Università degli Studi di Padova

Dipartimento di Ingegneria Civile, Edile e Ambientale

Master Thesis in Water and Geological Risk Engineering

Probability Model of Rainfall Extremes Under a Changing Climate

Supervisor
Marco Marani

Master Candidate
Santatriniaina O.F. Andriamanantena

Co-supervisor
Marco Borga

Student ID
2042283

Academic Year
2022-2023

To Mom and Dad.

Abstract

Expected increases in future extreme precipitation due to anthropogenic climate change highlight the need for robust projections. However, the main drivers of this increase and its precise rate remain uncertain. Numerous studies have sought empirical relationships to identify how precipitation percentiles scale with temperature, but the exact changes in the probability distribution of extremes under a warming climate remain poorly understood. To address this gap, this thesis proposes to reformulate the Metastatistical Extreme Value Distribution (MEVD), which accounts for the full distribution of underlying "ordinary" events, to explicitly incorporate their temperature dependence. Leveraging long historical records of precipitation and temperature, their probabilistic relationship and the influence of thermodynamics and large-scale circulation on their relationship are investigated. The results indicate that the change in daily precipitation extremes cannot be solely attributed to thermodynamics; rather, the dynamical effects must also be considered. Conversely, at the hourly scale, evidence emerges that changes in precipitation extremes, predominantly driven by thermodynamics, manifest as an exponential relationship between the parameters of the distribution of the "ordinary" events and temperature, leading to the intensification and the increase in the frequency of hourly precipitation extremes under warming conditions.

Sommario

Gli aumenti previsti delle future precipitazioni estreme a causa del cambiamento climatico antropogenico evidenziano la necessità di proiezioni robuste. Tuttavia, i principali drivers di questo aumento e il suo preciso tasso di incremento rimangono incerti. Numerosi studi hanno cercato relazioni empiriche per identificare come i percentili delle distribuzioni delle precipitazioni variano con la temperatura, ma gli esatti cambiamenti nella distribuzione della probabilità degli eventi estremi in condizioni di riscaldamento climatico rimangono poco compresi. Per colmare questa lacuna, questa tesi propone di riformulare la Distribuzione Metastatistica dei Valori Estremi (MEVD), che tiene conto dell'intera distribuzione degli eventi "ordinari" sottostanti, per incorporare esplicitamente la loro dipendenza dalla temperatura. Sfruttando lunghe serie storiche di precipitazione e temperatura, sono studiate la loro relazione probabilistica e l'influenza della termodinamica e della circolazione atmosferica a grande scala sulla loro relazione. I risultati indicano che il cambiamento nelle precipitazione giornaliere estreme non può essere attribuito esclusivamente alla termodinamica ma anche gli effetti dinamici devono essere considerati. Al contrario, alla scala oraria, emergono evidenze che i cambiamenti nelle precipitazioni estreme, prevalentemente guidati dalla termodinamica, si manifestano come una relazione esponenziale tra i parametri della distribuzione degli eventi "ordinari" e la temperatura, portando all'intensificazione e all'aumento della frequenza di eventi di precipitazioni orarie estreme in condizioni di riscaldamento.

Contents

List of figures

List of tables

1

Introduction

Understanding the evolving behavior of extreme rainfall is of critical significance. The Earth's climate system is undergoing substantial changes due to ongoing global warming driven by the increase in greenhouse gas concentrations. These changes are expected to deeply affect various environmental aspects. Specifically, they may result in modified patterns of extreme rainfall in terms of frequency, intensity, and spatial distribution. These alterations could increase the occurrence of hazards associated with such extremes, putting more people and assets at risk on a global scale. Flash floods, landslides, and damage to the agricultural system are among the hazards driven by rainfall extremes with far-reaching socio-economic impacts. Therefore, a thorough understanding of extreme rainfall dynamics under a global warming context becomes imperative. Such understanding is key for effective water resource man-

agement, robust infrastructure planning, accurate risk assessment, and the development of effective climate adaptation and mitigation strategies. Robust projections of these changes provide engineers and planners with indispensable insights enabling the implementation of measures that enhance infrastructure resilience against the growing threats of climate change. Furthermore, this knowledge serves as a foundation for informed risk management and policy decisions aimed at mitigating hydrogeological vulnerabilities posed by these changing extremes.

Several studies provide increasing evidence that a warmer climate will lead to more intense [1, 2, 3] and frequent [1, 4, 5] extreme rainfall events. However, the precise rate of increase and the roles of thermodynamic and atmospheric circulation processes remain uncertain. Much of the research addressing the empirical scaling of extreme precipitation with temperature is based on high precipitation percentiles. The relationship between extreme rainfall and temperature varies across studies depending on the region, with differing findings ranging from monotonic increasing trends at different scaling rates, and downward trends, to hook structures characterized by increasing trends followed by subsequent declines at higher temperatures [6, 7, 8, 9, 10]. However, the applicability of these observed scaling rates to explain global warming [11] or in the least, their suitability for characterizing future extreme rainfall behavior in the region from which they were derived is still subject of discussion.

In this thesis, I propose to look at the probabilistic relation between precipitation and temperature using long time series to investigate the role of thermodynamics in driving the generation of daily and hourly precipitation extremes. Additionally, I aim to reformulate an existing metastatistical extreme value model to investigate whether the change in the probability distribution of rainfall extremes under global warming can be accurately described by a parsimonious probability model.

2

2

Data and Methods

2.1 Extreme Precipitation

The term "extreme" lacks a universally agreed-upon definition. Readers are encouraged to refer to Gimeno et al. [12], which extensively addresses this issue by exploring disaster-related and statistical perspectives. In this thesis, our focus will center on examining extreme precipitation from a statistical parametric standpoint. This involves the fitting of extreme value distributions to estimate precipitation quantiles for given return periods, often much greater than the observation interval. This topic will be thoroughly discussed in section 2.2.

2.1.1 PRECIPITATION GENERATION MECHANISM

The space-time characteristics of precipitation events, and, hence, their probabilistic structure, depend on the precipitation generation mechanisms. We can distinguish three main types of precipitation generation mechanisms, namely stratiform, convective, and orographic.

At a larger scale, stratiform precipitation emerges as a consequence of broad, sustained processes. This form of precipitation occurs as widespread steady and long-duration events typically associated with synoptic-scale weather systems. This phenomenon occurs in the presence of relatively mild upward atmospheric motion [13].

In contrast, convective precipitation occurs on a more localized scale, resulting in shorter yet more intense episodes. This form of precipitation is closely linked with atmospheric instability and convection processes. Convective precipitation requires specific dynamic conditions, including moisture, atmospheric instability, and lift to trigger its occurrence. Notably, the dynamics driving convective precipitation are characterized by high non-linearity, making them remarkably sensitive to perturbations [8, 14, 5]. In Germany, observations show that convective precipitation exhibits a faster increase with temperature than stratiform precipitation [6]. Extensive datasets are required for a precise classification of precipitation events into these two distinct types. Radar data are typically used in that regard by looking at the drop size or reflectivity gradients. [15]. Other approaches involve the use of cloud type classification [6], or the occurrence of lightning or thunder as proxies to differentiate between convective and stratiform precipitation events [16].

In addition to the previously mentioned mechanisms, we can identify orographic precipitation. This type occurs when moist air is lifted to saturation due to the orographic effect, enhancing precipitation and triggering moist instabilities [17]. Therefore, it is strongly connected to the local topography and is typical in regions near mountain ranges. The response

4

of orographic precipitation to climate change depends on several factors. Notably, large-scale changes in wind speed and direction may alter atmospheric moisture transport, consequently impacting the spatial and temporal distribution of precipitation [18]. In addition, shifts from snow to rain can lead to alterations in the overall precipitation pattern [19].

2.1.2 Theoretical Basis of the Scaling of Extreme Precipitation

Let us consider a large-scale precipitation event. The conservation of moisture for an infinitesimal moist air parcel can be expressed as:

$$\frac{\partial q}{\partial t} + \nabla \cdot q\mathbf{v} = e - c \tag{2.1}$$

Here $\mathbf{v}$ represents the velocity field, q the specific humidity, e the rate of re-evaporation of clouds and rain-water, and c the rate of condensation per unit mass.

Let us now consider a saturated air parcel. The Lagrangian description of the net condensation rate needed to maintain the water vapor at saturation can be expressed as:

$$c = -\frac{Dq_s}{Dt} \tag{2.2}$$

where $q_s = q_s(p, T)$ is the saturation vapor mixing ratio and $\frac{D(.)}{Dt}$ is the material derivative operator. The condensation rate required to keep the rising air near saturation is given by [20]:

$$c = -\omega \frac{dq_s}{dp}\bigg|_{\theta_e^*} \tag{2.3}$$

where ω is the vertical velocity in the pressure coordinate. O'Gorman and Schneider [21] has shown that $\frac{\partial q_s}{\partial p}\big|_{\theta_e^*}$ and with it the condensation and the precipitation rate does not increase as rapidly with temperature as q_s.

Taking into account the hydrostatic equation given by:

$$\frac{\partial p}{\partial z} = \rho g \tag{2.4}$$

where g is the gravitational acceleration, the precipitation rate which can be assumed proportional to the total rate of condensation within an air column, can be written as:

$$P_e \propto - \int_{z_S}^{z_T} c\rho dz = \int_{p_S}^{p_T} c \frac{dp}{g} \tag{2.5}$$

where the subscripts S and T refer to the surface and the top of the troposphere respectively. We can therefore write:

$$P_e = \epsilon \int_{p_S}^{p_T} -\omega \frac{\partial q_s}{\partial p}\bigg|_{\theta_e^*} \frac{dp}{g} \tag{2.6}$$

where the constant of proportionality ϵ is the precipitation efficiency which accounts for the microphysical processes that convert the condensates into precipitation and potentially includes all approximations used to derive Equation 2.6.

If we denote the mass-weighted integral over the troposphere as $\langle . \rangle = \int_{p_S}^{p_T} (.) \frac{dp}{g}$ and neglect the changes in the precipitation efficiency, the fractional changes in P_e is given by [22]:

$$\frac{\delta P_e}{P_e} \approx \frac{\delta \langle \omega \frac{\partial q_s}{\partial p}\big|_{\theta_e^*} \rangle}{\langle \omega \frac{\partial q_s}{\partial p}\big|_{\theta_e^*} \rangle} \tag{2.7}$$

Neglecting the higher order terms, Equation 2.8 can be decomposed as:

$$\frac{\delta P_e}{P_e} \approx \frac{\langle \omega \delta(\frac{\partial q_s}{\partial p}\big|_{\theta_e^*}) \rangle}{\langle \omega \frac{\partial q_s}{\partial p}\big|_{\theta_e^*} \rangle} + \frac{\langle \delta(\omega) \frac{\partial q_s}{\partial p}\big|_{\theta_e^*} \rangle}{\langle \omega \frac{\partial q_s}{\partial p}\big|_{\theta_e^*} \rangle} \tag{2.8}$$

6

where the terms on the right-hand side refer to the thermodynamic and the dynamic contribution to the change in extreme precipitation respectively. This scaling relation is assessed using models [23, 21, 20, 24]. It has been shown that considering only the thermodynamic scaling for precipitation extremes is usually not enough to explain observed changes in precipitation extremes [21] since the dynamical contribution is a significant source of uncertainty.[24].

2.1.3   EMPIRICAL PRECIPITATION SCALING

The precipitation scaling in equation 2.6 is often assessed using a model where vertical velocity and temperature are explicitly computed. Due to the coarse resolution of the models typically used in the literature, all the relevant quantities are computed as average values over large areas. In contrast, observations are obtained from gauge stations which provide point measurements where local effects are dominant. In addition, taking into account factors such as initial-condition dependence and model uncertainty, one can expect discrepancies between scaling derived from models and observed scaling. Observed scaling relationships are derived by empirically relating observed extremes with temperature without the need for theoretical models such as the one described in Section 2.1.2. In this respect, a commonly used approach involves binning followed by an exponential regression[7, 25]. The precipitation depth during wet events (defined as days where P exceeds a certain threshold) is paired with same-day temperature and placed into a certain number of bins according to temperature. Typically, the bins are allowed to a variable width to allow an equal number of observations in each bin. The median temperature of each bin is used as its representative temperature.

The relevant statistics describing both the precipitation extremes and the underlying distribution (e.g., empirical percentiles, Weibull distribution parameters) are computed for each

bin. Assuming that all factors affecting precipitation remain equal, including relative humidity and atmospheric circulation, extreme precipitation which significantly exceeds the rate of evaporation, is primarily determined by the amount of water vapor present in the atmosphere. Consequently, it is expected to increase with temperature at the same rate as atmospheric-moisture content[26]. The change in the moisture-holding capacity of the atmosphere is governed by the Clausius-Clapeyron equation:

$$\frac{de_s}{e_s} = \frac{L_v dT}{R_v T^2} \tag{2.9}$$

where e_s is the saturation vapor pressure of water, L_v is the latent heat of vaporization, and R is the gas constant for water vapor. This rate of increase, known as the CC-scaling, is approximately $7\%/K$ and varies as a function of temperature [27].

Following the reasoning presented by Trenberth et al. [26], the relation between atmospheric moisture content and thereby extreme precipitation, with temperature can be described by an exponential relationship. Therefore, the following regression equation is defined to quantify the response of precipitation extremes to global warming, by relating the statistical properties of extreme precipitation with temperature:

$$S(T + \Delta T) = S(T)(1 + \alpha_S)^{\Delta T} \tag{2.10}$$

where $\alpha_S * 100\%$ represents the scaling rate.

2.2 Metastatistical Extreme Value Framework

The Metastatistical Extreme Value Distribution (MEVD) [28] is a non-asymptotic extreme value distribution based on the concept that extreme values emerge from the distribution of

underlying ordinary events. It therefore makes use of the full available data rather than of a small sub-sample of its tail, like in the case of the Generalized Extreme value distribution. A general expression of the MEVD can be derived as follows: Let us consider the cumulative distributions of the ordinary rainfall depth, denoted as $F_X(x) = P(X \leq x)$. Assuming independent events generating n realizations of rainfall in a year, the cumulative distribution of the yearly maximum rainfall, $Y_n = max\{X_1, \ldots, X_n\}$, is given by:

$$H_n(y) = P(Y \leq y) = P(X_i \leq y, i = 1, \ldots, n) = [F_X(y)]^n \qquad (2.11)$$

The MEVD relaxes the traditional assumption of $n \to \infty$ and considers it as a random variable. By use of the total probability theorem, we can obtain an expression of MEVD by considering all possible joint realization of the number of events n and of temperature T as follow:

$$\zeta(y) = P(Y \leq y) = \sum_n \int_T H_n(y|n,t) f_{N,T}(n,t) dT \qquad (2.12)$$

$f_{N,T}(n,t)$ denotes the joint probability distribution of the number of events per year, N, and the temperature, T. Let $f_N(n) = \sum_{n_j \in R_N} P_N(n_j) \delta(n - n_j)$ and $f_T(t)$ be the probability density function of N and T. Assuming that N and T are independent, we can write:

$$f(n,t) = f_N(n) \cdot f_T(t) = f_T(t) \sum_{n_j \in R_N} P_N(n_j) \delta(n - n_j) \qquad (2.13)$$

Equation 2.12 can be rewritten as:

$$
\begin{aligned}
\zeta(y) &= \sum_n \int_T H_n(y|n,t) f_T(t) (\sum_{n_j \in R_N} P_N(n_j) \delta(n - n_j)) dt \\
&= \int_T f_T(t) \sum_{n_j \in R_N} P_N(n_j) \sum_n H_n(y|n,t) \delta(n - n_j) dt
\end{aligned}
\qquad (2.14)
$$

If N assumes the values $R_N = \{n_1, n_2, \ldots, n_M\}$ with equal probability over the analysis period, we can write

$$\begin{aligned}
\zeta_{MEVP(T)}(y) &\approx \frac{1}{M} \sum_{j=1}^{M} \int_T H_{n_j}(y|n_j, t) f_T(t) dt \\
&\approx \frac{1}{M} \sum_{j=1}^{M} \int_T F_{X|T_j}(y|t)^{n_j} f_T(t) dt
\end{aligned} \tag{2.15}$$

where according to the total probability theorem, $\int_T H_{n_j}(y|n_j, t) f_{T(t)dt} = H_{n_j}(y)$ is a single sample of yearly maxima distribution with $j = 1, 2, \ldots, M$ representing the observational years, n_j is the number of events in the j$^{\text{th}}$ year, $F_{X|T}(x|t)$ the conditional probability distribution of ordinary precipitation given daily temperature.

In subsequent sections, I will refer to this model as MEV P(T) to emphasize its consideration of the dependence of precipitation on temperature, and to distinguish it from the standard MEV.

2.3 Quantile estimation

Quantile estimations of complex distributions, such as the reformulated MEVD in equation 2.15 pose computational challenges. These are due to the need to evaluate integrals over the entire range of possible realizations of temperature T. As a result, the computation of exact quantiles, which necessitates the determination of inverse CDFs, becomes unfeasible. To address these complexities and provide a practical solution, sampling-based approximations which offer an effective way to estimate quantiles without engaging directly with intricate integral computations are considered.

I assume that we know or have a statistical model of the probability distribution of temperature T and the number of events N, and the conditional probability distribution of

precipitation on T, from which we can generate, numerically, sample values. I then make use of the conditional dependence structure between these variables and the definition of block maxima in equation 2.12 to produce a sample value from the MEVD. An efficient sampling algorithm is as follows 2.1 In order to estimate the quantile ξ_p at a given level of

---

Algorithm 2.1 Sampling procedure from the MEV distribution of Precipitation

1: Take a sample n_i from the probability distribution of the number of wet events $f_N(t)$
2: Take a sample t_i from the probability distribution of temperature $f_T(t)$
3: Take n_i samples $x_1, \ldots, x_{n_i}$ from the conditional distribution of ordinary precipitation $F_{n_i}(x|T = t)$ whose parameters are given by $c = c(t_i)$ and $w = w(t_i)$ in the case of Weibull Distribution (see section 2.5.2)
4: A single sample of Y is given by $y_i = \max\{x_1, \ldots, x_{n_i}\}$

---

accuracy represented by a confidence interval, samples $Y_1, \ldots, Y_n$ are drawn from the MEV distribution according to algorithm 2.1 and ranked in increasing order as $Y_{(1)}, \ldots, Y_{(n)}$. An estimator of ξ_p is given by:

$$\hat{\xi}_p \equiv Y_{[np]} \qquad (2.16)$$

The criteria given by Briggs and Ying [29] to compute the minimum sample size that guarantees the required width ϵ_{abs} of the empirical confidence interval with a predefined probability, is summarized as follows.

The event $Y_{(k)} \leq \xi_p$ is true when at least k values in the sample are less than ξ_p. Since each realization of Y has an independent and identical probability of being smaller than ξ_p denoted by $\Pr(Y \leq \xi_p) = p$ by definition of a quantile; the number N_p of values in the sample less or equal that ξ_p follows a binomial distribution, i.e., $N_p \sim Binom(n, p)$. Thus,

$$\Pr\left(Y_{(k)} \leq \xi_p\right) = \Pr\left(N_p \geq k\right) = \sum_{i=k}^{n} \binom{n}{i} p^i (1 - p)^{n-i} \qquad (2.17)$$

From 2.17, the $(1 - \alpha)$-confidence interval $[X_{(r)}, X_{(s)}]$ of the estimator $\hat{\xi}_p \equiv Y_{[np]}$ is given

by:

$$\Pr\left(Y_{(r)} \leq \xi_p \leq Y_{(s)}\right) = \Pr\left(r \leq N \leq s\right) = \sum_{i=r}^{s-1} \binom{n}{i} p^i (1-p)^{n-i} > 1 - \alpha \quad (2.18)$$

where n is the required sample size.

The quantiles can then be computed by initially drawing a batch of samples of size n_{batch} from the MEVD. The ranks r and s that satisfy the condition given in equation 2.17 are iteratively computed by initially setting them to the mode of the binomial distribution $binom(n, p)$, and iteratively decrementing r and incrementing s until the condition is met. This procedure is repeated as more batches of samples are added to the previous ones until the resulting confidence interval matches the requirement. In practice, a batch size of $n_{batch} = 1000$ was arbitrarily chosen. In the case of the simultaneous estimation of multiple quantiles, samples are drawn in batches until all confidence intervals are narrower than the prescribed width.

2.4 Study Area and Dataset

In this thesis, the analysis at the daily scale is focused on data from the Padova and the Oxford Radcliffe observatory, chosen for their multi-centennial records of temperature and precipitation observations. The Padova time series consists of 284 years of daily rainfall records [30] along with 236 years of daily temperature observations (Figure 2.1i). Similarly, the Oxford dataset provides daily temperature records starting from 1815 and daily precipitation data from 1827 [31] (Figure 2.1ii). Detailed information about these time series can be found in the respective references. The analysis of these long-term instrumental observations holds significant value, as underscored by Marani and Zanetti [30]. Such multi-centennial records of daily precipitation provide direct insights into the statistical characteristics of rainfall, specifically in terms of extreme events and shifts in climatic patterns, allowing investiga-
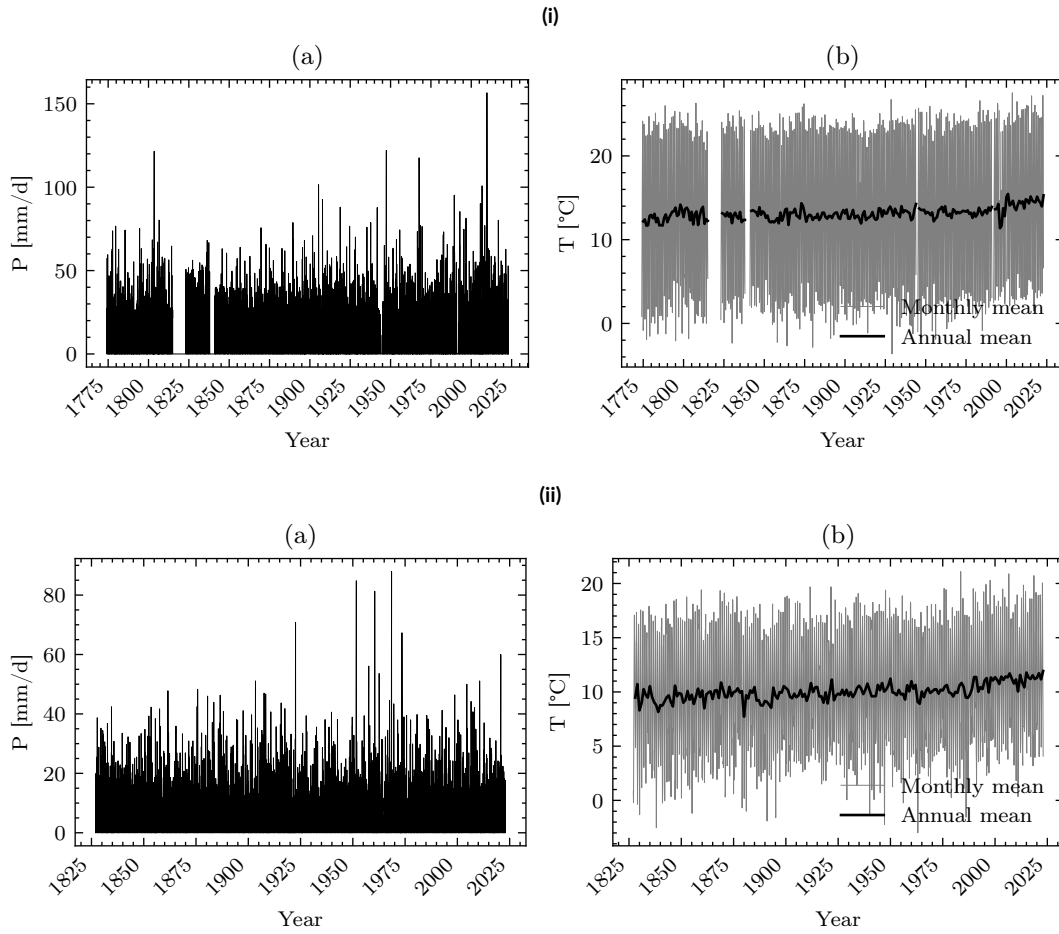
Figure 2.1: The daily rainfall and temperature observations in Padova (Italy) (i) and Oxford (U.K.) (ii)

tion into the extent to which trends in extremes can be explained by empirical scaling relationships between precipitation and temperature.

As for the hourly scale analysis, data from various continental US stations archived at the National Climatic Data Center (NCDC) are used. Building upon previous work by Shaw et al. [32], these stations have been identified to encompass a range of climatic and moisture availability conditions, as shown in the annual average total precipitation in Figure 2.2 In addition to the regions covered by Shaw et al., stations from the southeastern US, representative of the humid subtropical climate are also incorporated. A more detailed summary of the sta-
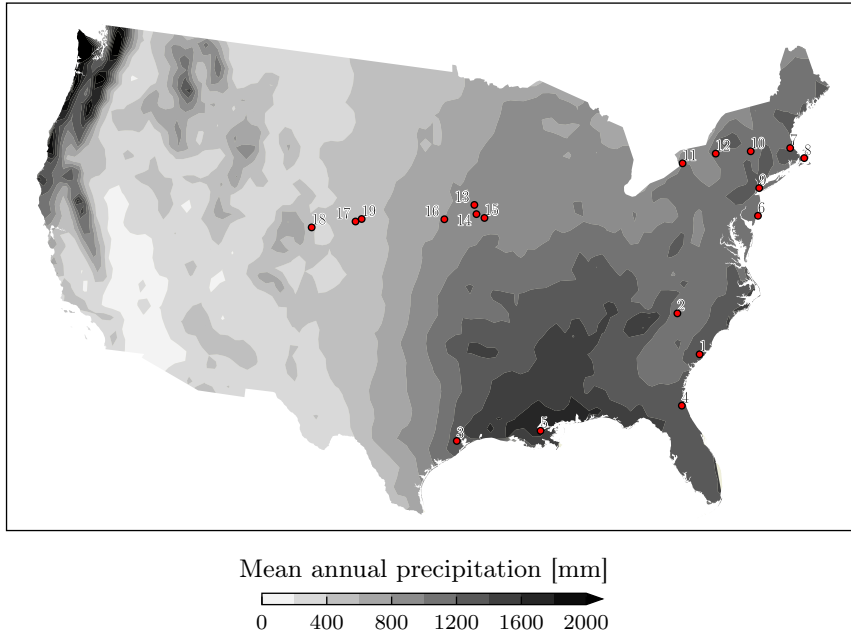
Figure 2.2: Map showing the U.S. annual mean total precipitation computed from the CPC dataset (1981 - 2010) and the location of the weather stations included in the analysis

tions can be found in Table 2.1 Particularly, our methodology excludes rainfall induced by tropical cyclones for the southeastern station.

Cyclone trajectories extracted from the HURDAT2 database [33] are used to classify instances of cyclonic rainfall, specifically when a cyclone is within a distance of 250 km [34].

Furthermore, it is noteworthy that daily mean temperature is used instead of hourly temperature to derive the scaling of hourly precipitation. As demonstrated in previous studies, hourly surface temperature, susceptible to fluctuation due to varying radiation intensity, is more related to boundary layer processes and may not well represent the conditions in the higher part of the atmosphere. [10, 32, 16]. Daily mean temperature, in contrast, provides a better representation of the temperature of the air mass associated with precipitation events. In subsequent analyses, a threshold of 1 mm (0.1 mm) is used to define daily (hourly) rainfall events.

14

ID	Region	COOP ID	Location	Period of analysis
1	Southeast	381544	Charleston, SC US	1948 - 2013
2		311690	Charlotte, NC US	1948 - 2013
3		414307	Houston, TX US	1948 - 2013
4		84358	Jacksonville, FL US	1948 - 2013
5		166660	New Orleans, LA US	1954 - 2014
6	Coastal Northeast	280325	Atlantic City, NJ US	1948 - 1993
7		190770	Boston, MA US	1948 - 2013
8		193821	Hyannis, MA US	1948 - 2001
9		305811	New York City, NY US	1948 - 2013
10	Interior NY	300042	Albany, NY US	1948 - 2013
11		301012	Buffalo, NY US	1948 - 2013
12		308383	Syracuse, NY US	1948 - 2013
13	Central Plains	130200	Ames, IA US	1964 - 2013
14		132203	Des Moines, IA US	1948 - 2013
15		134502	Knoxville, IA US	1948 - 2013
16		256255	Omaha, NE US	1948 - 2013
17	Western Plains	250865	Big Springs, NE US	1948 - 2013
18		53005	Fort Collins, CO US	1948 - 2013
19		254455	Kingsley Dam, NE US	1948 - 2013

Table 2.1: U.S. Stations summaries

2.5   STATISTICAL ANALYSIS

2.5.1   TIME SERIES DECLUSTERING

As discussed in Section 2.2, the Metastatistical Extreme Value Distribution is based on the statistical independence assumption among ordinary events. To guarantee this property, a careful declustering of the time series is essential. Following the methodology outlined by Marra et al. [35], event separation is carried out by computing one running parameter for each year of each station. This parameter defines the minimum time interval allowed between independent events [36]. For this purpose, the serial correlation of the time series is

computed across time lags up to 30 days. At each time lag, the long-lag noise is determined by the 90th quantile of the empirical autocorrelations observed for longer lags. The running parameter is then defined as the specific time lag at which the autocorrelation first aligns with the long-lag noise level. Subsequently, rainfall events occurring within intervals shorter than the running parameter are grouped into clusters while retaining only the cluster maxima. These cluster maxima define the independent ordinary events.

2.5.2 MODEL FITTING

Several studies have used the Weibull distribution to model rainfall depth at the daily [28, 37, 34] and the hourly [35] scale, following a global study by Wilson and Toumi [38]. Accordingly, I make the assumption that the conditional distribution of precipitation on temperature can be modeled as a Weibull distribution whose parameters vary with temperature. In this thesis, the dependence of the parameters on temperature is defined as a piecewise constant dependence, where one observation of the covariate T is, in turn, associated with specific values of the parameters estimated by the samples within a predefined temperature bin. All observations falling in one bin are then assumed to have common characteristics. The Weibull distribution is fitted to each bin by means of the Probability Weighted Moments (PWM) method which performs well for small samples and is less sensitive to the presence of outliers [39].

The goodness-of-fit of a Weibull distribution to data in each bin is assessed using the Kolmogorov-Smirnov hypothesis test [40]. This nonparametric goodness-of-fit test is used to compare a sample from a hypothesized distribution - the Weibull distribution in this context.

Finally, I use the relative frequencies of the temperature bins to model the frequency distribution of temperature T. Similarly, the relative frequency of the sample values of the

number of wet events per year N is used to estimate its frequency distribution.

2.5.3 MODEL VALIDATION

To validate the model and quantify the uncertainty in quantile estimation, a Monte-Carlo-based cross-validation procedure is carried out. Aside from MEV P(T), the standard MEV and the GEV are also tested for reference. To this end, the observations on record are split into yearly blocks, which are reshuffled randomly, thereby preserving the yearly frequency distribution of the observations. A L_{cal}-year-long calibration sample to which the EV models are fitted, and a L_{val}-year-long validation sample to assess the accuracy of the quantiles derived from the EV models are drawn from the reshuffled data. L_{val} should be equal to or greater than the return period of interest so that empirical cumulative frequencies of the annual maxima can be used as references for the exceedance probabilities inferred through the EV models. The empirical cumulative frequencies are estimated using the Weibull plotting position formula:

$$F_i = \frac{i}{L_{val} + 1} \tag{2.19}$$

where $i = 1, \ldots, L_{val}$ denotes the rank of a yearly maximum y_i in the sorted sample of annual maxima in the validation set in ascending order. The estimates $\hat{y}_i = \zeta_{EV}^{-1}(F_i)$ (where ζ_{EV}^{-1} denotes the inverse cumulative distribution function of the EV distribution considered) is then compared with y_i and the relative error is computed as:

$$\epsilon = \frac{\hat{y}_i - y_i}{y_i} \tag{2.20}$$

This process is repeated n_r times to provide a full statistical description of the error metric. The empirical return period associated with each quantile is given by:

$$T_{r_i} = \frac{1}{1 - F_i} \qquad (2.21)$$

Thus, for a given return period T_r, the Fractional Standard error (FSE) can be computed as the square root of the average across all the Monte Carlo realizations of the squared non-dimensional error associated with T_r as:

$$FSE(T_r) = [\frac{1}{n_r} \sum_{1}^{n_r} \epsilon(T_r)^2]^{\frac{1}{2}} \qquad (2.22)$$

3

Results

3.1 Kolmogorov-Smirnov Hypothesis Test

First and foremost, I carried out the Kolmogorov-Smirnov hypothesis test to identify an optimal number of bins for which the assumption that daily rainfall data observed in a given temperature interval (bin) are coming from a Weibull distribution is valid. The test is performed using a moving window approach, which allows the sampling of various segments of the time series while accounting for the interannual variabilities. For each window, the precipitation data is aggregated into n_{bins} discrete temperature bins. In this manner, I iteratively perform the Kolmogorov-Smirnov Hypothesis Test for windows of varying lengths for different number of bins $n_{bins} = 1, \ldots, n_{years}$ where n_{years} correspond to the number of years on record inside the window. For a given n_{bins}, the statistical test is performed at
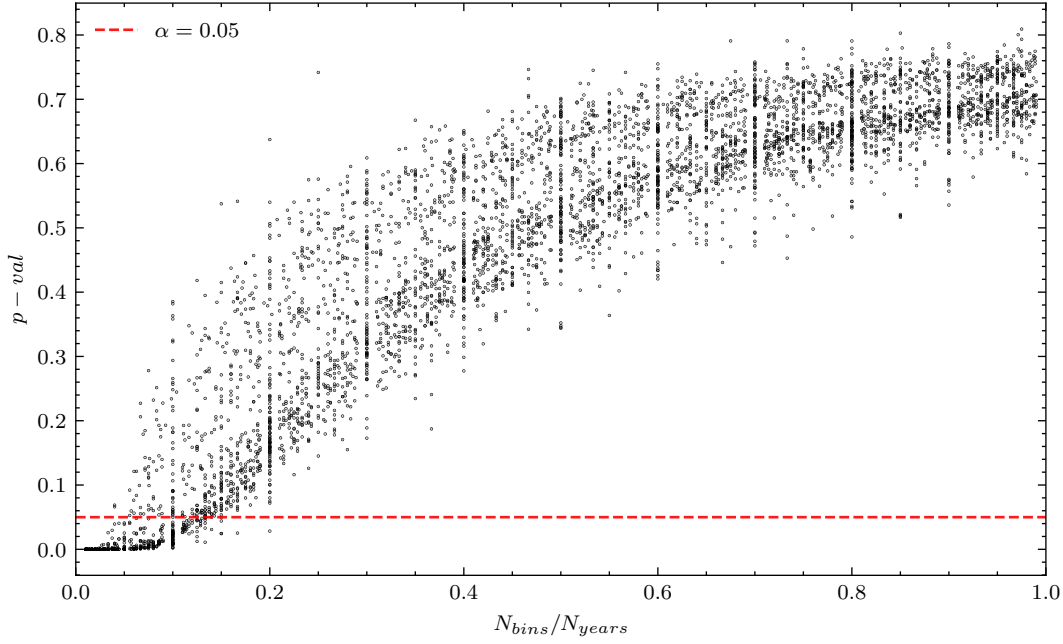
Figure 3.1: Scatter plot of the average p-values resulting from the Kolmogorov-Smirnov Hypothesis test against the number of bins normalized with the number of years in the estimation window.

each bin and the resulting p-values are subsequently averaged across all bins. This procedure is repeated for window sizes ranging from 20 to 100 years.

Figure 3.1 shows a scatter plot of the average p-values against the normalized n_{bins} defined as n_{bins}/n_{years}. To avoid overplotting, only a random subset of the complete result is displayed. We would expect the data to be consistent with the null hypothesis of the test if a good agreement exists between the empirical precipitation data and the Weibull distribution at each bin. This corresponds to a p-value larger than the significance level $alpha$. From the figure, we see that at 95% confidence level, the null hypothesis is rejected in favor of the alternative hypothesis for n_{bins}/n_{years} values below 0.02. This means that the Weibull assumption holds for a sufficiently large number of bins where $n_{bins} \gtrapprox 0.2\,n_{years}$. In subsequent analyses, unless specified, the number of bins n_{bins} is set to be equal to the number of years spanned by the estimation windows.

The scaling relationship was first derived for both Padova and Oxford. The precipitation and temperature data were paired and categorized into temperature bins, as detailed in Section 2.1.3. As illustrated in Figure 3.2, the 90th and the 99th daily precipitation percentiles generally scale at a lower rate than the CC-scaling rate, depicted by the dashed gray line, in both Padova and Oxford, except precipitation occurring at temperature below 5°C. Particularly in Padova, a hook-like structure is evident. The daily precipitation percentiles peak between 15°C and 20°C then begin to decline. The daily precipitation in Oxford, however, shows a monotonically increasing trend with temperature.
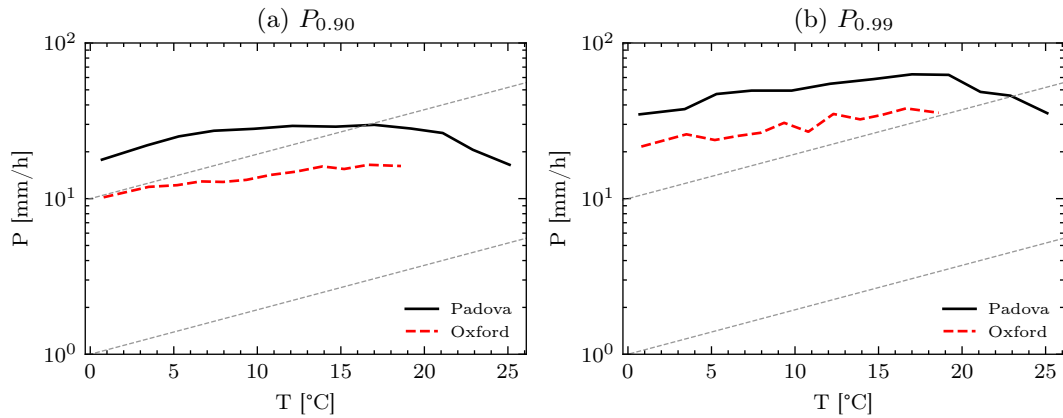
Figure 3.2: The 90 th (a) and 99th (b) percentiles of daily precipitation intensity as a function of daily mean temperature in Padova (Italy) and Oxford (b)

The P versus T relationship was also investigated for JJA and NDJ precipitation. The outcomes of these analyses are shown in Figure 3.3 and Figure 3.4. An exponential regression was fitted to the precipitation percentiles. The exponential regression appears to give a very reasonable fit except for the 99th percentile of JJA precipitation.

Summer and winter precipitation show very distinct characteristics in Padova. While NDJ precipitation percentiles display an increasing trend with T, JJA precipitation extremes
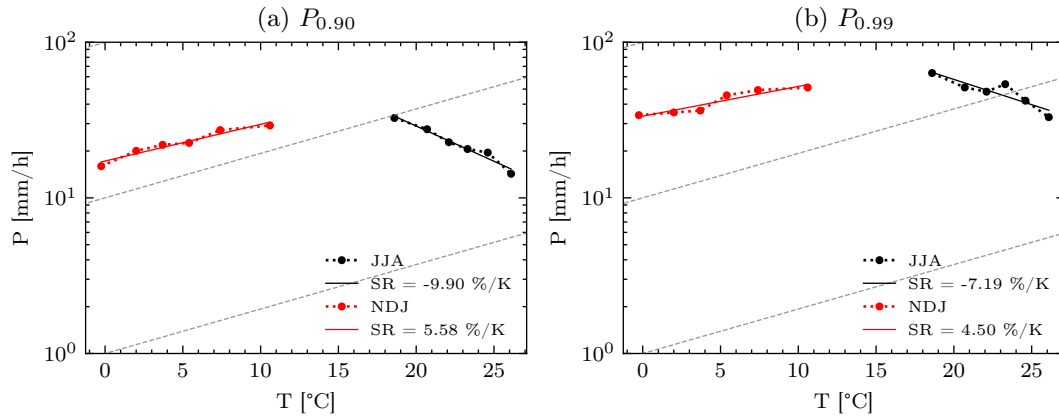
Figure 3.3: The 90 th and 99th percentiles of NDJ and JJA daily precipitation intensity as a function of daily mean temperature in Padova (Italy)
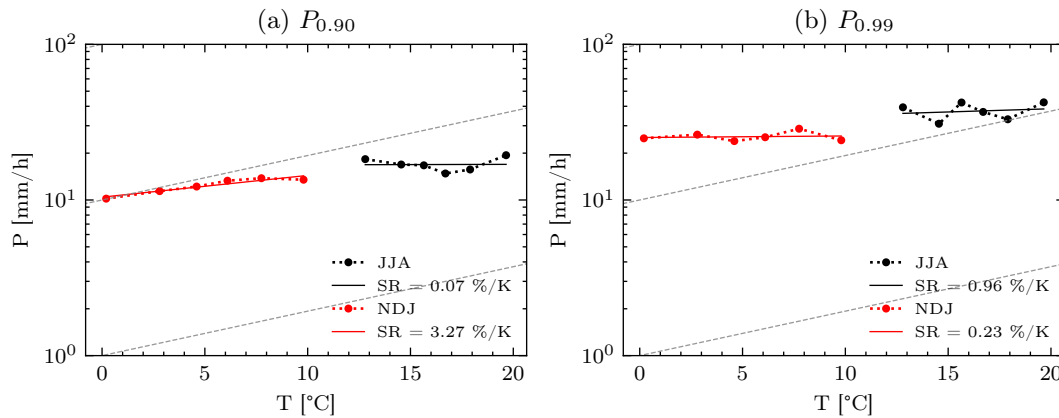
Figure 3.4: The 90 th and 99th percentiles of NDJ and JJA daily precipitation intensity as a function of daily mean temperature in Oxford (U.K.)

show the opposite. The corresponding scaling rates are also shown in the figures.

In the case of Oxford, examining the raw data, we generally observe a negative scaling of JJA precipitation with temperature at a temperature ranging between 12°C and 17°C. For temperatures above approximately 17°C, this scaling is positive. These opposite behaviors cancel out each other in the exponential regression. On the other hand, we observe an increasing trend only in the case of the 90th percentile. The 99th percentile seems not to display any trend as shown by the scaling rate close to zero.

Thereafter, the precipitation data were re-aggregated into n_{bins} equal to the number of years on record as discussed in Section 3.1 to assess the relationship between the Weibull parameters c and w and temperature T. The resulting plots for the Padova case are shown in semilogarithmic scale in Figure 3.5.

The scale parameter c values typically increase with T until a certain point beyond which c decreases. This point occurs at a temperature ranging between 16-18°C where the largest c values are observed. It also seems to correspond to the peak-point temperature observed earlier in Figure 3.2. The shape parameter w on the other hand, shows a decreasing trend with temperature.
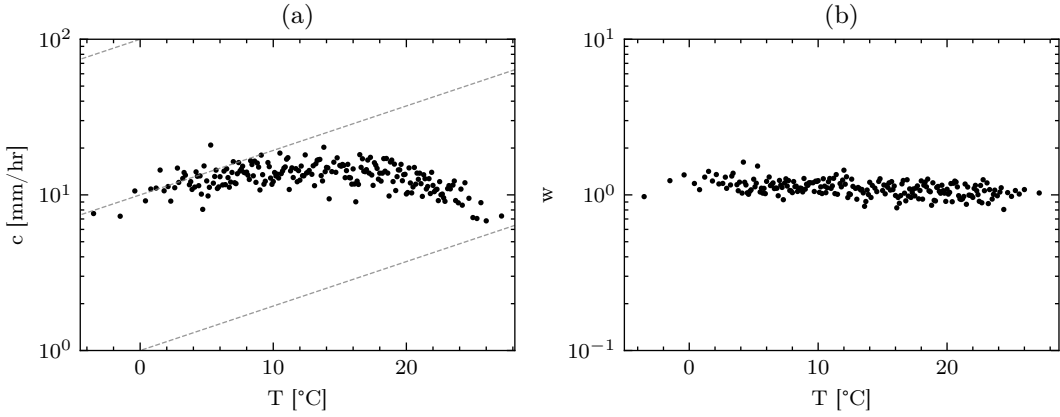
Figure 3.5: The Weibull distribution parameters of daily precipitation as a function of daily mean temperature in Padova (Italy)

This relationship was investigated closely by categorizing the data into NDJ and JJA precipitation (Figure 3.6). An exponential regression was fitted separately for each season to both $c - T$ and $w - T$. A marked distinction between the two seasons, similar to the one in Figure 3.3 can also be observed in the corresponding relationship between the Weibull distribution parameters and temperature, particularly for the scale parameter c. The overall shape of the $c-T$ plot looks very similar to the one in Figure 3.5. We see that the increasing trend in c we observed previously is associated with winter precipitation whereas the decreasing part
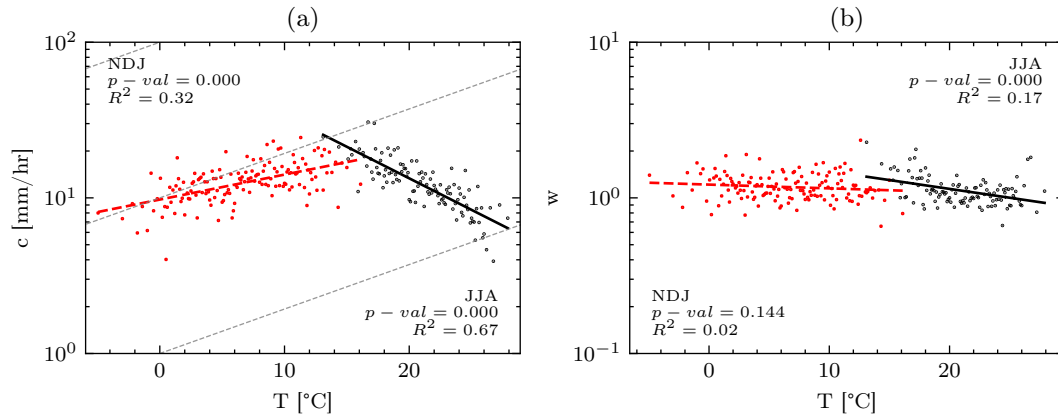
Figure 3.6: The Weibull distribution parameters of NDJ and JJA daily precipitation as a function of daily mean temperature in Padova (Italy)

corresponds to summer precipitation. The resulting p-values in both seasons make it evident that the (seasonal) $c - T$ relationship can be characterized by the statistically significant increasing exponential relationships in Figure 3.6a. As for the shape parameter w, a significant relationship is only observed in summer precipitation in which the w values decrease with temperature.

Figure 3.7 then depicts the dependence of the scale and shape parameters c and w of the Weibull distribution of daily precipitation on temperature in Oxford. The c seems to be positively correlated with temperature. Not much can be said about w.

As for Padova, the analyses were carried out separately for NDJ and JJA precipitation (Figure 3.8) At a 95% confidence level, a significant exponential relationship between c and T can only be seen in the NDJ precipitation. Interestingly, we observe a statistically significant increasing exponential relationship in $c - T$ for NDJ precipitation and a decreasing one for JJA precipitation. However, unlike the case of Padova only a small fraction of the variability in the data can be explained by an exponential relationship, as shown by the R^2 values. A statistically significant $w - T$ relationship is observed only in the JJA precipitation.

Figure 3.7: The Weibull distribution parameters of daily precipitation as a function of daily mean temperature in Oxford (U.K.)
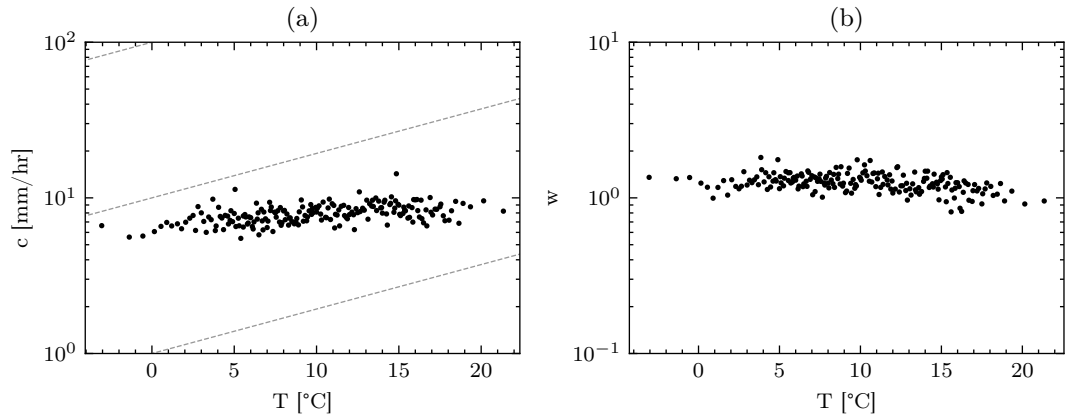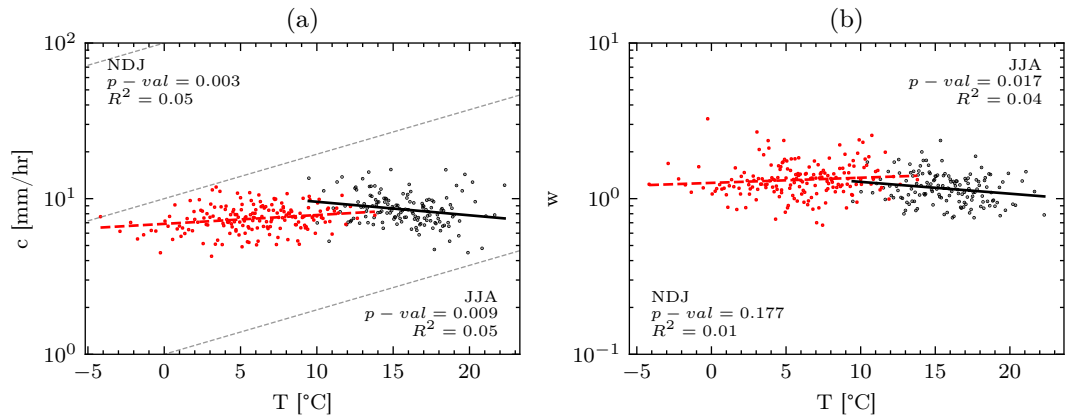

Figure 3.8: The Weibull distribution parameters of NDJ and JJA daily precipitation as a function of daily mean temperature in Oxford (U.K.)

3.3 Temporal Evolution of the Relationship Between Daily Precipitation and Temperature

A moving-window approach with a window length of 30 years was used to investigate the changes over time in the relationship between daily precipitation distribution and temperature. The choice of the window size is not trivial; while a larger window is associated with reduced parameter estimation uncertainties, it concurrently lowers the resolution at which

temporal changes are discerned due to the smoothing effect on high-frequency fluctuations. In the context of Padova, the analysis was conducted separately for JJA and NDJ precipitation. The changes over time in the relationship were indirectly estimated through the variations in the regression parameters of $c - T$ and $w - T$ as long as statistically significant relationships were established. At each window, an exponential regression was performed by applying a linear regression on the log-transformed parameter values that correspond to each temperature bin. The result of the analysis applied to the period between 1805 and 2022 in Padova is given in Figure 3.9. β_1 and β_0 in Figure 3.9b and Figure 3.9d represent the linear regression slope and intercept computed from the log-transformed data. The interpretation of their values is straightforward: β_1 represents the rate of growth or decay of the parameter with temperature, whereas e^{β_0} is the parameter value at $T = 0°C$. As depicted in Figure 3.9a and Figure 3.9c, the $c - T$ exponential relationship is statistically significant for the whole period between 1805 and 2022 for JJA precipitation. For NDJ precipitation, statistical significance is only observed between 1870 and 1985. On the other hand, there is often insufficient evidence to suggest a temperature dependence of the shape parameter w, except for very few instances in JJA precipitation, which displays a periodic recurrence as seen in Figure 3.9c.

The temporal variation of the regression parameters can be seen in Figure 3.9c and Figure 3.9d. Notably, β_1 and β_0 display opposing trends; when β_1 increases, β_0 decreases. Aside from the decadal fluctuations, oscillations with periodicities exceeding a century can be observed in both β_1 and β_0. Upon visual evaluation, an increasing (decreasing) trend in β_1 (β_0) can also be seen, particularly for JJA precipitation.

In the case of Oxford, the exponential regression was directly applied to the parameters of the Weibull distributions that were fitted within each temperature bin for each window between 1860 and 2022. The data were not categorized into seasons. This is because the
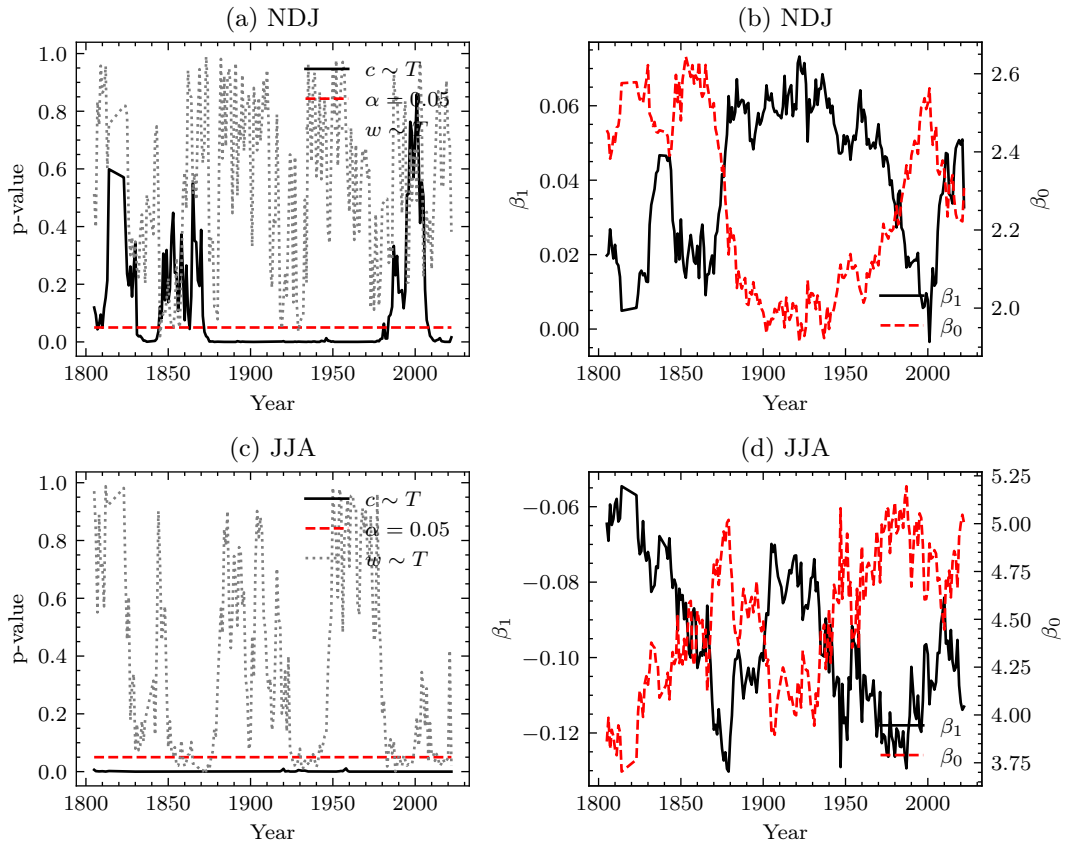
Figure 3.9: Temporal evolution of the relationship between Weibull distribution parameters of daily precipitation and temperature in terms of regression parameters in Padova (Italy). The p-values associated with the regression slopes are shown in (a) and (c). (b) and (d) show the evolution of the parameters of the regression applied to the scale parameter c

relationship between Weibull distribution parameters of daily precipitation and temperature can be described using a singular exponential relation. As illustrated in Figure 3.10a, the $c - T$ relationship is statistically significant for the majority of the time. An exponential relationship for $w - T$ on the other hand, started to become statistically significant after 1960. Analogous to Padova, an increasing and decreasing trend is observed for β_1 and β_0 respectively.
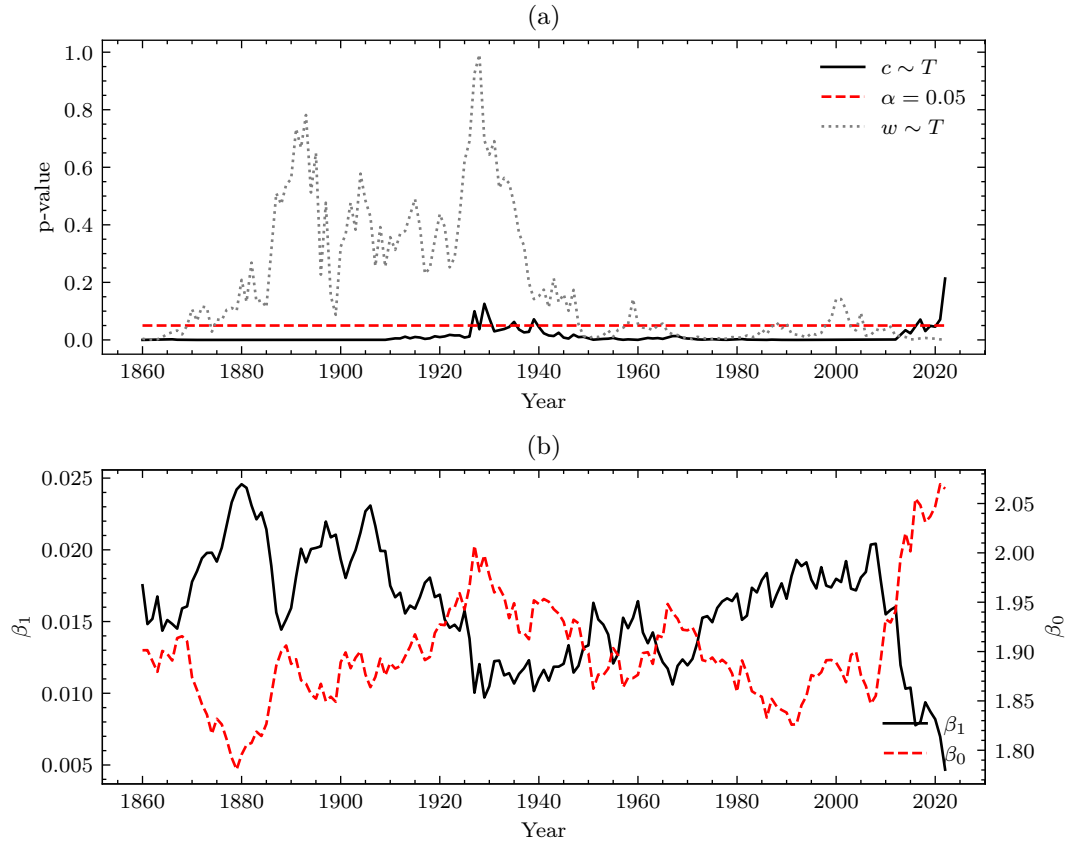
Figure 3.10: Temporal evolution of the relationship between Weibull distribution parameters of daily precipitation and temperature in terms of regression parameters in Oxford (U.K.). The p-values associated with the regression slopes are shown in (a). (b) shows the evolution of the parameters of the regression applied to the scale parameter c

3.4 MODEL VALIDATION

To validate the model and the simulation method that I proposed in Section 2.3, I conducted a thorough comparison of the MEV P(T) with the regular MEV approach and the traditional GEV method. Firstly, the MEV and GEV distributions were fitted to the Padova dataset using the Probability Weighted Moment method. Similarly, for the MEV P(T), the daily precipitation data corresponding to wet events were aggregated in temperature bins and a Weibull distribution was fitted to each bin. For reference, I proceeded by extracting the annual maxima y_i from the dataset, ranking them, and computing the corresponding

cumulative frequencies F_i using the Weibull plotting position formula (2.19). Then, the quantile estimates $\hat{y}_i = \zeta_{EV}{}^{-1}(F_i)$ are computed for each EV distribution. In the case of MEV P(T), quantile estimation is carried out by simulation in which the accuracy is set to be $\epsilon_{rel} = 0.05$. ϵ_{rel} corresponds to the pre-specified fraction of the whole (estimated) range of the quantile defined by the difference between the largest and smallest sampled values. During the analysis, an average sample size of $12,500$ was required during sampling to estimate the largest quantile observed in the Padova dataset at the pre-defined accuracy. The confidence intervals of the estimated quantiles were estimated by bootstrapping with replacement M random years from the full record, where M denotes the number of available years. The resulting return value plot is shown in Figure 3.11. As illustrated, the quantile estimates derived from the three methods are very close to each other. The uncertainty is largest in the GEV estimates and smallest in the standard MEV.
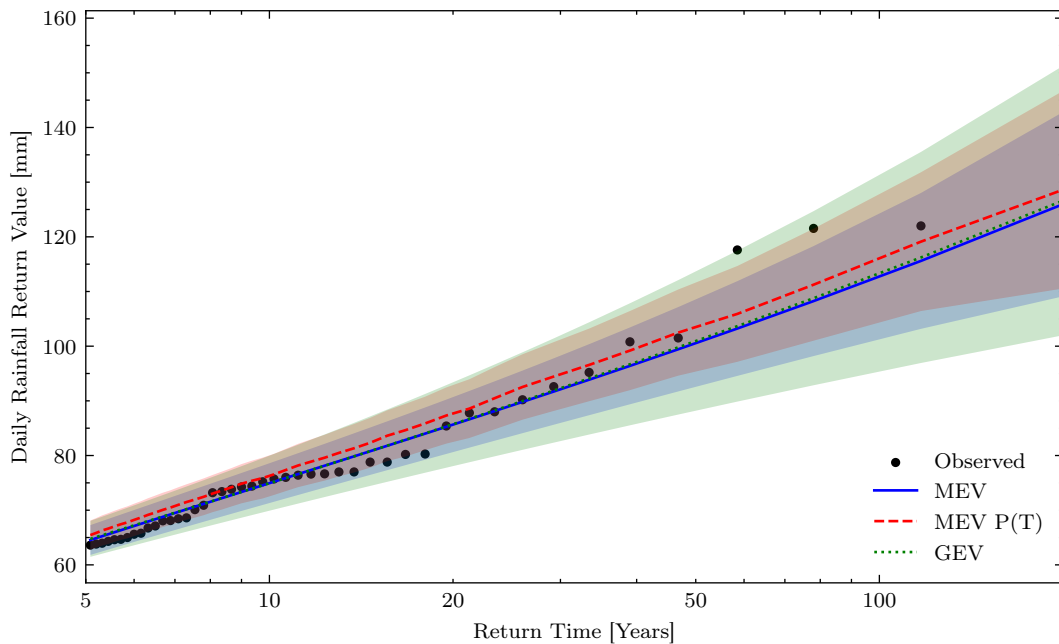
Figure 3.11: Return value plot of different estimates of the extreme value distribution $\zeta_{obs}(y)$

Subsequently, to address uncertainty in the estimation of high-return period quantiles

with limited data, I carried out a cross-validation Monte Carlo experiment with $n_r = 1000$ realizations. During this procedure, the observational years were reshuffled while preserving the yearly frequency distribution of the daily events. Then a 30-year-long calibration and a 100-year-long validation sample were extracted from the reshuffled series. The chosen length of the validation sample ensures that the value corresponding to the kth percentile within the sample serves as an approximation of the k-year return period quantile observed in the sample. To quantify the accuracy of the EV models, the evaluation metrics described in Section 2.5.3 were computed. (Figure 3.12). Figure 3.12 shows violin plots of the distribution of the relative error of the quantile estimates of each method. The corresponding FSE are shown in Figure 3.12. Both of these figures indicate that the MEV and the MEV P(T) exhibit relatively similar performance. They provide quite accurate estimates of the quantiles in the case of Padova. In the Oxford case, however, both methods tend to underestimate quantiles, particularly when the return period considered exceeds the observation window. Nevertheless, they both demonstrate a significantly reduced variability compared to the GEV.

I further explored model uncertainty by investigating the proposed model's ability to replicate temporal changes in quantiles as estimated by the MEV. This assessment is motivated by the idea that MEV P(T) estimates should closely align with MEV estimates, given that MEV P(T) is just a reformulated MEV. I computed the $P_0.90$ and $P_0.99$ corresponding to the 10 and 100-year return period quantile of precipitation using both the conventional MEV and the MEV P(T) employing a moving-window approach. I carried out this analysis using window sizes of both 20 and 30 years, to evaluate the impact of window size on quantile estimation. The results are presented in Figure 3.14a and 3.14b, respectively. It is important to highlight that the grey shading in the figures represents the 95% empirical confidence interval which depends on the value of the pre-specified ϵ_{rel}. It represents the uncertainty associated
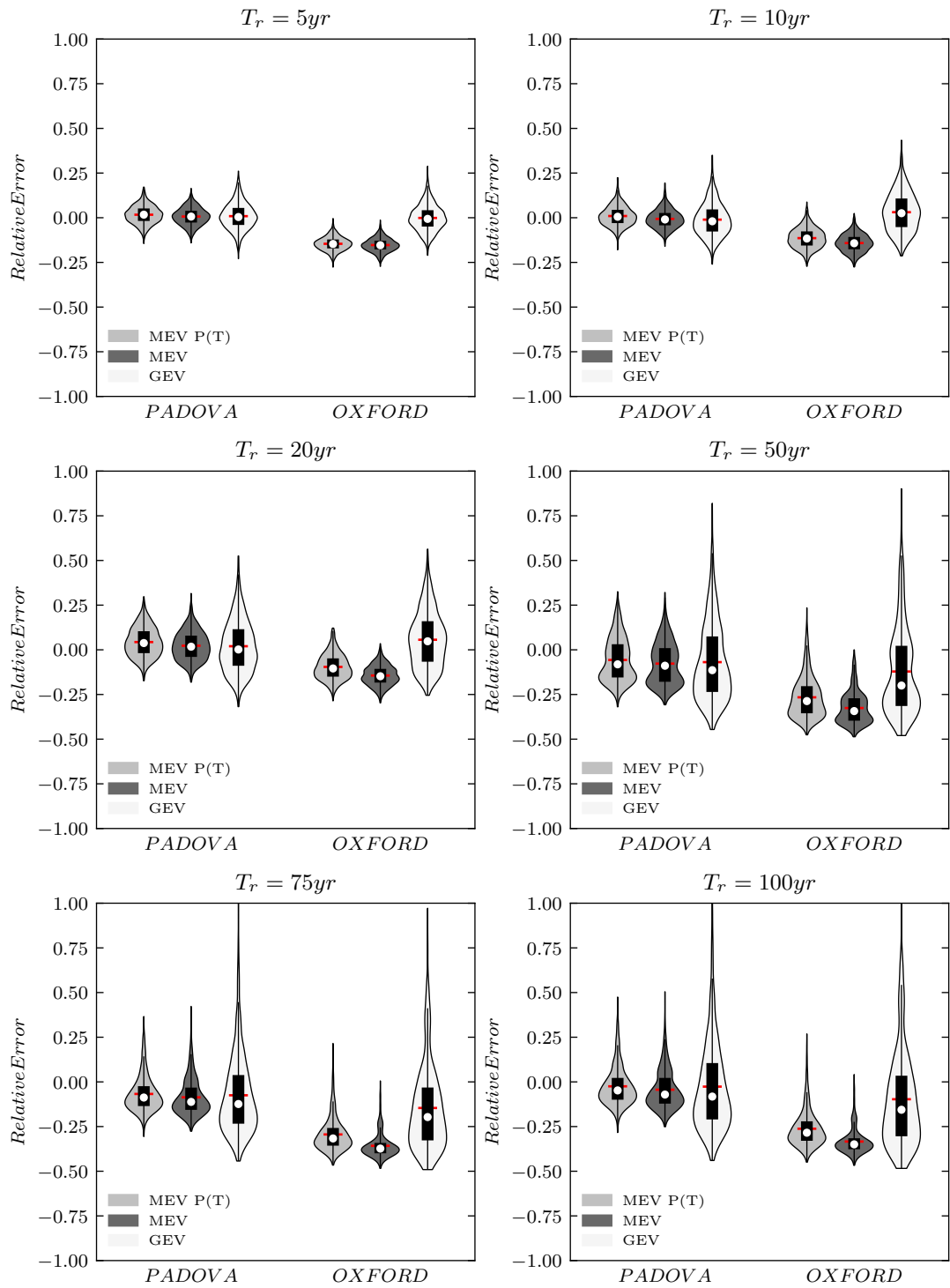
Figure 3.12: Distribution of the relative errors of different quantile estimates for different return periods

(a) Padova   (b) Oxford
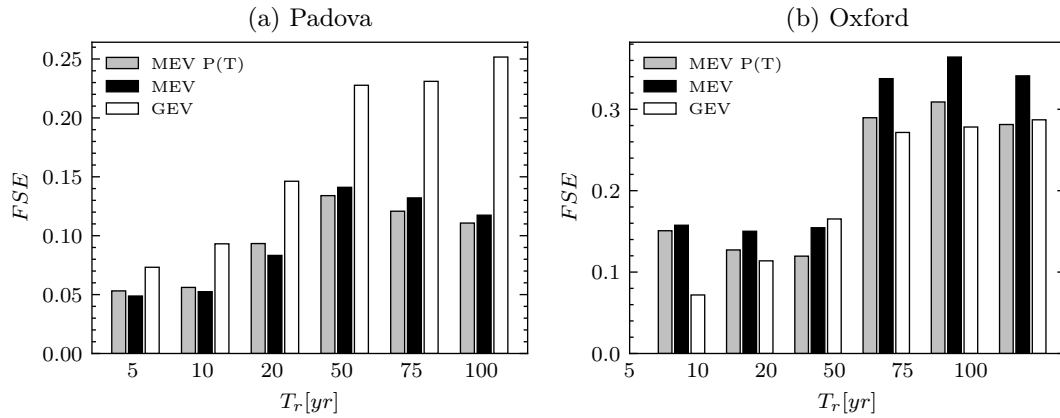
Figure 3.13: Bar chart of the FSE computed over all the Monte Carlo realizations of the relative error

with the quantile estimation via simulation, which stipulates that samples are drawn from the "True" distribution. Given the stochastic nature of the method used, the simulation results show inherent variability as seen in both figures. A close correspondence is evident between the MEV estimates MEV P(T), particularly for the 90th quantile. The results of the analysis carried out on the Oxford dataset are illustrated in Figure 3.15a and 3.15b. Analogous to the Padova dataset, the MEV P(T) estimates exhibit pronounced variability due to the stochastic nature of the method. Nonetheless, a close alignment between MEV and MEV P(T) estimates is observable.

3.5   Model Prediction

In this Section, the focus is not solely on assessing the predictive performance of the model; rather, it centers on exploring the potential behavior of precipitation extremes under a warming climate scenario, assuming a stationary relationship between precipitation distribution and temperature. To achieve this, I have derived the parameters of the conditional distribution of precipitation on temperature using data from the period prior to the year 1900 (Figure 3.16). Subsequently, I have assumed that the local physics is stationary and is captured by
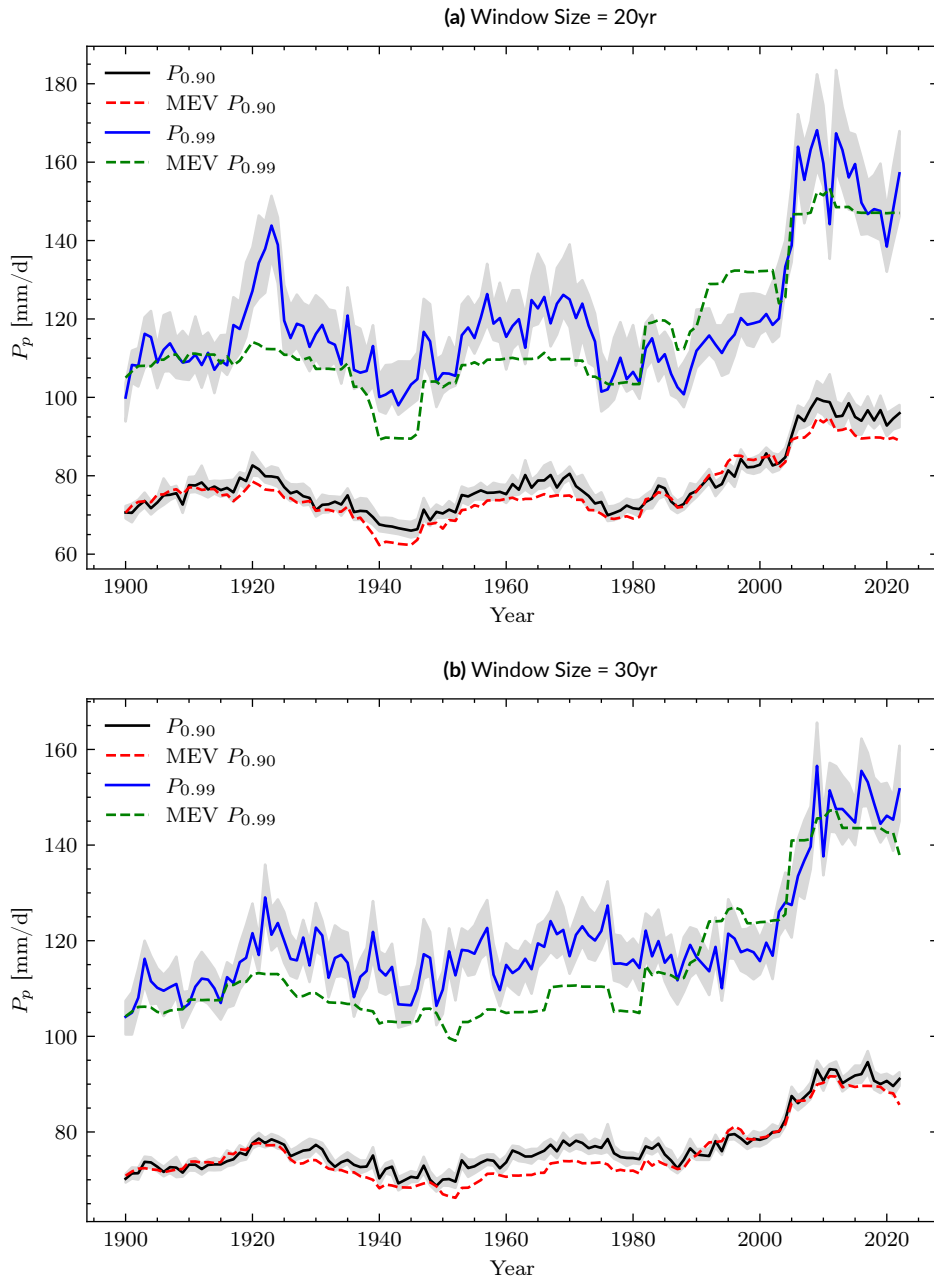
Figure 3.14: Validation of the proposed model by comparing its estimation of 10-year and 100-year daily precipitation quantiles with that of the regular MEVD on sliding window over the time series recorded in Padova (Italy) between 1900 - 2022.

(a) Window Size = 20yr

(b) Window Size = 30yr
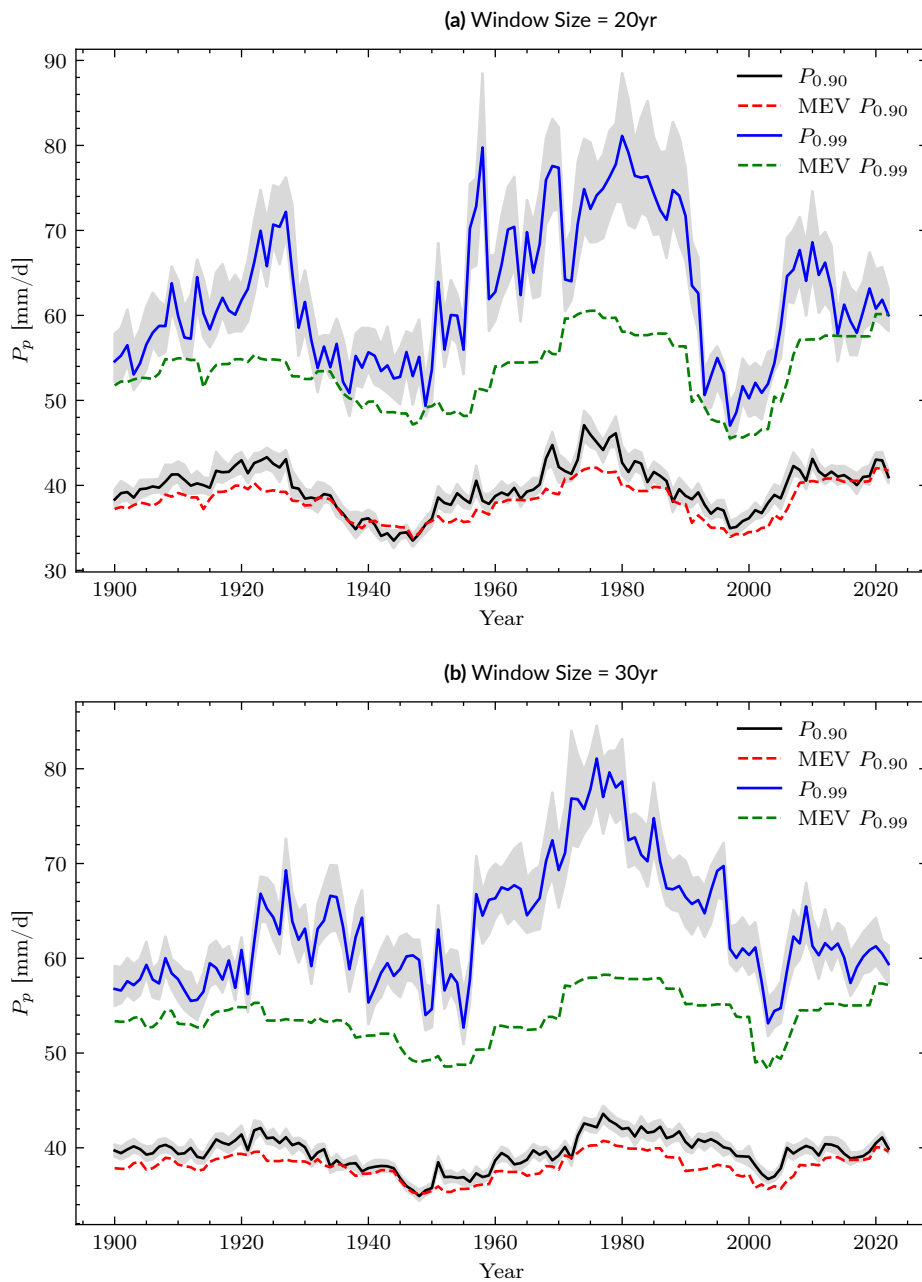
Figure 3.15: Validation of the proposed model by comparing its estimation of 10-year and 100-year daily precipitation quantiles with that of the regular MEVD on sliding window over the time series recorded in Oxford (U.K.) between 1900 - 2022.

the parameters' dependence on temperature. Then I use it to compute the 10- and 100-year return period quantiles using a moving window approach. Within each 30-year window, I have used the frequency distribution of temperature and the number of wet events per year of that window. This ensures that the variation of the temperature distribution captures the signal of global warming. The estimates of the standard MEV applied to the precipitation data within the 30-year-long window are used as references. The outcome of the analysis applied to the Padova dataset is shown in Figure 3.17. A contrast between the model pre-
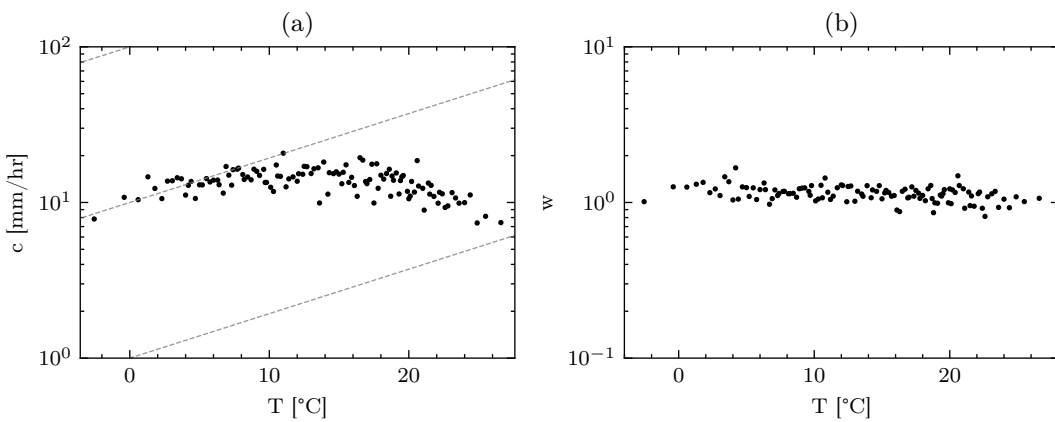
Figure 3.16: The Weibull distribution parameters of daily precipitation as a function of daily mean temperature in Padova (Italy) computed over the calibration period (1774 - 1899)

dictions and the reference value can be immediately noticed. Specifically, the model seems to predict a slight decline in the values of the 10- and 100-year return period quantiles, in contrast to the dramatic increase observed through MEV estimation after the year 1980.

To gain deeper insights into the MEV P(T) prediction, I carried out separate evaluations for winter (NDJ) and summer (JJA) precipitation. The relationship between the Weibull distribution parameter and temperature derived using data from the period prior to 1900 for each season is shown in Figure 3.18. The results of the sliding window analysis are presented in Figure 3.19. The empirical 10-year and 100-year return period quantiles computed using the entire dataset are depicted by the gray line in the Figures.
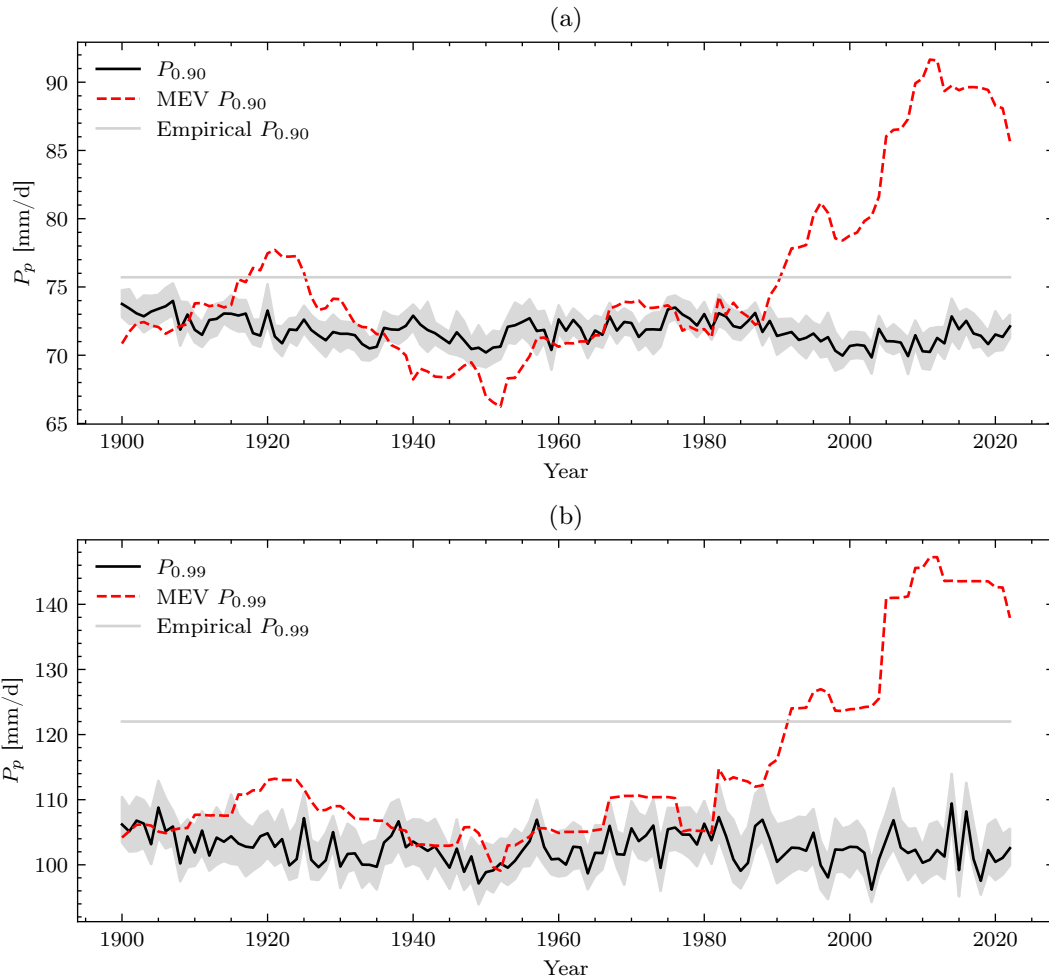
Figure 3.17: Predicted temporal evolution of 10-year and 100-year daily precipitation quantiles from 1940 to 2022 in Padova (Italy).

In the context of NDJ precipitation, the model is unable to replicate the low-frequency oscillations present in Figure 3.18a and Figure 3.18b. Although the predictions seem to display oscillations with the same periodicities, they are seemingly out of phase with the MEV estimates. Nevertheless, the predictions encapsulate the trend observed in the MEV reference (compare the quantile estimates at the beginning and the ones at the end of the analysis period in both Figure 3.19a and Figure 3.19b). Large discrepancies emerge in the case of JJA precipitation. The model predictions are not able to capture the periodic oscillation in

Figure 3.18: The Weibull distribution parameters of NDJ and JJA daily precipitation as a function of daily mean temperature in Padova (Italy) computed over the calibration period (1774 - 1899)
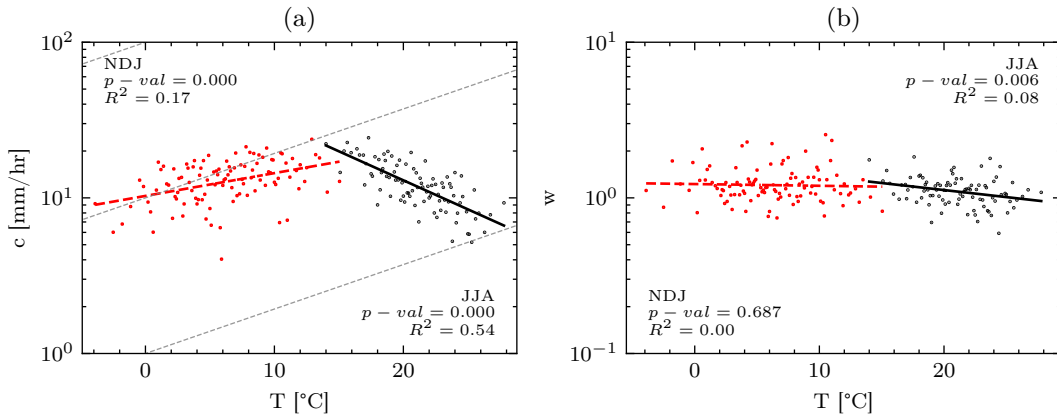
both quantiles. What is more striking is the opposite trends present in the reference MEV quantiles and the prediction. These trends are very clear in Figure 3.19c and more subtle in Figure 3.19d due to the difference in the plotting scale.

The same analysis was conducted on the Oxford dataset. The $c - T$ and $w - T$ values derived using data from the pre-1900 period that were used in the prediction are given in Figure 3.20. The moving-window analysis results are shown in and 3.21. As before, the predictions were directly computed without separating them into seasons. As seen from the figure, the model's predictions closely align with the reference values, with the exception of low-frequency oscillations evident in the reference quantiles. Particularly, the increasing trend observed in the MEV estimations of the two quantiles is discernible in the predictions. However, note that the magnitude of the "error" due to the oscillations is much lower compared to the Padova case.

3.6 ANALYSIS AT THE HOURLY SCALE

Shifting the focus to the hourly scale, the analyses were carried out using the NCDC dataset as discussed in Section 2.4. Expanding upon the work of [32], the relationship between pre-
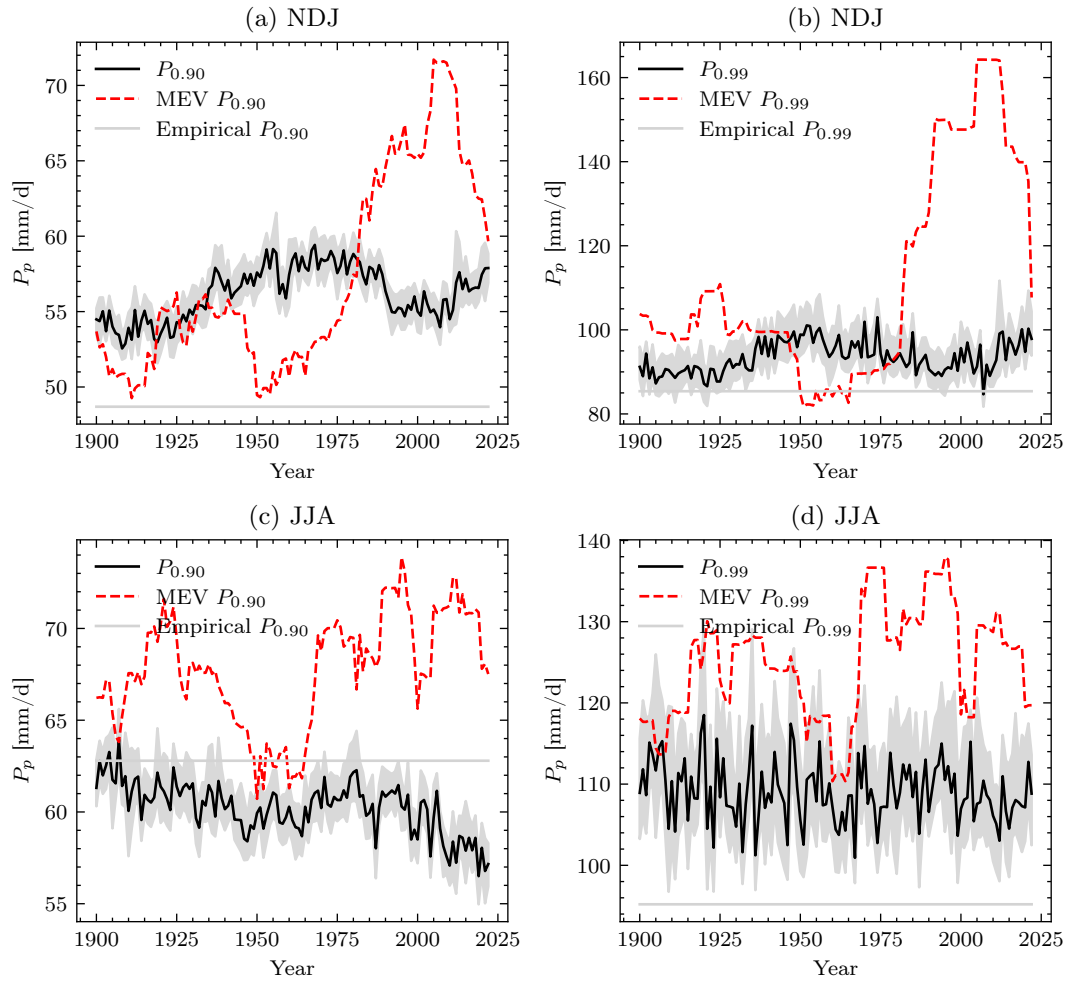
Figure 3.19: Predicted temporal evolution of 10-year and 100-year quantiles of NDJ and JJA daily precipitation from 1940 to 2022 in Padova (Italy).

cipitation distribution and temperature was investigated in five distinct regions within the contiguous United States: the Southeast, Coastal Northeast, Interior NY, Central Plains, and Western Plains. These regions differ in yearly total precipitation (see Figure 2.2). For each station in each region, a scaling analysis of precipitation percentiles with temperature was first carried out. The log-transformed 90th and 99th percentiles of hourly precipitation intensity are plotted against temperature and categorized by their respective regions in Figure 3.22. The dashed gray line in the figures, again, represents the reference C-C scaling rate
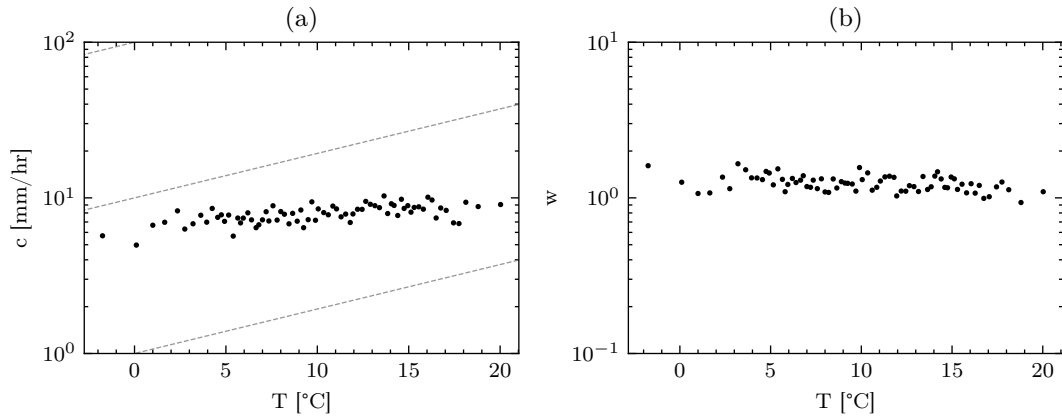
Figure 3.20: The Weibull distribution parameters of daily precipitation as a function of daily mean temperature in Oxford (U.K.) computed over the calibration period (1774 - 1899)

of $6.8\%/K$. Shaw et al. [32] carried out a comprehensive analysis of the P versus T relationship observed in various regions, excluding the Southeast. Summarizing their findings concerning the 99th percentile, they observed that the 99th precipitation percentiles generally scale with the CC-scaling rate, except Hyannis, MA. They attributed this discrepancy to the strong coastal influence specific to Hyannis, compared to the other stations they have studied. Expanding on their study, I have examined the Southeastern region (Figure 3.22a), characterized by a humid subtropical climate. Across most stations in this region, precipitation percentiles tend to increase at a similar rate close to the CC-scaling rate until reaching a peak around 24°C, beyond which it starts to decrease. This decrease is particularly pronounced in Houston, TX.

Following this, I analyzed the relationship between the precipitation distribution and temperature. The hourly precipitation data were grouped in temperature bins within which the Weibull distribution was fitted. Figure 3.23 shows plots of the log-transformed parameters against temperature with an exponential regression to characterize their relationship for each station in each respective region. The Weibull distribution parameters display a highly significant correlation with temperature. This is evident from the p-values associated with the

(a)

(b)
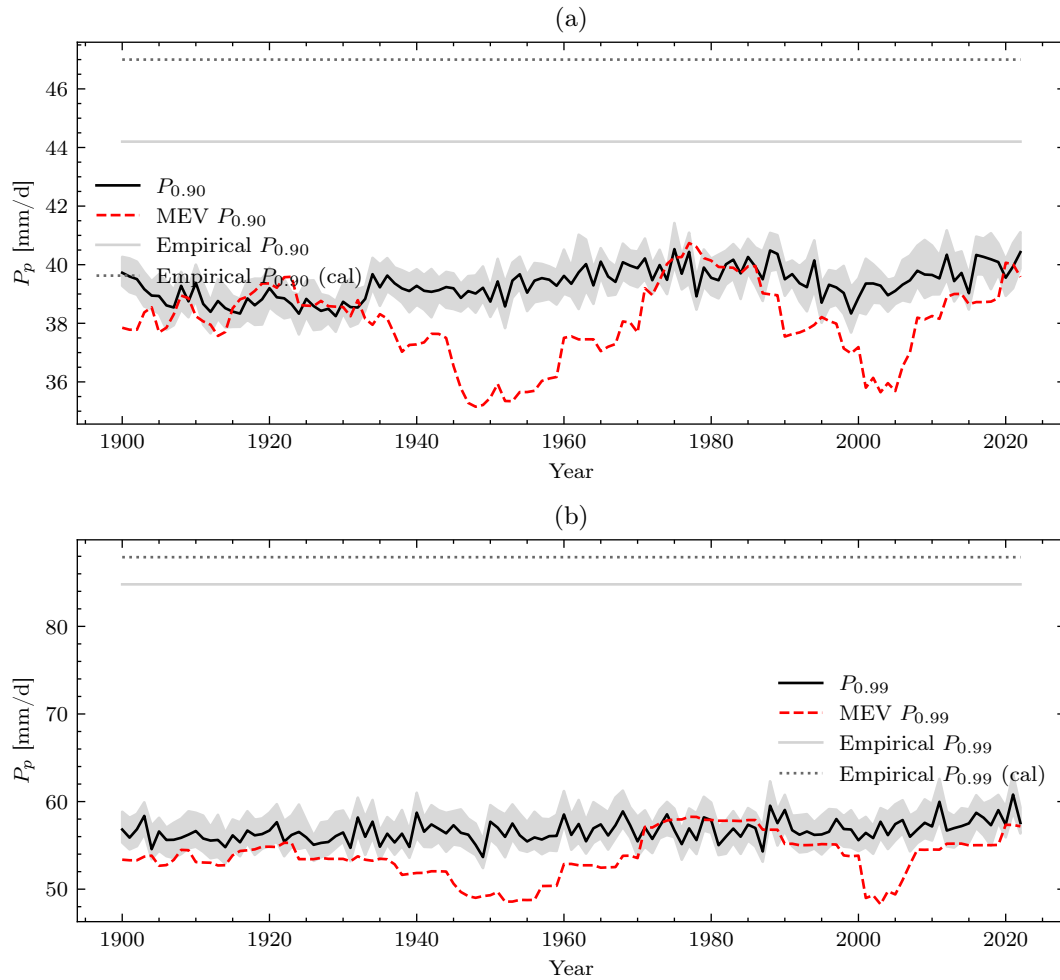
Figure 3.21: Predicted temporal evolution of 10-Year and 100-Yeardaily precipitation quantiles from 1940 to 2022 in Oxford (U.K.).

slope of the regression and the coefficient of determination. These p-values indicate that there is statistical evidence that the log-transformed scale parameter c increases with temperature. Conversely, the log-transformed shape parameter w tends to decrease as temperature rises. This is in contrast to the case of daily precipitation, where no statistical significance was observed. Interestingly, the exponential relationship accounts for a substantial fraction of the variance in the parameters, as shown by the R^2 values which are close to 0.9 in more than half of the stations analyzed. It is even more intriguing that R^2 is comparatively lower

in the station where the P versus T relationship deviates from the CC-relation.

3.6.1 By Precipitation Type

Hourly-Scale

The subsequent set of analyses was performed to explore the relationship between precipitation and temperature in different precipitation types. As elaborated in Section 2.1.1, a plethora of methods are available for classifying precipitation events into convective and stratiform types. In the context of this thesis, I use the occurrence of lightning as a proxy of convection. However, it is essential to recognize that observed lightning is a sufficient but not a necessary indicator of convective storm [16]. This consideration must be taken into account when interpreting the results.

In this part of the analysis, only a subset of stations were studied: Charlotte, NC, Houston, TX, and New Orleans, LA. Figure 3.24 illustrates the relationship between hourly precipitation percentiles and temperature for both convective and non-convective precipitation. In the case of non-convective precipitation, the scaling rates observed in New Orleans, LA for both $P_{0.90}$ and $P_{0.99}$ are very close to the CC-scaling rate. In contrast, the rate of increase for precipitation percentiles in Houston and Charlotte is considerably lower. Particularly, there is even a hint of reduced precipitation for temperatures exceeding 23°C in Houston. However, the number of rainfall observations beyond this point is limited and requires prudence in forming conclusions.

For convective precipitation, there is little to no relationship between $P_{0.90}$ and T. A nearly CC-scaling rate is apparent in Charlotte for the 99th percentile. In other cities, particularly in New Orleans, the raw $P_{0.99} - T$ data shows that intense precipitation can occur across a broad range of temperatures exceeding 15°C.

Finally, the relationship between the Weibull distribution parameters and temperature for different precipitation types is illustrated in Figure 3.25. In the case of non-convective precipitation, both c and w are very likely to increase and decrease respectively with T as corroborated by p-values near zero. The exponential model demonstrates a robust fit to the $c - T$ and $w - T$ relationships. However, for convective precipitation, this exponential relationship only aligns with data in Charlotte. Similar to the case of P versus T, there is no significant correlation between the Weibull distribution parameters and temperature.
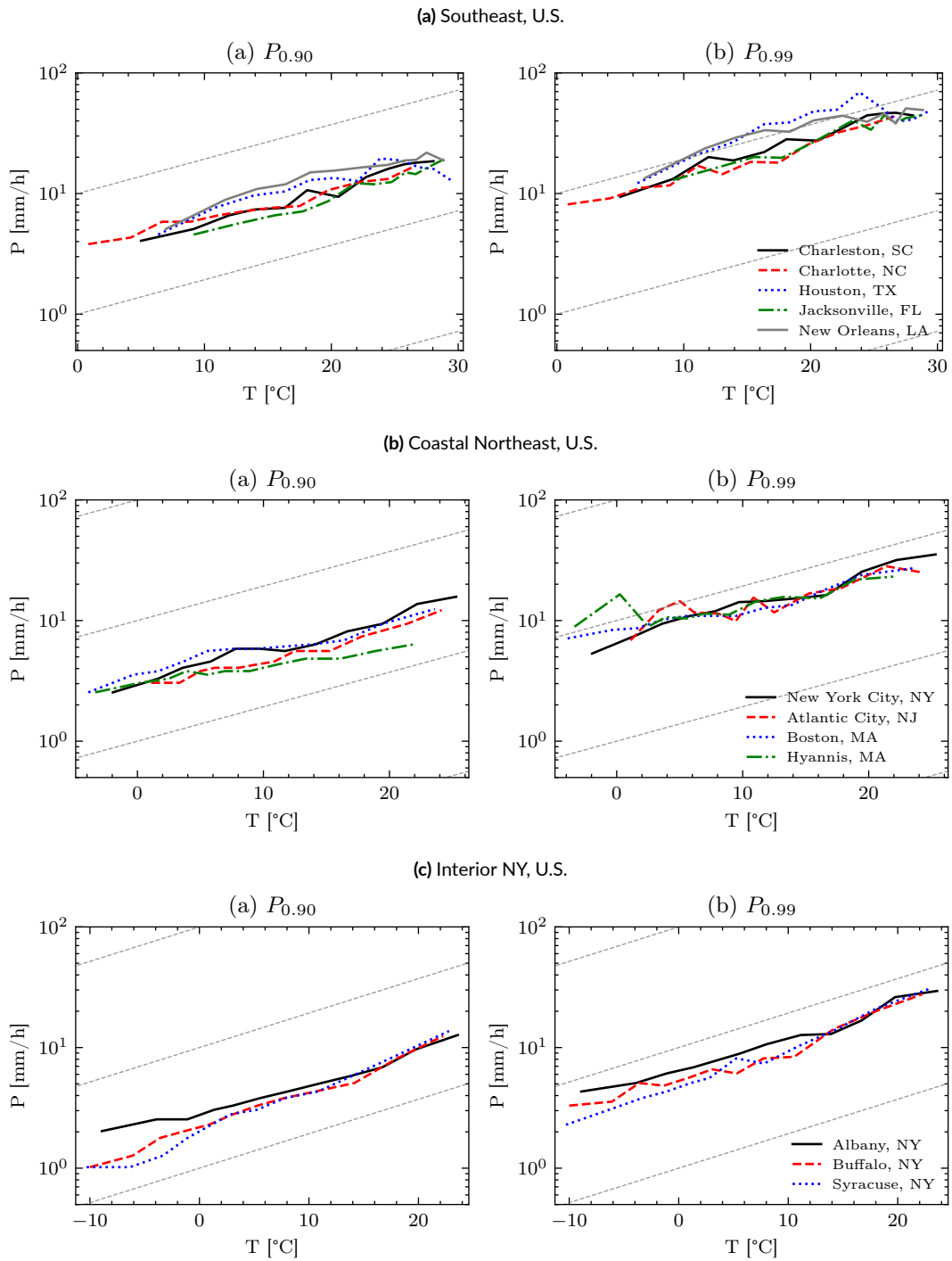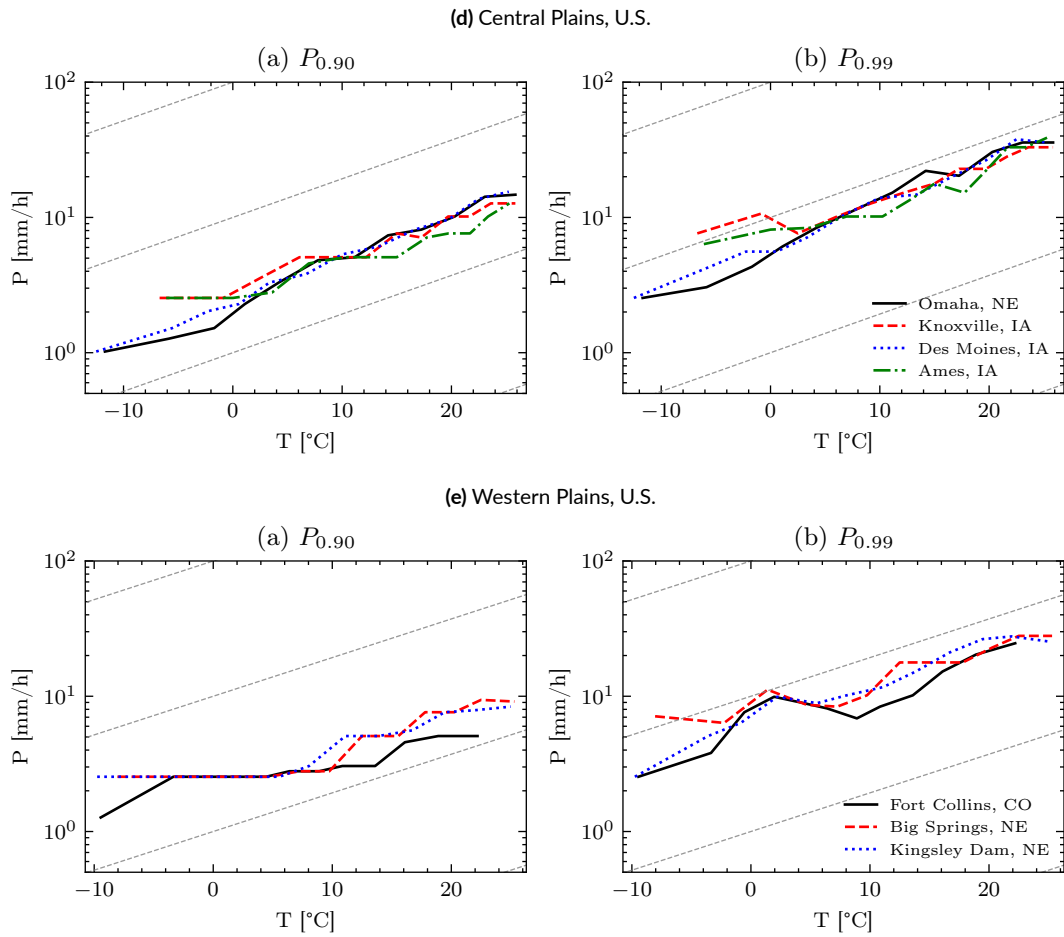
(a) Southeast, U.S.

(b) Coastal Northeast, U.S.

(c) Interior NY, U.S.

Figure 3.22: The 99th percentiles of hourly precipitation intensity as a function of daily mean temperature in different regions of the U.S.

(d) Central Plains, U.S.

(a) $P_{0.90}$

(b) $P_{0.99}$

(e) Western Plains, U.S.

(a) $P_{0.90}$

(b) $P_{0.99}$

Figure 3.22: The 99th percentiles of hourly precipitation intensity as a function of daily mean temperature in different regions of the U.S. (cont.)

Figure 3.23: The Weibull distribution parameters of hourly precipitation as a function of daily mean temperature in different regions of the U.S.

(iv) Central Plains, U.S.

(v) Western Plains, U.S.
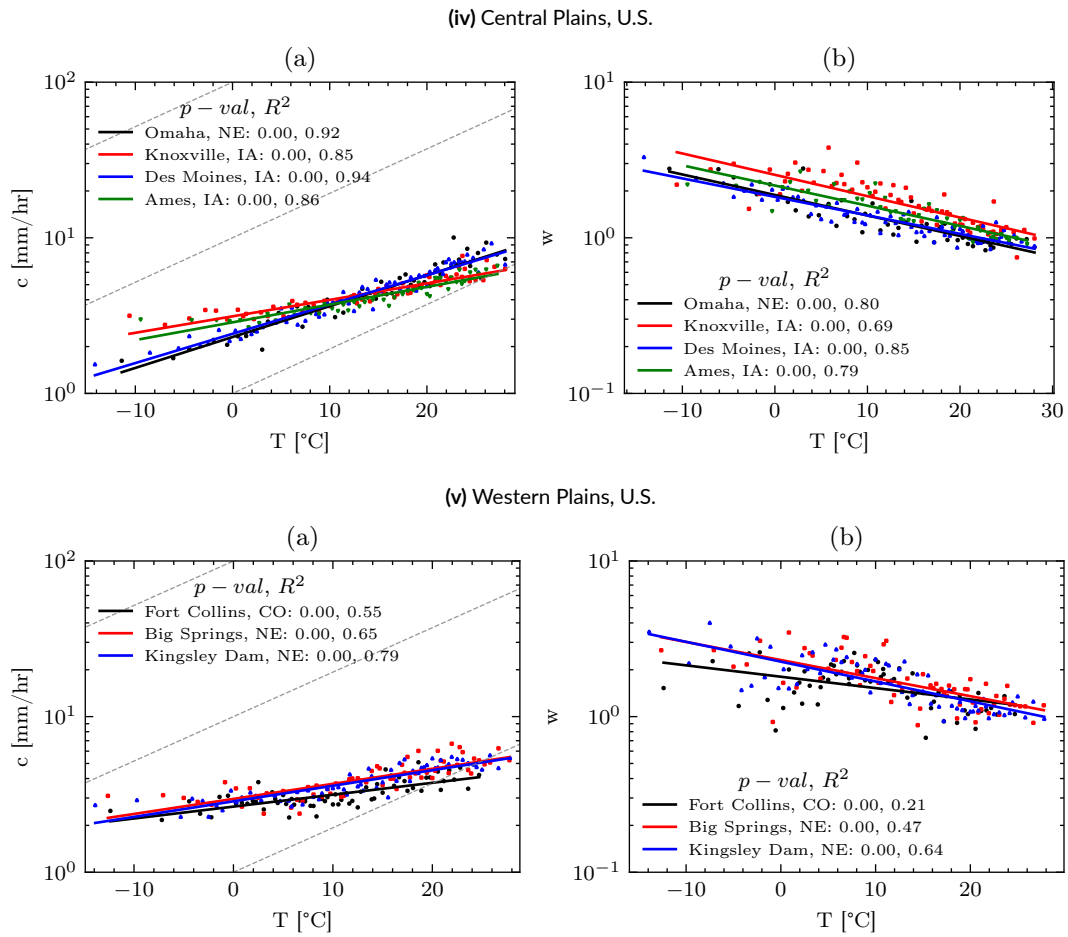
Figure 3.23: The Weibull distribution parameters of hourly precipitation as a function of daily mean temperature in different regions of the U.S. (cont.)
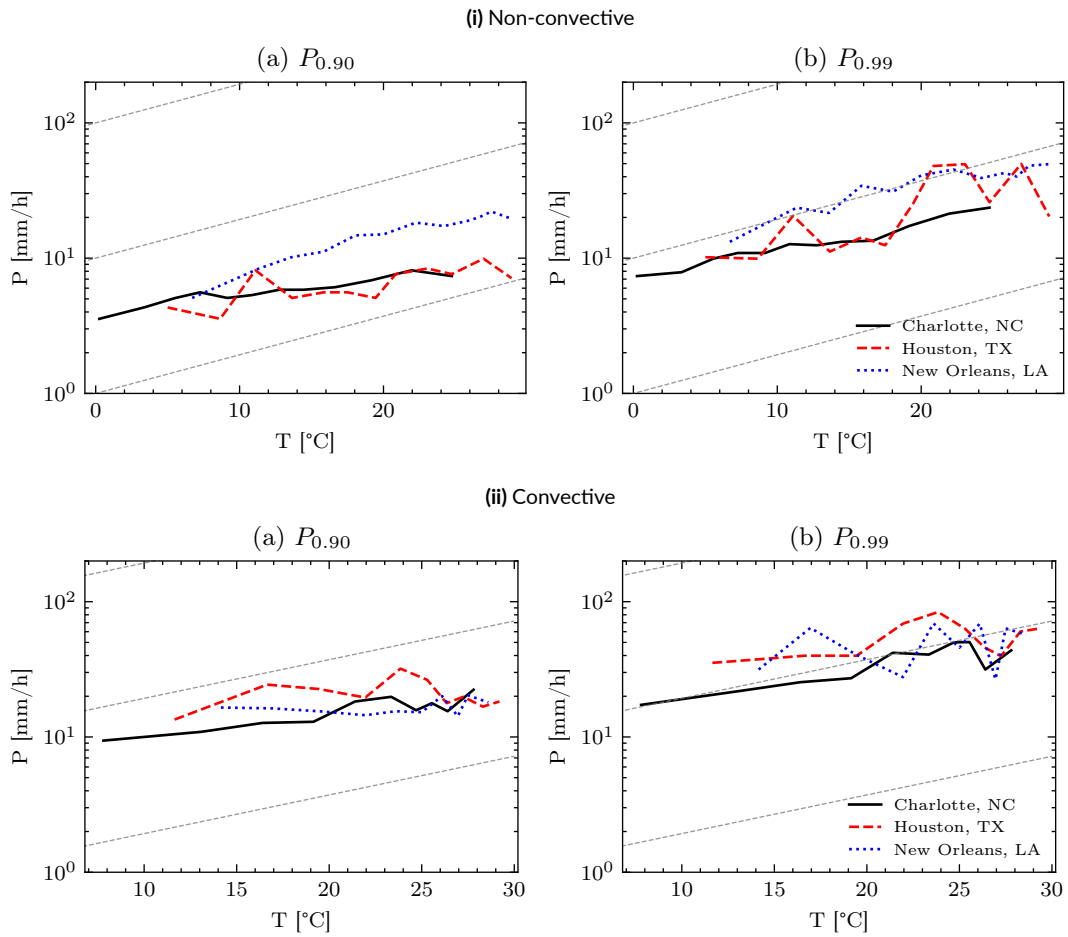
Figure 3.24: The 99th percentiles of hourly precipitation intensity as a function of daily mean temperature for each precipitation type in some stations in the U.S

(i) Non-convective

(a)

$p - val, R^2$
Charlotte, NC: 0.00, 0.61
Houston, TX: 0.00, 0.49
New Orleans, LA: 0.00, 0.73

(b)

$p - val, R^2$
Charlotte, NC: 0.00, 0.19
Houston, TX: 0.00, 0.40
New Orleans, LA: 0.00, 0.44

(ii) Convective

(a)

$p - val, R^2$
Charlotte, NC: 0.00, 0.48
Houston, TX: 0.67, 0.01
New Orleans, LA: 0.70, 0.00

(b)

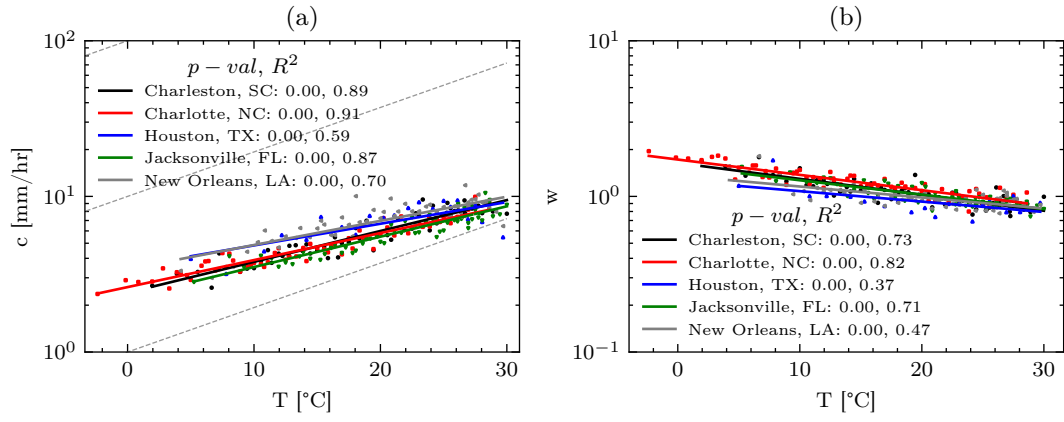$p - val, R^2$
Charlotte, NC: 0.00, 0.35
Houston, TX: 0.11, 0.12
New Orleans, LA: 0.04, 0.08
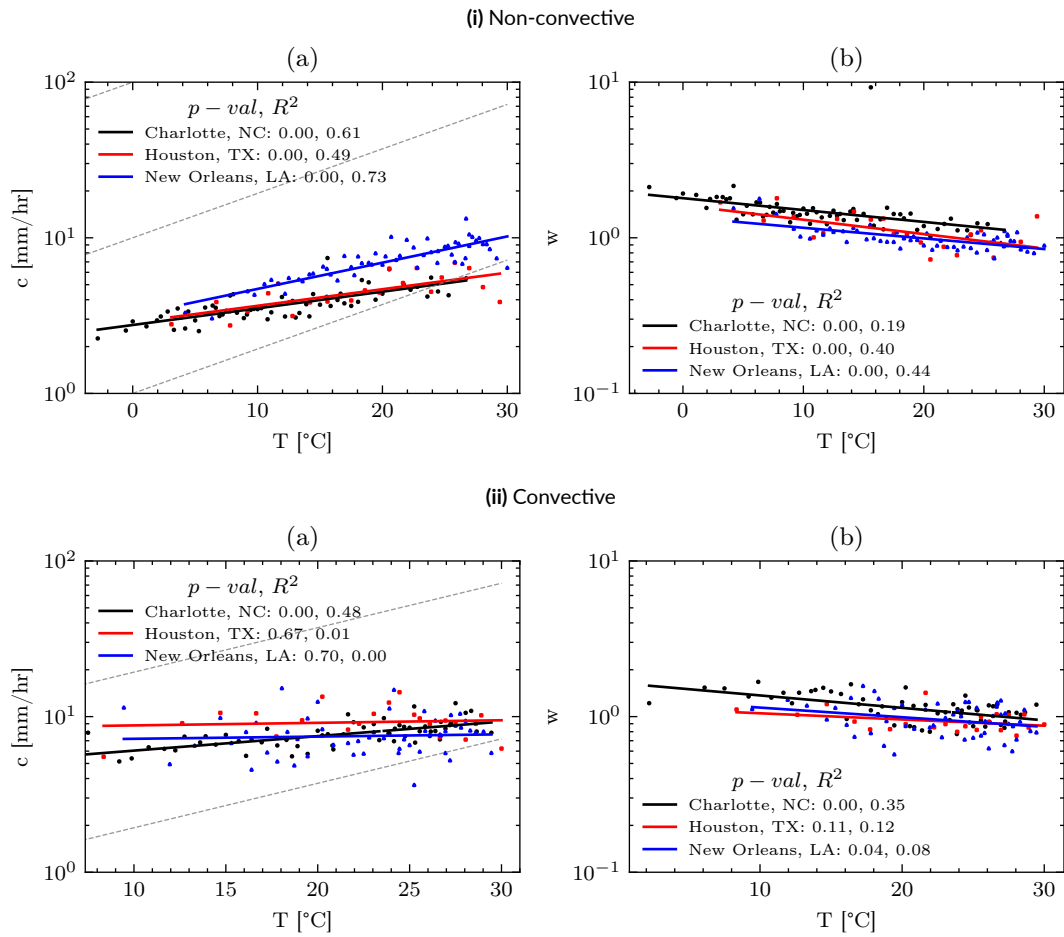
Figure 3.25: The 99th percentiles of hourly precipitation intensity as a function of daily mean temperature for each precipitation type in some stations in the U.S

4

Discussion

In this thesis, I have investigated the behavior of daily and hourly precipitation extremes under a warming climate, focusing on the scaling relationships between the parameters of the probability distribution of precipitation and temperature. The analyses have been conducted across various locations representing different climatic conditions, including, Padova (Italy), Oxford (U.K.) as well as multiple stations in the Contiguous United States.

4.1 Daily Precipitation Extremes

Regarding the scaling of daily precipitation extremes, my findings indicate that the conventional Clausius-Clapeyron (CC) scaling does not universally apply. In my observations, daily precipitation extremes show deviations from the CC-scaling relation in Padova and Oxford.

In particular, a decreasing trend in extreme percentiles with increasing temperatures was observed in Padova, specifically during the summer months. This pattern is consistent with the characteristics of convective summer precipitation as described by [41]. The decrease in extreme values is linked to limited available moisture for precipitation, which necessitates air parcels to ascend to greater altitudes to reach the condensation level. As a result, using surface temperature as a proxy for condensation temperature tends to overestimate the actual condensation temperature, contributing to the observed negative slope as highlighted by Drobinski et al. [41].

Furthermore, my analysis revealed that NDJ precipitation, which often originates from large-scale frontal systems, shows scaling patterns that align more closely with CC-scaling, especially in Padova.

In terms of the relationship between the parameters of ordinary event distribution and temperature T, I observed a significant negative correlation between the scale parameter c and T in the context of summer precipitation. This suggests that the scale parameter decreases as temperatures rise. This relationship will be explored further in Section 4.3. Winter precipitation, on the other hand, showed a correlation that is less conclusive due to high variability in c. According to Marani and Zanetti [30], the majority of NDJ rainfall events in Padova stem from Atlantic storms, which are significantly influenced by the phase of the North Atlantic Oscillation. These dynamic influences could potentially account for a substantial fraction of the variability in the scale parameter. However, further investigation is required.

Summer precipitation in Oxford showed a less pronounced negative correlation between the scale parameter c and temperature T, and a slight positive one in winter. However, the results were inconclusive due to the high variability in the data.

I also examined the time-variation of the relationship between the ordinary event distri-

bution parameters and temperature. The results indicate the existence of an exponential relationship given in Equation 2.10 between the scale parameter c and temperature. The rate of decay/increase of this relationship is shown to be variable in time. The oscillatory nature of these variations hints at the effects of large-scale atmospheric circulation.

4.2   HOURLY PRECIPITATION EXTREMES

Regarding hourly precipitation, I found that it generally aligns with the CC-scaling, with exceptions in some locations, which is consistent with the observations made by Shaw et al. [32]. Additionally, I identified consistent exponential relationships between the scale parameter c and temperature T across most stations, highlighting the robustness of these scaling patterns. These relationships accounted for a large fraction of the variability in the scale parameter, with higher R^2 values observed in stations that align with the CC scaling relation. These findings suggest that thermodynamics predominantly influences hourly precipitation extremes, except in regions with unique characteristics.

In this context, the analysis conducted by Martinez-Villalobos and Neelin [42] becomes relevant. They asserted that the scale parameter of the Gamma distribution (which they used to model the scale-dominated range of daily precipitation from which extremes emerge) is determined by the amplitude of moisture convergence fluctuations during rainy periods. Their research suggested that this scale parameter would exhibit CC-scaling-like behavior if only the thermodynamic component of change is considered and that this scaling could be amplified or offset by local variations in convergence.

For reference, the probability density function (pdf) of the gamma distribution can be expressed as follows:

$$f_\Gamma(x) \propto x^{k-1} \exp(-\frac{x}{\theta}) \tag{4.1}$$

where θ is the scale parameter, and k is the shape parameter of the gamma distribution. Similarly, the pdf of the Weibull distribution used in this thesis can be represented as:

$$f_W(x) \propto x^{w-1} \exp(-(\frac{x}{c})^w) \tag{4.2}$$

Among the outcomes of this thesis, the role of the shape parameter (w) emerges as a point of interest. A significant negative correlation was evident in sites that align with CC relation. While the physical interpretation of w remains challenging, its importance in shaping precipitation distribution cannot be overlooked, as discussed in the following section.

Lastly, I investigated the behavior of different precipitation types under warming. No discernible trend was seen in convective precipitation. In fact, the parameters did not correlate with temperature, and intense precipitation percentiles were observed to be likely across a wide range of temperatures. It is important to note that I only analyzed three stations and the classification method I employed might be inaccurate. Therefore, further investigation is necessary to ascertain whether this observed behavior is due to systematic errors or if it reflects a general characteristic of convection.

4.3 INTERPRETING CHANGES IN ORDINARY EVENT DISTRIBUTION PARAMETERS WITH TEMPERATURE AND THEIR ROLE IN SHAPING THE EXTREMES

The objective of this section is to understand the implications of the relationship between c, w, and temperature on ordinary event distribution and subsequently on the precipitation extremes. Considering the case of hourly precipitation first, let us focus on the temperature range between 5°C to 25°C and use the regression model for Des Moines, which exhibited the highest value among the U.S. stations, as the basis for our analysis.

Initially, when w fixed is kept constant (the average w value is used), as depicted in Fig-

ure 4.1a, it becomes evident that an increase in the scale parameter due to warming leads to a stretching of the probability distribution of precipitation. This means that intense hourly precipitation events become more frequent as temperature increases.

Now, considering the marginal role of w by keeping c constant (Figure 4.1b), the effect of the decrease in shape parameter becomes very pronounced for extremely rare precipitation ($F > 0.99$). This signifies an increase in the occurrence of rare precipitation events as temperature rises and w decreases.

The role of w becomes most apparent when we consider both c and w simultaneously. The anti-synergistic effect between c and w amplifies the previously observed dynamics.

Therefore, a significant negative correlation between w and T along with a positive one for c with T can cause an intensification of precipitation of given frequency and increased frequency of intense precipitation events (depicted in Figure 4.1c), corroborating with observations made by previous research [3, 1, 5, 4].

In the context of daily precipitation, let us consider the case of JJA precipitation in Padova. A significant negative correlation between c and T was observed (contrary to the case in Figure 4.1a), implying that under the assumption of constant w, intense daily precipitation events are projected to become less frequent. Although no significant temperature dependence of the shape parameter in the daily scale was observed, its variability still plays an important role. Figure 4.2 shows a histogram of the w values computed over the whole analysis period for both JJA and NDJ.

To illustrate a specific point, let us focus on the values of w that deviate by one standard deviation from the mean ($w_{left} = 0.87$ and $w_{right} = 1.41$). This survival function of ordinary JJA precipitation in Padova associated with each w value is shown in Figure 4.3.

Notice that the frequency of precipitation of a given intensity significantly changes due to the variability in w. In essence, while a decrease in the frequency of intense daily precipitation
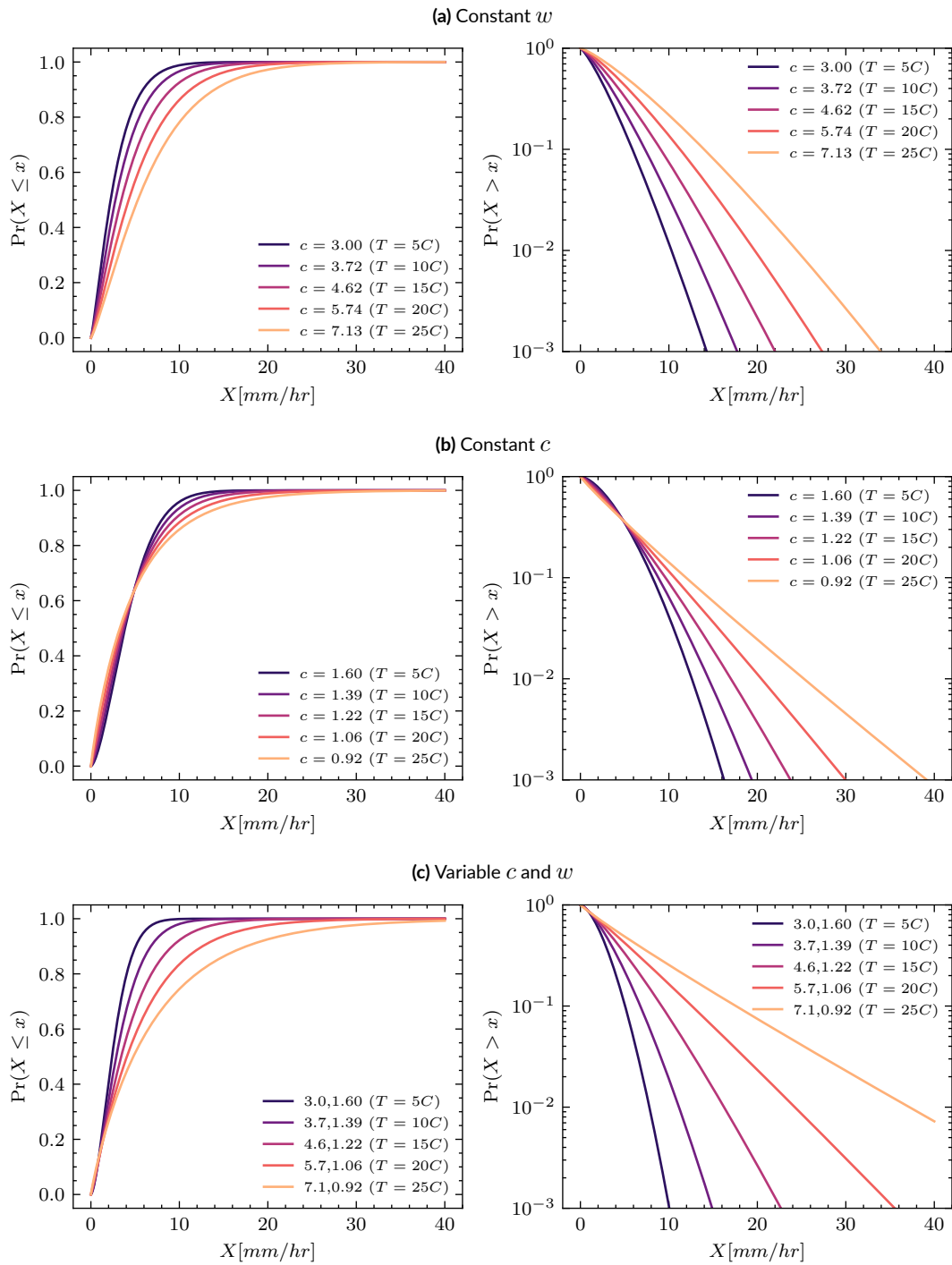
Figure 4.1: Effect of increasing T on hourly precipitation distribution
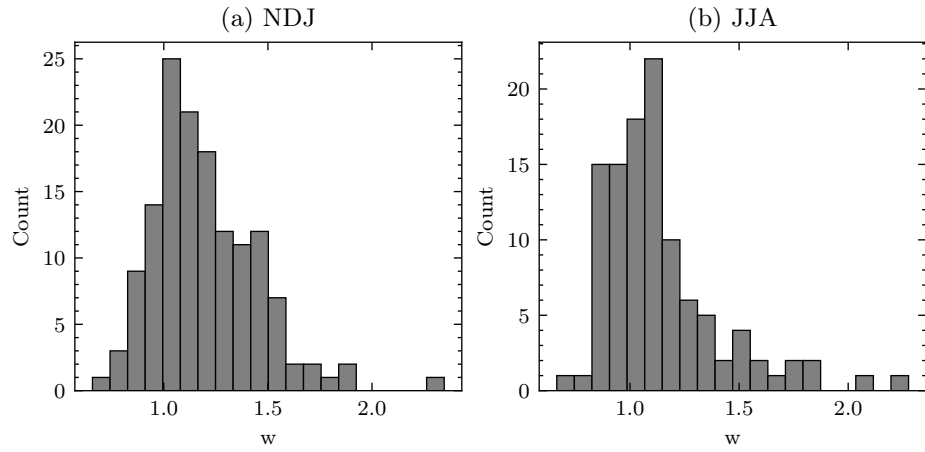
Figure 4.2: Histogram of the shape parameter w computed between 1774 - 2022 in Padova for NDJ and JJA precipitation

is projected if w were ignored, the introduction of w introduces substantial uncertainty into the prediction of extremes.

4.4 MEV P(T)

I proposed a model for rainfall extremes based on the MEV distribution, explicitly incorporating the temperature dependence of ordinary rainfall. Additionally, I proposed a straightforward quantile estimation method through simulation which effectively reduced uncer-
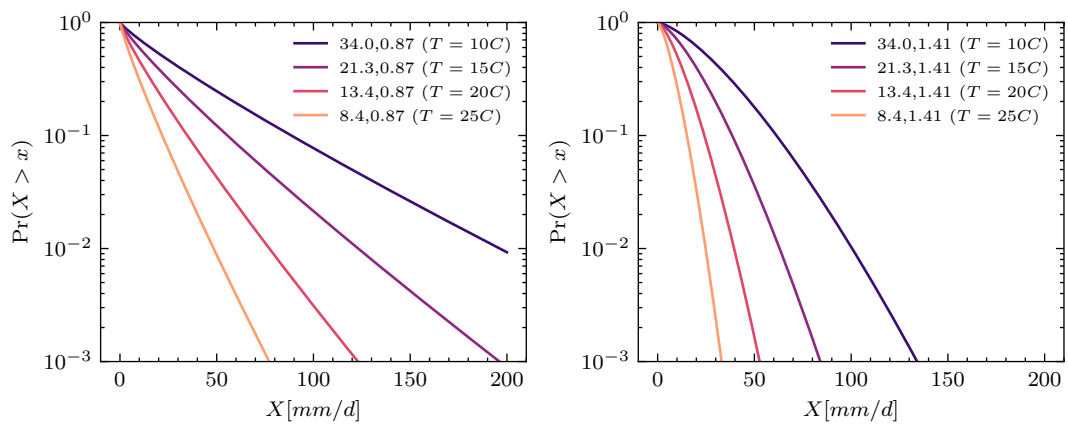
Figure 4.3: Effect of increasing T on daily JJA precipitation distribution in Padova

tainty, as anticipated from the MEV [28, 37, 35, 43], providing an alternative to the conventional cdf inversion.

Furthermore, my results indicated that the MEVPT model slightly outperforms the standard MEV, particularly in estimating quantiles for high return periods. Whether the difference in performance between the two is significant needs separate investigation. This potential advantage might come from the MEVPT's use of additional temperature information in conjunction with precipitation for quantile estimation.

Using this model, I also conducted predictive analyses wherein data from the pre-1900 period were used to calibrate the distribution of precipitation conditional on temperature. Subsequently, temperature and the number of events per year data from the period after the year 1900 were used to simulate climate change. The predictions successfully captured trends in extreme precipitation for Oxford and NDJ precipitation in Padova. However, they showed limitations in capturing oscillations. In JJA precipitation in Padova, contrary to observations, a decrease was predicted by the model aligning with the JJA precipitation scaling outcomes. This reinforces the assertion that while thermodynamic factors might explain a portion of changes in precipitation extremes, dynamic factors dominate daily-scale extremes.

5

Conclusion

The use of the Metastatistical Extreme Value (MEV) framework offered a novel way to incorporate temperature dependence into extreme precipitation analysis, providing basis for a probabilistic investigation of how extreme precipitation behaves under warming. The extent to which thermodynamics explains the changes in daily and hourly precipitation extremes was examined by focusing on the relations between probability distribution parameters of ordinary events and temperature across various locations with different climates. The MEV shows promise in reducing the uncertainty associated with high return period quantile estimation using small samples in the context of climate change.

The analysis of daily precipitation extremes has demonstrated that attributing the dynamics of extreme daily rainfall under warming solely to thermodynamics as governed by CC-

scaling can be misleading. Important deviations from CC-scaling were, in fact, observed. The negative correlation between the scale parameter of the ordinary-event distribution and temperature, found in summer observations at the relatively large set of study sites considered here, suggests that precipitation extremes should be expected to decrease as temperature rises, provided that there is no significant change in the shape parameter. It has been shown, however, that the shape parameter could play an important role in partially compensating or offsetting this decrease. Observations, in fact, show that summer precipitation extremes have increased in the last few decades, demonstrating the influence of other factors, related to large scale atmospheric circulation. Winter precipitation, on the other hand, exhibited a more complicated behavior, associated with a high variability in the parameters, potentially due to factors like the North Atlantic Oscillation. The variability of the probabilistic models describing ordinary and extreme rainfall emerged as a source of uncertainty in projecting daily extreme precipitation changes. The time-dependent correlation between precipitation distribution parameters and temperature suggests that large-scale circulation, alongside thermodynamics, plays a role in daily extreme changes. This finding highlights the need for a multidimensional approach to modeling daily extreme precipitation in a changing climate.

At the hourly scale, important observations have been made by directly relating ordinary-value distributions of precipitation with temperature, as opposed to the empirical approach of linking given quantile values with temperature. The scale parameter of the ordinary event distribution indicated a consistent CC-scaling-like exponential relationship with temperature, while the shape parameter showed a decreasing trend. The combined effect of these parameters can lead to intensified and more frequent hourly extreme precipitation events. This indicates the importance of thermodynamics in controlling hourly precipitation extremes. Stations that exhibited deviations from the CC-relation showed increased variability in the parameters, although this variability was less pronounced than that observed at the

daily scale. This increased variability was attributed to factors extending beyond thermodynamics, such as large-scale atmospheric circulation. Furthermore, the study of different precipitation types under warming indicated the need for additional research into the behavior of convective precipitation under warming.

In conclusion, this thesis contributes to a deeper understanding of the relationship between temperature and precipitation extremes. The insights gained have implications for understanding the main drivers of extreme precipitation changes at the daily and hourly scale, and for improving modeling and prediction of extreme precipitation events in a changing climate. As climate change continues, these findings will provide a useful contribution to inform adaptation strategies and policy decisions to mitigate the impacts of increasing precipitation extremes. Further research is needed to include dynamical effects to enhance extreme rainfall projection, ultimately advancing our ability to address the challenges posed by a warming world.

References

[1] E. M. Fischer and R. Knutti, "Observed heavy precipitation increase confirms theory and early models," *Nature Climate Change*, vol. 6, no. 11, pp. 986–991, Nov. 2016.

[2] V. V. Kharin, F. W. Zwiers, X. Zhang, and M. Wehner, "Changes in temperature and precipitation extremes in the CMIP5 ensemble," *Climatic Change*, vol. 119, no. 2, pp. 345–357, Jul. 2013.

[3] C. Wasko, R. Nathan, L. Stein, and D. O'Shea, "Evidence of shorter more extreme rainfalls and increased flood variability under climate change," *Journal of Hydrology*, vol. 603, p. 126994, Dec. 2021.

[4] G. Myhre, K. Alterskjær, C. W. Stjern, Ø. Hodnebrog, L. Marelle, B. H. Samset, J. Sillmann, N. Schaller, E. Fischer, M. Schulz, and A. Stohl, "Frequency of extreme precipitation increases extensively with event rareness under global warming," *Scientific Reports*, vol. 9, no. 1, p. 16063, Nov. 2019.

[5] S. M. Papalexiou and A. Montanari, "Global and Regional Increase of Precipitation Extremes Under Global Warming," *Water Resources Research*, vol. 55, no. 6, pp. 4901–4914, 2019.

[6] P. Berg, C. Moseley, and J. O. Haerter, "Strong increase in convective precipitation in response to higher temperatures," *Nature Geoscience*, vol. 6, no. 3, pp. 181–185, Mar. 2013.

[7] R. Hardwick Jones, S. Westra, and A. Sharma, "Observed relationships between extreme sub-daily precipitation, surface temperature, and relative humidity," *Geophysical Research Letters*, vol. 37, no. 22, 2010.

[8] G. Lenderink and E. van Meijgaard, "Increase in hourly precipitation extremes beyond expectations from temperature changes," *Nature Geoscience*, vol. 1, no. 8, pp. 511–514, Aug. 2008.

[9] E. E. Maeda, N. Utsumi, and T. Oki, "Decreasing precipitation extremes at higher temperatures in tropical regions," *Natural Hazards*, vol. 64, no. 1, pp. 935–941, Oct. 2012.

[10] V. Mishra, J. M. Wallace, and D. P. Lettenmaier, "Relationship between hourly extreme precipitation and local air temperature in the United States," *Geophysical Research Letters*, vol. 39, no. 16, 2012.

[11] J. Bao, S. C. Sherwood, L. V. Alexander, and J. P. Evans, "Future increases in extreme precipitation exceed observed scaling rates," *Nature Climate Change*, vol. 7, no. 2, pp. 128–132, Feb. 2017.

[12] L. Gimeno, R. Sorí, M. Vázquez, M. Stojanovic, I. Algarra, J. Eiras-Barca, L. Gimeno-Sotelo, and R. Nieto, "Extreme precipitation events," *WIREs Water*, vol. 9, no. 6, p. e1611, 2022.

[13] R. A. Houze Jr, "Stratiform precipitation in regions of convection: A meteorological paradox?" *Bulletin of the American Meteorological Society*, vol. 78, no. 10, pp. 2179–2196, 1997.

[14] S. Westra, H. J. Fowler, J. P. Evans, L. V. Alexander, P. Berg, F. Johnson, E. J. Kendon, G. Lenderink, and N. M. Roberts, "Future changes to the intensity and frequency of short-duration extreme rainfall," *Reviews of Geophysics*, vol. 52, no. 3, pp. 522–555, 2014.

[15] G. Penide, A. Protat, V. V. Kumar, and P. T. May, "Comparison of two convective/stratiform precipitation classification techniques: Radar reflectivity texture versus drop size distribution–based approach," *Journal of Atmospheric and Oceanic Technology*, vol. 30, no. 12, pp. 2788–2797, 2013.

[16] T. J. Ivancic and S. B. Shaw, "A U.S.-based analysis of the ability of the Clausius-Clapeyron relationship to explain changes in extreme rainfall with changing temperature," *Journal of Geophysical Research: Atmospheres*, vol. 121, no. 7, pp. 3066–3078, 2016.

[17] D. J. Kirshbaum, B. Adler, N. Kalthoff, C. Barthlott, and S. Serafin, "Moist orographic convection: Physical mechanisms and links to surface-exchange processes," *Atmosphere*, vol. 9, no. 3, p. 80, 2018.

[18] N. Siler and G. Roe, "How will orographic precipitation respond to surface warming? an idealized thermodynamic perspective," *Geophysical Research Letters*, vol. 41, no. 7, pp. 2606–2613, 2014.

[19] T. M. Pavelsky, S. Sobolowski, S. B. Kapnick, and J. B. Barnes, "Changes in orographic precipitation patterns caused by a shift from snow to rain," *Geophysical Research Letters*, vol. 39, no. 18, 2012.

[20] P. A. O'Gorman and T. Schneider, "Scaling of precipitation extremes over a wide range of climates simulated with an idealized gcm," *Journal of Climate*, vol. 22, no. 21, pp. 5676–5685, 2009.

[21] P. A. O'Gorman and T. Schneider, "The physical basis for increases in precipitation extremes in simulations of 21st-century climate change," *Proceedings of the National Academy of Sciences*, vol. 106, no. 35, pp. 14 773–14 777, Sep. 2009.

[22] C. J. Muller, P. A. O'Gorman, and L. E. Back, "Intensification of precipitation extremes with warming in a cloud-resolving model," *Journal of Climate*, no. 11, pp. 2784–2800, 2011.

[23] S. Emori and S. J. Brown, "Dynamic and thermodynamic changes in mean and extreme precipitation under changed climate," *Geophysical Research Letters*, vol. 32, no. 17, 2005.

[24] S. Pfahl, P. A. O'Gorman, and E. M. Fischer, "Understanding the regional pattern of projected future changes in extreme precipitation," *Nature Climate Change*, vol. 7, no. 6, pp. 423–427, Jun. 2017.

[25] C. Wasko, A. Sharma, and S. Westra, "Reduced spatial extent of extreme storms at higher temperatures," *Geophysical Research Letters*, vol. 43, no. 8, pp. 4026–4032, 2016.

[26] K. E. Trenberth, A. Dai, R. M. Rasmussen, and D. B. Parsons, "The Changing Character of Precipitation," *Bulletin of the American Meteorological Society*, vol. 84, no. 9, pp. 1205–1218, Sep. 2003.

[27] G. Panthou, A. Mailhot, E. Laurence, and G. Talbot, "Relationship between surface temperature and extreme rainfalls: A multi-time-scale and event-based analysis," *Journal of hydrometeorology*, vol. 15, no. 5, pp. 1999–2011, 2014.

[28] M. Marani and M. Ignaccolo, "A metastatistical approach to rainfall extremes," *Advances in Water Resources*, vol. 79, pp. 121–126, May 2015.

[29] K. Briggs and F. Ying, "How to estimate quantiles easily and reliably," *Mathematics Today*, vol. 2018, no. February, 2018.

[30] M. Marani and S. Zanetti, "Long-term oscillations in rainfall extremes in a 268 year daily time series," *Water Resources Research*, vol. 51, no. 1, pp. 639–647, 2015.

[31] S. Burt and T. Burt, *Oxford weather and climate since 1767*.   Oxford University Press, 2019.

[32] S. B. Shaw, A. A. Royem, and S. J. Riha, "The Relationship between Extreme Hourly Precipitation and Surface Temperature in Different Hydroclimatic Regions of the United States," *Journal of Hydrometeorology*, vol. 12, no. 2, pp. 319–325, Apr. 2011.

[33] C. W. Landsea and J. L. Franklin, "Atlantic hurricane database uncertainty and presentation of a new database format," *Monthly Weather Review*, vol. 141, no. 10, pp. 3576–3592, 2013.

[34] A. Miniussi, G. Villarini, and M. Marani, "Analyses through the metastatistical extreme value distribution identify contributions of tropical cyclones to rainfall extremes in the eastern united states," *Geophysical Research Letters*, vol. 47, no. 7, p. e2020GL087238, 2020.

[35] F. Marra, E. I. Nikolopoulos, E. N. Anagnostou, and E. Morin, "Metastatistical Extreme Value analysis of hourly rainfall from short records: Estimation of high quantiles and impact of measurement errors," *Advances in Water Resources*, vol. 117, pp. 27–39, Jul. 2018.

[36] S. Fukutome, M. Liniger, and M. Süveges, "Automatic threshold and run parameter selection: a climatology for extreme hourly precipitation in switzerland," *Theoretical and Applied Climatology*, vol. 120, pp. 403–416, 2015.

[37] E. Zorzetto, G. Botter, and M. Marani, "On the emergence of rainfall extremes from ordinary events," *Geophysical Research Letters*, vol. 43, no. 15, pp. 8076–8082, 2016.

[38] P. Wilson and R. Toumi, "A fundamental probability distribution for heavy rainfall," *Geophysical Research Letters*, vol. 32, no. 14, 2005.

[39] J. A. Greenwood, J. M. Landwehr, N. C. Matalas, and J. R. Wallis, "Probability weighted moments: definition and relation to parameters of several distributions expressable in inverse form," *Water resources research*, vol. 15, no. 5, pp. 1049–1054, 1979.

[40] F. J. Massey Jr, "The kolmogorov-smirnov test for goodness of fit," *Journal of the American statistical Association*, vol. 46, no. 253, pp. 68–78, 1951.

[41] P. Drobinski, B. Alonzo, S. Bastin, N. D. Silva, and C. Muller, "Scaling of precipitation extremes with temperature in the french mediterranean region: What explains the hook shape?" *Journal of Geophysical Research: Atmospheres*, vol. 121, no. 7, pp. 3100–3119, 2016.

[42] C. Martinez-Villalobos and J. D. Neelin, "Regionally high risk increase for precipitation extreme events under global warming," *Scientific Reports*, vol. 13, no. 1, p. 5579, 2023.

[43] A. Miniussi and M. Marani, "Estimation of Daily Rainfall Extremes Through the Metastatistical Extreme Value Distribution: Uncertainty Minimization and Implications for Trend Detection," *Water Resources Research*, vol. 56, no. 7, p. e2019WR026535, 2020.

Acknowledgments

I want to express my sincere gratitude to my thesis supervisor, Prof. Marani, and my co-supervisor, Prof. Borga, for their guidance and invaluable feedback throughout all the stages of writing this thesis. Their mentorship has been crucial in shaping this work.

I extend my heartfelt appreciation to the esteemed professors of the WGRE department for sharing their knowledge, which has greatly enriched this thesis.

A special thanks goes to Dr. Maria Francesca for all the resources and support she provided.

My sincere appreciation also goes to my friends. Their immeasurable support and encouragement have been an invaluable source of motivation.

Finally, I am profoundly grateful to my family - my parents and my sister - for their constant love and support throughout. I would not have reached this milestone if it were not for their dedication.