{-1}}(\cdot)$ refers to the quantile function of the specific model at the b-th MCMC iteration, $\mathbb{1}_A\left(x\right)$ is the indicator function that equals 1 if x belongs to A, and T_j is the empirical return time of y_j defined as $T_j = (1 - p_j)^{-1}$, with $p_j = \text{rank}(y_j)/(M_x + 1)$. M_x is the length in blocks of the sample of annual maxima used to compute the FSE. In the in-sample and out-of-sample validation performed here, $M_x = M_{train}$ and $M_x = M_{test}$ respectively. The value m_T represents the number of observations in the test set with empitical return time equal to or larger than $\tilde{T}$, i.e. $m_T = \sum_{j=1}^{M_x} \mathbb{1}_{(\tilde{T},\infty)}\left(T_j\right)$. Therefore, the FSE represents an average measure of a standardized distance between model-estimated quantiles and empirical quantiles for return times larger than $\tilde{T}$. In the following analysis I compute this measure for values of the return time larger than $\tilde{T} = 2$ years, thus focusing on the range of exceedance probability of interest in many practical applications.

To separately assess the precision and the variability of extreme value quantile estimates obtained from different models, I employ two additional measures, namely their average bias and the average width of the 90% posterior predictive credible intervals defined, respectively as

$$\begin{aligned} b_q &= \frac{1}{m_T} \sum_{j=1}^{M_x} \mathbb{1}_{(\tilde{T},\infty)} \left(T_j\right) \frac{1}{B} \sum_{b=1}^{B} \left(\frac{\zeta^{(b)^{-1}} \left(p_j\right) - y_j}{y_j} \right) \\ \Delta_{q90} &= \frac{1}{m_T} \sum_{j=1}^{M_x} \mathbb{1}_{(\tilde{T},\infty)} \left(T_j\right) \left(\hat{q}_{95}\left(p_j\right) - \hat{q}_5\left(p_j\right)\right), \end{aligned} \tag{3.14}$$

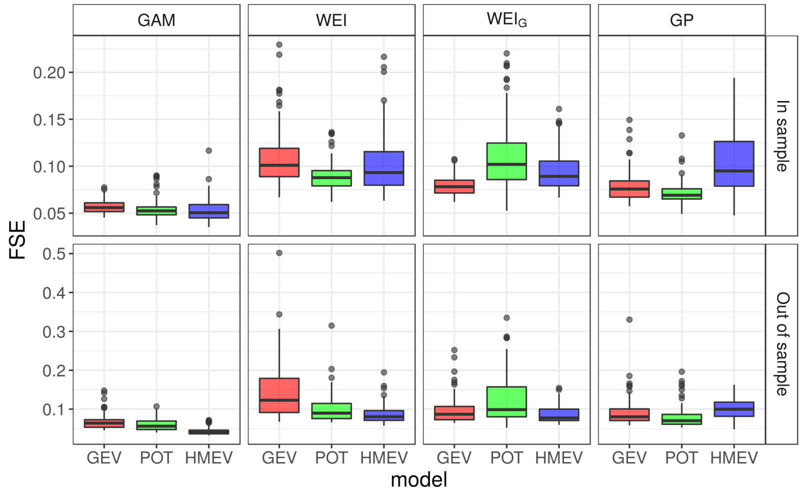


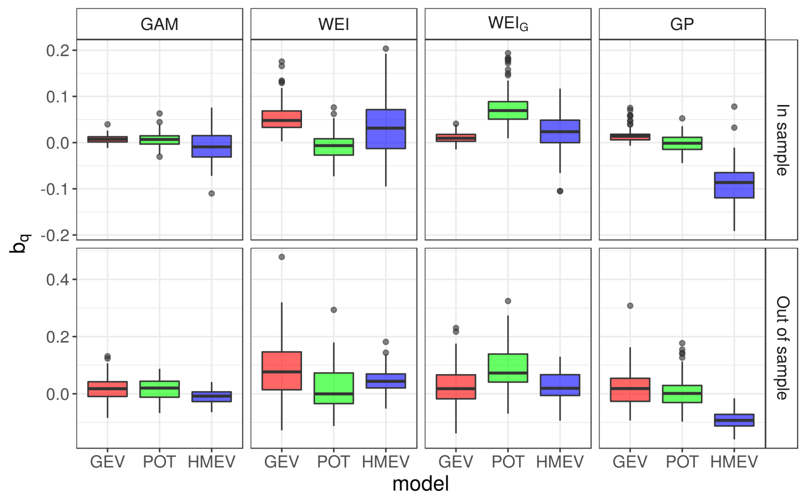
Figure 3.2: Fractional square error computed for the 4 different model specifications for in-sample data (upper panels) and for out-of-sample data (lower panels), computed for a sample size of 50 years.

where the quantities $\hat{q}_{95}(p_j)$ and $\hat{q}_5(p_j)$ are the upper and lower bounds of the posterior credibility interval for the quantile $\zeta^{(b)^{-1}}(p_j)$ estimated taking the empirical quantiles over the B MCMC draws.

3.3.2 Results

The results of the simulation study are illustrated in Figures 3.2–3.5. Specifically, Figure 3.2 shows the empirical distribution of the FSE over the $R_s = 100$ synthetic samples, training the model using 50 years of simulated data. The POT method appears to outperform the annual-maximum GEV in all cases examined, except in the case of WEI$_G$ specification, where arguably the inter-block variability of the x_{ij} distribution determines a variable rate of threshold exceedance, as well as a variable distribution of the excess magnitudes over the fixed threshold. While exhibiting a generally higher FSE for in-sample testing, HMEV cleary outperforms the competitors in the GAM, WEI, and WEI$_G$ scenarios in terms of out-of-sample performance.

(a)



(b)

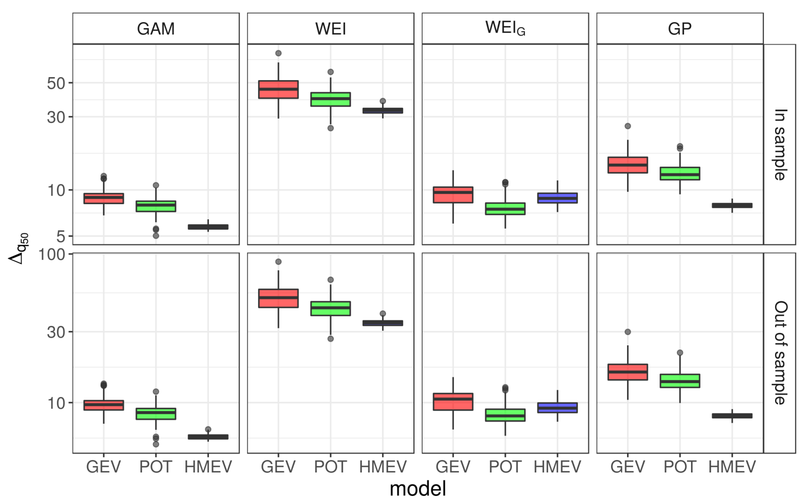


Figure 3.3: Mean bias (a) and mean credibility interval width (b) for the 4 different model specifications for in-sample data (upper panels) and for out-of-sample data (lower panels), computed for a sample size of 50 years.

In the GP scenario, POT remains the best model even in the case of out-of-sample testing.

To gain a deeper understanding of this general behavior, Figure 3.3a reports the results of the two measures introduced in (3.14). Generally, the best performance for the bias appears to be specification dependent, as is the case for the FSE, while for what concerns the width of the credibility interval, the HMEV is consistently the most efficient procedure, producing narrower credibility intervals. I note that the latent level temporal variability of the θ_j confer to HMEV a tail behavior which is intermediate between the lighter constant-parameter Weibull tail, and the Pareto model, as shown by the overestimation / underestimation of the posterior predictive quantiles in these two limiting cases.

The bias of the different models does not appear to vary significantly from in-sample to out-of-sample testing, suggesting than the difference observed in the FSE is primarily controlled by the variability of the different estimates.

To visualize this global behavior, Figure 3.4 shows a representative example of the performance of the methods. Specifically, it reports the quantile versus return time plots obtained for the different methods applied to a single dataset generated according to the WEI_G specification, with the yearly number of events $n_j \sim Bin(\lambda)$, with $\lambda = 0.3$. The results obtained for training datasets of 20 (panel a) and 50 (panel b) years show that HMEV yields quantile estimates characterized by narrower credibility interval and is characterized by a tail behavior which is lighter tailed compared to the other methods. Note that both the GEV and POT models, despite the informative prior used, appear to be more sensitive to the largest observations in the training samples and tend to overestimate the true function. This behavior is expected given the limited length of the training samples used here (20 to 50 years of data), which are however representative of sample sizes commonly available in many

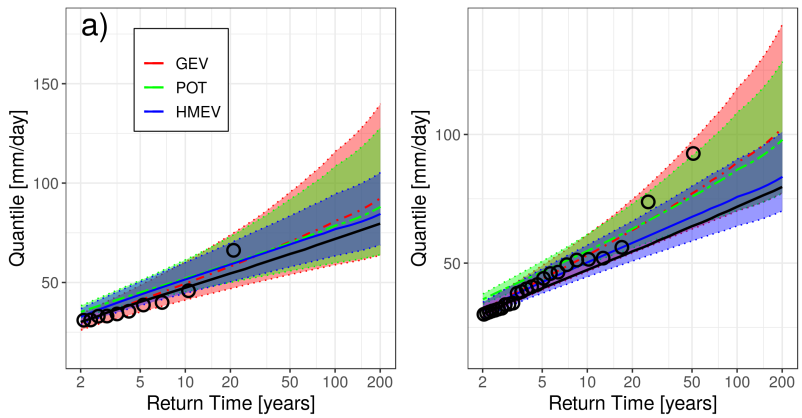


Figure 3.4: Quantiles predicted by the GEV (red), POT (green), and HMEV (blue) models based on training sets of 20 years (a) and 50 years (b) Lines show the expected value of the quantile for a given return time, while dashed lines represent the 5%-95% credibility intervals. Circles represent the observed return time of in-sample block maxima. The black lines report the quantiles computed from the true HMEV model.

applications in geophysics, engineering and environmental sciences. Representative plots for the remaining scenarios (model misspecification) are reported in Appendix B.

Note that the in-sample and out-of-sample tests illustrated in Figures 3.2 and 3.3 are characterized by different sample sizes. Therefore, the absolute difference between in-sample and out-of-sample metrics is not directly interpretable as a measure of overfitting. Therefore, to better quantify overfitting I study the effective number of parameters in each model, estimated as the difference between the in-sample lppd and the log posterior marginal likelihood, shown in Figure 3.5. HMEV displays a lower effective number of parameters in most of the specifications considered, suggesting that it is dramatically less prone to overfitting. While this behavior appears more markedly for two of the four data specifications (WEI and GAM), it is worth noting that this advantage increases when considering sample sizes smaller than $M_{train} = 50$ years examined here, as shown in Appendix B.

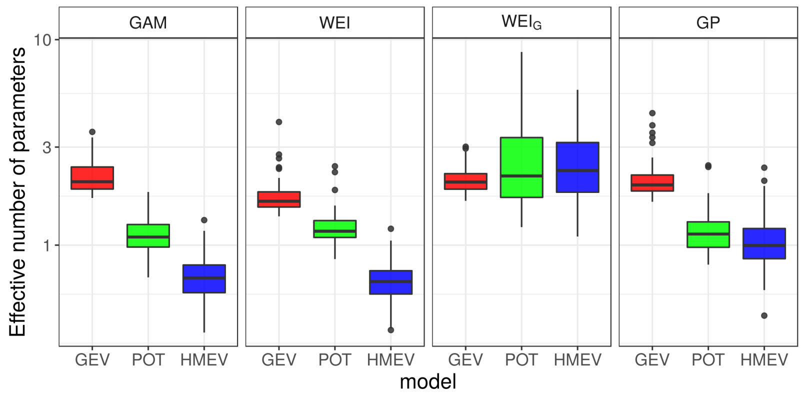


Figure 3.5: Effective number of parameters for the 4 different model specifications, evaluated for a sample size of $M_{train} = 50$ years of simulated data.

3.4 Application to the United States Historical Climatological Network Data

In this Section I analyze a rich collection of daily rainfall time series extracted from the United States Historical Climatological Network (USHCN) data. The data are freely available from the National Centers for Environmental Information (NCEI) of the National Oceanic and Atmospheric Administration (NOAA) [MDK$^+$12, MDV$^+$12]. The USHCN data set consists of 1218 long daily rainfall records covering the Conterminous Unites States (CONUS), with a significant fraction of the available records being longer than 100 years. This caracteristic makes this dataset particularly useful for the purpose of assessing the performance of this method by using only a portion of the data for the model fit, keeping the remainder as out-of-sample validation data. Moreover, since the CONUS spans a range of different climatic regimes, this datasets allows me to test the robustness of the model structure adopted here to different climates and precipitation types, which are expected to impact both the distribution of the x_{ij} and their arrival rate. The records characterized by non-blank quality flag were removed from the analysis, as well as the years characterized by more

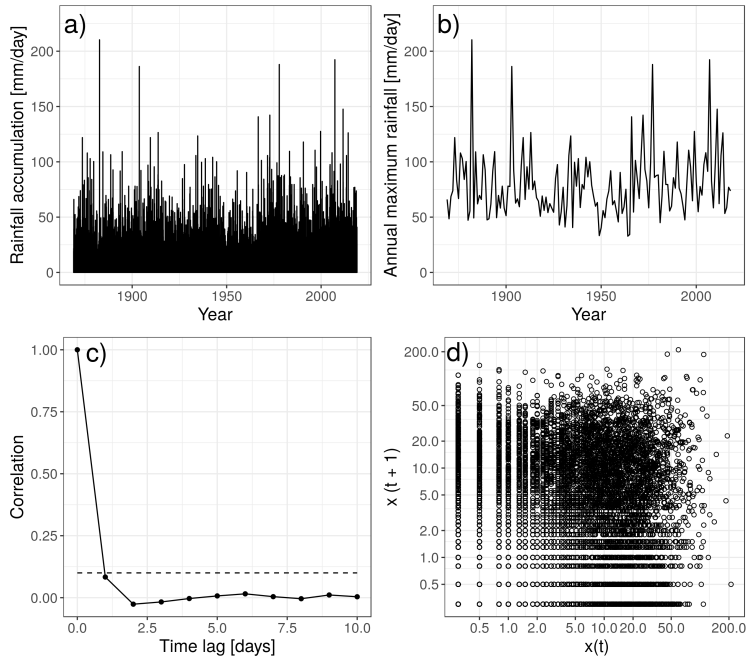


Figure 3.6: Rainfall time series measured at the New York Central Park (NYCP) station from 1969 to 2018 (Station ID USW00094728). (a), Time series of all daily rainfall accumulations, (b) annual maxima values only, (c) autocorrelation function of the daily rainfall accumulations,and (d) scatter plot of pairs of succeeding non-zero rainfall values.

than 30 daily missing observations. Therefore, for the subsequent analysis I select only stations with at least 100 years of record with enough non-missing, non-flagged observations, for a total of 479 stations.

3.4.1 New York Central Park station analysis

As a benchmark application I carefully discuss the results of the anlysis of the longest station in our data set, which was recorded in Central Park, New York City, from 1869 to 2018, for a total of 150 years of continuous observations (Station ID USW00094728). The entire series of daily event magnitudes as well as the 150 annual maxima values recorded at this station are reported in the top panels of Figure 3.6. Inspection of the autocorrelation function— panel c) of Figure 3.6—suggests that

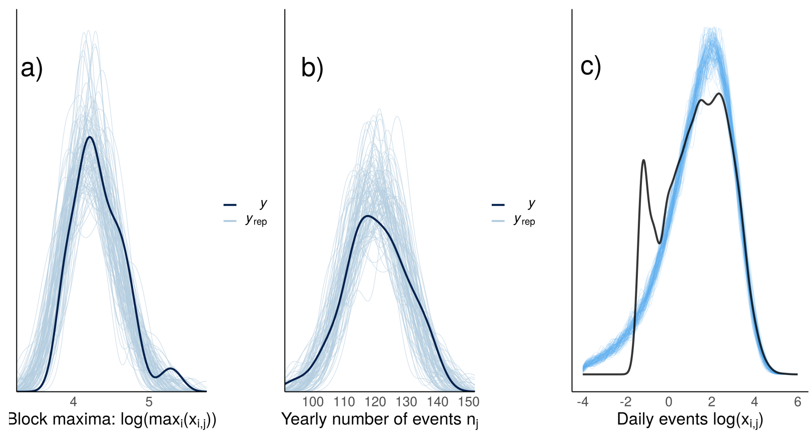


Figure 3.7: Posterior predictive distributions for the logarithm of the annual maximum daily rainfall accumulations (a), yearly number of events (b) and logarithm of non-zero daily rainfall events (c) computed by fitting HMEV to a 50-years sample extracted from the New York City time series. Black lines show the density of the observed values (obtained by kernel density estimation), while the light blue line show the kernel density estimates for 100 draws from the posterior predictive distributions.

the daily rainfall accumulations are not heavily correlated. However, when dealing with rainfall accumulations at shorter time scales, or in different climatic conditions, serial dependence may need to be accounted for when applying extreme value models based on the i.i.d. assumption. Hence, as commonly done in practice (see, e.g., [CNN07, MNAM18]), prior to model fitting I decluster the time series, by computing the autocorrelation of the daily magnitudes and determining the time lag τ_c in which the correlation decays below $c = 0.1$. Then, each rainfall record is declustered by only keeping the largest accumulation value observed within a neighborhood of length τ_c.

After this selection of pseudo-independent events, data are analyzed following the approach described in Section 3.2.2, with the same prior elicitation discussed in Section 3.2.3. Examination of the posterior predictive distributions for the annual maxima, number of events, and daily rainfall magnitudes, reported in Figure 3.7, shows that the pdfs of these variables are overall satisfactorily captured by HMEV. Note that a discrepancy appears for small values of daily rainfall magnitudes, where the censoring of actual values associated to the sensitivity of the instrument—0.3 mm here—introduces a threshold in the observed accumulations—clearly visible in panel d) of Figure 3.6. Despite this discrepancy for small magnitudes, the overall pdf of daily values, and in particular its right tail, appears to be satisfactorily captured by HMEV.

The considerable length of this particular time series allows me to explore the sensitivity of extreme value estimates to the specific sample used to train the model. In Figure 3.8 I compare extreme value quantiles obtained from the HMEV, GEV, and POT models trained on just the first 20 years or the first 50 years on record, respectively. Models estimates differ, with HMEV exhibiting—as previously observed from our simulation study—narrower credibility intervals with respect to POT and GEV models. HMEV predicts values slightly smaller than the Pareto model, but presents an overall good agreement with the empirical frequencies associated to the annual maxima extracted from the entire record (150 years of data). Interestingly, estimates from the GEV and POT models tend to fall between the frequencies computed from the training sets and those from the entire 150 year time series, showing their greater dependence on the specific training set used. This is shown even more clearly by the large differences, for GEV and POT estimates, between panel a) and panel b): when the length of the training set is increased such estimates significantly change, whereas HMEV estimates remain relatively insensitive to the increased available information.

3.4.2 Full USHCN data analysis

Building upon the insight gained from the simulation study I now turn my attention to quantifying the predictive ability of different models using real observations from a larger set of stations, and repeating the analysis for different sample sizes in order

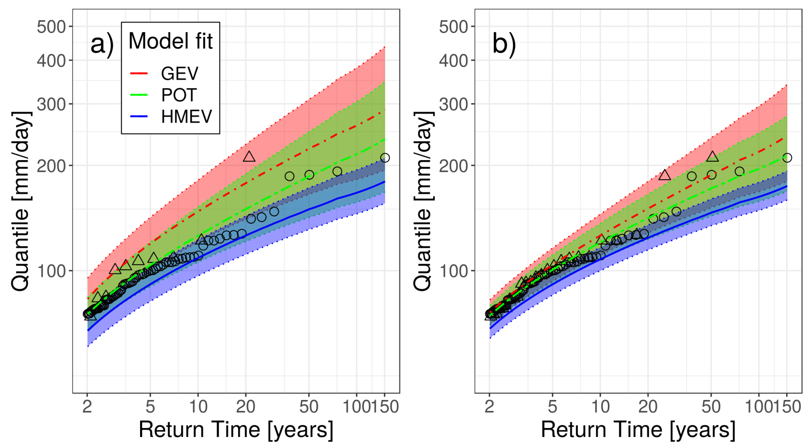


Figure 3.8: Extreme daily rainfall quantiles computed for the NYCP data set using for the fit only the first 20 (a) or 50 (b) years of the time series. Triangles represent the empirical cumulative frequency of data points in the training set, while black circles indicate the empirical frequencies computed from the entire 150-years time series. Predictions for the expected quantiles are indicated by red dashed line (Bayesian GEV), dashed green line (Bayesian POT), and continuous blue line (HMEV) with 5% - 95% credibility intervals reported as shaded areas for each model.

to test the sensitivity of the different models to sample size, a well-known issue in applications of extreme value models. To this end, for each station in our sample of 479 USHCN datasets, I extract M_{train} years to be used to train the extreme value model. Specifically, I repeat the analysis with M_{train} equal to 10, 20, 30, 40, 50 years. In each case, I then randomly extract 50 years of data from the remaining part of the time series to be used for independent validation. This procedure was repeated $R_g = 10$ times for each time series in the analysis reported here, thus producing a set of 4790 sample points. For each of these cases, I investigate the effect of the specific model used and of sample size, employing the set of different performance metrics introduced in Section 3.3.2.

Figure 3.9 shows the fraction of stations in which a specific competing model provides the best fit to the specific dataset. Examining this behavior for varying sample sizes, one can observe how HMEV becomes increasingly more competitive as

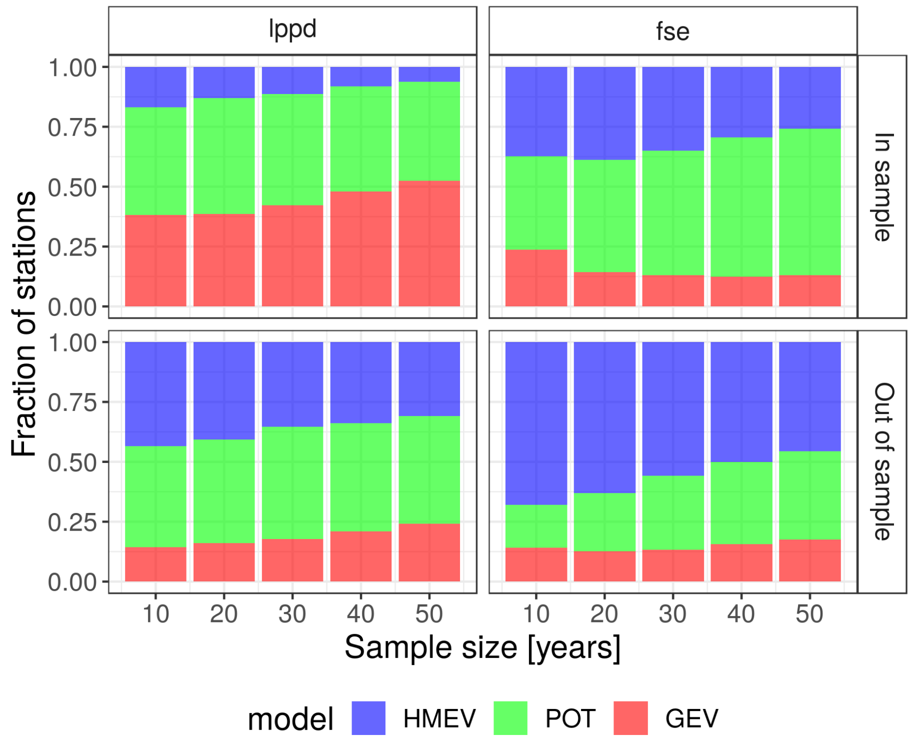


Figure 3.9: Fraction of cases in which each EV model exhibited the best estimation performance, evaluated with different metrics (FSE or lppd) using either in-sample or out-of-sample data.

the amount of available training data is decreased. As for simulation analyses, when one considers in-sample testing, POT most often is the best model. However, when out-of-sample performance is considered, the superior performance of the HMEV approach becomes clear. If one focuses on global measures of the probability distribution of estimation uncertainty for all yearly maxima, such as the lppd, the POT approach still seems superior for large training sample sizes. However, when the predictive uncertainty for extreme yearly maxima is examined (i.e. the FSE), arguably the main goal of extreme value analysis, the HMEV approach outperforms the other methods for all sample sizes. The difference between in-sample lpml and lppd, which provides a measure a model overfitting tendency, consistently depicts HMEV as the approach less prone to overfitting. See Figures B.1 and B.2 in Appendix B.

I also provide a spatially-explicit representation of model performances, by mapping, in Figures B.7 and B.8, reported in Appendix B, the best model for each station. The spatial distribution of the results overall appears to be consistent over the spatial domain of the study, even though some differences emerge between Western and Eastern USA. It is expected that, in specific locations, different precipitation regimes might produce distributions of daily rainfall which are not well captured by the stretched exponential model used in the present formulation of the HMEV. While producing location-specific models goes beyond the scope of the present study, the hierarchical structure of HMEV can be flexibly adapted by using different parametric families for the parent distribution, while benefiting from the high predictive performance outlined in our analysis.

As a representative application of the HMEV method to the computation of extreme value rainfall quantiles, I report with different colors in Figure 3.10 the magnitude of the 50-year daily rainfall event estimated for the set of stations analyzed here. For each station, following the analysis performed above, I randomly extract 50 years of record, repeating the procedure $R_g = 10$ times and averaging the results, so as to have a common same sample size for all records. As before, quantile estimates are obtained by numerically inverting the HMEV posterior predictive distribution, and computing the average quantile values over 4000 MCMC samples. This analysis provides a spatially explicit prediction for the 50-year event magnitude over the Continental United States, which as an example underlines the high quantile values corresponding to the South East, the Gulf coast, and the Pacific North West. The coherent probabilistic nature of the Bayesian HMEV can be exploited to assess the uncertainty of extreme value quantiles. For example, in Figure 3.10 the width of the 90% credibility intervals for a return time of 50 years—normalized for the corresponding quantile—is proportional to the size of each dot. This relative measure of uncertainty appears to be larger in the Western USA, characterized by a drier climate and lower values of 50-year quantiles.

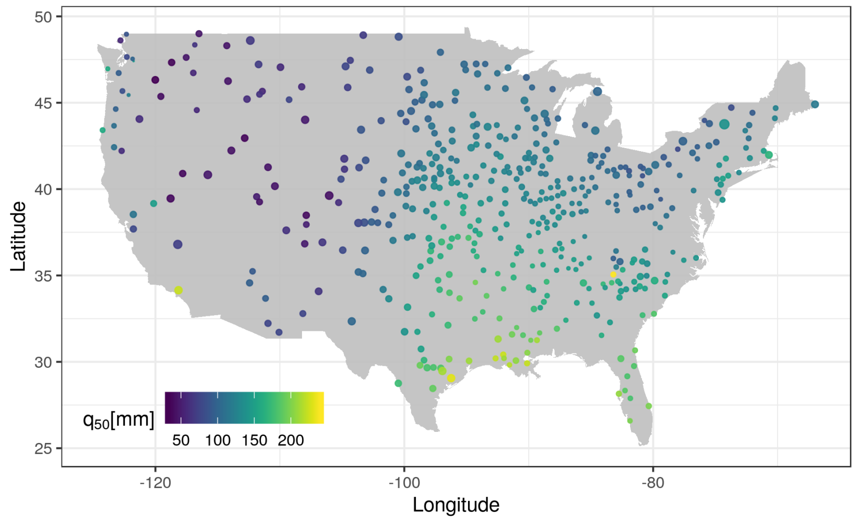


Figure 3.10: Spatial distribution of the HMEV quantiles (color shading) and related normalized uncertainty (dot dimension) corresponding to an average recurrence interval of $T_r = 50$ years computed from 50-years samples extracted from the 479 USHCN stations included in the analysis. Normalized uncertainty computed as the ratio between the width of the 90% credibility interval normalized and the posterior expected value of the quantile.

3.5 Discussion

I introduced a Bayesian hierarchical model to make inference on extreme values of intermittent sequences, with underlying model parameters possibly varying over time. I applied this approach both to synthetic and real data, testing its performance in estimating high quantiles, and provided a benchmark of its performance against some commonly used extreme value models. The proposed approach significantly reduced uncertainty in extreme value frequency estimation, which I attribute to the increased amount of observational information used and to the ability to leverage available information regarding the parent distribution describing the underlying physical process. This advantage becomes crucially important for short observational time series, and especially for large extremes in the right-most part of the distributional tail.

My findings show that, when the underlying process generating the observations x_{ij} is well approximated by a parametric model—such as the Weibull distribution adopted here— use of an asymptotic extremal model leads to the loss of a large amount of information and to the subsequent inflation of the posterior uncertainty.

While the ability of the proposed model to describe the tail of different processes appears to be dependent on the specific marginal distribution of x_{ij}, the model structure introduced here exhibits narrower posterior predictive intervals and a lower effective number of parameters when compared to other widely used extreme value models that do not attempt to account for the entire parent distribution. HMEV quantile estimates, in fact, exhibit reduced uncertainty even when the (synthetic) data being analyzed is not generated by the specific parent distribution of ordinary values chosen in the HMEV formulation. Therefore, a clear advantage in applying the HMEV methodology is that posterior predictive tests can be employed to check in-sample goodness-of-fit, and overfitting is minimized.

In addition to these advantages, HMEV, based as it is on the specification of a distribution for all observations, is also amenable to possible extensions and generalizations. For example, at locations where different event-generating mechanisms are present [LL13, MZAM19], one can quite naturally adopt more complex specifications for the distribution of event magnitudes, such as mixtures of parametric distributions, as was done in some MEVD formulations [MZAM19, MVM20]. Finally, the HMEV framework also naturally lends itself to extensions aimed at including possible systematic changes in the probability distributions of ordinary values, e.g. associated with trends in low-order moments derived from observations, climate model projections, or from physical principles that may provide insight into future rainfall characteristic magnitudes (e.g. Clausius-Clapeyron scaling of atmospheric water-holding capacity [AS08]).

Chapter 4

Downscaling of Extreme Rainfall Statistics from Satellite Observations

Adapted from: Zorzetto, Enrico, and Marco Marani. "*Downscaling of rainfall extremes from satellite observations*." Water Resources Research 55.1 (2019): 156-174.

4.1 Introduction

Even though systematic rainfall observations date back more than two centuries [NW08, CBD+13, MZ15], and rain-gauge networks are quite developed internationally [MDV+12], the global density of rainfall observations still exhibits large gaps over continents [KBH+16], with oceans remaining largely ungauged. In recent decades, advances in rainfall remote sensing technologies have contributed to attenuate this chronic lack of spatial information and have made available vast rainfall datasets, with unprecedented resolution in space and time. Rainfall satellite estimates from different sensors (chiefly radar, microwave imagery and infrared sensors) are now routinely combined to produce global grids of Quantitative Precipitation Estimates (QPE). Satellite QPEs, and in particular observations from the Tropical Rainfall Measuring Mission (TRMM) [HBN+07] and the Global Precipitation Measurement (GPM) mission [HBB+14], greatly improve our knowledge of global precipitation dynamics, with implications for a wide variety of water-related disciplines, from water resources engineering, to risk evaluation and management, to ecology and eco-hydrology. However, the quantification of the accuracy associated with satellite QPEs encounters the basic difficulty of quantitatively comparing them with reference observations at the ground, which are inevitably performed at much smaller spatial scales. In fact, it is well known that, when a rainfall field is aggregated over an area, its statistical properties change with the size of the averaging area and according to the correlation properties of the rainfall field itself. These scale properties of rainfall fields have been investigated both from the point of view of random processes [Bel87, CI88, CKO02, Mar05, Van10] and using the formalism of random cascades [SL87, GW93, OG96, NB15]. A central problem is developing relations between the properties of rainfall averaged over coarse spatial scales (as in the case of remote sensing QPEs, with common resolutions varying between $10^1\ km^2$ and $10^2\ km^2$), to those measured at a point in space, the reference source of rainfall observations.

Given the important implications associated with an accurate quantification of rainfall from space, calibrating, testing, and quantifying uncertainty in satellite QPEs using rain-gauge data is an important open problem [HH08, VK07, PLW10, MT13, LSLC16]. The comparison of ground-based radar and satellite sensor statistics has been object of extensive research. Kirstetter et al. [KHG+12] introduced a framework to evaluate the performance of space-borne precipitation sensors based on ground radar mosaics. They proposed a weighted average of ground radar observations which accounts for the power gain function of the space-born sensor, assumed to be Gaussian. Gebremichael et al. [GK04a] investigated to what extent radar-derived rainfall products can capture small scale rainfall variability. They employed a point to area conversion of the correlation function and found that radar-estimated correlations tend to be lower than those observed by rain gauges at spatial distances shorter than $5km$.

Müller and Thompson [MT13] proposed a bias correction procedure for TMPA 3b42 QPEs based on a stochastic representation of the rainfall field. Rainfall statistics observed at measuring stations are interpolated and used to estimate properties of the rainfall field at the TMPA pixel scale and correct TMPA QPEs accordingly. While the method can be used for extrapolations to ungauged pixels, its calibration requires observations from a sufficiently large number of measurement stations within the same TMPA pixel to fit the stochastic model to the rainfall field. A downscaling technique that only makes use of TMPA-measured rainfall statistics (mean, variance and number of wet days) was recently proposed [DJRRI15] based on a simple stochastic model of point rainfall [CI88]. While the proposed technique does not require knowledge of ground information for its calibration, the assumptions made on the point rainfall process (chiefly, the exponential distribution of rainfall duration and intensity) hinder its application to the study of rainfall extremes.

The estimation of extreme rainfall return levels is particularly affected by the uncertainty and measurement errors that characterize space-borne rainfall retrievals, and is further hindered by the short observational coverage provided by satellite sensors (currently less than 20 years for TRMM and GPM retrievals). Under these premises, the quantification of the frequency of occurrence of rainfall extremes is inherently difficult, as large quantiles are, by definition, poorly sampled in short observational time series. What is more, traditional extreme value analyses, based on the use of just annual maxima (AM), or of relatively few values over a high threshold [Col01], discard most of the information contained in already short QPE time series, and are thus extremely sensitive to the observational uncertainty of a small number of observations. As a result, standard extreme value analyses applied to QPE data are inevitably affected by large and difficult to quantify uncertainties, that severely limit its use for quantitative predictions [ZLH15].

The examination of the literature points to significant gaps in 1) relating the statistics of rainfall observed at coarse spatial scales with those observed at a point,

2) the use of remote sensing observations to derive extreme rainfall properties. The main objective of this Chapter is to bridge these gaps by introducing a novel statistical downscaling methodology with specific focus on extreme rainfall stastistics. To this end, I build a framework that 1) relates daily rainfall statistics from area-integrated remote sensing QPEs to those from point measurements at the ground, and 2) infers extreme rainfall statistics via the Metastatistical Extreme Value Distribution (MEVD) [MI15, ZBM16, MNAM18]. This approach simultaneously addresses the issues related to the short sample sizes and to the coarse spatial resolution of satellite QPEs. MEVD links the probability distribution of extreme events to the entire underlying distribution of "ordinary" daily events, here defined as all the daily rainfall accumulations greater than a fixed and low threshold. The MEVD approach has been shown to significantly reduce estimation uncertainty, with respect to traditional extreme value analysis methods, particularly for values of the average recurrence interval (*Return Time*, T_r) larger than the length of the time series used for calibration [ZBM16]. Furthermore, MEVD estimates are defined using the entire set of observations, rather than just a portion of the distributional tail, and thus produce estimates that are less sensitive to the observational uncertainty and to the presence of outliers in QPE datasets. Marra et al. [MNAM18] recently tested the MEVD framework using synthetic rainfall time series perturbed with errors typical of satellite observations, finding that the method is more robust to these source of error than traditional extreme value models, thus supporting this first application of the method to satellite QPEs.

This MEVD-based downscaling methodology is here tested using TRMM TMPA Research Version 7 3b42 QPEs over the Little Washita watershed, Oklahoma, where the dense Micronet rain gauge network allows an accurate description of the spatial distribution of rainfall [ESC+93, VK07, VKS09].

4.2 Materials and Methods

The instantaneous rainfall rate at the ground can be regarded as a three-dimensional random field $i(x, y, t)$, described by a suitable set of coordinates (here t indicates time, and x and y are the rectangular coordinates of a point on Earth's surface). Both rain gauges and remote sensing observations provide an integral representation of $i(x, y, t)$ on a finite space-time domain centered at a point (x_c, y_c, t_c), which can in general be expressed as

$$h_L(x_c, y_c, t_c) = \frac{1}{L_x L_y} \int_{x_c - \frac{L_x}{2}}^{x_c + \frac{L_x}{2}} \int_{y_c - \frac{L_y}{2}}^{y_c + \frac{L_y}{2}} \int_{t_c - \frac{T}{2}}^{t_c + \frac{T}{2}} i(x, y, t) dx dy dt \tag{4.1}$$

where the rainfall process $i(x, y, t)$ is averaged over a rectangular spatial domain, with sides L_x and L_y, and is integrated over a time interval T. For example, in the case of traditional rain gauges, rainfall volume is recorded as an integral over finite time intervals (e.g., hourly or daily), while the measurement can be regarded as being performed at point in space, given the small integration area (order of 10^{-2} m^2). Conversely, the reflectivity fields retrieved by a radar may be regarded as average values over a spatial domain corresponding to the size of a radar beam (order of square kilometers). For example, retrievals by the precipitation radar (PR) on board the TRMM mission are best interpreted as weighted averages within each radar beam, with weights that depend on the characteristics of the sensor [KHG+12]. The time scale of a radar retrieval is very short, as it is the result of the quasi-instantaneous detection of hydrometeors within layers of the atmospheric column. Here I focus my attention on the TMPA multi-sensor product, which can be regarded as pixel-average QPEs, and perform my analyses on precipitation retrievals aggregated at the fixed daily time scale ($T_d = 24$ hours)

$$h(x,y) = \int_{t_c-\frac{T_d}{2}}^{t_c+\frac{T_d}{2}} i(x,y,t)dt \tag{4.2}$$

so as to investigate the effects of spatial averaging alone on the statistical properties of satellite-sensed rainfall fields. To this end, in the next subsections (i) I introduce a framework for linking the daily rainfall distribution at a point and its counterpart averaged over an area of a given size, (ii) I present a methodology for inferring the correlation structure of the rainfall field from satellite QPEs, and (iii) I combine this information to estimate the MEVD at a point in space from satellite area-averaged observations.

4.2.1 Scale-wise variation of the distribution of daily rainfall

The occurrence of daily rainfall at a point can be described by a stochastic process alternating between dry and wet states. The probability distribution of this compound process is characterized by a finite atom of probability in zero and its moments, in particular its mean μ_{c_0} and variance $\sigma^2_{c_0}$, differ from the corresponding statistics of the wet process only (non-zero rainfall, with mean and variance μ_{r_0} and $\sigma^2_{r_0}$ respectively). Here the first subscript refers either to the *compound* process (c) or to the wet (rainfall being detected) component only (r), while the second subscript distinguishes the process at a point (0) from the process averaged at the spatial scale of satellite retrievals (L). L is characteristic linear scale of the satellite pixel measurement, defined as the square root of the pixel area ($L = \sqrt{L_x L_y}$). While I am interested in the distribution of 'wet' events when studying extremes, it is necessary to also consider the compound process as rainfall observations averaged at large scales also include zero-rainfall areas.

For the compound process, the reduction of variance connected with the spatial averaging of the rainfall field can be expressed by a variance function of the form

$$\gamma_0 \left(L \right) = \frac{\sigma_{c_L}^2}{\sigma_{c_0}^2} = \frac{4}{L_x^2 L_y^2} \int_0^{L_x} \int_0^{L_y} \left(L_x - x \right) \left(L_y - y \right) \rho(x, y) dx dy \tag{4.3}$$

where $\sigma_{c_L}^2$ is the variance of the (compound) rainfall field averaged in a space domain of area $A = L_x \cdot L_y$ corresponding to the QPE pixel size, $\sigma_{c_0}^2$ is its variance at a point and $\rho(x, y)$ is the spatial correlation function of the process, here assumed to be quadrant symmetric [Van10]. While γ_0 in general depends on the shape of the area over which the average is carried out, here I apply the method to TMPA time series, and therefore I assume that γ_0 only depends on the linear characteristic scale of the TMPA pixel, L.

The intermittent nature of rainfall fields implies that not only the variance of the process (and its higher-order moments), but also the yearly number of wet days N_L of the averaged process will in general differ from the yearly number of wet days N_0 observed at a point, as it is possible to have a pixel-averaged value of rainfall greater than zero when only part of the pixel is actually wet [VMKM08]. Here I characterize this effect by introducing a factor β_0, defined as the ratio between the probability of the pixel of size L being wet, p_{r_L}, and the probability of any given point inside it being wet p_{r_0}, i.e., $\beta_0 \left(L \right) = p_{r_L} / p_{r_0}$.

I require that the average of the rainfall process must remain constant when averaging in space, so that $\mu_{c_L} = \mu_{c_0}$, as a consequence of the conservation of mass. However, the average of the wet process will not in general be conserved across scales, such that

$$\mu_{r_0} = \mu_{r_L} \beta_0 \tag{4.4}$$

Analogously, the variance of the intermittent rainfall process can be linked to the variance of the wet events by means of the following relation, that holds at any spatial scale (see Appendix C for its derivation)

$$\sigma_c^2 = \sigma_r^2 p_r + \mu_r^2 \left(1 - p_r\right) p_r \tag{4.5}$$

whence using the definition of the variance function γ_0 from equation (4.3) I obtain

$$\sigma_{r_L}^2 p_{r_L} + \mu_{r_L}^2 \left(1 - p_{r_L}\right) p_{r_L} = \gamma_0 \left[\sigma_{r_0}^2 p_{r_0} + \mu_{r_0}^2 \left(1 - p_{r_0}\right) p_{r_0}\right] \tag{4.6}$$

and thus, using equations (4.4) and (4.6), one can obtain the statistics of the point process in terms of area-averaged quantities

$$\sigma_{r_0}^2 = \frac{\beta_0}{\gamma_0} \left[\sigma_{r_L}^2 + \mu_{r_L}^2 \left(1 - p_{r_L}\right)\right] - \mu_{r_0}^2 \left(1 - p_{r_0}\right) \tag{4.7}$$

Several parametric distributions have been proposed to model the non-zero daily rainfall accumulations, ranging from exponential-type distributions to stable distributions (e.g., see [GK54, MS00b]). Here I employ a Weibull distribution to describe the wet component of the daily rainfall process across spatial scales. I note that i) the Weibull distribution has been found to appropriately describe daily rainfall across many different climates globally, while sharing a formal analogy with the multiplicative nature of convective processes [FS97, WT05], and ii) the Weibull distribution describes both light- and heavy-tailed random variates with characteristic scale through a parsimonious 2-parameter model [LS98]. However, I note that the approach I propose here is quite general, and can be tailored to processes characterized by different 2-parameter distributions with only minor modifications.

The Weibull distribution is here parametrized with a scale parameter C and a shape parameter w, so that the cumulative probability distribution of the random variable H_L, representing daily rainfall accumulations at the spatial scale L, reads $P(H_L < h_L) = 1 - \exp\left[-\left(h_L/C_L\right)^{w_L}\right]$. The first two central moments of this distribution (i.e., its mean and variance) are respectively

$$\mu_{r_L} = \frac{C_L}{w_L}\Gamma\left(\frac{1}{w_L}\right) \tag{4.8}$$

and

$$\sigma^2_{r_L} = \frac{C_L^2}{w_L^2}\left[2w_L\Gamma\left(\frac{2}{w_L}\right) - \Gamma\left(\frac{1}{w_L}\right)^2\right] \tag{4.9}$$

where Γ denotes the Gamma function. Eq. (4.4) applied to the case of a Weibull variate yields a relation linking the shape and scale parameters of the process at two different scales

$$\left(\frac{C_0}{w_0}\right)^2 = \beta_0^2\left(\frac{C_L}{w_L}\right)^2\frac{\Gamma^2\left(\frac{1}{w_L}\right)}{\Gamma^2\left(\frac{1}{w_0}\right)} \tag{4.10}$$

A similar argument can be applied for the variance of the compound process, defined in equation (4.7), using the expression for the Weibull moments (equations (4.8) and (4.9)) and eq. (4.10)

$$\gamma_0\beta_0\frac{2w_0\Gamma\left(\frac{2}{w_0}\right)}{\Gamma^2\left(\frac{1}{w_0}\right)} = \frac{2w_L\Gamma\left(\frac{2}{w_L}\right)}{\Gamma^2\left(\frac{1}{w_L}\right)} + \left(\gamma_0 - 1\right)p_{r_L} \tag{4.11}$$

Eq. (4.11) is nonlinear and can be solved numerically to determine the value of the shape parameter, w_0, at a point in space. Finally, eq. (4.10) yields the value of the scale parameter C_0 at a point. This procedure can be used to infer the parameters of the probability distribution of non-zero rainfall values (the 'wet process') at a point from the parameters describing the distribution of area-averaged values, provided the values of γ_0 and β_0 are known.

Here I define a *wet day* (i.e., a day in which an 'ordinary' rainfall event occurs) as a 24-hour period characterized by rainfall amounts greater or equal to $q = 1$ mm. This threshold value is assumed constant across spatial scales and different data sources. Defining wet days using a fixed and low threshold is necessary when applying the method to data from both rain gauges and satellite sensors, which are inherently characterized by different detection thresholds. The 1 mm threshold value is coherent with the guidelines by the World Meteorological Organization [KTZZ09]. The analysis proceeds by applying the Weibull distribution of ordinary daily rainfall accumulations to the excess above the threshold, $y = h - q$. I note that the distribution of y corresponds to the distribution of the rainfall accumulations h conditional to being above threshold: $P(Y \leq y) = P(H \leq y + q | H > q)$. Accordingly, the number of wet events and their statistics are computed for the population above this detection threshold.

I also note that the above argument can be extended to link the pdf of daily rainfall between two spatial scales L_1 and L_2, with e.g., $L_1 > L_2$. In this case, the ratios of the variance and wet fraction of the process averaged at these two spatial scales can be respectively expressed as $\gamma(L_1, L_2) = \gamma_0(L_1)/\gamma_0(L_2)$ and $\beta(L_1, L_2) = \beta_0(L_1)/\beta_0(L_2)$. Next, I explore how the value of the two ratios, γ_0 and β_0, can be estimated from satellite retrievals.

4.2.2 A Model for the correlation structure of daily rainfall

The equations relating the Weibull distributional parameters at a point in space to those valid for areal-averaged rainfall (eqs. (4.10) and (4.11)) require knowledge of the variance function γ_0 (eq. (4.3)), which, in turn, depends on the correlation structure $\rho(x, y)$ of the rainfall field. Relating the correlation structure of the continuous stochastic field (whose realizations are the rain gauge point observations) to that of the same field averaged over finite areas (whose realizations are the satellite QPEs) requires additional attention to the issue of spatial scale. In fact, spatial averaging does not only affect the probability distribution of rainfall daily totals, but it also modifies the correlation between their values at two points in space when their distance is commensurate with the characteristic length scale L of the averaging area.

The covariance between the local averages h_L and h'_L of the precipitation field performed over two different pixel areas with the same characteristic size L can be expressed as [Van10]:

$$Cov\left[h_L, h'_L\right] = \frac{\sigma_{c0}^2}{4\left(L_x L_y\right)^2} \sum_{k=0}^{3} \sum_{l=0}^{3} \left(-1\right)^k \left(-1\right)^l \triangle\left(L_{x,k}, L_{y,l}\right) \tag{4.12}$$

where σ_{c0}^2 is the variance of the process at a point, and the quantities $\triangle(L_{x,k}, L_{y,l})$ are the analogues of the variance function in eq. (4.3) valid for the integral of the random field over a finite area of sizes $L_{x,k}$ and $L_{y,l}$

$$\triangle(L_{x,k}, L_{y,l}) = 4\int_0^{L_{x,k}} \int_0^{L_{y,l}} \left(L_{x,k} - s_1\right)\left(L_{y,l} - s_2\right)\rho(s_1, s_2) ds_1 ds_2 \tag{4.13}$$

The set of distances $L_{x,k}$ and $L_{y,l}$ (with $k, l = 0, 1, 2, 3$) encodes all the necessary information on the relative position of the two pixels over which the averages of the rainfall field are computed. If Δx and Δy are the distances, along the x and y direction respectively, between the centers of two pixels, $L_{x,k}$ is defined as

- $L_{x,0} = \Delta x - L_x$ (distance between the end of the first pixel and the beginning of the second, along coordinate x)

- $L_{x,1} = \Delta x$ (distance between the beginning of the first pixel and the beginning of the second, along the coordinate dimension x)

- $L_{x,2} = \Delta x + L_x$ (distance between the beginning of the first pixel and the end of the second, along the coordinate dimension x)

- $L_{x,3} = \Delta x$ (distance between the end of the first pixel and the end of the second, along the coordinate dimension x)

Analogous definitions hold for $L_{y,l}$ along the y direction.

The expression of the covariance (eq. (4.12)) can be used to obtain the correlation, ρ_{h_L,h'_L}, between the two pixel-averaged time series as a function of their relative position:

$$\rho_{h_L,h'_L} = \frac{Cov[h_L, h'_L]}{\sigma_{h_L}\sigma_{h'_L}} = \frac{\sum_{k=0}^{3}\sum_{l=0}^{3}(-1)^k(-1)^l \triangle(L_{x,k}, L_{y,l})}{4\triangle(L_x, L_y)} \tag{4.14}$$

where the covariance of local averages has been divided by the variances σ_{h_L} and $\sigma_{h'_L}$ of the process averaged over the two pixels respectively.

If a parametric analytical expression is available for the correlation $\rho(x, y)$ of the precipitation field at a point, then equation (4.14) can be used to determine its parameters by matching the right hand side with the inter-pixel correlation observed from a satellite-sensed rainfall field. Note that, under the hypothesis of an isotropic rainfall field, the two-point correlation $\rho(x, y)$ only depends on the distance between two points in space, so that it can be expressed as a function $\rho(d)$ with $d = \sqrt{x^2 + y^2}$.

Here I use a correlation structure characterized by an exponential kernel (EK) and a power-law tail [Mar03].

$$\rho(d) = \begin{cases} e^{-\frac{\alpha d}{\epsilon}} & d < \epsilon \\ \left(\frac{\epsilon}{ed}\right)^{\alpha} & d \geq \epsilon \end{cases} \tag{4.15}$$

This expression is continuous with continuous derivative in $d = \epsilon$, and can describe both light- and heavy-tailed families of correlation decay. In the following, I apply this model to test the ability of eq. (4.14) of reproducing the variation in the rainfall spatial correlation between the satellite pixel and the point scale.

The spatial correlation of the process averaged at the pixel scale is obtained by directly computing the value of the Pearson correlation coefficient between pairs of TMPA pixels located at different distances. While the Pearson correlation is known to be a possibly biased estimator when applied to a skewed and intermittent process such as rainfall [HKC01], alternative estimates would require ad-hoc hypotheses on the rainfall distribution and on the conditional probability of zero rainfall [VMKM08]. I therefore choose to use the classic Pearson correlation estimator, which entails the minimal number of additional assumptions.

If n_s TMPA grid cells are used in the estimation of the spatial correlation, $m = n_s(n_s - 1)/2$ estimates of the correlation (ρ_j, for $j = 1, 2, ...m$) are obtained, each corresponding to a distance d_j between the centers of the pair of pixels considered. The TMPA-observed spatial correlation is then assumed to match the correlation function of the area-averaged process, given by eq. (4.14), which depends on the unknown parameters ϵ and α defining the point correlation function. A Sum of Square Errors (SSE) is computed as

$$SSE\left(\epsilon, \alpha\right) = \sum_{j=1}^{m} \left[\rho_{h_L, h'_L}\left(d_j; \epsilon, \alpha\right) - \rho_j\right]^2 \tag{4.16}$$

The quantity $SSE(\epsilon, \alpha)$ in equation (4.16) is then minimized using the *L-BFGS-B* algorithm [BLNZ95] to obtain a best estimate of the parameters (ϵ, α). Once the correlation function at a point is known, it can be used in eq. (4.3) to compute the variance reduction function, $\gamma_0\left(L\right)$, necessary to obtain the distribution of rainfall accumulations at a point.

When computing the variance and covariance of local averages (eqs. 4.3 and 4.14)

I assume a square pixel so that $L_x \simeq L_y \simeq L$. Therefore, $\gamma_0 = \gamma_0(L)$ is simply a function of the linear characteristic scale L of the pixel.

4.2.3 Downscaling of the yearly number of rainfall events

The last piece of information necessary to reconstruct the pdf of daily rainfall at a point is the ratio β_0, or '*intermittency function*', which accounts for the variation in the yearly number of rainfall events when the rainfall field is averaged over the pixel area. Here I propose an application of *Taylor's Frozen Turbulence Hypothesis* [Tay38] to use TMPA information at finer temporal resolution (down to three hours) to infer the sub-pixel scale intermittency of the rainfall process at the daily scale. The Taylor hypothesis has been previously applied to study the space-time scaling of rainfall fields (e.g., [Dei00]) and was employed to compare precipitation products at different spatial scales [HEM$^+$15] . Here I apply a similar argument to estimate the spatial wet fraction of the compound rainfall field from its variability in time as inferred from TMPA 3b42 data only.

When applied to rainfall measurements, the Taylor hypothesis states that statistical properties of the rainfall field sampled at a spatial aggragtion scale X and instantaneously in time are equivalent to the same properties sampled at a temporal scale $T = X/U$ and at a point in space, where U has the meaning of an average 'advection' speed. Therefore, according to this hypothesis, properties of the field (such as the wet fraction p_r of the compound rainfall process), when advected past a rain gauge, do not change significantly over time. This assumption holds, for example, for turbulent flows characterized by small turbulent intensity (i.e., the root mean squared longitudinal velocity fluctuation must be small compared to the mean advection speed) [Stu12].

If the Taylor hypothesis holds exactly, then the contour lines where $p_r(X, T)$ is constant are straight lines in (X, T) space. More generally, the advection velocity may vary with the aggregation scale of the process, consistent with the rainfall process being described by a multifractal field [EBH+15]. Following this assumption, I define an advection velocity U as the ratio between differences of integration scales in space and time respectively that would produce the same observed difference in the quantity p_r.

Here I estimate the wet fraction p_r using the TMPA 3b42 precipitation time series integrated at different temporal ($T = 3, 6, 9, 12, 24, 36, 48$ hours) and spatial scales ($X = L, 2L, 3L$, corresponding to local averages over one TMPA pixel, 2x2 and 3x3 pixels respectively), as shown in Figure 4.1. The measured values of p_r are then interpolated in time (here using $n_p = 1000$ values of the temporal scale T in the range from 3 to 48 hours), and the local slope of the p_r contour levels is used to estimate the local advection velocity U.

I select a target aggregation scale (X_G, T_G) (e.g., the aggregation scale in space and time of rain gauge measurements) and identify the unique values $\hat{X}_0$ and $\hat{U}$ such that the line $X = \hat{X}_0 + \hat{U} \cdot T$ passes through the target scale and has a slope equal to the local advection speed. The local advection speed is evaluated as the slope between two points with spatial aggregation scales $X = L$ and $X = 2L$ (corresponding to the local averages over 1 pixel and 2x2 pixels respectively) and temporal scales determined by two conditions: i) the two points share the same value of the observed wet fraction $p{*}_r$, and ii) when extrapolating to the rain gauge scale, the resulting line passes through the target scale (X_G, T_G) (blue line in Figure 4.1). These two conditions together uniquely determine the two quantities $\hat{X}_0$ and $\hat{U}$, and thus can be used to compute the unknown property p_r^* at the rain gauge scale simply extrapolating TMPA observations. The intermittency function can then be evaluated as $\beta_0 = p_r(L, T_d)/p_r^*$, i.e., as the ratio between the wet fraction at the 1-pixel, daily scale $p_r(L, T_d)$, and the wet fraction p_r^* extrapolated at the rain gauge scale (X_G, T_G). This procedure assumes that the contour lines of the wet fraction in the (X, T) plane can be approximated to straight lines for spatial scales smaller than $2L$.

4.2.4 Extreme value model

I have seen how my hypotheses on the spatial structure of the rainfall fields yield a model for the marginal distribution of daily rainfall at a point. This information can now be used to estimate rainfall extremes. I base my analysis on the MEVD [MI15, ZBM16], which expresses the cumulative distribution function of block-maxima, $\zeta(h)$, of independent variates ("ordinary values") distributed according to an underlying parent distribution with cumulative probability function $P(H \leq h) = F(h; \vec{\theta})$, as

$$\zeta(h) = \sum_{N=0}^{\infty} \int_{\Omega_{\vec{\theta}}} g(N, \vec{\theta}) F(h; \vec{\theta})^N d\vec{\theta} \tag{4.17}$$

where $\vec{\theta}$ is the set of parameters describing the parent distribution $F(h, \vec{\theta})$, $\Omega_{\vec{\theta}}$ is their population, N is the number of events/block, and $g(N, \vec{\theta})$ is the joint probability density function of N and $\vec{\theta}$. In order to avoid ad-hoc assumptions about the expression for $g(N, \vec{\theta})$ [ZBM16], I use a sample mean in place of the ensemble mean when evaluating (4.17) from a sample time series of s years. I further adopt, as customary in extreme precipitation analysis, 1-year blocks, and, for $F(h; \vec{\theta})$, a Weibull distribution, such that the $\vec{\theta} = (C, w)$ and the MEVD cumulative distribution function of yearly maxima becomes:

$$\zeta(h) = \frac{1}{s} \sum_{j=1}^{s} \left[1 - e^{-\left(\frac{h}{C_j}\right)^{w_j}} \right]^{N_j} \tag{4.18}$$

where Weibull parameters and number of wet days (C_j, w_j, N_j) are allowed to vary across years. Zorzetto et al.[ZBM16] extensively study the properties of MEVD and find, over a large dataset of long daily rain-gauge time series, that it significantly reduces extreme-value estimation uncertainty with respect to traditional approaches hinged on the Generalized Extreme Value distribution when only relatively small samples (with respect to the return time of interest) are available for calibration. In the present context, the parameters of the Weibull distributions in eq. (4.17) are estimated for each year on record ($j = 1, 2, ..., s$) from the TMPA dataset and then downscaled to the point scale following the procedure described in the previous sections. The Weibull distribution is fitted by means of the probability weighted moments approach [GLMW79], following [ZBM16].

While the size of the estimation window over which Weibull parameters are estimated (not to be confused with the blocks over which maxima are determined, which typically remain yearly in applications) can in general be varied, I limit here my analysis to yearly estimates. I note that in the case of very dry climates (i.e., low number of events/year) this choice may not be optimal, and longer windows could improve parameter estimation. On the other hand, using estimation windows that are longer than necessary is not advantageous as it reduces the variability of extremes associated with interannual variabilities (and, possibly, with systematic long-term changes). I suggest that such metastatistical source of variability in the ordinary events plays an important role in the emergence of fat-tailed extreme values [ZBM16]. This is in general justified as the rainfall process is the result of a mixture of different mechanisms that appear every year with different frequencies and thus determine a variation in shape and scale parameters of the parent distribution of ordinary rainfall values.

4.2.5 Satellite rainfall data

The Tropical Rainfall Measurement Mission (TRMM) Multisatellite Precipitation Analysis (TMPA) 3b42 dataset provides a 19-years long, quasi-global coverage of tropical and subtropical regions (Latitudes between $-50°$S and $+50°$N), with remarkable spatial (0.25° x 0.25°) and temporal (3-hourly) resolution. The TMPA 3b42 version 7, research-quality, dataset provided by NASA and used in this study can be accessed at

`https://mirador.gsfc.nasa.gov/`. TMPA estimates are obtained by merging information from a set of different sensors (primarily passive microwave and infrared) characterized by differing accuracy and resolution, with the purpose of improving the overall quality and coverage of the final (level 3) product. As a final step, ground-based gridded precipitation data from the Global Precipitation Climatology Project (GPCP) are used to correct TMPA precipitation estimates to preserve ground-estimated monthly means (see [HBN+07] for details). For the purpose of this study, I extracted a lattice of $3x3$ TMPA pixels centered at the point with coordinates $34.785N, -98.125E$ (see Figure 4.2) to match the ground-based data used for testing and validation of the methods. TMPA 3b42 gridded rain rate fields are available every three hours. Here I regard these values as average quantities over the observation interval. After correcting for the local time zone (since data are reported at a nominal observation time), I compute daily totals based on the three-hourly rainfall rates. The complete TMPA record in the interval 1998-2015 was used in the analysis, after testing that for each QPE time series the yearly fraction of missing data was less than 10%. Missing data values were set to zero.

4.2.6 Ground-based observations

To test estimates of point-value extremes from TMPA data, I take advantage of a dense rain gauge network (ARS Micronet), located in the Little Washita River watershed, Oklahoma [ESC+93]. The network consists of 42 tipping bucket rain gauges, 23 of which fall within one single pixel of the TMPA 3b42 gridded product (see Figure 4.2). This particular network was selected to test my methodology because of (i) the remarkable spatial density of the stations, and (ii) the extensive characterization of TMPA performance available for this particular location (e.g., [VK07, HH08]). I identify and use for my analysis a subset of the rain gauges in the network that fall within a single TMPA pixel and for which a continuous 19-year record exists (1998-2016), providing a perfect temporal overlap with the TMPA dataset. The daily data from the Oklahoma Micronet were obtained from `http://ars.mesonet.org/` and were pre-processed by removing data that the were marked as affected by certain or highly probable instrumental error [ESC+93].

For testing downscaling results with ground observations, I use the Micronet time series for which an almost complete record exists in the interval 1998-2015, so as to match exactly the temporal range of TMPA QPEs. Of the 23 Micronet stations over the pixel centered in $34.785N, -98.125E$, only 7 have a continuous record in this range of years (with a maximum number of missing values/year less than 42). I use this subset of the stations to fit the distribution of daily rainfall accumulations at a point and to compare it with the corresponding TMPA pixel-average and downscaled distributions. For testing the TMPA-downscaled spatial correlation function, all stations in the Micronet network are used.

Due to the limited length of the TMPA dataset used in this study, it is desiderable to use an independent, and longer, set of observations in close proximity to the

Micronet, to be able to empirically evaluate quantiles with relatively large return periods. For this reason, I selected 4 stations from the National Oceanic and Atmospheric Administration (NOAA) Global Historical and Climatology Network Daily (GHCND) dataset (see Table 4.1). For each station, only years with less than 10% of missing years were included in the analysis. The stations were selected on the basis of (i) the length of the record (set to be at least 50 years), and (ii) their proximity to the Micronet stations. As shown in the following, the precipitation field is still highly correlated at distances larger than the maximum distance between these stations, suggesting that observations from the GHCND gauges and the Micronet stations can be regarded as samples of the same rainfall process.

4.3 Results

4.3.1 Correlation structure and downscaling of daily rainfall distribution

I start by comparing the spatial correlation function computed from the 9 TMPA time series corresponding to the pixels covering the study area, the downscaled point correlation function obtained by the minimization of eq. (4.16), and the correlation independently estimated using the Micronet rain gauges at the ground (Figure 4.2). The analysis of inter-station correlation for the Micronet network reveals that this site is characterized by a slowly decaying correlation (Figure 4.3). The values of the parameters defining the correlation model indicate a decay slower than predicted by a simple exponential (the transition to power-law behavior occurs at about ϵ=14 Km, see Table 4.2), suggesting caution in using the common exponential correlation functions arising from spatial Poisson models [CI88, DJRRI15]. The rainfall fields obtained from the TMPA 3b42 dataset exhibit greater spatial correlation than those from the point process, as expected. To test how well the correlation structure of the continuous rainfall field can be reconstructed based on TMPA data alone, I minimize the SSE (4.14) and determine the parameters (ϵ and α) describing the point to point correlation. Figure 4.3 shows that the proposed downscaling procedure yields correlation values which are consistent with the point correlation values estimated from the Micronet network observations (see Table 4.2 for the specific parameter values). For distances larger than 40 Km (extrapolating beyond the range of distances available for fitting the spatial correlation function), the downscaled correlation model appears to decay faster than the one obtained from fitting the point observations.

This evidence confirms the importance of correctly accounting for the area-averaged nature of remotely-sensed information when comparing correlation functions from satellite and ground-based precipitation observations. The downscaling exercise was also repeated by keeping the exponent α constant across scales, and by solving the optimization of eq. (4.16) only to determine the value of ϵ, the scale parameter of the correlation function at a point. This yielded a similar result to the one obtained by minimizing both parameters (Figure 4.3). Application of eq. (4.3) using the TMPA-downscaled spatial correlation function yielded a value of the variance reduction function $\gamma_0(L) = 0.89$ between the pixel and the point scale.

Next, I compare the values of the intermittency ratio β_0 obtained by applying Taylor's hypothesis as described in Section 4.2.3 with ground and satellite observations averaged at different scales. The procedure based on the Taylor hypothesis yields a value of $\beta_0(L) = 1.09$, slightly larger than the value obtained by averaging the intermittency ratio from Micronet time series measured at the ground ($N_L/N_0 = 1.05$). Figure 4.4 shows a comparison of the yearly number of wet days distribution observed at the ground (Micronet stations) with the corresponding values at the pixel scale.

The distribution of N_L obtained from TMPA QPEs (S_L) and by averaging rain gauges at the ground ($\overline{G_L}$) both exhibit a larger mode when compared to the point values observed at single rain gauges (G_i). I then apply the Taylor hypothesis to also obtain estimates of the number of wet days N_0 at a point in space. The distributions of N_0, obtained respectively from downscaling gauge average ($\overline{G_d}$) and TMPA QPEs (S_d), exhibit lower mean and median than the original distributions of N_L, yielding results consistent with the distribution of values observed by rain gauges at the ground.

I next illustrate the application of the downscaling approach to the transformation of the pdf of rainfall values averaged/observed at a coarse scale to the pdf of rainfall values at a finer spatial scale. Before applying the method to infer extreme value statistics at a point, I first test how well the method reproduces the distribution of ordinary rainfall values at different spatial scales. To do so, I apply the downscaling method to Weibull parameters obtained fitting the entire available satellite and rain gauge time series. To downscale these probability distributions of ordinary events, I proceed as follows. At the coarse scale, the Weibull parameters are estimated by means of the probability weighted moments technique [GLMW79]. I then estimate the values of the parameters at the point scale by use of the downscaling relations, eqs. (4.10) and (4.11), using the values of β_0 and γ_0 obtained from TMPA time series.

I first select the Micronet stations within the pixel centered in $34.875N, -98.125E$ (central pixel in Figure 4.2) with at least 18 years of data, in order to obtain a perfect overlapping with the time interval of the TMPA dataset, and compute their average time series. The downscaling procedure is then applied to this 'exact' areal average rainfall to test whether the approach can recover the distribution of daily rainfall accumulations at a point. Results show that the downscaling procedure provides a good estimation of the distribution at a point (Figure 4.5a), which is within the range of variability of the Micronet stations at the ground. Comparison with the TMPA daily time series over the same pixel shows that its distribution closely resemble the one obtained from data observed at a point (Figure 4.5b). In both cases, the magnitude of the scale correction directly depends on the decay of the correlation function, which exhibits a long tail.

Given the large spatial extent of the correlation over the study area, and in order to more stringently test the downscaling approach formulated here, I consider progressively coarser data obtained by averaging over 2 x 2 pixels (linear characteristic scale $2L$) and over 3 x 3 pixels (scale $3L$) centered in -98.125 E, 34.785 N (see lattice in Figure 4.2). Again, in order to test the downscaling method using homogeneous observations at different scales, I focus my attention on satellite estimates only, performing the downscaling from scale $3L$ to scale L (a single pixel) and from $3L$ to $2L$. This application of the downscaling method confirms that the methodology, even when applied to the coarse observations at the $3L$ scale, is able to correctly reproduce the exceedance probability distribution at a smaller spatial scales (downscaling from scale $3L$ to scales L and $2L$ are featured in figures 4.6a and 4.6b respectively). The shape of the exceedance probability distribution changes more markedly in this case, as large values become significantly less likely at coarser aggregation scales (see green and red lines in Figure 4.6). Overall, I find a better performance when the method is applied and tested using homogeneous data (using either TMPA QPEs or gauges averaged at different spatial scales), while some discrepancy exists when gauge and satellite data are compared, as in Figure 4.5b. These findings suggest that such added differences can be attributed to observational limitations affecting remote sensing QPE values.

4.3.2 Extreme value analysis

After testing how the distributions of ordinary values can be reconstructed at a small spatial scale given knowledge of the distribution at the TMPA pixel scale, I now turn to the inference of the extreme value distribution at the point scale. I first apply MEVD to the stations of the Little Washita Micronet, despite the short record length available. In this case, the empirical quantiles observed at the ground appear to be somewhat overestimated (Figure 4.7), similarly to what happens when downscaling the Weibull parameters from the pixel-average time series. I explain this behaviour by recalling that the MEVD yields an optimal performance when the return time for which a quantile is estimated is greater than the sample size used for the estimation [ZBM16]. In the present case the same sample is used for both calibration and validation, and the relatively short sample limits the range of return times that can be explored in the comparison with ground observations.

Hence, to more accurately investigate the performance of the proposed model over a wider range of return times, I compare MEVD downscaling results with those obtained from rain gauge stations from the GHCND network, with record lengths ranging from 72 to 115 years (Table 4.1). For each station, the corresponding TMPA pixel time series was used to fit the Weibull distribution and estimate the values of parameters C_L and w_L, which were subsequently downscaled to the point scale (C_0 and w_0). Downscaled parameter values were used to construct the MEVD according to equation (4.18) and estimate quantiles for a set of return times up to the length of the available time series. Comparisons are then performed with the empirical quantiles, i.e. the actual annual maxima observed. Results are also comparatively evaluated with estimates obtained by directly fitting GEVD and MEVD to rain gauge-measured time series (Figure 4.8). High quantile estimates obtained fitting

GEVD and MEVD (to time series of annual maxima and ordinary rainfall events, respectively) both exhibit a good match with the empirical quantiles extracted from the same time series. While some overestimation of extreme quantiles is still seen for one of the stations for estimates downscaled from TMPA QPEs, the comparison with longer observational time series shows a good match with empirical quantiles.

The effect of the specific detection threshold used to detect ordinary rainfall events is also evaluated (Figure 4.9). For small values of the threshold ($q < 1$ mm), the estimated quantiles seem to depend on the particular value of the threshold used in the analysis. In particular, this variation is more significant for the TMPA data when compared to the GHCND estimated quantiles. This difference is explained by considering that rainfall values are only recorded at finite intervals by the tipping bucket rain gauges, while discretization effects are negligible for TMPA QPEs. Significantly, variations in QPE-based quantile estimates are modest when one considers values of the threshold of one millimeter or larger. For this reason I have limited my analysis to the value $q = 1$ mm, within the range of values in which results are weakly dependent on the specific threshold value adopted and justified by a largely used definition of 'wet day'.

The considerable length of the GHCND records allows me to further quantify model performance using independent samples for calibration and testing. For each station, observations are resampled with resubstitution to generate realizations that preserve the set of parameters (N, C, w) from the original time series. Subsequently, the synthetic time series thus obtained are divided into two independent sub-samples, of which one is used for calibration and one for testing. The test was performed by extracting from test samples the annual maxima and by estimating the corresponding return times by means of the Weibull plotting position formula [ZBM16]. This procedure was repeated for a number of times $n_g = 100$. For each such bootstrap realization, observed (h^{obs}) and estimated quantiles for a given return time ($\hat{h}\,(Tr)$) were used to compute a Fractional Square Error (FSE), defined as:

$$FSE(s, Tr) = \left(\frac{1}{n_g} \sum_{i=1}^{n_g} \left[\frac{\hat{h}_i(Tr) - h_i^{obs}(s_s, T_r)}{h_i^{obs}(T_r)} \right]^2 \right)^{1/2} \tag{4.19}$$

where s is the calibration sample size and Tr is the return time. This procedure was here applied to rain-gauge calibration samples with length 20 and 30 years respectively, representative of the typical satellite record length, and for values of Tr up to 40 years (limited by the available test sample sizes).

The entire bootstrap procedure was then repeated using the downscaled parameters obtained from the TMPA record, and again using a set of $n_g = 100$ synthetic samples of 40 years, randomly extracted from the relevant GHCND station for testing. In this case only the validation sample (randomly extracted from a GHCND station at the ground) varied, while the calibration sample was kept constant (corresponding to the entire TMPA time series available).

The results of this analysis (Figure 4.10) show that the performance of MEVD and GEVD calibrated using rain gauge data are comparable for small values of the return time (up to about $Tr = 15$ years for the sample size of $s = 20$ years, and to about $Tr = 20$ years in the case of $s = 30$ years). In essence, the advantage in using the MEVD over the GEVD distribution becomes evident for values of Tr greater than the length of the sample used for calibration, as found in previous work [ZBM16]. The extreme values estimates obtained by downscaling TMPA statistics to the point scale (green line in Figure 4.10) yield values of the FSE which are consistently higher when compared to ground observations, as one would expect. However, as the return time increases, the error increases at a rate which is lower than that characterizing ground-based estimates, and results are comparable for return times of 20 years and larger. In this range of return times, estimation uncertainty is particularly large for the GEVD model, as observed when calibration and validation are performed using independent datasets. I note that this result was obtained without applying any bias correction to the TMPA-derived distribution of rainfall accumulations. Overall, these results support the robustness of the method proposed, when applied to relatively short time-series record, as is the case for the TMPA dataset, and when relatively high quantiles need to be estimated.

4.4 Discussion and Conclusions

I developed and tested a new downscaling approach to infer point rainfall extreme value distributions from satellite observations. The approach, outlined in sections 4.2.1 - 4.2.5, introduces a stochastic framework that provides estimates of "ordinary" and extreme value probability distributions at the point scale. The procedure is parsimonious and its application only requires 1) the specification and fitting of the probability distribution of "ordinary" values observed at the coarse aggregation scale, 2) knowledge of the correlation structure of the rainfall field as observed from remote sensing at the coarse scale, 3) knowledge of the intermittency structure of rainfall events as quantified by β_0, ratio of the wet fractions at the pixel and point scales, which can also be estimated from satellite observations using the Taylor hypothesis. A summary of the steps necessary for the application of the method is provided in Appendix C.

The use of the dense Little Washita Micronet rain gauge network allowed some detailed testing of the proposed approach. In particular, I performed downscaling tests from/to the following spatial scales: 75 km (3 x 3 TMPA pixels), 25 km (1 TMPA pixel), and the point scale (rain gauges). The ordinary and extreme value distributions downscaled from coarser scale observations exhibit a good agreement with those obtained from observations at target (smaller) scales when homogeneous observations (either rain gauges or TMPA QPE fields) averaged at different spatial scales are used. When the method is applied to TMPA QPEs and validated using independent ground observations, I observe some discrepancy in the downscaled probability distributions of daily rainfall. Comparison between these results (obtained by comparing downscaled TMPA statistics and rain gauges at the ground) with those obtained using homogeneous observations suggests that this behavior is primarily due to the performance of TMPA QPEs over this particular location. This results is appealing as extensive application of the method can lead to the evaluation of multi-sensor derived precipitation fields over regional to global scales.

For the study location, the application of the MEVD-downscaled method to TMPA data led to errors in quantile estimates that are comparable to those obtained by application of MEVD or GEVD distribution directly to the gauge data for the largest values of return time explored here. This result is quite encouraging in terms of 1) testing remote sensing rainfall observations (and TMPA data in particular) against point observations at the ground at the global scale; 2) evaluating with reasonable and quantifiable accuracy point extremes at the global scale using TMPA (and possibly other rainfall remote sensing) observations.

While this work builds a new framework for estimating rainfall extremes from satellite data, a significant result with many hydrological applications, a number of hypotheses are made. First, the assumption of isotropic correlation structure, while commonly used, could limit the application of the method e.g. in the presence of orographic forcing. However, the covariance expression, eq. (4.12), can in principle be adapted to different, and non-isotropic forms, if additional information on the spatial correlation structure is available/required. This is particularly important in applications over complex terrain. However, the method presented here can be directly applied to higher-resolution satellite products, and in particular to the IMERG dataset, as longer time series become progressively available.

The downscaling method proposed assumes that the form of the probability distribution of rainfall values (Weibull in this case) is preserved across different spatial scales. While it is known that this is not rigorously the case, my results suggest that this assumption approximately holds in the range of scales explored here (linear scales ranging from 0.1 m up to about 75 km).

I tested this approach in a region which is characterized by a relatively simple orography, where a dense network of ground observations is available for independent testing and where TMPA uncertainty is relatively well characterized in the absence of additional confounding factors (e.g., large variability in surface emissivity, high relief, etc.). It will be important to further test the proposed method, which is general in nature and can potentially be applied to any multi-sensor satellite QPE product, by exploring a wider set of locations with different rainfall regimes. In particular, applications to coastal areas, tropical climates and locations with high relief will be particularly challenging, for the coexistence of different precipitation mechanisms and severe storms likely to affect the shape of the daily rainfall pdf as well as the performance of TRMM sensors [RMPG13, LWB18].

In applying the proposed approach at larger scales, I note that, because the downscaling procedure connects satellite-scale and point-scale probability distributions, it can also be used to correct the satellite-inferred ordinary distribution and MEVD in data-scarce regions where only sparse rain gauges are available. Even in the presence of a single rain gauge in a given location, it would be possible to compare satellite-derived point rainfall statistics (β_0, γ_0 and Weibull parameters) with the corresponding values observed at a point. The approach can thus potentially lead to a self-contained procedure, yielding internal bias correction as well as high-quantile rainfall estimates.

The scale-wise dependence of the distribution of daily rainfall was here combined with the MEVD framework to infer extreme value properties of the rainfall field at a point in space. Together, these two steps provide a link between statistical properties of the rainfall process at different spatial scales, with broad implications for hydrological and ecological watershed studies, for the study of water resources in poorly gauged areas, and, ultimately, for better understanding the global distribution of hydrologic extremes.

Station	GHCND Station ID	Latitude	Longitude	Elevation	Time span
Anadarko	USC00340224	35.0667	-98.25	354.8	1893 - 2016
Chickasha	USC00341747	35.05	-97.95	332.5	1901-1965
Chickasha	USC00341750	35.05	-97.91667	331.9	1954 - 2016
Duncan	USC00342660	34.5011	-97.9591	343.2	1936 - 2016
Marlow	USC00345581	34.6368	-97.9786	393.8	1900 - 2016

Table 4.1: Summary of the GHCND stations used to validate downscaled high-return period rainfall values obtained from TMPA data.

Dataset	$\epsilon[Km]$	α
Micronet stations network	13.75	0.13
TMPA 3b42 pixels	53.14	0.28
Downscaling from TMPA, α and ϵ	26.50	0.23
Downscaling from TMPA, only ϵ	34.79	0.28

Table 4.2: Parameters ϵ (scale) and α (shape) of the EK model describing the spatial correlation structure of daily rainfall.

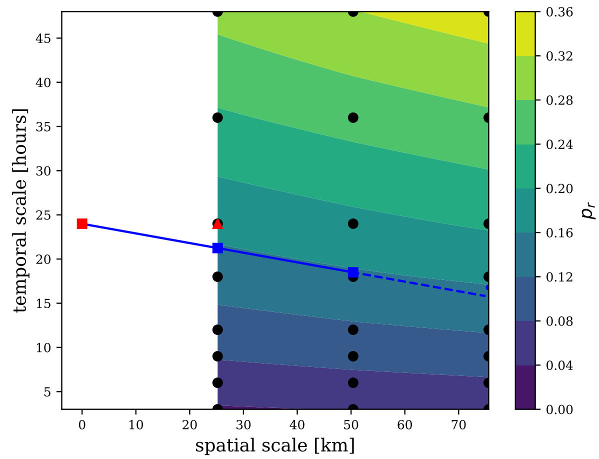


Figure 4.1: Extrapolation of the wet fraction at different integration scales. The black circles represent scales in space and time at which TMPA 3b42 precipitation estimates are aggregated. The density plot shows interpolated p_r values within the range of scales covered by the TMPA product. Blue squares represent the points used to compute the local advection velocity (slope of the blue line), and the red square shows the target scale at which extrapolation of p_r is performed (rain gauge measurement). The red triangle shows the time-space scale of the TMPA pixel-average time series, used to compute the β_0 ratio.

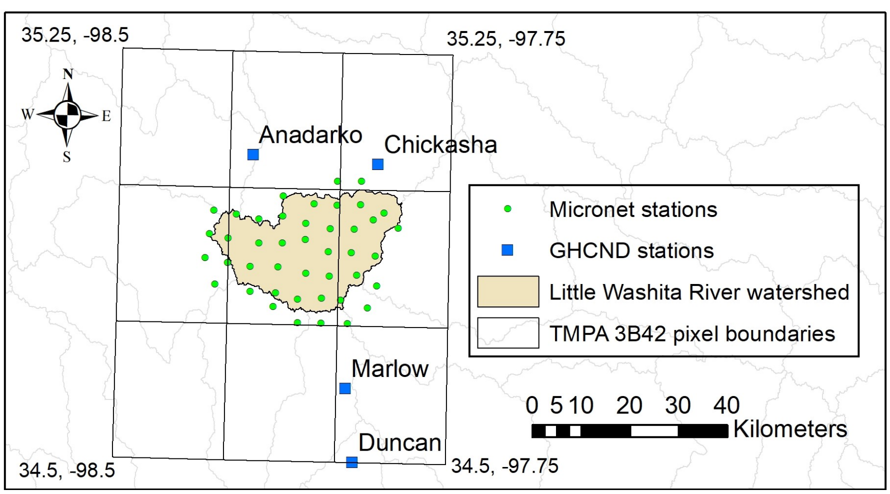


Figure 4.2: Schematic map of TMPA pixels and of the rain gauge network used in this study.

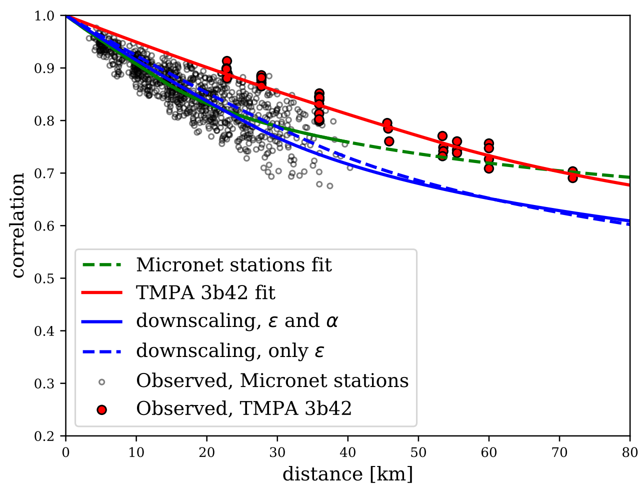


Figure 4.3: Comparison of rainfall correlation function in space as computed from the rain gauges in the Micronet network (open circles) and TMPA gridded precipitation (red cirles). Fit of the EK model is shown by the dashed black and red line for rain gauges and TMPA correlations, respectively. Result of the downscaling scheme are also indicated (blue line). The result obtained by only downscaling the scale parameter while is also reported for comparison (blue dashed line).

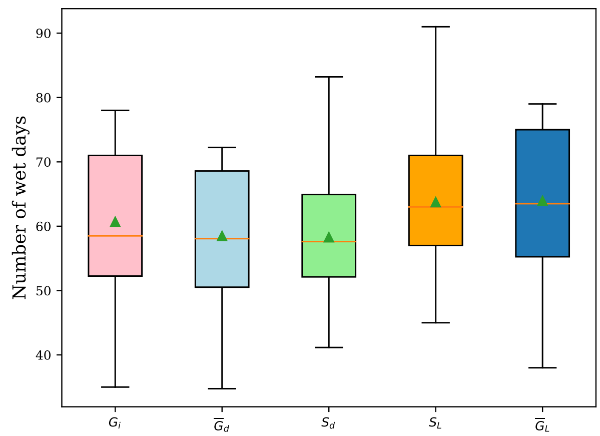


Figure 4.4: The results of downscaling for the wet fraction p_r. The distribution of the yearly number of wet days observed for the Micronet stations (pink, G_i), for their pixel average (blue, $\bar{G_L}$), and for the TMPA time series at the same location (orange, S_L). The downscaled values at a point obtained from the gauge average $\bar{G_d}$ and from the TMPA data S_d are included in cyan and green respectively. The boxplots report the mean (green triangles), and median (orange line) of each distribution. The bars extend from the lower to the upper distributional quartiles, and the whiskers represent the range of each sample.

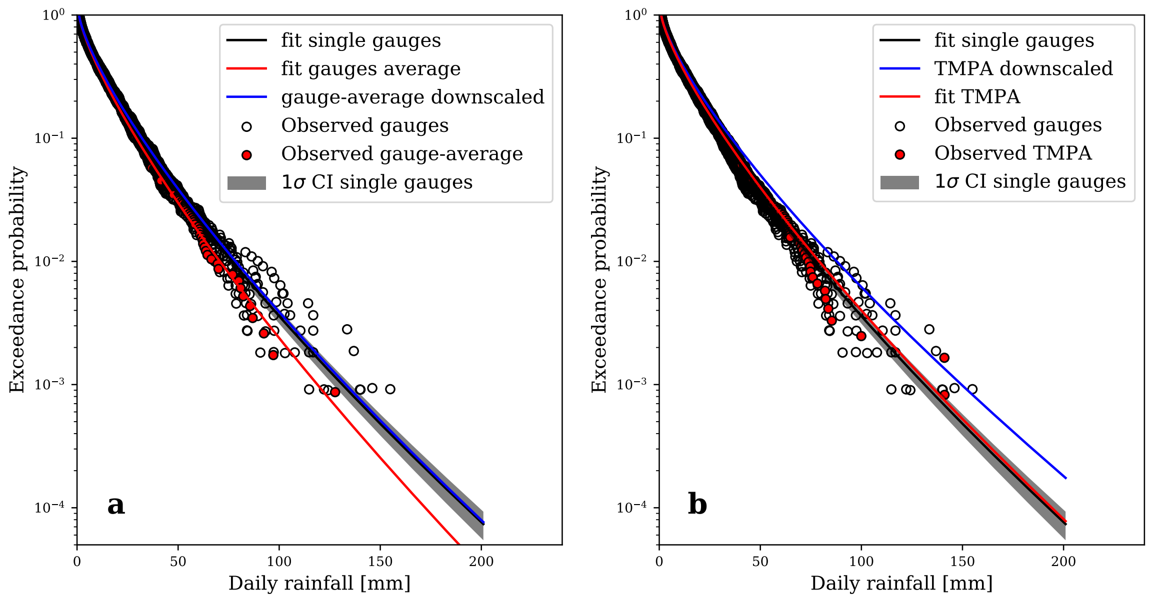


Figure 4.5: (a) Exceedance probability distributions observed for single Micronet stations with at least 19 years of data (open circles) and from their pixel average (red circles). The Weibull fits are reported in red for the average, and black for the single station (mean and 1σ standard deviation confidence interval). The blue line shows the distribution obtained by downscaling from the pixel scale to a point. (b) Comparison of the same exceedance probability distribution for Micronet stations (open circles) and satellite observations (red circles), and the respective Weibull fit (red line and black area respectively). The blue line represent the probability distribution downscaled from the satellite dataset.

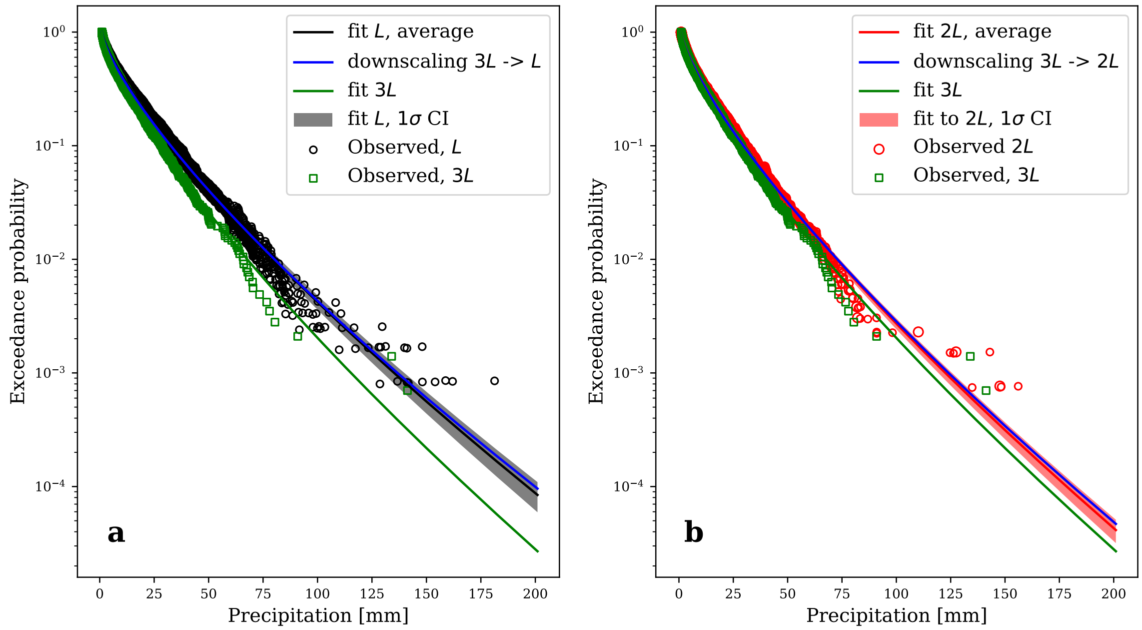


Figure 4.6: (a) Exceedance probability distributions observed for TMPA time series aggregated spatially at different scales. (a) results for the downscaling from scale $3L$ to scale L, and (b) downscaling from scale $3L$ to $2L$.

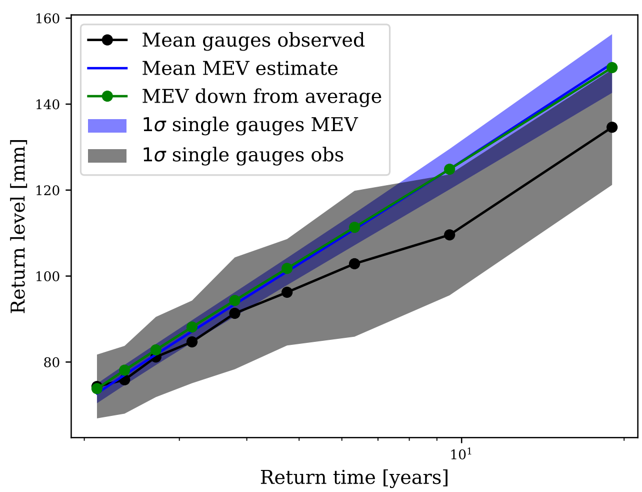


Figure 4.7: Extreme values observed and estimated for the Micronet stations within the TMPA pixel centered in $34.785N, -98.125E$. Observed annual maxima from the Micronet stations with 19 years of data (black, mean and 1σ confidence interval), quantiles estimated from MEVD fitted directly to time series at a point (blue, mean and 1σ confidence intervals), and from downscaling Weibull parameter values from the pixel-average time series (green circles).

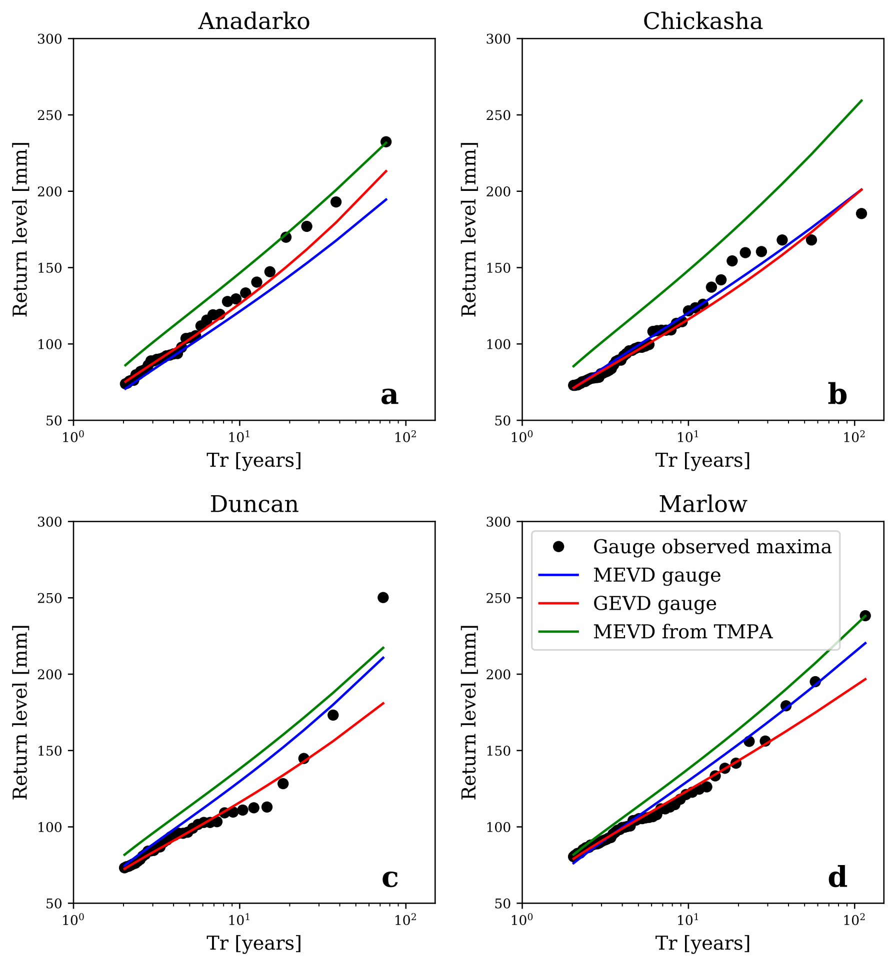


Figure 4.8: Extreme quantiles observed for the 4 GHCN stations in the vicinity of the Micronet network (black markers). The corresponding estimated quantiles have been obtained from directly fitting the extreme value models GEVD (red lines) and MEVD (blue lines) to the entire observed time series. Green lines show the corresponding MEVD quantiles obtained by fitting the Weibull distribution to corresponding TMPA pixel time series and by downscaling the yearly Weibull parameters in the MEVD expression to the point scale.

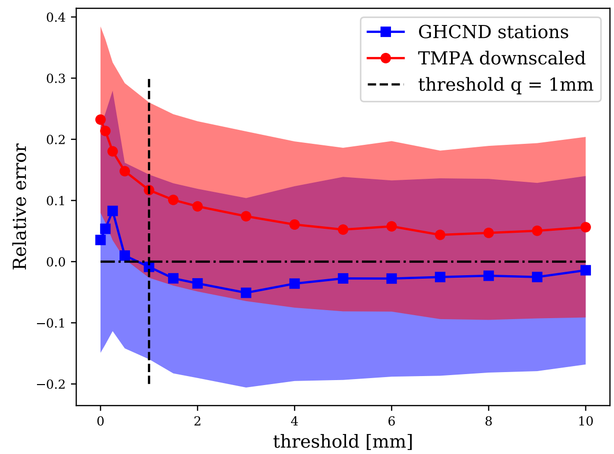


Figure 4.9: Effect of the detection threshold used to define ordinary rainfall values on extreme value estimates for a return times of $Tr = 15$ years. The relative error is reported for TMPA-downscaled estimates (red) and GHCND stations (blue). Solid lines and shaded areas refer to average and 1σ intervals over the 4 stations.

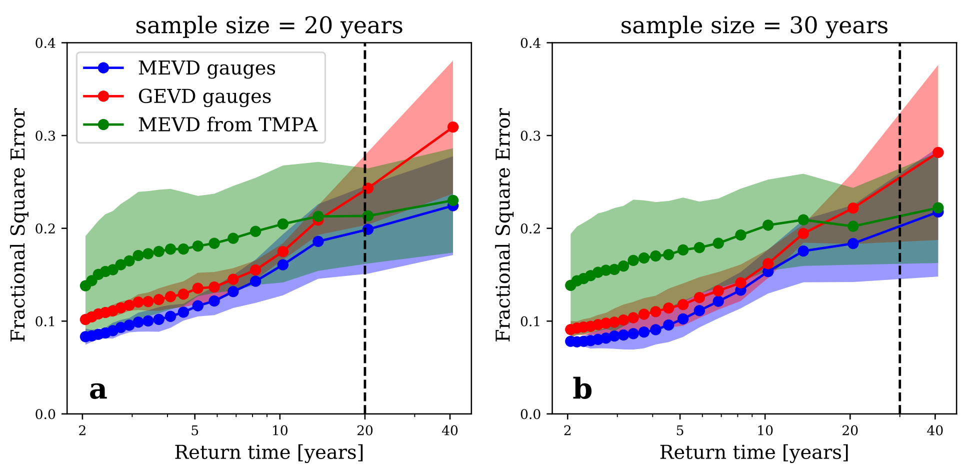


Figure 4.10: RMSE obtained from the cross-validation test using samples of 20 years (a) and 30 years (b) in fitting MEVD (blue) and GEVD (red) to rain gauge GHCND data. The results from the application of the MEVD distribution calibrated from TMPA observations and downscaled at a point is reported in green for both panels. Lines and shaded areas depict averages and standard deviations respectively, computed by repeating the procedure for the set of GHCND stations. Vertical dashed lines indicate the value of the return time corresponding to the sample size used for calibration of MEVD and GEVD with rain gauge data.

Chapter 5

Extreme Value Analysis of Remotely-Sensed Rainfall in Ungauged Areas: Spatial Downscaling and Error Modelling

Adapted from: Zorzetto, Enrico, and Marco Marani. "*Extreme value metastatistical analysis of remotely sensed rainfall in ungauged areas: Spatial downscaling and error modelling.*" Advances in Water Resources 135 (2020): 103483.

5.1 Introduction

Quantitative Precipitation Estimates (QPE) from satellite-borne sensors provide much needed information on the water cycle at the global scale, and are an essential source of observations over large areas worldwide, particularly where the density of rain gauge stations at the ground is low [KBH+16, CW18]. However, while current algorithms try to optimally merge information from radar, passive microwave and infrared sensors [HBN+07, HBB+14], uncertainty in the rainfall retrievals from these instrumental sources inevitably propagates to the final multi-sensor rainfall QPEs. The error structure of the resulting QPEs is thus difficult to characterize, as it depends on numerous variables, including the specific sensor, the precipitation type, and surface characteristics [HBN+07, THA+13, MSA+14]. For this reason, considerable effort has been devoted to the ground validation of satellite QPEs, based on the available observations at the ground. Among many others, see [NSM$^+$03, SHL08, SXWX10, Vil10, CHG$^+$13, BS15, MPBH$^+$17, MKT$^+$18]. However, the density of observational networks at the ground varies significantly around the world, so that a spatially consistent characterization of the QPE error structure is to date very challenging over extended regions. Ideally, the validation of satellite QPEs should be performed by comparing the estimated rainfall rates or rainfall accumulation volumes with a ground-truth dataset aggregated at the same integral domain in space and time [PMF$^+$18]. However, this is in practice only possible with the use of exceptionally dense rain gauge networks, of which a limited number exist worlwide (e.g., see [ESC$^+$93, VK07, FKF$^+$17, PJCM10, PBAM13, AUG$^+$12, VMKM08, WW10, DWB15]) or by comparing QPEs with ground radar estimates where these are available [TB10, KHG$^+$12, AMNB12, KHG$^+$13, MMP$^+$17]. These requirements severely hinder the QPE validation effort over complex terrain and poorly gauged areas. The problem of validating QPE rainfall statistics is particularly challenging when the frequency of heavy rainfall is the variable of interest [HHA09, PMPA16]. While the global-scale coverage and fine temporal scale of satellite QPEs make them invaluable datasets for studying extreme rainfall at the global scale (see e.g., [ZCL$^+$06]), the inference on rainfall distributional tail properties is challenging due to the short length of homogenous records, and to the observational uncertainty in satellite retrievals which inevitably produces large uncertainty in the estimated extreme rainfall quantiles. Recently, the use of satellite QPE datasets for rainfall frequency analysis over ungauged regions, and particularly for the estimation of intensity-duration-frequency curves, has received increasing attention [GHS17, ONSH18, FYH$^+$18, MNA$^+$19]. The main objective of this work is to contribute to bridging this gap by coupling a recently introduced downscaling technique [ZM19] with a model of the error for QPE-derived rainfall statistics. The key steps performed in this combined technique are i) evaluation of key rainfall properties (probability density function (pdf), spatial correlation structure, intermittency and frequency of extremes) from a gridded precipitation dataset, ii) downscaling these quantities so as to make a comparison possible with point-measurements at the ground where these are available, and iii) developing a model of the error aimed at extrapolating the QPE error structure to locations where ground measurements are not available. A key feature of the approach developed here is that it is specifically targeted towards the study of extreme rainfall frequency, with important consquences for the use of satellite QPEs for hydrological analysis and for studying the risk of natural hazards. To this end, I employ the *Metastatistical Extreme Value* distribution (MEVD) [MI15, ZBM16, MNAM18], an extreme value model which provides a connection between the pdf of rainfall accumulations and the frequency of extremes. Global-scale applications of statistical value models to multi satellite QPE datasets have been performed before. For example, Zhou et al. [ZLH15] apply the Generalized Extreme Value distribution in order to map the average recurrence intervals of precipitation for real-time precipitation monitoring; Demirdjian et al. [DZH18] apply a recursive clustering algorithm which, combined with the Peak Over Threshold Method (e.g., [Col01]) reduces estimation uncertainty when applied to short, remotely-sensed datasets. Here I produce the first global-scale maps of extreme value quantiles estimated with MEVD, which i) reduces estimation uncertainty from short rainfall datasets, as shown in [ZBM16] and ii) by providing a link between extreme and non-extreme rainfall statistics, introduces a new way to study the effects of spatial averaging and QPE bias correction on the statistics of rainfall extremes. Overall, the combined method of analysis developed here provides a novel framework for understanding how biases in QPE statistics propagate to rainfall frequency analysis, and for designing suitable site-specific corrections. This last point has significant impact for ungauged areas worldwide where the density of stations (or the length of the available records) is not sufficient for obtaining reliable extreme value estimates with traditional statistical techniques. The statistical framework outlined above is here developed and applied to a large domain covering the entire Conterminous United States (CONUS). This domain is covered by a large number of gauges and ground-radar derived precipitation products, so that the performance of satellite sensors and QPE datasets have been extensively studied here (e.g., [ABS+11, CHG+13, PN15]). This makes it the ideal domain for testing my methodology, which does not make use of any ground observation for the downscaling, and only requires sparse gauge observations, or radar data, for constructing the model for the error. Therefore, ground rain-gauge observations and comparison with previous studies can be used for independently testing the performance of the method and for validating my findings. Moreover, the CONUS domain spans a broad range of climatic regimes and terrain types, thus allowing me to extend my previous work [ZM19] and explore the variability of QPE rainfall statistics and the associated error structure over this broader domain. Here I focus my attention on the Tropical Rainfall Measuring Mission (TRMM) Multisatellite Precipitation Analysis (TMPA) 3B42 research version dataset, which to date includes 20 years of continous data and, even after the end of the TRMM mission, is a well-tested source of information for hydrological applications. However, I note that the approach developed here is not product-dependent, and is suitable for a straightforward application to any gridded precipitation product, such as the Integrated Multi-Satellite Retrievals for GPM (IMERG) product [HBB+14], or even reanalysis products and climate model outputs. The Chapter is organized as follows: In Section 5.2 I present the statistical technique used to downscale rainfall statistics from gridded QPE datasets, so as to enable a direct comparison with point gauge measurements. Building on this method, in Section 5.3 I develop a non parametric model for the error based on a quantile regression forest algorithm [Mei06]. The error model is then tested using a cross validation scheme so as to effectively simulate the prediction of QPE rainfall statistics over ungauged areas. In Section 5.4 I then present the statistical framework aimed at the large-scale extreme rainfall value analysis based on gridded QPE precipitation products, and discuss the current limitations in evaluating the results over poorly instrumented areas. The extreme value estimates are then corrected and tested using the model of the error developed in Section 5.3. Results of these steps are presented throughout these sections, and are followed by a discussion of my findings and by the conclusions drawn by this study (Section 5.5). A list of acronyms and abbreviations used throughout the Chapter is featured in Appendix D.

5.2 Spatial Downscaling of the Probability Distribution of QPE Magnitudes

I start by addressing the scale gap between daily rainfall statistics estimated from TMPA gridded QPE fields and their counterpart derived from rain gauge point measurements. For this purpose, I employ a recently developed statistical downscaling technique [ZM19] based on the theory of random fields [Van10, MT13]. While this technique entails a number of assumptions, it greatly extends the areas where validation of gridded QPEs datasets is possible as it does not require the presence of a dense network of rainfall gauges or ground radars covering the location of interest for downscaling QPE statistics. The main assumptions made in deriving this methodology are: i) Local spatial homogeneity and isotropy of the rainfall field, ii) a Taylor frozen-turbulence hypothesis for obtaining the wet fraction at a point in space, and iii) the assumption that the distribution of daily rainfall events can be represented as a stretched exponential both at the point scale and at the grid cell scale (albeit with varying parameters). In the following I briefly summarize the downscaling method applied here; for additional information on the derivation of the methods, the reader is referred to [ZM19]. I define as *ordinary rainfall events* all daily rainfall accumulations at a given temporal integration scale in excess of a fixed low threshold value, here set to $q = 1$ mm/day. I choose a fixed threshold to define ordinary daily rainfall events which is larger than both the detection limits of gauges and QPE data, when aggregated at the daily timescale, while at the same time low enough so as to include the bulk of the daily rainfall distribution in my analysis. I denote with N_s the yearly number of such events, or *wet days*. The wet fraction is thus defined as as $p_{r_s} = N_s/N_t$, with $N_t = 366$ the number of daily observations in each year. The suffix s here indicates the linear spatial scale S at which a rainfall time series is averaged. Here, $s = 0$ indicates the *point scale* L_0 corresponding to rain gauge measurements (the characteristic size of a rain gauge is of the order of $L_0 \simeq 10^{-4}$ km), and $s = L$ corresponds to the linear characteristic scale of the gridded QPE to be downscaled ($L = \sqrt{L_x L_y}$, where L_x and L_y are the dimension of a QPE grid cell along the zonal and meridianal directions respectively). Since I am dealing with spatial domains which are well characterized by a single length scale, here I use a single linear length scale S to characterize a random field averaged over an area S^2. I assume that in each year the ordinary rainfall event magnitudes observed at scale s, $h_s = \widetilde{h}_s - q$ (defined for daily rainfall totals $\widetilde{h}_s \geq q$), are realizations of a random variable H_s with population $\Omega_{H_s} = (0, +\infty)$ and marginal distribution described by a *Weibull* or stretched exponential distribution, defined by the cumulative probability function

$$P(H_s < h_s) = F(h_s) = 1 - e^{-\left(\frac{h_s}{C_s}\right)^{w_s}} \tag{5.1}$$

here C_s is a scale parameter with the same units of h_s (mm/day here) and w_s is the dimensionless shape parameter of the distribution. A recent study analyzed the tail properties of hourly rainfall accumulations over the CONUS [PAFG18], and found that Weibull provides a satisfactory description of hourly rainfall observations, complementing and reinforcing previous evidence for rain-gauge daily accumulations at the global scale [WT05]. While the Weibull parameter C_s represents the characteristic magnitude of daily rainfall events, w_s describes the decay of the tail of $F(h_s)$, such that values of $w_s < 1$ correspond to a sub exponential behavior (i.e., the exceedance probability exhibits a heavy tail, albeit with a characteristic scale). An exponential decay is recovered for $w_s = 1$, and values of w_s larger than unity correspond to faster-than-exponential decays. This simple parametric model is assumed to describe the shape of the daily rainfall distribution of both rain gauge and satellite-derived rainfall accumulations, such that their respective difference is encoded in differences between the values of C_s and w_s at the different spatial scales.

5.2.1 Downscaling scheme

Zorzetto and Marani [ZM19] showed that the parameters of the distribution of ordinary rainfall events in eq. (5.1), aggregated at a fixed temporal scale and averaged at two different spatial scales can be linked by two equations which depend on the intermittency and spatial correlation of the rainfall field. For example, the stretched exponential parameters C_L, w_L describing the ordinary rainfall pdf at the grid cell scale ($s = L$) of the QPE dataset can be expressed in terms of the parameters C_0 and w_0 of the rainfall field at the point scale ($s = 0$) by the relations

$$\gamma_0 \beta_0 \frac{2 w_0 \Gamma\left(\frac{2}{w_0}\right)}{\Gamma^2\left(\frac{1}{w_0}\right)} = 2 w_l \frac{\Gamma\left(\frac{2}{w_L}\right)}{\Gamma^2\left(\frac{1}{w_L}\right)} + (\gamma_0 - 1) p_{r_L} \tag{5.2}$$

$$\left(\frac{C_0}{w_0}\right)^2 = \beta_0^2 \left(\frac{C_L}{w_L}\right)^2 \frac{\Gamma^2\left(\frac{1}{w_L}\right)}{\Gamma^2\left(\frac{1}{w_0}\right)} \tag{5.3}$$

where Γ denotes the Gamma function, p_{r_L} is the wet fraction of the ordinary rainfall process at the grid cell scale L, $\gamma_0 = \gamma_0(L)$ is the *variance reduction factor* i.e., the ratio of the variances σ_L^2 and σ_0^2 of the process averaged at the two scales $S = L$ and $S = L_0$ respectively, and $\beta_0 = \beta_0(L) = p_{r_L}/p_{r_0}$, or *intermittency function*, is the ratio of the wet fraction at the two scales considered here. While p_{r_L} can be directly computed from QPE time series, γ_0 and β_0 are not known in the absence of rain gauge measurements, since they do not only depend on the areal average process but also on the rainfall process at scale L_0, i.e., at '*a point in space*'.

5.2.2 Scalewise variation of the spatial correlation function

However, the variance reduction function γ_0 can be obtained from the spatial correlation of the rainfall field $\rho\left(s_1, s_2\right)$, here assumed to be quadrant-symmetric [Van10], where s_1 and s_2 are distances between two points measured along two coordinate axes, as

$$\gamma_0 = \frac{\sigma_L^2}{\sigma_0^2} = \frac{4}{L_x^2 L_y^2} \int_0^{L_x} \int_0^{L_y} \left(L_x - s_1\right)\left(L_y - s_2\right) \rho\left(s_1, s_2\right) ds_1 ds_2 \tag{5.4}$$

where L_x and L_y are the spatial dimensions of the grid cell size. Ideally, the point correlation $\rho\left(s_1, s_2\right)$ should be estimated using a sufficient number of rain gauges distributed in space. As these are not available everywhere (in fact, they are lacking in most areas) here I estimate it using the correlation between QPEs time series sampled at grid cells within a 3 x 3 local neighborhood. This correlation between spatially averaged values can be linked to the unknown correlation at the point scale by means of the following relation [Van10]

$$\rho_{h_L,h_{L'}} = \frac{\sum_{k=0}^{3}\sum_{l=0}^{3}(-1)^k(-1)^l\Delta\left(L_{x,k},L_{y,l}\right)}{4\Delta\left(L_x,L_y\right)} \tag{5.5}$$

where the set of distances $L_{x,k}$ and $L_{y,l}$ for $k,l = 0,1,2,3$ contains the information on the relative positions of pairs of cells within the neighborhood considered. Letting Δx and Δy be the distances between the two grid cells along the s_1 and s_2 coordinate directions, then these distances are defined as

- $L_{x,0} = \Delta x - L_x,\ L_{y,0} = \Delta y - L_y$
- $L_{x,1} = \Delta x,\ L_{y,1} = \Delta y$
- $L_{x,2} = \Delta x + L_x,\ L_{y,2} = \Delta y + L_y$
- $L_{x,3} = \Delta x,\ L_{y,3} = \Delta y$

and with the function $\Delta\left(a,b\right)$ defined as

$$\Delta\left(a,b\right) = 4\int_0^a\int_0^b\left(a-s_1\right)\left(b-s_2\right)\rho(s_1,s_2)ds_1,ds_2 \tag{5.6}$$

Since the QPE dataset allows the computation of estimates of $\rho_{h_L,h_{L'}}$, I can obtain an estimate of the point scale correlation by assuming a parametric form for the spatial correlation function $\rho(s_1,s_2)$ and by numerically minimizing the Sum of Squared Errors (SSE) between its areal average value at the grid cell scale given by eq. (5.5) and the values estimated from QPEs. I assume here an isotropic correlation function, such that $\rho(s_1,s_2) = \rho(\sqrt{{s_1}^2+{s_2}^2}) = \rho(d)$

$$\rho(d;\epsilon,\alpha) = \begin{cases} e^{-\frac{\alpha d}{\epsilon}} & d<\epsilon \\ \left(\frac{\epsilon}{ed}\right)^{\alpha} & d\geq\epsilon \end{cases} \tag{5.7}$$

This formulation has an exponential kernel and a power-law tail characterized by the exponent α; the smooth transition between the two regimes occur at a distance $d = \epsilon$ [Mar03, ZM19]. Depending on the values of the parameters α and ϵ it allows for the description of both light- and heavy-tailed correlations. With this assumption, the SSE becomes a function of the parameters α and ϵ:

$$SSE\left(\epsilon, \alpha\right) = \sum_{j=1}^{m} \left[\rho_{h_L, h_{L'}}\left(d_j; \epsilon, \alpha\right) - \rho_j\right]^2 \tag{5.8}$$

where the ρ_j are correlation estimates obtained from the QPE dataset, using a local lattice of 3x3 grid cells centered over the location of interest. Using a larger lattice would increase the number of points available to estimate the correlation function, but, on the other hand, would increase the likelihood of including in the analysis non-homogeneous rainfall statistics, which certaintly are present at the regional scale, especially over a complex terrain. For each pair of grid cells within this local domain, the Pearson correlation was computed between the respective QPE time series, and the resulting values binned over a set m of distances d_j. Here I minimize the SSE in eq. (5.8) by means of the *differential evolution* stochastic minimization algorithm [SP97], as opposed to the deterministic algorithm used in [ZM19]. The differential evolution minimization is particularly suited to avoid possible local minima in the (α, ϵ) parametric space, and thus more robust in minimizing a function which depends upon the experimental points ρ_j. I seek the global minimum of eq. (5.8) within the rectangular domain $\alpha \in (0, 1)$ and $\epsilon \in [0, 1000]$ km, which are physically meaningful ranges of values. While quantifying the uncertainty in the downscaled correlation parameters is a challenging task, as the shape of objective function eq. (5.8) varies with the observed values ρ_j, I note that the downscaling methods only requires the functional of the correlation function defined in eq. (5.4), which is an integral property of the correlation function over the pixel size. Therefore, I expect γ_0 to be less sensitive to observational uncertainty than the single parameter values α and ϵ.

5.2.3 Scalewise variation of the wet fraction

To obtain an estimate of the intermittency function $\beta_0(L)$, i.e. of the ratio between the yearly number of events for the ordinary rainfall process averaged at the grid cell scale L_0 with respect to the point scale, I rely on an application of the *Taylor frozen turbulence hypothesis* [Tay38, Dei00, HEM+15] introduced by [ZM19]. This approximation enables me to use of information from the QPE dataset at smaller temporal scales (up to 3-hours for the TMPA dataset) to infer a property of the rainfall field, namely the wet fraction, at spatial scales smaller than the grid cell scale and at the daily timescale. I aggregate TMPA rainfall fields at increasing spatial scales S and temporal scales T (integrating in time and averaging in space), and at each space-time scale the wet fraction $p_r(S,T)$ is computed as the fraction above q of the time series at scale (S,T). I then extrapolate this quantity to the point scale L_0 by assuming that the function $p_r(S,T)$ is locally linear in the (S,T) plane; i.e., that integration in time and averaging in space have the same effect on a property of the rainfall field (the wet fraction here) up to a constant factor which has the meaning of a local advection velocity. This assumption holds exactly only in the case of a perfectly fractal rainfall field [HEM+15] but, even though only approximate in general, it offers a conceptually satisfactory way to infer the point-scale wet fraction using only QPE data at the grid-cell scale. Through this approach I obtain $\beta_0 = p_{r_L}/p_{r_0}$, where p_{r_L} is simply obtained from the remote-sensing QPE, and p_{r_0} is the value that a linear extrapolation in the (S,T) plane yields for daily rainfall at the rain gauge spatial scale L_0. Note that in order to apply this technique I need a sufficiently fine-scale resolution in time for the QPE dataset compared to the scale at which the downscaling analysis is performed (3-hrs versus daily here). The reader is referred to [ZM19] for a more detailed description and application of the methodology.

5.2.4 Data

An independent evaluation of the statistical structure of a QPE dataset using the approach described above requires observations from rain gauges at the ground. I focus on the CONUS domain, defined here as the domain over land within 22°N - 50°N and -130°W - -60°W. Over this domain I use observations from the network of hourly precipitation data (HPD) rain gauges of the National Oceanogaphic and Atmospheric Administration (NOAA), which covers the CONUS starting from 1948 [NCE]. Precipitation data were aggregated at the daily time scale by summing hourly accumulations in each day of the record, defined starting from midnight. I excluded from the analysis days with quality-flagged or missing precipitation amounts. The remotely-sensed gridded precipitation product used in this study is *TMPA 3b42 version 7, research version* [HBN$^+$07, HABN10, HB13], of which I use the entire record 1998-2018 so as to obtain the longest possible dataset for extreme value analysis. Note that starting in 2014, after the end of the TRMM era, rainfall retrievals from the GPM mission are used in TMPA estimates, introducing some heterogeneity in the dataset. Rainfall rates obtained from the TMPA dataset are here assumed to represent the average rainfall rate for each 3-hr timestep. These values were aggregated so as to compute rainfall accumulations at time scales raging from $3-$hrs to $48-$hours, as needed for estimating the intermittency of the rainfall field as discussed in Section 5.2.3, with the $24-$hours totals used for my analysis at the daily time scale. I note that discrepancies between gauge and TMPA daily totals may potentially arise in the presence of a pronounced daily precipitation cycle, and as a consequence of the instantaneous nature as well as the timing of satellite rainfall retrievals [LSLC16]. Here I do not anticipate that this issue will play a relevant role in my analysis, since I do not study the timing of specific events but only average statistics. The spatial resolution of the data is 0.25° x 0.25°, corresponding to a characteristic grid cell size of about 25km over the CONUS domain. While here I interpret the TMPA rainfall rates as areal averages over the grid cell area (see, e.g., [VK07]), this is an approximation, as these estimates are obtained by merging retrievals from different instrumental sources (passive microwaves and infrared), with different footprints. Additionally, the monthly-scale gauge correction is performed using precipitation gridded at scales larger than the spatial reolution of TMPA, and therefore introduces larger-scale information in the pixel-scale QPE time series. In essence, any discrepancy stemming from this assumption will be included in the structure of the error that will be estimated as a result of my analysis.

In order to evaluate QPE statistics, after downscaling, I select a set of QPE grid cells from the TMPA dataset that are characterized by the presence of (i) at least one rain gauge within the grid cell with at least a 10-year record of daily rainfall, and (ii) at least 4 rain gauges in a 5x5 pixel neighborhood around the location of interest, which are used for producing an estimate of the local spatial correlation of the rainfall field. This information is not used to train the downscaling method, but only for validation purposes. To be selected, these rain gauge records must overlap for a temporal window of at least 2000 observations (i.e., about $\simeq$ 7 years of record). This condition was chosen as a tradeoff between having enough data for reliably estimating the correlation between sites, and including a large enough number of sites for testing purposes. My results are not very sensitive to this choice, as for most of the gauged sites in the dataset the record length is significantly longer than 7 years. In the following I will refer to the set of sites meeting these conditions as the *Set of Gauged Sites*, or SGS.

5.2.5 Results of the downscaling method

To study the spatial correlation of the rainfall field, I focus on two metrics: the ratio between the two parameters of the correlation function ϵ/α, and the variance reduction function γ_0. The first is a characteristic spatial scale describing how the correlation decays with distance (for very large values of ϵ this is exactly the spatial integral scale of the field, which is not necessarily finite depending on the values of α), while γ_0 is an integral property of the correlation function between the point and the grid-cell scale. γ_0 is the main quantity of interest here as it connects the variance of the rainfall accumulations at different spatial averaging scales, and it is the only quantity dependent on the correlation function that is directly needed in downscaling the pdf of ordinary rainfall in eqns. (5.2) and (5.3). I note that three different cases can be identified when downscaling the spatial correlation function from gridded QPEs. When the correlation scale of the rainfall field at the ground is small compared to the spatial averaging scale ($\epsilon/\alpha L < 1$), then the correlation structure of the continuous process is in large part hidden by the averaging process and I expect it cannot be completely recoverable from the QPE dataset alone. This occurs only in a very limited subset of the SGS sites examined here. I find that the ratio ϵ/α is smaller than 25km only at 33 of the 860 SGS sites where downscaling is performed (3.8% of all cases). These are primarily locations characterized by a complex terrain, and it is reasonable to think that in these cases the correlation at the point scale is not retrievable from QPE alone. I therefore exclude these locations from the following analysis. Conversely, when the correlation of the field is much larger than the grid scale over which averaging is performed, most of the information about the correlation function is preserved even after averaging. In this case I argue that the effect of, and need for, a downscaling process may be limited because the variables being averaged are highly correlated, and thus they tend to behave very similarly to their spatil average. This case also does not occur frequently, with only 8 out of the 860 sites examined here exhibiting ratios ϵ/α larger than 200km, the largest ratio being around 8. In the intermediate case between these two situations, when the correlation distance of the rainfall field is comparable to or larger than the scale of averaging, a substantial part of the information about the correlation structure of the continuous process at the point is preserved by the averaging process. The application of the downscaling approach, i.e. the minimization of eq. ((5.8)), can produce in this case good estimates of the correlation structure at the point scale. To provide a global metric describing the effect of correlation downscaling, I compare the estimated values of ϵ/α and γ_0 with the corresponding metric $\epsilon^{(L)}/\alpha^{(L)}$ obtained from fitting eq. (5.7) to the empirical correlation estimates $(\rho_j,\, d_j)$ obtained from a local lattice of QPE time series. In turn, using $\epsilon^{(L)}$ and $\alpha^{(L)}$ instead of ϵ and α in eq. (5.4) I can infer an estimate, $\gamma_{0,L}$, of the variance reduction function value that one would obtain by using the grid-cell scale correlation function instead of its downscaled version. Examination of the values of rainfall spatial correlation over the CONUS domain (here indexed by longitudinal position, see Figure 5.1a) shows that the TMPA-estimated correlation obtained by fitting eq. (5.7) to the gridded QPEs is larger than its estimates from point measurements at the ground. This is as expected, since I are comparing a correlation between averaged values with its point-scale counterpart. When I compare the downscaled correlation values obtained from the QPE dataset by minimizing eq. (5.8), I see that some underestimation occurs primarily at the boundaries of the domain, and chiefly for grid cells located in the proximity of the west coast. This result is consistent with a previous study which investigated the ability of radar rainfall retrievals to reproduce the small-scale rainfall variability [GK04a]. However, in the central part of the CONUS domain, the zonal variability of the correlation appears to be reasonably captured by the TMPA estimates. The application of the downscaling approach produces correlation values that are much closer to those from gauge observations, suggesting that, indeed, the proposed downscaling procedure yields statistical properties that are close to those of point observations. A similar result emerges if I turn my attention to the variance reduction function (Figure 5.1b), for which I again notice how the downscaled results are closer to values from ground observations for the central part of the CONUS, and less so along the east and west coasts.

A summary of the variability of the downscaled parameters over the SGS is provided by the scatter plots in Figure 5.2. While no apparent bias appears for the values of $\gamma_{0,d}$ with respect their counterpart estimated at the ground, a significant variability is detected for the lowest values of γ_0, which correspond to grid cells located primarity in the mountainous part of the Western United States (Figure 5.2a). This scatter for low values of γ_0 confirms the increasing difficulty of correctly estimating the point correlation when its characteristic length scale becomes smaller compared to the averaging length L; However, note that the large majority of the SGS sites have values of γ_0 larger than 0.9, range in which the discrepancy between ground and downscaled values is limited, as shown in panel 5.2a. The comparison of the number of wet days, $N_{0,d}$, obtained from downscaling TMPA QPE's with gauge-estimated values, $N_{0,g}$, indicates that the QPE dataset consistently underestimates the number of events recorded at the ground (Figure 5.2b), and that this underestimation increases with the value of $N_{0,g}$. While the number of events $N_{0,d}$ is sistematically underestimated, the scale parameter $C_{0,d}$ appears to be significantly larger than its corresponding ground values (Figure 5.2c). I argue that this feature of the QPE dataset is at least partially explained by the gauge correction applied to the TMPA

3b42 research version dataset: gauge data are used to rescale TMPA monthly totals, and thus if the number of missed events is appreciable, such a correction will result in a deformation of the pdf of ordinary events with respect to the 'real' one at the ground, and will lead to the overestimation of the characteristic event size $C_{0,d}$ so that monthly totals measured at the ground are preserved. The downscaled shape parameter $w_{0,d}$ appears to be quite variable with respect to its gauge-estimated counterpart $w_{0,g}$ (Figure 5.2d). As was the case for γ_0, even though most of the values here are close to the identity line, for some points corresponding to locations in the Western United States the downscaled values of w_0 are significantly lower than gauge estimates, meaning that for these locations the downscaled ordinary rainfall statistics exhibit a heavier tail, i.e., an overestimated probability of intense events. To study the spatial distribution of the discrepancy between QPE downscaled statistics and the corresponding values at the ground, I define the relative errors

$$\eta_z = \frac{z_{0,d} - z_{0,g}}{z_{0,g}} \tag{5.9}$$

where the variable of interest can be one of the following: $z = C, w, N$ or γ, the subscript 0 again refers to values at the "*point*" spatial scale L_0, and the subscripts d and g refer to the downscaled or rain-gauge observed quantities respectively. Figure 5.3 reports the spatial variability of the relative errors η_z for the four parameters and the set of SGS sites. The map of the relative errors η_γ in the variance reduction function over the CONUS is featured in panel 5.3a. The variance reduction function γ_0 appears to be well captured by the downscaled correlation, especially over the East coast and in the midwest regions of the CONUS. Some overestimation appears corresponding to the West Coast, while γ_0 tends to be generally underestimated with respect to ground values in the Western CONUS. The error in the yearly number of events appears to be larger on the West coast and in the North-East, whereas it is smaller in the South-East and Mid-West regions of the USA (Figure 5.3b). Once the variance reduction function and intermittency functions are known, the parameters describing the pdf of ordinary rainfall events at the point scale can be estimated by means of eqns. (5.2) and (5.3). For the scale parameter C_0, some overestimation is found to occur, especially in the northeastern and northwestern sectors of the CONUS, while mostly underestimation occurs in the west (Figure 5.3c). Conversely, the shape parameter w_0 appears to be underestimated in the west and overestimated in the east (Figure 5.3d). Together, these results suggest that in the West downscaled QPE statistics exhibit heavier tails and lower mean compared to their ground counterparts, while the opposite is true for the Eastern USA and the Pacific coast, where both parameters are overestimated: the characteristic event magnitude is overestimated, but the tail of the distribution is lighter than ground estimates suggest. These results are coherent with previous work [PN15], which found that TMPA 3b42 shows a consistent underestimantion of the yearly number of events, especially in the North East and Middle Atlantic regions of the CONUS. Therefore, the results obtained here for the Weibull scale parameter are also to be expected, as the gauge correction applied to TMPA by rescaling the monthly total accumulations will determine an overestimation of the characteristic event size in the number of events is underestimated. As I will discuss in the following section, these distortions of the pdf of daily rainfall have relevant consequences for the estimation of extremes.

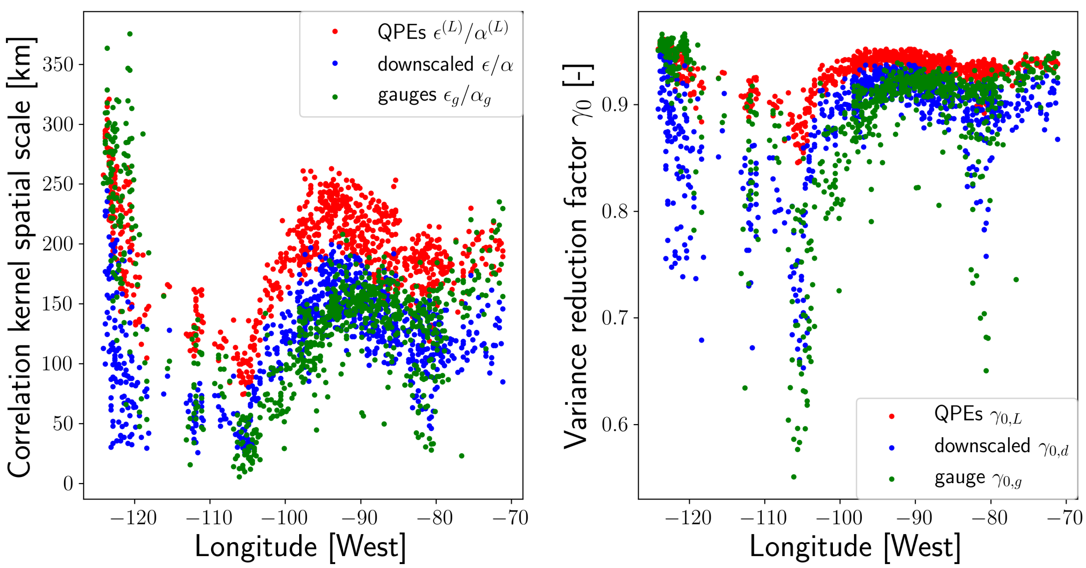


Figure 5.1: Results obtained by downscaling the daily precipitation spatial correlation function from the grid cell to the point scale. (a) Correlation kernel spatial scale, and (b) variance reduction factor γ_0 as a function of Longitude. Red circles denote the values obtained by fitting the correlation function directly to the empirical correlation between TMPA grid cells, blue circles refer to the correlation downscaled to the point scale, and green circles are obtained by estimating the correlation function from rain gauges of the NOAA HPD dataset.

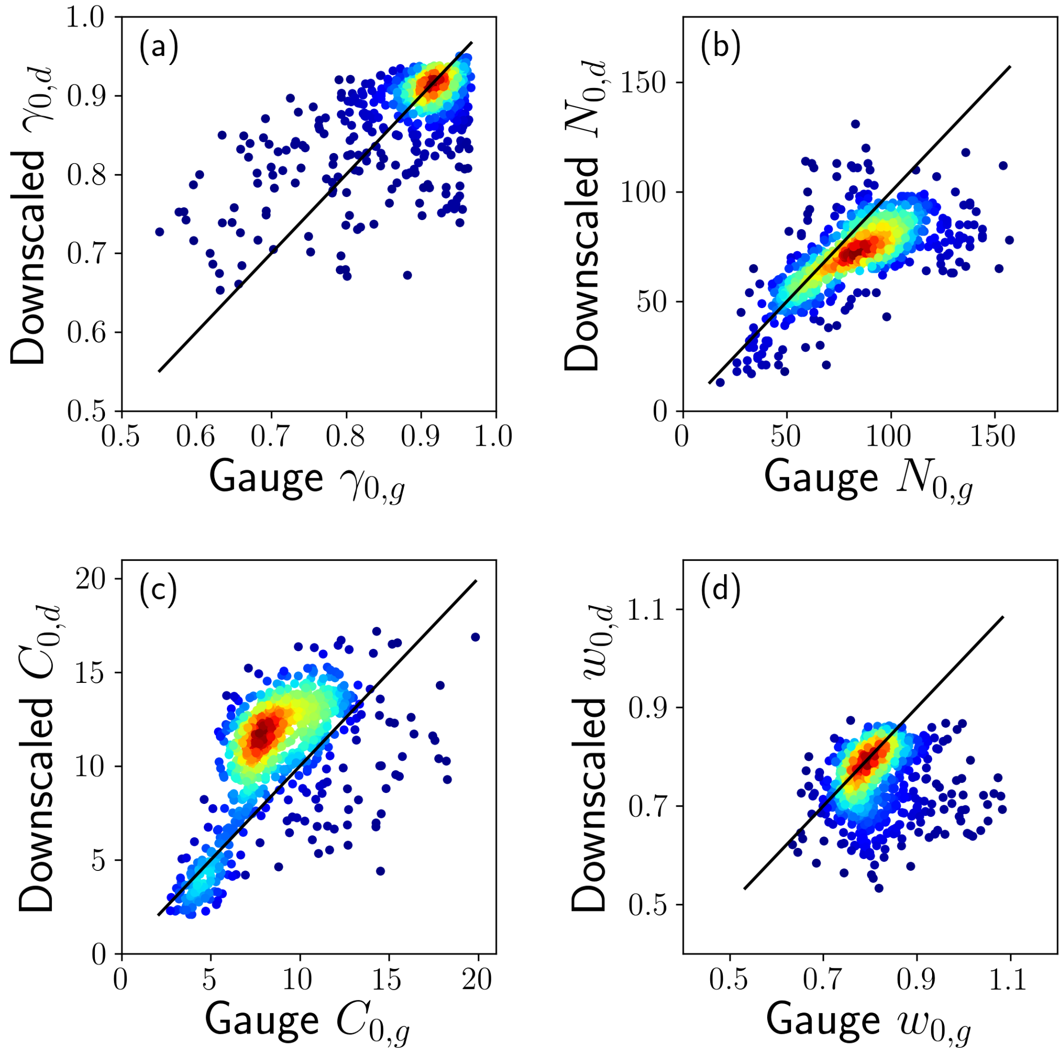


Figure 5.2: Scatter plots of parameter values downscaled from the TMPA dataset vs. the corresponding values estimated from rain gauges at the ground. (a) Variance reduction function γ_0, (b) Average yearly number of events N_0, (c) Weibull scale parameter C_0, and (d) Weibull shape parameter w_0. The color indicates the relative density of points in the graph computed by means of kernel density estimation, with warmer colors indicating higher point density.

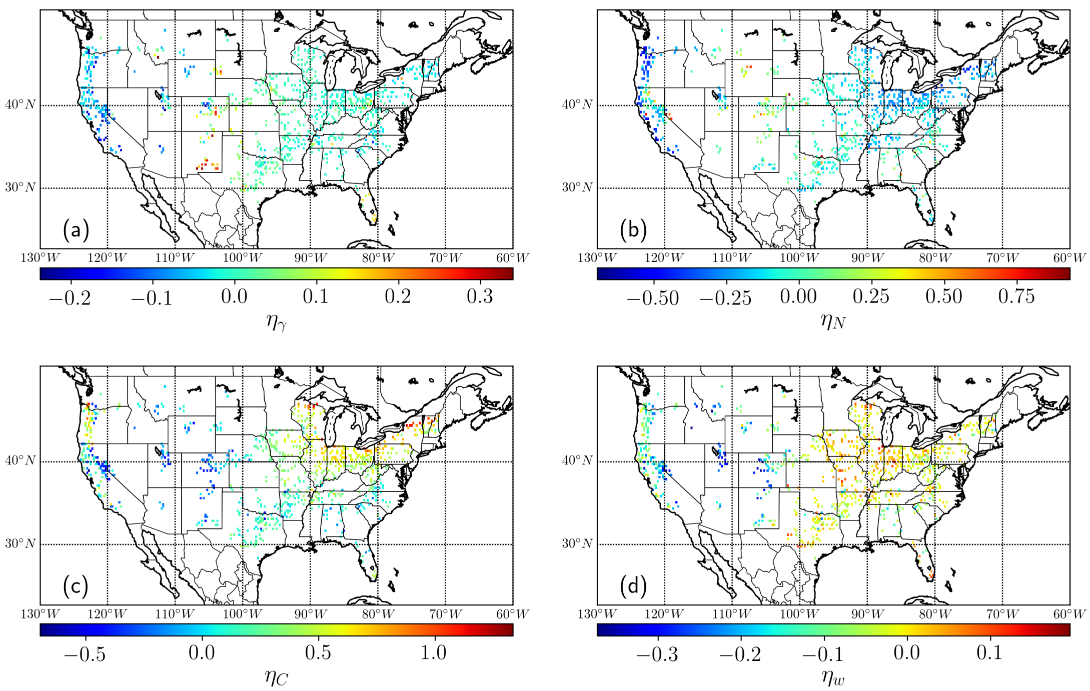


Figure 5.3: Spatial distribution of the relative errors between parameter values downscaled from the TMPA gridded precipitation dataset with respect to their counterparts estimated from rain gauge records at the ground. The figure features in panel (a) the error in the variance reduction function η_γ, in (b) the error in the yearly number of events η_N, and in panels (c) and (d) the errors in the Weibull scale η_C and shape η_w parameters respectively.

5.3 A Model of the Error for QPE Downscaled Statistics

As discussed above, downscaling the pdf of rainfall accumulations allows for the direct validation and correction of remotely sensed QPE's statistics at gauged sites. This result is useful in itself, as downscaling allows for proper comparisons of quantities defined at the same scale, and it reduces the density of the gauge networks required for validating QPE-derived rainfall estimates. However, for many areas worldwide which are characterized by sparse or altogether absent ground stations, it is of primary importance to provide the error analyses that can be applied to ungauged locations. In the case of target sites located at a limited distance from gauged locations, this objective can be pursued by building a geo-spatial model of QPE error statistics. Here I investigate how QPE errors can be predicted from the information collected at gauged sites located at considerable distance from the target sites, so that they are characterized by weak or no direct correlation. To this end, I develop a non-parametric model of the error based on the *quantile regression forest* (QRF) algorithm [Mei06] with the objective of inferring the relative errors in the downscaled parameters ($y = \eta_C, \eta_w$ or η_N) at ungauged locations, based on a set of variables describing the local rainfall regime and terrain type. QRF has been applied before to study the error structure of passive microwave rainfall retrievals [BAK17], and constitutes a modification of the classic *Random Forest* (RF) algorithm introduced by Breiman [Bre01]. This modification of the RF algorithm is chosen for its ability to 1) describe the generally nonlinear relations between the error in the downscaled parameters and a set of explanatory features [BJKK12], 2) limit the problem of overfitting [Bre01, TPL19], and 3) deal with possibly correlated predictor variables [ZK14]. Predictions from these models are based on regression trees, structures which encode the recursive partitioning of the predictor variable space in distinct and non overlapping regions. In each terminal node of a tree, the predicted value of the target variable is estimated as the average response in that region. Predictions of RF and QRF are ensemble values over a large number k of decision trees, each trained on a bootstrap sample from the original dataset. Instead of predicting the expected value of the variable of interest for a given value of the predictor, as is the case for the RF method, QRF estimates the full conditional distribution of the response variable Y given a value of X, a possibly multi-dimensional predictor vector. Reconstructing the entire conditional distribution can be useful here for i) limiting the possible effect of outliers on the estimated conditional mean using the median value instead, and ii) constructing confidence intervals for the estimates of interest. QRF estimates the empirical distribution function of Y for a given value x of the explanatory variable X as [Mei06]

$$P\left(Y < y \mid X = x\right) = \sum_{i=1}^{n} \overline{\omega}_i\left(x\right) 1_{\{y_i \leq y\}} \tag{5.10}$$

where

$$\overline{\omega}_i\left(x\right) = k^{-1} \sum_{t=1}^{k} \omega_i\left(x, \theta_t\right) \tag{5.11}$$

here k is the number of trees, and each tree is built from an independent and identically distributed vector $\theta_t, t = 1, \ldots, k$ which encodes the information on which predictors are used as split point in each node of the tree. The indicator function $1_{\{y_i \leq y\}}$ assumes the value 1 when the observation $y_i \leq y$, and is equal to 0 otherwise. The quantities $\omega_i(x, \theta_t)$ are weights obtained from observation y_i of the response variable, for the t-th tree. They are defined as

$$\omega_i(x, \theta_t) = \frac{1_{\left\{x_i \in R_{l(x,\theta_t)}\right\}}}{\sum_{j=1}^{n} 1_{\left\{j : X_j \in R_{l(x,\theta_y)}\right\}}} \tag{5.12}$$

where again $1_{\{\cdot\}}$ denotes the indicator function, so that they are zero if observation x_i does not belong to leaf $l(x, \theta_t)$ and are positive otherwise. The leaf $l(x, \theta_t)$ refers to the terminal node of tree t individuating the region $R_{l(x,\theta_y)}$ obtained by partitioning the predictor variable space at each node split.

Here I apply this algorithm as implemented in the Scikit-garden Python package [Kum17]. For the present application, I select a set of explanatory features which are representative of the local rainfall climatology, and which can all be obtained from a remotely-sensed dataset, without the need of observations at the ground. For this purpose I identify as predictors the stretched exponential scale (C_L) and shape (w_L) parameters, the average yearly number of events N_L, and the variance reduction function γ_0 as estimated from the QPE dataset. Additional predictors I consider to describe local environmental conditions are the mean elevation averaged over the QPE grid cell (μ_e), and the standard deviation of the elevation σ_e computed over a domain of size 1.25°x1.25° degrees centered over the location of interest, in order to account nearby orographic features. Predictor variables were normalized by subtracting their averages and diving by their standard deviations. Elevation data were obtained from the *global topographic map* obtained from NOAA National Center for Environmental Information [AE09], which provides elevation with respect to mean sea level globally at 1 arc-minute resolution. This information is here averaged so as to match the TMPA grid over the CONUS. The QRF model is here applied by training $k = 2000$ decision trees, although experiments were carried out with different values of k to make sure the results were not overly sensitive to this choice. I note that for their nature, RF and QRF models should not be used in extrapolation outside the range of the predictors used for training. Instead, they should be trained using a set of observations representative of the range in which I wish to obtain predictions. Therefore, when applied to a heterogeneous region such as the CONUS, one would like to understand if error predictions are possible based on sparse and low-density network of training sites, which is the working condition over many areas worlwide.

Here I address this question by designing a cross-validation procedure in which I take advantage of the relatively dense observational network available in this study over the CONUS to simulate how the method performs when trained using a low-density gauge network and tested on an independent dataset. First, locations within the set of gauged sites (SGS) are randomly resampled. Starting from a random location I construct a sample by examining sequentially (and randomly) inspecting and removing all gauged sites located at a distance less than 100km from any sites already included in the sample. By doing so, I obtain a sample size of about 175 sites (which vary at each realization), and obtain a dataset in which selected sites are at most weakly spatially correlated, thus allowing for proper calibration and testing of the error model. Second, I randomly divide the stations obtained in stage 1 into two independent sets, used for training the model and for independent testing respectively. Here I use 50% of the sites for calibration and the remining 50% for validation of the QRF model. I repeat these two steps a number of times ($n_g = 20$ in the analysis reported here), and pool together the results to obtain a global measure of performance over the test sites extracted in each realization. In each testing or training site, I performed the downscaling of QPE statistics as described in the previous Section 5.2, obtaining estimates of the parameters $C_{0,d}$ and $w_{0,d}$, as well as of the yearly number of events $N_{0,d}$ at a point. Moreover, at these sites independent rain gauge estimates of these quantities at the ground ($C_{0,g}$, $w_{0,g}$ and $N_{0,g}$) are also available. Therefore, I can obtain a measure of the error in downscaling these variables by simply applying eq. (5.9). Results of this analysis are featured in Figure 5.4, which compares the values of the errors η_C, η_W, and η_N predicted for the test sites with the 'true' values of such errors obtained using gauge records at the ground. In addition, the lower panels in Figure 5.4 show the values $C_{0,c}$, $w_{0,c}$, and $N_{0,c}$ obtained by correcting the downscaled parameters using the QRF error model. The QRF-corrected values show a good agreement with their counterpart from rain gauges at the ground, especially for the scale parameter $C_{0,c}$ and the number of events $N_{0,c}$. A somewhat larger scatter occurs for the shape parameter $w_{0,c}$. Overall, this result constitutes a clear improvement with respect to the downscaled parameter values shown in Figure 5.2. While using an ensemble of decision trees as done in the QRF algorithm allows to reduce prediction uncertainty and improve stability with respect to single decision trees, these improvements occur at the expense of the interpretability of the model. Nevertheless, it is possible to obtain an estimate of the relative importance of the predictor variables. In the regression model, the importance of a given predictor is computed as the mean decrease in the sum of squared residuals achieved at each node split by selecting that predictor. [JWHT13]. The importance of the features used in the QRF model to predict the three errors of interest η_C, η_w and η_N is quantified in Figure 5.5. The error in the number of events seems to depend primarily on the characteristic scale of ordinary rainfall events C_L, whereas for the error in the scale parameter η_C, the most important features are the parameters of the ordinary rainfall distribution at the QPE grid cells scale. In the case of the error in the shape parameter η_w, the most important predictors are those representing local orography, and the scale parameter C_L. The average number of events does not play a primary role for either of the three error variables, and is only somewhat relevant for predicting the values of η_C. However, I note that these parameter importances should be interpreted with caution, as the predictor variables are not completely independent between each other. While this is not an issue for the QRF error predictions, this should be kept in mind when examining parameter importances. The dependence of the shape parameter error from elevation suggested by the QRF analysis is not a surprise, given its spatial distribution shown examined in Figure 5.3. However, it is interesting to note that the role of orography seems to be more substantial in explaining errors in the tail of the distribution (i.e., in w_0) rather than on the characteristic rainfall magnitude or in the yearly average number of events.

5.4 The Statistical Distribution of Extreme Events from Remotely-sensed Rainfall

I turn here my attention to studying the probability distribution of extreme events based on TMPA QPE's, building on the tools developed and tested in the previous sections. To this end, here I provide the first large-scale application of the *Metastatistical Extreme Value Distribution* (MEVD) to remotely sensed QPE fields. Following [MI15, ZBM16, ZM19] I study the distribution of the annual maximum, $H_s^{(m)}$, among a variable number of independent and identically-distributed ordinary rainfall events, here at the daily time scale. The parameters describing the distribution of ordinary events, as well as the yearly number of events, are themselves considered to be random variables, whose realizations are estimated in each year of a rainfall record. From a sample of M years of daily rainfall observations, I thus estimate the cumulative probability distribution of the annual maximum according to the MEVD:

$$P\left(H_s^{(m)} < h\right) \simeq \zeta_s(h) = \frac{1}{M}\sum_{i=1}^{M}\left[F_s^{(i)}(h)\right]^{N_s^{(i)}} \tag{5.13}$$

where, for each year $i = 1, ..., M$ in the rainfall time series the yearly number of ordinary events is $N_s^{(i)}$ and: $P_i\left(H_s < h\right) = F_s^{(i)}(h)$ is the distribution of ordinary events in year i of the rainfall record, which here I assume to be a stretched exponential

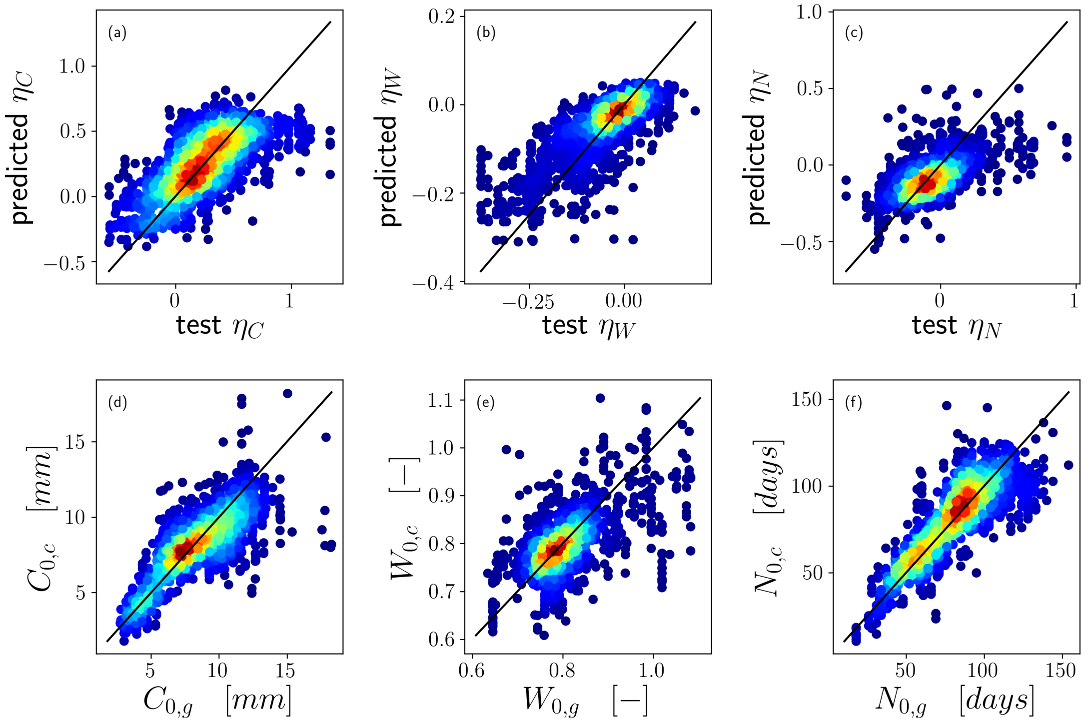


Figure 5.4: Results from the application of the quantile regression forest algorithm to predict parameter values at the ground and relative errors. Performance is quantified only for grid cells located at least 100 Km apart, in order to avoid the effects of spatial correlated grid cells. Upper panels (a), (b), and (c) report the dimensionless errors η_C, η_w and η_N for scale and shape parameters and for the number of events respectively. Lower panels report scatter plots of the corrected parameters compared to the respective values estimated at the ground: (d) Weibull scale parameter C_0, (e) Weibull shape parameter w_0, and (f) average yearly number of events N_0. Again, colors indicate the density of the points in the graph as computed by means of kernel density estimation, with warmer colors indicating a higher density of points.

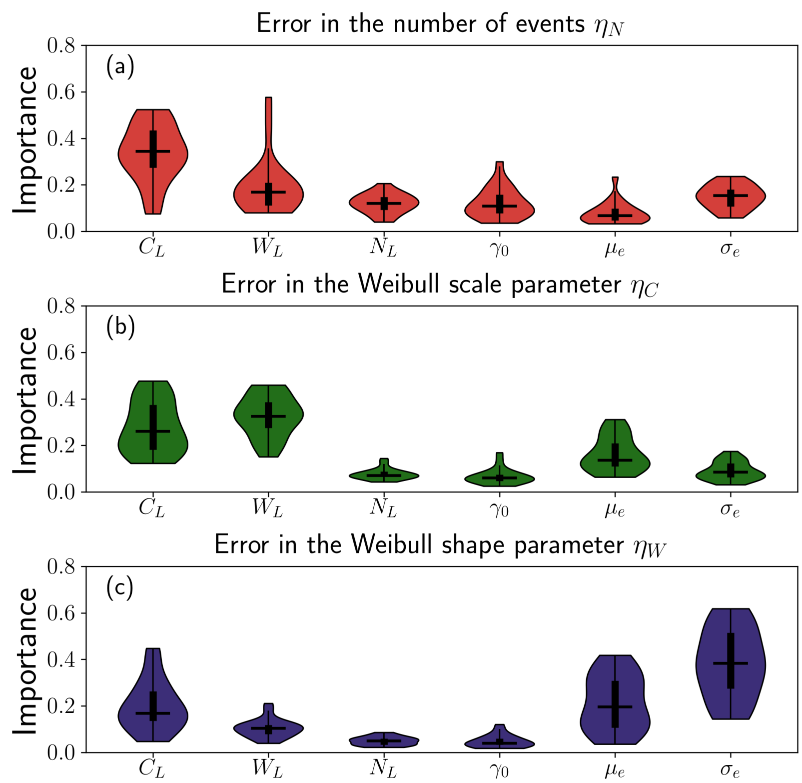


Figure 5.5: Distributions of importance for the predictor variables used in the QRF model for predicting relative errors in the average number of events (panel a), Weibull scale parameter (panel 2) and shape parameter (panel 3) as estimated in the cross validation analysis used to test the error model. Horizontal and vertical black bars represent the medians and the interquantile range respectively.

with cumulative probability

$$F_s^{(i)}(h) = 1 - e^{-\left(\frac{h}{C_s^{(i)}}\right)^{w_s^{(i)}}} \tag{5.14}$$

with yearly scale and shape parameters $C_s^{(i)}$ and $w_s^{(i)}$, and spatial averaging scale s. The main idea behing the use of eq. (5.13) for modelling daily rainfall accumulations extreme values is the recognition that low frequency variability in the rainfall generating mechanism can produce heavier tails than would be otherwise observed. This effect is captured in eq. (5.13) by averaging over parameter values estimated independently for each year in the record [ZBM16]. While the Weibull distribution is widely used for hourly to daily accumulations [WT05, PAFG18], I note that Eq. (5.13) can be tailored to different parametric models for the probability distribution of ordinary rainfall events, thus conferring to the MEVD significant application flexibility. For computing rainfall quantiles for a given cumulative probability p_{ne} (or return time $T_r = 1 - 1/p_{ne}$) I can invert equation (5.13) and evaluate quantiles as

$$\hat{h_s}\left(Tr\right) = q + \zeta_s^{-1}\left(p_{ne}\right). \tag{5.15}$$

Parameters for the MEVD and GEV distributions were computed by means of probability weighted moments [GLMW79] and L-moments [HWW85] respectively. For extreme value analysis, years with more than 10% of missing observations (36 data points at the daily scale) were not included in the analysis, as the selection of annual maxima values as well as the estimation of daily rainfall statistics would be potentially biased in the presence of relevant fractions of missing data.

5.4.1 The MEVD applied to TMPA QPE's

The expected value of the daily rainfall accumulation corresponding to a 50-years return time was obtained by applying the MEVD to the TMPA 3B42 V7 Research version for the CONUS. For comparison, I produced the same estimate by fitting the Generalized Extreme Value (GEV) distribution (e.g., [Col01]) to the series of Annual Maxima (AM) extracted from the same dataset. The large-scale spatial features of extreme rainfall frequency over the CONUS are similarly captured by the two models (Figure 5.6). Interestingly, the spatial distribution of MEVD-estimated quantiles is significantly smoother when compared with its GEV counterpart. This is consistent with the notion that MEVD, by using all the observations available, produces more stable estimates [ZBM16] that are also less sensitive to outliers, inevitably present in remote-sensing estimates, as also noted in [MNA+19]. GEV-estimates are, on the contrary, dominated by a few large outlier values, as manifest in the "grainy" texture of Figure 5.6b. I argue that this behavior originates from the fact that MEVD inferences are based upon the entire distribution of ordinary events, determining a decreased sensitivity to biases in QPE estimates of large rainfall values compared to annual-maxima or peaks-over-threshold approaches. As discussed in [MA14], while TMPA is able to capture rainfall from intense convective events, the accuracy of the estimates progressively descreses when limiting the analysis to rainfall rates exceeding increasingly high threshold values. A second observation is that the 50-year quantiles estimated by MEVD appear to be larger than those estimated by GEV in many areas over the CONUS, especially in the midwest. This feature does not appear to be uniquely inherent to the statistical models used: If I repeat the same analysis for the rain gauges from the NOAA hourly precipitation dataset (HPD) aggregated at the same daily scale (Figure 5.6 c, d) I find that the two statistical models yield much more similar results in this case, although the 50-year quantile field obtained from MEVD remains spatially smoother that its GEV counterpart. It is reasonable that, being the MEVD and GEV distribution fitted using different parts of the dataset (annual maxima in the case of the GEV, and the bulk of the distribution in the case of MEV), the extreme values estimated through the two formulations should respond differently to distortions in the pdf of the QPE magnitudes. However, a direct comparison of the extreme value quantiles reported in panels a) -b) and c-d) of Figure 5.6 is not feasible because of the scale discrepancy between QPE's and point rain gauge measurements. Eq. (5.13) provides a direct link between properties of the ordinary daily rainfall distribution and the frequency of extreme values. Hence, using downscaled parameters of the ordinary rainfall distribution inferred from remote sensing, it can be used to directly compare the MEVD-derived extremes at a point with those from independent rain gauge records, and to possibly correct for emerging discrepancies. Here I test this idea by extending the cross validation method employed in Section 5.3 for testing how the correction applied to the stretched exponential parameters affects extreme rainfall quantile estimates. For each iteration in the cross-validation scheme, once the TMPA parameters $C_L^{(i)}$, $w_L^{(i)}$ and $N_L^{(i)}$ for $i = 1, ..., M$ have been computed over the test sites, they are used to provide an estimate of the $Tr = 50$ years quantile $\hat{h}_L\left(Tr\right)$ by means of eq. (5.15). Analogously, the downscaled ($C_{0,d}^{(i)}$, $w_{0,d}^{(i)}$,$N_{0,d}^{(i)}$) and corrected ($C_{0,c}^{(i)}$, $w_{0,c}^{(i)}$,$N_{0,c}^{(i)}$) values of the parameters are used to estimate the corresponding quantiles at a point, $\hat{h}_{0,d}\left(Tr\right)$ and $\hat{h}_{0,c}\left(Tr\right)$ respectively. For each testing site I independently compute the MEVD parameters ($C_{0,g}^{(i)}$, $w_{0,g}^{(i)}$,$N_{0,g}^{(i)}$) and quantiles $\hat{h}_{0,g}\left(Tr\right)$ from the rain gauge record at the ground. Note that here I estimate the parameters $C_s^{(i)}$, $w_s^{(i)}$ and the number of events $N_s^{(i)}$, appearing in eq. (5.13), separately for each year in the rainfall record so as to account for their inter-annual variability. I then downscale each yearly value of the parameters using the equations described in Section 5.2, and correct them using the error model described in Section 5.3. Note that the downscaling and correction relations used are the same for each year in the record. When downscaling the Weibull parameters for estimating MEVD at subgrid scales, I apply eqns. (5.2) and(5.3) to each set of yearly parameter values separately, using the constant values of the variance reduction function and intermittency function derived from the entire QPE time series available. In principle, the spatial correlation of precipitation may also vary year to year, e.g. as a consequence of the varying frequency of different precipitation types. However, given the known difficulty of estimating values of correlation for skewed, non-Gaussian processes such as Precipitation at the daily time scale [GK04a], I did not include this possible source of inter-annual variability in the model, and computed correlation values from the entire QPE time series. Therefore, in this application the intermittency of the rainfall field and the decrease of its variance with averaging scale are the same for each year of observations. The downscaled MEVD parameters are then corrected for each year of the record using the median value of the relative error predicted by the QRF algorithm. As for the parameter downscaling, I do not explicitly correct for possible biases in the inter-annual variability of the parameter estimates. The result of this comparison between TMPA, downscaled and corrected quantiles with the corresponding rain gauge quantiles $\hat{h}_{0,g}$ is reported in Figure 5.7, in the form of scatter plots (panel 5.7a) and pdf of the relative error between estimated values and reference rain gauge quantiles (panel 5.7b). One can see that QPE quantiles tend to consistently overestimate rain gauge quantiles, even though they have a different nature as they are best interpreted as areal average quantiles over the pixel as opposed to values at the point. This is clearly a byproduct of the biases observed in the TMPA-estimated Weibull parameters, especially the overestimation of the scale parameter observed in Figure 5.2c. Turning our attention to the downscaled values, they exhibit an even more substantial overestimation with respect to the rain gauge benchmark. This is expected, as moving towards smaller averaging scales the fluctuations of the daily rainfall process tend to become more energetic, as quantified by the increase in variance and decrease in the number of events captured by eqns. (5.3) and (5.2). Lastly, the quantile values corrected by applying the QRF error model appear to be much closer to the ground estimates. Despite some variability remaining in the distribution of the estimates, QPE statistics bias is greatly reduced and the pdf of relative errors exhibit in this case a clear peak around zero. This encouraging result supports the choice of applying the same correction to each yearly parameter in the MEVD distribution, without applying any specific correction to the inter-annual variability of the parameters. I note that the application of downscaling and bias correction techniques tend to have opposite effects on the estimated quantiles: The downscaling tends to increase the magnitude of the $50 - year$ event in this case, and the bias correction tends to decrese it (on average). Even though this might seem counterintuitive, and perhaps unnecessary, I stress that this distinction between scale effects and errors is quite important, as it is the difference between the downscaled and gauge values that should be minimized, for example when designing new sensors and algorithms: areal averages and point values should not be directly compared. Overall, the combined application of downscaling and bias correction appears to provide good results, and confirms that the bias correction performed on the parameters is relevant for correcting extreme value quantiles. As a representative application of the error model, I trained the QRF model with data from all the SGS sites and predicted the spatial distribution of the error over the CONUS. Figure 5.8 features the median relative errors in the Weibull parameters (panels 5.8(a) and 5.8(b)) and in the yearly number of events (panel5.8(c)). Consistently with the observations from gauged sites, these results clearly show spatial features such as the overestimation of the characteristic

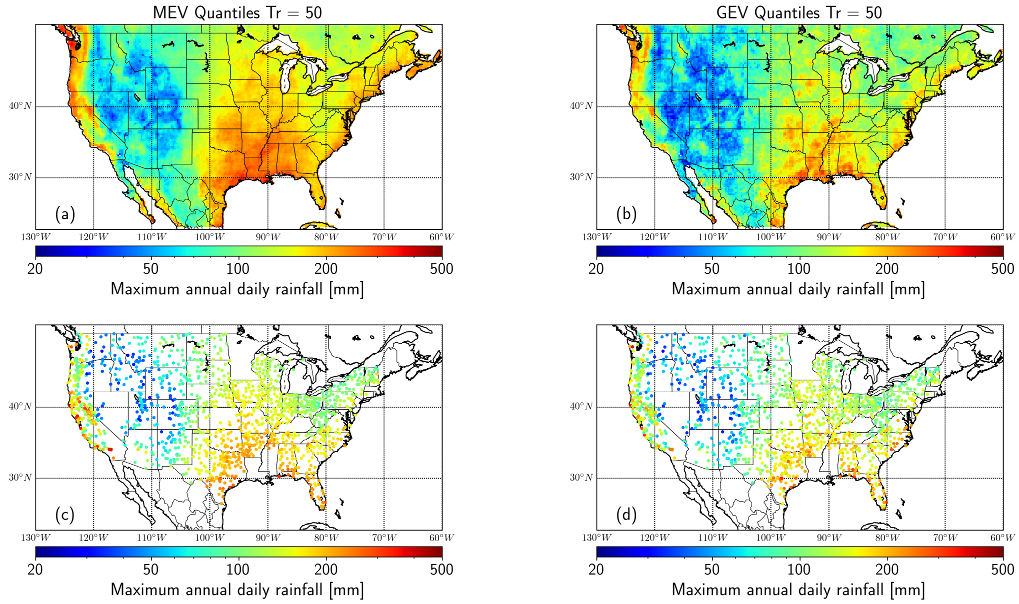


Figure 5.6: Results of the extreme value analysis over the study domain. Expected value for the annual maximum daily rainfall quantile corresponding to a return time of 50 years. (a) MEVD quantiles computed from the TMPA dataset over the CONUS, (b), GEV quantiles computed from TMPA over the same domain, (c), MEVD quantiles computed from the NOAA HPD rain gauge stations, and (c) GEV quantiles for the NOAA HPD stations. Only stations with at least 10 years of data were included in the analysis.

scale C_d over the Northeast US and the underestimation of the shape parameter w_d in the Western US, predominantly controlled by orography. After correcting the yearly MEVD parameters with the median QRF relativer error estimate, one can compute the errors in MEVD estimate quantiles. The spatial distribution of the error in the 50-year quantile is featured in Figure 5.8(d), which shows the quantile being generally overestimated over the CONUS. This overestimation is particularly relevant in the Western US, where underestimation of the shape parameter over complex terrain produces extreme value distributions with tails significantly heavier than the corresponding ground estimates.

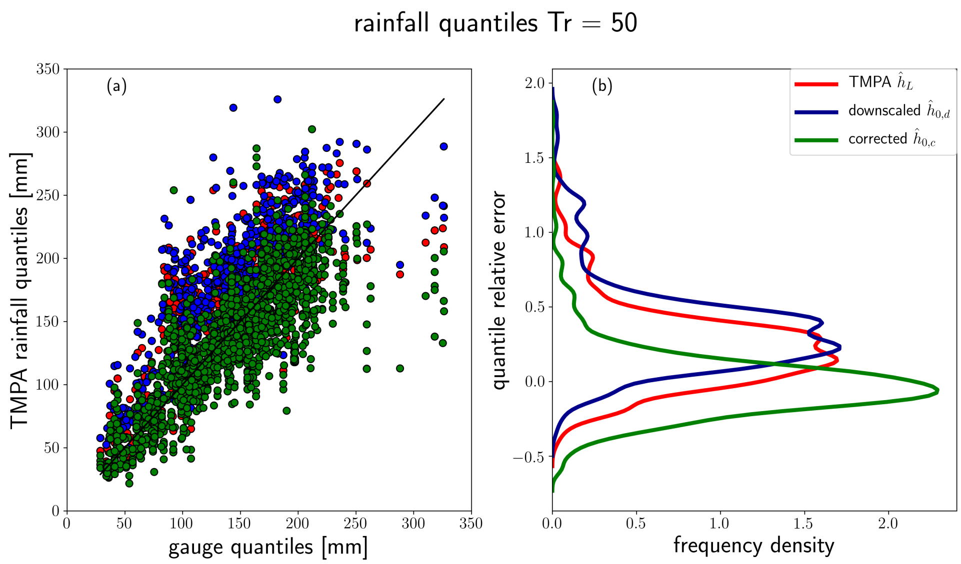


Figure 5.7: (a) Scatter plot comparing extreme rainfall quantiles computed from TMPA data to corresponding rain-gauge values for a return time $Tr = 50$ years, as a result of the cross validation scheme used to test the QRF error model. For each random extraction of the test sites, I report the TMPA extreme values quantiles estimated at the grid-cell scale (red circles), the corresponding values downscaled to a point in space (blue circles), and the QRF-corrected quantiles (green circles). In panel (b) I show the corresponding histogram of rainfall quantiles from gauge records (grey bars), and kernel density estimates for the distributions of TMPA quantiles (red line), downscaled quantiles (blue line) and corrected quantiles (green line).

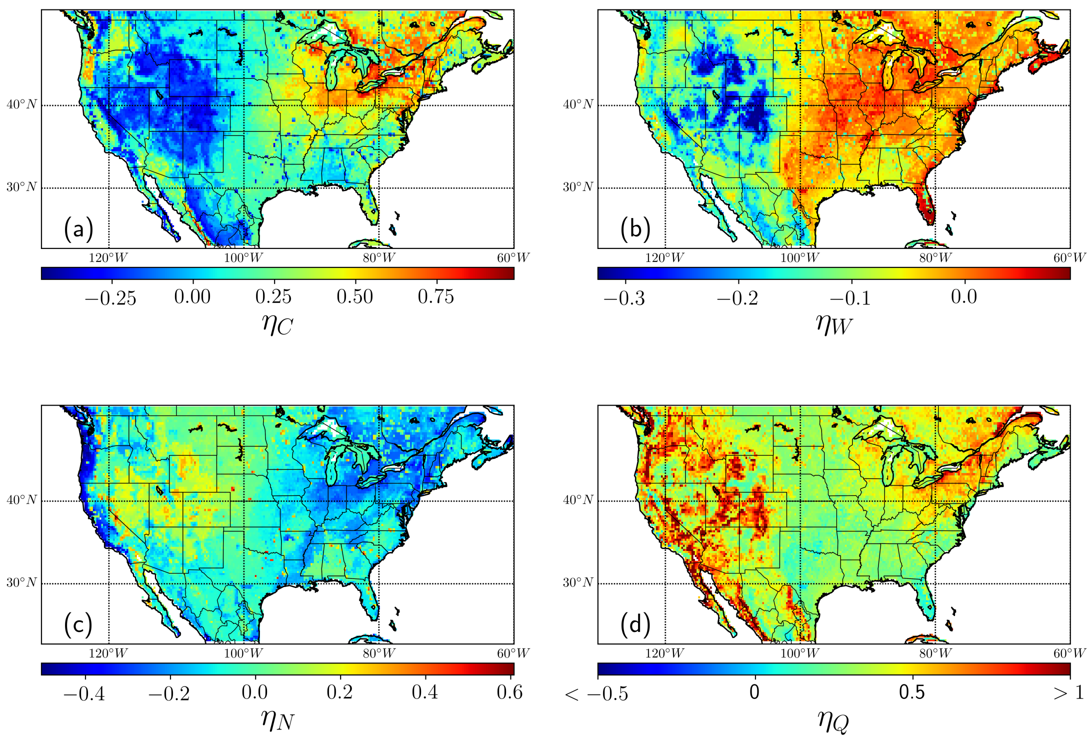


Figure 5.8: Spatial distributions of the median relative errors predicted for the Weibull scale (a) and shape (b) parameters, and for the yearly number of events (c). Panel (d) features the relative error in the 50-years MEVD estimated quantile, obtained correcting MEVD parameters using QRF-drived median error predictions.

5.4.2 The global distribution of rainfall extremes

I am now in a position to extend the analysis of rainfall extremes over the CONUS to the entire TMPA domain, which covers tropical, subropical and mid-latitude areas between $50°S$ and $50°N$. Application of the MEVD to the TMPA dataset is performed for each pixel independently, as done for the CONUS, and at the grid cell spatial scale. The global distribution of extreme daily rainfall obtained from computing quantiles for a 50-year return period with the MEVD is featured in Figure 5.9. The most extreme rainfall values occur over subtropical regions, with several hotspots over land which include South America and South-East Asia. While the most intense quantiles are estimated to occur over Ocean. However, the results obtained here by applying the error model should be regarded as representative of conditions over land only, and validation over ocean should be pursued separately (e.g., see [SA09, PG14]). Moreover, the gauge correction of TMPA dataset performed at the monthly scale can hardly be performed over the ocean, for lack of sufficient data. This circumstance is expected to negatively impact the accuracy of QPE's over the ocean with respect to the accuracy over land. When comparing quantiles estimated with GEV and MEVD, one can observe that i) the main quantitative patterns of the two models estimates are quite coherent at the global scale; and ii) MEVD estimates exhibit a remarkable spatial coherence when compared to GEV estimates. While the estimation uncertainty of GEV is heavily affected by the short dataset available here (20 years of annual maxima), regionalization techniques are expected to decrease this uncertainty [HW05, BRC+16, DZH18, SLH19]. However, I note that the reduction in estimation uncertainty achieved by using MEVD does not use information from nearby grid cells, but uses the entire distribution of ordinary values from a single grid cell. As an example, I show a zoomed-in representation of the 50-year event

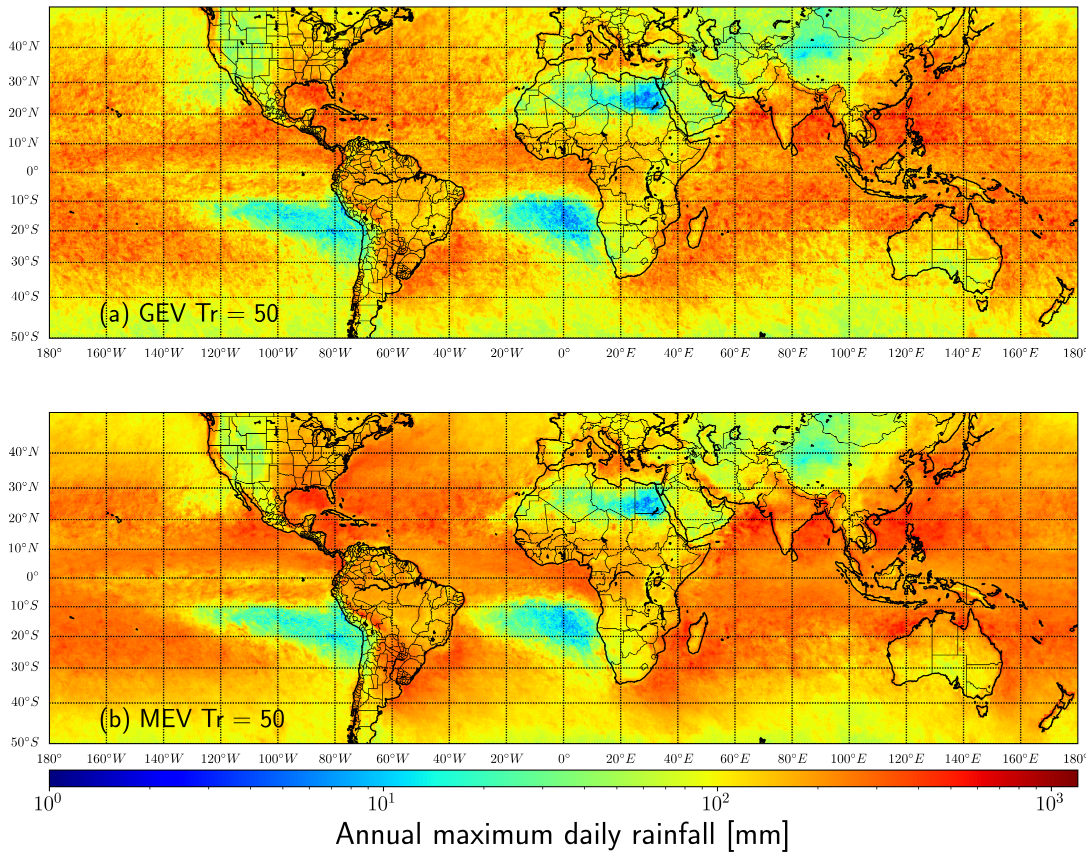


Figure 5.9: Distribution of extreme daily rainfall magnitude over the global-scale TMPA domain over land and ocean surfaces. Rainfall quantiles corresponding to a 50 years return time are reported as estimated by the GEV distribution (panel (a)) and MEVD (panel (b)).

estimated for a domain in South-East asia, one of the regions characterized by most intense maximum rainfall accumulation according to the global-scale analysis above 5.10). Again, MEVD estimates appear very coherent in space, suggesting that the uncertainty intervals in quantile estimates are significantly reduced with respect to the GEV model.

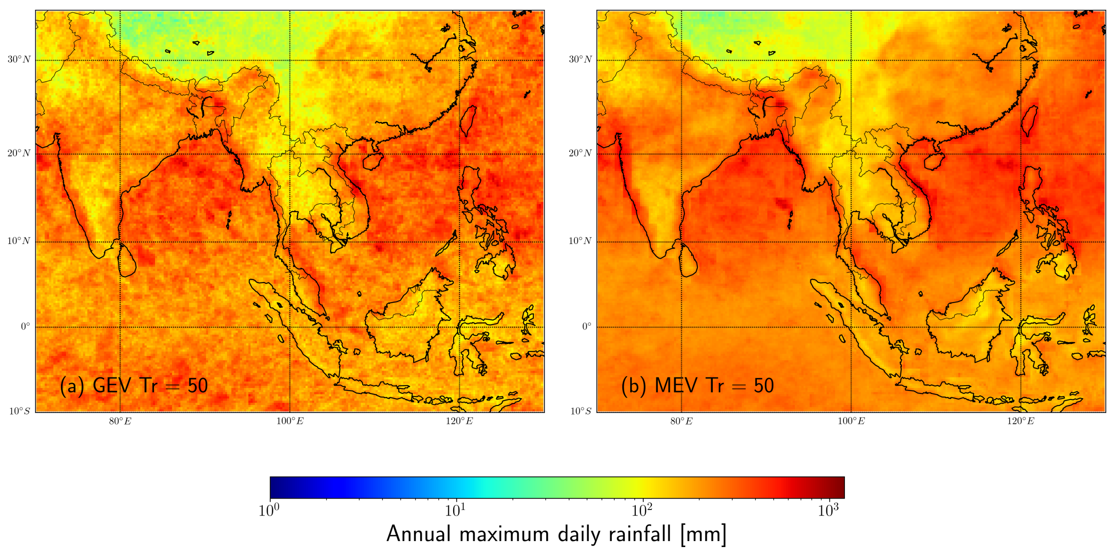


Figure 5.10: Distribution of extreme daily rainfall magnitude over South-East Asia. Rainfall quantiles corresponding to a 50 years return time are reported as estimated by the GEV distribution (panel (a)) and MEVD (panel (b)).

5.5 Discussion and Conclusions

In this Chapter I have described an approach for downscaling and validating satellite QPEs over ungauged areas, and for correcting and extending extreme-value estimates over ungauged locations. The methodology proposed here separately accounts for the scale difference between QPE gridded statistics and reference rain gauges at the ground, and for errors in QPE estimates. My cross-validation results quantify estimation uncertainty and support the downscaling and correction procedures proposed, suggesting that they can be applied to the entire CONUS and, by extrapolation, to many mid-latitude areas at the global scale, thus providing error estimates also where ground information is scarce. The analysis over the CONUS provides useful insight over a range of different climatic regimes, while optimally exploiting the ground observations available there for testing purposes. I also note that the two main assumptions adopted here, namely the adoption of the Weibull distribution for daily rainfall accumulations across scales and of eq. (5.7) for the spatial correlation, can in principle be relaxed and tailored to local rainfall regimes as needed, thus conferring to the proposed approach a greater flexibility. While the Weibull distribution has been advocated as a general model for the tail of the daily rainfall distribution at the global scale [WT05], the goodness of fit to daily values is expected to vary in different climate regions. However, the approach proposed here can be applied to other parametric models for daily rainfall, provided that the first two moments exist. The assumption of isotropic correlation function would be more difficult to relax, as is the case for the *Taylor* hypothesis used here to downscale the wet fraction of the rainfall process to the point. While these are approximations, a more refined description of the rainfall field would require the availability of information at the ground. As my main objective here is to propose a methodology for application to data-scarce regions, I utilize these assumptions on the rainfall field to infer sub-grid scale properties of the rainfall field from the QPE dataset alone. One limitation encountered in Section 5.2.5 concerns the downscaling of the spatial correlation function at locations where the correlation decays very rapidly with distance, as is the case for sites characterized by particularly complex terrain. While this situation occurs only for a small fraction of the SGS sites examined here, these conditions may occur in many mountainous regions worldwide. A possible way to address this limitation could be to develop spatially explicit models of the correlation based on local studies over complex terrain (e.g., [DWB15]). The results obtained in this study suggest that the gauge correction applied in order to obtain the TMPA research version dataset can lead to a significant deformation of the shape of the daily rainfall distribution. This is especially relevant in areas where remote sensors fail to detect a significant fraction of events. This consideration suggests the possibility of applying the QRF error model directly to the non-adjusted real time datasets. An interesting avenue of future research will be the application of this framework separately to the non-adjusted precipitation statistics computed for different seasons. This analysis would outline the performance of microwave retrievals in recovering rainfall statistics for different precipitation types, and quantify their contributions to the discrepancies observed in the yearly rainfall distributions. The aim of the downscaling approach proposed here is to infer rainfall statistics across spatial averaging scales at a fixed temporal integration scale (daily in this application). Applications of the methodology to different temporal scales are possible and deserve further investigation. Extending the downscaling procedure to finer scales may encounter some limitations. First, the estimation of the spatial correlation becomes increasingly uncertain when moving to short time scales, and the characteristic correlation length scale of the rainfall field decreases, as discussed by Gebremichael et al. [GK04a]. At temporal scales where the correlation characteristic distance becomes smaller than the grid cell size, the downscaling becomes unfeasible, as discussed in Section 5.2.5. Second, the Taylor hypothesis approach introduced by Zorzetto and Marani [ZM19] and discussed here in Section 5.2.3 requires information at temporal scales smaller that the temporal scale of interest in order to infer the wet fraction of the rainfall field at sub-grid spatial scales. Therefore, the shortest temporal scale at which the spatial downscaling analysis can be performed needs to be larger than the temporal resolution of the QPE dataset. Finally, a third limitation concerns the use of eq. (5.13): Moving to progressively finer temporal scales, the assumption of independent ordinary rainfall events becomes increasingly questionable as the temporal autocorrelation of the process becomes more significant. Therefore, application of MEVD would require declustering of the observations, e.g., as proposed by [MNAM18]. I also note that, for the QRF error model to work adequately, one should limit extrapolation of the QPE errors only to areas whose characteristics are represented in the training dataset. However, as I show in the cross validation experiment, the model can be successfully trained over large areas using low-density networks of rain gauges. This was indeed the main objective of the study, and it extends the domain where QPE validation is possible beyond the existing dense networks of rain gauges, which tend to be available only in developed countries, in densely populated areas, and are lacking in arid locations and locations characterized by complex topography. However, where this information is not available, the approach presented here provides a coherent quantification of estimation uncertainty, accounting for the different intrinsic nature of spatially-averaged satellite QPEs and point observations at the ground. The analyses proposed here can be used to inform the development of rain-gauge networks covering a targeted set of sample climatic location to maximize the inference potential over ungauged sites. An interesting research direction that builds on the results presented here is the possible extension of the methodology to a seasonal analysis of rainfall pdf's and correlation structure, or, possibly, to a classification based on different precipitation types. This type of differentiation could be useful either to assess more specifically the performance of remotely sensed QPEs in different seasons and for different precipitation types, and to study the distribution of extremes originated from multiple mechanisms.

Chapter 6

Extremes, Intermittency, and Time Directionality of Atmospheric Turbulence at the Crossover from Production to Inertial Scales

Adapted from: Zorzetto, E., A. D. Bragg, and G. Katul. "*Extremes, intermittency, and time directionality of atmospheric turbulence at the crossover from production to inertial scales.*" Physical Review Fluids 3.9 (2018): 094604.

6.1 Introduction

Turbulence in fluids is prototypical of spatially extended nonlinear dissipative systems characterized by large fluctuations that are active over wide ranging scales [Sre99]. The dynamics of a substance or scalar advected by a turbulent flow (often termed '*scalar turbulence*' [SS00]) is by no means an exception to this description. Scalar turbulence shares many phenomenological parallels with the much studied turbulent velocity fluctuations, especially in the inertial subrange. However, scalar turbulence also exhibits distinctive large- and fine-scaled temporal patterns (e.g. ramp-cliff) that are usually weak or all together absent from their component-wise turbulent velocity counterparts [ACFVA79, SS00, War00]. This finding is particularly true in the atmospheric surface layer (ASL) [Gar94, Stu12], a layer within the atmospheric boundary layer (ABL) that is sufficiently far above roughness elements but not too far from the ground to be directly impacted by the Coriolis force. In the ASL, the frictional Reynolds number $Re_* = u_* z/\nu$ can readily exceed 10^5, where z is the distance above the ground surface, u_* is the friction velocity related to the kinematic turbulent stress, and ν is the kinematic viscosity of air. A direct consequence of this large Re_* is a wide separation between scales over which turbulent kinetic energy (k) is produced and dissipated. In the absence of thermal stratification, k is produced at scales commensurate with z; however, the action of fluid viscosity responsible for the dissipation of k occurs at scales commensurate to or smaller than the Kolmogorov microscale $\eta_K = (\nu^3/\langle\epsilon\rangle)^{1/4}$, where $\langle\epsilon\rangle$ is the mean turbulent kinetic energy dissipation rate that is proportional to u_*^3/z for a neutrally stratified ASL [Stu12]. These estimates of $\langle\epsilon\rangle$ and η_K result in $z/\eta_K \sim Re_*^{3/4} > 5000$ in the ASL, which is rarely achieved in direct numerical simulations or laboratory studies. Embedded in this wide ranging scale separation is the inertial subrange [Kol41], where self similar scaling of velocity and air temperature structure functions is expected to hold for eddy sizes much larger than η_K but much smaller than z. Integral scales or scales comparable to z are directly influenced by boundary conditions imposed on the flow including surface heating (or cooling) in the ASL, whereas small scales (e.g. η_K) may attain universality and local isotropy after a large number of cascading steps away from the energy injection scales.

Much attention has been historically dedicated to the inertial subrange and the subsequent cross-over to the viscous or molecular regimes precisely because of the possible universal character of turbulence at such fine scales [Kra68, KPS92, War00, SSK+14, YZS15, KMP+15]. However, it is now accepted that some coupling between small and large scales exists, especially for passive scalars [War00, SS00, KPCS06], that act to enhance intermittency buildup across scales and distort any universal behavior by injecting the effects of the boundary conditions (or the k generation mechanism). Along similar lines of inquiry, it has been conjectured that the presence of coherent ramp-cliff patterns in concentration (or temperature) time series are responsible, to some degree, for this coupling [War00]. Ramp-cliff structures are characterized by local intense scalar gradients separated by large quiescent regions. The presence of ramp-cliff structures in scalar time series has been shown to break locality of eddy interactions and determine some departures from small scale isotropy.

Sweep-ejection dynamics connected to the presence of ramps are likely to play a major role in observed extreme value statistics, as shown e.g., for Lagrangian velocity sequences in plant canopy turbulence [Rey12]. Moreover, ramps are asymmetric and produce non-zero odd ordered structure functions, sharing striking resemblance with flight-crash events recently reported for the turbulent kinetic energy of Lagrangian particles [XPF$^+$14]. Even though ramps have been extensively observed experimentally [ACFVA79], studied as surface renewal processes [KPCS06], and from a Lagrangian perspective [SS00, FGV01], a unified picture describing their effects on inertial scales statistics remains lacking and motivates the work here.

My main objective is to investigate two questions about scalar turbulence at scales spanning production to inertial subranges: How do ramp-cliff patterns modify (i) the probability of extreme scalar concentration or air temperature excursions and its corollary intermittency buildup, and (ii) symmetry and time reversibility of scalar turbulence. These two questions are explored for differing turbulent energy injection mechanisms (mechanical and buoyancy forces) in the ASL. Here I focus on the production-to-inertial scales instead of the usual inertial to viscous ranges for the following reasons. First, any cross-scale coupling with ramp-cliff patterns is likely to be sensed at large scales commensurate with the ramp durations. Second, these scales are deemed most relevant when constructing sub-grid scale models for improving Large Eddy Simulations [MLC96, PAMP00, HPM03, SPA06]. Third, these scales encode much of the scalar variance that is needed when deriving phenomenological theories for the bulk flow properties based on the spectral shapes of the turbulent velocity and air temperature [KKP11, KLCBZ13, LKBZ12, KPSBZ14, LKZ15], especially for the ASL.

To achieve the study objectives, high frequency measurements of the three velocity components and air temperature fluctuations in the ASL are used to explore flow statistics at the transition from production to inertial scales. In particular, the focus is on the first two decades dominated by approximate inertial subrange effects, where the transition from the large eddies to the universal equilibrium or inertial range occurs. The statistical properties of temperature increments within this range of scales is examined with the goal of addressing to what extent the tail properties (and thus the probability of extreme events) at fine scales still carry signatures from the production ranges and in particular of large coherent structures such as ramp-cliffs. The experiments here span several atmospheric stability regimes that dictate to what degree turbulent kinetic energy is mechanically or buoyantly generated (or dissipated) depending on surface heating (or cooling) and on the turbulent shear stress near the ground [MO54b]. However, due to the large Reynolds number encountered in the ASL, the stable stratification is not sufficiently severe to allow for a transition to non-turbulent regimes. Therefore, the turbulence can be studied as three dimensional and fully developed.

This Chapter is organized as follows: In Section 6.2, the budget for turbulent kinetic energy forced by a mean velocity gradient and buoyancy is reviewed so as to define the key variables and dimensionless quantities pertinent to ASL flows. Then, the statistical tools used to characterize intermittency and time directionality of the scalar field are introduced. Section 6.3 presents the experimental setup, data process-ing, and compares the outcome of this experiment with predictions from traditional turbulence theory in the inertial subrange. The results obtained investigating extreme values and time directional properties for velocity and temperature are then presented in Section 6.4. In Section 6.5 the main conclusions are featured. Appendix E shows that distortions of the inertial range due to stable stratification are not relevant for the range of scales studied here.

6.2 Theory

6.2.1 Overview of ASL similarity at large- and small-scales

The turbulent kinetic energy budget for a stationary and planar homogeneous flow in the absence of subsidence is given by

$$\frac{\partial k}{\partial t_0} = 0 = -\overline{u'w'}\frac{dU}{dz} + \beta_o g\overline{w'T'} + P_D + T_T - \epsilon, \tag{6.1}$$

where $k = (\overline{u'^2 + v'^2 + w'^2})/2$ is the turbulent kinetic energy, u', v', and w' are the turbulent velocity components along the mean wind (or x), lateral (or y), and vertical (or z) directions, respectively, t_0 is time, and the five terms on the right-hand side of Eq. (6.1) are mechanical production, buoyant production (or destruction), pressure transport, turbulent transport of k, and viscous dissipation of k, respectively, β_o is the thermal expansion coefficient for gases ($\beta_o = 1/T$, T is air temperature here), g is the gravitational acceleration, $-\overline{u'w'} = u_*^2$ is the turbulent kinematic shear stress near the surface, and $\overline{w'T'}$ is the kinematic sensible heat flux from (or to) the surface. When $\overline{w'T'} > 0$, buoyancy is responsible for the generation of k and the ASL is classified as unstable. When $\overline{w'T'} < 0$, the ASL is classified as stable and buoyancy acts to diminish the mechanical production of k. The relative significance of the mechanical production with respect to the buoyancy generation (or destruction) may be expressed as

$$-\overline{u'w'}\frac{dU}{dz} + \beta_o g\overline{w'T'} = \frac{u_*^3}{\kappa z}\left[\phi_m(\zeta) + \frac{\kappa z \beta_o g\overline{w'T'}}{u_*^3}\right] = \frac{u_*^3}{\kappa z}\left[\phi_m(\zeta) - \zeta\right], \tag{6.2}$$

where

$$\frac{dU}{dz} = \frac{u_*}{\kappa z}\phi_m(\zeta), \quad \zeta = \frac{z}{L}, \quad L = -\frac{u_*^3}{\kappa g \beta_o \overline{w'T'}}, \tag{6.3}$$

and $\phi_m(\zeta)$ is known as a stability correction function reflecting the effects of thermal stratification on the mean velocity gradient ($\phi_m(0) = 1$ recovers the von Karman-Prandtl log-law), $\kappa \approx 0.4$ is the von Karman constant, and L is known as the Obukhov length as described by the Monin and Obukhov similarity theory [MO54b]. The physical interpretation of L is that it is the height at which mechanical production balances the buoyant production or destruction when $\phi_m(\zeta)$ does not deviate appreciably from unity. For a neutrally stratified ASL flow, $|L| \to \infty$ and $|\zeta| \to 0$. The sign of L reflects the direction of the heat flux, with negative values of L corresponding to upward heat fluxes (unstable atmospheric conditions) and positive values of L corresponding to downward heat flux (stable atmosphere).

Several bulk flow statistics in the ASL can be reasonably described by the aforementioned Monin-Obukhov similarity theory, including the mean air temperature gradient dT/dz and the air temperature variance $\overline{T'^2}$, both varying with ζ when normalized by a temperature scale $T_* = -\overline{w'T'}/u_*$. However, the statistics of some large-scale features within the temperature time series traces, such as the statistics of ramp-cliff patterns, do not scale with z. For starters, the ramp characteristic dimension is generally larger than z and their duration exceeds the mean vorticity time scale $(\kappa z \phi_m(\zeta)^{-1})u_*^{-1}$. Ramps have been observed within canopies, near the canopy atmosphere interface, and other types of flows as reviewed elsewhere [KPCS06, War00].

While z/L may not be the proper scaling parameter for ramps, it does indirectly impact many of their features in air temperature time traces sampled within the ASL. For example, in stably stratified ASL flows, the temperature ramps appear 'inverted' when compared to their near-neutral counterparts. The amplitudes and durations of ramps can increase with increasing instability due to weaker shearing and intense buoyant updrafts [CNBL97, TF07].

At small scales associated with the inertial subrange, the velocity and temperature second-order structure functions are commonly described by the Kolmogorov 1941 (K41) theory [Kol41] given as

$$S_u^2(r) = \overline{[\Delta u(r)]^2} = 4C_{o,u}(\langle\epsilon\rangle r)^{2/3}, \tag{6.4}$$

$$S_w^2(r) = \overline{[\Delta w(r)]^2} = 4C_{o,w}(\langle\epsilon\rangle r)^{2/3}, \tag{6.5}$$

$$S_T^2(r) = \overline{[\Delta T(r)]^2} = 4C_{o,T}\langle\epsilon_T\rangle\langle\epsilon\rangle^{-1/3}r^{2/3}, \tag{6.6}$$

where $\Delta u(r) = u(x+r) - u(x)$, $\Delta w(r) = w(x+r) - w(x)$, and $\Delta T(r) = T(x+r) - T(x)$ are the velocity and temperature increments at separation distance (or scale) r, $\langle\epsilon\rangle$ and $\langle\epsilon_T\rangle$ are the k and temperature variance dissipation rates respectively, $C_{o,u}$ and $C_{o,w}$ are the Kolmogorov constants for the longitudinal and vertical velocity components, and $C_{o,T}$ is the Kolmogorov-Obukhov-Corrsin (KOC) constant. These scaling laws, obtained under the assumptions of similarity and local isotropy, appear to hold reasonably in the ASL for scales smaller than $z/2$ [KHS97]. Moreover, the normalized third order structure functions

$$S(r) = \frac{S_u^3}{(S_u^2)^{3/2}} = \frac{\langle\Delta u(r)^3\rangle}{\langle\Delta u(r)^2\rangle^{3/2}} \tag{6.7}$$

and

$$F(r) = \frac{S^3_{TTu}}{S^2_T\,[S^2_u]^{1/2}} = \frac{\langle \Delta u(r)\Delta T(r)^2\rangle}{\langle \Delta T(r)^2\rangle \langle \Delta u(r)^2\rangle^{1/2}} \tag{6.8}$$

must be constant to recover K41 predictions for S^2_u and S^2_T in the inertial range [Obu49].

However, relevant deviations from K41 scaling have been reported for higher order structure functions, especially for the scalar fluctuations. These deviations arise as (i) Eqs. (6.4) - (6.6) do not account for intermittency related to spatial variability of the actual ϵ and ϵ_T, and (ii) the hypothesis of local isotropy might not hold for scalars due to non-local interactions across scales [Sre91]. A signature of the latter is the large structure skewness for temperature determined by ramp structures [KHS97, War00]. Many models, starting from Kolmogorov's log-normal dissipation rate refinement [Kol62], seek to relax some of the restrictive assumptions of K41 so as to explain the anomalous scaling observed in higher order moments. For scalars, these corrections are commonly expressed as

$$S^n_T = C_n\,(\epsilon r)^{n/3}\,(r/L_I)^{\zeta'_n - n/3} \tag{6.9}$$

where the exponent ζ'_n implies a scaling different from K41 that depends on the moment order n. The presence of an integral time scale L_I suggests an explicit dependence on large scale eddy motion within the inertial subrange. One estimate of L_I may be derived from the integral length scale of the flow given by

$$L_I = U \cdot I_w = U \cdot \int_0^\infty \rho_w(\tau_0) d\tau_0, \tag{6.10}$$

where $\rho_w(\tau_0)$ is the vertical velocity autocorrelation function and τ_0 is the time lag. Here, I_w is presumed to be the most restrictive scale given that w' is the flow variable most impacted by the presence of the boundary.

The statistics of air temperature increments across scales (τ_0/I_w) for different ζ conditions are explored with a lens on two primary features: buildup of heavy tails and destruction of asymmetry originating from ramp-cliff structures at the cross-over from $\tau_0/I_w > 1$ to $\tau_0/I_w \approx 0.1$. Because changes in ζ do result in changes in I_w, the time (or space) lags are presented in dimensionless form as $\tau = \tau_0/I_w$, so that the increments of a flow variable Δs, with $\Delta s = \Delta u, \Delta w, \Delta T$ at a given dimensionless scale τ, can be expressed as $\Delta s(\tau) = s(t+\tau) - s(t)$, where $t = t_0/I_w$.

6.2.2 Probabilistic description of intermittency across scales

The intermittent behavior of ASL turbulent flows has been documented by several experiments [KPC94, KVA01], and a number of models have been proposed to capture the effects of intermittency on the flow statistics in the inertial range of scales (e.g., lognormal, bi- and multi-fractals - beta model, log-stable, She-Leveque vortex filaments, etc). Common to all these models is the hypothesis of local isotropy and the accounting for uneven distribution of eddy activity in the space/time domain, which explains the anomalous scaling of higher order even structure functions.

Here, a statistical description of scalar increments is used to fingerprint large-scale signatures across scales τ for different ζ. If such fingerprints exist, the dissipation rates ϵ and ϵ_T need not be sufficient to describe all aspects of the inertial range statistics. The one-time probability density function (pdf) of the increments $\Delta s(\tau)$ of the flow variable $s = u, w, T$ at a given dimensionless scale τ, can be expressed as [PC93]

$$p(\Delta s) = \frac{N}{q_o(\Delta s)} \exp \int_0^{\Delta s} \frac{r_o(\Delta s')}{q_o(\Delta s')} d\Delta s'. \tag{6.11}$$

This expression is exact when Δs are realizations of a stationary stochastic process

$S(t)$ under the condition $p(\Delta s) \to 0$ as $\Delta s \to \infty$. Here $q_o(\Delta s) = \langle \dot{S}^2 | \Delta s \rangle / \langle \dot{S}^2 \rangle$ and $r_o(\Delta s) = \langle \ddot{S} | \Delta s \rangle / \langle \dot{S}^2 \rangle$ are the normalized averages of the first and second order conditional derivatives of the process $S(t)$, and N is a normalization constant. Eq. (6.11) generalizes previous results obtained by Sinai and Yakhot [SY89] and Ching [Chi93] for the pdf of temperature fluctuations and their increments, where the term $r_o(\Delta s)$ was linear ($r_o(\Delta s) = -\Delta s$). Eq. (6.11), while derived for a twice-differentiable process, can be interpreted as the steady-state solution of a Fokker Planck equation with $p(\Delta s)$ vanishing at infinite boundaries, with drift and diffusion coefficient equal to r_0 and q_0 respectively [Gar24, PKC+11].

Although Eq. (6.11) can be directly computed from an observed time series, the estimation of the conditional derivatives in $q_o(\Delta s)$ and $r_o(\Delta s)$ becomes inevitably uncertain as Δs approaches the tails of the pdf. However, a number of parametric distributions commonly used in statistical mechanics arise as particular cases of Eq. (6.11) when $r_o(\Delta s) = -\Delta s$, such as Gaussian ($q_o$ constant), power-laws ($q_o(\Delta s) \sim \Delta s^2$) and stretched exponentials ($q_o(\Delta s) \sim \Delta s^a, 0 < a < 2$). To facilitate estimation and comparisons with data, two different parametric models for the tails of Eq. (6.11) are here adopted: a Stretched Exponential (SE) and a q-Gaussian distribution (QG). The first arises from multiplicative processes of normal-distributed random variates [FS97], while the second maximizes a generalized measure of information entropy proposed by Tsallis [Tsa88, TLSM95, SVKA05]. While QG does not have a clear physical basis in the context of turbulent flows [GK04b], it has been widely used in the analysis of turbulence simulations and data [RRN+01, AA02, BRS+02, KPCS06]. I employ these two models to infer tail behavior as well as to test the independence of my findings from the particular parametric distribution used to characterize $p(\Delta s)$. The QG and SE pdfs are given as

$$p_{QG}(\Delta s) = N(q) \cdot \left(1 + (q-1)\frac{\Delta s^2}{2\psi^2}\right)^{\frac{1}{1-q}}, \tag{6.12}$$

$$p_{SE}(\Delta s) = \frac{\eta}{\lambda}\left(\frac{\Delta s}{\lambda}\right)^{\eta-1} \cdot \exp\left(\frac{\Delta s}{\lambda}\right)^{\eta}. \tag{6.13}$$

Both pdf models have two degrees of freedom corresponding to a scale (ψ,λ) and shape (η, q) parameter. I adopt the (symmetric) QG model and the SE fitted separately to right and left tails of $p(\Delta T)$.

6.2.3 Probabilistic description of asymmetry and irreversibility across scales

The presence of ramp-cliff structures has been conjectured to result in non-local interactions of different size eddies within the inertial subrange [War00]. This non-locality affects both even and odd moments of higher order. A statistical framework to investigate the effects of ramps on the asymmetric nature of velocity and scalar increments for different atmospheric stability classes is now discussed. Sharp edges associated with cliffs might directly inject scalar variance at much smaller scales and thus alter the magnitude and sign of odd order moments within the inertial range (depending on z/L). The presence of asymmetry has been investigated based on odd-ordered structure functions [War00] or multipoint correlators [MPS$^+$98]. In particular, a simple measure for the persistence of asymmetry at small scales is the skewness of the scalar increments $S_T^3 = \langle \Delta T(\tau)^3 \rangle / \langle \Delta T(\tau)^2 \rangle^{3/2}$. The structure skewness of air temperature has been found to scale as $Re_\lambda = \sigma_u \lambda/\nu$ (where λ is the Taylor microscale and σ_u is the root mean square of the longitudinal velocity fluctuations) and thus for a boundary layer $S_3^T \sim Re_*^{1/2}$. However, for large values of Re_λ experimental evidence suggests that S_3^T tends to plateau and become independent of Re_λ [Sre91, War00].

A further signature of ramp-cliff structures is that increments $\Delta T(\tau)$ may exhibit a time directional (or 'irreversible') behavior. Time reversibility implies that the trajectories of a stationary process Θ_t exhibit the same statistical properties when considered forward or backward in time. In particular, for a reversible time series the n-points joint pdf of $(\Theta_1, \Theta_2, ...\Theta_n)$ is equal to the joint pdf of the reversed sequence $(\Theta_n, \Theta_{n-1}, ...\Theta_1)$ for every n. While testing this general definition of reversibility would require perfect knowledge of the phase space trajectories, a weaker definition is the so called lag-reversibility. This condition only requires the two-points pdfs to be equal: $f_{\Theta_t,\Theta_{t+\tau}}(\Theta_1,\Theta_2) = f_{\Theta_{t+\tau},\Theta_t}(\Theta_2,\Theta_1)$. While this definition is less general, it still provides a necessary condition for testing time reversibility. Moreover, it is consistent with the traditional descriptions of turbulence that are primarily based on two-point statistics. Lag reversibility implies that [Law91]

$$R_\tau = \rho_c(\Theta_t^2, \Theta_{t+\tau}) - \rho_c(\Theta_t, \Theta_{t+\tau}^2) = 0. \tag{6.14}$$

where ρ_c denotes a correlation coefficient between two variables. This condition can be directly tested across different τ and ζ using a conventional correlation analysis.

A second test for reversibility of scalar trajectories is here performed based on the Kullback-Leibner measure, a form of relative entropy that determines the average distance between the entire pdf of forward and backward trajectories [CT18, PRD07, PKC+11]. Again, the analysis here is restricted to the inspection of lag-reversibility ($n = 2$) across scales τ. In such a restricted form, this measure reduces to

$$\langle Z_\tau \rangle = \int_{\Omega_\Theta} \int_{\Omega_{\Theta'_\tau}} p(\Theta'_\tau|\Theta) p(\Theta) \log \frac{p(\Theta'_\tau|\Theta)}{p(-\Theta'_\tau|\Theta)} d\Theta'_\tau d\Theta, \tag{6.15}$$

where $\Theta'_\tau = \Delta\Theta(\tau)/\tau$, and the domains of integration Ω_Θ and $\Omega_{\Theta'_\tau}$ correspond to the populations of the random variables Θ and Θ'_τ respectively. Eq. (6.15) determines, at each dimensionless scale τ, the average distance between the probability of the transition $\Delta\Theta(\tau)$ and its inverse, at every given level Θ.

A statistical mechanics interpretation of Eq. (6.15) would imply that for a system in non-equilibrium steady state, the *Fluctuation Theorem* must hold so that

$$\log \frac{p(-Z_\tau)}{p(Z_\tau)} = -Z_\tau \tag{6.16}$$

for the variable Z_τ computed at some level Θ

$$Z_\tau(\Theta) = \log \frac{p(\Theta'_\tau|\Theta)}{p(-\Theta'_\tau|\Theta)}. \tag{6.17}$$

Note here the usage of conditional probabilities instead of their unconditional forms employed in recent flight-crash studies of Lagrangian fluid particles [XPF+14] that also made use of Fluctuation Theorem and time-reversibility. Eq. (6.15) has been shown to have general validity [PRD07] independent of the underlying dynamics or statistical-mechanics interpretations, when considering conditional statistics.

6.3 Data and Methods

The three velocity components and air temperature measurements were sampled at 56 Hz using an ultra-sonic anemometer positioned at z =5.2 m above a grass-covered surface at the Blackwood Division of the Duke Forest, near Durham, North Carolina, USA. The anemometer samples the air velocity in three non-orthogonal directions by transmitting sonic waves in opposite directions and measuring their travel times along a fixed 0.15 m path length. Temperature fluctuations are then computed from measured fluctuations in the speed of sound assuming air is an ideal gas. The non-orthogonal sonic anemometer design used here has proven to be the most effective at reducing flow distortions induced by the presence of the instrument.

The experiment resulted in 123 time series records (henceforth termed 'runs') each having a duration of 19.5 minutes (65536 data points at $56Hz$), covering a range of different atmospheric stability conditions [KHS97]. Of these, only 34 runs passed a stationarity test and were included in the analysis (see Table 6.1 for a summary of the properties of the flow for these runs). The assumption of stationarity is necessary so as to (i) decompose the flow variables into a mean and fluctuating part, (ii) adopt Eqs. (6.11) and (6.15) so as to describe intermittency and time irreversibility respectively, and (iii) compute the integral scales needed in delineating the transition from production to inertial. To test the dataset for stationarity, I employ the second order structure functions of velocity components (u, w) and air temperature T. Runs were included only if the slope of $S_s^2 = \langle [s(t+\tau) - s(t)]^2 \rangle$ for time delays larger than about 9 minutes (30000 sample points) was smaller than a fixed value (0.01). For the 34 runs meeting this strict stationarity criterion, second order structure functions for w and T are featured in Fig. 6.1. As expected, structure functions exhibit an approximate 2/3 scaling at fine scales and transition to a constant value as the autocorrelation weakens at large separation distances.

The presence of a stable stratification is known to produce distortions on the spectral properties of turbulence at scales commensurate with (and larger than) the Dougherty-Ozmidov length scale [RMP15]. I investigated this issue (see Appendix E for more details) finding that stable stratification effects are only relevant at scales larger than the integral scale I_w considered here and not in the inertial range.

As earlier noted, the most restrictive (i.e. smallest) integral time scale is I_w associated with the vertical velocity w due to ground effects. I assume that this time scale characterizes the transition from production to inertial ranges for all three flow variables u, w, T. Eq (6.10) is here evaluated by integrating $\rho_w(\tau)$ up to the first zero crossing so as to avoid the effects of low frequency oscillations. Figure 6.1 illustrates the integral time scales of w and T as a function of ζ, where the aforementioned integral time scales are normalized by the mean vorticity time scale $dU/dz = \phi_m(\zeta)u_*(\kappa_v z)^{-1}$. It is clear that such normalized I_w is approximately constant across stability regimes and suggests I_w to be proportional to the duration of vortices most efficient at transporting momentum to the ground for all ζ. Conversely, the temperature integral time scale is much longer than I_w for near-neutral conditions and only approaches I_w for strongly unstable conditions.

A known limitation of sonic anemometry is the presence of distortions at high frequencies due to instrument path-averaging. For this reason, the smallest time scale considered in the analysis is $0.05 \cdot I_w$, which corresponds to a minimum travel path of $30cm$ (or twice the sonic anemometer path length). Taylor's frozen turbulence hypothesis [Tay38] ($r = -\overline{U}t$) was employed to convert values of τ to separation distances r within the inertial subrange even though the turbulent intensity σ_u/U is not small as shown in Table 6.1. For this reason, I adopt the dimensionless lag τ for analysis and presentation. The τ can be interpreted as temporal or spatial noting that distortions due to the use of Taylor's hypothesis impact similarly the numerator and denominator.

For every run, ζ was computed using Eq. (6.3) and then employed to classify the ASL stability condition. Most of the runs in the dataset are unstable with a wide range of $|\zeta|$, while only 4 runs are characterized by $\zeta > 0$. To ensure a balanced statistical design, two stability classes are selected with the same number of runs (8) in each class: strongly unstable ($|\zeta| > 0.5$) and near neutral runs ($|\zeta| < 0.072$). A summary of the bulk flow properties for these runs are featured in Table (6.1).

In the analysis, each flow variable s ($s = u, w, T$) is normalized to zero-mean and unit-variance (labeled as s_n). Then, at scale τ, a time series of $\Delta s(\tau) = s_n(t + \tau) -$

$s_n(t)$ is constructed and again normalized to have unit variance.

For illustration purposes, Fig. 6.2 shows sequences of fluctuations u', w', T' extracted from runs in unstable and stable atmospheric regimes. In the first case, temperature fluctuations clearly exhibit ramp-cliff structures occurring with time scales larger than I_w. In the stable/near neutral case, large scale scalar structure are still present even though their structure is qualitatively different from the unstable case, and may include inverted ramp structures as in Fig. 6.2(B) when $\overline{w'T'} < 0$.

To test the effects of these coherent structures on inertial subrange statistics, and in particular to isolate the effect of temperature ramps on intermittency and asymmetry, synthetic time series are used and are constructed as follows. First, a phase-randomization of the original temperature records [PT94] is performed by preserving the amplitudes of the Fourier coefficients while destroying coherent patterns encoded in the phase angle. A synthetic sawtooth time series is then superimposed on the time series obtained by phase-randomization. Here a coefficient α measures the relative weight of the ramps with respect to the phase-randomized sequence. This combination yields realizations of a renewal process (see Fig. 6.2(C) for a representative example with $\alpha = 0.5$) that is unconnected with Navier-Stokes scalar turbulence, but mimics sweep-ejection dynamics [KPCS06]. Synthetic ramps are here generated with exponentially distributed durations and with a mean duration set to a multiple of the integral time scale ($2 \cdot I_w$ in Figure 6.2(C)). The resulting time series is again normalized to have zero mean and unit variance.

Table 6.1: Bulk flow properties for the runs in the dataset. The table reports the atmospheric stability parameter ζ, the Obukhov length $L \quad [m]$, the sensible heat flux $H = \rho C_p \overline{w'T'} \quad [Wm^{-2}]$ (where ρ is the mean air density and C_p is the specific heat capacity of dry air at constant pressure), the mean air temperature $T \quad [°C]$ and mean velocity $U \quad [m/s]$, and the integral time scale for w $[s]$, the turbulent intensity σ_u/U, the temperature standard deviation $\sigma_T \quad [°C]$, and vertical velocity standard deviation $\sigma_w \quad [m/s]$.

Run	ζ	L	H	T	U	I_w	σ_u/U	u^*	σ_T	σ_w
1	-11.56	-0.4	93.2	33.9	2.1	2.62	0.44	0.08	0.48	0.40
2	-1.31	-4.0	121.6	26.9	1.0	7.58	0.72	0.17	0.54	0.30
3	-0.89	-5.8	73.1	27.8	0.5	6.62	0.91	0.16	0.37	0.30
4	-0.81	-6.4	79.9	32.7	0.7	5.75	1.05	0.17	0.61	0.29
5	-0.80	-6.5	138.1	27.4	0.8	8.18	0.48	0.21	0.57	0.31
6	-0.67	-7.7	149.8	31.4	0.9	11.64	1.04	0.23	0.63	0.38
7	-0.59	-8.8	118.1	34.8	1.5	3.43	0.71	0.22	0.58	0.34
8	-0.52	-10.0	85.4	32.5	2.1	1.74	0.37	0.21	0.44	0.37
9	-0.45	-11.5	78.6	31.7	1.1	7.44	0.61	0.21	0.43	0.30
10	-0.44	-11.7	110.7	31.9	1.2	5.89	0.65	0.24	0.49	0.37
11	-0.44	-11.8	39.4	34.4	1.3	3.19	0.45	0.17	0.32	0.29
12	-0.40	-13.0	36.6	34.1	1.7	2.30	0.39	0.17	0.37	0.28
13	-0.37	-14.0	65.1	25.2	1.6	2.91	0.39	0.21	0.35	0.27
14	-0.33	-15.6	48.0	28.9	1.4	2.58	0.41	0.20	0.27	0.30

Run	ζ	L	H	T	U	I_w	σ_u/U	u^*	σ_T	σ_w
15	-0.33	-15.8	4.8	33.4	1.6	1.59	0.35	0.09	0.09	0.23
16	-0.29	-18.2	115.2	32.1	2.7	2.16	0.37	0.28	0.44	0.47
17	-0.28	-18.5	136.2	29.2	0.9	6.88	1.11	0.30	0.56	0.37
18	-0.27	-19.1	108.6	30.5	1.7	3.56	0.62	0.28	0.54	0.34
19	-0.17	-29.7	70.5	29.5	2.6	2.22	0.29	0.28	0.36	0.42
20	-0.15	-33.8	63.2	32.9	2.2	2.97	0.39	0.28	0.36	0.40
21	-0.14	-37.9	30.9	34.2	1.6	4.17	0.49	0.23	0.34	0.32
22	-0.12	-44.4	118.6	31.0	2.6	3.78	0.42	0.38	0.49	0.42
23	-0.09	-56.5	26.7	33.9	1.9	3.39	0.31	0.25	0.15	0.31
24	-0.08	-61.7	49.7	31.7	2.0	3.50	0.41	0.31	0.27	0.39
25	-0.08	-65.1	17.6	34.0	2.2	3.22	0.29	0.23	0.13	0.31
26	-0.07	-72.5	28.8	31.5	1.8	2.71	0.41	0.28	0.29	0.30
27	-0.04	-126.2	45.1	31.0	4.3	1.21	0.33	0.39	0.35	0.71
28	-0.03	-171.8	3.9	31.3	1.7	3.18	0.39	0.19	0.15	0.30
29	-0.02	-261.4	46.1	31.2	3.8	1.37	0.39	0.50	0.23	0.72
30	-0.02	-304.3	47.1	29.4	5.0	0.84	0.31	0.53	0.21	0.80
31	0.002	2397.4	-0.4	31.2	1.9	1.94	0.44	0.22	0.69	0.32
32	0.01	525.5	-1.3	32.9	0.9	3.00	0.51	0.19	0.18	0.23
33	0.05	93.8	-20.7	29.8	2.6	1.52	0.30	0.27	0.23	0.39
34	0.07	71.4	-14.2	30.4	1.9	2.18	0.37	0.22	0.25	0.28

6.4 Results

The main questions to be addressed here require determination of (i) the probability of extreme scalar concentration excursions and concomitant intermittency, and (ii) scalar asymmetry and time irreversibility across scales. Here, tools introduced in sections 6.2.2 and 6.2.3 are used to investigate how these two features vary from production to inertial scales for temperature traces, and to compare this behavior with the observed velocity components. Comparison of these quantities for runs recorded in different atmospheric stability conditions allows to test whether significant coupling across scales exists, and to what extent velocity and temperature statistics are universal at the smallest scale examined here.

6.4.1 Probabilistic description of intermittency across scales

I first investigate the intermittent behavior of both scalar and velocity components by assessing to what extent the scaling of even-order structure functions departs from K41 predictions. Inspection of scaling exponents ζ'_n in Eq. (6.9) for u, w, T confirms that K41 predictions significantly overestimate scaling exponents for structure functions of order higher than 2, as shown in figure 6.3(A). The scaling exponents obtained for the scalar T show reasonable agreement with previous experimental results (Fig. 6.3(B)), with values systematically lower than predicted by the Kraichnan model in the limiting case of time-uncorrelated velocity field [Kra94]. The values of ζ'_n averaged over the set of runs observed during the experiment are lower for the scalar, especially when compared to the longitudinal velocity components. From this analysis, intermittency effects appear stronger for the scalar than for the longitudinal velocity.

The empirical pdfs of velocity and air temperature increments ($\Delta s = \Delta u, \Delta w, \Delta T$)

for runs in the near-neutral ($|\zeta| < 0.072$) and strongly unstable ($\zeta < -0.5$) classes (Fig. 6.4) show clear transitions from a quasi-Gaussian regime at large lags ($\tau = 2$ in figure) to distributions with sharper peaks and longer tails at scales well within the inertial subrange ($\tau = 0.05$). This behavior has been documented for a wide range of turbulent flows [Men91] and is associated with the build up of intermittency [Kol62] due to self-amplification inertial dynamics [LM05].

The bulk of the pdf of temperature increments at any given scale can also be characterized by the coefficients of Eq. (6.11). Results show some differences between runs with differing $|\zeta|$ (Fig. 6.5). Namely, for runs in the strongly unstable class, q_0 exhibits a more pronounced peak around the origin and is characterized by larger asymmetry at the cross-over scale $\tau = 1$ compared to their near-neutral counterparts (Fig. 6.5(A)). Moreover, the results here confirm that a choice of linear $r_0(\Delta T)$ and quadratic $q_0(\Delta T)$ appear reasonable for ASL flows. In the case of an unstable ASL, the term $r_0(\Delta T)$ remains linear, while inspection of $q_0(\Delta T)$ suggests that a dependence on s with an exponent smaller than 2 might be more appropriate, corresponding to stretched exponential tails for $p(\Delta T)$ for small lags τ in unstable ASL flows. Comparison with the same data after run-by-run spectral phase randomization [PT94] shows that the latter exhibits almost Gaussian behavior, confirming that the emergence of long tails at inertial scales is primarily a consequence of non linear structures in the original time series.

The variation of the tail parameters η and q with decreasing scale τ (Fig. 6.6) provides a robust measure of how the distributional tails of $p(\Delta T)$ evolve at the onset of the inertial range. For temperature differences, the rates of change across scales of both η and q appear to be dependent on the magnitude of the stability parameter ζ. Consequently, while at large scales - where the pdf closely resembles a Gaussian - neither η nor q exhibit a significant dependence on ζ, for scales well within the inertial subrange stability is clearly impacting the tail behavior of ΔT (Fig. 6.7).

This evidence suggests that the observed intermittency is not only internal (i.e., not only due to variability in the instantaneous dissipation rate [KPS92]) but is also directly impacted by the larger scale eddy motion that sense boundary conditions. In particular, when buoyancy generation is significant, the heat flux $\overline{w'T'}$ is connected with the sweep and sudden ejection of air parcels, corresponding with the sharp edges of the temperature ramps [ACFVA79, KPCS06]. The resulting sawtooth behavior could be responsible for the injection of scalar variance at small scales (instead of a gradual cascade), acting in particular on the negative tail of the ΔT pdf, as evident from Fig. 6.5(A). On the other hand, the buildup of non-Gaussian statistics for velocity increments is not as impacted by the stability regime, and therefore the dominant effects are in this case primarily an effect of internal intermittency.

6.4.2 Probabilistic description of asymmetry across scales

To compare the data sets used here with laboratory studies, I first test the validity of Obukhov's constant skewness hypothesis, which would require the third order structure function of the longitudinal velocity component being constant within the inertial range. Figure 6.8 reports the values of the third order structure functions (Eqs. (6.7) and (6.8)) evaluated at the onset of the inertial subrange as delineated by the w time series. Both are approximately constant for scales smaller than I_w. While comparison with experiments shows good agreement for $S(\tau) \simeq -0.25$, $F(\tau)$ is systematically smaller than its anticipated value [KHS97] (-0.4) for all ζ.

For the scalar T, The presence of a finite third order temperature structure function signifies that local isotropy is not fully attained in the range of scales explored here. The temperature skewness S_T^3 exhibits a plateau for scales smaller than I_w

(Fig. 6.9(A)) similar to previous measurements reported in grid turbulence forced by a mean temperature gradient [MW98]. Moreover, S_T^3 levels off to positive values for $\zeta > 0$, while it becomes negative for $\zeta < 0$. This finding is consistent with the presence of ramp-like structures when $\zeta > 0$ (mildly stable conditions) that are inverted when compared to their unstable counterparts.

The findings here confirm that at the cross-over from production to inertial, imprints of ramp structures persists well into the inertial subrange. The consequence of these imprints on time-reversibility is now considered for temperature sequences. The irreversibility analysis detects strong irreversbility at large scales that slowly decreases at the onset of the inertial range (Fig. 6.9). This finding is consistent with the idea that atmospheric stability determines a preferential direction for the large-scale scalar structures, which becomes progressively weaker at scales smaller than $\tau = 1$. Here, the sign of the heat flux has a primary effect on the orientation of the ramps, as captured by R_τ. Furthermore, phase randomization is shown to destroy much of this time irreversibility (Fig. 6.9(B)) while the addition of synthetic ramps, either with positive or negative orientation, produces values of R_τ that closely resemble observations of stable and unstable ASL respectively. These synthetic experiments also recover the sign of the third order moment S_T^3 (Fig. 6.9(A)) but not its magnitude at smaller scales. As one would expect, a sawtooth time series does not fully reproduce inertial scale scalar dynamics, even though it does clearly capture the qualitative effect of boundary conditions on scalar ramp-cliffs.

Additional insight can be obtained by the relative entropy measure defined in Eq. (6.15), which was here evaluated by integrating the relative entropy over the joint frequency distribution of normalized temperature fluctuations and their increments at each scale τ. I used a coarse binning for estimating the joint pdf $p(T'(\tau), T)$ and assumed [PRD07] that only finite probability ratios contribute to $\langle Z_\tau \rangle$. To check the consistency of this approach, calculations of Eq. (6.15) were repeated using a phase space reconstruction technique based on embedding sequences $(T_t, T_{t+\tau})$ with delay time τ and embedding dimension 2, which confirmed the validity of this approach (results not shown).

The averaged relative entropy $\langle Z_\tau \rangle$, while insensitive to the ramp orientation, at every given level T quantifies the imbalance between forward and backward probability fluxes of temperature trajectories (Fig. 6.10(A)). Again, irreversibility of scalar records increases with the lag τ and here tend to plateu at larger scales ($\tau > 1$).

Phase-randomized time series, by comparison, exhibit smaller values of $\langle Z_\tau \rangle$ in the inertial range. As one would expect, the excess is thus likely a direct result of the presence of scalar ramps. The presence of asymmetric patterns in temperature time traces further suggests that in the inertial range scalar turbulence is more time-irreversible than velocity, as confirmed by the larger values of $\langle Z_\tau \rangle$ at inertial scales (Fig. 6.10(B)).

Time-irreversibility of phase space trajectories was further investigated by testing if a significant difference exists between the probability distribution $p(T'_\tau|T)$ and $p(-T'_\tau|T)$. To this end, a Kolmogorov-Smirnov (KS) test was performed at the significance level 0.05. At every scale τ, results were averaged over different values of T and across runs within the same stability class. The results from the KS test confirm the picture obtained from the relative entropy measure $\langle Z_\tau \rangle$: The pdf of forward and backward temperature diverge significantly as the scale τ increases as shown in figure 6.10, panels (C) and (D). While this test does not capture the sign of the ramps, the behavior of near neutral runs exhibit some difference from the case of relevant heat flux: near neutral runs appear on average more reversible than unstable runs at the same dimensionless scale τ.

6.5 Discussion and Conclusions

In this work, statistical measures for the frequency of extreme fluctuations and the time-directional behavior of observed time series were applied to scalar turbulence in the ASL. It was demonstrated that i) the extreme value properties of the scalar markedly depend on the external forcing, and ii) scalar dynamics is characterized by time-irreversible behavior at the scales of injection of scalar variance in the turbulent flow. This time-irreversibility propagates down to the smaller scales of the flow examined here, thus carrying fingerprint of the energy injection mechanism.

It is well known that the pdfs of scalar increments develop heavier tails with decreasing scales in the inertial range when compared to their velocity counterparts. The analysis here shows that within the first two decades of the inertial subrange, this buildup of tails also carries the signature of turbulent kinetic energy generation. The direct injection of scalar variance from large scales seem to hinder any universal description of ΔT statistics within this range of scales. Instead, the pdf of $\Delta T(r)$ for ASL flows appear to be conditional on the value of ζ at scale r. This finding reinforces previous experimental results [LM09] obtained for a different type of flow (turbulent wake). In this case, the scalar injection mechanism was shown to impact higher order scaling exponents of the temperature structure functions.

This dependence on atmospheric stability regime for $p(\Delta T)$ further suggests that the topology of large eddies, and in particular the presence of ramp-cliff scalar structures, may be responsible for the scale-wise evolution of intermittency and the persistent time directionality at fine scales. This intermittency excess observed in the transition from production to inertial scales is consistent with self-amplification dynamics taking place that further excite the excess of scalar variance injected by the ramps.

However, while measures of intermittency appear to be dependent on the absolute value of ζ, i.e., on the relative magnitude of shear and buoyancy production terms (regardless on the sign of the heat flux), the analysis of asymmetry and time reversibility clearly sense the sign of the heat flux H more than the magnitude of ζ itself. This effect is arguably a product of the preferential orientation that the external temperature gradient imposes on the scalar ramp-cliffs, as explained by sweep-ejection dynamics. This hypothesis was here further tested by comparisons with synthetic time series that mimic ramp-cliff patterns observed in the scalar time series. The analysis confirmed that much of the observed time irreversibility, as well as its dependence on the sign of H, are recovered by these surrogate time series (Fig. 6.9).

The analysis of time directional properties showed that time-irreversible behavior for the scalar is stronger at the large scales of the flow where boundary conditions, and in particular the sign of H, determine the orientation and structure of the eddies. At finer scales, time irreversibility as quantified by both $\langle Z_\tau \rangle$ and R_τ progressively decreases as advection destroys the preferential eddy orientation imposed by boundary conditions. Note that this behavior is not captured by a simple measure of skewness such as S_T^3 (Fig. 6.9(A)), which is small at large scales and plateaus in the inertial range consistent with previous experiments [War00] and numerical simulations [CLMV00], thus showing that local isotropy is not fully attained at the finer scales examined here.

Turbulent flows exist in a state far from thermodynamic equilibrium, with the flow statistics exhibiting irreversibility. This irreversibility is typically described in terms of fluxes of energy or asymmetries in the pdfs of the fluid velocity increments [Fal09]. Similar methods could be used to describe irreversibility in the scalar field, e.g. using S_T^3, and this would imply that the irreversibility of the scalar field is stronger at smaller scales than it is at larger scales. However, in this paper I have used alternative measures to quantify the irreversibility, namely $\langle Z_\tau \rangle$ and R_τ. These quantities paint a different picture, namely that it is the largest scales, not the smallest (inertial) scales in the scalar field that exhibit the strongest irreversibility. A potential cause for these differing behaviors is that whereas fluxes and quantities such as S_T^3 are multi-point, single-time quantities, $\langle Z_\tau \rangle$ and R_τ are single-point, multi-time quantities. Thus, these two ways of describing irreversibility provide different perspectives about the nature of irreversibility in turbulence, which involves fields that evolve in both space and time. This difference in perspectives is a topic for future inquiry.

Collectively, the results presented in this paper suggest the following picture for ASL turbulence at the cross-over from production to inertial. Increasing instability in the ASL leads to increases in the mean turbulent kinetic energy dissipation rate (as evidenced by Eq. (6.1)) and its spatial autocorrelation function and pdf. The consequences of this increased dissipation with increased instability has different outcomes for velocity and scalar turbulence. For velocity, refinements to K41 appear sufficient to explain the observed scaling in the inertial subrange. For scalar turbulence, the picture appears more complicated. Intermittency buildup with decreasing (inertial) scales is more rapid when compared to their velocity counterparts, and the signature of the temperature variance injection mechanism persists at even the finer scales explored here.

Turbulence and scalar turbulence are characterized by a constant flux of energy and scalar variance from the scales of production down to dissipation. While early theories hypothesized a cascade only depending on these quantities, experimental evidence to date supports a more complicated picture. The multi-time information encoded in $\langle Z_\tau \rangle$ reveal that time-reversibility is not constant across scales, as do the fluxes of information entropy. Probability fluxes forward and backward in time are not balanced in general for air temperature increments, especially at the cross-over from production to inertial. Furthermore, these fluxes carry the signature of the external boundary conditions (i.e. H) and show that dissipation rates themselves are not independent of the large-scale dynamics. Although a formal analogy between Eq. (6.15) and the thermodynamics of microscopic non-equilibrium steady state systems exists, I stress that in the present application turbulent fluctuations are macroscopic and are the result of non-linear and non-local interactions.

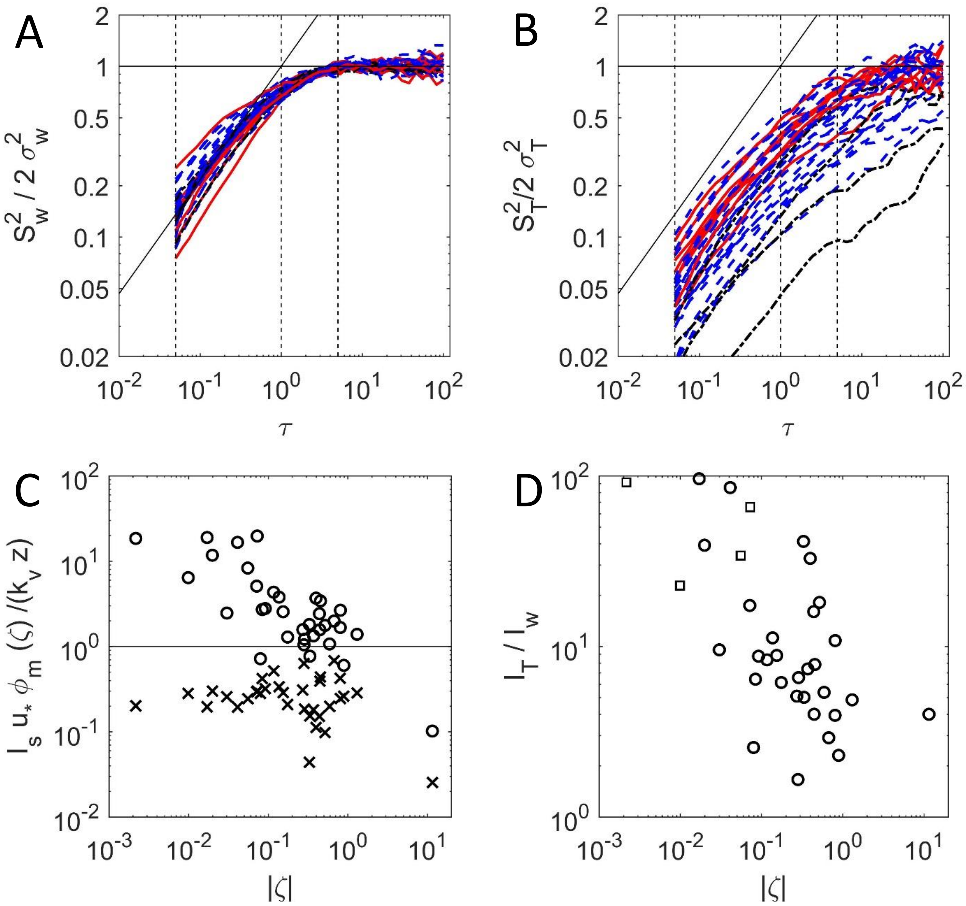


Figure 6.1: In the upper panels, the normalized second order structure functions for vertical velocity (A) and temperature (B) are shown for runs that are weakly unstable (blue dashed lines), strongly unstable (red lines), and stable (black dash-dot lines). Black lines indicate the value 1 and the 2/3 power law for reference; vertical dashed lines correspond to the dimensionless scales $\tau = 0.05$ (smallest scale not impacted by instrument path length), $\tau = 1$ (integral scale of the flow), and $\tau = 5$ (typical scale larger than I_w, while small enough not to be impacted by statistical convergence issues in structure functions calculations). Lower panels illustrate (C) the integral scales of the flow for $s = T$ (circles) and $s = w$ (crosses) as a function of the stability parameter $|\zeta|$, and (D) their ratio I_T to I_w again as a function of $|\zeta|$, where stable runs ($\zeta > 0$) are indicated by black squares.

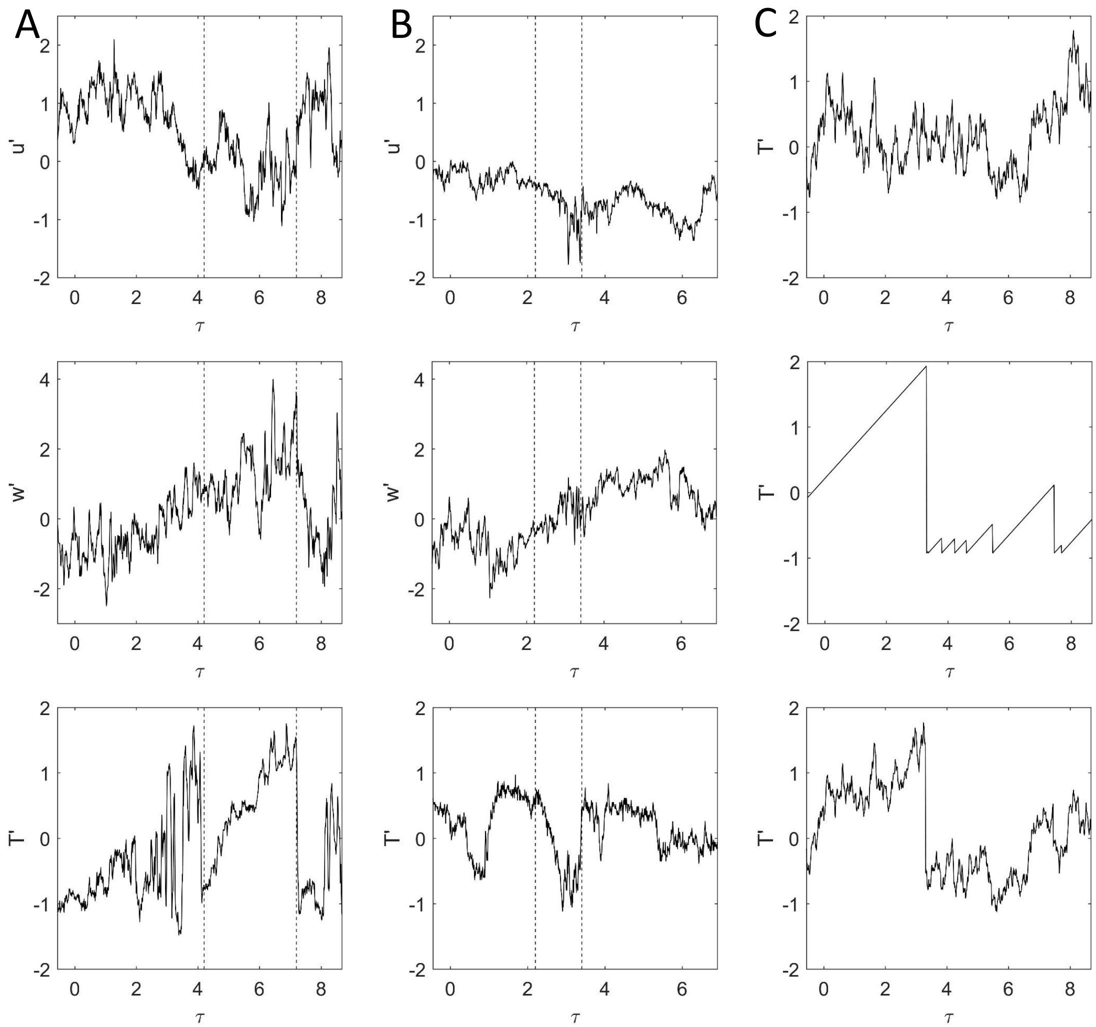


Figure 6.2: Sequences of velocity and temperature fluctuations extracted from a strongly unstable run (run 8, $\zeta = -0.52$, $I_w = 1.74s$, column A) and a stable/near neutral one (run 34, $\zeta = 0.07$, $I_w = 2.18s$ column B). The presence of ramps and inverted-ramp like structures respectively is marked by dashed vertical lines. Column (C) illustrates a phase-randomized sequence obtained from run 34 (top), a series of synthetic ramps with durations exponentially distributed with mean $2I_w$ (middle) and the surrogate time series obtained merging the above sawtooth pattern with the phase randomized time series (bottom), where the relative weight of the ramps α was set equal to 0.5.

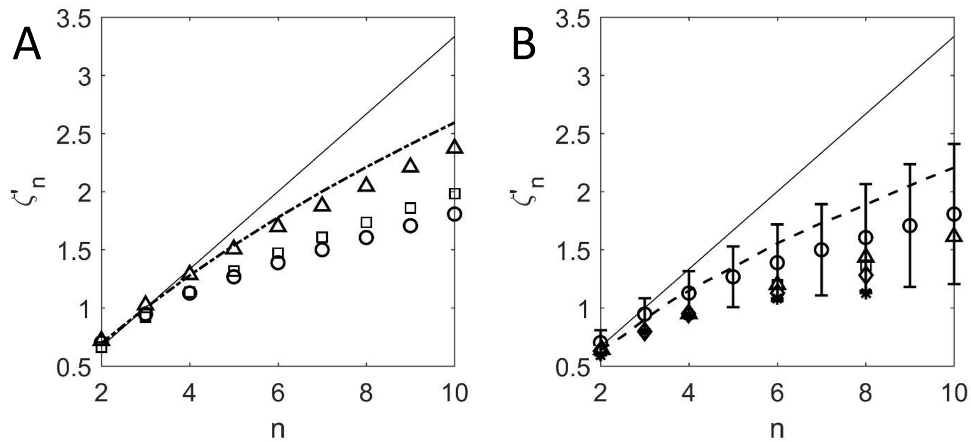


Figure 6.3: (A) Average values of the scaling exponents for longitudinal velocity u (triangles), vertical velocity w (squares), and temperature T (circles). Black continuous line and dashed line show respectively the K41 and the She-leveque predictions for the longitudinal velocity structure functions. Exponents are computed from scales ranging between $\tau = 0.05$ and $\tau = 0.2$. (B) Scaling exponents for temperature only; Mean and standard deviation values are computed over all the runs and are indicated by circles and vertical bars, respectively. Data from Mydlarsky and Warhaft (1990) [MW98] (squares), Antonia et al. (1984) [AHGA84] (triangles), Meneveau et al. (1990) [MSKF90] (*) and Ruiz et al. (1996) (diamonds) [RCBC96] are shown for comparison. The KOC scaling (black line) and results from the Kraichnan model (1994) [Kra94] (dashed line) as reported in [War00] are also presented for reference.

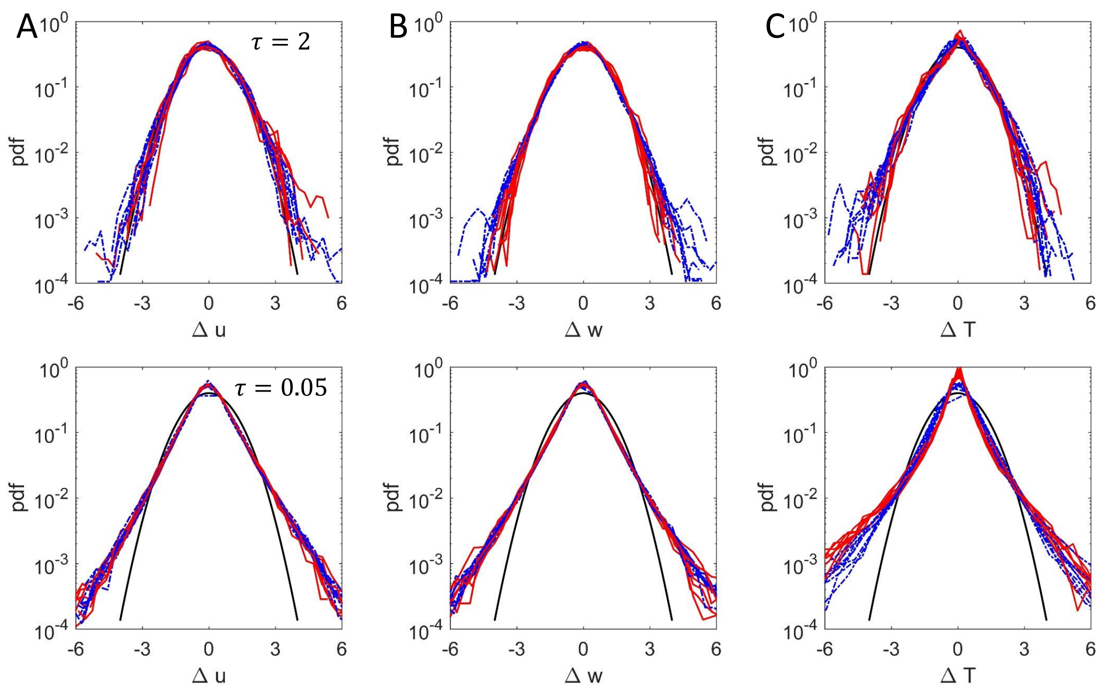


Figure 6.4: Normalized probability density functions observed for increments of longitudinal velocity (A), vertical velocity (B) and air temperature (C) at large scales ($\tau = 2$, top panels) and small scales ($\tau = 0.05$, lower panels). The figure includes data from runs in the strongly unstable class ($\zeta < -0.5$, shown in red), and near neutral class ($|\zeta| < 0.072$, blue). Black lines show the standard Gaussian distribution for reference.

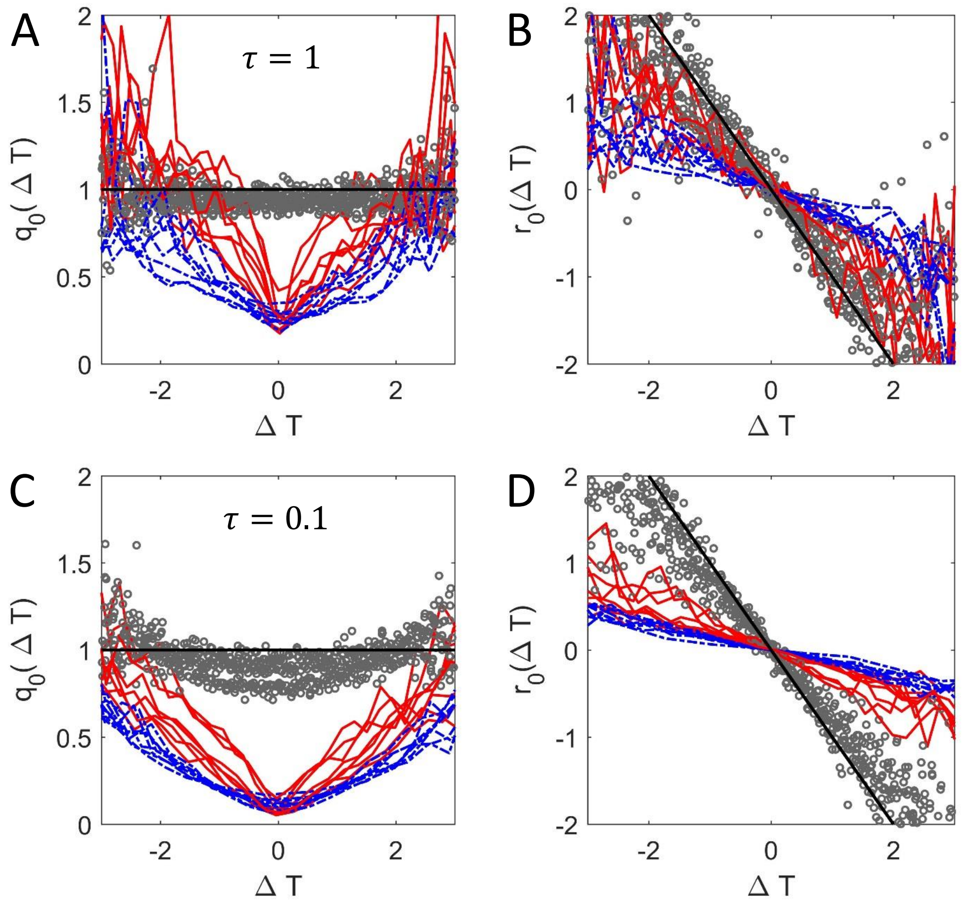


Figure 6.5: Functions $q_0(\Delta T)$ and $r_0(\Delta T)$ estimated from the conditional derivatives of the original temperature time series, for the two classes of strongly unstable (red lines) and near neutral runs (blue dashed lines). The same quantities are reported for phase-randomized surrogate time series for comparison (grey circles). Results are shown for the central body of the pdf (within 3σ from the mean) for illustration purposes. Top panels (A,B) are computed for a lag equal to the integral time scale of the flow $\tau = 1$, while the bottom panels (C,D) correspond to the smaller time lag $\tau = 0.1$. Black lines $q_0 = 1$ and $r_0 = -\Delta T$ correspond to the standard Gaussian distribution.

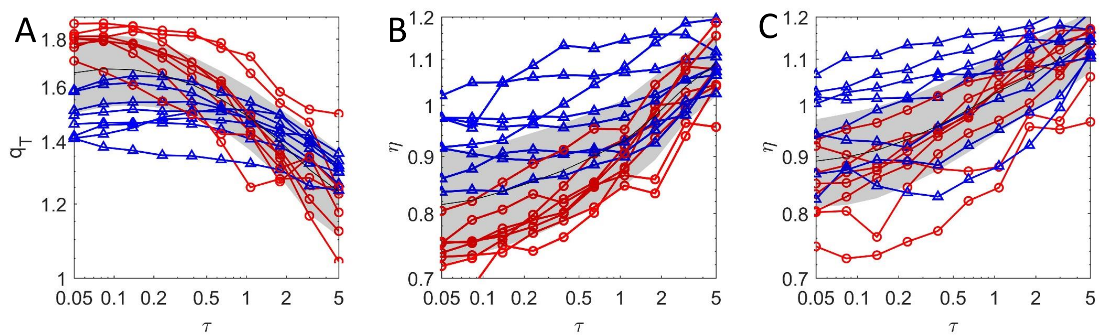


Figure 6.6: Evolution across scales τ of the q-Gaussian tail parameter q (A), and of the stretched exponential shape parameter η obtained from separate fit to the left (B) and right (C) tails of the distribution of temperature increments. Data from two stability classes are included: strongly unstable ($\zeta < -0.5$, red cirles) and near neutral conditions ($|\zeta| < 0.072$, blue triangles). Black lines and shaded areas indicate average values and standard deviations respectively computed over the entire dataset.

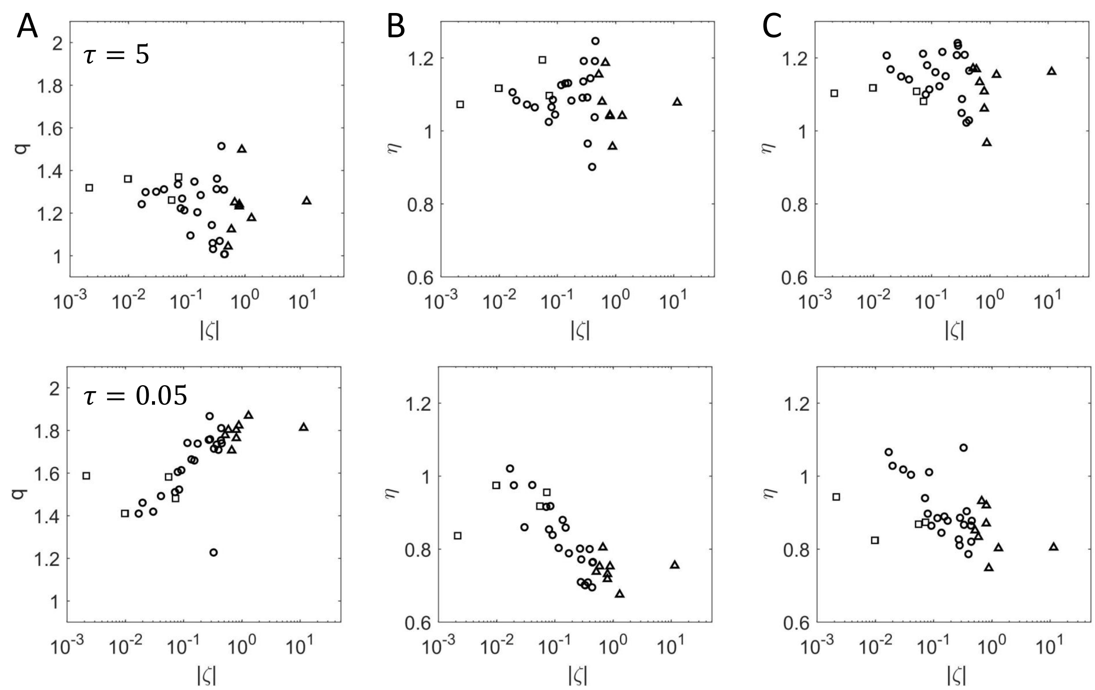


Figure 6.7: Tail parameters of the pdf of temperature increments across stability conditions ζ. Results include the q-Gaussian tail parameter q (column A) and the stretched exponential shape parameter η, obtained from fitting the left (column B) and right tail (column C) of the distribution $p(\Delta T)$. Values of q and η are reported for large scales ($\tau = 5$, upper panels) and small scales ($\tau = 0.05$, lower panels). Triangles denote strongly unstable runs ($\zeta < -0.5$), squares denote stable runs ($\zeta > 0$) and circles refer to slightly unstable runs ($-0.5 < \zeta < 0$).

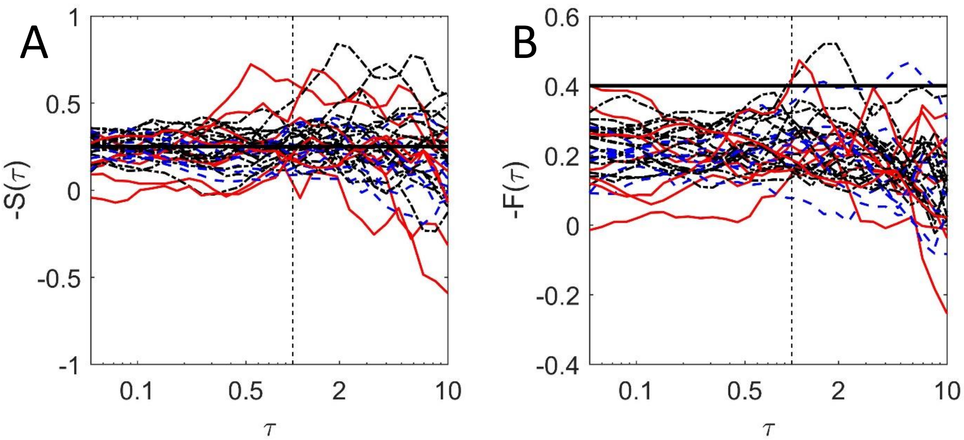


Figure 6.8: Normalized third order structure functions $S(\tau)$ and $F(\tau)$ at the crossover from inertial to production scales. Vertical dashed line indicates the integral time scales, horizontal lines show the constant values 0.25 (A) and 0.4 (B). Results are shown for near neutral runs ($|\zeta| < 0.072$, blue dashed lines), strongly unstable runs ($|\zeta| > 0.5$, red lines), and runs with intermediate values of $|\zeta|$ (black dash-dot lines).

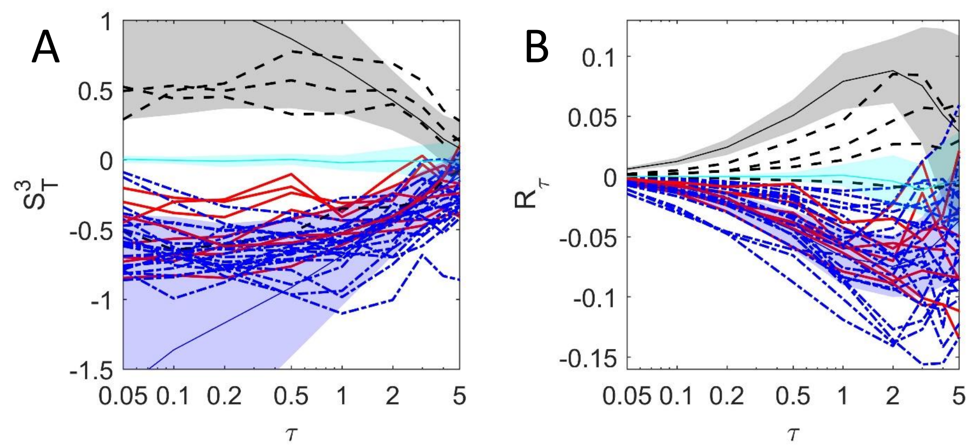


Figure 6.9: Measures of asymmetry S_T^3 (A) and time irreversibility R_τ (B) computed for temperature increments for scales varying from $\tau = 0.05$ to $\tau = 5$. The plots include stable runs (black dashed lines), weakly unstable runs (blue dash-dot lines) and strongly unstable runs (red lines). For reference, the same quantities are computed for phase-randomized time series (cyan), and sythetic time series with sawtooth positive (blue) and inverted ramps (black). Shaded regions correspond to the 1σ-confidence intervals over 34 realizations of the surrogate time series. Relative weight and mean duration of the synthetic ramps were set to $\alpha = 0.4$ and $2I_w$ respectively.

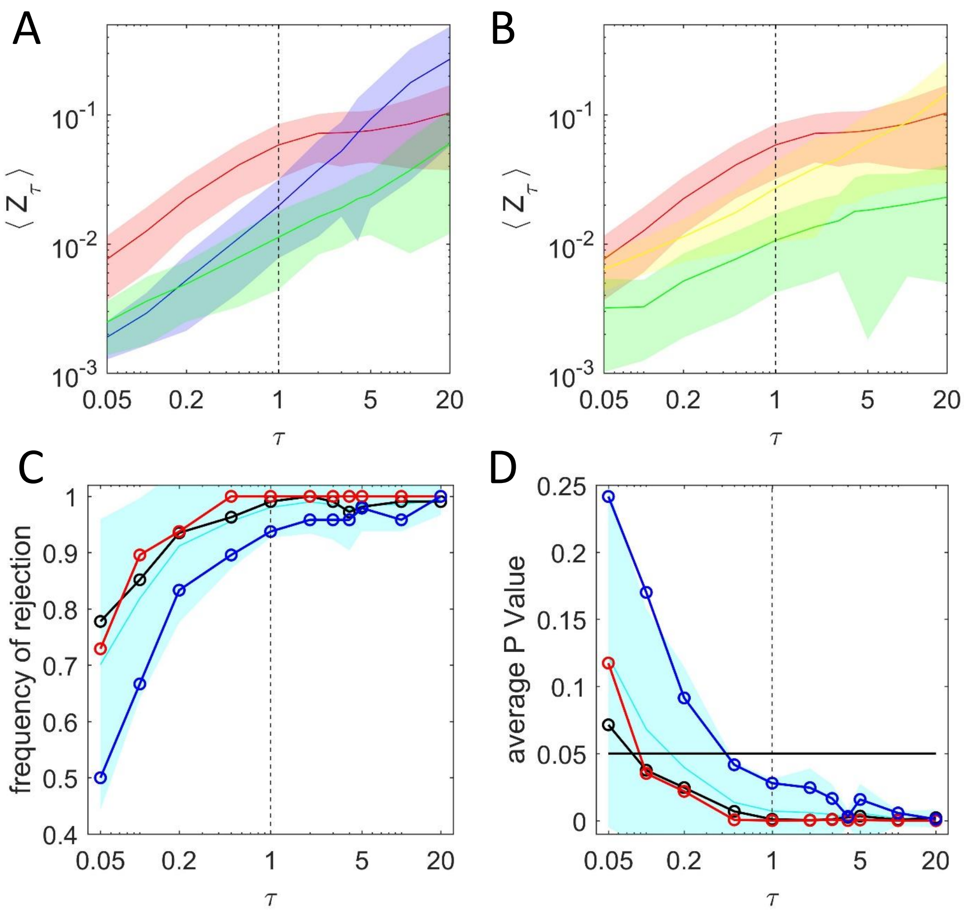


Figure 6.10: (A) Mean and standard deviation over 34 time series of $\langle Z_\tau \rangle$ computed for scales varying from $\tau = 0.05$ to $\tau = 20$. Values of $\langle Z_\tau \rangle$ are shown for original temperature records (red), and surrogate time series obtained by phase-randomization (green). For comparison, the same analysis is reported for fractional brownian motion with Hurst exponent $H = 1/3$ (blue). (B) A comparison of $\langle Z_\tau \rangle$ for temperature (red), longitudinal velocity (yellow) and vertical velocity (green). The lower panel shows the Kolmogorov-Smirnov test average rejection rate (C) and average P-value (D) computed for all the temperature time series (cyan for mean value and 1σ confidence interval), and for different stability classes: strongly unstable runs ($\zeta < -0.5$, red), near-neutral runs ($|\zeta| < 0.072$, blue) and intermediate values ($0.072 < |\zeta| < 0.5$, black). KS test was performed at the 0.05 significance level, corresponding to the horizontal line in (D). The vertical dashed line marks the integral time scale I_w.

Chapter 7

Intermittent Surface Renewals and Methane Hotspots over Natural Wetlands

Adapted from: Zorzetto, E., Peltola, O., Grönholm, T., and G. G. Katul. "*Intermittent Surface Renewals and Methane Hotspots over Natural Wetlands.*" Manuscript in preparation.

7.1 Introduction

Methane (CH_4) fluxes are emitted in many types of natural and managed ecosystems including wetlands, peatlands, and rice fields [BH93, Aul01, CSY20], but their magnitude is highly variable [Lai09, BCC$^+$04, HMD$^+$14, KJP$^+$19]. The CH_4 fluxes are conventionally measured using gas chambers positioned at the ground surface [BLG$^+$13, BMR93, MDYB$^+$11] or by means of micrometeorological methods such as the eddy covariance technique [BDS$^+$12, HFRS11, RRP$^+$07, HMD$^+$14]. The release of methane from the ground occurs via three main mechanisms: Molecular diffusion through the saturated soil surface, diffusion through the aerenchymatous tissue in plants [RKL$^+$20], and release of methane through bubbles (ebullition). The latter is by far the more variable mechanism, as the gas bubbles are emitted sporadically and can lead to significant mass transfer rates that are localized in time and in space. Because of the co-occurrence of these different sources at the ground, the overall CH_4 fluxes in these environments can exhibit a marked variability in space and time. This peculiarity can cause significant difficulties in measuring CH_4 exchange with conventional gas chambers covering a small fraction of the ground area. Rapid advancements in gas analyzers have been made over the past two decades, allowing for direct measurements of methane turbulent fluxes above the ground using the eddy-covariance technique. These measurements are able to resolve methane fluxes on sub-hourly time scale over a sufficiently large 'footprint', but on the other hand do not distinguish between ebullition, plant-mediated and background emissions. The spatially intermittent source characterizing CH_4 fluxes has been observed to lead to scalar concentration fluctuations that are much more 'heavy-tailed' than their carbon dioxide (CO_2), water vapor (H_2O), or air temperature counterparts even above ecosystems that appear to be reasonably uniform [PMH$^+$13, KPG$^+$18]. These differences in concentration statistics may be exploited to progress on partitioning measured CH_4 fluxes into a background and an intermittent ("*hotspot*") contribution. A partition methodology based on the wavelet transform was recently proposed [IHT$^+$18] to partition CH_4 fluxes based on the scalar similarity with a reference scalar (water vapor). Here, I apply a similar technique to characterize the fraction of flux, scalar variance, and footprint area that carry the signature of localized CH_4 hotspots. Then, based on these quantities, I extend the surface renewal (SR) theory [Hig35, Dan51] to characterize mass exchange for scalars such as CH_4 that are characterized by an intermittent source at the ground. The SR theory has a long history in atmospheric sciences and has been applied to turbulent transfer of heat and water vapor [Bru65, Bru75, CFC96, HO13, HO15b, HO15a, KL17] assuming that Kolmogorov size eddies dominate interfacial transport. However, prior models based on continuous renewal of the surface are not directly applicable to characterizing CH_4 emissions originating from intermittent sources at the ground. In fact, as early as 1961, it was already pointed out that the presence of zero-contact times (no renewal) can have a significant effect on estimated fluxes, and in particular larger than the effect of the specific distribution of contact time used [Per61]. It is precisely this intuition that is to be exploited here to model CH_4 interfacial mass transfer. That is, parcels of air may be continually refreshed via ejections and sweeps, but the absence of the 'hotspot' source may be viewed as a no-renewal time when compared to other scalars such as H_2O or CO_2. To be clear, the SR method developed here must be regarded as a pragmatic approach to partitioning measured CH_4 fluxes into their background component and the contribution from spatially intermittent "*hotspots*". I expect these hotspots to be statistically related to the occurrence of ebullition events, or potentially to spatial variability in the distribution of plants and microbial activity. These hypotheses are examined by relating the partitioned CH_4 fluxes to environmental variables previously documented as controlling gas exchange as seasonality, atmospheric pressure and water table level. Hence, with imminent proliferation of methane eddy-covariance flux measurements worldwide, the proposed SR scheme can be complementary to other measurement approaches and modeling efforts aimed at quantifying the magnitude and intermittency of CH_4 fluxes across various ecosystems and environmental conditions. The chapter is organized as follows: In Section 7.2 key definitions and a brief review of turbulent scalar exchange in the atmospheric surface layer are provided with a focus on the classic SR theory (section 7.2.1) and its limitations in describing fluxes of scalar quantities characterized by intermittent and sparse sources or sinks at the ground. Section 7.3 describes the experimental setup and data used. Section 7.4 describes the partition scheme based on the wavelet decomposition, which is employed for partitioning scalar fluxes and scalar variances into contributions connected to background or hotspot fluxes respectively. Based on this partition, Section 7.5 introduces the SR scheme used to model scalar fluxes for the setup here. The classic SR is extended to describe scalars such as CH_4 that are characterized by spatially and temporally inhomogeneous fluxes. In section 7.6, implications of the findings are discussed with a focus on the frequency and the transport efficiency of the methane hotspot fluxes as well as an investigation of their potential dependence on the aforementioned slowly evolving environmental variables.

7.2 Turbulent mass transfer and surface renewal theory

Throughout, x, y, z and the corresponding velocity components u, v, and w are respectively defined with x being aligned along the direction of the mean wind, z the vertical direction with the ground set to $z = 0$, and y the direction orthogonal to both. The transport of momentum, heat, and scalar quantities in the lower atmosphere is primarily driven by turbulent eddies generated by the interplay between shear and buoyancy forces [Wyn10, Stu12]. For stationary and planar-homogeneous flow at high Reynolds numbers, in the absence of mean subsidence and mean pressure gradients, the mean momentum and continuity equations for any scalar c reduce to

$$\frac{d\overline{u'w'}}{dz} = 0; \frac{d\overline{w'c'}}{dz} = 0, \tag{7.1}$$

where overline represents averaging over coordinates of statistical homogeneity (time in this case), s' denotes the instantaneous fluctuation from the mean of a generic variable s, so that $\overline{s'} = 0$. Here $\overline{u'w'}$ represents the net downward flux of momentum to the ground, $\overline{w'c'}$ the turbulent flux of any scalar with positive values signifying emissions and negative values signifying uptake, and c is the scalar concentration representing CH_4, H_2O, CO_2 or temperature T. Integrating these equations with respect to z leads to the so-called constant flux assumption, meaning that turbulent fluxes measured at some reference height z_r above a surface represent the emissions or uptake at the surface [LKL18, FN08]. The presence of surface heating or cooling at the ground produces thermal stratification within the atmospheric surface layer (ASL) that contributes to the production or destruction of turbulent kinetic energy (TKE) depending on the sign of the sensible heat flux $\overline{w'T'}$. Monin-Obukhov similarity theory (MOST) [MO54a, Fok06] argues that in the case of a horizontally homogeneous and stationary flow, the flow statistics only depend on one dimensionless stability parameter $\zeta = z/L_{mo}$, such that the equation for the mean velocity profile $\overline{u}\,(z)$ is

$$\frac{d\overline{u}}{dz}\frac{k_v z}{u_*} = \phi_m\,(\zeta)\,, \tag{7.2}$$

where $k_v \simeq 0.4$ is the von Kárman constant, $u_* = \sqrt{-\overline{u'w'}}$ is the friction velocity, and $\phi_m\,(\zeta)$ is a stability correction function, L_{mo} is the *Obukhov* length defined as

$$L_{mo} = -\frac{u_*^3}{k_v g \frac{\overline{w'T'}}{T_a}}, \tag{7.3}$$

where T_a is the mean virtual temperature (in K), and T' are temperature fluctuations. The well-known Businger-Dyer stability correction functions for ϕ_m can be used to determine its value [Fok06, Bru13]. Likewise, for a generic scalar quantity c, the mean concentration gradient is given as

$$\frac{d\overline{c}}{dz}\frac{k_v z}{c_*} = \phi_c\,(\zeta)\,, \tag{7.4}$$

where $c_* = \overline{w'c'}/u_*$ and $\phi_c\,(\zeta)$ is a stability correction function for a scalar described elsewhere [Bru13]. In near-neutral conditions ($\zeta \to 0$) and the usual logarithmic mean velocity and mean scalar concentration profiles are recovered with $\phi_m\,(0) = 1$. In general, eddy-covariance measurements of methane assume that measured turbulent fluxes are constant with height but that ϕ_c for methane differs from its water vapor or carbon dioxide counterpart simply due to dissimilarities in sources and sinks at the ground or entrainment from the atmosphere. In the case of methane, the source strength at the ground is statistically patchy whereas for water vapor or carbon dioxide, the ground sources or sinks are comparatively uniform (or at least better blended). Hence, the effective eddy diffusivity from MOST given by $K_t = k_v z u_* / \phi_c(\zeta)$ may differ among scalars because the same air parcels making contact with the ground and subsequently ejecting continually exchanges water vapor and carbon dioxide with the surface but less so for methane in the absence of ebullition. This assumption forms the basis of the extended surface renewal theory proposed here, where the presence of source hotspots is linked to a breakdown in the continuous surface renewal process described next.

7.2.1 Surface Renewal Theory

Surface Renewal (SR) theory has a long tradition in turbulence research and interfacial mass transfer. It was first introduced in the mid 1930s [Hig35] and has been extensively applied to numerous interfacial mass and heat transfer problems [Dan51, Han56, Per61, Bru65, GSJ04, AP91]. In its original form, mass exchange is driven by a surface concentration c_s and a background concentration of the same substance c_b. The transport of any quantity c (e.g., mass, or possibly other quantities such as heat and momentum) is carried out by the arrival of elements of fresh fluid (i.e., turbulent eddies characterized by a background concentration c_b) to the surface, where the contact with the surface concentration c_s triggers unsteady molecular diffusion of the substance (or property) c on a semi-infinite domain. The conditions imposed on this transient diffusion are $c = c_b$ at $z > 0$, $t = 0$, $c = c_b$ for large z and $t > 0$, and $c = c_s$ at $z = 0$, $t > 0$. Solving this problem for a scalar with molecular diffusivity D_m yields the following expression for the instantaneous mass flux for a surface element characterized by age t

$$F(t) = D_m \left(\frac{\partial c}{\partial z} \right)_{z=0} = \Delta c \sqrt{\frac{D_m}{\pi t}}. \tag{7.5}$$

After a certain residence time, each fluid element is ejected from the surface and immediately replaced by a fresh element with background concentration. This idealized model assumes that i) only molecular diffusion into the air parcel in contact with the surface is allowed, ii) the freshly renewed surface is characterized by a scalar concentration equal to the background concentration, and iii) the time scales of sweep and ejection of the fluid parcel are small compared to the residence times of the parcel at the surface. The distribution of contact times t on the surface depends on the characteristics of the flow. In the original Higbie model [Hig35], each fluid parcel had the same residence time τ_c (Plug Flow). The opposite situation is that studied by Danckwerts [Dan51], in which the surface renewal is completely random (i.e., the renewal rate does not depend on eddy age). This assumption has been shown to lead to an exponential distribution of surface ages $p(\tau) = (1/b) \exp(-\tau/b)$, where $1/b$ is a surface renewal rate and $p(\tau)$ is the probability density function of τ. The following expression for the overall flux can now be derived based on an exponential $p(\tau)$:

$$F_m = \int_0^\infty F(\tau) p(\tau) d\tau = \Delta c \int_0^\infty \sqrt{\frac{D_m}{\pi \tau}} \frac{1}{b} \exp\left(\frac{-\tau}{b} \right) d\tau = \Delta c \sqrt{\frac{D_m}{b}}. \tag{7.6}$$

Other functional forms have been considered for the distribution of residence times, including Gamma [Har62, BD72] (that would arise in the case in which younger surface eddies are less likely to be renewed when compared with the Danckwerts model) and Lognormal [GJH02] that is consistent with turbulence theories with internal intermittency in the dissipation rate accommodated. Several studies pointed out that the exact functional form of $p(\tau)$ is not important provided that its mean value is correct [KPH66, Per61, SL88, KL17] though others dispute this claim for water surfaces where local vertical advection can be large [KS09]. To be clear, water surfaces do not admit a 'no-slip' condition, meaning that velocity is finite at the air-water interface. This is not the case for the surfaces studied here. A large literature exists on the application of SR techniques to turbulent flows and reviewing all this literature is outside the scope here. A model based on the surface renewal of eddies characterized by residence times of the order of the Kolmogorov time scale, first proposed by Brutsaert [Bru65, Bru75, KL17, KLM19] to estimate evaporation rates is used for illustration. The SR theory has also been applied to sensible heat flux assuming that large-scale coherent structures (often referred to as temperature ramps or ramp-cliff patterns [ACFVA79, War00, ZBK18] dominate the heat transfer. A large eddy model was also proposed, recognizing that in the case of limited Reynolds numbers, scales much larger than the Kolmogorov microscale can be the primary contributors to the mass or heat transfer at the surface [FP67, LS70, THB76, KNM90, KMGV18]. The SR theory has been applied to the flux of momentum and promising results were reported from channel flow experiments [Han56, TS72, MHF92]. In the case of momentum, the role of the pressure term complicates the picture of the renewal processes when compared with passive scalar quantities, where molecular diffusion dominates the smallest scales of the flow. However, common to all these studies is the assumption that the renewal process at the surface is continuous. After presenting the experimental setup (7.3) and discussing the Wavelet partitioning scheme 7.4, in Section 7.5 this approach is generalized to the case of intermittent source at the ground, so as to tailor it to the case of CH_4 fluxes over peatlands.

7.3 Experiment

A published dataset [KPG$^+$18] collected at Siikaneva, a boreal wetland located in Southern Finland, is analyzed. Briefly, wind velocity components and virtual temperature fluctuations were measured by an ultra-sonic anemometer. Two independent gas analyzers were used to measure H_2O, CO_2 and CH_4 concentrations, respectively. All quantities were sampled at a frequency f_s =10 Hz and stored in 30-minutes records with each 30-minute record termed as a 'run'. A total of 4416 runs were recorded between June 1st, 2013, and August 31st, 2013. Measurements were carried out at a height $z_r = 2.8\,m$ from the surface. I focus on data from a specific wind sector (from 230 to 270 degrees, i.e., primarily the South-West wind direction) to avoid wind sectors in which nearby forested areas could be part of the footprint and possibly contaminate the signal from the peatland. In the wind sector used here, the terrain is a reasonably homogeneous boreal bog for several hundred meters. For each run, the turbulent time series were despiked following the method proposed in [Bro86], and recommended in [SMF$^+$16]. For CH_4, an additional check was performed: Any remaining fluctuations exceeding 60 standard deviations were removed and set equal to the median of the signal. Coordinates were double-rotated such that $\overline{u} = U$ equals the average wind speed, with $\overline{v} = 0$ and $\overline{w} = 0$. A cross correlation analysis was used to compute and correct the lag of the gas concentration time series. A stationarity check [FW96] was then applied to the fluxes of momentum $\overline{u'w'}$, temperature $\overline{w'T'}$, and water vapor $\overline{w'c'}$, with c' here denoting the fluctuations of the scalar H_2O. Runs passing the test for all three quantities and with at least $N = 2^{14}$ data points were included in the analysis. The stationarity check was not applied to the CH_4 concentration time series since the objective is precisely the characterization of CH_4 fluxes possibly connected with short term localized emission events.

Additionally, the runs in the dataset were filtered by requiring a minimum turbulence mixing ($u_* > 0.2\ m/s$). I also required that the H_2O flux be upward ($\overline{w'c'}_{H_2O} > 0$) so as to avoid runs characterized by little or no evaporation, and that the H_2O flux approximately follows the MOST variance-flux relation so as to avoid runs characterized by non-local effects and entrainment from the top of the atmospheric boundary layer as done in [IHT+18]. These conditions of H_2O were prompted by the fact that water vapor is to be used as a reference scalar to partition the CH_4 signal based on gas exchange mechanism as discussed in Section 7.4.

7.4 Partitioning methane fluxes using concentration measurements

Recent attention has been devoted to determining scalar fluxes originated from short turbulent events [SGF17] and to the partitioning of CH_4 fluxes connected with ebullition using wavelet transforms [IHT+18, SKFG19]. Building upon these previous studies, the CH_4 time series is first analyzed in the wavelet domain with the objective of partitioning turbulent fluxes into their background (B) and hotspot (H) components. This information will be used to calibrate the intermittent surface renewal model. Let c denote the concentration of the scalar of interest (e.g., methane here) and r the concentration of a reference scalar quantity, here chosen to be water vapor. If scalar similarity holds [Hil89], then the two scalar time series should be perfectly correlated and the wavelet coefficients of the normalized time series $c_n = \left(c - \overline{c}\right)/\sigma_c$ and $r_n = \left(r - \overline{r}\right)/\sigma_r$ should also be equal. To localize CH_4 hotspot events both in the time and frequency domains, the orthonormal wavelet transform [Mal89] is applied. The scalar series are decomposed using the *Haar* wavelet basis, which has been widely applied to the study of turbulent flows [Men91, KPC94, SA01]. The use of an orthonormal basis here is preferred because it does not introduce spurious information in the transformed series, and thus the scale wise dependence between the two scalar time series is preserved by the transform [KPC94, KFG97]. The *Haar* basis functions are of the form $\psi^{(m)}(x) = 2^{-m/2}\psi^{(0)}((x - 2^m i)/2^m)$ with $i, m \in Z$ the position and scale indices defining dilated and translated version of the *Haar* mother wavelet $\psi^{(0)}(x)$, defined as

$$\psi^{(0)}(x) = \begin{cases} 1 & \text{if } 0 < x < 1/2 \\ -1 & \text{if } 1/2 \leq x < 1. \\ 0 & \text{otherwise} \end{cases} \tag{7.7}$$

The wavelet coefficients $WT_f^{(m)}[i]$ can be computed by iteratively coarse-graining the generic series $f(t)$ as shown elsewhere [KPC94]

$$WT^{(m)}[i] = \int_{-\infty}^{+\infty} f(t)\, \psi^{(m)}(t - i)\, dt. \tag{7.8}$$

In the wavelet domain, the flux of a scalar quantity c can be computed from its concentration measurement time series and vertical velocity as

$$\overline{F_c} = \overline{w'c'} = \frac{1}{N} \sum_{m=1}^{M} \sum_{i=1}^{2^{M-m}} WT_w^{(m)}[i] WT_c^{(m)}[i], \tag{7.9}$$

while the variance of the scalar can be obtained as

$$\sigma_c^2 = \overline{c'c'} = \frac{1}{N} \sum_{m=1}^{M} \sum_{i=1}^{2^{M-m}} \left(WT_c^{(m)}[i]\right)^2. \tag{7.10}$$

For each N observations per unit averaging interval ($N = 2^{14} = 16384$ data points per averaging period), the original series can be fully described from the $N - 1$

wavelet coefficients $WT^m[i]$ (and the remaining coarse grained series for wavelets with support longer that the *Haar* wavelet). To exploit the similarity between the scalar of interest and a reference scalar, both series are filtered thereby excluding from the analysis small scales within the inertial subrange and large scales for which (i) scalar similarity does not necessarily hold, and (ii) the poor time localization of the wavelet decomposition at the lowest frequencies may not provide sufficient information. Therefore, I focus on the range of scales between $0.5\,T_w$ and $200\,T_w$, where T_w is the integral time scale of the vertical velocity fluctuations w'. This range of scales explains most of the variances and covariances for both CH_4 and H_2O. I define $\delta_s(m)$ as the indicator function selecting this range of scales (i.e., $\delta_s(m) = 1$ if $0.5 < 2^m/(f_s T_w) < 200$, and $\delta_s(m) = 0$ otherwise). To quantify the distance between the coefficients $WT_c^{(m)}[i]$ and the corresponding reference scalar coefficients $WT_r^{(m)}[i]$, the slope and intercept of the linear relation $WT_c^{(m)}[i]/\sigma_c = \alpha_w WT_r^{(m)}[i]/\sigma_r + \beta_w$ are first determined. The α_w and β_w are computed using a Huber regressor algorithm, which is particularly robust to the presence of outliers. The absolute differences $\Delta WT^{(m)}[i] = |WT_c^{(m)}[i]/\sigma_c - \alpha_w WT_r^{(m)}[i]/\sigma_r - \beta_w|$ between the wavelet coefficients of the two scalars $c(t)$ and $r(t)$ are then determined, and their standard deviation evaluated over the range of scales for which $\delta_s(m) = 1$:

$$\Sigma_{cr} = \left(\overline{\left(\Delta WT^{(m)}[i] - \overline{\Delta WT^{(m)}[i]}\right)^2} \right)^{1/2} \tag{7.11}$$

As expected, Σ_{cr} tends to decrease for increasing values of correlation R_{cr} between the two scalars as shown in Figure 7.1(a) for $c = CH_4$ and $r = H_2O$. In the case of perfect scalar similarity, the correlation between the two scalars should be $R_{cr} = 1$ and this difference should approach zero. Moreover, panel 7.1(b) shows that runs with larger Σ_{cr} are also characterized by CH_4 traces having a larger skewness M_{30} compared to their H_2O counterpart, as previously noted [KPG+18]. In the presence of identical sources at the ground, this should not be the case, as noted by Hill [Hil89]. This observation will be used to partition hotspot and background CH_4 coefficients. For this purpose, I select a value of Σ_{cr} that is representative of the variability between the two time series that one would experience in the absence of differences in the sources or sinks at the ground. In the dataset here, this limiting standard deviation value is taken to be $\hat{\Sigma}_{rc} = 0.3$. As shown in Figure 7.1, this threshold corresponds to runs characterized by R_{cr} closest to unity with CH_4 and H_2O exhibiting a comparable skewness. As done in the study by Iwata et al., [IHT+18], a threshold of $3\hat{\Sigma}_{rc}$ was adopted to determine whether or not any given CH_4 wavelet coefficient $WT_c^{(m)}[i]$ carries signature of localized hotspots. This partition was applied with this hard threshold to all the runs. For each run, the subset of the scale-time wavelet domain characterized by values of $\Delta WT^{(m)}[i] < 3\hat{\Sigma}_{cr}$ is characterized by a homogeneous ground source for the two scalar. Conversely, the condition for detecting hotspots is given by $\Delta WT^{(m)}[i] > 3\hat{\Sigma}_{rc}$. Denoting $\delta_x(i,m)$ as the indicator function that selects these CH_4 hotspots events only ($x = h$, equal to 1 when $\Delta WT^{(m)}[i] > 3\hat{\Sigma}_{rc}$, else 0) or background events only ($x = b$, equal to $\delta_x(i,m) = 0$ when $\Delta WT^{(m)}[i] > 3\hat{\Sigma}_{rc}$, else $\delta_x(i,m) = 1$), then CH_4 fluxes and variances can be decomposed based on the type of source at the ground exploiting similarity with water vapor concentration time series.

The fraction A^+ of time in which attached eddies identify localize hotspots is then obtained by first back-transforming the filtered series of wavelet coefficients to the time domain, using only the non-zero coefficients for the hotspot and background series respectively as selected by the indicators $\delta_x(i,m)$ and $\delta_s(m)$:

$$c_x(t) = \sum_{m=1}^{M} \sum_{i=1}^{2^{M-m}} \psi_{i,m}(i)\, WT_c^{(m)}[i]\, \delta_s(m)\, \delta_x(i,m), \tag{7.12}$$

Then, CH_4 hotspots are individuated in the time domain imposing the condition that the local power of the hotspot time series is larger than the variance of the background time series. Therefore, the fraction of time in which hotspots are active is obtained by the condition that $(c_e - median\,(c_e))^2 > \overline{c_b^2}$, i.e., whenever the squared fluctuations from the median of the hotspot concentration component (c_e) are larger than the variance of the background time series c_b. The dimensionless quantity A^+ also corresponds to the fraction of source area where integral-scale eddies carry the signature of hotspots at any given instant in time on average. This partition of the original time series in hotspot and background contributions can be used to compute the contribution of each component to the overall scalar variance and flux. The fractions of total flux and total variance of each component can be expressed (for either hotspots, $x = h$ or background, $x = b$) as

$$f_x = \frac{1}{\overline{w'c'}} \frac{1}{N} \sum_{i=1}^{N} c'_x(i) w'(i) \delta_{t,x}(i) \tag{7.13}$$

and

$$v_x = \frac{1}{\sigma_c^2} \frac{1}{N} \sum_{i=1}^{N} c'_x(i)^2 \delta_{t,x}(i) \tag{7.14}$$

where $\delta_{t,x}(i) = 1$ if the $i-th$ observation has been classified as hotspot $(x = h)$ or background $(x = b)$, and $\delta_{t,x}(i) = 0$ otherwise. As anticipated, the *active* fractional area of the footprint is

$$A^+ = \frac{1}{N} \sum_{i=1}^{N} \delta_{t,h}(i). \tag{7.15}$$

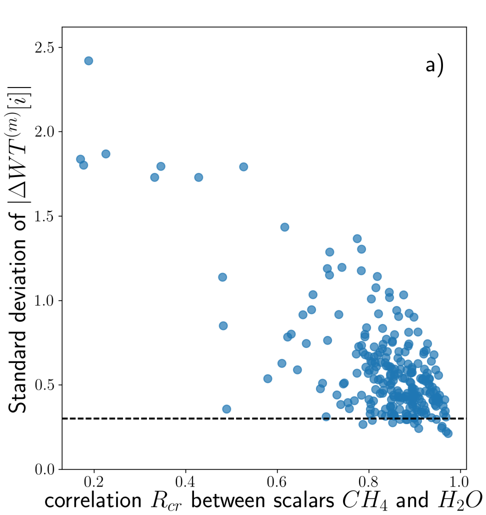


Figure 7.1: Standard deviation of the difference of wavelet coefficients (Σ_{cr}) between CH_4 and H_2O as a function of (a) the correlation R_{cr} between the two scalars, and (b) the difference in skewness M_{30} between CH_4 and H_20 for each run included in the analysis. The threshold value $\Sigma_{cr} = 0.3$ is indicated by the dashed horizontal lines.

To test this partition methodology, the moments of the joint distribution of the scalar and vertical velocity fluctuations are computed. This is done both for the original and for the reconstructed time series with background coefficients only. Previous studies have shown that 4^{th}−order cumulant expansions suffice to explain the joint pdfs of CH_4 and w [KPG+18]. Therefore, the focus is on moments M_{ij} up to the 4^{th} order where M_{ij} is defined as $\overline{c^i w^j}/\sigma_c^i \sigma_w^j$ where again σ_c, σ_w are the root-mean-squared (rms) fluctuation of a scalar c and of the vertical velocity w respectively. To characterize the statistical properties of the hotspot or background components of any scalar concentration, a partial mixed moments can be defined (for either $x = h, b$) as

$$M_{ij}^x = \frac{1}{\sigma_c^i \sigma_w^j} \frac{1}{N} \sum_{k=1}^{N} (c'(k))^i \, (w'(k))^j \, \delta_{t,x}(k). \qquad (7.16)$$

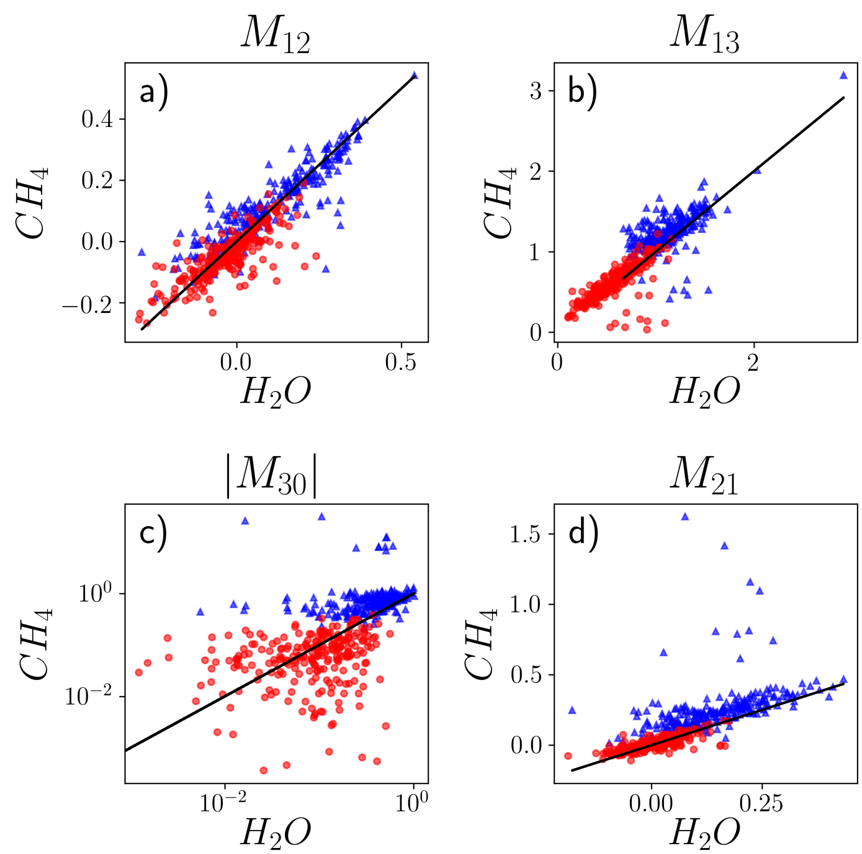


Figure 7.2: Comparison of moments of the joint distribution of scalar quantities and vertical velocity, for both the original CH_4 and H_2O time series (blue data points) and the corresponding background-only values (red data points) for each run included in the analysis.

While the partition method only uses values of the scalar wavelet coefficients, this test includes the interaction of the scalar with the vertical velocity fluctuations. Overall, this check shows as expected that the background component moments are closer to the 1:1 line compared to the original time CH_4 series, and in particular the imbalance in the scalar skewness M_{30} appears reduced (Figure 7.2). An additional check can be performed by considering the transport efficiency e_T, which quantifies the ratio between direct and indirect flux as evaluated through quadrant analysis. Here, e_T is evaluated in the time domain for hotspot and background CH_4 components as $e_T = 1 - |\overline{F}_{ind}/\overline{F}_{dir}|$ where $\overline{F}_{ind}$ and $\overline{F}_{dir}$ are the covariances evaluated by including in the sum only indirect (in the case of CH_4, downward, i.e., inner and outer interactions only) or direct (upward, i.e., sweeps and ejections only) contributions to the flux respectively [KPG+18]. The CH_4 fluxes are characterized by a larger e_T when compared to other scalars (Figure 7.3), and this behavior may be connected to the signature of localized hotspots leading to an alignment between large w' and c'. In the case of CO_2, e_T appears lower than its H_2O counterpart. Also in this case this behavior may be the direct effect of the inhomogeneity of the spatial distribution of CO_2 sources and sinks at the ground. The relatively inefficient CO_2 flux could originate from areas characterized by less dense vegetation, or from the interplay of photosynthethic activity (net sink) with the emission of bubbles also containing CO_2 from the peat, which would produce localized contributions of opposite sign. As shown in Figure 7.4, e_T exhibits a clear dependence on the atmospheric stability as quantified by ζ. Overall, e_T tends to be larger for CH_4 when compared to H_2O regardless of the atmospheric stability conditions. However, when examining separately background and hotspot components, e_T appears lower for the former, and more in line with the corresponding e_T values observed for H_2O, as shown in Figures 7.3 and 7.4. Even though the partition methodology used here only employs information on the scalar concentrations, its results appear consistent even when considering the interaction with the vertical velocity fluctuations, and in particular the scalar transport efficiency e_T. Based on these observations, an extension of surface renewal theory is now introduced for characterizing the intermittent nature of CH_4 fluxes. The extension is based on the quantities v_x, f_x, and A^+ obtained through the wavelet partition performed here.

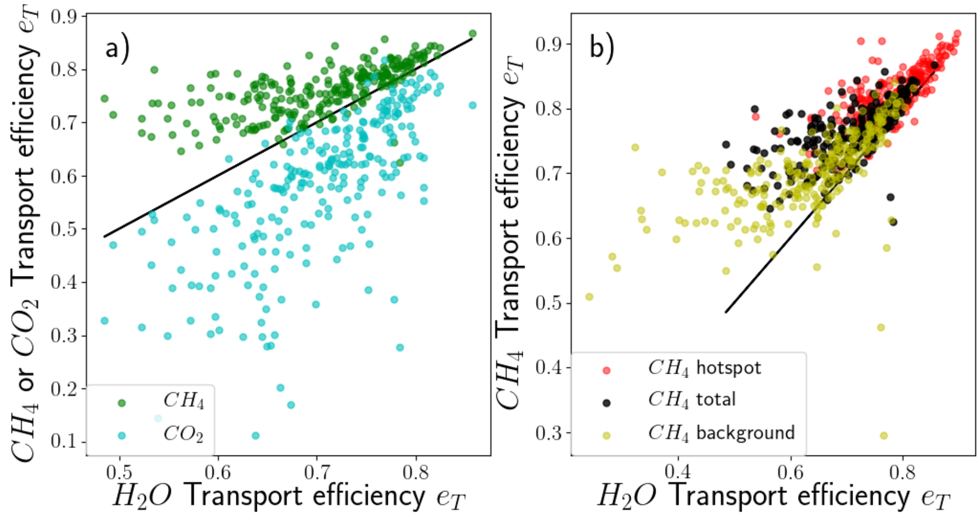


Figure 7.3: Comparison of transport efficiencies e_T for different scalars (CH_4, CO_2) with the background scalar (H_2O), and the same analysis for the partition of CH_4 in hotspot and background components.

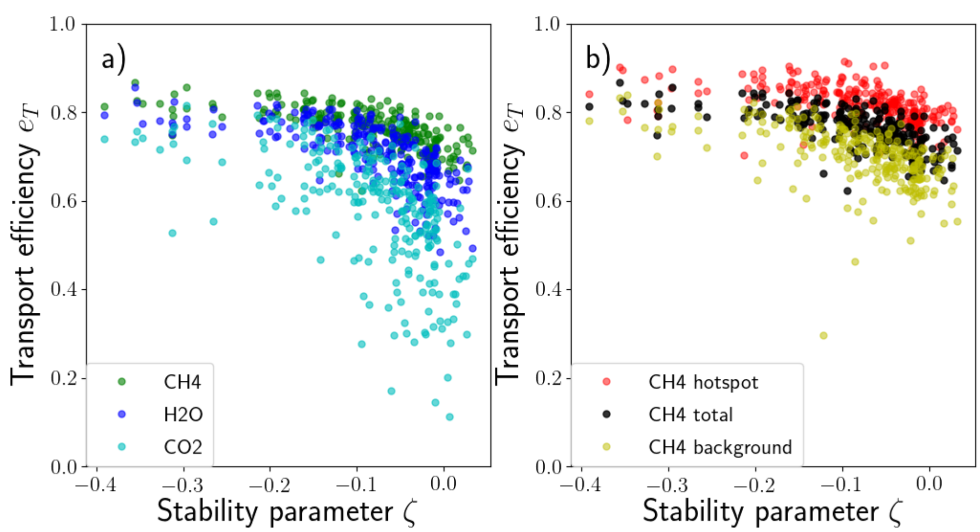


Figure 7.4: Dependence of scalar transport efficiencies e_T on the atmospheric stability parameter ζ for H_2O, CH_4 and its hotspot and background components.

7.5 A SR approach for continuous and intermittent scalar sources

The SR approach begins by assuming homogeneous scalar source at the ground as may be the case for H_2O (at least when compared to the remaining scalars). The turbulent transport can be adequately described by an SR scheme in which turbulent eddies are continuously touching the ground and are randomly renewed so that the distribution of residence times in contact with the ground is exponential [Dan51]. In conventional SR theories, the characteristic residence time is of the order of the Kolmogorov time scale so that $b = \tau_\eta = (\nu/\epsilon)^{1/2}$ where ν is the kinematic viscosity of air and ϵ is the mean turbulent kinetic energy (TKE) mean dissipation rate. For a stationary, planar-homogeneous flow in the absence of subsidence and upon neglecting turbulent transport and pressure redistribution terms [KKP11], the TKE budget is given as

$$\epsilon \simeq \frac{u_*^3}{k_v z}\phi_m\left(z/L_{mo}\right) + \frac{g}{T_a}\overline{w'T'}, \tag{7.17}$$

where the two terms on the right hand side of eq. (7.17) correspond to mechanical production and buoyancy production (or destruction, depending on the sign of L_{mo}) of TKE. Atmospheric surface layer flows over peatlands are characterized by high $Re_* = z_r u_*/\nu$ and a dynamically rough surface (i.e. z_o is independent of Re_*). Therefore, ϵ was computed by evaluating the balance in eq. (7.17) at $z = h$, with $h = 7.5 z_0$ [Bru13] and z_0 computed from the mean velocity profile as $z_0 = z_r e^{-k_v U/u_*}$, where U is the mean longitudinal velocity measured at a distance z_r from the ground ($z_r = 2.8\,m$ in the experimental setup here). The sign of the flux is not predicted by the SR model. Therefore, the skewness of the flux time series, its time directionality properties [ZBK18], or simply the sign of the eddy correlation flux (when this information is available) can be used. Here, the analysis is restricted to the case of positive (i.e., upward) fluxes (H_2O and CH_4). When comparing results with $c = T$ or $c = CO_2$ time series, the sign of $\overline{w'c'}$ will be used. In field studies of methane, the source strength at the ground Δc is not usually known. Here, I estimate Δc as a function of the energy and skewness of the scalar fluctuations as $\Delta c = k_{sr}\sigma_c^{(I)} a_{sr,c} = k_{sr}\sigma_c^{(I)}\left(1 + |M_{30,c}|^{1/3}\right)$, where $M_{30,c}$ is the skewness of the scalar time series, and $\sigma_c^{(I)}$ the square root of the spectral energy density of the scalar c integrated over a range of scales around the integral time scale T_w (here set to the range of scales between 0.5 T_w and 5 T_w due to the limited frequency localization of the *Haar* decomposition). I found empirically that the proportionality constant $k_{sr} \approx 4$ applies for all scalar quantities examined here. Based on previous results obtained for temperature, the quantity $k_{sr}\sigma_c^{(I)}$ may be interpreted as a difference between the maximum and minimum concentration fluctuations expected in a run [FB73], as driven by the difference between source at the ground and background concentrations. Since the scalar variance varies with turbulent intensity $u*$, the correction factor $a_{sr,c}$ accounts for the possible asymmetry in the scalar time series, which is a statistical quantity expected to retain information on the injection of scalar variance in a turbulent flow [War00, ZBK18]. With these assumptions, the following expression for scalar fluxes can be derived

$$v_g = \frac{\overline{w'c'}}{\Delta c} = \frac{\overline{w'c'}}{k_{sr}\sigma_c^{(I)} a_{sr,c}} = \sqrt{\frac{D_m}{\tau_\eta}} = Sc^{-1/2}(\nu\epsilon)^{1/4}, \tag{7.18}$$

where v_g is a gas transfer velocity and $Sc = \nu/D_m$ is the molecular Schmidt number. The v_g emerging from this analysis is, as expected, consistent with the micro-eddy model given by $v_g \propto [Sc^{-1/2}(\nu\epsilon)^{1/4}]$ [KMGV18], where $(\nu\epsilon)^{1/4}$ is the Kolmogorov velocity. The $v_g \sim Sc^{-1/2}$ has received wide support from experiments and direct numerical simulations [TKKK16]. The D_m used here for different scalar quantities are reported in table 7.1 for convenience. Unlike water, diffusion of scalars in the atmosphere leads to an Sc close to unity and is suggestive that Sc adjustments to v_g are not as crucial for the scalars of interest here. The surface renewal scheme proposed was then applied to different scalars: methane, water vapor, temperature and longitudinal velocity to further test the robustness of the approach. Figure 7.5 shows a comparison of this SR scheme with Eddy covariance flux estimates for all 4 quantities. The method, after calibration of a single constant k_{sr} (which is the same for each scalar and constant across all runs) appears to reasonably reproduce all scalars, and surprisingly momentum fluxes (as noted in prior studies by [Han56]). Note however that the SR is not fully prognostic when applied to momentum, since u^* is needed in the TKE budget to evaluate the Kolmogorov time scale. For CH_4 SR and EC flux estimates exhibit a larger scatter, and the presence of some very large outliers hints at the possible differences in ground sources when compared to the other scalars, which will be examined next.

7.5.1 Extension to intermittent surface renewals

Since the complete SR theory requires the description of small-scale quantities at the Kolmorogov scale and their spatial distribution, it is not feasible to reconstruct them completely using conventional eddy covariance measurements. The current eddy-covariance measurements simply do not resolve such fine scales due to instrument separation and volume-averaging. To progress on an SR theory that is mindful of such instrument limitations, a simplified intermittent model that will be based on surface renewal micro eddy model (i.e., the bottleneck in the transport is dominated by kolmogorov scale eddies in the vicinity of the surface) while the intermittency of the process is described at the integral scale - i.e., at the scales of eddies touching the ground, which may carry - or not - the signature of one or more CH_4 hotspots events.

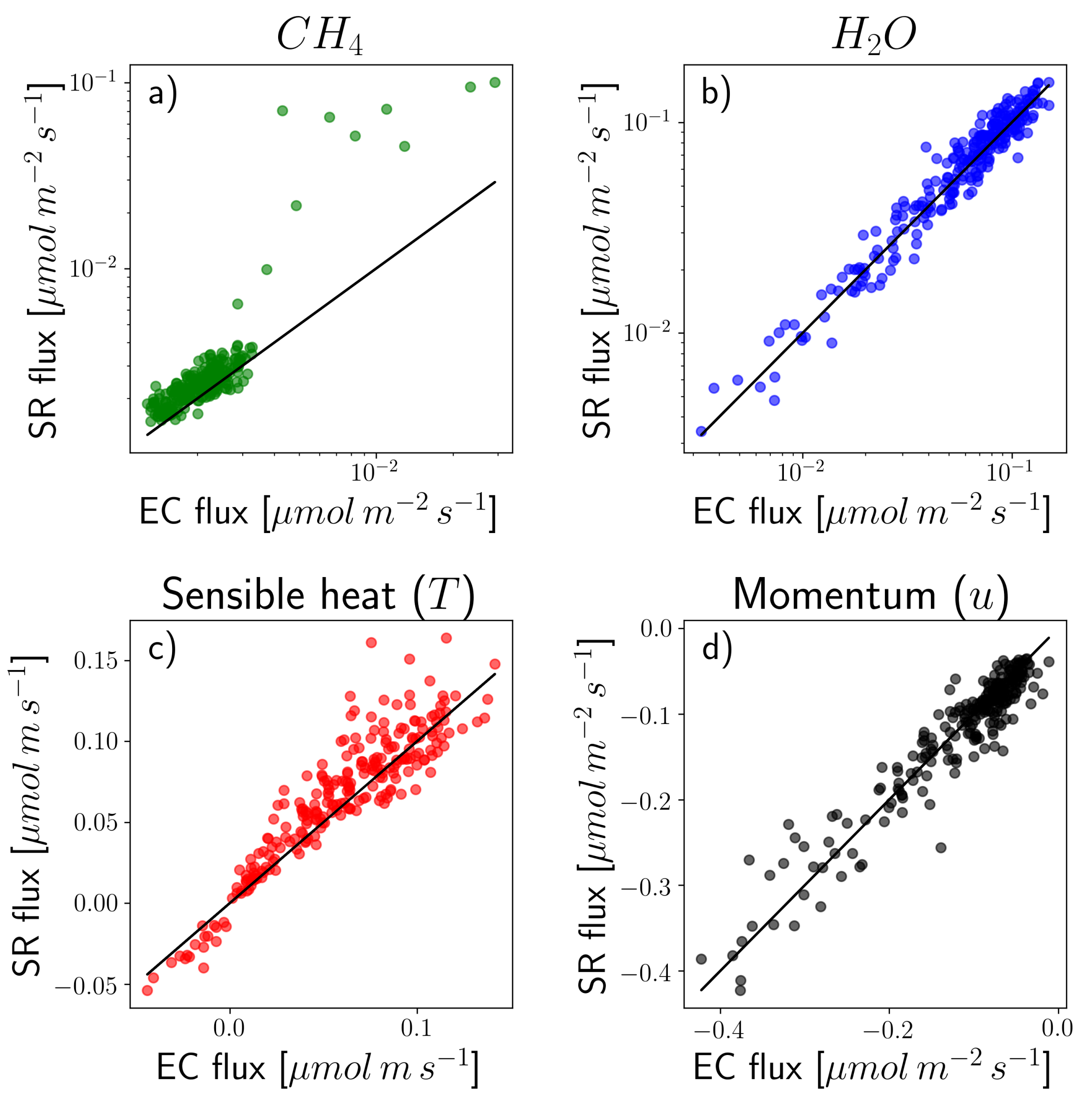


Figure 7.5: Comparison between eddy covariance (EC) and surface renewal (SR) fluxes for methane, sensible and latent heat fluxes, and (kinematic) momentum flux.

Given this separation of scales relevant for the overall scalar fluxes, the large-scale depiction of the intermittent source does not directly impact the shape of the distribution of eddy ages at the ground, and thus the scaling of the fluxes with Re_*. As is the case for the surface renewal micro-eddy model, both hotspot and background diffusive fluxes should scale as $Sc^{-1/2}Re_*^{-1/4}$. However, since the relative contributions of the two mechanisms is expected to change with environmental conditions, the overall flux may diverge from the $Re_*^{-1/4}$ scaling as anticipated from the larger scatter for CH_4 flux estimates when compared to other scalars (Figure 7.5). In this formulation, the overall flux is given by $\overline{F} = (1 - A^+)\overline{F_b} + (A^+)\overline{F_h}$ i.e., is a weighted average of an hotspot flux $\overline{F_h}$ and a background flux $\overline{F_b}$, each weighted by the respective fractional contributing area of the footprint. Each of these two components has expression of the type derived in eq. (7.18), but the two mass transfers are driven by different source strengths denoted as Δc_b or Δc_h for the background and hotspot components respectively. Since the source strength is proportional to the square root of the scalar energy (over a given range of scales around T_w), one obtains $\Delta c_b = k_{sr} a_{sr,b} \sigma_b^{(I)}/\sqrt{1 - A^+}$ and $\Delta c_h = k_{sr} a_{sr,h} \sigma_h^{(I)}/\sqrt{A^+}$, taking $\sigma_b^{(I)} = v_b \cdot \sigma_c^{(I)}$ and $\sigma_h^{(I)} = v_e \cdot \sigma_h^{(I)}$. Analogously to $a_{sr,c}$, the quantities $a_{sr,h}$ and $a_{sr,b}$ are evaluated for the two components using the skewness of the hotspot and background components alone as obtained following the wavelet partitioning scheme discussed in section 7.4. Here, k_{sr} has the same value determined for the non-intermittent SR, and again assumes a constant value across different runs and different scalars. Therefore, the overall flux can be expressed as

$$\overline{F} = \left(\sqrt{1 - A^+} + \sqrt{A^+} \frac{v_h}{v_b} \frac{a_{sr,h}}{a_{sr,b}} \right) k_{sr} v_b a_{sr,b} \sigma_c^{(I)} \sqrt{\frac{D_m}{\tau_\eta}}, \tag{7.19}$$

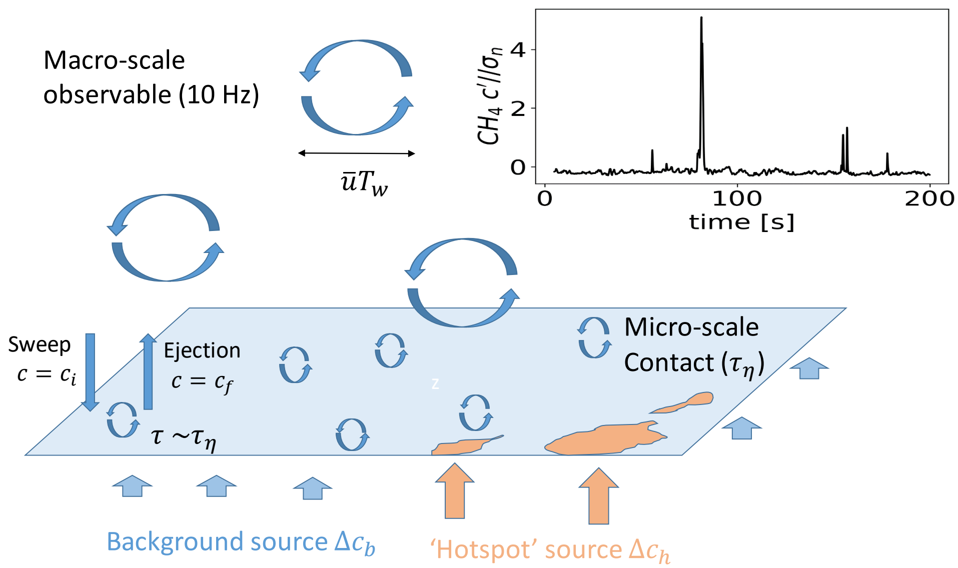


Figure 7.6: Scheme of the transport mechanism for CH_4 near the surface. The figure represents an idealized patchy CH_4 source at the ground, emphasizing the small scale η of eddies dominating the interfacial gas transfer, and the integral scale $u \cdot T_w$ of the vertical velocity representing the large scales 'resolved' by the tower measurements, which carry the integrated signature of the patchy CH_4 source at the ground.

where A^+ is the average fractional area characterized by integral-scale eddies carrying the signature of CH_4 hotspots. These quantities (A^+, Δc_h and Δc_b) are all large-scale quantities and do not correspond to the Kolmogorov-scale quantities governing the mass exchange at the interface, but rather are linked to their integral-scale effect as captured by the scalar concentration measurement time series. For a schematic representation of the intermittent surface renewal scheme featuring the range of scales involved, see Figure 7.6. Scalar fluxes over a rough surface for near neutral atmospheric conditions as described by a micro-eddy type model [Bru75] can be expressed in term of a dimensionless *Dalton* number $Da = F/(u_* \Delta c)$ as $Da = Sc^{-1/2} Re_*^{-1/4}$. Similarly, a dimensionless expression for the intermittent surface renewal fluxes can be derived here as

$$Da_b = \left(\sqrt{1 - A^+} + \sqrt{A^+}E^+\right) Sc^{-1/2} Re_*^{-1/4}, \tag{7.20}$$

where two additional dimensionless numbers are now required describing the interplay of diffusive and hotspot contributions to the overall CH_4 fluxes. Again, the number A^+ represents the average fraction of area (or time) characterized by hotspot source (at the integral time scale); $E^+ = (a_{sr,h}v_h)/(v_b a_{sr,b})$ represents the relative strength of ebullition sources at the ground with respect to the concentration gradient driving the diffusion process, and $Da_b = F/(u_* \Delta c_b)$. The results of the CH_4 flux wavelet partition and its comparison with the (intermittent) surface renewal fluxes are reported in Figure 7.7. The overall CH_4 fluxes exhibit a larger scatter when compared with H_2O, suggesting that the transport efficiency is more variable as determined by the interplay of hotspot and background fluxes. Comparing separately the hotspot and background components to the fluxes with their counterpart obtained through the wavelet partition scheme described in section 7.4, again one can see that a micro-model type SR scheme is appropriate for each component separately even though a larger scatter is present for the hotspot flux, which is responsible for the run-to-run variability observed for the overall CH_4 flux. In particular, the largest outliers observed for the overall flux are predominantly characterized by hotspot-type fluxes, as captured both by the wavelet partition scheme and the ISR-estimated flux.

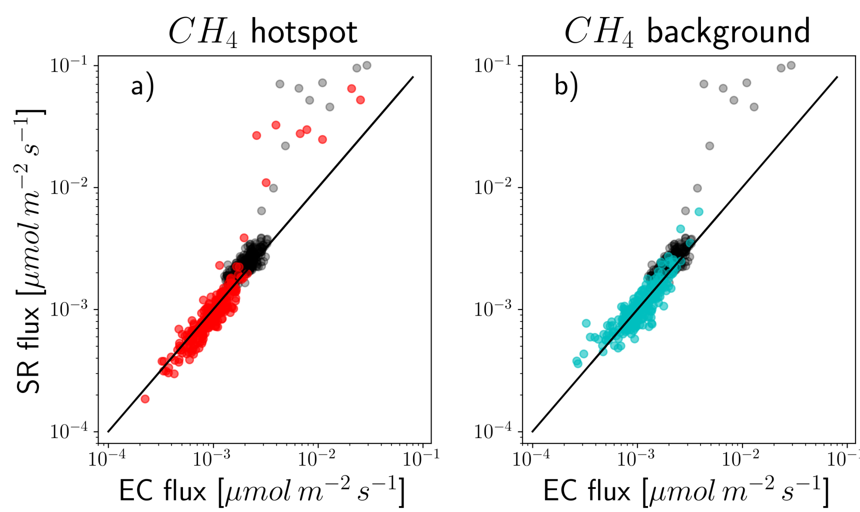


Figure 7.7: Surface renewal estimates for the hotspot (red) and background (cyan) components of $CH4$ compared with the overall CH4 flux (black circles). SR fluxes are compared compared with the corresponding wavelet - eddy covariance fluxes partitioned through the wavelet scheme described in section 7.4.

7.6 Discussion

7.6.1 Transport efficiency and Interpretation of ISR parameters

The gas transfer velocity and transport efficiency of the different gas transfer mechanisms are now considered. Generally, transport efficiency can be computed from eddy covariance flux estimates as $\overline{w'c'}/\sigma_c/\sigma_w$ and is related to the asymmetry in the c', w' components by quadrant analysis (through the quantity e_T analyzed in section 7.4), or through the slope β in the relaxed eddy accumulation method [BO90, BNB92, PDR93]. Transport efficiencies and gas transfer velocities ($\overline{w'c'}/\sigma_c$) are reported in Figure 7.8. Based on the wavelet partition scheme, the gas transfer velocities for hotspot and background CH_4 components can be computed as $(A^+\overline{F}_h)/(\sqrt{v_h/A^+}\sigma_c)$ and $((1-A^+)\,\overline{F}_b)/(\sqrt{v_b/\,(1-A^+)}\sigma_c)$ respectively. From this analysis, the hotspot component tends to have a larger transport efficiency compared to the background, as foreshadowed by the difference between H_2O and CH_4 in the moments of the joint pdf of (c' and w'). However, the overall dependence on u_* appears consistent between hotspot and background components, thus supporting the structure of the SR model used here, where both components of the flux are assumed to scale with $u_*^{3/4}$. As shown in panel 7.8b, for both components the gas transfer velocity exhibits the same dependence on u_*. A similar result is obtained from the analysis of SR fluxes (Figure 7.9. In this case, the gas transfer velocity is estimated dividing the SR flux by the standard deviation of the scalar. SR estimates appear to satisfactorily reproduce CH_4 gas transfer velocities and, as in the case of EC, assigns a larger transport efficiency to the CH_4 intermittent component. Note that a discrepancy exists for the runs characterized by a very intense intermittent flux, for which SR predicts a larger gas transfer velocity when compared to EC estimates. This can also be seen by directly comparing SR and EC gas transfer velocity values as reported in Figure 7.10, and is an effect of the large values of scalar skewness characterizing the runs with intense CH_4 hotspot events.

7.6.2 Effects of intermittent sources on CH_4 and CO_2 scalar concentrations

To elaborate on the effects of the intermittent CH_4 fluxes, and the results of the wavelet partition procedure, I focus my attention on a single run. I choose one of the most '*extreme*' runs in the dataset, recorded on $2013-07-09$ at $13:30$, which is a day characterized by an exceptionally intense CH_4 flux, and for which the partition scheme identifies multiple source hotspots. For this run, the CH_4 time series is particularly asymmetric with several 'spikes' of the order of 10 standard

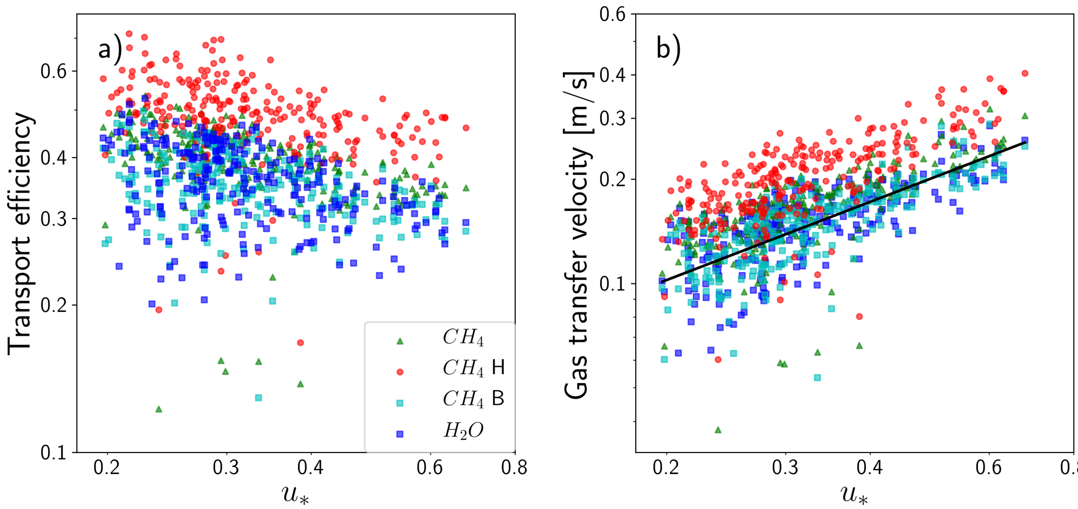


Figure 7.8: Eddy-covariance estimated transport efficiency (a) and gas transfer velocity (b) for H_2O (blue), CH_4 (green) and its hotspot (red) and background (cyan) flux components as a function of u_*, obtained through the wavelet partition of the total CH_4 flux. For each component, each data point corresponds to a different 30 minutes run in the dataset. The black line shows the 3/4 slope expected from the SR scaling.

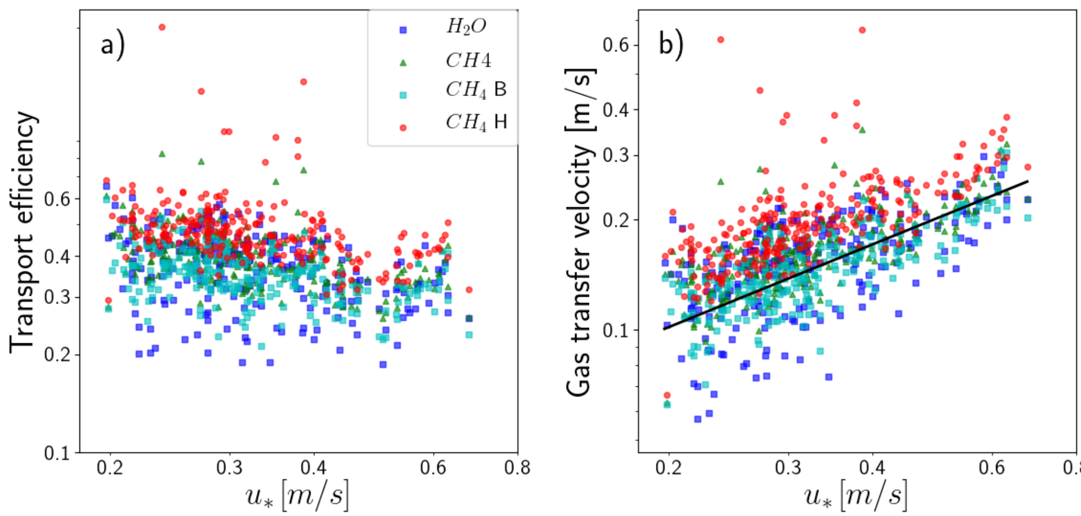


Figure 7.9: Surface renewal estimated transport efficiency (a) and gas transfer velocity (b) for H_2O (blue), CH_4 (green) and its hotspot (red) and background (cyan) flux components as a function of u_*. For each component, each data point corresponds to a different 30 minutes run in the dataset. The black line shows the 3/4 slope expected from the SR scaling.

deviations (Figure 7.11) while the water vapor series exhibits ordinary statistical properties. As a consequence, the wavelet partition clearly classifies most of these large events as intermittent (Figure 7.11c). Interestingly, CO_2 concentrations exhibit some positive concentration fluctuations that appear synchronous with their CH_4 counterparts. However, these CO_2 events are much less energetic when normalized by the variance of their respective series. This findings is consistent with the presence of localized ebullition events, with bubbles containing not only CH_4 but also CO_2. Gas chamber measurements indeed suggest that bubbles may be composed of about $20-80\%$ methane, and partially filled with other gases such as N_2 or CO_2 [PRLV17]. If this is the case, the relative weakness of CO_2 bubble-related signal can be explained by considering the atmospheric mixing ratio of these two gases. The ratio between bubble CH_4 concentration (~ 0.5) and CH_4 concentration in air ($\sim 2e-6$) is two

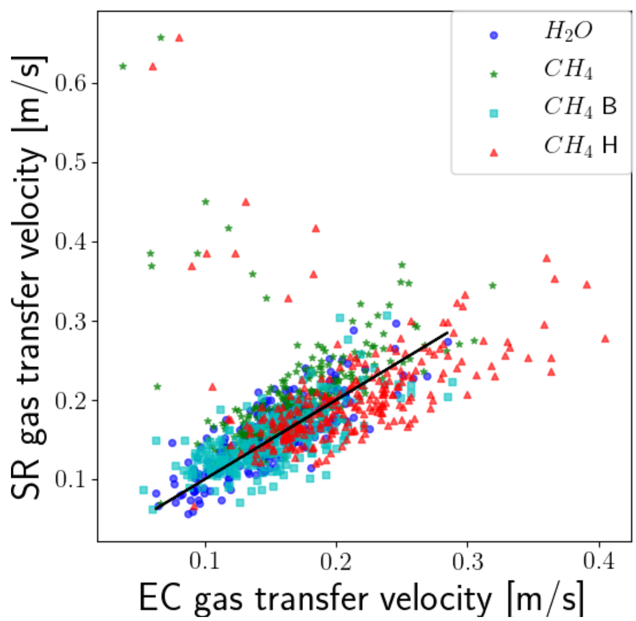


Figure 7.10: Comparison of gas transfer velocities computed with the eddy covariance (EC) and surface renewal (SR) approaches for H_2O (blue), CH_4 (green) and its hotspot (red) and background (cyan) flux components. The $1:1$ line is reported as reference.

order of magnitudes larger when compared to the ratio between bubble CO_2 (~ 0.3) and air CO_2 concentration ($\sim 4e-4$). Hence, the occurrence of ebullition would determine spikes in CH_4 concentrations that are significantly more intense and easily detected compared to their CO_2 counterparts. Moreover, these considerations may be used to explain the initial findings that CO_2 transport efficiency is lower than H_2O, a behavior opposite to what was observed for CH_4 (Figures 7.3a and 7.4a). In the case of CH_4, the occurrence of ebullition determines an increased transport efficiency with the most intense buildup in methane concentration being transported in few intense ejection events. These events significantly add to the already positive background diffusive flux. However, during daytime (which is the case for most of the runs analyzed here) the overall CO_2 flux is downward due to photosynthetic activity. However, the sporadic release of CO_2-containing bubbles determines an inflated fraction of positive CO_2 ejection and thus an overall reduced e_T for the total downward flux. This behavior can also be seen in Figure 7.13, where the daily cycle of scalar fluxes is shown as estimated by either EC or non-intermittent SR. Methane fluxes do not show any clear diurnal variability and are dominated by the few already noted highly intermittent runs. On the other hand, H_2O, CO_2 and sensible heat fluxes exhibit diurnal variability as expected. However, it is interesting to note that if one focuses on SR estimates and on the few runs supposedly characterized by intense ebullition, a large discrepancy can be observed between EC and SR fluxes for CO_2. This finding again supports the picture regarding intermittent CO_2 flux, which again would require an ISR model to be described as done for CH_4. In particular, it is the skewness term used in approximating Δc that produces this behavior. This behavior thus suggests the possibility to observe runs in which the ground is overall a CO_2 sink but the CO_2 skewness is positive due to the effect of localized ejections. A ISR scheme analogous to that developed for CH_4 here could describe this behavior after partitioning the time series in its background (downward) and intermittent (with a possible upward) CO_2 flux components. The results of such an intermittent SR model applied to CO_2 fluxes are shown in Figure 7.12, where one can see the results are (a) overall consistent with the EC results, and (b) CO_2 hotspots are characterized by more intense downward fluxes during most days, but in some cases these the sign of hotspots and backround fluxes differ. As seen in Figure 7.11, in some cases positive CO_2 spikes are coherent with their CH_4 counterparts and it is therefore possible that the interplay of these different transport mechanism reduces the overall transport efficiency for CO_2 when compared to the reference H_2O. These results however are not as clear as those obtained for CH_4, and may represent the combined effect of both release of bubbles and inhomogeneity in the CO_2 sinks at the surface (the vegetation is patchy).

Table 7.1: Values used for the molecular diffusivity D_m and Schmidt number $Sc = \nu/D_m$ for the scalar quantities of interest.

Quantity	$D_m[\cdot\ 10^{-4}\ m^2/s]$	$Sc\ [-]$
u	0.151	-
T	0.212	0.71
H_2O	0.24	0.59
CO_2	0.157	0.96
CH_4	0.222	0.68

7.6.3 Relation between partitioned CH_4 fluxes and environmental parameters

How the parameters of the SR scheme derived above describing the ebullition process vary with differing environmental and flow conditions is now discussed. The seasonal variability of the methane fluxes is reported in Figure 7.14a. Panel 7.14b shows the seasonal variability of ebullition fractions of flux (f_e), variance (v_e), and footprint area

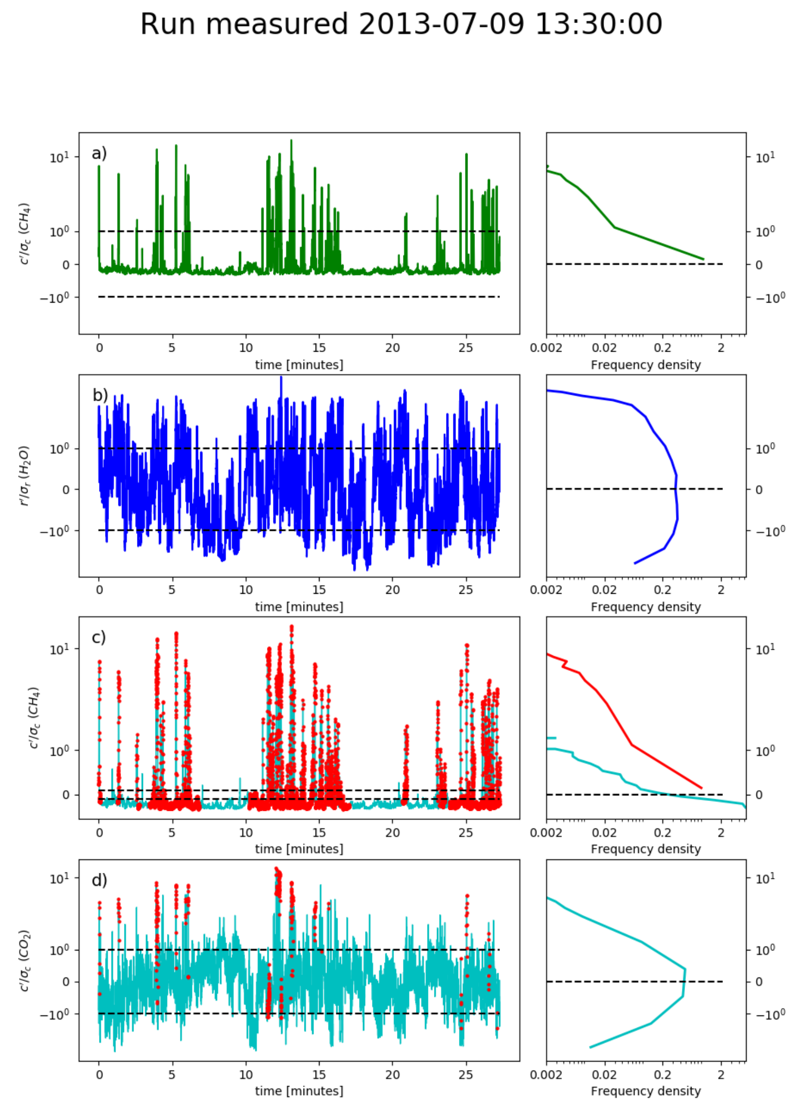


Figure 7.11: Example of run measured 2013-07-09 13:30:00, characterized by very intense CH_4 hotspots. The figure features (a) the normalized methane concentration (green), (b) water vapor (blue) and (c) the results of the wavelet partition (cyan and red indicating background and hotspot components respectively). A comparison with the corresponding CO_2 time series is featured in the lower panel, where red circles denote points for which the corresponding CH_4 concentration exceedes $3\sigma_c$ (d). For each time series, the dashed horizontal lines show the 1 standard deviation bands around the mean (computed for the background component only in panel (c)). The pdfs of normalized scalar fluctuations are included for comparison.

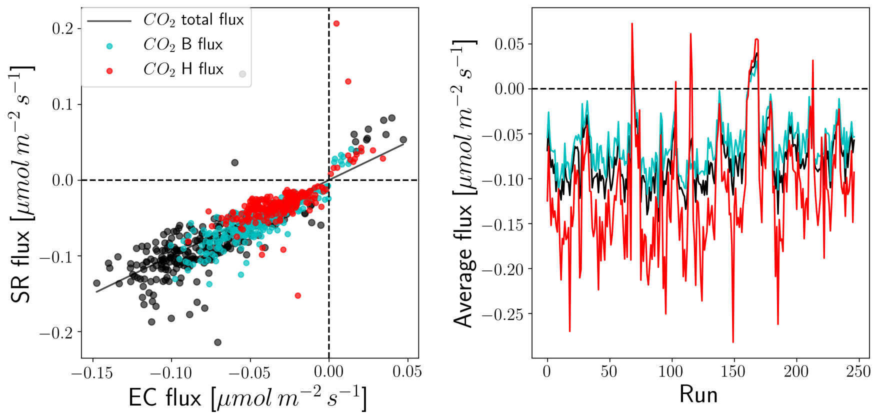


Figure 7.12: Partition of the CO_2 flux in background and hotspot components. Left panel shows a comparison of SR and EC estimates, while the right panel shows the average fluxes measured during hotspot and background only times, respectively.

A^+. The clear outliers occurring around the day 7-9-2013 are all characterized by intense CH_4 emissions, and by hotspots dominating the mass transfer mechanism. As an example, I have already examined one of these runs in Figure 7.11. I also show the relation of CH_4 fluxes with key environmental parameters. In particular, I study its variability with water table depth relative to the peatland surface, and with changes in atmospheric pressure (Figure 7.14). While overall there is no clear dependence, it is interesting to note that the runs characterized by very high CH_4 fluxes all occur for relatively high water table levels and in decreasing atmospheric pressure conditions. This picture is consistent with the conditions for which CH_4 ebullition is expected to occur. I conjecture that is it likely that for most of the runs, the mechanisms determining the intermittent fluxes may be variable, and include non-homogeneous terrain, variable microbial activity, and some ebullition. However, in about 10 cases, the runs are characterized by unusually intense CH_4 positive fluctuations that give rise to fluxes an order of magnitude larger than median seasonal values. For these runs, the evidence analyzed here would suggest ebullition is the primary cause for

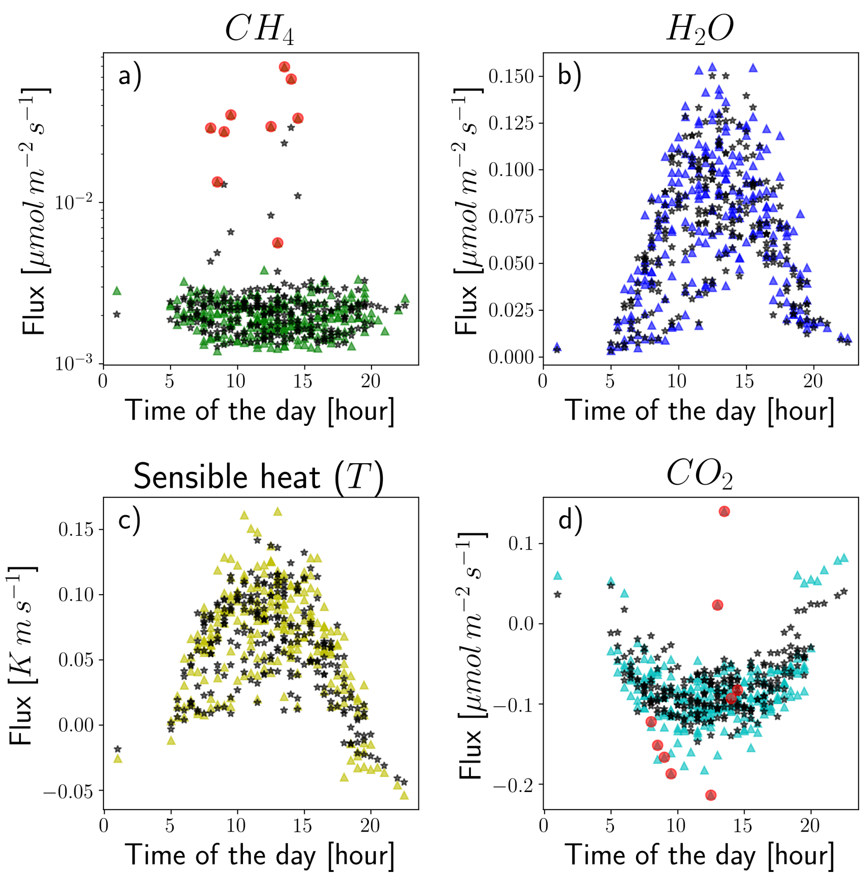


Figure 7.13: Daily cycles of surface renewal (black stars) and eddy covariance flux (colored circles) estimates for various scalars, for all the runs in the dataset. Note the outliers for CH_4 and CO_2 corresponding to large fluctuations in the concentration time series (marked in red for SR CH_4 fluxes larger than $0.004\, \mu mol\, m^{-2}\, s^{-1}$).

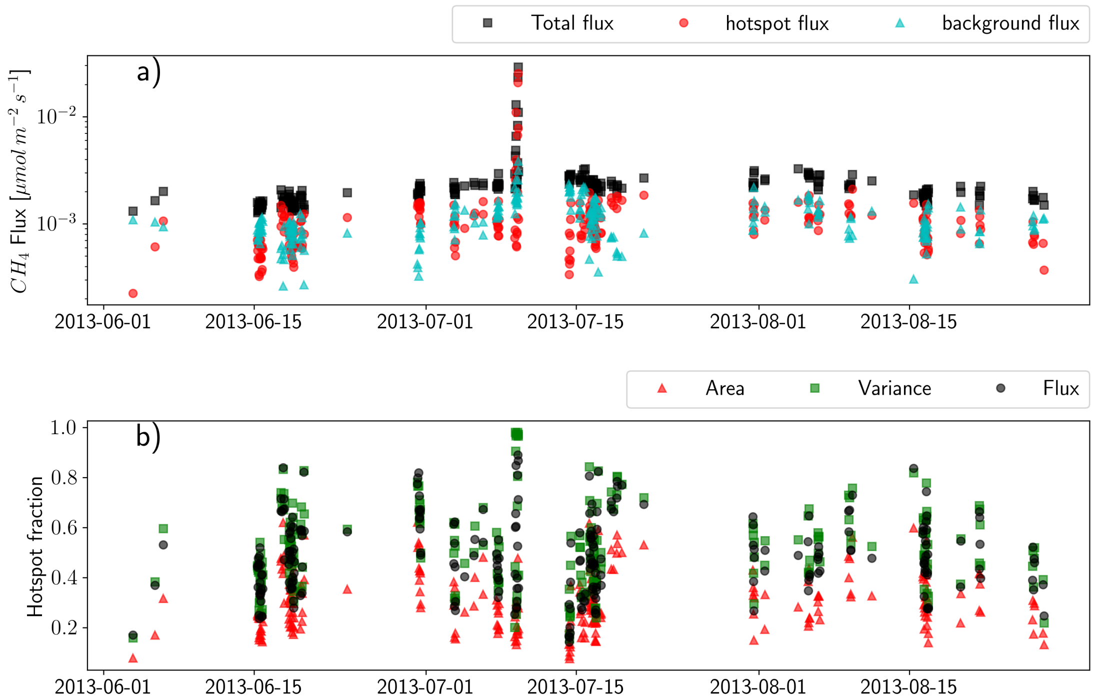


Figure 7.14: Panel (a): Seasonal variability of the total methane flux and of its ebullition and background components throughout the period of measurement analyzed. Panel (b): Fraction of ebullition flux (black circle markers), scalar variance (green squares) and fractional footprint area (red triangles) during the same time interval.

these events. This appears to be the case because i) they all occur for high water table levels, ii) they all appear for decreasing atmospheric pressure values, and iii) they are accompanied by a similar (albeit weaker) behavior observed for the CO_2 time series, which is consistent with a mixture of the two gases being release at the interface in the presence of bubbles.

7.7 Conclusions

An intermittent surface renewal scheme was proposed with the objective of characterizing the interplay and relative importance of diffusive and intermittent fluxes of

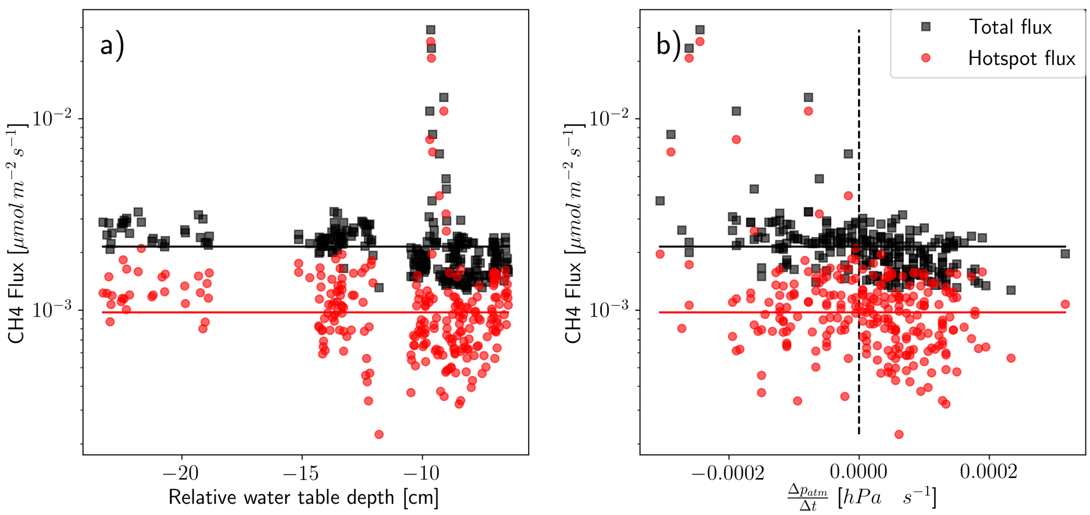


Figure 7.15: Dependence of CH_4 flux on (a) water table depth relative to the peatland surface, and (b) atmospheric pressure tendency (increment in atmospheric pressure computed for the 30 minutes corresponding to each run).

methane over boreal peatlands. This model, while still describing the interfacial mass transport as a function of Schmidt number Sc and Reynolds number Re_*, depends on additional parameters related to the spatial intermittency of the fluxes (A^+) and the relative strength of intermittent and background (i.e., continuous) scalar sources at the interface (E^+). In the context of CH_4 emissions from boreal wetlands, the intermittent hotspots detected with this framework can be linked either to the occurrence of ebullition, or to the non-homogeneous distribution of plants and microbial activity. However, the analysis of the CH_4 events characterized by the largest magnitudes, together with relevant environmental parameters, suggest that the sporadic release of bubble may be at the origin of the most intense hotspot events. Partitioning CH_4 fluxes with respect to a reference scalar (H_2O) suggested that spatial non-homogeneity of the CH_4 source, and in particular hotspots possibly related to ebullition, may be the primary cause for the observed statistical properties of CH_4 concentration traces: In particular, the skewed and non-Gaussian character of their pdfs, and the larger transport efficiency e_T compared to other scalars (e.g. temperature, water vapor). In the case of CO_2, the lower values of e_T appear to originate for the same reason in some of the runs analyzed here: The presence of local ejections, possibly related to sporadic bubble release, determines a decrease in the transport efficiency for the (downward) CO_2 flux. The results here suggest that SR, as well as the ISR extension for spatially inhomogeneous sources, can be complementary to EC measurements in quantifying scalar fluxes, and instrumental in detecting contributions to CH_4 fluxes from intermittent sources. This information is practically relevant for the purpose of upscaling point measurements of CH_4 fluxes performed with the use of gas chambers. Given the limited time and spatial coverage of such field measurements, the information provided by methods such as ISR can be relevant for planning and interpreting the results obtained from field campaigns, now increasing in number and coverage for CH_4.

Chapter 8

Conclusion

In this dissertation I have investigated the statistical properties of extreme values observed in rainfall as well as in atmospheric surface layer (ASL) turbulent flows. In the context of rainfall, this work challenges the classical adoption of asymptotic extreme value models, and proposes instead techniques which use all available observations to infer the frequency of large, and possibly yet unobserved, extremes. This point of view appears particularly suited for applications to rainfall estimates from remote sensing, where observational uncertainty hinders the use of models which only include the few largest observed events. The advantage of using the non-asymptotic extreme value framework proposed was further quantified by studying extremes from satellite quantitative precipitation estimates (QPEs). This application, in turn, paved the way toward additional lines of inquiry, since proper QPE validation requires a careful examination of (i) observational biases, which may deform the probability distribution of rainfall intensities, with important consequences for extreme-value statistics, and (ii) issues related to spatially-averaged rainfall fields, whose statistics, and extremes, are dependent on the averaging scale, which needs to be accounted for when comparing observations/estimates at different scales. These two directions of investigation were separately studied in Chapters 4 and 5, where a framework for downscaling and validating QPE statistics in data-scarce regions was developed. This dissertation also investigated the role of extremes in ASL flows. From this perspective, my work clarified the role of boundary conditions in determining the frequency of extreme fluctuations of scalar quantities transported in the ASL. By studying the transport of sensible heat, I found that the frequency of intense temperature fluctuations depends on the interplay of buoyancy associated with heating and shear-driven turbulence, even well within the inertial subrange, a range of scale at which the classical paradigm assumes that the detailed characteristics of the turbulence-generation mechanism are lost as energy cascades through smaller and smaller spatial scales. The small-scale statistical properties of scalars in ASL turbulence are not only impacted the turbulence generation mechanism, but also contain relevant information on whether the scalar is released in the flow intermittently, or in a spatially heterogeneous way. Here I studied this phenomenon in the context of methane (CH_4) emissions from boreal peatlands. These environment are characterized by the presence of natural CH_4 'hotspots', related to either the sporadic release of methane-containing bubbles in air, or to the heterogeneous distribution of plants and microbial activity, which impress a clear signature to the statistical properties of CH_4 concentrations measured in the ASL. In Chapter 7, this signature was exploited to partition CH_4 fluxes based on the different gas exchange mechanisms (i.e., 'hotspots' possibly related to bubble release events, as opposed to steady background diffusion through the water and peat column), and to study the role of these extreme gas exchange events in relation to the overall ecosystem greenhouse gases budget. While different from a physical standpoint, the analysis of remotely-sensed rainfall and intermittent gas fluxes from gas analyzer and sonic anemomenter measurements pose a very similar challenge to the investigator. In both cases, the physical processes of interest, i.e., either the generation of extreme rainfall on the one hand, or the gas transfer rate on the other, is dominated by small-scale spatial distribution of the relevant physical quantities, at scales that are, in particular, smaller than the 'resolution' of the observations. This resolution is the grid cell size in the case of rainfall, and the sampling frequency of the sonic anemometer in the case of ASL turbulence. One of the main objectives of the investigations described in Chapters 4, 5, and 7 was indeed bridging this gap between the scale of available data products, and the smaller scales of the physical processes under investigation. In both cases, as shown in Chapter 4 for rainfall, and in Chapter 7 for ASL turbulence, when information at the integral time scale is available, this can be combined with assumptions on the 'subgrid' behavior in order to infer some properties of the process at the smaller scales. In Chapter 7 this is achieved through a simplified turbulent kinetic energy balance [KKP11], coupled with the idealized representation of interfacial gas transfer provided by the surface renewal theory [ZPGK20]. Similarly, in the case of rainfall, the effects of spatial averaging on the probability distribution of large events can be inferred by assuming a spatial correlation function consistent with the observed spatial scaling of rainfall [Mar05, OG94, NB15, ZM19]. In both cases, particular care is needed when applying this technique to new data. As an example, in the case of rainfall areas with complex topography, the associated spatial heterogeneity and anisotropy of the rainfall field can severely limit this transfer of information across scales, as found and discussed in Chapter 5. With respect to this issue, the results obtained here can be used to inform the design of future studies and field campaigns, by combining point-measurements (e.g., from rain gauges or gas chambers) with 'remote' observations of integral properties of the two processes (e.g., remote sensing or eddy-covariance measurements).

The main contributions of this dissertation can be summarized as follows:

- Chapter 2 tested a statistical model of extreme daily rainfall which, by describing the entire distribution of daily values, proves to be a more robust alternative to traditional methods in the case of short sample sizes. By analyzing a set of long ($S >$ 100 years) rainfall records, my analyses pointed out two important aspects to be considered when evaluating extreme-value models. The first aspect concerns the importance of using cross validation techniques to evaluate and select alternative models. The second aspect regards the relevance of the ratio T_r/S, i.e. the ratio of the average recurrence interval of interest to the sample size available. This ratio quantifies the degree of extrapolation involved in the statistical estimation being performed and, as found in Chapter 2, different extreme-value statistical models exhibit quite different performances when analyzed as a function of T_r/S.

- In Chapter 3, I developed an extension of the model described in Chapter 2 using the formalism of Bayesian hierarchical models. This extension proved useful in quantifying estimation uncertainty, and in incorporating physical information on the process at hand through the elicitation of informative priors. While the use of Bayesian models is not new in atmospheric science and hydrological applications, here we introduced a novel latent level approach which can be instrumental in describing extreme fluctuations of processes varying over multiple, much different, temporal scales. By comparing measures of in-sample and out-of-sample predictive accuracy, I found that the model structure developed here can significantly improve robustness with respect to overfitting the specific sample of extreme values observed.

- Chapter 4 focused on the estimation of extreme rainfall from remotely-sensed datasets. To this end, I worked to address two key issues: the limited length of rainfall datasets on one hand, and the scale disparity between gridded and rain-gauge precipitation products on the other, a relevant issue when validating results using ground observations. Here I addressed this problem by developing an approach to downscale key statistical properties of the rainfall field: The frequency of rainfall events, their spatial correlation, and the distribution of their intensities. Testing this methodology using data from the Tropical Rainfall Measuring Mission (TRMM) over a single study site in Oklahoma, I have found that this approach satisfactorily reproduced downscaled daily rainfall pdfs. Then, the methodology was combined with the extreme value model presented in Chapter 2. By applying this combined approach over the test site, I found that, despite the short record length available, its application can prove relevant for improving the estimation of extreme rainfall frequency over poorly instrumented regions.

- In Chapter 5 I applied the downscaling technique developed in Chapter 4 to perform a large-scale comparison of statistical properties of rainfall observed from single rain gauges at the ground and a gridded precipitation product over the Conterminous US. After evaluating discrepancies between downscaled and ground-measured rainfall statistics, the methodology was then extended by developing a spatial model of the error. This extension is intended for interpolating and characterizing the performance of remotely sensed rainfall products over poorly gauged areas. Towards the long-term objective of producing and validating a global-scale representation of extreme rainfall frequency, here I produced a first global estimate based on the tools developed in Chapter 2. I then discussed the implications and the possible approach for a large-scale, global, validation effort based on the results obtained from the United States.

- Chapter 6 experimentally explored the effects of mechanical generation of turbulent kinetic energy and buoyancy forces on the statistics of air temperature in the atmospheric boundary layer, focusing on a range of scales from production to the inertial range. I found that within this range of scales the turbulence generation mechanism leaves a signature on both the probability of extreme temperature fluctuations, and on the imbalance between forward and backward phase-space trajectories of the temperature field. In particular, I found that the absolute magnitude of the atmospheric stability impacts the distribution of scalar increments at separation scales well within the inertial sub-range. Moreover, the sign of the heat flux fingerprints the observed time-directionality properties of the temperature field in the first two decades of inertial sub-range scales. These combined findings demonstrate that external boundary conditions, and in particular the magnitude and sign of the sensible heat flux, have a significant impact on temperature advection-diffusion dynamics well within the inertial range.

- Chapter 7 introduced a technique for partitioning measured methane fluxes over natural wetlands based on the spatial intermittency of the gas transfer mechanism at the interface. For a study site in Finland, I found that differently from water vapor fluxes, the exchange of CH_4 exhibits short-term extreme flux *"hotspots"* events which can be difficult to characterize with the usual eddy covariance technique, as well as to measure directly at the ground. By developing an intermittent extension of the classical surface renewal theory, this work introduces a novel approach for inferring the intermittent nature of scalar sources at the ground and for exploring how their heterogeneity impacts the efficiency of gas turbulent transport in the atmospheric surface layer.

These results suggest multiple future research directions. In recent years, increasing attention has been devoted to the concept of probabilistic QPEs, i.e., rainfall estimates characterized by a complete probabilistic description of the different sources of uncertainty deriving from the specific retrievals and algorithms used [MSA+14, KGH+15, KKHH18]. However, when using these products to train statistical models, an essential step is providing a unified quantification of the uncertainty deriving from these different sources, together with the uncertainty deriving from inference on the statistical model itself. The Bayesian hierarchical model developed in Chapter 3 is a promising tool in this respect. By adding one additional layer to the model for representing the probabilistic nature of the QPE, the overall uncertainty can be estimated for any functional quantity of interest and then used in hydrometeorological studies.

In a broader perspective, the research described in this dissertation hopes to be a contribution to long-standing research questions relevant for hydroclimatology. Heavy precipitation and flooding events are expected to become more frequent in a warming climate, although the strength of these changes is difficult to characterize, especially at the finest temporal scales [AS08, WFE+14, MV15]. This difficulty arises both from a data-driven perspective, due to the need of long and homogeneous records [PM19], and from a modelling standpoint, where the correct representation of future changes in precipitation is one of the *'real holes of climate science'* [Sch10]. Moreover, the inter-annual variability in the occurrence of extreme precipitation events, as modulated by the internal variability of the climate system, is also expected to be impacted by the warming conditions [PKL+17]. To address this challenge, hydrologists and water managers have long advocated the adoption of non-stationary extreme value models [MBF+08]. However, their application is still actively debated in the hydrological literature [LC11, MK14], as the benefit of including climate-informed covariates in the statistical analysis of extremes is often overshadowed by the inflated uncertainty associated with both the inference process and model selection. Toward this long-standing problem, my dissertation offers two contributions. First, the results obtained in Chapters 2 and 3 underline the importance of accounting for the time-scale separation between possible climatic effects and the *'fast'* timescale representing the actual phenomenon of interested (extreme daily rainfall in this case). Second, the hierarchical structure developed in Chapter 3 is particularly suited to be expanded to include the effect of climate controls on precipitation statistics. Combining the tools developed here with physical understanding of climate controls on precipitation processes will lead to an improved ability to characterize variability and future changes in the frequency of extreme rainfall, a fundamental step for improve preparation to and response against the hazards associated with precipitation extremes.

Appendices

Appendix A

Appendix to Chapter 2

In this section I provide supporting information on data sets and methods used in Chapter 2 to compare the performance of traditional GEV methods and Metastatistical Extreme Value distributions.

A.1 Fitting of the GEV Distribution

The GEV distribution, depending on the value of its shape parameter ξ, encompasses the entire range of limiting distribution *types*, as stated in the Extremal Types Theorem [FT28]. The three cases of upper bounded, exponential and heavy tailed distribution are thus included in this formulation. The Von Mises parametrization of the GEV Cumulative Distribution Function (CDF) of a random variable X reads [VM36]:

$$F\left(x; \mu, \psi, \xi\right) = \exp\left\{-\left(1 + \frac{\xi}{\psi}\left(x - \mu\right)\right)^{-1/\xi}\right\} \tag{A.1}$$

Where μ, ψ, ξ are respectively location, scale and shape parameters. The GEV distribution has become the standard approach to model rainfall extremes and much of the recent literature in the field focuses on different approaches to most effectively estimate its parameters from a sample of block-maxima. Several methods have been developed for this task, some of the main contenders being the Maximum Likelihood (ML) [Col01, MS00a], L-Moments (LMOM) [Hos90, HWW85], and Mixed Methods [MS02]. It has been shown that the ML estimators for the GEV distribution have a larger variance compared to the LMOM-family estimators, especially for positive values of ξ. On the contrary, LMOM estimators exhibit smaller variance and are affected from bias [MS02, HWW85]. Hence, LMOM has better performances when applied to small samples (due to the reduced estimator variance) and therefore it is the most suited candidate to be applied in the range of sample sizes explored in this work. I implemented the ML approach by numerical minimization [GSU01] of the log likelihood function:

$$\lambda_{x_i}(\mu,\psi,\xi) = -N\log\psi - \left(\frac{1}{\xi}+1\right)\sum_i \log\left(1+\xi\frac{x_i-\mu}{\psi}\right) - \sum_i\left(1+\xi\frac{x_i-\mu}{\psi}\right)^{-1/\xi} \tag{A.2}$$

For a complete description of the MLE for GEV and its asymptotic properties, see Coles [Col01]. As per the LMOM method, GEV parameters were estimated using the approach introduced by Hosking [Hos90]. L-moments are defined as linear combinations of the Probability Weighted Moments (PWM). The estimate for the r-th order PWM from a ordered sample $x_{1:n}, .., x_{n:n}$ reads:

$$\beta_r = \frac{1}{n}\frac{\sum_{i=1}^{n}\binom{i-1}{r}x_{1:n}}{\binom{n-1}{r}} \tag{A.3}$$

Hence, L moments estimates are evaluated from PWM estimates as $\hat{\lambda}_1 = \beta_0$, $\hat{\lambda}_2 = 2\beta_1 - \beta_0$ and $\hat{\lambda}_3 = 6\beta_2 - 6\beta_1 + \beta_0$. Following this approach, it is possible to determine GEV parameters as a function of the first two L moments $\hat{\lambda_1}$ and $\hat{\lambda_2}$ and the L skewness $\hat{\tau}_3 = \hat{\lambda_3}/\hat{\lambda_2}$:

$$\frac{\hat{\tau}_3+3}{2} = \frac{1-3^{\hat{\xi}}}{1-2^{\hat{\xi}}} \tag{A.4}$$

$$\hat{\psi} = \frac{\hat{\lambda_2}\hat{\xi}}{\left(2^{\hat{\xi}} - 1\right)\Gamma\left(1 - \hat{\xi}\right)} \tag{A.5}$$

$$\hat{\mu} = \hat{\lambda_1} + \frac{\hat{\psi}}{\hat{\xi}}\left[1 - \Gamma\left(1 - \hat{\xi}\right)\right] \tag{A.6}$$

Solution of equation (A.4) would require a numerical procedure. However, for the range of shape parameter encountered in this application ($-0.5 < \xi < 0.5$) the following polynomial approximation can be used (see [HWW85] for details):

$$\hat{\xi} = 7.8590 \cdot c + 2.9554 \cdot c^2 \tag{A.7}$$

Where

$$c = \frac{2}{3 + \hat{\tau}_3} - \frac{\log 2}{\log 3} \tag{A.8}$$

The results obtained from the Monte Carlo analysis (Fig. A.4D) show that for the range of sample sizes considered here, LMOM has in general better performances than ML estimators. In particular for return times bigger than the available sample size, LMOM yields on average estimates with a Root Mean Square Error (RMSE) smaller than ML. The GEV distribution, obtained as asymptotic distribution, would ideally be a proper model only for maxima extracted from infinitely large blocks. In practical applications the selection of larger blocks inevitably reduces the sample size available for the fitting procedure. On the other hand, too small blocks violate GEV fundamental hypothesis and in many environmental application fail to provide stationary series (e.g. sub annual blocks in the case of daily rainfall).

A.2 The Peak Over Threshold Method

The Peak Over Threshold (POT) method was implemented following [DS90] to test its performance against the GEV and MEV approaches. In the POT framework, the excesses $y = x - q$ over a sufficiently high threshold (q) are assumed to be drawn from independent and identically distributed random variables, which obey the following Generalized Pareto Distribution (GPD):

$$F(x; \kappa, \sigma) = 1 - \left(1 + \frac{\kappa y}{\sigma}\right)^{-1/\kappa}; \tag{A.9}$$

The GPD shape and scale parameters (κ and σ, respectively) were estimated by means of the ML method [Col01, GSU01]. The arrival rate (number of exceedances over q in one year) is modeled using a Poisson distribution with mean λ (estimated here as the sample mean of the yearly number of exceedances $\hat{\lambda} = \bar{n}$). This POT model yields a CDF for the annual maximum which has a GEV form. Hence, the parameters of the latter can be estimated on the basis of the POT parameters $(\kappa, \sigma, \lambda)$ as follows:

$$\xi = \kappa \tag{A.10}$$

$$\sigma = \psi + \xi\,(q - \mu) \tag{A.11}$$

$$\lambda = \left(1 + \xi \frac{q - \mu}{\psi}\right)^{-1/\xi} \tag{A.12}$$

As the threshold q tends to ∞ [Lea83] (i) the distribution of excesses tends to a GPD form and, (ii) their arrival process tends to be Poisson. When fitting a POT model, the selected threshold must be high enough such that these two conditions are approximately met. Interestingly, if a relatively low threshold is selected, such that the bulk of $F(x)$ is also included rather than just its tail, the shape parameter of the GPD is seen to decrease. This implies that, for lower thresholds, the distribution of the excesses is only marginally different from an exponential [CH04, SK14]. In this case a stretched exponential distribution is more suitable to capture the observed rainfall values. In every realization of the Monte Carlo analysis a threshold was selected that is exceeded on average 5 times per year (thus, the total number of excesses is $5 \cdot s$). This guarantees that a variable and sufficiently high threshold is automatically selected at every generation. Figure A.1 shows the performance of MEV and POT in the (s, Tr) space. The result is consistent with the GEV-MEV contour (Fig. 2 in Chapter 2), confirming that MEV on average produces better estimates for return time bigger than the available sample size. The same results are also reported as a function of the dimensionless quantity (Tr/s) (Fig A.4C).

A.3 On the Asymptotic Assumption

It has been shown that, in hydrological applications, the number of yearly events may not be sufficient for the asymptotic assumption to hold [CH04, Kou04]. This problem could be circumvented by increasing the size of the blocks (in the Block Maxima approach) or by considering higher thresholds (for POT), which would in turn require disproportionately large samples. On the other hand, considering lower thresholds allows a reduction of the estimation uncertainty, but too low a threshold might lead to a bias when the GPD and/or Poisson assumptions are no longer justified. Hence an essential task is to determine what the tail of the distribution exactly is, i.e. what fraction of the available observations should be incorporated in the EV analyses. The optimal solution is a trade off between bias and variance. For some underlying distributions (e.g. in the case of a Weibull parent) the convergence rate to its limiting extremal pdf is so slow [CH04] that the use of a different asymptotic distribution (Frechet instead of Gumbel), though not theoretically justified, is suggested by empirical evidence [Kou04, PK13].Papalexiou and Koutsoyiannis[PK13] and Serinaldi and Kilsby [SK14] explored respectively the dependence of the GEV/GPD estimated shape parameter on different features of samples and estimation procedures. In particular, they find that, even in the presence of a heavy-tailed underlying stochastic process, the sub-exponential tail tends to remain undetected when small samples are used to fit and asymptotic EV distribution.

A.4 Fitting of the MEV Distribution

The discrete expression of the MEV-Weibull distribution, obtained by means of Monte Carlo integration, requires the fitting of the Weibull distribution to the non zero daily rainfall values for every single year of the sample. Since the sample size is relatively small (it is equal to the yearly number of wet days) the Probability Weighted Moments (PWM) approach was used. The k-th order PWM for the Weibull distribution reads [GLM79]:

$$M_{1,0,k} = M_{(k)} = \frac{C \cdot \Gamma\left(1 + \frac{1}{w}\right)}{(1+k)^{1+\frac{1}{w}}} \tag{A.13}$$

For an ordered sample of size n, $x_{1:n}, ..., x_{n:n}$, the estimated PWM of order k $\hat{M}_{(k)}$ is:

$$\hat{M}_{(k)} = \frac{1}{n} \sum_{j=1}^{n} \frac{(n-j) \cdot (n-j-1) \cdot ... \cdot (n-j-k+1)}{\cdot (n-1) \cdot (n-2) \cdot ... \cdot (n-k)} \cdot x_{j:n} \tag{A.14}$$

So that the first two moments read $\hat{M}_{(0)} = \frac{1}{n} \sum_{j=1}^{n} x_{j:n}$ and $\hat{M}_{(1)} = \frac{1}{n} \sum_{j=1}^{n} \left(\frac{n-j}{n-1}\right) x_{j:n}$.

The Weibull shape and scale parameters can thus be explicitly expressed as a function of the first two PWM sample estimates:

$$\hat{w} = \frac{\hat{M}_{(0)}}{\Gamma\left(\log \frac{\hat{M}_{(0)}}{\hat{M}_{(1)}}\right)} \tag{A.15}$$

$$\hat{C} = \frac{\log 2}{\log \frac{\hat{M}_{(0)}}{2\hat{M}_{(1)}}} \tag{A.16}$$

Wilson and Toumi [WT05] provided a physical justification for the adoption of a stretched exponential as fundamental distribution to model heavy daily rainfall. They expressed the daily precipitated depth as a product of three independent Gaussian distributed variables (mass flux, specific humidity and precipitation efficiency) and recognized the pdf of such a product to be stretched exponential with shape parameter $w = 2/3$ and scale parameter $C \in R$, thus implying a sub exponential tail behaviour with characteristic scale [LS98]. The MEV framework assumes daily rainfall values to be realizations of independent and identically distributed Weibull random variables. To corroborate this assumption, I analyzed the correlation of daily rainfall from all the datasets, finding that the average value for the 1-day lag autocorrelation is about 0.2 (Fig A.3B). The Weibull assumption was also tested by means of Kolmogorov-Smirnov (KS) and Cramer Von-Mises (CVM) statistical tests. These tests were performed for all the single years for every dataset, treated as independent samples (coherently with their fitting in estimating the MEV parameters). Averages of the p-value were then computed for all the 37 datasets (Fig A.3A). Both tests yielded very similar result, showing good agreement between empirical data and the stretched exponential model. Convergence of the discrete MEV formulation (3) to eq. (2) is ensured by the Bayesian Monte Carlo integration [Wei00]:

$$\int_{\vec{\theta}_* \in \Omega_{\vec{\theta}_*}} f(\vec{\theta}_*) p(\vec{\theta}_*) d\vec{\theta} \simeq \frac{1}{M} \sum_{j=1}^{M} f(\vec{\theta}_j) \tag{A.17}$$

Convergence is ensured as long as $\vec{\theta}_j$ $j = 1..M$ are sampled from the target distribution $p(\vec{\theta})$, as any value $\vec{\theta}_* \in \Omega_{\vec{\theta}}$ appears in the summation (A.17) with frequency $Mp(\vec{\theta}_*)d\vec{\theta}$. Hence, the right hand side of eq. (A.17) is a discretization of the integral:

$$\frac{1}{M} \sum_{j=1}^{M} f(\vec{\theta}_j) = \sum_{\vec{\theta}_* \in \Omega_{\vec{\theta}}} f(\vec{\theta}_*) p(\vec{\theta}_*) d\vec{\theta} \tag{A.18}$$

The accuracy of this approximation increases with M (in particular, the error scales like $1/\sqrt{M}$), where in the present application M is the length (in years) of the available rainfall record.

A.5 Distribution of the Meta-Parameters

The nature of the distribution of the meta parameters allows to better understand the properties of the MEV distribution. It has been shown [BC03, PVF06] that by compounding an exponential with a gamma distribution, one can obtain a power-law behaviour. A similar theoretical argument has been applied to the case of a compound Weibull-Gamma distribution [Dub68]. In the case of the MEV distribution, a similar argument can be used if one assumes the quantity $k = C^{-w}$ to be a gamma-distributed random variable (with shape and scale parameters, say, α and β respectively) and n to be fixed. In this case the expression of the compound distribution is:

$$\zeta(x) = \int_{-\infty}^{+\infty} \left(1 - n \cdot e^{k \cdot x^w}\right) \frac{k^{\alpha-1} e^{-\beta k}}{\Gamma(\alpha) \beta^{-\alpha}} \cdot dk \tag{A.19}$$

Note that here I applied the *Penultimate Approximation* [CH04, MI15] to eq. (2), thus assuming that $H_n(x) \cong 1 - n \cdot G(x) \cong \exp[-n \cdot G(x)]$ (Where $G(x)$ is the Weibull survival distribution function and $G(x) << 1$ for large x). Hence, it is possible to rewrite eq. A.19 as:

$$\zeta(x) = 1 - n \cdot \left(1 + \frac{x^w}{\beta}\right)^{-\alpha} \tag{A.20}$$

Hence, for sufficiently large values, the distribution of X tends to a power law with exponent $-w \cdot \alpha$. The higher-order variability in the Weibull scale parameter produces, in this case, a compound distribution with a heavier tail.

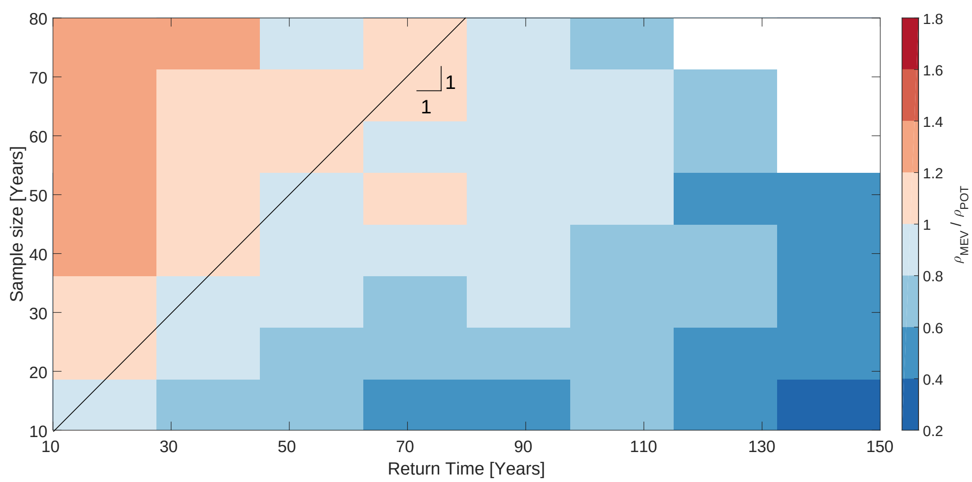


Figure A.1: Ratio ρ_{MEV}/ρ_{POT}, of the Root Mean Square Errors of quantile estimates from the MEV-Weibull and POT approaches as a function of return period and size of the sample. Individual ρ_{MEV}/ρ_{GEV} values (obtained from $n_r = 100$ Monte Carlo reshuffled time series) from all the stations are pooled together and averaged over rectangular tiles on an uniform grid. White cells do not contain data.

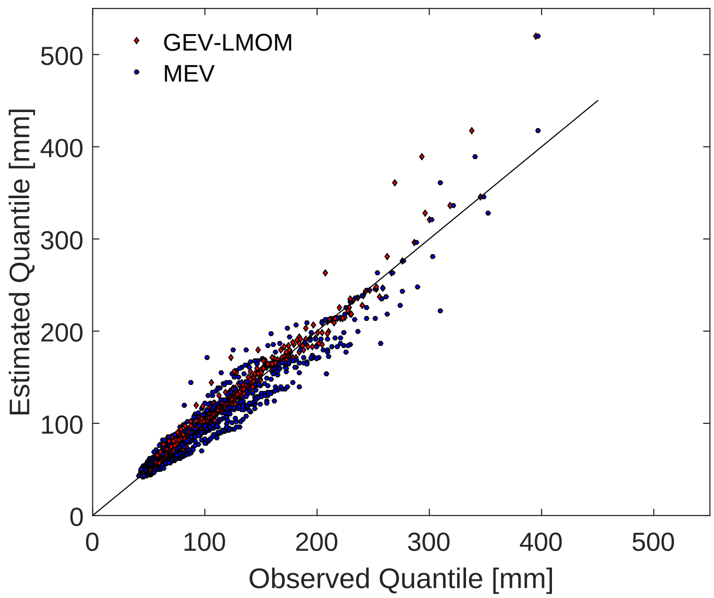


Figure A.2: Quantile-quantile plot for MEV and GEV distributions fitted to all the available time series after reshuffling (to ensure stationarity). The same sample is here used for calibration and performance evaluation. The GEV distribution shows a better performance when evaluated using the same sample of annual maxima used for its calibration. Comparisons with proper independent testing, where the MEV approach outperforms the GEV distribution, suggest that the latter is less capable to infer the general properties of the population.

Table A.1: Mean and standard deviation of the estimation errors for a sample of size s=30 years.

Return Time		GEV	POT	MEV
Tr=10	μ	-8.51	-8.54	-5.37
	σ	0.224	0.223	0.236
Tr=20	μ	-7.87	-7.95	-3.59
	σ	0.223	0.225	0.228
Tr=50	μ	-5.99	-6.27	-0.76
	σ	0.232	0.234	0.217
Tr=100	μ	-5.46	-5.67	-3.71
	σ	0.262	0.252	0.185

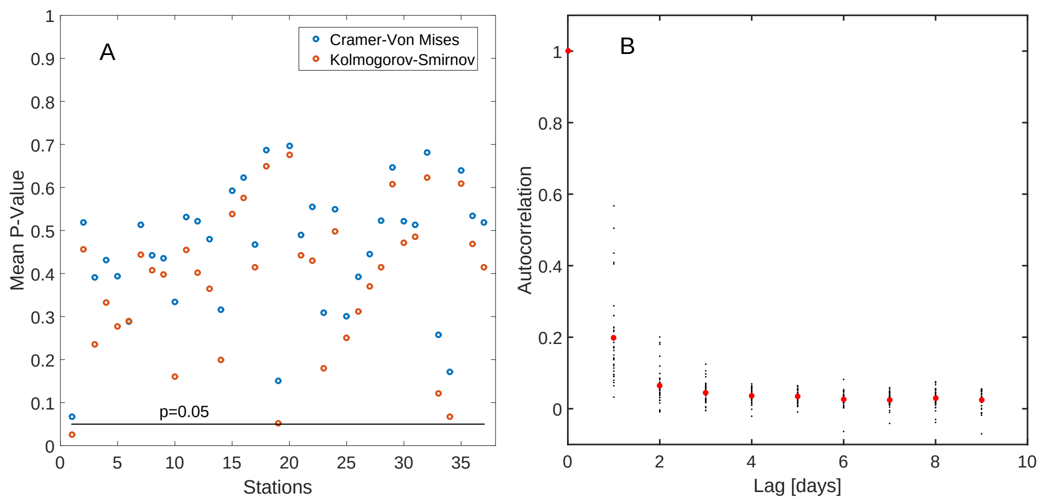


Figure A.3: Results of goodness-of-fit tests for the Weibull distribution (A), and rainfall autocorrelation (B). Kolmogorov-Smirnov and Cramer-Von Mises tests were performed by fitting the Weibull distribution to every single year on record for every station. Yearly p-values were then averaged for every station. The line corresponding to the critical value $\alpha = 0.05$ is reported as a reference. Note that for most of the stations the mean p-value significantly exceeds α. In panel (A) the autocorrelation of daily rainfall is reported for different time lags. Black points represent the value computed for each station, whereas red closed circles are averaged over all the available datasets. The autocorrelation for a 1-day lag is, on average, about 0.2.

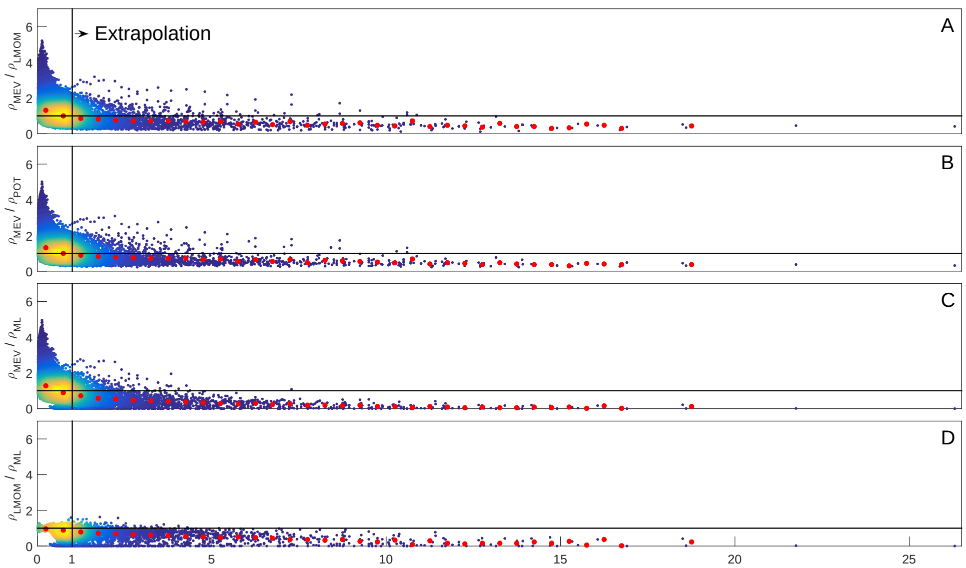


Figure A.4: Comparative performance of different EV models as a function of Tr/s, obtained for $n = 100$ Monte Carlo generations. Ratio between the Root Mean Square Errors for (A) MEV and GEV fitted with LMOM, (B) MEV and POT, (C) MEV and GEV fitted with ML, and (D) GEV fitted with LMOM and ML. The colors represent the density (points/unit area of the plot) of the values falling in each area of the scatter plot (blue indicating the lowest density and yellow the highest one). The horizontal line marks the equal performance condition between two methods. The vertical lines correspond to the value $Tr/s = 1$.

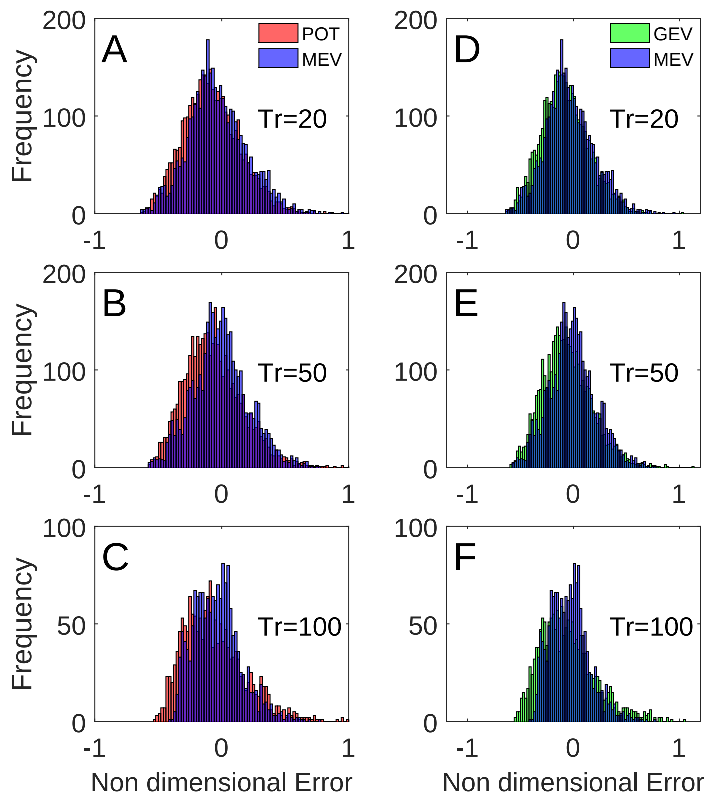


Figure A.5: Comparisons of the relative error distributions for the MEV and GEV-LMOM (panels A,B,C), and POT (panels D,E,F). The relative error was computed over all the stations and Monte Carlo realizations ($n_r = 100$) using a fixed sample size ($s = 30$ years) and different return times. As the return time increases with respect to the sample size, the MEV distribution exhibits a remarkably smaller variance compared to GEV and POT methods (see table A.1 for the values of error mean and standard deviation). Histograms for $Tr = 100$ were obtain pooling together data from all the stations longer than 130 years.

Table A.2: Summary information about the records used in the analysis.

Station ID	Station Name	Country/ US State	Elevation [m m.s.l.]	Latitude	Longitude	Period of record	Missing data [years]	Length [years]
FR000007560	Mont Aigoual	FR	1567	44.1167	3.5831	1896-2014	3	116
USC00090140	Albany	GA	54.9	31.5333	-84.1333	1893-2014	2	120
UK000047811	Armagh	UK	62	54.35	-6.65	1838-2001	0	164
-	Asheville	NC	682.1	35.5954	-82.5568	1903-2006	0	104
GM000004063	Bamberg	GM	240	49.8753	10.9217	1879-2014	0	136
ASN00063118	Bilpin Fern Grove	AS	610	-33.5156	150.4892	1895-2004	0	110
ITE00100550	Bologna	IT	53	44.50	11.3458	1813-2007	0	195
SF000030320	Bredasdorp	SF	50	-34.533	20.033	1875-1991	9	108
ASN00068007	Brownlow Hill	AS	61	-34.025	150.645	1883-2014	4	128
EI000003969	Dublin	EI	49	53.3639	-6.3192	1881-2014	8	126
ITE00100552	Genova	IT	55	44.4144	8.9264	1833-2008	0	176
NLE00100502	Heerde	NL	6	52.3958	6.0514	1893-2014	0	122
NLE00100503	Hoofdoorp	NL	-3	52.3108	4.7042	1866-2014	1	148
GM000004204	Jena Sternwarte	GM	155	50.9267	11.5842	1827-2014	10	178
AU000005010	Kremsmuenster	AU	383	48.05	14.1331	1876-2014	3	136
USC00044997	Livermore	CA	149	37.6666	-121.767	1903-2014	2	110
SIM00014015	Ljubljana	SL	299	46.0656	14.5169	1900-2014	0	115
SZ000009480	Lugano	SZ	300	46.01	8.9667	1901-2014	0	114
086071 *	Melbourne	AS	31.2	-37.8075	144.9700	1856-2013	0	158
USC00025467	Mesa	AZ	374.9	33.4166	-111.867	1897-2014	3	115
ITE00100554	Milano	IT	150	45.4717	9.1892	1858-2008	0	151
GM000004199	Muenchen	GM	515	48.1642	11.5442	1879-2014	4	132
†	Oxford	UK	63	51.77	-1.27	1853-2008	0	156
‡	Padova	IT	12	45.398	11.880	1725-2013	17	272
FR000007747	Perpignan	FR	42	42.7381	2.8731	1901-2014	2	112
ITE00115584	Pesaro	IT	11	43.9108	12.9042	1871-2008	1	137
-	Philadelphia	PA	11	39.95	-75.15	1901-2006	0	106
NLE00101991	Putten	NL	14	52.2517	5.62	1868-2014	0	147
USC00027281	Roosvelt	AZ	672.1	33.6666	-111.15	1906-2014	5	104
SF000208660	Royal Obs.	SF	40	-33.93	18.48	1850-1997	23	125
SZ000002220	Saentis	SZ	2502	47.25	9.35	1901-2014	0	114
SZ000006717	C.d.G. San Bernard	IT	2472	45.8667	7.1667	1901-2014	9	105
†	Sheffield	UK	115	53.38	1.494	1883-2008	0	126
ASN00066062	Sydney	AS	39	-33.8607	151.205	1859-2014	0	156
SF000227590	Worcester	SF	270	-33.617	19.467	1880-2008	13	116
HR000142360	Zagreb	HR	157	45.8167	15.9781	1862-2004	0	143
SZ000003700	Zurich	SZ	556	47.3831	8.5667	1901-2014	0	114

*Australian Bureau of Meteorology.

†British Atmospheric Data Centre, UK Meteorological Office

‡See [MZ15]

Appendix B

Appendix to Chapter 3

B.1 Details on the Bayesian GEV and POT Models Implemented

Here I briefly review the main extreme value statistical models consistently with the notation used in Chapter 3. The two approaches detailed below are the Generalized Extreme Value distribution used as a model for block maxima series, and the Peak Over Threshold (POT) model with the frequency of excesses over threshold described through a Poisson point process. For a complete discussion, see Coles [Col01] or De Haan and Ferreira [DHF07]. These models are commonly used for EV analysis and statistical software that implement these technique is available, such as for example the *extRemes* R package [GK^{+}16]. The Bayesian (Hamiltonian Monte Carlo) estimation for the GEV and POT models used in our study is described below and implemented in the *hmevr* R package, available at `https://github.com/EnricoZorzetto/hmevr`.

B.2 The Generalized Extreme Value Distribution

The (GEV) distribution [VM36] has cdf

$$Pr\left(Y \leq y\right) = F_{GEV}\left(y \mid \mu, \sigma, \xi\right) = \exp\left\{ - \left(1 + \frac{\xi}{\sigma}\left(y - \mu\right)\right)_{+}^{-1/\xi} \right\}. \tag{B.1}$$

where $\mu \in R$ and $\sigma \in R^+$ are the location and scale parameters respectively, while $\xi \in R$ is a shape parameter, and $(\cdot)_+ = \max\{0, \cdot\}$. Depending on the value of ξ, the GEV family encompasses a double exponential, an heavy-tailed, and an upper-bounded distribution. The GEV parameters are generally estimated by means of Maximum Likelihood (ML), Penalized ML [MS00a], L-Moments [Hos90] or Bayesian methods [CT96, CPS03]. Generally the L-moments approach performs better than ML in the case of small samples. Bayesian methods generally allow for a better characterization of the variability of estimated extreme values [CT96, CP96, CPS03, ST04]. Here I use Bayesian methods for fitting the GEV model (I implemented a Stan model, sampling from the posterior using the Hamiltonian Monte Carlo sampler as done for HMEV). This yields Bayesian probability intervals for the GEV quantiles for any given return time. In this analysis, I elicit the prior distributions for the three GEV parameters as follows: For the shape parameter, I select a normal distribution centered in 0.114 with a standard deviation $\sigma = 0.125$. This choice matches the expected value suggested globally for daily rainfall extremes [Kou04], while the overall shape of the prior distribution closely matches the *Geophysical prior* proposed by Martins and Stedinger [MS00a] in order to guide inference towards realistic values of the shape parameter in the present application. For the shape and scale parameters, I select informative gamma prior distributions centered around the mean and standard deviation of the annual maxima samples respectively, in order to exploit the available knowledge on the expected value and characteristic variability of the maxima observed at each site.

B.3 The Peak Over Threshold Method

The GEV distribution also arises as limiting model for the block maxima of a point process with Poisson-distributed arrival of events, and with magnitudes distributed according to a Generalized Pareto Distribution (GPD). The GPD model is often used to model exceedance over a high threshold [DS90]; If the random variable $X \in R^{+}$ represents the daily rainfall magnitude, and $Y = X - q$ its excess over a fixed threshold q, the cumulative distribution of excesses over threshold reads

$$\begin{aligned} Pr\left(Y > y\right) = Pr\left(X > y + q | X > q\right) = \frac{1 - F(y+q)}{1 - F(q)} = \\ = 1 - F_{GPD}\left(x | q, \beta, \kappa\right) = \left(1 + \frac{\kappa}{\beta}\left(x - q\right)\right)^{-1/\kappa} \end{aligned} \tag{B.2}$$

where the GPD scale and shape parameters are $\sigma \in R^{+}$ and $\xi \in R$ respectively. In this case, the distribution of block maxima reads

$$\begin{aligned} F_{PP}\left(x\right) = \sum_{n=0}^{\infty} p_n(n|\lambda) F_{GPD}\left(x | q, \beta, \kappa\right) = \\ = 1 - \sum_{n=0}^{\infty} \frac{\lambda^n e^{-\lambda}}{n!} \left(1 + \frac{\kappa}{\beta}\left(x - q\right)\right)^{-1/\kappa} = F_{GEV}\left(x | \mu, \sigma, \xi\right) \end{aligned} \tag{B.3}$$

where the parameters of the GEV are obtained as follows: $\kappa = \xi$, $\beta = \sigma + \xi\left(q - \mu\right)$, and $\lambda = \left(1 + \frac{\xi}{\beta}\left(q - \mu\right)\right)^{-1/\xi}$. Here I use Bayesian inference for this Poisson-GPD model for threshold exceedances. I elicit the prior distribution for the model parameters as follows: For the Pareto shape parameter, I choose the same informative prior elicited for the GEV - annual maxima model. For the Poisson rate, I select a weakly informative prior distribution centered on 4 events/year, which appears reasonable value since I select the threshold based on a fixed number of average exceedances in each block. For the analysis over the entire USHCN dataset, I choose an automated threshold such that 5% of the non-zero values in the entire record are above threshold.

Once I draw from the posterior samples for these parameters $\lambda^{(s)}$, $\beta^{(s)}$, and $\kappa^{(s)}$, for each posteror draw $s = 1, \ldots, S$, I can compute the corresponding posterior probability distributions for the GEV parameters through the relations

$\mu^{(s)} = q + \frac{\beta^{(s)}}{\kappa^{(s)}}\left(\lambda^{(s)\kappa^{(s)}} - 1\right)$, $\sigma^{(s)} = \beta^{(s)}\lambda^{(s)\kappa^{(s)}}$, $\xi^{(s)} = \kappa^{(s)}$

B.4 Definitions of *lppd* and *lpml*

Evaluating the predictive accuracy of extreme value models in estimating the right tail of the distribution is indeed an inherently challenging task, as high quantiles are, by definition, poorly represented in the available samples. For this reason, cross validation techniques are rarely used to assess the performance of fitted extreme value models. In the analysis discussed in Chapter 3, however, I harnessed the considerable length of the synthetic data sets available here in order to extensively test the performance of different methods using both in-sample and out-of-sample validation techniques. The log pointwise predictive density (lppd) [GCS+13] computed both for the in-sample data and for the out-of-sample data is used as a measure of global performance of the models. For a sample of block maxima y_i, $i = 1, \ldots, M$, this measure can be directly estimated from S MCMC draws as

$$\widehat{lppd} = -\frac{1}{M}\sum_{i=1}^{M}\log\left(\frac{1}{S}\sum_{s=1}^{S}p(y_i|\theta^{(s)})\right) \tag{B.4}$$

where here I multiply by the factor $-1/M$ so as to reduce its variation with sample size M. This quantity, if computed for in-sample annual maxima data y_i, $i = 1, \ldots, M$ is expected to overestimate the expected log predictive density (elpd) for the same data points. This overestimation is generally corrected by quantifying the overfit of the model using some estimate of its effective number of parameters. Common corrections used in practice include the Deviance Information Criterion (DIC), the Watanabe-Aikake information criterion (WAIC) [GCS+13], or leave-one-out techniques such as the log posterior marginal likelihood (LPML) [GD94], or a leave-one-out based on Pareto Smoothed Importance Sampling (PSIS) [VGG17].

Here I used the logarithm of the pseudo-marginal likelihood (lpml), a convenient index that directly accounts, at no additional computational cost, for a leave-one-out cross validation measure [GD94]. Notably, since the lpml approximates the expected log pointwise predictive density, the difference between the in-sample lppd and the lpml represents the number of effective parameters of a model [VGG17] and thus will be used to quantify the tendency of different models to overfitting. The log posterior marginal likelihood is

$$lpml = -\frac{1}{M} \sum_{i=1}^{M} \log \left(CPO_i \right) \tag{B.5}$$

where again I multiply by the factor $-1/M$ to reduce its variation with sample size. Here CPO_i is the *Conditional Predictive Ordinate* statistics introduced by [GDC92] and [GD94], which estimates the probability of observing a value y_i given that $\mathbf{y}_{-i}$ has been observed. CPO_i can be obtained as follows:

$$CPO_i = \left\{ \int \frac{1}{p\left(y_i \mid \theta\right)} p\left(\theta \mid \mathbf{y_{-i}}\right) d\theta \right\}^{-1} \tag{B.6}$$

The CPO can be computed as the geometric mean of the likelihood of the data (annual maxima y_i) given the model. Sampling from the posterior, one can compute

Table B.1: Values of the constants used to elicit the prior hyperparameters of the model for event magnitudes.

Parameter	Value	
$\alpha_{\mu\gamma0} =$	$i_{\mu\gamma0}$	
$\beta_{\mu\gamma0} =$	$i_{\mu\gamma0} \cdot e_{\mu\gamma0}$	
$\alpha_{\mu\delta0} =$	$i_{\mu\delta0}$	
$\beta_{\mu\delta0} =$	$i_{\mu\delta0} \cdot e_{\mu\delta0}$	
$\alpha_{\sigma\gamma0} =$	$i_{\sigma\gamma0}$	
$\beta_{\sigma\gamma0} =$	$i_{\sigma\gamma0} \cdot e_{\sigma\gamma0} \cdot v_{\sigma\gamma0}$	
$\alpha_{\sigma\delta0} =$	$i_{\sigma\delta0}$	
$\beta_{\sigma\delta0} =$	$i_{\sigma\delta0} \cdot e_{\sigma\delta0} \cdot v_{\sigma\delta0}$	
Constant	Value	Meaning
$i_{\mu\gamma0}$	10	shape informativeness
$i_{\mu\delta0}$	10	scale informativeness
$i_{\sigma\gamma0}$	10	shape informativeness
$i_{\sigma\delta0}$	10	scale informativeness
$e_{\mu\gamma0}$	0.7	expected value shape
$e_{\mu\delta0}$	$10/\Gamma\left(1+1/0.7\right)$	expected value scale
$v_{\sigma\gamma0}$	0.05	expected variability shape
$v_{\sigma\delta0}$	0.25	expected variability scale

CPO_i as follows:

$$\widehat{CPO_i} = \left[\frac{1}{S}\sum_{s=1}^{S}\frac{1}{p\left(y_i \mid \theta^{(s)}\right)}\right]^{-1} \tag{B.7}$$

and thus $\widehat{lpml} = -\frac{1}{M}\sum_{i=1}^{M}\log\left(\widehat{CPO_i}\right)$.

B.5 Effective Number of Parameters for USHCN Station Data

As shown in Figure B.1, I report the effective number of paramaters of the EV models fit to the observed station data, computed as the difference between *lppd* and *lpml*. Additionally, I include the same quantity computed in an alternative way (using the

Table B.2: Summary of 4 model specifications used to generate synthetic datasets in the simulation study.

Model for x_{ij}	parameters
GP	$\xi = 0.1, \quad \sigma = 8$
GAM	$\alpha = 1.2, \quad \beta = 0.12$
WEI	$\gamma = 0.6, \quad \delta = 8$
WEI_G	$\mu_\delta = 6, \quad \sigma_\delta = 1, \quad \mu_\gamma = 1, \quad \sigma_\gamma = 0.1,$
Model for n_j	**parameters**
BBN	$\mu_n = 100, \quad \sigma_n^2 = 150$

Table B.3: Summary of the model for the number of arrivals.

Model	Outer level	Inner Level	Prior
Binomial	$n_j \sim Bin(\pi_0)$		$\pi_0 \sim Beta\,(2,2)$
Beta Binomial	$n_j \sim Bin(\pi_0)$	$\pi_0 \sim Beta\,(\alpha_n, \beta_n)$	$\mu_n\ \Gamma\,(10, 0.1)$ $\omega_n\ \Gamma\,(0.2, 0.2)$ $\alpha_n = \frac{\left(1 - \frac{(\omega_n+1)}{(N_t-\mu_n)}\right)}{\left(\frac{(\omega_n+1)}{(N_t-\mu_n)} N_t - 1\right)}$ $\beta_n = \frac{\alpha_n}{\mu_n}\,(N_t - \mu_n)$

Watanabe-Aikake information criterion, WAIC, in Figure B.2 - see e.g. Gelman et al. [GCS+13]) to show that this result is not overly sensitive to the specific choice of information criterion adopted.

Here I also show the effective number of parameters computed in the simulation study for a sample size of 10 and 20 years, in Figures B.3 and B.4 respectively. These complement the result obtained for a sample size of $M_{train} = 50$ years, which is featured in Chapter 3.

To complete the results in Chapter 3, here I include results of the simulation study obtained for a sample size of $M_{train} = 20$ years. These include fractional square error (Figure B.5), mean bias (Figure B.5), and average width of credibility intervals for the extreme value quantiles (Figure B.6).

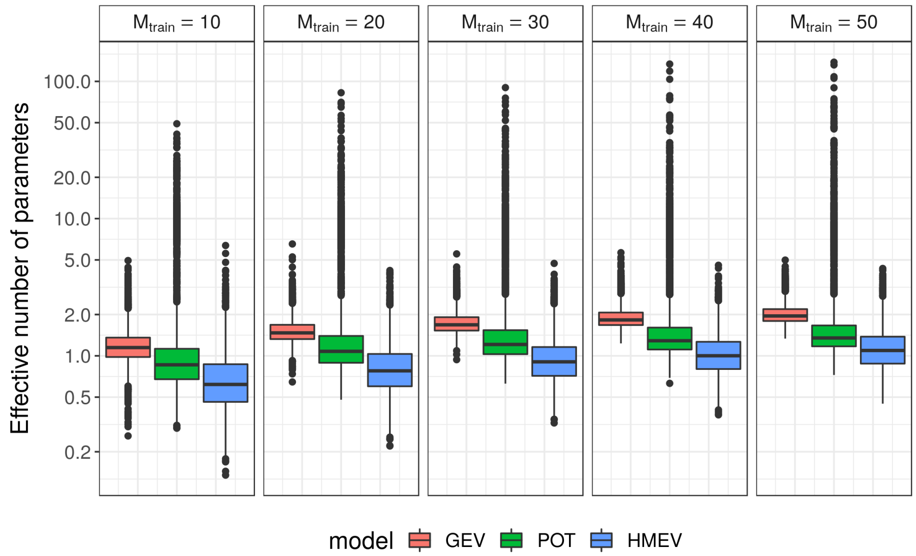


Figure B.1: Effective number of parameters for the stations in the USHCN dataset.

B.6 Spatial Distribution of the Results for *lppd* and *FSE*

Here I provide a spatially-explicit representation of model performances by mapping, in Figures B.7 and B.8, the best model for each station as evaluated through the *lppd* and *FSE* measures respectively. Each metric is evaluated both for in-sample and out-of-sample data, and the result is averaged over the $R_g = 10$ resampling of the original data, according to the procedure described in Chapter 3. This representation of the results of the analysis shows again the interesting difference observed for the in-sample analysis, which tends to favor the POT method, and the out-of-sample results, where HMEV appears to be selected more often as preferred model. The frequency of HMEV being the model of choice is higher for smaller sample sizes, as discussed in Chapter 3.

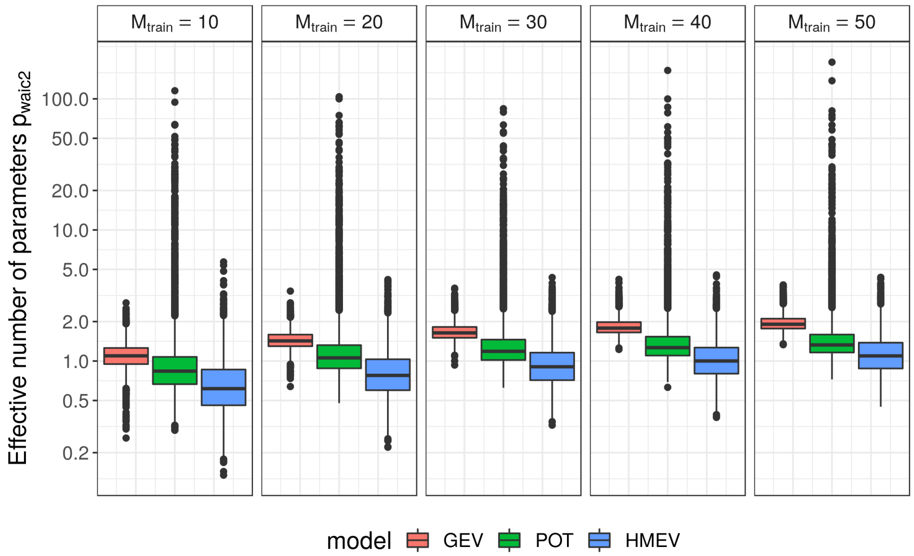


Figure B.2: Effective number of parameters for the stations in the USHCN dataset, computed using the Watanabe-Aikake information criterion (WAIC).

B.7 Examples of Fit to Simulated Samples

in Figures B.9, B.10, and B.11 I report examples of EV models fit to simulated data generated using the Weibull, Gamma, and Generalized Pareto specifications, respectively.

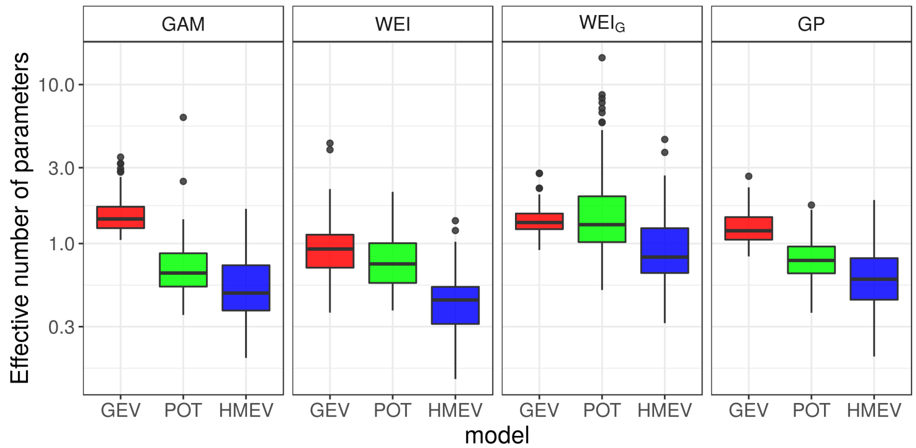


Figure B.3: Effective number of parameters computed in the simulation study for a sample size of 10 years.

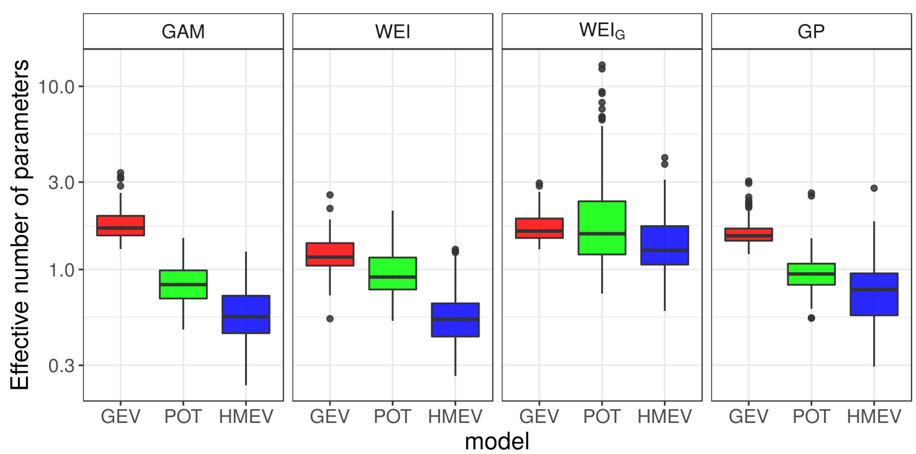
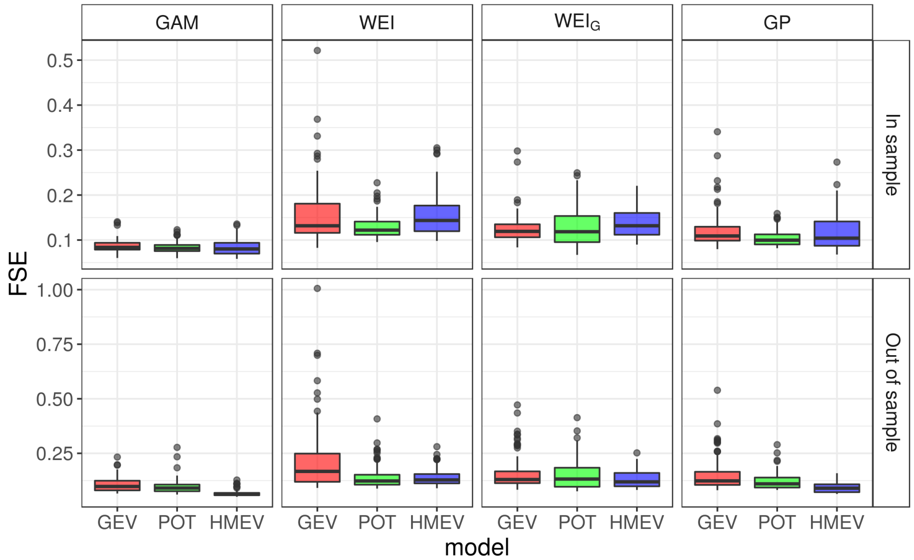


Figure B.4: Effective number of parameters computed in the simulation study for a sample size of 20 years.



[Fractional square error computed for the 4 different model specifications]Fractional square error computed for the 4 different model specifications for in-sample data (upper panels) and for out-of-sample data (lower panels), computed for sample size of 20 years.

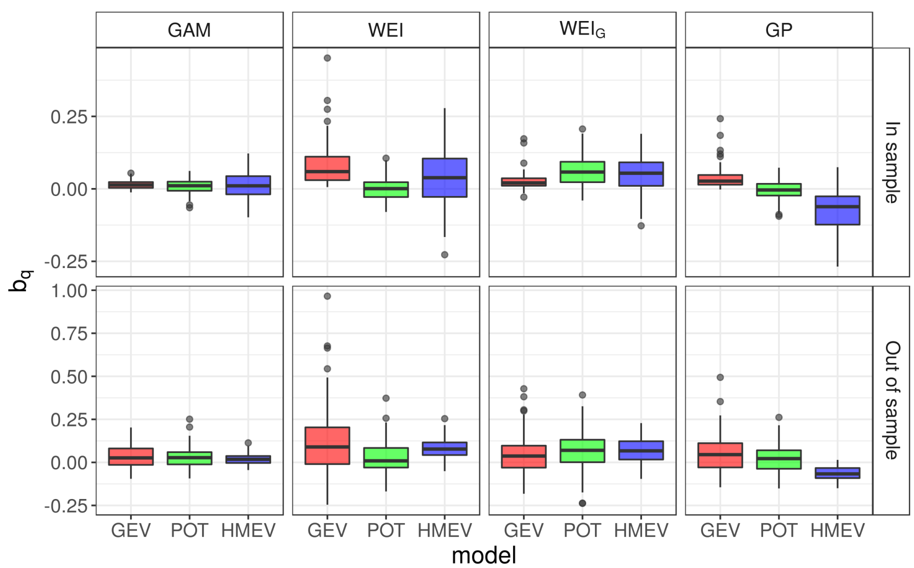


Figure B.5: Mean bias for the 4 different model specifications for in-sample data (upper panels) and for out-of-sample data (lower panels), computed for sample size of 20 years.

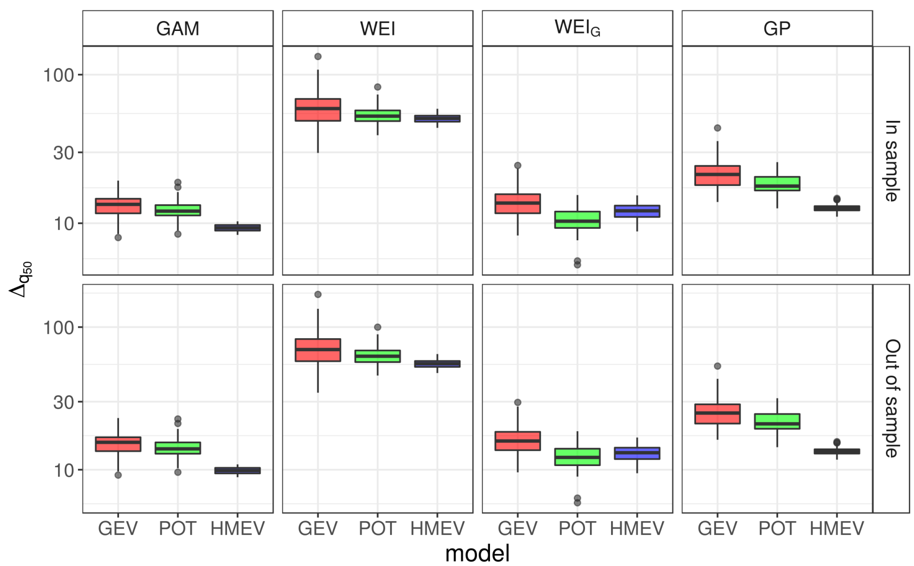


Figure B.6: Mean credibility interval width for the 4 different model specifications for in-sample data (upper panels) and for out-of-sample data (lower panels), computed for sample size of 20 years.

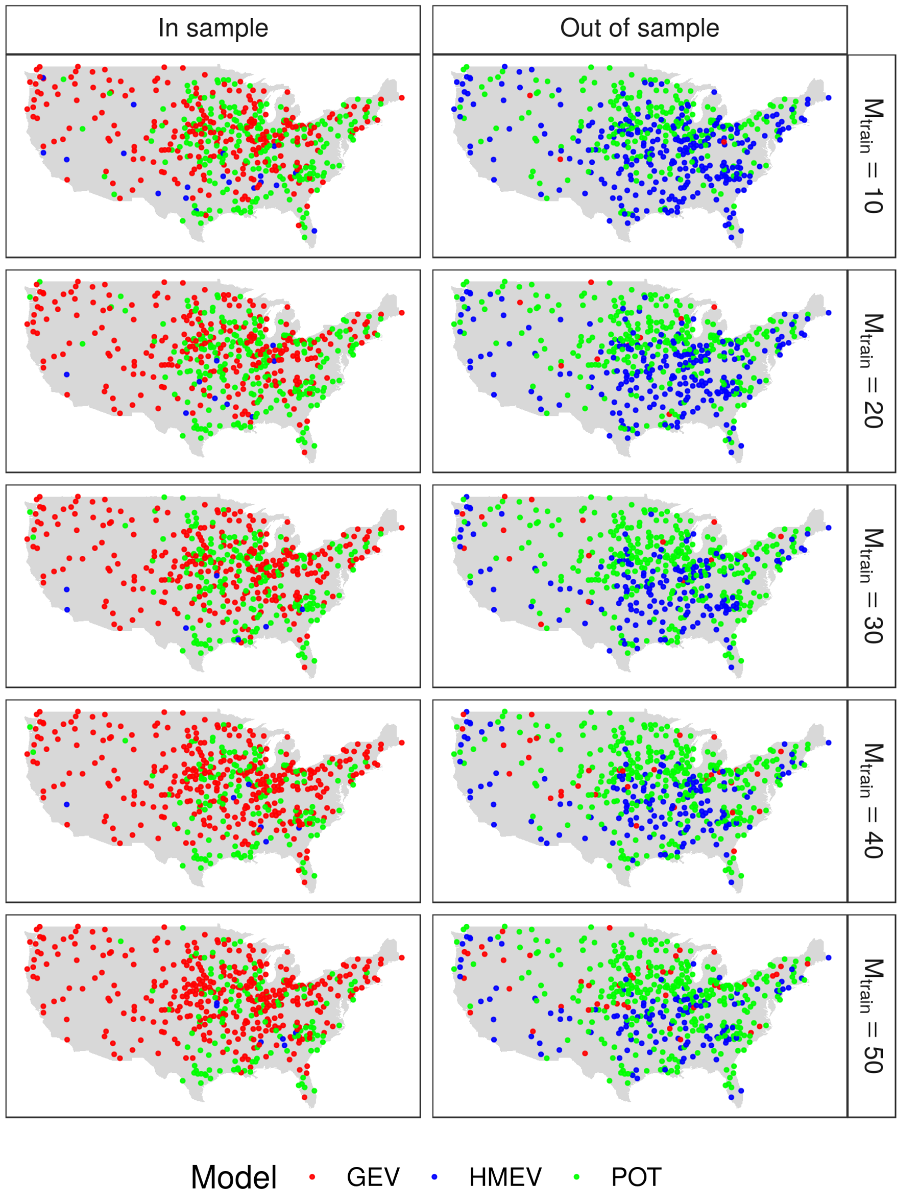


Figure B.7: Best model for each station, as evaluated through the lppd measure for in-sample and out-of-sample data.

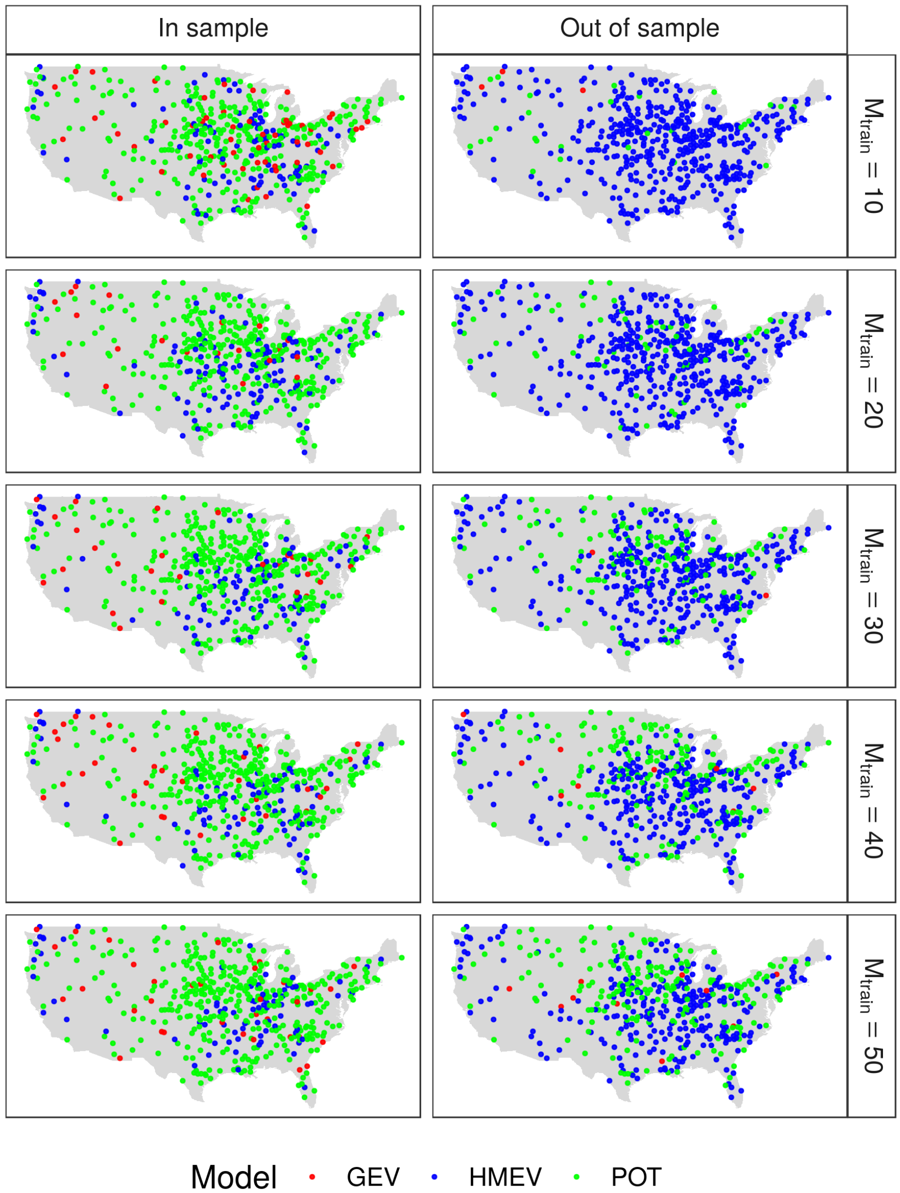


Figure B.8: Best model for each station, as evaluated through the fractional square error (FSE), evaluated for in-sample and out-of-sample data.

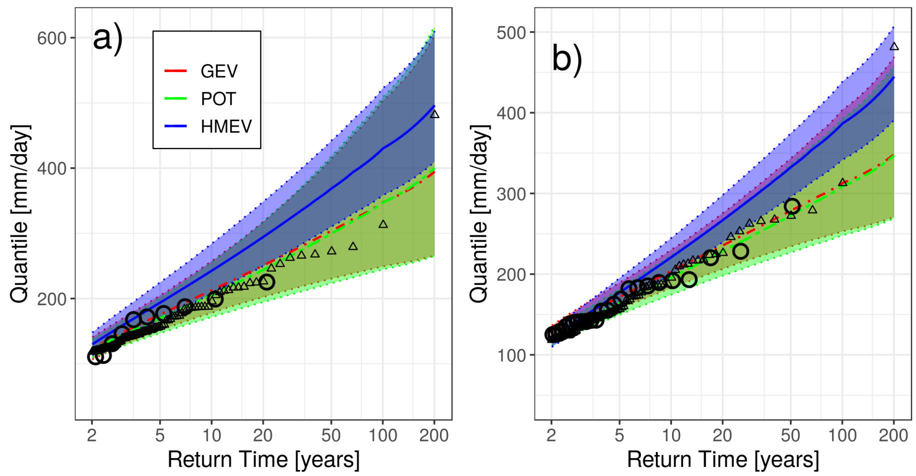


Figure B.9: Example of fit to samples of 20 (left panel) and 50 (right panel) yearly blocks of data generated according to the Weibull specification.

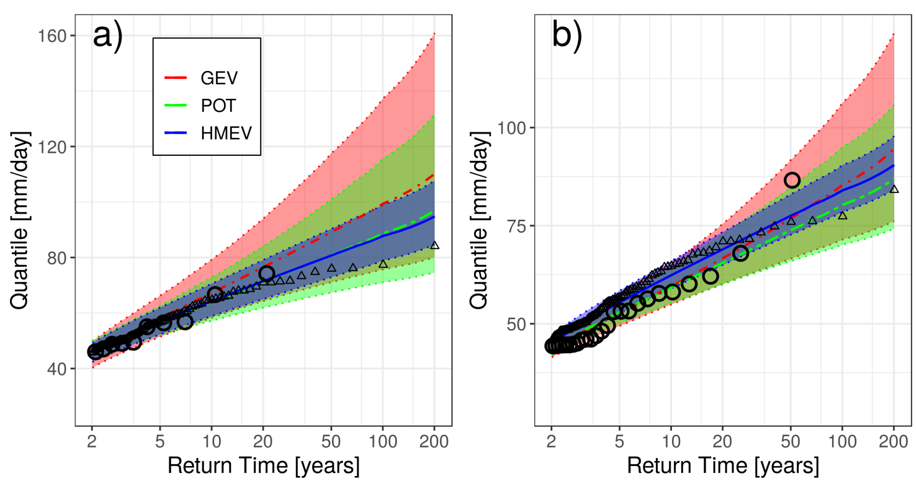


Figure B.10: Example of fit to samples of 20 (left panel) and 50 (right panel) yearly blocks of data generated according to the Gamma specification.

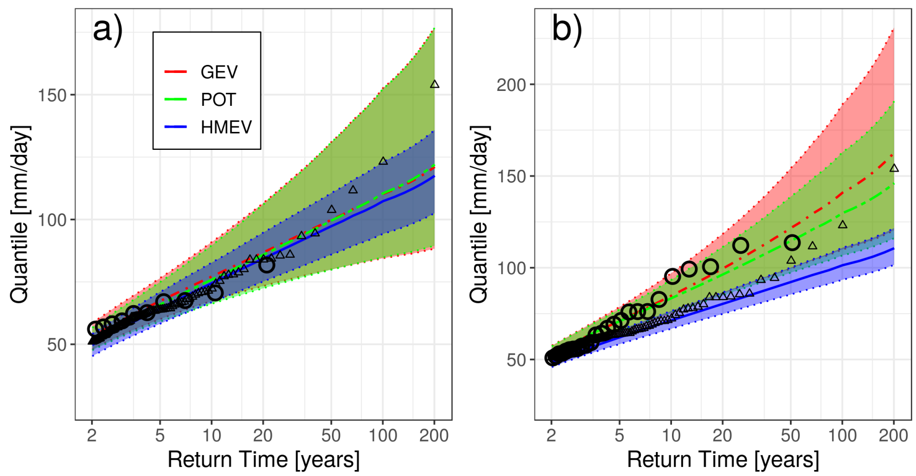


Figure B.11: Example of fit to samples of 20 (left panel) and 50 (right panel) yearly blocks of data generated according to the GPD specification.

Appendix C

Appendix to Chapter 4

C.1 Variance of the Compound Rainfall Process

The compound rainfall process (wet and dry periods) is characterized by a pdf $f_c(h)$ that has a finite atom of probability in $h = 0$, such that $f_c(h) = (1-p_r)\delta(h)+p_r f_r(h)$, where $f_r(h)$ is the pdf of wet events only, $\delta(h)$ is the Dirac delta function centered in 0, and p_r is the probability of a day being wet. Therefore, the mean of the compound process is $\mu_c = \mu_r p_r$ as there is no contribution from the atom of probability in zero. The variance is

$$\sigma_c^2 = E\left[(h-\mu_c)^2\right] = E\left[h^2\right]_c - E\left[h\right]_c^2 \tag{C.1}$$

where $E\left[\cdot\right]$ is the expected value operator. Therefore

$$\sigma_c^2 = (1-p_r)\int_0^{\nu} h^2\delta(h-0)dh + p_r\int_{\nu}^{+\infty} h^2 f_r(h)dh - E\left[h\right]_c^2 \tag{C.2}$$

in the limit $\nu \to 0$. Since the value of the first integral in eq. (C.2) is zero, one obtains:

$$\sigma_c^2 = p_r E\left[h^2\right]_r - p_r^2\mu_r^2 \tag{C.3}$$

from which, summing and subtracting $p_r\mu_r^2$

$$\sigma_c^2 = p_r E\left(\left[h^2\right]_r - \mu_r^2\right) + p_r\mu_r^2 - p_r^2\mu_r^2 \tag{C.4}$$

Thus proving eq. (4.5)

$$\sigma_c^2 = \sigma_r^2 p_r + \mu_r^2 (1 - p_r) p_r \tag{C.5}$$

C.2 Summary Description of the Downscaling Methodology

1. Extraction of the local lattice of TMPA 3b42 pixels QPEs time series at the 3-hourly time scale, centered over the location of interest.

2. Aggregation of TMPA data at the daily scale, construction of the time series of exceedances over the detection threshold, evaluation of the cross correlation between the TMPA QPEs time series, and minimization of eq. (4.16) in order to estimate the parameters α and ϵ defining the point correlation function.

3. Application of the procedure detailed in section 4.2.3 to estimate the quantity β_0 using Taylor's frozen turbulence hypothesis.

4. Evaluation of the the variance reduction function eq. (4.3) by integrating the correlation function $\rho(d; \alpha, \epsilon)$ at a point.

5. Estimation of the yearly parameters C_L and w_L by fitting the Weibull distribution to the TMPA QPE time series over the location of interest.

6. Downscaling of the yearly parameters of the Weibull distribution at a point (C_0 and w_0) using eqs. (4.10) and (4.11).

7. Numerical inversion of the MEVD non exceedance probability expression eq. (4.18) to compute extreme value quantiles $\hat{h}\,(Tr)$ for the desired return time.

C.3 Notation Used in Chapter 4

i **:** instantaneous rainfall rate at a point $[mm/hour]$

L_x, L_y **:** effective linear dimensions of a TMPA pixel $[km]$

L **:** characteristic linear dimension of a TMPA pixel $[km]$

T **:** generic time integration interval $[hours]$

X **:** generic linear characteristic averaging scale $[km]$

U **:** advection speed $[km/hour]$

T_d **:** (= 24 hours) daily time integration interval

h **:** daily rainfall accumulation at a point $[mm]$

h_L **:** daily rainfall accumulation at the pixel scale $[mm]$

p_r **:** yearly fraction of rainy days

μ_r^2 **:** mean of wet process $[mm]$

σ_r^2 **:** variance of wet process $[mm^2]$

$\gamma_0(L)$ **:** variance function

$\beta_0(L)$ **:** intermittency function

$\beta(L_1, L_2)$ **:** intermittency function between two generic scales

$\gamma(L_1, L_2)$ **:** variance function between two generic scales

$\rho(x, y)$ **:** auto correlation function at a point

d **:** distance between two points in space $[Km]$

$\rho_{h_L,h'_L}(d)$ **:** correlation of local averages

ρ_j **:** Pearson correlation between two TMPA time series at distance d_j.

Δ **:** Integral variance function

$F(h)$ **:** cumulative distribution function of ordinary rainfall

$\zeta(h)$ **:** cumulative distribution function of extreme rainfall

w **:** Weibull shape parameter

C **:** Weibull scale parameter $[mm]$

$\vec{\theta}$ **:** Generic set of parameters of the MEV distribution

N **:** yearly number of rainy days $[days]$

ϵ **:** auto correlation function scale parameter $[km]$

α **:** auto correlation function shape parameter

s **:** sample size $[years]$

T_r **:** return time $[years]$

FSE **:** fractional square error

SSE **:** sum of square errors

q **:** rainfall detection threshold $[mm]$

y **:** excesses over threshold $[mm]$

n_g **:** number of resamplings for bootstrapping

Appendix D

Appendix to Chapter 5

D.1 List of Symbols Used in Chapter 5

- S generic linear spatial scale, corresponding to spatial average over an area S^2.
- T generic integration time scale.
- L spatial scale of the QPE grid cell
- L_0 spatial scale of rain gauge measurements
- $\widetilde{h}_s$ daily rainfall accumulation magnitude at spatial averaging scale S.
- h_s *ordinary rainfall* amount (realization) at scale S, i.e. rainfall magnitude in excess of the threshold q.
- H_s *ordinary rainfall* amount (random variable).
- q threshold used to determine ordinary rainfall events.
- N_s average yearly number of events at the spatial scale S
- N_t number of daily rainfall observations in one year.
- p_{r_s} wet fraction at spatial scale S
- w_s Weibull shape parameter at spatial scale S
- C_s Weibull scale parameter at spatial scale S

- σ_s^2 variance of the rainfall process at spatial scale s
- $\beta_0 = \beta_0(L)$ intermittency function
- $\gamma_0 = \gamma_0(L)$ variance reduction function
- α shape parameter of the spatial correlation function
- ϵ transition parameter of the spatial correlation function
- d generic distance between two points in space
- $\rho_s(s_1, s_2) = \rho_s(d)$ spatial correlation function at spatial scale s and distance d
- ρ_{h_L, h'_L} correlation between local averages at scale L
- ρ_j areal-averaged correlation value sampled between grid cells located at distance d_j
- SSE Sum of Squared Errors
- $\alpha^{(L)}$ shape parameter of the correlation function obtained from fit to TMPA gridded data
- $\epsilon^{(L)}$ transition parameter of the correlation function obtained from fit to TMPA gridded data
- α_g shape parameter of the correlation function obtained from fit to rain gauge data
- ϵ_g transition parameter of the correlation function obtained from fit to rain gauge data

- $\gamma_{0,L}$ variance reduction function computed from TMPA-fitted correlation (areal-averaged).
- $\gamma_{0,d}$ variance reduction function computed from TMPA-downscaled correlation.
- $\gamma_{0,g}$ variance reduction function computed from rain-gauge-fitted correlation.
- $C_{0,g}$ Weibull scale parameter at a point, from rain gauge data
- $w_{0,g}$ Weibull scale parameter at a point, from rain gauge data
- $N_{0,g}$ Yearly number of events at a point, from rain gauge data
- $C_{0,d}$ Weibull scale parameter at a point, from downscaling
- $w_{0,d}$ Weibull scale parameter at a point, from downscaling
- $N_{0,d}$ Yearly number of events at a point, from downscaling
- $C_{0,c}$ Weibull scale parameter at a point, corrected using the QRF error model.
- $w_{0,c}$ Weibull scale parameter at a point, corrected using the QRF error model.
- $N_{0,c}$ Yearly number of events at a point, corrected using the QRF error model.
- η_z Relative error between downscaled and gauge-estimated values of a quantity ($z = C, w, N$ or γ)
- y response variable in the quantile regression forest algorithm, realization of the r.v. Y.
- x multivariate explanatory variable in the quantile regression forest algorithm, realization of the r.v. X.

- ω_i weights for the QRF algorithm.
- k number of trees.
- θ vector describe the tree structure.
- σ_e standard deviation of elevation.
- μ_e mean elevation.
- $H_s^{(m)}$ annual maxima daily rainfall at spatial scale s (random variable).
- $C_s^{(i)}$ Weibull scale parameter estimated for year i at spatial scale s.
- $w_s^{(i)}$ Weibull shape parameter estimated for year i at spatial scale s.
- $N_s^{(i)}$ Yearly number of events estimated for year i at spatial scale s.
- p_{ne} non exceedance probability.
- Tr return time or average recurrence interval.
- M number of years in the rainfall record used for extreme value analysis.

D.2 List of Acronyms Used in Chapter 5

- **TRMM** Tropical Rainfall Measuring Mission
- **TMPA** TRMM Multisatellite Precipitation Analysis
- **GPM** Global Precipitation Measurement [Mission]
- **IMERG** Integrated Multi-SatellitE Retrievals for GPM
- **CONUS** Conterminous United States

- **QPE** Quantitative Precipitation Estimates
- **MEV** Metatstatistical Extreme Value [Distribution]
- **GEV** Generalized Extreme Value [Distribution]
- **QRF** Quantile Regression Forest
- **RF** Random Forest
- **POT** Peak Over Threshold
- **AM** Annual Maxima
- **HPD** Hourly Precipitation Dataset
- **NOAA** National Oceanogaphic and Atmospheric Administration
- **NASA** National Aeronautics and Space Administration
- **SGGC** Set of gauged grid cells [used in the analysis]

Appendix E

Appendix to Chapter 6: Stable stratification and distortions of the inertial subrange

In general, stable stratification limits the onset and extent of the inertial subrange given its damping effect in the vertical direction [RMP15]. Here, I show that the scales for which these effects are relevant occur at scales larger than the inertial range examined here. The Ozmidov length scale [Ozm65] (originally suggested by Dougherty [Dou61] in 1961), is defined as the scale above which buoyancy forces significantly distort the spectrum of turbulence.

This length scale, sometimes labeled as the Dougherty-Ozmidov scale, can be expressed as

$$L_0 = \sqrt{\frac{\epsilon}{N^3}}, \tag{E.1}$$

where ϵ is, as before, the mean turbulent kinetic energy dissipation rate and N is the *Brunt Väisälä* frequency, defined as

$$N = \sqrt{\frac{g}{T}\frac{dT}{dz}}. \tag{E.2}$$

In the study used here, no information was provided about the actual mean potential temperature gradient dT/dz. However, an approximated estimate of L_0 for the runs collected in case of stable atmospheric stratification may be conducted. Note that only 4 runs follow this stability class as runs not meeting strict stationarity requirements were excluded from the analysis (and they were mainly collected in unstable atmospheric conditions). The mean dT/dz was computed using Monin- Obukhov similarity theory as

$$\frac{dT}{dz} = -\left(\frac{T^*}{K_v z}\right)\phi_T\left(\frac{z}{L}\right) \tag{E.3}$$

where $k_v = 0.41$ is the *von Karman* constant, $z = 5.1$ m is the distance from the ground, $T^* = \frac{\langle w'T'\rangle}{u^*}$, and for mildly stable stratification

$$\phi_T = \phi_m = 1 + 4.7\left(\frac{z}{L}\right). \tag{E.4}$$

The mean turbulent kinetic energy dissipation rate was computed as

$$\epsilon = \frac{u^{*3}}{k_v z}\left(\phi_m - \frac{z}{L}\right) \tag{E.5}$$

Figure E.1(A) shows that the quantity

$$I_s = \frac{I_w u^* \phi_m}{k_v z} = constant \simeq 0.4 \tag{E.6}$$

is almost constant across runs and exhibits a value slightly lower than the expected 0.4.

The estimated values of the dimensionless Ozmidov number $L_0/\left(I_w u^* \phi_m\right)$ are reported in Figure E.1(B). L_0 decreases with increasing stability ζ as the effect of buoyancy is felt by eddies of sizes progressively smaller. However, the values of the Ozmidov scale are consistently larger than the integral scale of the flow I_w for the 4 stable runs here. Hence, ignoring distortions caused by stable stratification on inertial subrange scales for the aforementioned 4 runs may be deemed plausible.

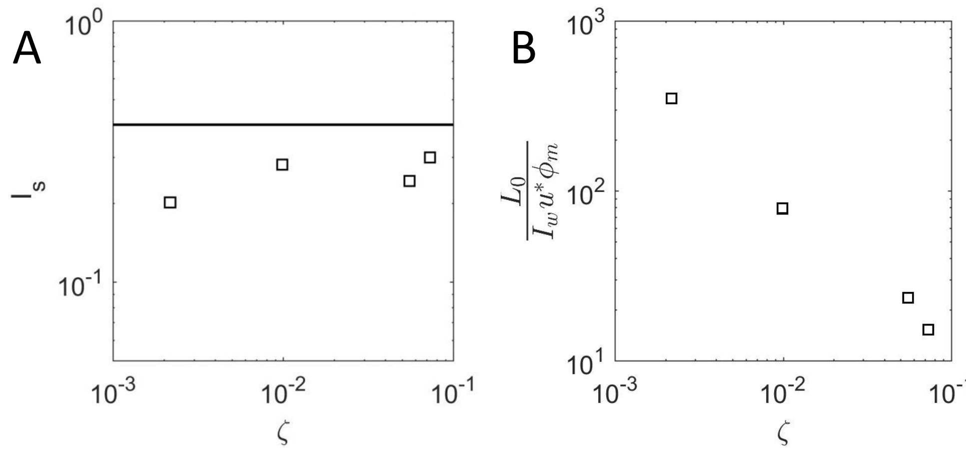


Figure E.1: (A) Quantity Is and its expected value 0.4 (black horizontal line) for the 4 stable runs in the dataset. (B) Normalized Ozmidov length for the same runs.

Bibliography

[AA02] T Arimitsu and N Arimitsu. Tsallis statistics and turbulence. *Chaos, Solitons & Fractals*, 13(3):479–489, 2002.

[ABS+11] A AghaKouchak, A Behrangi, S Sorooshian, K Hsu, and Eyal Amitai. Evaluation of satellite-retrieved extreme precipitation rates across the central united states. *Journal of Geophysical Research: Atmospheres*, 116(D2), 2011.

[ACFVA79] RA Antonia, AJ Chambers, CA Friehe, and CW Van Atta. Temperature ramps in the atmospheric surface layer. *Journal of the Atmospheric Sciences*, 36(1):99–108, 1979.

[AE09] Christopher Amante and Barry W Eakins. Etopo1 arc-minute global relief model: procedures, data sources and analysis. 2009.

[AHGA84] RA Antonia, EJ Hopfinger, Y Gagne, and F Anselmet. Temperature structure functions in turbulent shear flows. *Physical Review A*, 30(5):2704, 1984.

[AK05] Eduardo G Altmann and Holger Kantz. Recurrence time analysis, long-term correlations, and extreme events. *Physical Review E*, 71(5):056106, 2005.

[AMNB12] Amir AghaKouchak, Ali Mehran, Hamidreza Norouzi, and Ali Behrangi. Systematic and random error components in satellite precipitation data sets. *Geophysical Research Letters*, 39(9), 2012.

[AP91] William E Asher and James F Pankow. Prediction of gas/water mass transport coefficients by a surface renewal model. *Environmental science & technology*, 25(7):1294–1300, 1991.

[AS08] Richard P Allan and Brian J Soden. Atmospheric warming and the amplification of precipitation extremes. *Science*, 321(5895):1481–1484, 2008.

[AUG+12] Eyal Amitai, Carl L Unkrich, David C Goodrich, Emad Habib, and Bryson Thill. Assessing satellite-based rainfall estimates in semiarid watersheds using the usda-ars walnut gulch gauge network and trmm pr. *Journal of Hydrometeorology*, 13(5):1579–1588, 2012.

[Aul01] Milkha S Aulakh. Methane emissions from rice fields-quantification, mechanisms, role of management, and mitigation options. *Adv. Agron.*, 70:193–260, 2001.

[BAK17] Md Abul Ehsan Bhuiyan, Emmanouil N Anagnostou, and P-E Kirstetter. A nonparametric statistical technique for modeling overland tmi (2a12) rainfall retrieval error. *IEEE Geoscience and Remote Sensing Letters*, 14(11):1898–1902, 2017.

[BBIR13] G. Botter, S. Basso, I. Rodriguez Iturbe, and A. Rinaldo. Resilience of river flow regimes. *Proc. Natl. Acad. Sci. USA.*, 110(32):12925–12930, 2013.

[BC03] Christian Beck and Ezechiel GD Cohen. Superstatistics. *Physica A: Statistical mechanics and its applications*, 322:267–275, 2003.

[BCC+04] Kenneth A Byrne, Bogdan Chojnicki, Torben R Christensen, Matthias Drosler, Steve Frolking, Anders Lindroth, Julia Mailhammer, Nils Malmer, Pirkko Selin, Jukka Turunen, et al. Eu peatlands: Current carbon stocks and trace gas fluxes. Technical Report Report 4/2004 to 'Concerted action: Synthesis of the European Greenhouse Gas Budget, Geosphere-Biosphere Centre, University of Lund, Sweden, 2004.

[BD72] JA Bullin and AE Dukler. Random eddy models for surface renewal: formulation as a stochastic process. *Chemical Engineering Science*, 27(2):439–442, 1972.

[BDH74] August A Balkema and Laurens De Haan. Residual life time at great age. *The Annals of probability*, pages 792–804, 1974.

[BDS+12] Dennis Baldocchi, Matteo Detto, Oliver Sonnentag, Joe Verfaillie, Yit Arn Teh, Whendee Silver, and N Maggi Kelly. The challenges of measuring methane fluxes and concentrations over a peatland pasture. *Agricultural and Forest Meteorology*, 153:177–187, 2012.

[Bel87] Thomas L Bell. A space-time stochastic model of rainfall for satellite remote-sensing studies. *Journal of Geophysical Research: Atmospheres*, 92(D8):9631–9643, 1987.

[BG14] Paola Bortot and Carlo Gaetan. A latent process model for temporal extremes. *Scandinavian Journal of Statistics*, 41(3):606–621, 2014.

[BG16] Paola Bortot and Carlo Gaetan. Latent process modelling of threshold exceedances in hourly rainfall series. *Journal of agricultural, biological, and environmental statistics*, 21(3):531–547, 2016.

[BH93] Karen B Bartlett and Robert C Harriss. Review and assessment of methane emissions from wetlands. *Chemosphere*, 26(1-4):261–320, 1993.

[BHRM18] C Bracken, KD Holman, B Rajagopalan, and Hamid Moradkhani. A bayesian hierarchical approach to multivariate nonstationary hydrologic frequency analysis. *Water Resources Research*, 54(1):243–255, 2018.

[BJKK12] Anne-Laure Boulesteix, Silke Janitza, Jochen Kruppa, and Inke R König. Overview of random forest methodology and practical guidance with emphasis on computational biology and bioinformatics. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 2(6):493–507, 2012.

[BLG+13] S Beetz, H Liebersbach, S Glatzel, G Jurasinski, U Buczko, and H Höper. Effects of land use intensity on the full greenhouse gas balance in an atlantic peat bog. *Biogeosciences*, 10(2):1067–1082, 2013.

[BLNZ95] Richard H Byrd, Peihuang Lu, Jorge Nocedal, and Ciyou Zhu. A limited memory algorithm for bound constrained optimization. *SIAM Journal on Scientific Computing*, 16(5):1190–1208, 1995.

[BMR93] JL Bubier, TR Moore, and NT Roulet. Methane emissions from wetlands in the midboreal region of northern ontario, canada. *Ecology*, 74(8):2240–2254, 1993.

[BNB92] JM Baker, JM Norman, and WL Bland. Field-scale application of flux measurement by conditional sampling. *Agricultural and Forest Meteorology*, 62(1-2):31–52, 1992.

[BO90] Joost A Businger and Steven P Oncley. Flux measurement with conditional sampling. *Journal of Atmospheric and Oceanic Technology*, 7(2):349–352, 1990.

[BRC+16] Cameron Bracken, Balaji Rajagopalan, Linyin Cheng, Will Kleiber, and Subhrendu Gangopadhyay. Spatial bayesian hierarchical modeling of precipitation extremes over a large domain. *Water Resources Research*, 52(8):6643–6655, 2016.

[Bre01] Leo Breiman. Random forests. *Machine learning*, 45(1):5–32, 2001.

[Bro86] Fred V Brock. A nonlinear filter to remove impulse noise from meteorological data. *Journal of Atmospheric and Oceanic Technology*, 3(1):51–58, 1986.

[BRS+02] Maurício JA Bolzan, Fernando M Ramos, Leonardo DA Sá, Camilo Rodrigues Neto, and Reinaldo R Rosa. Analysis of fine-scale canopy turbulence within and above an amazon forest using tsallis' generalized thermostatistics. *Journal of Geophysical Research: Atmospheres*, 107(D20), 2002.

[Bru65] Wilfried Brutsaert. A model for evaporation as a molecular diffusion process into a turbulent atmosphere. *Journal of Geophysical Research*, 70(20):5017–5024, 1965.

[Bru75] Wilfried Brutsaert. A theory for local evaporation (or heat transfer) from rough and smooth surfaces at ground level. *Water resources research*, 11(4):543–550, 1975.

[Bru13] Wilfried Brutsaert. *Evaporation into the atmosphere: theory, history and applications*, volume 1. Springer Science & Business Media, 2013.

[BS15] Vidhi Bharti and Charu Singh. Evaluation of error in trmm 3b42v7 precipitation estimates over the himalayan region. *Journal of Geophysical Research: Atmospheres*, 120(24):12458–12473, 2015.

[BT98] Paola Bortot and Jonathan A Tawn. Models for the extremes of markov chains. *Biometrika*, 85(4):851–867, 1998.

[CBD+13] Dario Camuffo, C Bertolin, N Diodato, C Cocheo, M Barriendos, F Dominguez-Castro, E Garnier, MJ Alcoforado, and MF Nunes. Western mediterranean precipitation over the last 300 years from instrumental observations. *Climatic change*, 117(1-2):85–101, 2013.

[CFC96] CA Clayson, CW Fairall, and JA Curry. Evaluation of turbulent fluxes at the ocean surface using surface renewal theory. *Journal of Geophysical Research: Oceans*, 101(C12):28503–28513, 1996.

[CGH+17] Bob Carpenter, Andrew Gelman, Matthew D Hoffman, Daniel Lee, Ben Goodrich, Michael Betancourt, Marcus Brubaker, Jiqiang Guo, Peter Li, and Allen Riddell. Stan: A probabilistic programming language. *Journal of statistical software*, 76(1), 2017.

[CH04] Nicholas J Cook and R Ian Harris. Exact and general ft1 penultimate distributions of extreme wind speeds drawn from tail-equivalent weibull parents. *Structural Safety*, 26(4):391–420, 2004.

[CHG+13] Sheng Chen, Yang Hong, Jonathan J Gourley, George J Huffman, Yudong Tian, Qing Cao, Bin Yong, Pierre-Emmanuel Kirstetter, Junjun Hu, Jill Hardy, et al. Evaluation of the successive v6 and v7 trmm multisatellite precipitation analysis over the continental united states. *Water Resources Research*, 49(12):8174–8186, 2013.

[Chi93] Emily SC Ching. Probability densities of turbulent temperature fluctuations. *Physical Review Letters*, 70(3):283, 1993.

[CI88] DR Cox and Valerie Isham. A simple spatial-temporal model of rainfall. In *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, volume 415, pages 317–328. The Royal Society, 1988.

[CKO02] PSP Cowpertwait, CG Kilsby, and PE O'Connell. A space-time neyman-scott model of rainfall: Empirical analysis of extremes. *Water Resources Research*, 38(8), 2002.

[CLMV00] A Celani, A Lanotte, A Mazzino, and M Vergassola. Universality and saturation of intermittency in passive scalar turbulence. *Physical Review Letters*, 84(11):2385, 2000.

[CNBL97] Wenjun Chen, Michael D Novak, T Andrew Black, and Xuhui Lee. Coherent eddies and temperature structure functions for three contrasting surfaces. part i: Ramp model with finite microfront time. *Boundary-Layer Meteorology*, 84(1):99–124, 1997.

[CNN07] Daniel Cooley, Douglas Nychka, and Philippe Naveau. Bayesian spatial modeling of extreme precipitation return levels. *Journal of the American Statistical Association*, 102(479):824–840, 2007.

[Col01] Stuart Coles. *An introduction to statistical modeling of extreme values*, volume 208. Springer, 2001.

[CP96] Stuart G Coles and Elwyn A Powell. Bayesian methods in extreme value modelling: a review and new developments. *International Statistical Review/Revue Internationale de Statistique*, pages 119–136, 1996.

[CPS03] Stuart Coles, Luis Raúl Pericchi, and Scott Sisson. A fully probabilistic approach to extreme rainfall modeling. *Journal of Hydrology*, 273(1-4):35–50, 2003.

[CSY20] Salvatore Calabrese, Rodolfo Souza, and Jun Yin. Role of drainage timing on mitigating methane emissions from rice paddy fields. *Earth and Space Science Open Archive*, page 22, 2020.

[CT18] Thomas M Cover and Joy A Thomas. *Elements of Information Theory.* John Wiley & Sons, 2012, p.18.

[CT90] S.G. Coles and J.A. Tawn. Statistics of coastal flood prevention. *Philos Trans A Math Phys Eng Sci*, 332(1627):457–476, 1990.

[CT96] Stuart G Coles and Jonathan A Tawn. A bayesian analysis of extreme rainfall data. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 45(4):463–478, 1996.

[CW18] Lajiao Chen and Lizhe Wang. Recent advance in earth observation big data for hydrology. *Big Earth Data*, 2(1):86–107, 2018.

[Dan51] PV Danckwerts. Significance of liquid-film coefficients in gas absorption. *Industrial & Engineering Chemistry*, 43(6):1460–1467, 1951.

[Dav84] Anthony C Davison. Modelling excesses over high thresholds, with an application. In *Statistical extremes and applications*, pages 461–482. Springer, 1984.

[Dei00] Roberto Deidda. Rainfall downscaling in a space-time multifractal framework. *Water Resources Research*, 36(7):1779–1794, 2000.

[DHF07] Laurens De Haan and Ana Ferreira. *Extreme value theory: an introduction*. Springer Science & Business Media, 2007.

[DJRRI15] Manuel Del Jesus, Andrea Rinaldo, and Ignacio Rodríguez-Iturbe. Point rainfall statistics for ecohydrological analyses derived from satellite integrated rainfall measurements. *Water Resources Research*, 51(4):2974–2985, 2015.

[Dou61] JP Dougherty. The anisotropy of turbulence at the meteor level. *Journal of Atmospheric and Terrestrial Physics*, 21(2-3):210–213, 1961.

[DPR+12] Anthony C Davison, Simone A Padoan, Mathieu Ribatet, et al. Statistical modeling of spatial extremes. *Statistical science*, 27(2):161–186, 2012.

[DS90] Anthony C Davison and Richard L Smith. Models for exceedances over high thresholds. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 393–442, 1990.

[Dub68] Satya D Dubey. A compound weibull distribution. *Naval Research Logistics Quarterly*, 15(2):179–188, 1968.

[Dup99] Debbie J Dupuis. Exceedances over high thresholds: A guide to threshold selection. *Extremes*, 1(3):251–261, 1999.

[DWB15] Yajuan Duan, Anna M Wilson, and Ana P Barros. Scoping a field experiment: error diagnostics of trmm precipitation radar estimates in complex terrain as a basis for iphex2014. *Hydrology and Earth System Sciences*, 19(3):1501–1520, 2015.

[DZH18] Levon Demirdjian, Yaping Zhou, and George J Huffman. Statistical modeling of extreme precipitation with trmm data. *Journal of Applied Meteorology and Climatology*, 57(1):15–30, 2018.

[EBH+15] B Eggert, Peter Berg, JO Haerter, D Jacob, and Christopher Moseley. Temporal and spatial scaling impacts on extreme precipitation. *Atmospheric Chemistry and Physics*, 15(10):5957–5971, 2015.

[EKM13] Paul Embrechts, Claudia Klüppelberg, and Thomas Mikosch. *Modelling extremal events: for insurance and finance*, volume 33. Springer Science & Business Media, 2013.

[ESC+93] RL Elliott, FR Schiebe, KC Crawford, KD Peter, and WE Puckett. A unique data capability for natural resources studies. In *International Winter Meeting of the American Society of Agricultural Engineers, Chicago, IL*, pages 14–17, 1993.

[ET09] E.F. Eastoe and J.A. Tawn. Modelling non-stationary extremes with application to surface level ozone. *J. Roy. Statist. Soc. Ser. C*, 58(1):25–45, 2009.

[ET10] Emma F Eastoe and Jonathan A Tawn. Statistical models for overdispersion in the frequency of peaks over threshold data for a flow series. *Water resources research*, 46(2), 2010.

[Fal09] Gregory Falkovich. Symmetries of the turbulent state. *Journal of Physics A: Mathematical and Theoretical*, 42(12):123001, 2009.

[FB73] A Shelby Frisch and Joost A Businger. A study of convective elements in the atmospheric surface layer. *Boundary-Layer Meteorology*, 3(3):301–328, 1973.

[FD96] Uriel Frisch and Russell J Donnelly. Turbulence: the legacy of an kolmogorov, 1996.

[FG18] Lee Fawcett and Amy C Green. Bayesian posterior predictive return levels for environmental extremes. *Stochastic Environmental Research and Risk Assessment*, 32(8):2233–2252, 2018.

[FGV01] G Falkovich, K Gawedzki, and Massimo Vergassola. Particles and fields in fluid turbulence. *Reviews of Modern Physics*, 73(4):913, 2001.

[FHR02] Arnoldo Frigessi, Ola Haug, and Håvard Rue. A dynamic mixture model for unsupervised tail estimation without threshold selection. *Extremes*, 5(3):219–235, 2002.

[FKF+17] Ulrich Foelsche, Gottfried Kirchengast, Juergen Fuchsberger, Jackson Tan, Walter A Petersen, et al. Evaluation of gpm imerg early, late, and final rainfall estimates using wegenernet gauge data in southeastern austria. *Hydrology and Earth System Sciences*, 21(12):6559–6572, 2017.

[FN08] Thomas Foken and Carmen J Napo. *Micrometeorology*, volume 2. Springer, 2008.

[Fok06] Thomas Foken. 50 years of the monin–obukhov similarity theory. *Boundary-Layer Meteorology*, 119(3):431–447, 2006.

[FP67] GE Fortescue and JRA Pearson. On gas absorption into a turbulent liquid. *Chemical Engineering Science*, 22(9):1163–1176, 1967.

[FS97] Uriel Frisch and Didier Sornette. Extreme deviations and applications. *Journal de Physique I*, 7(9):1155–1171, 1997.

[FT28] Ronald Aylmer Fisher and Leonard Henry Caleb Tippett. Limiting forms of the frequency distribution of the largest or smallest member of a sample. In *Mathematical Proceedings of the Cambridge Philosophical Society*, volume 24, pages 180–190. Cambridge University Press, 1928.

[FW96] Th Foken and Bodo Wichura. Tools for quality assessment of surface-based flux measurements. *Agricultural and forest meteorology*, 78(1-2):83–105, 1996.

[FYH+18] Mohammad Faridzad, Tiantian Yang, Kuolin Hsu, Soroosh Sorooshian, and Chan Xiao. Rainfall frequency analysis for ungauged regions using remotely sensed precipitation information. *Journal of hydrology*, 563:123–142, 2018.

[Gar94] John Roy Garratt. Review: the atmospheric boundary layer. *Earth-Science Reviews*, 37(1-2):89–134, 1994.

[Gar24] Crispin W Gardiner. *Handbook of Stochastic Methods*. Springer-Verlag, Berlin–Heidelberg–New York–Tokyo, 1985, p.124.

[GCS+13] Andrew Gelman, John B Carlin, Hal S Stern, David B Dunson, Aki Vehtari, and Donald B Rubin. *Bayesian data analysis*. Chapman and Hall/CRC, 2013.

[GD94] Alan E Gelfand and Dipak K Dey. Bayesian model choice: asymptotics and exact calculations. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 501–514, 1994.

[GDC92] Alan E Gelfand, Dipak K Dey, and Hong Chang. Model determination using predictive distributions with implementation via sampling-based methods. Technical report, STANFORD UNIV CA DEPT OF STATISTICS, 1992.

[GHS17] Tamer A Gado, Kuolin Hsu, and Soroosh Sorooshian. Rainfall frequency analysis for ungauged sites using satellite precipitation products. *Journal of hydrology*, 554:646–655, 2017.

[GJH02] Christoph S Garbe, Bernd Jähne, and Horst Haußecker. Measuring the sea surface heat flux and probability distribution of surface renewal events. *Gas Transfer at Water Surfaces*, pages 109–114, 2002.

[GK54] BV Gnedenko and AN Kolmogorov. *Independent Random Variables.* Cambridge, Massachusetts: Addison-Wesley, 1954.

[GK04a] Mekonnen Gebremichael and Witold F Krajewski. Assessment of the statistical characterization of small-scale rainfall variability from radar: Analysis of trmm ground validation datasets. *Journal of Applied Meteorology*, 43(8):1180–1199, 2004.

[GK04b] Toshiyuki Gotoh and Robert H Kraichnan. Turbulence and tsallis statistics. *Physica D: Nonlinear Phenomena*, 193(1):231–244, 2004.

[GK+16] Eric Gilleland, Richard W Katz, et al. extremes 2.0: an extreme value analysis package in r. *Journal of Statistical Software*, 72(8):1–39, 2016.

[GLM79] J. Greenwood, J. Landwehr, and N. Matalas. Probability weighted moments: Definition and relation to parameters of several distributions expressible in inverse form. *Water Resour. Res.*, 15(5):1049–1054, 1979.

[GLMW79] J Arthur Greenwood, J Maciunas Landwehr, Nicolas C Matalas, and James R Wallis. Probability weighted moments: definition and relation to parameters of several distributions expressable in inverse form. *Water Resources Research*, 15(5):1049–1054, 1979.

[GMFG+10] Vamsi Ganti, Mark M Meerschaert, Efi Foufoula-Georgiou, Enrica Viparelli, and Gary Parker. Normal and anomalous diffusion of gravel tracer particles in rivers. *Journal of Geophysical Research: Earth Surface*, 115(F2), 2010.

[Gne43] Boris Gnedenko. Sur la distribution limite du terme maximum d'une serie aleatoire. *Annals of mathematics*, pages 423–453, 1943.

[GSJ04] Christoph S Garbe, Uwe Schimpf, and Bernd Jähne. A surface renewal model to analyze infrared image sequences of the ocean surface for the study of air-sea heat and gas exchange. *Journal of Geophysical Research: Oceans*, 109(C8), 2004.

[GSU01] R. Gençay, F. Selçuk, and A. Ulugülyagci. Evim: A software package for extreme value analysis in matlab. *Stud. Nonlinear Dyn. E.*, 5(3), 2001.

[Gum58] E.J. Gumbel. *Statistics of Extremes*. Columbia Univ. Press, New York, 1958.

[GW93] Vijay K Gupta and Edward C Waymire. A statistical analysis of mesoscale rainfall as a random cascade. *Journal of Applied Meteorology*, 32(2):251–267, 1993.

[HABN10] George J Huffman, Robert F Adler, David T Bolvin, and Eric J Nelkin. The trmm multi-satellite precipitation analysis (tmpa). In *Satellite rainfall applications for surface hydrology*, pages 3–22. Springer, 2010.

[Han56] Thomas J Hanratty. Turbulent exchange of mass and momentum with a boundary. *AIChE Journal*, 2(3):359–362, 1956.

[Har62] Peter Harriott. A random eddy modification of the penetration theory. *Chemical Engineering Science*, 17(3):149–154, 1962.

[HB13] George J Huffman and David T Bolvin. Trmm and other data precipitation data set documentation. *NASA, Greenbelt, USA*, 28(2.3):1, 2013.

[HBB+14] George J Huffman, David T Bolvin, Dan Braithwaite, Kuolin Hsu, Robert Joyce, Pingping Xie, and Soo-Hyun Yoo. Nasa global precipitation measurement (gpm) integrated multi-satellite retrievals for gpm (imerg). *Algorithm Theoretical Basis Document (ATBD), NASA/GSFC, Greenbelt, MD, USA*, 2014.

[HBN+07] George J Huffman, David T Bolvin, Eric J Nelkin, David B Wolff, Robert F Adler, Guojun Gu, Yang Hong, Kenneth P Bowman, and Erich F Stocker. The trmm multisatellite precipitation analysis (tmpa): Quasi-global, multiyear, combined-sensor precipitation estimates at fine scales. *Journal of Hydrometeorology*, 8(1):38–55, 2007.

[HC14] R Ian Harris and Nicholas J Cook. The parent wind speed distribution: Why weibull? *Journal of wind engineering and industrial aerodynamics*, 131:72–87, 2014.

[HEM+15] Jan O Haerter, Bastian Eggert, Christopher Moseley, Claudio Piani, and Peter Berg. Statistical precipitation bias correction of gridded model data using point measurements. *Geophysical Research Letters*, 42(6):1919–1929, 2015.

[HFRS11] Mathias Herbst, Thomas Friborg, Rasmus Ringgaard, and Henrik Soegaard. Interpreting the variations in atmospheric methane fluxes observed above a restored wetland. *Agricultural and Forest Meteorology*, 151(7):841–853, 2011.

[HH08] Faisal Hossain and George J Huffman. Investigating error metrics for satellite rainfall data at hydrologically relevant scales. *Journal of Hydrometeorology*, 9(3):563–575, 2008.

[HHA09] Emad Habib, Amy Henschke, and Robert F Adler. Evaluation of tmpa satellite-based research and real-time rainfall estimates during six tropical-related heavy rainfall events over louisiana, usa. *Atmospheric Research*, 94(3):373–388, 2009.

[Hig35] Ralph Higbie. The rate of absorption of a pure gas into a still liquid during short periods of exposure. *Trans. AIChE*, 31:365–389, 1935.

[Hil89] Reginald J Hill. Implications of monin–obukhov similarity theory for scalar quantities. *Journal of the Atmospheric Sciences*, 46(14):2236–2244, 1989.

[HKC01] Emad Habib, Witold F Krajewski, and Grzegorz J Ciach. Estimation of rainfall interstation correlation. *Journal of Hydrometeorology*, 2(6):621–629, 2001.

[HMD+14] Janina Hommeltenberg, Matthias Mauder, Matthias Drösler, Katja Heidbach, Peter Werle, and Hans Peter Schmid. Ecosystem scale methane fluxes in a natural temperate bog-pine forest in southern germany. *Agricultural and forest meteorology*, 198:273–284, 2014.

[HNZ19] Whitney K Huang, Douglas W Nychka, and Hao Zhang. Estimating precipitation extremes using the log-histospline. *Environmetrics*, 30(4):e2543, 2019.

[HO13] E Haghighi and Dani Or. Evaporation from porous surfaces into turbulent airflows: Coupling eddy characteristics with pore scale vapor diffusion. *Water Resources Research*, 49(12):8432–8442, 2013.

[HO15a] Erfan Haghighi and Dani Or. Linking evaporative fluxes from bare soil across surface viscous sublayer with the monin–obukhov atmospheric flux-profile estimates. *Journal of Hydrology*, 525:684–693, 2015.

[HO15b] Erfan Haghighi and Dani Or. Thermal signatures of turbulent airflows interacting with evaporating thin porous surfaces. *International Journal of Heat and Mass Transfer*, 87:429–446, 2015.

[Hos90] Jonathan RM Hosking. 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.

[HPM03] Chad W Higgins, Marc B Parlange, and C Meneveau. Alignment trends of velocity gradients and subgrid-scale fluxes in the turbulent atmospheric boundary layer. *Boundary-Layer Meteorology*, 109(1):59–83, 2003.

[HSM20] SR Hosseini, M Scaioni, and M Marani. Extreme atlantic hurricane probability of occurrence through the metastatistical extreme value distribution. *Geophysical Research Letters*, 2020.

[HW87] Jonathan RM Hosking and James R Wallis. Parameter and quantile estimation for the generalized pareto distribution. *Technometrics*, 29(3):339–349, 1987.

[HW05] Jonathan Richard Morley Hosking and James R Wallis. *Regional frequency analysis: an approach based on L-moments*. Cambridge University Press, 2005.

[HWW85] J. R. M Hosking, J.R. Wallis, and E.F. Wood. Estimation of the generalized extreme-value distribution by the method of probability-weighted moments. *Technometrics*, 27(3):251–261, 1985.

[IHT+18] Hiroki Iwata, Ryuichi Hirata, Yoshiyuki Takahashi, Yuichi Miyabara, Masayuki Itoh, and Kotaro Iizuka. Partitioning eddy-covariance methane fluxes from a shallow lake into diffusive and ebullitive fluxes. *Boundary-layer meteorology*, 169(3):413–428, 2018.

[JRM+19] Maxwell B Joseph, Matthew W Rossi, Nathan P Mietkiewicz, Adam L Mahood, Megan E Cattau, Lise Ann St. Denis, R Chelsea Nagy, Virginia Iglesias, John T Abatzoglou, and Jennifer K Balch. Spatiotemporal prediction of wildfire size extremes with bayesian finite sample maxima. *Ecological Applications*, page e01898, 2019.

[JWHT13] Gareth James, Daniela Witten, Trevor Hastie, and Robert Tibshirani. *An introduction to statistical learning*, volume 112. Springer, 2013.

[KBH+16] Chris Kidd, Andreas Becker, George J Huffman, Catherine L Muller, Paul Joe, Gail Skofronick-Jackson, and Dalia B Kirschbaum. So, how much of the earth's surface is covered by rain gauges? *Bulletin of the American Meteorological Society*, (2016), 2016.

[KBP05] R.W. Katz, G.S. Brush, and M.B. Parlange. Statistics of extremes: Modeling ecological disturbances. *Ecology*, 86(5):1124–1134, 2005.

[KFG97] Praveen Kumar and Efi Foufoula-Georgiou. Wavelet analysis for geophysical applications. *Reviews of geophysics*, 35(4):385–412, 1997.

[KGH+15] Pierre-Emmanuel Kirstetter, Jonathan J Gourley, Yang Hong, Jian Zhang, Saber Moazamigoodarzi, Carrie Langston, and Ami Arthur. Probabilistic precipitation rate estimates with ground-based radar networks. *Water Resources Research*, 51(3):1422–1442, 2015.

[KHG+12] Pierre-Emmanuel Kirstetter, Y Hong, JJ Gourley, S Chen, Z Flamig, J Zhang, M Schwaller, W Petersen, and Eyal Amitai. Toward a framework for systematic error modeling of spaceborne precipitation radar with noaa/nssl ground radar–based national mosaic qpe. *Journal of Hydrometeorology*, 13(4):1285–1300, 2012.

[KHG+13] Pierre-Emmanuel Kirstetter, Y Hong, JJ Gourley, M Schwaller, W Petersen, and J Zhang. Comparison of trmm 2a25 products, version 6 and version 7, with noaa/nssl ground radar–based national mosaic qpe. *Journal of Hydrometeorology*, 14(2):661–669, 2013.

[KHS97] Gabriel Katul, Cheng-I Hsieh, and John Sigmon. Energy-inertial scale interactions for velocity and temperature in the unstable atmospheric surface layer. *Boundary-Layer Meteorology*, 82(1):49–80, 1997.

[KJP+19] Sara H Knox, Robert B Jackson, Benjamin Poulter, Gavin McNicol, Etienne Fluet-Chouinard, Zhen Zhang, Gustaf Hugelius, Philippe Bousquet, Josep G Canadell, Marielle Saunois, et al. Fluxnet-ch4 synthesis activity: Objectives, observations, and future directions. *Bulletin of the American Meteorological Society*, (2019), 2019.

[KKHH18] Pierre-Emmanuel Kirstetter, Negar Karbalaee, Kuolin Hsu, and Yang Hong. Probabilistic precipitation rate estimates with space-based infrared sensors. *Quarterly Journal of the Royal Meteorological Society*, 144:191–205, 2018.

[KKP11] Gabriel G Katul, Alexandra G Konings, and Amilcare Porporato. Mean velocity profile in a sheared and thermally stratified atmospheric boundary layer. *Physical Review Letters*, 107(26):268502, 2011.

[KL17] Gabriel Katul and Heping Liu. A kolmogorov-brutsaert structure function model for evaporation into a turbulent atmosphere. *Water Resources Research*, 53(5):3635–3644, 2017.

[KLCBZ13] Gabriel G Katul, Dan Li, Marcelo Chamecki, and Elie Bou-Zeid. Mean scalar concentration profile in a sheared and thermally stratified atmospheric surface layer. *Physical Review E*, 87(2):023004, 2013.

[KLM19] Gabriel Katul, Dan Li, and Costantino Manes. A primer on turbulence in hydrology and hydraulics: The power of dimensional analysis. *Wiley Interdisciplinary Reviews: Water*, page e1336, 2019.

[KMGV18] Gabriel Katul, Ivan Mammarella, Tiia Grönholm, and Timo Vesala. A structure function model recovers the many formulations for air-water gas transfer velocity. *Water Resources Research*, 54(9):5905–5920, 2018.

[KMP+15] Gabriel G Katul, Costantino Manes, Amilcare Porporato, Elie Bou-Zeid, and Marcelo Chamecki. Bottlenecks in turbulent kinetic energy spectra predicted from structure function inflections using the von kármán-howarth equation. *Physical Review E*, 92(3):033009, 2015.

[KNM90] Satoru Komori, Ryuichi Nagaosa, and Yasuhiro Murakami. Mass transfer into a turbulent liquid across the zero-shear gas-liquid interface. *AIChE journal*, 36(6):957–960, 1990.

[Kol41] Andrei N Kolmogorov. The local structure of turbulence in incompressible viscous fluid for very large reynolds numbers. In *Dokl. Akad. Nauk SSSR*, volume 30, pages 301–305. JSTOR, 1941.

[Kol62] Andrey Nikolaevich Kolmogorov. A refinement of previous hypotheses concerning the local structure of turbulence in a viscous incompressible fluid at high reynolds number. *Journal of Fluid Mechanics*, 13(01):82–85, 1962.

[Kou04] Demetris Koutsoyiannis. Statistics of extremes and estimation of extreme rainfall: I. theoretical investigation/statistiques de valeurs extrêmes et estimation de précipitations extrêmes: I. recherche théorique. *Hydrological sciences journal*, 49(4), 2004.

[KPC94] Gabriel G Katul, Marc B Parlange, and Chia R Chu. Intermittency, local isotropy, and non-gaussian statistics in atmospheric surface layer turbulence. *Physics of Fluids*, 6(7):2480–2492, 1994.

[KPCS06] Gabriel Katul, Amilcare Porporato, Daniela Cava, and Mario Siqueira. An analysis of intermittency, scaling, and surface renewal in atmospheric surface layer turbulence. *Physica D: Nonlinear Phenomena*, 215(2):117–126, 2006.

[KPG+18] Gabriel Katul, Olli Peltola, Tiia Grönholm, Samuli Launiainen, Ivan Mammarella, and Timo Vesala. Ejective and sweeping motions above a peatland and their role in relaxed-eddy-accumulation measurements and turbulent transport modelling. *Boundary-Layer Meteorology*, 169(2):163–184, 2018.

[KPH66] LB Koppel, RD Patel, and JT Holmes. Statistical models for surface renewal in heat and mass transfer: Part i. dependence of average transport coefficients on age distribution. *AIChE Journal*, 12(5):941–946, 1966.

[KPN02] R.W. Katz, M.B. Parlange, and P. Naveau. Statistics of extremes in hydrology. *Adv. Water Resour.*, 25:1287–1304, 2002.

[KPS92] VR Kuznetsov, AA Praskovsky, and VA Sabelnikov. Fine-scale turbulence structure of intermittent shear flows. *Journal of Fluid Mechanics*, 243:595–622, 1992.

[KPSBZ14] Gabriel G Katul, Amilcare Porporato, Stimit Shah, and Elie Bou-Zeid. Two phenomenological constants explain similarity laws in stably stratified turbulence. *Physical Review E*, 89(2):023007, 2014.

[Kra68] Robert H Kraichnan. Small-scale structure of a scalar field convected by turbulence. *The Physics of Fluids*, 11(5):945–953, 1968.

[Kra94] Robert H Kraichnan. Anomalous scaling of a randomly advected passive scalar. *Physical Review Letters*, 72(7):1016, 1994.

[KS09] A Kermani and L Shen. Surface age of surface renewal in turbulent interfacial transport. *Geophysical Research Letters*, 36(10), 2009.

[KTZZ09] A M G Klein Tank, F W Zwiers, and X Zhang. Guidelines on analysis of extremes in a changing climate in support of informed decisions for adaptation. *World Meteorological Organization*, 2009.

[Kum17] Manoj Kumar. Scikit-garden, 2017. [Online].

[KVA01] Gabriel Katul, Brani Vidakovic, and John Albertson. Estimating global and local scaling exponents in turbulent flows using discrete wavelet transformations. *Physics of Fluids*, 13(1):241–250, 2001.

[Lai09] DYF Lai. Methane dynamics in northern peatlands: a review. *Pedosphere*, 19(4):409–421, 2009.

[Law91] AJ Lawrance. Directionality and reversibility in time series. *International Statistical Review/Revue Internationale de Statistique*, pages 67–79, 1991.

[LC11] Harry F Lins and Timothy A Cohn. Stationarity: wanted dead or alive? 1. *JAWRA Journal of the American Water Resources Association*, 47(3):475–480, 2011.

[Lea83] M Ross Leadbetter. Extremes and local dependence in stationary sequences. *Probability Theory and Related Fields*, 65(2):291–306, 1983.

[LKBZ12] Dan Li, Gabriel G Katul, and Elie Bou-Zeid. Mean velocity and temperature profiles in a sheared diabatic turbulent boundary layer. *Physics of Fluids*, 24(10):105105, 2012.

[LKL18] Dan Li, Gabriel G Katul, and Heping Liu. Intrinsic constraints on asymmetric turbulent transport of scalars within the constant flux layer of the lower atmosphere. *Geophysical Research Letters*, 45(4):2022–2030, 2018.

[LKZ15] Dan Li, Gabriel G Katul, and Sergej S Zilitinkevich. Revisiting the turbulent prandtl number in an idealized atmospheric surface layer. *Journal of the Atmospheric Sciences*, 72(6):2394–2410, 2015.

[LL59] Lev Davidovich Landau and Evgeny Mikhailovich Lifshitz. Fluid mechanics. *flme*, 1959.

[LL13] Laifang Li and Wenhong Li. Southeastern united states summer rainfall framework and its implication for seasonal prediction. *Environmental Research Letters*, 8(4):044017, 2013.

[LLR12] Malcolm R Leadbetter, Georg Lindgren, and Holger Rootzén. *Extremes and related properties of random sequences and processes*. Springer Science & Business Media, 2012.

[LM05] Yi Li and Charles Meneveau. Origin of non-gaussian statistics in hydrodynamic turbulence. *Physical Review Letters*, 95(16):164502, 2005.

[LM09] J Lepore and L Mydlarski. Effect of the scalar injection mechanism on passive scalar structure functions in a turbulent flow. *Physical review letters*, 103(3):034501, 2009.

[LS70] John C Lamont and DS Scott. An eddy cell model of mass transfer into the surface of a turbulent liquid. *AIChE Journal*, 16(4):513–519, 1970.

[LS98] Jean Laherrere and Didier Sornette. Stretched exponential distributions in nature and economy: "fat tails" with characteristic scales. *The European Physical Journal B-Condensed Matter and Complex Systems*, 2(4):525–539, 1998.

[LSLC16] Andrea Libertino, Ashish Sharma, Venkat Lakshmi, and Pierluigi Claps. A global assessment of the timing of extreme rainfall from trmm and gpm for improving hydrologic design. *Environmental Research Letters*, 11(5):054003, 2016.

[LWB18] Xin Li, Xuan Wang, and Vladan Babovic. Analysis of variability and trends of precipitation extremes in singapore during 1980–2013. *International Journal of Climatology*, 38(1):125–141, 2018.

[MA14] Ali Mehran and Amir AghaKouchak. Capabilities of satellite precipitation datasets to estimate heavy precipitation rates at different temporal accumulations. *Hydrological Processes*, 28(4):2262–2270, 2014.

[Mal89] Stephane G Mallat. A theory for multiresolution signal decomposition: the wavelet representation. *IEEE transactions on pattern analysis and machine intelligence*, 11(7):674–693, 1989.

[Mar03] Marco Marani. On the correlation structure of continuous and discrete point rainfall. *Water Resources Research*, 39(5), 2003.

[Mar05] M. Marani. Non-power-law-scale properties of rainfall in space and time. *Water Resour. Res.*, 41:W08413, 2005.

[MBF+08] Paul CD Milly, Julio Betancourt, Malin Falkenmark, Robert M Hirsch, Zbigniew W Kundzewicz, Dennis P Lettenmaier, and Ronald J Stouffer. Stationarity is dead: Whither water management? *Science*, 319(5863):573–574, 2008.

[MDK+12] Matthew J Menne, Imke Durre, Bryant Korzeniewski, Shelley McNeal, Kristy Thomas, Xungang Yin, Steven Anthony, Ron Ray, Russell S Vose, Byron E Gleason, et al. Global historical climatology network-daily (ghcn-daily), version 3. *NOAA National Climatic Data Center*, 10, 2012.

[MDV+12] Matthew J Menne, Imke Durre, Russell S Vose, Byron E Gleason, and Tamara G Houston. An overview of the global historical climatology network-daily database. *Journal of Atmospheric and Oceanic Technology*, 29(7):897–910, 2012.

[MDYB+11] Tim R Moore, Allison De Young, Jill L Bubier, Elyn R Humphreys, Peter M Lafleur, and Nigel T Roulet. A multi-year record of methane flux at the mer bleue bog, southern canada. *Ecosystems*, 14(4):646, 2011.

[Mei06] Nicolai Meinshausen. Quantile regression forests. *Journal of Machine Learning Research*, 7(Jun):983–999, 2006.

[Men91] Charles Meneveau. Analysis of turbulence in the orthonormal wavelet representation. *Journal of Fluid Mechanics*, 232:469–520, 1991.

[MHF92] EE Musschenga, PJ Hamersma, and JMH Fortuin. Momentum, heat and mass transfer in turbulent pipe flow: the extended random surface renewal model. *Chemical engineering science*, 47(17-18):4373–4392, 1992.

[MI15] Marco Marani and Massimiliano Ignaccolo. A metastatistical approach to rainfall extremes. *Advances in Water Resources*, 79:121–126, 2015.

[MK14] Alberto Montanari and Demetris Koutsoyiannis. Modeling and mitigating natural hazards: Stationarity is immortal! *Water Resources Research*, 50(12):9748–9756, 2014.

[MKT+18] Elise Monsieurs, Dalia Bach Kirschbaum, Jackson Tan, Jean-Claude Maki Mateso, Liesbet Jacobs, Pierre-Denis Plisnier, Wim Thiery, Augusta Umutoni, Didace Musoni, Toussaint Mugaruka Bibentyo, et al. Evaluating tmpa rainfall over the sparsely gauged east african rift. *Journal of Hydrometeorology*, 19(9):1507–1528, 2018.

[MLC96] Charles Meneveau, Thomas S Lund, and William H Cabot. A lagrangian dynamic subgrid-scale model of turbulence. *Journal of Fluid Mechanics*, 319:353–385, 1996.

[MMP+17] Francesco Marra, Efrat Morin, Nadav Peleg, Yiwen Mei, and Emmanouil N Anagnostou. Intensity-duration-frequency curves from remote sensing rainfall estimates: comparing satellite and weather radar over the eastern mediterranean. *Hydrology and Earth System Sciences*, 21(5):2389–2404, 2017.

[MMV20] Arianna Miniussi, Marco Marani, and Gabriele Villarini. Metastatistical extreme value distribution applied to floods across the continental united states. *Advances in Water Resources*, 136:103498, 2020.

[MNA+19] Francesco Marra, Efthymios I Nikolopoulos, Emmanouil N Anagnostou, Andras Bardossy, and Efrat Morin. Precipitation frequency analysis from remotely sensed datasets: A focused review. *Journal of Hydrology*, 2019.

[MNAM18] Francesco Marra, Efthymios I Nikolopoulos, Emmanouil N Anagnostou, and Efrat Morin. Metastatistical extreme value analysis of hourly rainfall from short records: Estimation of high quantiles and impact of measurement errors. *Advances in Water Resources*, 117:27–39, 2018.

[MO54a] AS Monin and AM Obukhov. Dimensionless characteristics of turbulence in the surface layer. *Akad. Nauk SSSR, Geofiz. Inst., Tr*, 24:163–187, 1954.

[MO54b] AS Monin and AMF Obukhov. Basic laws of turbulent mixing in the surface layer of the atmosphere. *Contrib. Geophys. Inst. Acad. Sci. USSR*, 151(163):e187, 1954.

[MPBH+17] Bastian Manz, Sebastián Páez-Bimos, Natalia Horna, Wouter Buytaert, Boris Ochoa-Tocachi, Waldo Lavado-Casimiro, and Bram Willems.

Comparative ground validation of imerg and tmpa at variable spatiotemporal scales in the tropical andes. *Journal of Hydrometeorology*, 18(9):2469–2489, 2017.

[MPS+98] Laurent Mydlarski, Alain Pumir, Boris I Shraiman, Eric D Siggia, and Zellman Warhaft. Structures and multipoint correlators for turbulent advection: predictions and experiments. *Physical review letters*, 81(20):4373, 1998.

[MS00a] Eduardo S Martins and Jery R Stedinger. Generalized maximum-likelihood generalized extreme-value quantile estimators for hydrologic data. *Water Resources Research*, 36(3):737–744, 2000.

[MS00b] Merab Menabde and Murugesu Sivapalan. Modeling of rainfall time series and extremes using bounded random cascades and levy-stable distributions. *Water Resources Research*, 36(11):3293–3300, 2000.

[MS02] J.E. Morrison and J.A. Smith. Stochastic modeling of flood peaks using the generalized extreme value distribution. *Water Resour. Res.*, 38(12):1305, 2002.

[MSA+14] Viviana Maggioni, Mathew RP Sapiano, Robert F Adler, Yudong Tian, and George J Huffman. An error model for uncertainty quantification in high-time-resolution precipitation products. *Journal of Hydrometeorology*, 15(3):1274–1292, 2014.

[MSKF90] Charles Meneveau, KR Sreenivasan, P Kailasnath, and MS Fan. Joint multifractal measures: Theory and applications to turbulence. *Physical Review A*, 41(2):894, 1990.

[MT13] Marc F Müller and Sally E Thompson. Bias adjustment of satellite rainfall data through stochastic modeling: Methods development and application to nepal. *Advances in Water Resources*, 60:121–134, 2013.

[MV15] Iman Mallakpour and Gabriele Villarini. The changing nature of flooding across the central united states. *Nature Climate Change*, 5(3):250–254, 2015.

[MVM20] Arianna Miniussi, Gabriele Villarini, and Marco Marani. Analyses through the metastatistical extreme value distribution identify contributions of tropical cyclones to rainfall extremes in the eastern us. *Geophysical Research Letters*, 2020.

[MVN19] Cristian Martinez-Villalobos and J David Neelin. Why do precipitation intensities tend to follow gamma distributions? *Journal of the Atmospheric Sciences*, (2019), 2019.

[MW98] L Mydlarski and Z Warhaft. Passive scalar statistics in high-péclet-number grid turbulence. *Journal of Fluid Mechanics*, 358:135–175, 1998.

[MZ15] M. Marani and S. Zanetti. Long term oscillations in rainfall extremes in a 268-year daily time series. *Water Resour. Res.*, 51(1):639–647, 2015.

[MZAM19] Francesco Marra, Davide Zoccatelli, Moshe Armon, and Efrat Morin. A simplified mev formulation to model extremes emerging from multiple nonstationary underlying processes. *Advances in water resources*, 127:280–290, 2019.

[NB15] M Nogueira and AP Barros. Transient stochastic downscaling of quantitative precipitation estimates for hydrological applications. *Journal of Hydrology*, 529:1407–1421, 2015.

[NCE] NOAA NCEP. Hourly precipitation data (hpd) network, version 1.0 beta. NOAA National Centers for Environmental Information. https://data.noaa.gov/dataset/dataset/u-s-hourly-precipitation-data.

[NHRH16] Philippe Naveau, Raphael Huser, Pierre Ribereau, and Alexis Hannart. Modeling jointly low, moderate, and heavy rainfall intensities without a threshold selection. *Water Resources Research*, 52(4):2753–2769, 2016.

[NSM+03] Sharon E Nicholson, B Some, J McCollum, E Nelkin, D Klotter, Y Berte, BM Diallo, I Gaye, G Kpabeba, O Ndiaye, et al. Validation of trmm and other rainfall estimates with a high-density gauge dataset for west africa. part ii: Validation of trmm rainfall products. *Journal of Applied Meteorology*, 42(10):1355–1368, 2003.

[NSSB17] J David Neelin, Sandeep Sahany, Samuel N Stechmann, and Diana N Bernstein. Global warming precipitation accumulation increases above the current-climate cutoff scale. *Proceedings of the National Academy of Sciences*, 114(6):1258–1263, 2017.

[NW08] Victor Ntegeka and Patrick Willems. Trends and multidecadal oscillations in rainfall extremes, based on a more than 100-year time series of 10 min rainfall intensities at uccle, belgium. *Water Resources Research*, 44(7), 2008.

[Obu49] AM Obukhov. The local structure of atmospheric turbulence. In *Dokl. Akad. Nauk. SSSR*, volume 67, pages 643–646, 1949.

[OG94] Thomas M Over and Vijay 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, 1994.

[OG96] Thomas M Over and Vijay K Gupta. A space-time theory of mesoscale rainfall using random cascades. *Journal of Geophysical Research: Atmospheres*, 101(D21):26319–26331, 1996.

[ONSH18] Mohammed Ombadi, Phu Nguyen, Soroosh Sorooshian, and Kuo-lin Hsu. Developing intensity-duration-frequency (idf) curves from satellite-based precipitation: Methodology and evaluation. *Water Resources Research*, 54(10):7752–7766, 2018.

[Ozm65] RV Ozmidov. On the turbulent exchange in a stably stratified ocean. izv. acad. sci. ussr. *Atmos. Oceanic Phys.*, 1:861–871, 1965.

[P+75] III James Pickands et al. Statistical inference using extreme order statistics. *the Annals of Statistics*, 3(1):119–131, 1975.

[PAFG18] Simon Michael Papalexiou, Amir AghaKouchak, and Efi Foufoula-Georgiou. A diagnostic framework for understanding climatology of tails of hourly precipitation extremes in the united states. *Water Resources Research*, 54(9):6725–6738, 2018.

[PAMP00] Fernando Porté-Agel, Charles Meneveau, and Marc B Parlange. A scale-dependent dynamic model for large-eddy simulation: application to a neutral atmospheric boundary layer. *Journal of Fluid Mechanics*, 415:261–284, 2000.

[PBAM13] N Peleg, M Ben-Asher, and E Morin. Radar subpixel-scale rainfall variability and uncertainty: lessons learned from observations of a dense rain-gauge network. *Hydrology and Earth System Sciences*, 17(6):2195, 2013.

[PC93] SB Pope and Emily SC Ching. Stationary probability density functions: An exact result. *Physics of Fluids A: Fluid Dynamics*, 5(7):1529–1531, 1993.

[PDR93] ERLP Pattey, RL Desjardins, and P Rochette. Accuracy of the relaxed eddy-accumulation technique, evaluated using co 2 flux measurements. *Boundary-Layer Meteorology*, 66(4):341–355, 1993.

[Per61] DD Perlmutter. Surface-renewal models in mass transfer. *Chemical Engineering Science*, 16(3-4):287–296, 1961.

[PG14] Satya Prakash and RM Gairola. Validation of trmm-3b42 precipitation product over the tropical indian ocean using rain gauge data from the rama buoy array. *Theoretical and applied climatology*, 115(3-4):451–460, 2014.

[PJCM10] Lisbeth Pedersen, Niels Einar Jensen, Lasse Engbo Christensen, and Henrik Madsen. Quantification of the spatial variability of rainfall based on a dense network of rain gauges. *Atmospheric research*, 95(4):441–454, 2010.

[PK13] Simon Michael Papalexiou and Demetris Koutsoyiannis. Battle of extreme value distributions: A global survey on extreme daily rainfall. *Water Resources Research*, 49(1):187–201, 2013.

[PKC+11] A Porporato, PR Kramer, M Cassiani, E Daly, and J Mattingly. Local kinetic interpretation of entropy production through reversed diffusion. *Physical Review E*, 84(4):041142, 2011.

[PKL+17] Angeline G Pendergrass, Reto Knutti, Flavio Lehner, Clara Deser, and Benjamin M Sanderson. Precipitation variability increases in a warmer climate. *Scientific reports*, 7(1):1–9, 2017.

[PKM13] SM Papalexiou, D Koutsoyiannis, and C Makropoulos. How extreme is extreme? an assessment of daily rainfall distribution tails. *Hydrology and Earth System Sciences*, 17(2):851–862, 2013.

[PLW10] Ming Pan, Haibin Li, and Eric Wood. Assessing the skill of satellite-based precipitation estimates in hydrologic applications. *Water Resources Research*, 46(9), 2010.

[PM19] Simon Michael Papalexiou and Alberto Montanari. Global and regional increase of precipitation extremes under global warming. *Water Resources Research*, 55(6):4901–4914, 2019.

[PMF+18] Nadav Peleg, Francesco Marra, Simone Fatichi, Athanasios Paschalis, Peter Molnar, and Paolo Burlando. Spatial variability of extreme rainfall at radar subpixel scale. *Journal of Hydrology*, 556:922–933, 2018.

[PMH+13] Olli Peltola, Ivan Mammarella, Sami Haapanala, George Burba, and Timo Vesala. Field intercomparison of four methane gas analyzers suitable for eddy covariance flux measurements. *Biogeosciences*, 10(6):3749–3765, 2013.

[PMPA16] Satya Prakash, Ashis K Mitra, DS Pai, and Amir AghaKouchak. From trmm to gpm: How well can heavy rainfall be detected from space? *Advances in Water Resources*, 88:1–7, 2016.

[PN15] Olivier P Prat and Brian R Nelson. Evaluation of precipitation estimates over conus derived from satellite, radar, and rain gauge data sets at daily to annual scales (2002–2012). *Hydrology and Earth System Sciences*, 19(4):2037–2056, 2015.

[PRD07] A Porporato, JR Rigby, and E Daly. Irreversibility and fluctuation theorem in stationary time series. *Physical Review Letters*, 98(9):094101, 2007.

[PRLV17] O Peltola, M Raivonen, X Li, and T Vesala. Technical note: Comparison of methane ebullition modelling approaches used in terrestrial wetland models, biogeosciences discuss, 2017.

[PSSR14] V.F. Pisarenko, A. Sornette, D. Sornette, and M.V. Rodin. Characterization of the tail of the distribution of earthquake magnitudes by combining the gev and gpd descriptions of extreme value theory. *Pure Appl. Geophys.*, 171(8):1599–1624, 2014.

[PT94] Dean Prichard and James Theiler. Generating surrogate data for time series with several simultaneously measured variables. *Physical Review Letters*, 73(7):951, 1994.

[PT13] Ioannis Papastathopoulos and Jonathan A Tawn. Extended generalised pareto models for tail estimation. *Journal of Statistical Planning and Inference*, 143(1):131–143, 2013.

[PVF06] Amilcare Porporato, Giulia Vico, and Philip A Fay. Superstatistics of hydro-climatic fluctuations and interannual ecosystem productivity. *Geophysical Research Letters*, 33(15), 2006.

[RCBC96] G Ruiz-Chavarria, C Baudet, and S Ciliberto. Scaling laws and dissipation scale of a passive scalar in fully developed turbulence. *Physica D: Nonlinear Phenomena*, 99(2):369–380, 1996.

[Rey12] AM Reynolds. Gusts within plant canopies are extreme value processes. *Physica A: Statistical Mechanics and its Applications*, 391(21):5059–5063, 2012.

[RICI87] Ignacio Rodriguez-Iturbe, David Roxbee Cox, and Valerie Isham. Some models for rainfall based on stochastic point processes. *Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences*, 410(1839):269–288, 1987.

[RKL+20] Terhi Riutta, Aino Korrensalo, Anna M Laine, Jukka Laine, and Eeva-Stiina Tuittila. Interacting effects of vegetation components and water level on methane dynamics in a boreal fen. *Biogeosciences*, 17(3):727–740, 2020.

[RMP15] C Rorai, PD Mininni, and A Pouquet. Stably stratified turbulence in the presence of large-scale forcing. *Physical Review E*, 92(1):013003, 2015.

[RMPG13] William B Rossow, Ademe Mekonnen, Cindy Pearl, and Weber Goncalves. Tropical precipitation extremes. *Journal of Climate*, 26(4):1457–1466, 2013.

[RRN+01] Fernando M Ramos, Reinaldo R Rosa, Camilo Rodrigues Neto, Mauricio JA Bolzan, Leonardo D Abreu Sá, and Haroldo F Campos Velho. Non-extensive statistics and three-dimensional fully developed turbulence. *Physica A: Statistical Mechanics and its Applications*, 295(1):250–253, 2001.

[RRP+07] Janne Rinne, Terhi Riutta, Mari Pihlatie, Mika Aurela, Sami Haapanala, Juha-Pekka Tuovinen, Eeva-Stiina Tuittila, and Timo Vesala. Annual cycle of methane emission from a boreal fen measured by the eddy covariance technique. *Tellus B: Chemical and Physical Meteorology*, 59(3):449–457, 2007.

[S+89] Richard L Smith et al. Extreme value analysis of environmental time series: an application to trend detection in ground-level ozone. *Statistical Science*, 4(4):367–377, 1989.

[SA01] Todd M Scanlon and John D Albertson. Turbulent transport of carbon dioxide and water vapor within a vegetation canopy during unstable conditions: Identification of episodes using wavelet analysis. *Journal of Geophysical Research: Atmospheres*, 106(D7):7251–7262, 2001.

[SA09] MRP Sapiano and PA Arkin. An intercomparison and validation of high-resolution satellite precipitation estimates with 3-hourly gauge data. *Journal of Hydrometeorology*, 10(1):149–166, 2009.

[Sch10] Quirin Schiermeier. The real holes in climate science: like any other field, research on climate change has some fundamental gaps, although not the ones typically claimed by sceptics. quirin schiermeier takes a hard look at some of the biggest problem areas. *Nature*, 463(7279):284–288, 2010.

[SGF17] Carsten Schaller, Mathias Göckede, and Thomas Foken. Flux calculation of short turbulent events–comparison of three methods. *Atmospheric Measurement Techniques*, 10(3):869–880, 2017.

[SHL08] Fengge Su, Yang Hong, and Dennis P Lettenmaier. Evaluation of trmm multisatellite precipitation analysis (tmpa) and its utility in hydrologic prediction in the la plata basin. *Journal of Hydrometeorology*, 9(4):622–640, 2008.

[SK14] Francesco Serinaldi and Chris G Kilsby. Rainfall extremes: Toward reconciliation after the battle of distributions. *Water resources research*, 50(1):336–352, 2014.

[SK15] F. Serinaldi and C.G. Kilsby. Stationarity is undead: Uncertainty dominates the distribution of extremes. *Adv. Water Resour.*, 77:17–36, 2015.

[SKFG19] Carsten Schaller, Fanny Kittler, Thomas Foken, and Mathias Göckede. Characterisation of short-term extreme methane fluxes related to non-turbulent mixing above an arctic permafrost ecosystem. *Atmospheric Chemistry and Physics*, 19(6):4041–4059, 2019.

[SL87] Daniel Schertzer and Shaun Lovejoy. Physical modeling and analysis of rain and clouds by anisotropic scaling multiplicative processes. *Journal of Geophysical Research: Atmospheres*, 92(D8):9693–9714, 1987.

[SL88] Yang Gon Seo and Won Kook Lee. Single-eddy model for random surface renewal. *Chemical Engineering Science*, 43(6):1395–1402, 1988.

[SLH19] Harald Schellander, Alexander Lieb, and Tobias Hell. Error structure of metastatistical and generalized extreme value distributions for modeling extreme rainfall in austria. *Earth and Space Science*, 6(9):1616–1632, 2019.

[SMF+16] Derek Starkenburg, Stefan Metzger, Gilberto J Fochesatto, Joseph G Alfieri, Rudiger Gens, Anupma Prakash, and Jordi Cristóbal. Assessment of despiking methods for turbulence data in micrometeorology. *Journal of Atmospheric and Oceanic Technology*, 33(9):2001–2013, 2016.

[Smi84] Richard L Smith. Threshold methods for sample extremes. In *Statistical extremes and applications*, pages 621–638. Springer, 1984.

[Smi92] Richard L Smith. The extremal index for a markov chain. *Journal of applied probability*, 29(1):37–45, 1992.

[SN14] Samuel N Stechmann and J David Neelin. First-passage-time prototypes for precipitation statistics. *Journal of the Atmospheric Sciences*, 71(9):3269–3291, 2014.

[Sor06] Didier Sornette. *Critical phenomena in natural sciences: chaos, fractals, selforganization and disorder: concepts and tools*. Springer Science & Business Media, 2006.

[SP97] Rainer Storn and Kenneth Price. Differential evolution–a simple and efficient heuristic for global optimization over continuous spaces. *Journal of global optimization*, 11(4):341–359, 1997.

[SPA06] Rob Stoll and Fernando Porté-Agel. Dynamic subgrid-scale models for momentum and scalar fluxes in large-eddy simulations of neutrally stratified atmospheric boundary layers over heterogeneous terrain. *Water Resources Research*, 42(1), 2006.

[Sre91] KR Sreenivasan. On local isotropy of passive scalars in turbulent shear flows. In *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, volume 434, pages 165–182. The Royal Society, 1991.

[Sre99] Katepalli R Sreenivasan. Fluid turbulence. *Reviews of Modern Physics*, 71(2):S383, 1999.

[SS00] Boris I Shraiman and Eric D Siggia. Scalar turbulence. *Nature*, 405(6787):639–646, 2000.

[SSK+14] Jörg Schumacher, Janet D Scheel, Dmitry Krasnov, Diego A Donzis, Victor Yakhot, and Katepalli R Sreenivasan. Small-scale universality in fluid turbulence. *Proceedings of the National Academy of Sciences*, 111(30):10961–10965, 2014.

[ST04] Alec Stephenson and Jonathan Tawn. Bayesian inference for extremes: accounting for the three extremal types. *Extremes*, 7(4):291–307, 2004.

[Ste93] Jery R Stedinger. Frequency analysis of extreme events. *Handbook of hydrology*, 18, 1993.

[Stu12] Roland B Stull. *An introduction to boundary layer meteorology*, volume 13. Springer Science & Business Media, 2012.

[SVKA05] Bin Shi, Brani Vidakovic, Gabriel G Katul, and John D Albertson. Assessing the effects of atmospheric stability on the fine structure of surface layer turbulence using local and global multiscale approaches. *Physics of Fluids*, 17(5):055104, 2005.

[SXWX10] Yan Shen, Anyuan Xiong, Ying Wang, and Pingping Xie. Performance of high-resolution satellite precipitation products over china. *Journal of Geophysical Research: Atmospheres*, 115(D2), 2010.

[SY89] Ya G Sinai and Victor Yakhot. Limiting probability distributions of a passive scalar in a random velocity field. *Physical Review Letters*, 63(18):1962, 1989.

[TAO06] Andrea Tancredi, Clive Anderson, and Anthony O'Hagan. Accounting for threshold uncertainty in extreme value estimation. *Extremes*, 9(2):87, 2006.

[Tay38] Geoffrey Ingram Taylor. The spectrum of turbulence. *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences*, pages 476–490, 1938.

[TB10] Kun Tao and Ana P Barros. Using fractal downscaling of satellite precipitation products for hydrometeorological applications. *Journal of Atmospheric and Oceanic Technology*, 27(3):409–427, 2010.

[TF07] Christoph Thomas and Thomas Foken. Organised motion in a tall spruce canopy: temporal scales, structure spacing and terrain effects. *Boundary-Layer Meteorology*, 122(1):123–147, 2007.

[THA+13] Yudong Tian, George J Huffman, Robert F Adler, Ling Tang, Mathew Sapiano, Viviana Maggioni, and Huan Wu. Modeling errors in daily precipitation measurements: Additive or multiplicative? *Geophysical Research Letters*, 40(10):2060–2065, 2013.

[THB76] TG Theofanous, RN Houze, and LK Brumfield. Turbulent mass transfer at free, gas-liquid interfaces, with applications to open-channel, bubble and jet flows. *International Journal of Heat and Mass Transfer*, 19(6):613–624, 1976.

[TKKK16] Naohisa Takagaki, Ryoichi Kurose, Atsushi Kimura, and Satoru Komori. Effect of schmidt number on mass transfer across a sheared gas-liquid interface in a wind-driven turbulence. *Scientific reports*, 6:37059, 2016.

[TLSM95] Constantino Tsallis, Silvio VF Levy, André MC Souza, and Roger Maynard. Statistical-mechanical foundation of the ubiquity of lévy distributions in nature. *Physical Review Letters*, 75(20):3589, 1995.

[TPL19] Hristos Tyralis, Georgia Papacharalampous, and Andreas Langousis. A brief review of random forests for water scientists and practitioners and their recent history in water resources. *Water*, 11(5):910, 2019.

[TS72] GM Thomson and RS Silver. Reynolds flux and danckwerts surface renewal theory. *International Journal of Heat and Mass Transfer*, 15(6):1284–1287, 1972.

[Tsa88] Constantino Tsallis. Possible generalization of boltzmann-gibbs statistics. *Journal of Statistical Physics*, 52(1):479–487, 1988.

[Van10] Erik Vanmarcke. *Random fields: analysis and synthesis*. World Scientific Publishing Co Inc, 2010.

[VGG17] Aki Vehtari, Andrew Gelman, and Jonah Gabry. Practical bayesian model evaluation using leave-one-out cross-validation and waic. *Statistics and computing*, 27(5):1413–1432, 2017.

[Vil10] Gabriele Villarini. Evaluation of the research-version tmpa rainfall estimate at its finest spatial and temporal scales over the rome metropolitan area. *Journal of Applied Meteorology and Climatology*, 49(12):2591–2602, 2010.

[VK07] Gabriele Villarini and Witold F Krajewski. Evaluation of the research version tmpa three-hourly 0.25× 0.25 rainfall estimates over oklahoma. *Geophysical Research Letters*, 34(5), 2007.

[VKS09] Gabriele Villarini, Witold F Krajewski, and James A Smith. New paradigm for statistical validation of satellite precipitation estimates: Application to a large sample of the tmpa 0.25° 3-hourly estimates over oklahoma. *Journal of Geophysical Research: Atmospheres*, 114(D12), 2009.

[VLL15] D. Veneziano, A. Langousis, and C. Lepore. New asymptotic and preasymptotic results on rainfall maxima from multifractal theory. *Water Resour. Res.*, 45(11):W11421, 2015.

[VM36] Richard Von Mises. La distribution de la plus grande de n valeurs. *Rev. math. Union interbalcanique*, 1(1), 1936.

[VMKM08] Gabriele Villarini, Pradeep V Mandapaka, Witold F Krajewski, and Robert J Moore. Rainfall and sampling uncertainties: A rain gauge perspective. *Journal of Geophysical Research: Atmospheres*, 113(D11), 2008.

[VS10] Gabriele Villarini and James A Smith. Flood peak distributions for the eastern united states. *Water Resources Research*, 46(6), 2010.

[War00] Z Warhaft. Passive scalars in turbulent flows. *Annual Review of Fluid Mechanics*, 32(1):203–240, 2000.

[Wei00] Stefan Weinzierl. Introduction to monte carlo methods. *arXiv:hep-ph/0006269v1*, 2000.

[WFE+14] S Westra, HJ Fowler, JP Evans, LV Alexander, Peter Berg, F Johnson, EJ Kendon, G Lenderink, and NM Roberts. Future changes to the intensity and frequency of short-duration extreme rainfall. *Reviews of Geophysics*, 52(3):522–555, 2014.

[WT05] PS Wilson and R Toumi. A fundamental probability distribution for heavy rainfall. *Geophysical Research Letters*, 32(14), 2005.

[WT12] JL Wadsworth and JA Tawn. Likelihood-based procedures for threshold diagnostics and uncertainty in extreme value modelling. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 74(3):543–567, 2012.

[WW10] Jianxin Wang and David B Wolff. Evaluation of trmm ground-validation radar-rain errors using rain gauge measurements. *Journal of Applied Meteorology and Climatology*, 49(2):310–324, 2010.

[Wyn10] John C Wyngaard. *Turbulence in the Atmosphere*. Cambridge University Press, 2010.

[XPF+14] Haitao Xu, Alain Pumir, Gregory Falkovich, Eberhard Bodenschatz, Michael Shats, Hua Xia, Nicolas Francois, and Guido Boffetta. Flight–crash events in turbulence. *Proceedings of the National Academy of Sciences*, 111(21):7558–7563, 2014.

[YZS15] PK Yeung, XM Zhai, and Katepalli R Sreenivasan. Extreme events in computational turbulence. *Proceedings of the National Academy of Sciences*, 112(41):12633–12638, 2015.

[ZBK18] Enrico Zorzetto, Andrew D Bragg, and Gabriel G Katul. Extremes, intermittency, and time directionality of atmospheric turbulence at the crossover from production to inertial scales. *Physical Review Fluids*, 3(9):094604, 2018.

[ZBM16] Enrico Zorzetto, Gianluca Botter, and Marco Marani. On the emergence of rainfall extremes from ordinary events. *Geophysical Research Letters*, 43(15):8076–8082, 2016.

[ZCL+06] Edward J Zipser, Daniel J Cecil, Chuntao Liu, Stephen W Nesbitt, and David P Yorty. Where are the most intense thunderstorms on earth? *Bulletin of the American Meteorological Society*, 87(8):1057–1072, 2006.

[ZCM20] Enrico Zorzetto, Antonio Canale, and Marco Marani. Bayesian hierarchical models for environmental time series. *submitted to Bayesian Analysis*, 2020.

[ZK14] Andreas Ziegler and Inke R König. Mining data with random forests: current options for real-world applications. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 4(1):55–63, 2014.

[ZLH15] Yaping Zhou, William KM Lau, and George J Huffman. Mapping trmm tmpa into average recurrence interval for monitoring extreme precipitation events. *Journal of Applied Meteorology and Climatology*, 54(5):979–995, 2015.

[ZM19] Enrico Zorzetto and Marco Marani. Downscaling of rainfall extremes from satellite observations. *Water Resources Research*, 55(1):156–174, 2019.

[ZM20] Enrico Zorzetto and Marco Marani. Extreme value metastatistical analysis of remotely sensed rainfall in ungauged areas: Spatial downscaling and error modelling. *Advances in Water Resources*, 135:103483, 2020.

[ZPGK20] Enrico Zorzetto, Olli Peltola, Tiia Grönholm, and Gabriel G. Katul. Intermittent surface renewals and methane hotspots over natural wetlands. *manuscript in preparation*, 2020.

Biography

Enrico Zorzetto obtained a B.Sc. and M.Sc. in Civil and Hydraulic Engineering at the University of Padova (Padova, Italy, 2012, 2015). He then moved to Duke University to pursue a Ph.D. in Earth and Ocean sciences (EOS) (Durham, NC, 2015 - 2020). At Duke, he was first supported by Teaching Assistantship from the Division of Earth and Ocean Sciences (2015 - 2017) and was later supported by the NASA NESSF program (2017 - 2020).