UNIVERSITÀ
DEGLI STUDI
DI PADOVA


Università degli Studi di Padova

Dipartimento di Ingegneria Civile, Edile e Ambientale

Corso di Dottorato di Ricerca in
Scienze dell'Ingegneria Civile, Ambientale e dell'Architettura
XXXV°ciclo


# Extremes in the hydrological cycle: A metastatistical framework applied to extreme coastal flooding and droughts


**Coordinatore:** Ch.mo Prof. Carmelo Majorana
**Supervisore:** Ch.mo Prof. Marco Marani

**Dottoranda:** Maria Francesca Caruso



Anno Accademico 2022/2023

**Abstract**

The study of extremes, maxima and minima of random variables, has a long history in applied hydrology for engineering purposes. The magnitude and the frequency of extreme events are described by the right-most part of a probability distribution, usually referred as "tail". Therefore, the task for hydrologists is to extract as much information as possible from the data set to assess the tail behavior correctly and reduce the uncertainty in the estimates.

Traditionally, the most widely used probabilistic methods are based on the asymptotic results of the Extreme Value (EV) theory, which is commonly applied using block maxima or via peak-over-threshold analysis. However, despite its theoretical basis, a number of scientific contributions highlight the limits of these traditional approaches. In particular, classical EV methods are considered to waste much of the available observational information, which is reflected into increased estimation uncertainty for quantiles that are large with respect to the observed largest values. This notion is leading to the development of alternative modeling approaches that make a better use of the observations. Among these methods, non-asymptotic models, i.e. statistical models which do not assume the block maximum to arise from a large number of ordinary values, promise to lead to more robust estimates of high quantiles. In particular, this dissertation estimates the probability of extremely large events using a non-asymptotic approach based on the Metastatistical Extreme Value Distribution, MEVD, and its simplified versions, including SMEV, or Simplified Metastatistical Extreme Value distribution. A comparative assessment of the predictive performance of the proposed non-asymptotic model and conventional approaches based on the three-parameter Generalized Extreme Value distribution, GEV, is the focus here. The present dissertation, therefore, investigates the potential of MEVD-based approaches to characterize the probabilistic structure of the tail distribution that governs opposing phenomena such as coastal flooding and drought occurrence. In fact, the proposed "general" model (i.e., fewer a-priori assumptions on the properties of the event occurrence process and efficient use of the data) yields a reduction of estimation uncertainty in the quantification of extremely rare quantiles. Even though the studied extreme events are different from a process perspective, the results confirm the advantages and flexibility of these novel extreme value distributions.

# Contents

# List of Figures

# List of Tables

# Chapter 1

# Introduction

## 1.1 Background

Extreme values of a large variety of physical parameters describing the state of the Earth System (simply referred to as "extremes" here) have long attracted the interest and attention of scientists, policymakers, and stakeholders due to their impacts on society and ecosystems. Indeed, modeling, and quantifying, extreme-value probability of occurrence is an active field of theoretical and applied research in many disciplines, such as hydrology and climatology (Katz et al., 2002; Cancelliere, 2017; Mekonnen et al., 2021; Miniussi and Marra, 2021; Dallan et al., 2022a; Palazzolo et al., 2022), ecology (Katz et al., 2005; Rypkema et al., 2019), ocean wave modeling (Rueda et al., 2016; Benetazzo et al., 2017; Barbariol et al., 2019), transport engineering (Songchitruksa and Tarko, 2006), geophysical processes (Pisarenko et al., 2014a,b; Elvidge and Angling, 2018; Hosseini et al., 2020), biomedical data analysis (De Zea Bermudez and Mendes, 2012; Chiu et al., 2018), insurance and financial applications (Embrechts et al., 1997; Chan et al., 2022), and many others.

In addition to natural variability and fluctuations internal to the climate system, anthropogenic forcing (e.g., greenhouse gas emissions and direct impact of human activities on the water cycle through water abstraction, irrigation, and land use change; Figure 1.1) is regarded as a driver of possible abrupt changes in the global water cycle (Williams et al., 2007; Allan et al., 2020; Zhang and Gillies, 2022). In fact an increasingly wide literature (see e.g., chapter 8 of the Sixth Assessment Report of the IPCC, AR 6; Douville et al., 2022) indicates that the hydrological cycle appears to have intensified, along with the frequency and duration of extreme events (e.g., heavy precipitation, flood and drought conditions, Loaiciga et al., 1996; Huntington, 2010). As indicated in Trenberth (1998), the term intensification of the hydrological cycle is used to describe an acceleration in the rate of change in atmospheric water vapor content, through changed precipitation, evaporation, and evapotranspiration. Further, the relationship between atmospheric warming and the intensification of the hydrological cycle

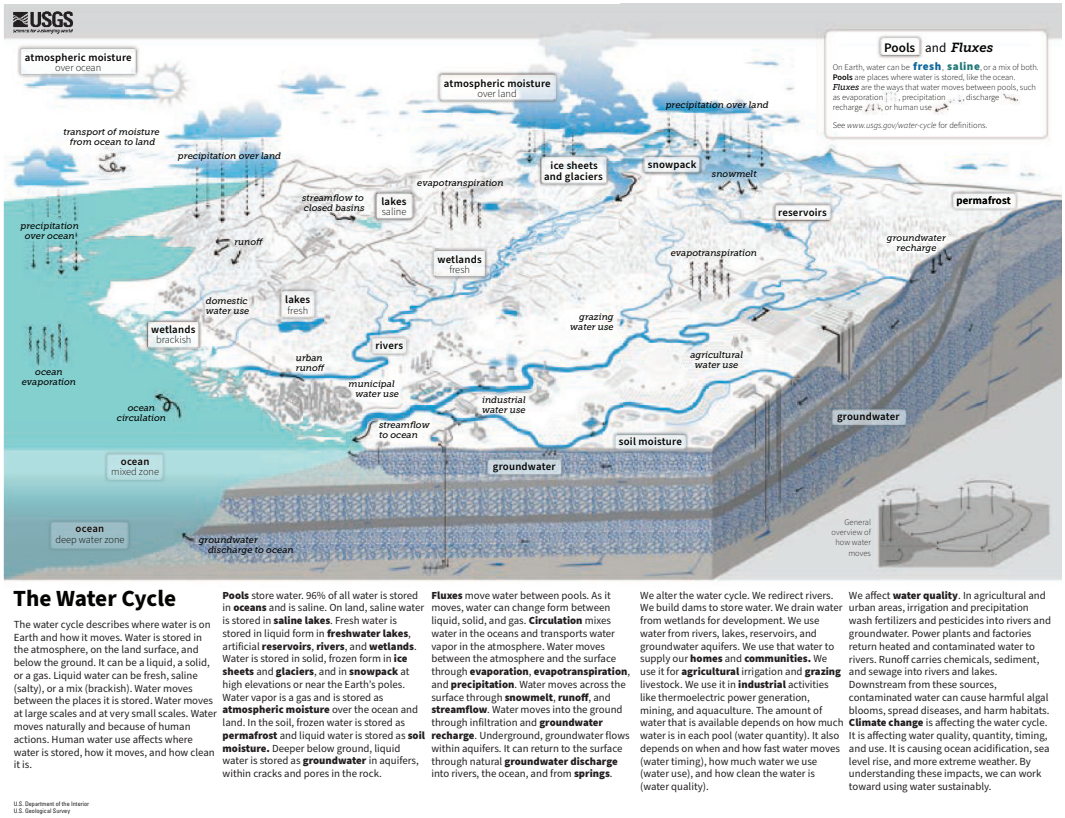


**Figure 1.1:** Global water cycle with human influence. Image credit: USGSC - United States Geological Survey (2022).

.

is physically consistent with the thermodynamic constraint imposed by the Clausius–Clapeyron (CC) relation, which defines the increase in the saturated atmospheric vapor content with increasing temperature. Changes in atmospheric vapor dynamics from the local to the global scale can weaken or reinforce the thermodynamic effect, thereby leading to non-trivial changes in the hydrologic fluxes and in the frequency and intensity of the extreme events over the globe.

Despite the growing number of studies, observing, modeling, and understanding extremes in a changing climate remains a challenge. Improving our understanding of extreme phenomena and our ability to quantify their impacts indeed remains a key societal need to develop adaptation strategies and support risk management approaches. The definition of an "extreme" varies widely across different fields, depending on applications and purposes. McPhillips et al. (2018) reviewed 10 years of academic literature engaged in research and management of extreme events to elucidate how six disciplines (i.e., climatology, earth sciences, ecology, engineering, hydrology, and social sciences are considered in this review) define and communicate such events, in terms of their causes, attribution, and impacts. Frequently, extremes are drivers of natural disasters causing loss of life, infrastructure damage, and economic hardship (Rashid and Wahl,

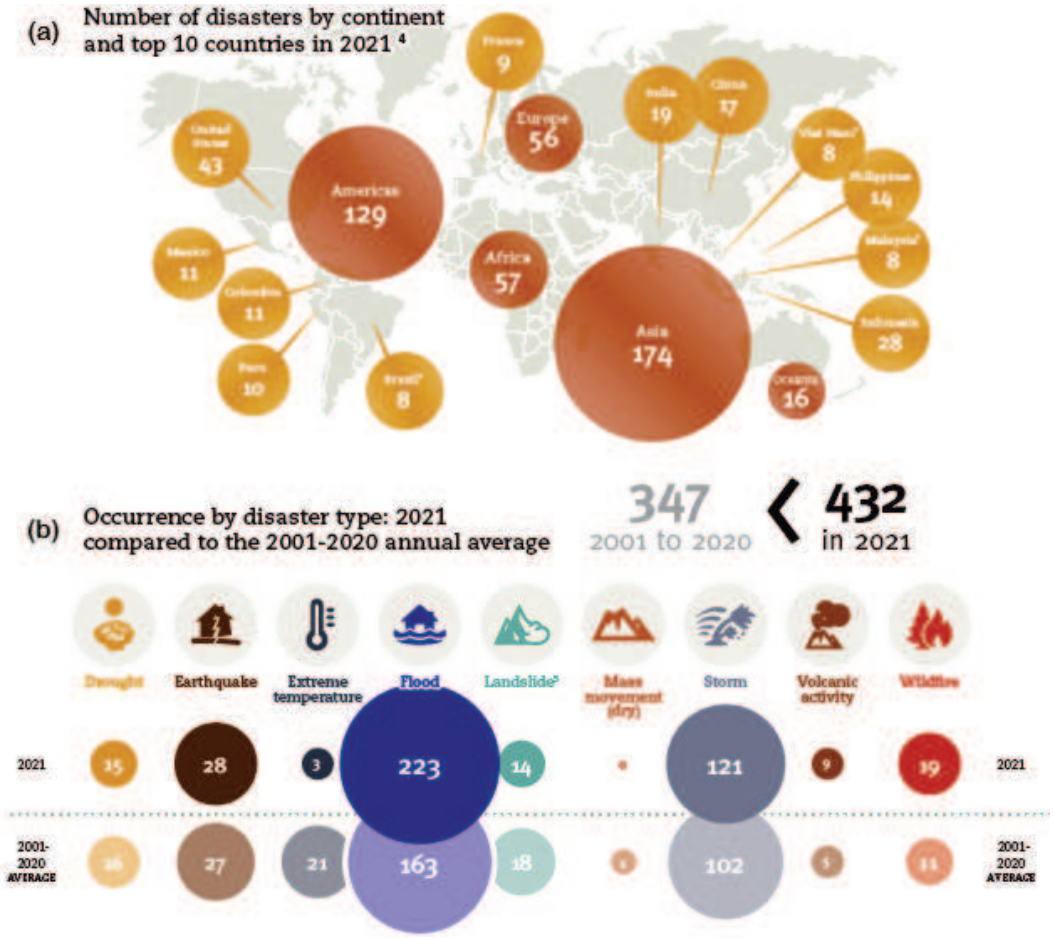


**Figure 1.2:** Occurrence of disasters. Image credit: CRED - Centre for Research on the Epidemiology of Disasters (2022).

2022). As an example, in 2021 the Emergency Event Database (EMDAT; CRED - Centre for Research on the Epidemiology of Disasters, 2022) recorded 432 disastrous events related to natural hazards worldwide (Figure 1.2b). Overall, these accounted for 10,492 deaths, affected 101.8 million people and caused approximately 252.1 billion US$ of economic losses. Floods and storms – with 223 and 121 occurrences, respectively – are the most frequent and impactful events around the world. In contrast to floods and storms, worldwide 15 droughts were recorded in 2021; Africa (South Africa, Somalia, Ethiopia, Kenya) and Asia (Afghanistan, Iraq, Syria, Iran) were most heavily impacted in terms of the number of people affected. Among continents, Asia was the most severely impacted (Figure 1.2a), and globally, the EM-DAT marked 2021 with an increase in the number of disaster events and extensive economic losses.

This general overview highlights the importance of understanding, modeling and quantifying the occurrence of extremely large events. Motivated by this awareness, this thesis represents a contribution to the statistical description of extremes using novel approaches that optimize the use of observational information and allow the estimation

of large extremes on the basis of small samples.

## 1.2 State of the art

In the context of hydrological applications, the statistical properties of the largest or smallest event within a sequence have long attracted the attention of researchers dealing with floods and droughts. Hydraulic structures are built to resist to a "design event", generally defined by a prescribed probability of occurrence, typically measured in terms of return period (Cipollini et al., 2021; Volpi et al., 2015; Vogel and Castellarin, 2017). The return period is defined as the inverse of the exceedance probability (i.e. the complementary of the non-exceedance probability) of the random variable under consideration (e.g., storm event, wind speed, river discharge, low stream flow). Under the assumption of stationarity, the return period of a set magnitude value of the variable of interest is equal to the mean time between two successive exceedances of this value. Although the concept and application of return period are well established in the literature, reliable estimates of return period and corresponding magnitude still require the definition of a suitable probabilistic model (Bonaccorso et al., 2003; Emanuel and Jagger, 2010; Volpi et al., 2015; Marani and Ignaccolo, 2015; Zorzetto et al., 2016; Volpi et al., 2019; Miniussi and Marra, 2021; Bateni et al., 2022). Usual approaches, under the assumption of event independence and process stationarity, are based on the classical Extreme Value (EV) theory (Fréchet, 1927; Dalrymple, 1960; Coles, 2001; Woodworth and Blackman, 2002; Hamdi et al., 2014, 2015), which identifies the Generalized Extreme Value (GEV) distribution as a general model for the distribution of maxima (or minima) extracted from fixed time periods of equal length (i.e., "blocks", usually assumed to be 1 year in hydrological applications). The GEV, according to its original formulation, arises as a limiting distribution for block-maxima (or minima) of a sequence of independent and identically distributed random variables when the number of events per block is very large. GEV-based models are commonly used in practical hydrological applications and are often fitted to time series composed just of block maxima (or minima) (GEV–BM). The fitted GEV distribution is then used to estimate the probability of occurrence of events of various magnitudes. Another widely used approach is the Peaks-Over-Threshold (POT) method, which extends the original GEV formulation by modeling all events exceeding a high threshold, as opposed to considering just block maxima. The POT approach again recovers the GEV distribution as the distribution of the annual maxima if two assumptions are valid: (1) the number of events per year is Poisson-distributed; (2) exceedances over the threshold come from a generalized Pareto distribution (GPD). Several papers in hydrological literature discuss the advantages and disadvantages of methods based on block maxima or partial duration series (see e.g., Martins and Stedinger, 2001; Villarini et al., 2011; Nerantzaki and Papalexiou, 2022). Both approaches are only strictly valid asymptoti-

cally, i.e. when a large number of events ($n \to \infty$) occurs in each year. Furthermore, GEV-based approaches, by construction, discard most of the observations and do not attempt to optimize the use of the available information (Volpi et al., 2019).
Oftentimes, available data are insufficient to precisely quantify the hazard associated to a particular design event, especially when the return period of interest lies in the "extrapolation range", i.e. it is greater than the observation duration characterizing the data set. Here, the challenge lies in choosing an appropriate statistical distribution to best capture the statistical behavior of the tail that is not explicitly represented in the observational record (Stedinger et al., 1993). Given that large extreme events are, by definition, poorly represented in observational records, recent work on the statistical description of extremes has focused on methods that best exploit the available information. One contribution to overcoming these limitations of the traditional EV theory is the Metastatistical Extreme Value Distribution (MEVD), proposed by Marani and Ignaccolo (2015). The MEVD models the distribution of yearly maxima starting from the distribution of "ordinary events", i.e. all the available data, in contrast to just considering annual maxima or a few values above a threshold. Moreover, the MEVD framework (i) is a non-asymptotic extreme-value distribution, which does not require the number of events per year to be large, and (ii) makes no a-priori assumptions on the properties of the event occurrence process as in the block maxima and partial duration series approaches. In existing applications (e.g., rainfall, flood-frequency analysis, and hurricane intensities), the MEVD has been shown to significantly reduce estimation uncertainty compared to traditional approaches, especially when considering return periods greater than the sample size used for parameter estimation (see e.g., Zorzetto et al., 2016; Marra et al., 2018; Zorzetto and Marani, 2020; Miniussi and Marani, 2020). Extreme value analysis approaches based on both frameworks, the GEV distribution and the MEVD, will be developed and used in the present thesis. Hence, overviews of both theories are discussed in Chapter 2.

## 1.3 Objectives

The research community has investigated the performance of the MEVD framework in (1) quantifying rainfall frequency and intensity (Marani and Ignaccolo, 2015; Zorzetto et al., 2016; Marra et al., 2018), (2) capturing the spatial variability in rainfall (Miniussi and Marra, 2021; Hu et al., 2020), (3) evaluating the accuracy in the estimates by using data set with different spatial and temporal resolutions (Marra et al., 2018, 2020; Amponsah et al., 2022; Dallan et al., 2022b), (4) frequency analysis of flood events (Miniussi et al., 2020a), (5) quantifying the probability of occurrence of extreme hurricanes (Hosseini et al., 2020). On the other hand, phenomena for which no knowledge or experience exists regarding the applicability and performance of MEVD include coastal flooding and drought occurrence. Notoriously, these opposing phenomena often recur in

environmental hazard evaluation, decision-making, and the design of mitigation measures. In particular, in the context of global warming, these extremes are recurring features of a changing climate around the world. Recent history has shown the vulnerability of many areas to these rare events. For example, Europe experienced several prolonged periods of dry spells in 2015 (Ionita et al., 2017; Laaha et al., 2017), 2018 (Brunner et al., 2019; Bakke et al., 2020), and during the summer of 2022 (Commission and Centre, 2022). Regarding examples of coastal flooding, Venice is increasingly more frequently exposed to floods; the series of floods occurred in November 2019 (Ferrarin et al., 2021; Zanchettin et al., 2021), as well as the recent events in 2022 are an example and resound as an alarm.

Therefore, motivated by the above considerations, coastal flooding and drought occurrence will be the focus of this thesis. Even though these phenomena are quite different from a physical process perspective, traditionally, similar methods are used for inferring information on the tail of their distribution, such as block maxima and peak-over-threshold approaches. The use of novel extreme value distributions, namely MEVD-based approaches, are proposed here as beneficial formulations to improve the estimation of high quantiles.

## 1.4 Outline

The rest of the present thesis is structured as follows. Chapter 2 recalls the theoretical basis of the traditional EV theory and presents the recent alternative statistical models based on the MEVD approach. The latter category includes approaches with differing levels of simplification, from the full-fledged Metastatistical Extreme Value Distribution (MEVD; Marani and Ignaccolo, 2015; Zorzetto et al., 2016), to the Simplified Metastatistical Extreme Value (SMEV) distribution (SMEV; Marra et al., 2019, 2020). The problem of extreme coastal water level estimation and projection is tackled in Chapter 3. Chapter 4 focuses on the analysis of extreme drought events by using natural climatic archives based on tree-ring climate proxies. Finally, in Chapter 5 some conclusions are drawn.

# Chapter 2

# Statistics of extremes: tradition and novelty

Extreme value analysis is concerned with modeling and quantifying the occurrence probability of rare events. Events that, by definition, are poorly represented in observational time series. The theoretical basis of the theory of extreme events date back to the first half of the last century. The key result obtained by Fisher and Tippett (1928) on the limit forms of the frequency distribution of the largest or smallest member of a sample has attracted the attention of many statisticians. Only after the publication of Gumbel's book ("Statistics of Extremes"; Gumbel, 1958), the extreme value theory began to be applied to engineering fields.

Among the numerous contributions dealing with statistical aspects of extremes, the work of de Haan (1970) and subsequent publications (see e.g., Balkema and de Haan, 1974; Pickands, 1975) have shaped the main theoretical results in the traditional extreme value theory. Although the probabilistic and stochastic properties of extreme values were developed into an attractive theory, the statistical modeling of extremes still remains a subject of debate and research. The interested reader can refer to Coles (2001) for a detailed description of statistical methods for extremes in hydrology or Papalexiou and Koutsoyiannis (2013); Marani and Zorzetto (2019) for a recent overview of the history of the EV theory.

This Chapter summarizes the theoretical basis of the traditional extreme value theory and presents alternative approaches (so-called non-asymptotic formulations) to derive information from the extreme value distributions.

## 2.1 Extreme value theory

Extreme Value (EV) theory is a set of results in the field of probability that permits the study of the distribution of extreme values and to estimate their probability of occurrence over time under fairly general assumptions. It deals with the limiting

distributions for large collections of independent random variables from the same unknown underlying distribution. The foundation of the EV theory can be found in the three-types theorem, or Fisher–Tippett–Gnedenko theorem, introduced by Fisher and Tippett (1928) and later proven by Gnedenko (1943). The basic idea is to study the statistical behavior of:

$$M_n = max(X_1, \ldots, X_n) \tag{2.1}$$

where $X_i$ (for $i = 1 \ldots, n$) are a sequence of independent and identically distributed (i.i.d.) random variables having a common distribution function ($F$). The distribution function of $M_n$ is given by the $n^{th}$ power of $F$:

$$\begin{aligned} Pr\{M_n \leq x\} &= Pr\{X_1 \leq x, X_2 \leq x, \cdots, X_n \leq x\} \\ &= Pr\{X_1 \leq x\} Pr\{X_2 \leq x\} \cdots Pr\{X_n \leq x\} \\ &= F^n(x) \end{aligned} \tag{2.2}$$

By studying the asymptotic behavior of this distribution and introducing two scaling constants $a_n > 0$, $b_n \in R$, a non-degenerate distribution function, $G(x)$, is obtained:

$$P(\frac{M_n - b_n}{a_n} \leq x) = F^n(a_n x + b_n) \rightarrow G(x), \ as \ n \rightarrow \infty$$

According to the Fisher–Tippett–Gnedenko theorem, the corresponding normalized variable $M_n^* = \frac{M_n - b_n}{a_n}$ has a limiting distribution that must be one of only three types of extreme value distributions (Gumbel, Fréchet, and Weibull), characterized by different shapes of tail behavior. Von Mises (1936) proposed to combine the above three limiting distribution laws into a single family of three-parameter distributions known as the Generalized Extreme Value (GEV) distribution given by:

$$G(x; \mu, \psi, \xi) = exp\{-[1 + \frac{\xi}{\psi} \cdot (x - \mu)]\}^{-1/\xi} \tag{2.3}$$

defined on the region for which $\{x : 1 + \frac{\xi}{\psi} \cdot (x - \mu) > 0\}$, where $\mu \in (-\infty, +\infty)$, $\psi > 0$, and $\xi \in (-\infty, +\infty)$ are the location, scale, and shape parameters, respectively. Among these, the shape controls the tail decay behavior:

1. $\xi \rightarrow 0$ describes the exponential tail typical of the Gumbel family or type I;
2. $\xi > 0$ identifies the heavy-tailed case of the Fréchet family or type II;
3. $\xi < 0$ gives the short-tailed case (bounded upper tail) characterized by the reverse-Weibull case or type III.

The GEV distribution is often regarded as a general probability distribution to describe extreme values in applications to environmental and hydrology processes. Two main

approaches exist for practical extreme value analysis, the block-maxima and the peaks-over-threshold methods.

### 2.1.1 Block maxima (minima) method

Interpreting the limit in Eq. (2.2) as an approximation for large value of $n$ suggests the use of the GEV family for modeling the distribution of maxima from long sequences. This leads to the GEV-block maxima (GEV–BM) formulation for modeling extremes of a series of i.i.d. observations. It consists in dividing the observation sample into blocks of equal length and in extracting the maximum value from each block. Block maxima are then used to estimate the parameters of the GEV distribution (Eq. (2.3)). The choice of a suitable block size is may affect estimation error bias and variance. Blocks that are too small may cause a deviation from the hypothesis that the number of observations is large underlying the three-type theorem. If large blocks are chosen, the sample will contain only a few maxima, leading to a large parameter estimation uncertainty. In most environmental processes a one year block length is usually adopted, leading to the study of annual maxima time series.
The GEV–BM method is widely used for its simplicity. However, this method, by construction, discards most of the available observations, without even attempting to optimize its use (Volpi et al., 2019).
Some applications require models for extremely small values. Therefore, after appropriate changes, the GEV distribution can be also applied for minima.

### 2.1.2 Peaks-over-threshold method

To address the limitations of the block-maxima method an alternative approach is widely used in practical applications, which extends the number of observations used in fitting the GEV distribution. The Peaks-Over-Threshold (POT) method is based on the Pickands–Balkema–de Haan theorem (viz., the second theorem of EV theory; Balkema and de Haan, 1974; Pickands, 1975) (Davison and Smith, 1990). The theorem states that, for a large enough threshold value, $u$, the distribution of exceedances over the threshold, $y = X - u$ (where $X$ is a i.i.d. random variable), is described by a generalized Pareto distribution (GPD).
Suppose $X_i$ (for $i = 1, \ldots, n$) is a sequence of i.i.d. random variables whose distribution function is $F$. If $Y_i = X_i - u$ are the excesses over $u$, conditioned on $X_i > u$, the cumulative distribution of exceedances is defined by:

$$Pr\{Y_i \leq y\} = Pr\{X_i \leq u + y | X_i > u\} = F_u(y) = \frac{F(u+y) - F(u)}{1 - F(u)}$$

Pickands (1975) showed that when the threshold $u$ approaches the finite or infinity upper endpoint of the parent distribution $F$, hence exists a set of parameters $\xi$ and $\sigma_u$

(which depend on $u$) such that the GPD is a good approximation for the distribution of the excesses $F_u$. Therefore, for large enough $u$, the approximation of the form is:

$$F_u(y) \approx G(y; \sigma_u, \xi) = 1 - (1 + \frac{\xi}{\sigma_u} \cdot y)^{(-1/\xi)} \tag{2.4}$$

defined on $\{y : y > 0 \text{ and } (1 + \frac{\xi}{\sigma_u} \cdot y > 0)\}$, where $\mu$, $\sigma_u = \sigma + \xi(u - \mu)$ and $\xi$ are the location, scale, and shape parameters, respectively.
This result implies that, if block maxima have approximate distribution $G$, then threshold excesses have a corresponding approximate distribution within the generalized Pareto family. In this case, the parameters of the GPD of threshold exceedances are determined by those of the associated GEV distribution of block maxima. The duality between the GEV and GPD means that the shape parameter is dominant in determining the behavior of the GPD tail (as it is for the GEV distribution):

1. for $\xi > 0$ the distribution has no upper limit (equivalent to the Pareto distribution) and the tail distribution function satisfies $1 - H(y) \sim cy^{(-1/\xi)}$ with $c > 0$, i.e. the polynomial distribution;

2. for $\xi < 0$ the distribution of excesses has an upper endpoint at $\omega_F = \sigma_u / |\xi|$);

3. for $\xi = 0$ the distribution is unbounded. This case is interpreted as $\xi \to 0$, i.e. the exponential distribution with mean $\sigma$.

For a fixed value of $u$, the number of exceedances is assumed to be a random variable itself and is often modeled with Poisson distribution, leading to the so-called Poisson–GPD model. According to this model, if (1) the number of events per year is Poisson-distributed, and (2) the exceedances over the threshold come from the GPD distribution (with mean $\lambda$), therefore the probability that the annual maximum of the process is less than a certain value $x$ is:

$$\begin{aligned} Pr\{max_{1 \le i \le N} Y_i \le x\} &= Pr\{N = 0\} + \sum_{n=1}^{+\infty} Pr\{N = n, Y_1 \le x, \cdot, Y_n \le x\} = \\ &= e^{-\lambda} + \sum_{n=1}^{+\infty} \frac{\lambda^n \cdot e^{-\lambda}}{n} \cdot \{1 - (1 + \frac{\xi}{\sigma_u} \cdot (x - u))^{(-1/\xi)}\}^n = \\ &= exp\{-\lambda[1 + \frac{\xi}{\sigma} \cdot (x - u)]\}^{-1/\xi} \end{aligned} \tag{2.5}$$

This property shows that the probability distribution of the annual maxima of a GPD–Poisson model is a GEV distribution (see Eq. (2.3)). Therefore, the GEV and GPD models are consistent with each other if $\xi^{GEV} = \xi^{GPD}$, $\sigma = \psi + \xi(u - \mu)$ and $\lambda = [1 + \frac{\xi}{\psi} \cdot (u - \mu)]^{-1/\xi}$. Under these suitable conditions, in the present dissertation the acronym POT–GPD is used when referring to the GPD–Poisson model.

The POT–GPD framework exploits more of the available information with respect to the GEV–BM formulation (e.g., Coles, 2001; Bernardara et al., 2014). However, the choice of a suitable threshold to retain a few above-threshold events per year is a critical step, and the estimation uncertainty significantly depends on threshold selection (Önöz and Bayazit, 2001; Li et al., 2012; Solari et al., 2017). The selected threshold value implies a balance between bias and estimation error variance. In fact, too low a threshold will violate the independence hypothesis of the framework, leading to bias, while too high a threshold will retain just few values above the threshold, leading to high estimation error variance.

## 2.2 Non-asymptotic statistical framework

The need to overcome key limitations in the field of EV theory led to the development of alternative approaches that allow for a more efficient use of the available data to infer the distribution of extremes. As highlighted by Serinaldi et al. (2020), there is a renewed research interest in non-asymptotic distributions of block maxima by using the entire data set rather than only the extreme events (see e.g., Caruso and Marani, 2022; Gründemann et al., 2020; Hosseini et al., 2020; Marani and Ignaccolo, 2015; Marani and Zorzetto, 2019; Marra et al., 2018, 2019; Miniussi et al., 2020a,b; Miniussi and Marra, 2021; Volpi et al., 2019; Zorzetto et al., 2016).

Among the non-asymptotic formulations reported in the literature to characterize hydro-climatic phenomena, this section introduces distributions based on a metastatistical approach, also sometimes referred to as superstatistics (Beck and Cohen, 2003), doubly stochastic processes (Dubey, 1968, 1970) or compound distributions (Cox and Isham, 1980), depending on disciplinary context. These methods include the Metastatistical Extreme Value Distribution (Marani and Ignaccolo, 2015; Zorzetto et al., 2016) and its possible simplifications, such as the Simplified Metastatistical Extreme Value formulation (Marra et al., 2019, 2020). These approaches are based on the statistical analysis of "ordinary events": i.e. all the independent realizations of the variable of interest, as opposed to just block maxima or values above a threshold. For example, in daily rainfall estimates, the set of ordinary events is defined as any observed daily rainfall depth greater than 1 mm (see e.g., Zorzetto and Marani, 2019; Miniussi and Marani, 2020). For fluvial floods, Miniussi et al. (2020a) identify as the sample of ordinary events all the uncorrelated peak discharge values occurring in the record, irrespective of their magnitude.

### 2.2.1 Metastatistical Extreme Value Distribution

The typical EV theory derivation starts from the premise that the maximum value among $n$ realizations of a random variable ($M_n$) is distributed according to the cumu-

lative distribution function $P(M_n \leq x) = G(x) = F(x; \overrightarrow{\theta})^n$ (where, as customary, a capital letter indicates the random variable and a lower-case letter indicates a value of the random variable). This approach assumes that the $n$ values of the random variable of interest are generated by the same distribution, the "ordinary value" distribution $F(x; \overrightarrow{\theta})$, and are independent; $n$ is the number of events in a block, such that $G(x)$ is the cumulative distribution of the block maxima. The classical EV theory assumes that either the number of events per block is large (asymptotic hypothesis, leading to the GEV–BM formulation) or that the number of events per block above a high threshold is distributed according to a Poisson distribution (POT–GPD formulation). The recently-proposed Metastatistical Extreme Value Distribution (MEVD; Marani and Ignaccolo (2015); Zorzetto et al. (2016)), a non-asymptotic extreme value approach, has been shown to offer several advantages compared to the traditional EV theory. In particular, the MEVD treats as stochastic variables both the parameters ($\overrightarrow{\theta}$) of the ordinary value probability distribution and the number of events per block. As highlighted by Marani and Zorzetto (2019) and Serinaldi et al. (2020), this formulation can be considered as a compound version of the distribution of block maxima in finite-size blocks resulting from the theory of order statistics for independent random variables. Hence, the MEVD cumulative distribution of block maxima (estimated using a much larger sample than just the yearly maxima used in the GEV–BM approach) is then defined as the compound probability:

$$G(x) = \sum_{n=1}^{+\infty} \int_{\Omega_{\overrightarrow{\Theta}}} F(x; \overrightarrow{\theta})^n g(n, \overrightarrow{\theta}) d\overrightarrow{\theta} \tag{2.6}$$

where $g(n, \overrightarrow{\theta})$ is the joint probability distribution of the number of events in a block and of the parameters vector (discrete in $N$ and continuous in $\overrightarrow{\Theta}$, where the capital letters indicate the random variables), $\Omega_{\overrightarrow{\Theta}}$ is the population of all possible parameter values. For practical applications, the MEVD can be approximated by substituting the ensemble average in Eq. (2.6) with the sample average computed over all the blocks in the time series, obtaining:

$$G(x) \cong \frac{1}{M} \sum_{j=1}^{M} F(x; \overrightarrow{\theta_j})^{n_j} \tag{2.7}$$

where $M$ is the number of blocks in the historical record, $F(x; \overrightarrow{\theta_j})$ is the cumulative distribution of ordinary values in the $j^{th}$ block, and $n_j$ is the number of events in the $j^{th}$ block. The interested reader can refer to Figure 2.1 for a visual representation of the proposed non-asymptotic statistical model. A common choice for the block length is 1 year. However, when particular hydro-climatic processes are being studied – as, for example, drought events, characterized by long durations and long inter-arrival times

– the size of the block may be greater than 1 year, to sample a sufficient number of events in each block. Note that the values of the parameters $\overrightarrow{\theta_j}$ may be estimated on the Estimation Windows (EW) with length that is different from block length. For example, if the block length is 1 year, it may be advantageous to estimate parameter values on longer time slices to ensure, depending on the rate of event occurrence, a reliable estimation of the parameters. Miniussi and Marani (2020) in applications to daily rainfall extremes find that, when the number of events per year is less than 20-25, then the optimal EW length may be greater than one year. It is interesting to note that the POT approach, briefly described above, can be thought of as a particular case of MEVD. In fact, Zorzetto et al. (2016) highlight that if one assumes (i) $x$ to be the excess over a high threshold, (ii) $F(x; \overrightarrow{\theta_j})$ 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 MEVD by means of the peak-over-threshold approach.

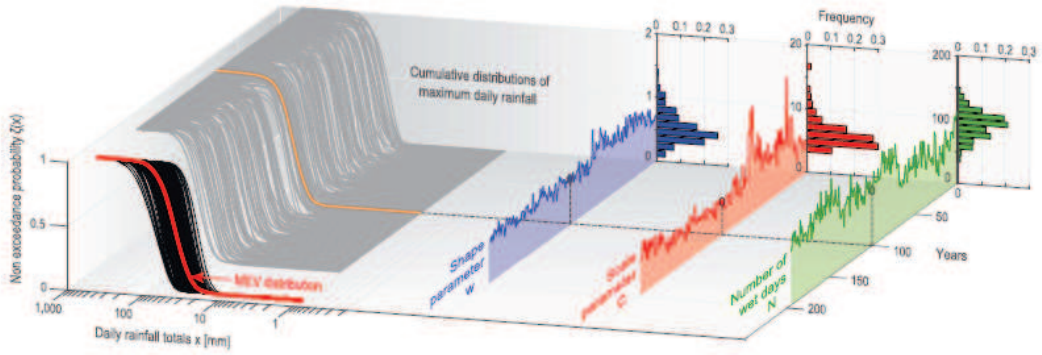


**Figure 2.1:** Visual representation of the MEVD, as introduced in Eq. 2.7, applied to a series of daily rainfall considering a block of 1 year. In this context, the two parameters Weibull distribution is used for modeling rainfall ordinary events. Yearly Weibull parameters (scale, C, and shape, w, in red and blue, respectively), and the number of ordinary events (n, in green), define the cumulative distribution function of ordinary values in each block. The MEVD (in red in the vertical xy plane in the foreground) is then computed by averaging over the empirical frequency distribution of the parameters. Image credit: Zorzetto et al. (2016).

The MEVD has been applied to several Earth-science contexts. In each application, the initial step is the identification of the form of the underlying ordinary value distribution. For example, in rainfall extremes estimates, the ordinary value distribution is assumed to be Weibull when applied to point daily rainfall (Marani and Ignaccolo, 2015; Zorzetto et al., 2016; Schellander et al., 2019; Miniussi and Marani, 2020; Miniussi et al., 2020b), point sub-daily rainfall (Marra et al., 2018), and satellite rainfall estimates (Zorzetto and Marani, 2019, 2020; Hu et al., 2020; Mekonnen et al., 2021). For fluvial floods, Miniussi et al. (2020a) propose to adopt a Gamma distribution for $F(x; \overrightarrow{\theta_j})$. Hosseini et al. (2020) describe Atlantic hurricane intensities using a generalized Pareto ordinary value distribution. Marani et al. (2021) suggest to describe the epidemic intensity by a generalized Pareto ordinary event distribution. In this dissertation, the appro-

priate form for the underlying ordinary value distribution (discussed in the following Chapters), is identified by exploring a variety of candidate distribution and by minimizing the estimation uncertainty within a cross-validation approach briefly introduced in Section 2.3.

### 2.2.2 Simplified Metastatistical Extreme Value

Marra et al. (2019) used a simplified version of the MEVD, in which the inter-annual variability, expressed by the dependence on $j$ in Eq. (2.7)), is neglected both in the ordinary event distributions (i.e. assuming a constant distribution of the ordinary events) and in the number of yearly event occurrences. Under these conditions, the Simplified Metastatistical Extreme Value (SMEV) formulation describing yearly exceedance probabilities, reads as follows:

$$G(x) \approx F(x; \overrightarrow{\theta})^n \tag{2.8}$$

where $F(x; \overrightarrow{\theta})$ is the cumulative distribution of ordinary values (described by the parameters $\overrightarrow{\theta}$), and $n$ is the average number of ordinary events per year. SMEV was originally adotped to improve parameter estimation in the presence of a small yearly number of ordinary events. In several cases this approach was applied by left-censoring the available data using a case-dependent threshold (Marra et al., 2019), which was shown to improve parameter estimation. This formulation was used to characterize daily (Schellander et al., 2019; Miniussi and Marra, 2021; Amponsah et al., 2022) and sub-daily extreme rainfall durations (Marra et al., 2018, 2020) over diverse climatic settings, including continental, temperate, Mediterranean, tropical and arid/semi-arid climates. In this context, in order to investigate the coastal and horographic effects on extreme precipitation of durations between 10 min and 24 h, Marra et al. (2022) propose a methodology for extreme precipitation frequency analysis by adjusting relatively short archives of weather radar precipitation estimates using rain gauges as reference. Araujo et al. (2022) recently propose to adopt the SMEV framework for the analysis of the current 100-year drought in Australia. More current research activity is focused on the application of MEVD and SMEV to evaluate the performance of convection-permitting climate models in representing extreme rainfall and it possible horographic controls at the hourly scale (Dallan et al., 2022b).

## 2.3 The quantification of estimation uncertainty

Statistical modeling aims to use sample information to infer the probability distribution of the population from which the data are extracted. This inference is uncertain due to imperfect parameter estimates and to the possible inability of the chosen distribution to capture the statistical properties of the underlying population. Although

these sources of uncertainty are inherent in any statistical model, their impact can be minimized by a careful choice of the model and by an effective use of all sources of information (Coles, 2001). In many applications uncertainty is estimated by means of goodness-of-fit measures, which quantify how well the distribution compares to the sample on which it was fitted. However, this procedure does not provide a measure of the predictive uncertainty encountered when trying to estimate the probability of occurrence of the "next", yet unobserved value. The quality of return periods estimates by using the non-asymptotic distributions and GEV-based approaches are here quantified by adopting a cross-validation (CV) procedure, in which model predictions of the quantiles corresponding to prescribed values of the probability of occurrence are compared to their empirical counterparts obtained from data that were not used in the estimation of model parameters. This is possible by dividing observations into two sets of independent data: the estimation set is the sample from which model parameters are estimated and the test set is the sample with which model predictions are compared.
This thesis focuses in the following on the accurate frequency analyses of two different climate-induced hydrological extremes (i.e., extreme-coastal-water-level estimates and drought extremes). Given the different nature of the processes and, therefore, of the corresponding CV procedures, the latter will be described in each of the corresponding Chapters, i.e. Chapters 3 and 4.

# Chapter 3

# Extreme-coastal-water-level estimation and projection: a comparison of statistical methods

Since millions of people live in low-lying coastal zones and many economical activities take place here, the occurrence of extreme coastal water levels and estimates of its future changes in frequency are a matter of debate in the research communities. In fact, the problem is pivotal for assessing risk and for designing coastal defense structures. In this regard, climate change and anthropogenic pressures are widely expected to exacerbate coastal hazards such as episodic coastal flooding (Almar et al., 2021).
In this context, this Chapter focuses on the accurate estimates of the probability of extreme sea levels. For practical applications, this probability is typically estimated by modeling observed sea-level records using one of a few statistical approaches. In this study, the generalized extreme value distribution, based on block maxima and peaks-over-threshold formulations, and the recent metastatistical extreme value distribution are comparatively applied to four long time series of sea-level observations distributed along European coastlines. This application studies extreme values of the coastal water level – the sum of the water level setup induced by meteorological forcing and of the astronomical tide. A cross-validation approach, dividing available data into separate calibration and test sub-samples, is used to compare the performances of the three formulations explored in high-quantiles estimation. To address the limitations posed by the length of the observational time series, the estimation uncertainty associated with different calibration sample sizes, from 5 to 30 years, is quantified. Finally, this Chapter investigates the influence of end-of-century projected mean sea levels, on the probability of occurrence of extreme total water levels (the sum of the instantaneous water level and the increasing mean sea level) frequencies.
The results were published in 2022 in the journal Natural Hazards and Earth System Sciences (see Caruso and Marani, 2022). Therefore, this Chapter reports the sections

of the article with some minor rearrangements and with the following organization. Section 3.1 presents the aim of the study. In Section 3.2 the data and methods for (a) pre-processing the time series of sea levels, (b) selecting the independent events, (c) applying the cross-validation procedure, (d) estimating the probability of future total water levels along European coats, and (e) assessing the return period, are presented here. Section 3.3 describes and discusses the results of the application. Finally, in Section 3.4 the conclusive remarks are drawn.
An overview of statistical approaches (i.e., GEV-based approaches and MEVD framework) used in this application is given in Chapter 2.

## 3.1 Introduction

The statistical analysis of extreme values of random variables is of wide conceptual and applicative importance in science and engineering (Coles, 2001; Beirlant et al., 2004; Castillo et al., 2005; Finkenstädt and Rootzén, 2004). In particular, the reliable estimation of the occurrence probability of coastal flooding events of large magnitude is crucial to environmental hazard evaluation (Coles and Tawn, 2005; Hamdi et al., 2018) and to decision-making and mitigation measure design. In fact, coastal flooding hazard has been increasing at the global scale in recent decades, a trend expected to continue as a result of climate change (Meehl et al., 2007; Church et al., 2013; Fortunato et al., 2016). Several studies highlight that global sea-level rise will continue accelerating in the 21st century as a consequence of climate change (Church and White, 2006; Jevrejeva et al., 2008; Church and White, 2011; Haigh et al., 2014b; Hay et al., 2015). Additionally, changes in storminess may have an important role in modifying the frequency and magnitude of water level extremes (Lowe et al., 2010; Menéndez and Woodworth, 2010; Woodworth et al., 2011).
Much of the current work on extreme-coastal-flooding events is based on the classical Extreme Value (EV) theory (Fréchet, 1927; Dalrymple, 1960; Coles, 2001; Woodworth and Blackman, 2002; Hamdi et al., 2014, 2015, and references therein), which identifies the family of distribution functions known as Generalized Extreme Value (GEV) distribution (Von Mises, 1936) as a general model for the distribution of maxima (or minima) extracted from fixed time periods of equal length ("blocks", most commonly with length of 1 year). The GEV, according to its original formulation, arises as a limiting distribution for maxima (or minima, not considered here) of a sequence of independent and identically distributed (i.i.d.) random variables. The Peaks-Over-Threshold (POT) formulation (Balkema and de Haan, 1974; Pickands, 1975), extends the original GEV derivation by modeling all events exceeding a high threshold, as opposed to considering just yearly maxima as in the GEV–block maxima formulation (GEV–BM). The POT approach again recovers the GEV distribution as the distribution of the annual maxima if two assumptions are valid (Davison and Smith, 1990): 1) the number of events per

year is Poisson-distributed; 2) exceedances over the threshold come from a generalized Pareto distribution (GPD). Under these suitable conditions, in the following we will refer to the POT framework as POT-GPD formulation. For a brief overview of the theory underlying EV theory and the two main methods based on the GEV distribution (i.e., BM and POT approaches), the reader can refer to Chapter 2.

For extreme-sea-level studies in particular, Coles and Tawn (2005) and Haigh et al. (2010) recognize two weaknesses in the use of the GEV–BM analysis: 1) sea level is the combination of tide-driven (deterministic) and storm-driven (stochastic) components. The presence of a deterministic component is suggested to violate the i.i.d. assumption required in the GEV–BM derivation; 2) sea-level data are collected frequently (e.g., hourly), while the GEV–BM approach only studies annual maxima, with an extremely inefficient use of the data.

More generally, GEV-based approaches, by construction, discard most of the observations and do not attempt to optimize the use of the available information (Volpi et al., 2019). Furthermore, the traditional EV theory derives the GEV distribution either as the asymptotic distribution when the number of events per block becomes very large or through the ad-hoc GPD–Poisson assumptions underlying the POT approach. Whether these hypotheses do apply to the case of sea levels is a matter of discussion, but it seems beneficial to adopt methods that require the least amount of a-priori assumptions on the properties of the event arrival process. As a contribution to overcoming the limitations of the traditional EV theory, here we explore the use of an alternative approach for modeling extreme sea levels, the Metastatistical Extreme Value Distribution (MEVD; Marani and Ignaccolo, 2015). As previously introduced in Chapter 2, the MEVD models the distribution of yearly maxima starting from the distribution of "ordinary values", i.e. all the available data, in contrast to just considering annual maxima or a few values above a threshold. Moreover, the MEVD framework (i) is a non-asymptotic extreme value distribution, which does not require the number of events per year to be large as in the traditional theory, and (ii) makes no a-priori assumptions on the properties of the event occurrence process. In previous applications, the MEVD has been shown to significantly reduce estimation uncertainty compared to traditional approaches, especially when considering return periods greater than the sample size used for parameter estimation (Zorzetto et al., 2016; Marra et al., 2018; Miniussi and Marani, 2020; Miniussi et al., 2020a,b).

Here we comparatively analyze the performance of GEV-based approaches and MEVD in high-quantile estimations with application to extreme sea levels at different observation sites. The aim is to: 1) identify the statistical tool affording minimal uncertainty in the estimate of extreme sea levels with an assigned probability of exceedance, and 2) model and understand how climate change will affect the extreme sea-level occurrence. To achieve these objectives, we analyze selected sea-level time series along the European coastline and evaluate extreme-sea-level predictive uncertainty by adopting a

cross-validation approach, in which calibration and test samples are kept separate and independent. Subsequently, we use the optimized estimation method to infer possible changes in coastal flooding hazard under Intergovernmental Panel on Climate Change (IPCC) climate change scenarios RCP4.5 and RCP8.5.

## 3.2 Materials and Methods

### 3.2.1 Data

The analyses were performed using daily and hourly sea-level records from four tide gauge stations (see Table 3.1) distributed along European coastlines: Venice (Italy), Hornbæk (Denmark), Marseille (France), and Newlyn (United Kingdom). The study sites span a variety of geographical locations, coastal morphologies and storm regimes. Venice sea-level data (maximum and minimum daily observations) were obtained from the "Centro Previsioni e Segnalazioni Maree" of the Venice Municipality (`https://www.comune.venezia.it/it/content/centro-previsioni-e-segnalazioni-maree`; Città di Venezia, 2020) for the Punta della Salute gauge station. The remaining water level data, all at the hourly scale, were downloaded from the University of Hawaii Sea Level Center (UHSLC) repository (`http://uhslc.soest.hawaii.edu/data/?rq#uh745a/`; Caldwell et al., 2015).

All sea-level data set span long observational periods: 148 years for Venice, 122 years for Hornbæk, 115 years for Marseille (ca. 19 missing years), and 102 years for Newlyn. The raw data for all stations were pre-processed to eliminate: (1) years with less than 6 months of water level observations, and (2) days with less than 24 h of data (for the case of hourly data). This process yields four quality-controlled time series that were subsequently used in the analyses (see Table 3.1). Figure 3.1 shows daily maximum sea levels at the gauge stations explored after pre-processing.

### 3.2.2 Methods

#### Mean sea level removal

The sea-level sequence is highly correlated and is generated by a non-stationary process due to long term trends in mean sea level, the deterministic tidal component, surge seasonality, and interactions between the tide and surge (Dixon and Tawn, 1999). Tide–surge interactions may change amplitude and phase of the surges, mostly in shallow estuarine areas (Johns and Ali, 1980; Bernier and Thompson, 2007; Zhang et al., 2010). Therefore, this effect needs to be taken into account when separating the surge and tide components. However, here, we do not attempt to separate these contributions; we only analyze the sum given by the combination of the water level setup, induced by meteorological forcing, and the astronomical tide. Hence, we simply study such

**Table 3.1:** Information of sea-level data used in this application.

| *Site* | *Country* | *Location (degree, min.)* | | *Period* | *Missing years (%)* | *Deleted years* | *Number of years* |
|---|---|---|---|---|---|---|---|
| | | *Lat.* | *Long.* | | | | |
| Venice | Italy | 45°25.0'N | 12°20.0'E | 1872–2019 | – | – | 148 |
| Hornbæk | Denmark | 56°06.0'N | 12°28.0'E | 1891–2012 | – | 1985 | 121 |
| Marseille | France | 43°16.7'N | 5°21.2'E | 1885–2018 | 14.2 | 1897, 1918, 1919, 1928, 1937, 1940, 1998, 2009, 2010 | 106 |
| Newlyn | United Kingdom | 50°06.1'N | 5°32.5'W | 1915–2016 | – | 1984, 2010 | 100 |

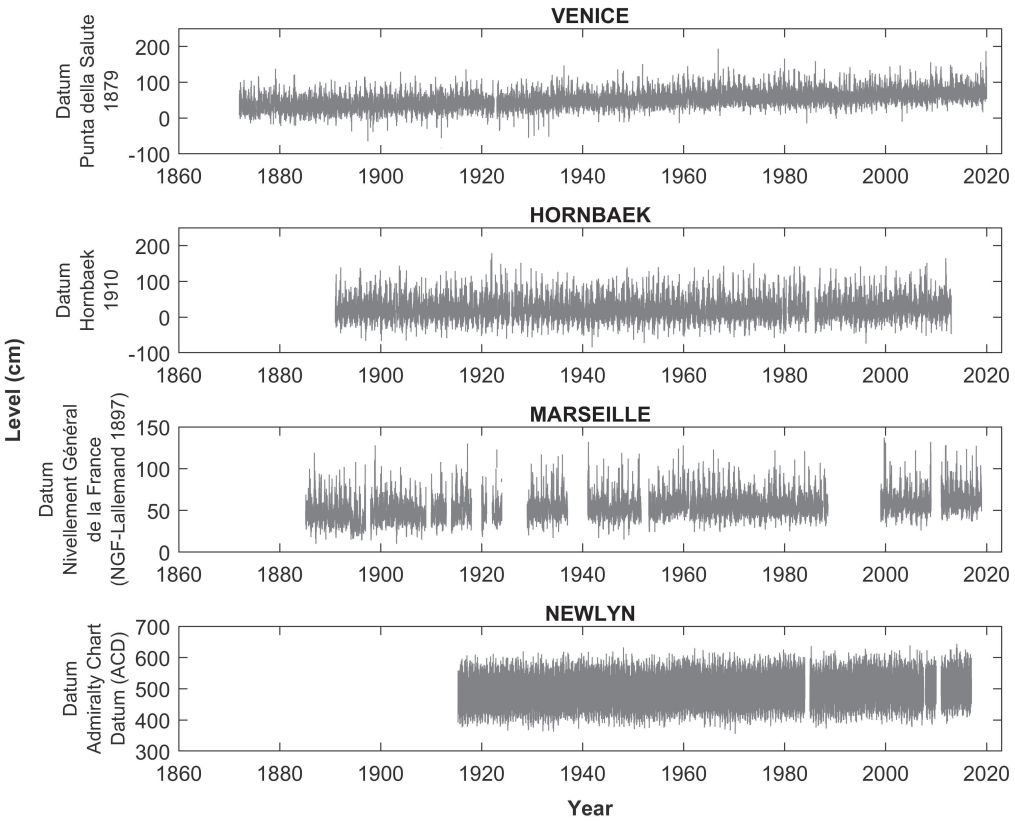


**Figure 3.1:** Daily maximum sea levels at different gauge stations explored after pre-processing: Venice (IT), Hornbæk (DK), Marseille (FR), and Newlyn (UK).

sum as the final result of the nonlinear interactions between individual components. Under this premise, for a given site and at any instant of time $t$, the observed sea level $z(t)$ (after averaging out waves), can be split into three components (Pugh and Vassie, 1979): mean sea level, $m.s.l.(t)$; astronomically induced tidal level, $x(t)$; and meteorologically induced surge level, $y(t)$:

$$z(t) = m.s.l.(t) + x(t) + y(t) \tag{3.1}$$

The term $m.s.l.(t)$ represents the long-term variations of water levels and of the elevation datum (i.e. possible land subsidence or uplift). Local $m.s.l.(t)$ does not change uniformly over time and its calculation is affected by many factors, such as tidal phases, long-term wind and atmospheric pressure patterns, vertical land motion (subsidence or uplift). The tidal contribution to the instantaneous sea level, $x(t)$, caused by the gravitational forces exerted by the moon and the sun is deterministic in nature, and can be predicted with a good degree of accuracy. This tidal variability occurs with characteristic periodicities between 12 h and 18.61 years (Eliot, 2010; Haigh et al., 2011; Pugh and Woodworth, 2014; Peng et al., 2019; Valle-Levinson et al., 2021). This latter longest tidal periodicity corresponds to the precession of the lunar nodal cycle. The storm-surge contribution, $y(t)$, is the meteorologically induced change in the water level generated by a combination of factors, such as the magnitude and direction of the wind, spatial gradients in atmospheric pressure, storm size, fetch, bathymetry, and storm duration (Hall and Sobel, 2013).

Two classes of methods are widely used to estimate the probability of occurrence of extreme sea levels: direct and indirect methods. Indirect methods model separately the deterministic and the stochastic components of $z(t)$, followed by a convolution to obtain the joint probability distribution of their sum. Examples are the joint probability method (Pugh and Vassie, 1979, 1980), the revised joint probability method (Tawn and Vassie, 1989), the exceedance probability method (Middleton and Thompson, 1986; Hamon and Middleton, 1989), and the empirical simulation technique (Scheffner et al., 1996; Goring et al., 2011). Direct methods, such as the one adopted here, analyze observed values compounding the astronomical and stochastic storm-surge component. Direct methods mostly differ based on the analysis approach adopted, such as the annual maxima method (Jenkinson, 1955; Gumbel, 1958), the peaks-over-threshold method (Davison and Smith, 1990), or the $r$-largest method (Smith, 1986; Tawn, 1988). Here, we study the distribution of the sum, $h(t)$, of the contributions from the deterministic tide and the stochastic surge:

$$h(t) = z(t) - m.s.l.(t) \tag{3.2}$$

From a statistical point of view, this choice is justified by the fact that the random arrival of storms adds a stochastic surge contribution at unpredictable times, thereby causing $h(t)$ to be values from a random variable, even though it contains a deterministic component. The presence of a deterministic component of course does imply a strong auto-correlation in the observed signal, which will be subsequently filtered out by suitable signal processing described below.

Here, $m.s.l.(t)$, is computed as the yearly average of daily levels. The yearly average is chosen rather than the customary 19-year average that eliminates all tidal periodicities, however small in amplitude, to better capture the surge contribution that causes the water level to deviate during a storm with respect to the "current" yearly value of $m.s.l.(t)$. Once $h(t)$ is computed by removing $m.s.l.(t)$ from recorded levels, all local maxima of $h(t)$, or water level peaks, are identified and their values constitute the basis for subsequent analyses of (i) long-term trends study of maximum yearly departures from the average mean sea level (two-tail Mann–Kendall test; Mann, 1945), and (ii) statistical inference of past coastal flooding events and their potential future changes.

In the following discussion, we will use the terms "total water level" and "coastal water level" when referring to the quantities $z(t)$ and $h(t)$, respectively.

### Selection of independent events

The GEV-based approaches are fit on either annual peak maxima (GEV–BM) or on a few water level peaks over a high threshold (POT–GPD), which can be assumed to be realizations of independent stochastic variables. The MEVD requires that all ordinary values (coastal-water-level peaks in this case) within one block may be assumed to be realizations from independent random variables. This hypothesis, in turn, requires that observed peaks are filtered to only retain events that may be considered to be independent, through a declustering process (Coles, 2001; Ferro and Segers, 2003; Beirlant et al., 2004; Bommier, 2014; Marra et al., 2018). Several criteria have been developed for such processing of the data. A common criterion sets the minimal time separation, or lag ($\tau$), for two events to be considered independent. Intuitively, high-water-level events separated by a sufficiently long time period are reasonably caused by distinct storm events. However, when analyzing the water level with respect to current mean sea level, a quantity that contains the deterministic tidal contribution, dependence may be expected to be present also for large lags. In theory, some dependence is present for lags up to the longest periodicity in the tidal signal (18.61 years). In practice, as the dependence in the tidal signal decreases for increasing lag, one expects that a much shorter threshold time lag will be sufficient to make sure that only independent events are considered. The analysis of the correlograms of selected coastal-water-level peaks shows that some correlation persists also for long time lags and also in the declustered time series. Even though the strength of this correlation is relatively small (the

autocorrelation function, ACF, is always less than 0.3), it could impact the ability of the MEVD, which assumes independence, to capture observed extreme behavior. The declustering process does significantly decrease correlation, as may be seen by comparing Figure 3.2 (ACF prior to declustering) and Figure 3.3 (after declustering). Interestingly, it is seen that the tidal contribution (that generates periodicities in the ACF) is strongly visible in Venice and Newlyn, while it is quite small in Hornbæk and Marseille. The underlying tidally induced correlation becomes more clearly visible after declustering also in Hornbæk and Marseille.
It is interestingly to note that the existing literature implementing declustering approaches to coastal level signals normally focuses on studying the storm-surge component only. As result, it uses threshold time lag values that are smaller than those adopted here because characteristic correlation times of the surge component are significantly smaller than those associated with the sum given by the combination of surge and tidal components. For example, the independence between two consecutive storm surge events in southern Europe has been found to be achieved with a threshold lag of 3 days (Cid et al., 2015). A threshold separation of 1 day between consecutive events is imposed by Tebaldi et al. (2012) in their analysis of storm surges along the US coast. Haigh et al. (2010) adopt a threshold lag of 30 h in the English Channel, while Bernardara et al. (2011) assume a 72 h independence criterion. After exploring values between 24 h and several days, we adopt a threshold lag of 30 days, which yielded the minimum estimation error under the MEVD approach, and is consistent with the main lunar periodicity. The result of this declustering process is a set of independent events with magnitudes $h_k$, whose number $n_j$ in year - or block - $j$ is a realization of a random variable as illustrated in Eqs. (2.6) and (2.7).

#### Cross-validation procedure

In this particular application to extreme coastal water levels, the CV procedure can be summarized as follows: (a) the observational years on record are randomly reshuffled while keeping all the water level independent peaks in their original year to (1) preserve both the ordinary value frequency distribution in each year and the distribution of the number of events per year, and (2) remove possible non-stationarity and correlation in the time series; (b) the observational sample is divided into two independent subsamples obtained by randomly selecting $S$ years from the original time series of length $M$: this sub-sample (in the following "calibration sample") is used for parameter estimation, while data in the remaining $V = M - S$ years are used for testing (in the following "validation sample" or "test sample"); (c) as usual in frequency analysis, an empirical frequency value given by Weibull's estimator $F_i = i/(V + 1)$ – where $i$ is the rank of $x_i$ in the list of yearly maxima in the validation sub-sample sorted in ascending order, and $V = M - S$ is the sample size in the validation sub-sample – is associated to each ob-

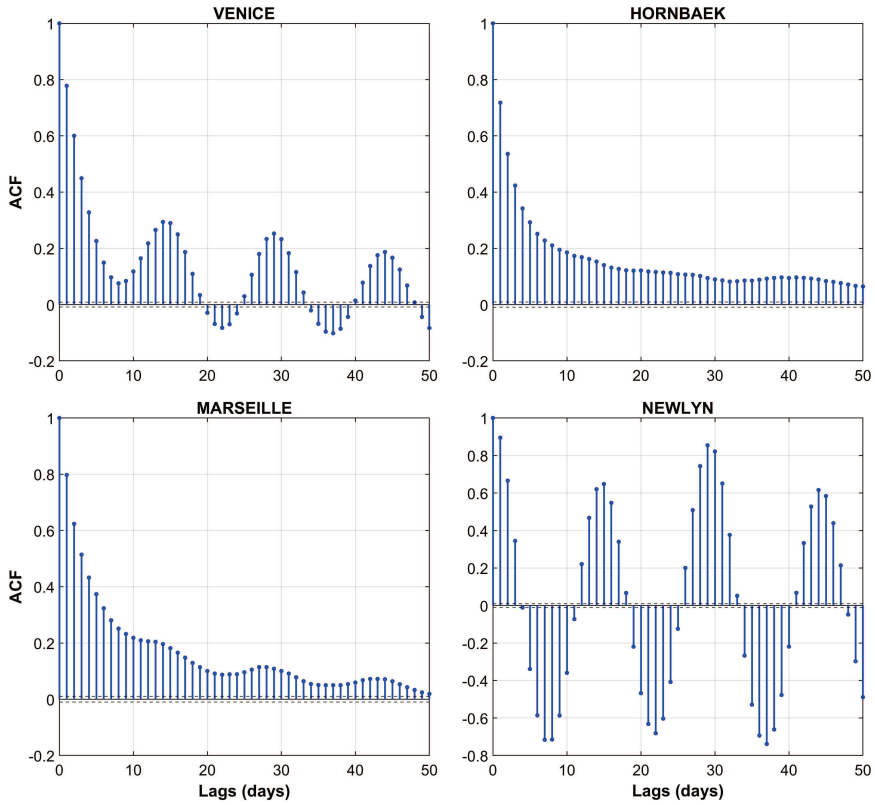


**Figure 3.2:** Autocorrelation function (ACF) prior to the declustering process.

served yearly maximum $x_i$. The return period $T_r$ associated with each yearly maximum is then simply $T_{r,i} = 1/(1-F_i)$; (d) the GEV and MEVD quantiles are computed using the parameter values estimated in step (b) from the calibration sub-sample; (e) focusing on the validation sub-sample, in every realization (for $p = 1, \ldots, N_r$; $N_r = 1000$ here) and for a fixed mean recurrence time ($T_r$), the nondimensional estimation error (NDE) – i.e., the relative error between the estimated and observed quantiles – is computed as follows: $NDE_p(S, T_r) = [h_{(est,p)}(S, T_r) - h_{(obs,p)}(S, T_r)]/h_{(obs,p)}(S, T_r)$; (f) the CV scheme introduced above is repeated $N_r$ times. This procedure is performed for different calibration sample sizes ($S$ = 5, 10, 20, and 30 years) to evaluate how estimation uncertainty varies with return period and calibration sample size.

### Future total water level projections

Future increases in the frequency of extreme total water levels (i.e. the variable previously referred as $z(t)$) due to climate change will have serious impacts on coastal regions. These impacts will vary temporally and regionally, depending on (i) the local relative mean-sea-level rise (including possible subsidence or uplift), (ii) current storm-surge intensity probability distributions, and (iii) changes in the dominant meteorological dynamics. In this particular application to extreme coastal water levels (i.e.

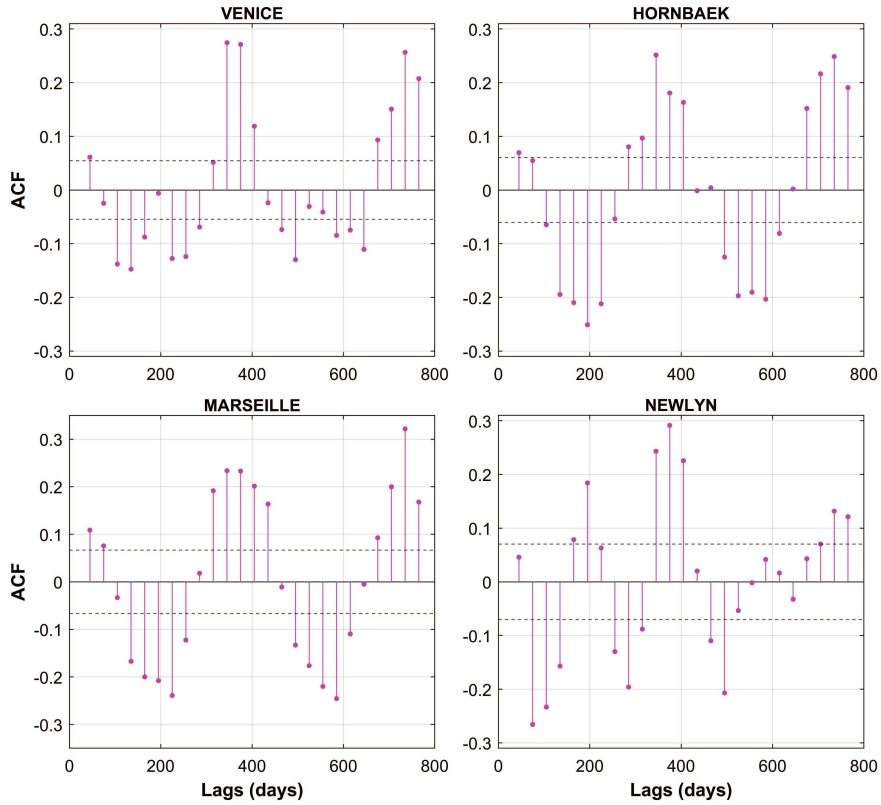


**Figure 3.3:** Autocorrelation function (ACF) for the declustered time series.

the sum given by the combination of the water level setup, induced by meteorological forcing, and the astronomical tide), only the first two factors are considered.

It is very likely that sea-level rise will continue to accelerate over time, thereby increasing the frequency of extreme sea level events, leading to severe flooding in many low-lying coastal cities and small islands (Oppenheimer et al., 2019). Various techniques have been used to study possible changes in coastal flooding hazard (e.g., McInnes et al., 2013; Vousdoukas et al., 2016). Several authors have found that past variations in the frequency of occurrence of extreme sea levels have been primarily determined by changes in mean sea level (e.g., Zhang et al., 2000; Woodworth and Blackman, 2004; Lowe et al., 2010; Menéndez and Woodworth, 2010; Haigh et al., 2014b; Wahl et al., 2017). This implies that effects of variations in storminess (e.g., magnitude, trajectories, and frequency) have been small in the observational record, compared to the dominant effects of mean sea-level changes (Haigh et al., 2014a). This notion is also confirmed by the trend analyses of maximum yearly departures from the average sea level (see Section 3.2.2), which fail to detect trends in the maximum difference between total sea level and concurrent mean sea level except at one of the sites (Venice), where it is smaller (0.7 mm $yr^{-1}$) than past and projected rates of sea-level rise ($\sim$3.0 mm $yr^{-1}$ and $\sim$8.0 mm $yr^{-1}$, respectively, by the end of the century, according to the RCP8.5 IPCC scenario).

Based on these elements, the probability of future total water levels along European coastlines is here estimated by assuming that changes in the tidal and storm-surge components are negligible with respect to changes in mean sea level, an assumption common to previous approaches (Araújo and Pugh, 2008; Haigh et al., 2010; Tebaldi et al., 2012). To assess how the exceedance probabilities of extreme total water levels might change in the future, the projections of sea-level rise through 2100 from the IPCC's Fifth Assessment Report (AR5) are used. In particular, an intermediate (RCP4.5) and an extreme scenario (RCP8.5) are explored, using CMIP5 model outputs from the "Integrated Climate Data Center" (ICDC) database (University of Hamburg: `https://icdc.cen.uni-hamburg.de/en/ar5-slr.html`; Church et al., 2013).
For each tide gauge, the approach can be summarized as follows: (1) the probability distribution of extreme coastal water levels (annual maxima extracted from observed independent events whose intensity – i.e., maximum coastal water level attained, $h_k$ – is defined with respect to the concurrent mean sea level computed on a yearly basis) is inferred; (2) the future probability of extreme total water levels is estimated by translating extreme-level quantile estimates upward according to location-specific projections of mean sea level in the year 2100 (thereby implicitly assuming subsidence/uplift to be negligible).

#### Return period and sensitivity measure

One of the main objectives of frequency analysis is to calculate the average recurrence interval or return period. It is a widely used concept in hydrological and geophysical risk analysis. If a process is stationary, the return period ($T_r$) of an event magnitude is defined as the average time elapsing between two consecutive exceedances of this magnitude. Alternatively, it may be said that a magnitude value is expected to be exceeded, on average, in each return period. If the yearly maximum magnitude $h$ is exceeded on average once in $T_r$ years, then its exceedance probability, $E(h) = 1-G(h)$, in a given year is:

$$E(h) = P[H \geq h] = \frac{1}{T_r(h)} \tag{3.3}$$

Therefore, the return period of the level value $h$ is the inverse of the probability of exceedance and can be expressed as a function of the cumulative distribution, $G(h)$, of annual maxima, e.g., through the MEVD introduced in Eq. (2.7):

$$T_r(h) = \frac{1}{E(h)} = \frac{1}{1 - G(h)} \tag{3.4}$$

Because for a fixed value of mean sea level there is a one-to-one relation between the value of the sum of the astronomical and the storm surge contribution, $h$, and the total water level, $z = h + m.s.l.$, one can write $G_h(h) = P[H > h] = P[H > z - m.s.l.] = P[Z - m.s.l. > z - m.s.l.] = P[Z > z] = G_z(h)$, such that Eq. (3.4) can be used, once

the cumulative distribution is known and for each (time-dependent) value of $m.s.l.$, to determine the return period of the total water level (at the time when $m.s.l.$ is evaluated):

$$T_r(z) = \frac{1}{1 - G_z(h)} = \frac{1}{1 - G_h(h)} = \frac{1}{1 - G(z - m.s.l.)}. \quad (3.5)$$

Based on the hypothesis introduced in Section 3.2.2 that mean sea-level rise is the dominant effect in future coastal flooding, the characteristics of the extremes (i.e. the parameters of the GPDs defining the MEVD) remain valid in future scenarios. Eq. (3.5) clarifies that the return period of a fixed value $z$ decreases as $m.s.l.$ increases, basically because for higher values of $m.s.l.$ a smaller value of $h$ is needed to achieve the same total water level $z$. This decrease is nonlinear, due to the nonlinear form of the right-hand side in Eq. (3.5).

Changes in sea-level extremes can also be studied by focusing on changes in the return period of a fixed value of the total water level. To this end, one can define a sensitivity measure as:

$$SM = \frac{1}{T_r} \cdot \frac{dT_r}{dm.s.l.} = -\frac{1}{T_r} \cdot \frac{1}{[1 - G(z - m.s.l.)]^2} \cdot f(z - m.s.l.) = -f(z - m.s.l.) \cdot T_r, \quad (3.6)$$

which is obtained by derivation of Eq. 3.5, and where $f(z) = \frac{dG}{dz}$ is the probability density function associated with $G(z)$. Eq. (3.6) shows that, at a given site and for a set value of $z$, the relative change in return period grows linearly with $T_r$.

## 3.3 Results and discussion

### 3.3.1 Mann–Kendall trend analysis of the deviation of yearly maxima from yearly mean sea level

This application starts by computing mean sea level on yearly basis and by subtracting it from observed total water level. The first question that this Chapter wants to address is the presence of long-term trends, unrelated to sea-level rise and associated with other factors (e.g., human-induced factors, morphological variations), in the "cleaned up" signal, i.e. the observed measurements without mean sea level. To answer this question, this application focuses on the deviation of yearly maxima from yearly mean sea level and tests for the presence of trend by the two-tail Mann–Kendall test (Mann, 1945). Figure 3.4 summarizes results for each location explored. From a first visual inspection of Figure 3.4, the Venice (1872–2019) and Hornbæk (1891–2012) time series appear to show an increasing trend in the deviations of yearly maxima from yearly mean sea level (blue line) of different magnitudes. In contrast, Marseille sea-level observations (1985–2018) seem to be characterized by a decreasing trend. Finally, the Newlyn historical record (1915–2016) displays a fairly constant signal with no noticeable variations. The application of the Mann–Kendall test reveals a partly different story.

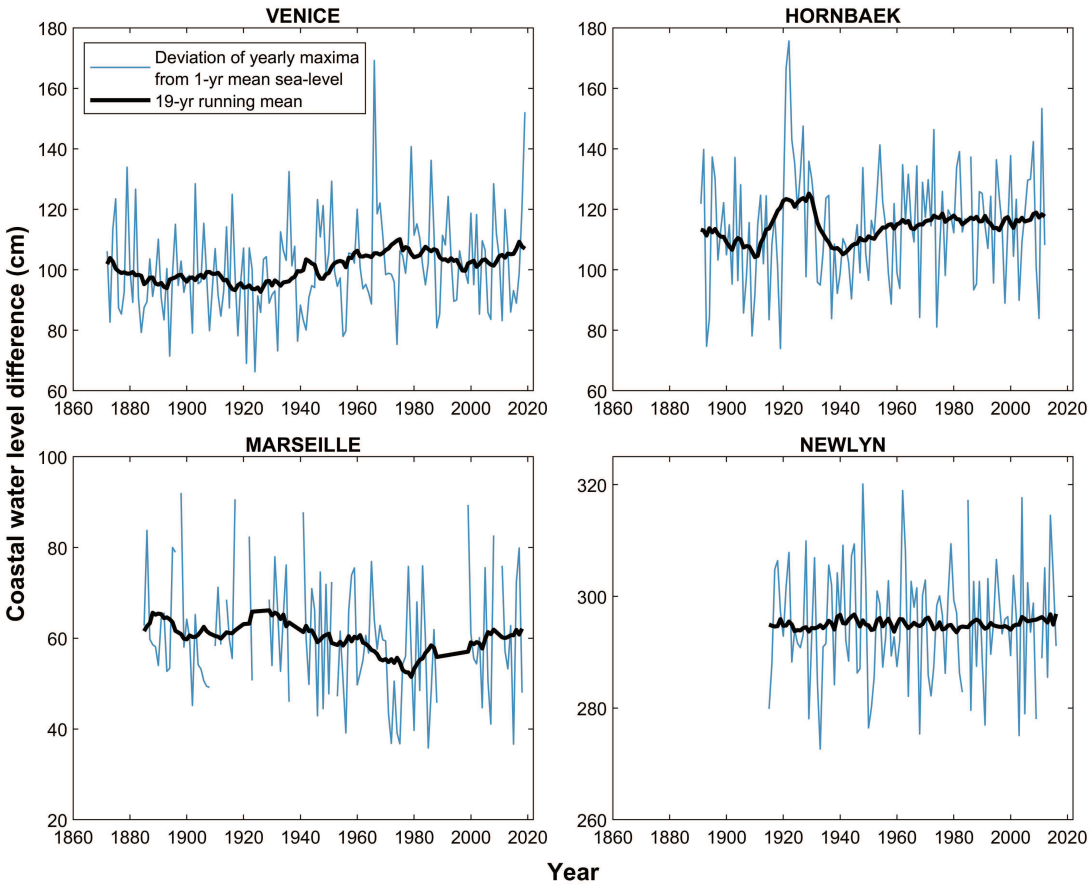


**Figure 3.4:** Deviation of yearly maxima from yearly mean sea level (blue line) and 19-yr running mean (black line) calculated for Venice (IT), Hornbæk (DK), Marseille (FR), and Newlyn (UK).

The test rejects the hypothesis of the absence of trend at the 95 % confidence level, only for the Venice site ($\text{p-value}^{\text{Venice}} = 0.014$). This result suggests that the increase of the yearly maximum deviations from yearly mean sea level may be a direct result of the local morphological variations in lagoon channels where the tidal wave propagates (whereby dissipation of the wave is reduced), and/or land subsidence. In contrast, at the remaining locations, the null hypothesis of no trend cannot be rejected ($\text{p-value}^{\text{Hornbæk}} = 0.352$, $\text{p-value}^{\text{Marseille}} = 0.110$, and $\text{p-value}^{\text{Newlyn}} = 0.997$). The results obtained from these analyses support the validity of the hypothesis that mean sea-level rise is the dominant factor in determining the future frequency of coastal flooding (see Section 3.2.2). For the tests performed here to compare different extreme-value statistical models, the possible presence of trends (e.g., in Venice) is irrelevant, since such tests are performed by first reshuffling observed values, thereby eliminating any existing trend, albeit small.

### 3.3.2 Extreme value analysis

As mentioned in Section 3.2.2, the GEV-based approaches and the MEVD framework require the independence between two consecutive coastal water level events. This hypothesis is guaranteed by imposing a minimum time lag. Firstly, the daily maxima sea levels from the original record are selected; secondly, independent events are defined as those that are separated by at least 30 days. Subsequently, the samples used for statistical inference are built as follows: (1) GEV–BM – the yearly maxima are selected; (2) POT–GPD – as proposed by Coles (2001), the optimal threshold ($u$) is determined by studying the stability of the GPD shape ($\xi$) and modified scale ($\sigma^* = \sigma_u - \xi u$) parameters estimated using a wide range of values of $u$; using this method, threshold values of 65 cm (Venice), 50 cm (Hornbæk), 35 cm (Marseille), and 260 cm (Newlyn) were identified; (3) MEVD – all the independent coastal water level events above a low threshold are used to fit the probability distributions of ordinary values. In the present context, the ordinary values are defined as any coastal water elevation (i.e. the maximum water level reached during a storm event) greater than a site-specific threshold value. This threshold is chosen to be as small as possible (differently from the POT approach), to retain as much of the observational information as possible, and will be dependent on the magnitude of the local tidal range (sea level difference between high and low water level over a tidal cycle) and of storm contributions. Additionally, the threshold is set to be large enough to filter out coastal water level peaks that are likely fully determined by tidal fluctuation, in the absence of any storm contribution. Given the above constraints, the optimal threshold value is chosen (a) by testing several threshold values and evaluating the goodness-of-fit of the distribution using diagnostic graphical plots and (2) so as to minimize the estimation error under the MEVD framework. According to this selection criteria, the low thresholds adopted in the four study sites are 59 cm for Venice, 40 cm for Hornbæk, 25 cm for Marseille, and 250 cm for Newlyn. For every observed site, Table 3.2 and Figure A.1 display the gradual increase in the number of independent events (i.e., annual maxima, exceedances over the threshold, and ordinary values) used to infer the distributions when moving from GEV–BM, POT–GPD to MEVD approaches.

As previously introduced in Section 2.2.1, the MEVD formulation requires the choice of an optimal distribution of ordinary values that can represent the characteristics of the natural phenomenon under analysis. Different candidate distributions for the $F(x; \overrightarrow{\theta_j})$ in Eq. (2.7) are evaluated and the most suitable distribution is selected on the basis of the CV procedure comparing the MEVD-estimated quantiles with the observed ones in each selected sub-sample. In this particular application to extreme coastal water levels, three candidate probability distributions for $F(x; \overrightarrow{\theta_j})$ in Eq. 2.7 are tested, i.e., the Gamma, Weibull, and generalized Pareto distributions. Based on the comparative evaluation of the performance of these distributions, e.g., using diagnostic quantile–

**Table 3.2:** Total number of independent events and average number of events per year for all the gauge stations explored.

| *Site name* | *Independent events* | | | |
|---|---|---|---|---|
| | | *BM* | *POT* | *MEVD* |
| Venice | Total | 148 | 605 | 775 |
| | no. events/year | 1 | 4.08 | 5.23 |
| Hornbæk | Total | 121 | 595 | 736 |
| | no. events/year | 1 | 4.91 | 6.08 |
| Marseille | Total | 106 | 275 | 489 |
| | no. events/year | 1 | 2.57 | 4.61 |
| Newlyn | Total | 100 | 399 | 520 |
| | no. events/year | 1 | 3.99 | 5.20 |

quantile scatter plots, the generalized Pareto distribution emerged as the best model for the "ordinary" coastal-water-level values. It is important to highlight again that the GPD used in the MEVD framework is obtained by imposing a small threshold (differently from the high threshold adopted in the POT–GPD approach) to capture the distribution of the main body of the probability distribution of the ordinary events and does not require the event arrival process to be Poisson (Marani and Zorzetto, 2019).

As suggested by several rainfall applications, ordinary distribution parameters are here estimated using the Probability Weighted Moments (PWMs) method in non-overlapping estimation windows of 5 years. In the present application, the optimal estimation window length was set to 5 years to obtain a more robust parameters estimation, especially when few values in each year are available. PWMs estimation, introduced by Greenwood et al. (1979), is widely applied because of its good performance, particularly in the presence of small sample sizes, its reduced estimation bias and sensitivity to the presence of outliers in the data (Hosking et al., 1985; Hosking and Wallis, 1987; Hosking, 1990).

Considering the above threshold values, the observed and estimated distributions of coastal water level are compared by plotting their quantiles against each other. By comparing measures of in-sample and out-of sample test predictive accuracy, the results are presented by means of quantile–quantile (QQ) plots. The reader can refer to Figure 3.5 (or Figures A.2, A.3, A.4, A.5, and A.6 in Appendix A) to compare the results obtained with the MEVD framework (or the GEV-based approaches – GEV–BM and POT–GPD – vs. the MEVD formulation) for the four sites analyzed. QQ plots are obtained as a result of the CV procedure with 1000 random realizations and sample size: (a) $S = 30$ years (in-sample-test on the left column); (b) $V = M - S$ years

(out-of-sample test on the right column). The colors represent the density of points around the 45° line (i.e. the line of equality). This highlights how the estimated quantiles are closely comparable with the observed ones for all the three approaches tested and for both the sample size explored ($S$ and $V$). In particular, if the reader looks at the supplementary Figures A.2–A.6 in Appendix A and if out-of-sample performance is considered, it is difficult to quantify which distribution is the best due to a large variability in the estimates. Overall, if only the MEVD performance is investigated, the reader can look to the right column (out-of-sample test) in Figure 3.5, where the results display that the MEVD formulation performs similarly for all sites analyzed. In particular, it proves to be a good model for lower and intermediate quantiles but shows variability in the estimates for higher quantiles.

Regarding the evaluation of the performance of the three approaches (GEV–BM, POT–GPD and MEVD) in high-quantile estimation, this application explores the predictive performance of the MEVD and GEV distribution as a function of the NDE (Section 2.3) computed for the maximum return period, $T_{r,max} = M - S + 1$, associated with the maximum value in each test sub-samples that is randomly extracted in the CV approach. The use of NDE metric allows to easily characterize and compare models estimation uncertainty associated with fixed return time of interest and the variation of the calibration sample size (from 5 to 30 years). The results are summarized by means of box plots (Figure 3.6) and kernel density estimates computed for a calibration sample size of 30 years (Figure 3.7). Table 3.3 summarizes the main results underlying the chosen evaluation metric. When we focus on the case of a short sample (5 years), different sites display variable results: I) the GEV and MEVD approaches perform similarly for Venice (Figure 3.6a) and Hornbæk (Figure 3.6b) with similar interquartile ranges and underestimations of the actual quantile; II) for the Newlyn gauge station (Figure 3.6d) the GEV–BM distribution yields better results, even though the POT and MEVD median error are also close to zero. In contrast, when considering longer calibration sample sizes (from 10 to 30 years), the MEVD-based estimates outperform the traditional approaches for most gauge stations explored: (I) results for the Venice site confirm the robustness of the MEVD with respect to the GEV distribution especially for calibration sample size equal to 30 years; in this case, the median error in the MEVD estimates tends to be closer to zero (–0.004), corresponding to approximately unbiased estimates; (II) the Hornbæk station displays similar results to those for Venice and the MEVD-based estimates become more reliable when we consider a calibration sample size greater than 10–20 years; (III) Newlyn estimation errors show a trade-off between the BM method and MEVD for calibration sample size equal to 20 and 30 years.

Results for the Marseille site show a peculiar behavior that requires a specific discussion. In this case, the application of the traditional extreme value theory is advantageous when compared with the MEVD (Figure 3.6c). In order to better understand the

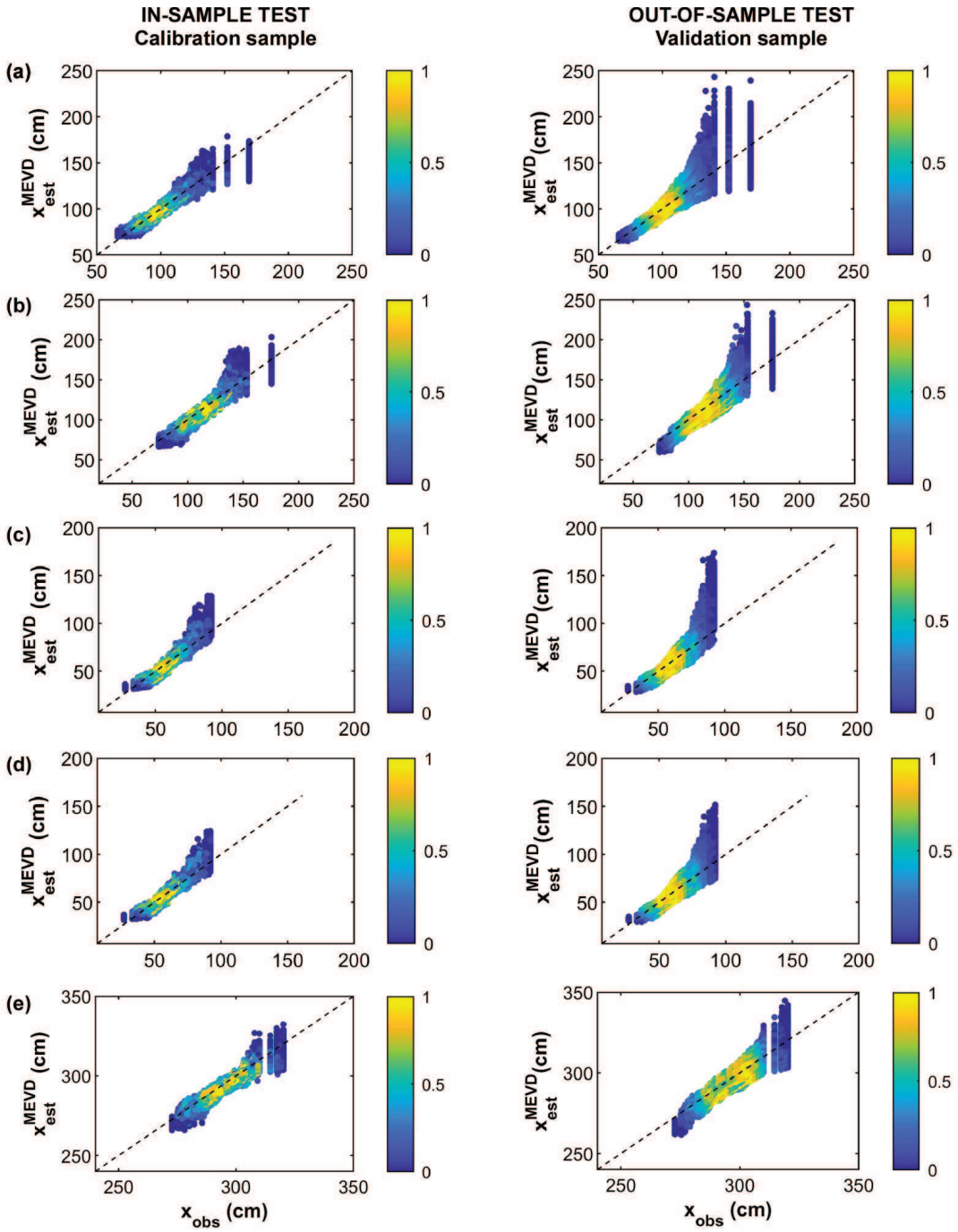


**Figure 3.5:** QQ plots of extreme-coastal-water-level quantiles, computed with the MEVD framework, for the (**a**) Venice (IT), (**b**) Hornbæk (DK), (**c, d**) Marseille (FR), and (**e**) Newlyn (UK) sites. The MEVD parameters estimations are based on non-overlapping sub-samples of fixed size (5 years), while subplots indicated with the letter (**d**) display the QQ plots obtained with MEVD parameters estimations based on data from the whole calibration sample size. The plots are obtained as a result of the cross-validation method used to test the global performance of the models and are estimated for 1000 random realizations and for sample size: (1) $S = 30$ years (in-sample-test in the left column); (2) $V = M - S$ years (out-of-sample test in the right column). The colors represent the point density around the 45° line (black dashed line) corresponding to the best fit.

application to the Marseille site, the MEVD parameter estimation is performed using two approaches: (1) estimation based on non-overlapping calibration samples of fixed size (5 years as for the other sites); (2) parameter estimation on data from the whole calibration sample. The comparison of the results from these two setups confirms that when longer time slices are used for estimating GPD parameters (black color in Figure 3.6c), the MEVD performance is improved (for example when we consider $S = 30$ years, MEVD median$_{[\text{S-year}_{\text{window}}]} = 0.17$ vs. MEVD median$_{[\text{5-year}_{\text{window}}]} = 0.35$), though it does not yet match the results obtained from GEV–BM approach (GEV–BM median error $= 0.016$). This can be explained by considering sea-level peaks occur in Marseille about once every year on average. In this case GEV–BM is advantageous because the small number of events per year does not provide a more numerous calibration sample with respect to the sample of annual maxima. This result confirms the conclusion by Miniussi and Marani (2020), according to which the selection of the estimation window size for fitting the ordinary value distribution strongly depends on the average number of extreme events per year.

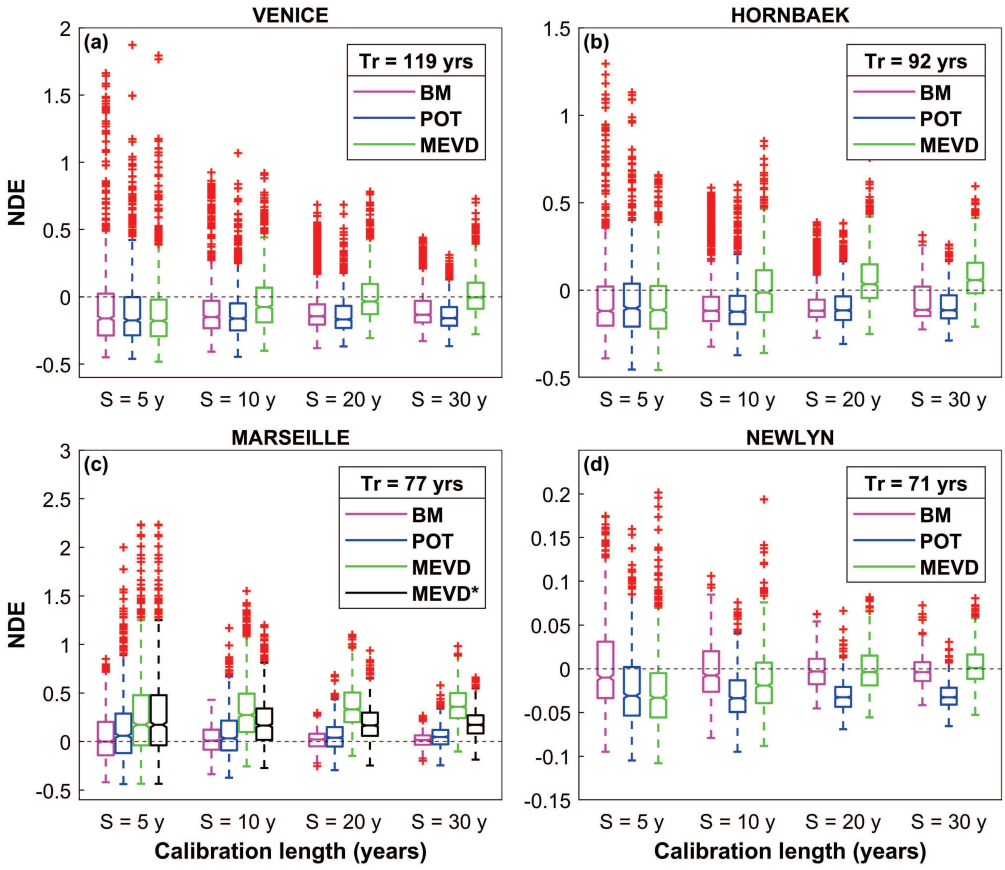


**Figure 3.6:** Distribution of the nondimensional estimation error (NDE) for maximum sample return period ($T_r$) represented by means of box plots at given gauge stations explored: (**a**) Venice (IT), (**b**) Hornbæk (DK), (**c**) Marseille (FR), (**d**) Newlyn. In the case of the Marseille (FR) site, MEVD parameter estimation is based (1) on non-overlapping sub-samples of fixed size (5 years; green color), and (2) on data from the whole calibration sample (black color).

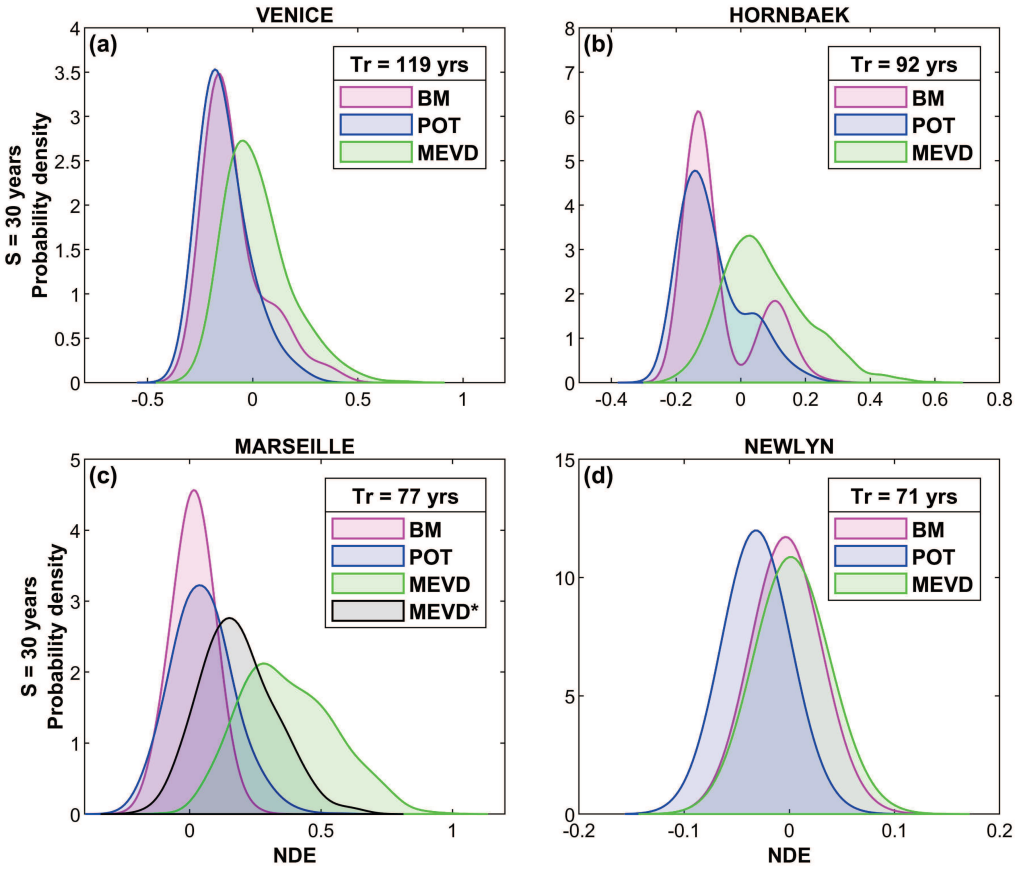


**Figure 3.7:** Kernel density estimates for the nondimensional estimation error (NDE) distributions obtained with a calibration sample size ($S$) of 30 years and maximum return period ($T_r$) at given gauge stations explored: (**a**) Venice (IT), (**b**) Hornbæk (DK), (**c**) Marseille (FR), (**d**) Newlyn (UK). In the case of the Marseille (FR) site, MEVD parameter estimation is based (1) on non-overlapping sub-samples of fixed size (5 years; green color), and (2) on data from the whole calibration sample (black color).

In this Chapter, a comparative analysis between the three methods is also provided to evaluate if the tested extreme-value distributions are representative of the entire range of return times of interest. To achieve this purpose, the performance of the methods is also evaluated for intermediate $T_r$ values, greater than the calibration sample size, since for $T_r < S$ the empirical quantiles can be used. This additional analysis is performed for the Venice, Hornbæk, and Newlyn sites. Figure 3.8 summarizes the results obtained by estimating the probability distribution parameters on 30-year calibration sub-samples. The analyses suggest that when we focus on the median error associated with moderate values of the return period, GEV–BM displays an overall greater robustness (e.g., in the case of Venice and Hornbæk sites) with respect to POT–GPD and MEVD, which exhibit greater fluctuations. In particular, results show that MEVD is a good model for the highest values of the return period, but exhibit a greater absolute value of the estimation error for smaller $Tr$. Overall, the results suggest that no single approach is clearly superior at all values of $T_r$, due to a large variability in the estimates. For example, for

**Table 3.3:** Results of the evaluation metric obtained for all the gauge stations and for calibration sample sizes ($S$) equal to 5 and 30 years. In the case of the Marseille site, text in bold refers to MEVD parameter estimation based on data from the whole calibration sample size.

| *Site name* | *Variables* | *S = 5 years* | | | *S = 30 years* | | |
|---|---|---|---|---|---|---|---|
| | | *BM* | *POT* | *MEVD* | *BM* | *POT* | *MEVD* |
| Venice | NDE median | – 0.160 | – 0.175 | – 0.178 | – 0.133 | – 0.158 | – 0.004 |
| | NDE mean | – 0.069 | – 0.101 | – 0.116 | – 0.087 | – 0.133 | 0.024 |
| | NDE SD | 0.366 | 0.274 | 0.267 | 0.156 | 0.113 | 0.155 |
| Hornbæk | NDE median | – 0.119 | – 0.104 | – 0.113 | – 0.113 | – 0.115 | 0.056 |
| | NDE mean | – 0.069 | – 0.101 | – 0.116 | – 0.068 | – 0.087 | 0.077 |
| | NDE SD | 0.366 | 0.274 | 0.267 | 0.113 | 0.100 | 0.131 |
| Marseille | NDE median | – 0.0003 | 0.059 | 0.172 | 0.016 | 0.047 | 0.357<br>**0.172** |
| | NDE mean | 0.045 | 0.129 | 0.262 | 0.013 | 0.050 | 0.374<br>**0.183** |
| | NDE SD | 0.252 | 0.350 | 0.421 | 0.072 | 0.115 | 0.178<br>**0.140** |
| Newlyn | NDE-median | – 0.010 | – 0.030 | – 0.033 | – 0.003 | – 0.032 | 0.0008 |
| | NDE mean | 0.003 | – 0.022 | – 0.026 | – 0.002 | – 0.031 | 0.002 |
| | NDE SD | 0.050 | 0.042 | 0.042 | 0.016 | 0.014 | 0.021 |

the Venice site there is a decrease (in many cases an unbiased estimate) in the MEVD-NDE values for intermediate $T_r$ (between 85 and 105 years), while for greater $T_r$ values (but smaller than $T_{r,max}$) the error shows an overestimation of the actual quantile with respect to traditional approaches (which exhibit an underestimation tendency). To be more specific, if $T_r > 105$ years are considered, MEVD yields error estimates between 0 % and $< 10$ %, while errors associated with GEV–BM and POT–GPD lie between 0 % and $< -20$ %. The Hornbæk site shows similar results to the Venice site, while Newlyn's results exhibit more fluctuations for large $T_r$ values with much reduced smaller amplitudes and values of the NDE.

### 3.3.3 Future projections of extreme total water levels

The final aim that this Chapter wants to pursue is how sea-level rise may influence the frequency of extreme total water levels across the sites analyzed. As described in Section 3.2.2, only the influence of an increased mean sea level is evaluated, i.e. possible changes in storm regimes are not addressed (see, e.g., Tebaldi et al., 2012).
Site-specific sea-level projections from IPCC's AR5 (Church et al., 2013) are used, which indicate an accelerating sea-level rise at all four observation sites (for each gauge station

under analysis, the reader can refer to the panels (a), (c), (e), and (g) in Figure 3.9), with expected water level increases by the end of the century (RCP8.5) of 48 cm in Venice, 52 cm in Hornbæk, 59 cm in Newlyn, and 54 cm in Marseille. The panels (b), (d), (f), and (h) in Figure 3.9 show observed (green line) and future (blue and red lines) changes in the return period associated with maximum water level events due to sea-level rise. These curves were obtained by using, in Eq. (3.5), the MEVD with parameters estimated on 5-year sliding windows. As noted above, changes in return levels are nonlinear: relative changes are more significant for smaller extremes than for larger ones. The $T_r$ vs. $z$ curves are concave downward and display varying slopes depending on the site explored. When a fixed return period is considered (e.g., 500 years), the mean-sea-level projections quantify the expected increase in extreme-total-water-level peaks for that particular return period. These changes vary heterogeneously across the different coastal sites explored. By comparing the percentage changes associated with the two emission scenarios and the two return periods (Table 3.4), Venice and Marseille are seen to experience the greatest changes in extreme total water levels (e.g., with reference to $T_r = 100$ years and RCP8.5, the variations at Venice and Marseille

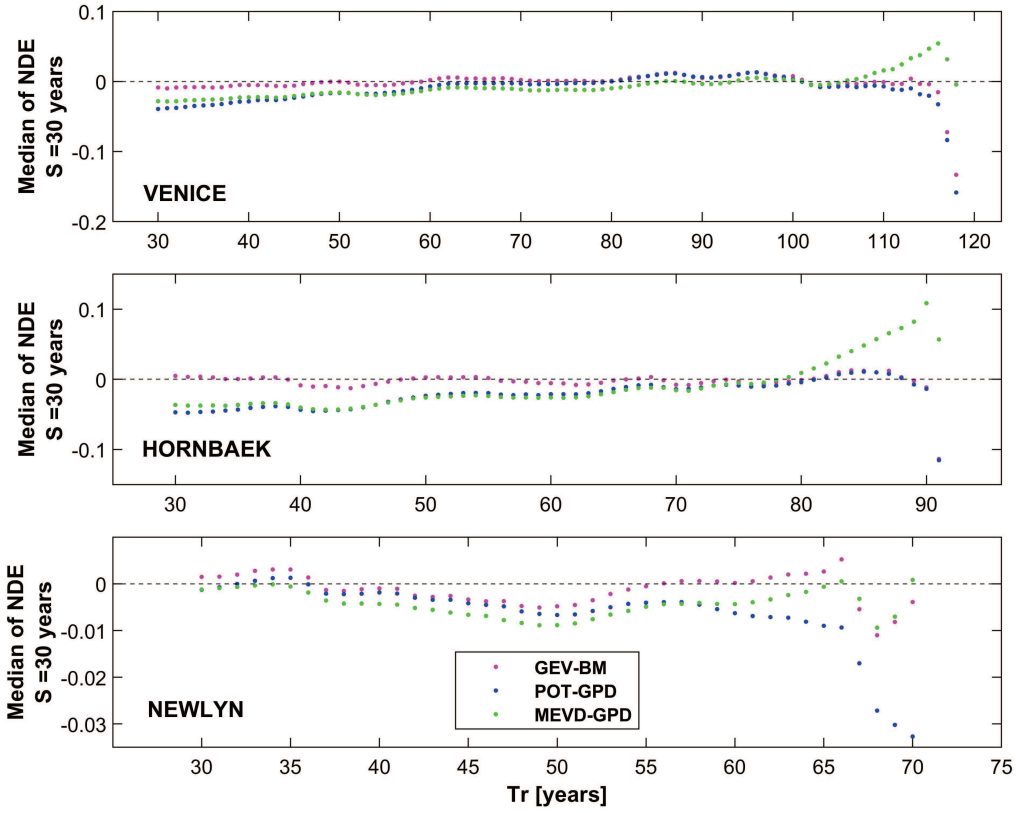


**Figure 3.8:** Median of the nondimensional estimation error (NDE) for return period greater than the calibration sample size in test sub-sample for the GEV–BM, POT–GPD, and MEVD approaches (magenta, blue and green dots, respectively). The results are obtained for the Venice (IT), Hornbæk (DK), and Newlyn (UK) sites and by estimating the distribution parameters on 30-year calibration sub-samples.

**Table 3.4:**
Results of the percentage changes in total water level (Δz) obtained with the two future scenarios (RCP4.5 and RCP8.5) and the return periods (100 and 500 years) for the four sites under analysis.

| $T_r$ | RCP | Δz (%) | | | |
|---|---|---|---|---|---|
| | | Venice | Hornbæk | Marseille | Newlyn |
| 100 | 4.5 | 16.75 % | 14.22% | 22.62 % | 11.29 % |
| | 8.5 | 22.82 % | 21.76 % | 29.73 % | 15.26 % |
| 500 | 4.5 | 14.60 % | 11.70 % | 16.24% | 11.23 % |
| | 8.5 | 19.92 % | 18.18 % | 21.91 % | 15.09 % |

sites are approximately 23 % and 29 %, respectively). All sites display greater percent changes for the lower 100-year return period in each scenario, i.e., "less-infrequent" extremes will be most impacted by sea-level changes in the near future.
Focusing on changes in the return period of a fixed value of the total water level, for example, the reader can see in Figures 3.9b, d, f, h how, for a given value of $z$, changes (horizontal spacing between the curves) are greater for $T_r$ = 1000 years than for $T_r$ = 500 years. Applying the expression for the sensitivity measure, $SM$ (Eq. (3.6)), also tells us that changes in $T_r$ are more significant, everything else being equal, for values of $z - m.s.l.$ near the mode of the distribution, where $f(z - m.s.l.)$ is maximum (e.g., compare changes at the Venice or Hornbæk sites with those at Newlyn for a same initial value of $T_r$). Finally, Eq. (3.6) shows that percentage changes in $T_r$ are highly site-dependent through the shape of $f(z - m.s.l.)$.

## 3.4 Conclusions

In this Chapter the comparative examination of extreme-value distributions applied to observed sea levels at several sites along European coasts provides insights into the predictive performance of traditional and new approaches. The analyses confirm some practical and conceptual advantages of the MEVD with respect to traditional methods. A cross-validation scheme (with 1000 realizations for each site) was used to compare model performance in high-quantile estimation. The use of two independent sub-sample (calibration and test sample) allows the quantification of actual predictive uncertainty. The results highlight that the MEVD approach provides reliable estimates of high quantiles – i.e. associated with the maximum value in each test sub-samples – for almost all the gauge stations explored, particularly when sufficiently long calibration sample sizes are considered (e.g., from 10 to 30 years). Differences in performance between the MEVD framework and GEV-based approaches are subtle, and a definitive conclusion on an optimal solution independent of the return period of interest remains elusive. However, small differences in the estimation accuracy are relevant for engineering applications when dealing with rare extreme events. If we focus on high-return-period

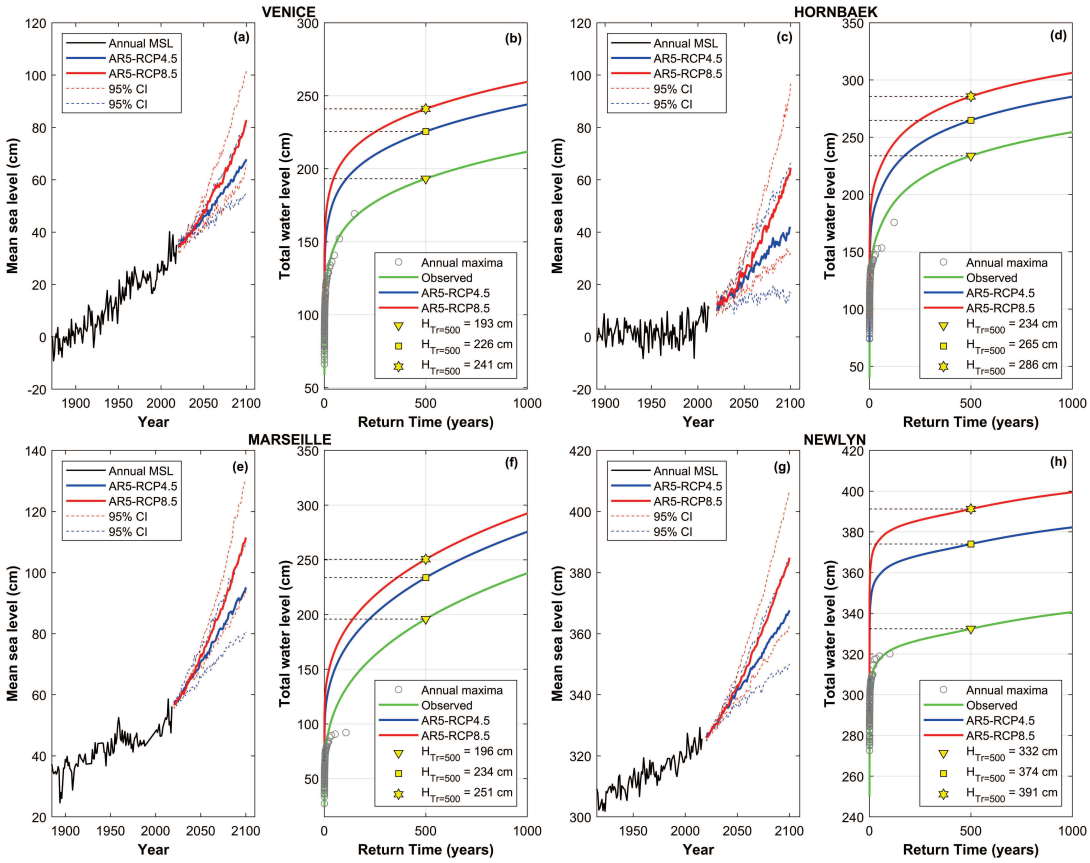


**Figure 3.9:** Future total water level projections, with respect to the current mean sea level, in Venice (IT; panels **a** and **b**), Hornbæk (DK; panels **c** and **d**), Marseille (FR; panels **e** and **f**), and Newlyn (UK; panels **g** and **h**). (**a, c, e, g**) Annual (black line) and future mean sea level until 2100 with RCP4.5 (blue line) and RCP8.5 (red line). Dashed lines represent the 95 % confidence intervals. (**b, d, f, h**) Return period curves for extreme total water level. The green curve represents the estimates obtained with the observed record; the blue and red curves represent the estimates obtained with the projected sea-level rise in the year 2100 with RCP4.5 (blue) and RCP8.5 (red), respectively; the gray dots indicate the observed annual maxima. The triangle, square and pentagon highlight the heights of extreme total water levels for a fixed return period equal to 500 years.

quantile estimation, the analyses show that the MEVD approach provides reliable estimates for almost all the gauge stations explored. Data from the Marseille gauge station exhibit a behavior that deviates from those from other sites, showing an inferior performance of the MEVD with respect to GEV-based approaches. The cause is to be found in the small average number of sea-level peaks every year: the small sample of yearly ordinary events available prevents the MEVD from adding significant information with respect to GEV-BM and POT–GPD. Conversely, when methods performance is evaluated for intermediate return period values, GEV–BM displays an overall greater robustness, and MEVD exhibits a greater absolute value of the estimation error.
Unfortunately, the size of the available data set does not allow to explore model performance for greater values of the return period. Future work could investigate if the estimation error can be reduced, with respect to what was found here, by using different approaches, e.g., by assuming "time-invariant" parameters in the ordinary distribution, whose estimation would thus be performed on the entire calibration data set, rather than on relatively short sliding windows. Synthetic water level time series may be produced by one of the several existing numerical models to extend analyses to arbitrarily long return period values.
Finally, projections of the frequency of extreme total water levels driven by changes in mean sea level are explored. The sensitivity of extreme-water-level frequency to sea-level rise is location-dependent and the results show that, at a given site and for a set value of the total water level extreme, the relative change in return time grows linearly with return time itself.

**Data availability.** All data set used are publicly available from data sources cited throughout this Chapter. Venice sea-level data were obtained from the "Centro Previsioni e Segnalazioni Maree" of the Venice Municipality (`https://www.comune.venezia.it/it/content/centro-previsioni-e-segnalazioni-maree`; Città di Venezia, 2020). Hornbæk (Denmark), Marseille (France), and Newlyn (United Kingdom) water level data were downloaded from the University of Hawaii Sea Level Center (UHSLC) repository (`http://uhslc.soest.hawaii.edu/data/?rq#uh745a/`; Caldwell et al., 2015). The CMIP5 model outputs used for the future total water level projections are available at the "Integrated Climate Data Center" (ICDC) database of the University of Hamburg (`https://icdc.cen.uni-hamburg.de/en/ar5-slr.html`; Church et al., 2013).

# Chapter 4

# Modeling extreme meteorological droughts from paleo-climatic reconstructions: a statistical approach based on the metastatistical framework

Large areas worldwide are exposed to a multitude of natural hazards, such as earthquakes, volcanic eruptions, hurricanes, storms, tornadoes, floods, and droughts. Among these, drought affects all the components of the water cycle and many sectors can be affected. Therefore, assessing the rarity and magnitude of extreme drought events is challenging. In fact, drought studies suffer from the lack of sufficiently long observations (e.g., hydro-climatic records rarely span more than 100 years) and consolidated methods of analysis. Accordingly, extreme quantile estimates are subject to large uncertainty. Since the hydrological time series do not fulfill the basic requirements of the EV theory, this Chapter proposes to (1) use the reconstructions of Palmer's drought index based on paleoclimate data, (2) adopt non-asymptotic formulations to accurately estimate rare return periods associated with drought events. The structure of this Chapter includes 4 Sections as follows. Section 4.1 presents a general introduction of the drought phenomenon; Section 4.2 focuses on the data and methods used in this application, Section 4.3 describes and discusses the results obtained, and finally Section 4.4 draws the conclusive remarks.

## 4.1 Introduction

A drought is one of the most complex, insidious and multifaceted natural phenomena whose effects can have harmful environmental and socio-economic impacts on communities around the world (Ji and Peters, 2003; Van Loon, 2015). As a recurrent feature of climate, it can occur in any hydroclimatological regimes (Wilhite, 2000), with heterogeneous spatial and temporal characteristics across the regions. Recent history has shown the vulnerability of many areas to drought episodes – e.g., the droughts in Southern Africa from 1901 to 1999 (Rouault and Richard, 2005), the 2012–2017 event in California (Ullrich et al., 2018), the so-called Millennium drought (1997–2010) in Australia (Ummenhofer et al., 2009; Aghakouchak et al., 2014; Cai et al., 2014), the 2015 (Ionita et al., 2017; Laaha et al., 2017), 2018 (Brunner et al., 2019; Bakke et al., 2020), 2022 (Commission and Centre, 2022) dry spells in Europe – and their adverse effects on both water supply systems and environment (Rossi et al., 1992). Additionally, the observed and projected increases in severity, frequency, duration, and spatial extent of droughts worldwide (Cook et al., 2014; Forzieri et al., 2014; Bouras et al., 2019; Tramblay et al., 2020) are a cause of significant concern in the scientific community.

A drought is considered as an elusive phenomenon that differs substantially from other natural hazards (e.g., floods, tropical cyclones, tornadoes, earthquakes, and volcanic eruptions), making its management a challenging task. Due to its complex nature and multi-layered impacts, there are many obstacles in its study. Firstly, a universal definition of the term "drought" still remains elusive (Lloyd-Hughes, 2014), and is possibly conceptually impossible and ill-posed. Many drought definitions are adopted in different fields (e.g., meteorology, hydrology, economy), and, indeed, reflect different objectives. Reviews of drought definitions can be found in Palmer (1965); Dracup et al. (1980); Wilhite and Glantz (1985); Tallaksen and van Lanen (2004); Mishra and Singh (2010). Among them, a general definition of drought is given by Tallaksen and van Lanen (2004), which define drought as "a sustained and regionally extensive occurrence of below average natural water availability and can thus be characterized as a deviation from normal conditions of variables such as precipitation, soil moisture, groundwater, and streamflow". In applying this definition, the question is: "what are normal conditions?". As discussed in Van Loon (2015), what should be regarded as the "normal" situation strongly depends on what the water is used for. Realistically, definitions of a drought and its impact depend on the focus of the study. Secondly, drought events are rarely confined to a single point but, conversely, affect large areas. Thirdly, drought evolution is influenced by a variety of hydrometeorological variables (e.g., precipitation, evapotranspiration, runoff), which implies an even greater complexity in assessing drought properties and dynamics (Mishra and Singh, 2010). Moreover, since this phenomenon develops slowly over a prolonged period of time (months, years, or even decades), and may linger for years after the conclusion of the event, the drought

onset and termination is difficult to determine. Finally, droughts affect different components of the hydrological cycle, ecosystems, and sectors of society, which makes their direct, indirect, and intangible impacts often difficult to quantify. As result, the research community sometimes uses the term "creeping disaster" to describe a drought (Gillette, 1950).

Droughts are generally classified according to four types (Wilhite and Glantz, 1985): meteorological, hydrological, agricultural, and socio-economic drought. A meteorological drought occurs due to a prolonged lack of precipitation over a large area and is the focus here. Hydrological drought is related to a period with inadequate surface and subsurface water uses under a given water management scheme. Agricultural drought refers to a period with declining soil moisture and consequent crop failure. Finally, socio-economic drought is linked to the impact of the above-mentioned types of droughts on society.

Due to the persistence and often long interarrival times between droughts, occurring on time scales of years to decades or more, very long observational time series are necessary to study their statistical properties. It is rare that such a large amount of observational information, at appropriate space-time resolutions and consistency, is available (Monteleone et al., 2020), which is a major limitation to the development of appropriate drought models and to the quantification of the probability of occurrence of future droughts. One possible approach to overcome this problem relies on the use of proxy climatic data to extend the instrumental record. Additionally, since relatively few droughts occur even within records of several hundred years, techniques which optimally use available information, such as the MEVD framework, are highly beneficial in the analysis of drought probability.

Motivated by the above considerations, this study exploits the long tree ring reconstruction of the summer self-calibrating Palmer Drought Severity Index to investigate the probabilistic structure of droughts. The main aim of the study is twofold: (1) to identify a relationship between tree ring-derived index and local hydrologic balance, (2) to estimate extreme return periods using a novel statistical approach based on the distribution of the ordinary events, i.e. all the independent realizations of the variable of interest, as opposed to just block maxima or values above a threshold.

## 4.2 Materials and Methods

### 4.2.1 Palmer Drought Severity Index

Unlike other extremes, the exact quantification of drought characteristics in terms of intensity, magnitude, duration, and spatial extent is very difficult. For this reason, several drought indices have been proposed in recent decades to quantify a drought, each with its own strengths and weaknesses. Examples are the Palmer drought severity

index (PDSI; Palmer, 1965), the rainfall anomaly index (RAI; van Rooy, 1965), deciles (Gibbs and Maher, 1967), crop moisture index (CMI; Palmer, 1968), the standardized precipitation index (SPI; McKee et al., 1993), the standardized precipitation evapotranspiration index (SPEI; Vicente-Serrano et al., 2010), the standardized streamflow index (SSI; Peña-Gallardo et al., 2019), or the standardized snowmelt and rain index (SMRI; Staudinger et al., 2014).

In this application the self-calibrating Palmer drought severity index (sc-PDSI; Wells et al., 2004) is the focus. It is a revised version of the original PDSI proposed by Palmer (1965). Briefly, PDSI is a normalized indicator of drought that represents a measure of the balance between atmospheric moisture supply (precipitation) and demand (evapotranspiration) from month to month. PDSI is, hence, based on the supply-and-demand concept of the water balance equation by considering antecedent precipitation, moisture supply, and moisture demand (based on the Thornthwaite formulation) into the hydrologic accounting system. Palmer used a two-layer bucket-type model for soil moisture computations and made some assumptions concerning field capacity and transfer of moisture to and from the layers. Many limitations have been highlighted in the Palmer model and among these, the following are mentioned. Firstly, the assumptions used, for example: (1) the top soil layer has a field capacity of 1 in. (i.e. 2.45 cm), (2) moisture is not transferred to the bottom layer until the top layer is saturated, (3) runoff occurs only when both layers are saturated, and (4) all of the precipitation occurring in a month is utilized during that month to meet evapotranspiration and soil moisture demand or be lost as runoff. Secondly, the use of the Thornthwaite parametrization for potential evapotranspiration (PET) rather than the more physically realistic Penman–Monteith formulation. Thirdly, other factors such as changes in surface solar radiation due to changes in cloudiness or aerosol concentrations are not explicitly considered in the Palmer model. Fourthly, PDSI cannot reflect soil moisture conditions in presence of snow, melt, or frozen. Lastly (probably the main critical point), the PDSI values are not comparable between diverse climatic regions. In fact, in the PDSI formulation, Palmer computed a set of empirical constants and duration factors by averaging the values from only a few locations in the United States. These averaged values directly affect the spatial comparability of the index.

Nevertheless, PDSI is an approximate measure of the cumulative effect of atmospheric moisture supply and demand (Dai et al., 2004). Thus, this index can be used to monitor long-term evolution of drought. A detailed description of the PDSI is given in Palmer (1965), Alley (1984), and Karl (1986).

One of the PDSI deficiencies was resolved by the development of the sc-PDSI (Wells et al., 2004) which is spatially comparable across diverse climatic regions. In fact, Wells et al. (2004) replace the empirically climatic characteristics and duration factors in the index computation with dynamically calculated values, the latter based on the local historical climatic records. Starting from the formulation proposed by Wells

et al. (2004), van der Schrier et al. (2013) present a new global sc-PDSI data set that uses the Penman–Monteith formulation for PET and includes a simple snow melt and accumulation process in the water balance. These global maps of monthly sc-PDSI based on gridded meteorological time series available from the Climate Research Unit (CRU; Harris et al., 2014), are used as target points for the tree-ring-based Old World Drought Atlas used here.

### 4.2.2 Data

#### The Old World Drought Atlas

The Old World Drought Atlas (OWDA; Cook et al., 2015) is a set of year-to-year maps of tree ring-based reconstructions of summer season (June–July–August, JJA) self-calibrating Palmer Drought Severity Index values (sc-PDSI; van der Schrier et al., 2013). The reconstruction is an indicator of the spring-summer soil moisture on 5414 half-degree grid points (small black grid points in Figure 4.1) for the entire Common Era over the European-Mediterranean domain. OWDA data set is publicly available at `https://www.ncei.noaa.gov/access/paleo-search/study/19419`.

The PDSI is estimated using a regression method on tree-ring properties. The method of regression is analogous to that proposed to reconstruct the drought over North America (North American Drought Atlas, NADA; Cook et al., 2004, 2010b), monsoon Asia (Monsoon Asia Drought Atlas, MADA; Cook et al., 2010a) and northwestern Africa (Touchan et al., 2011). In particular, the reconstruction uses the point-by-point regression (PPR) method described in Briffa et al. (1986), Cook et al. (1994), and Cook et al. (1999), with an extension of the procedure producing ensembles of climate reconstructions (Cook et al., 2010a, 2013).

PPR is a well-tested regression method based on the premise that ≪only those tree-ring chronologies proximal to a given PDSI grid point are likely to be predictors of drought, where "true" implies a casual relationship between tree rings and drought that is stable through time≫ (Cook et al., 1999). As originally proposed, PPR had two variables to optimize: (1) a fixed search radius around each grid point, and (2) screening probability. The first defines the local zone control used to find candidate tree-ring predictors of PDSI. The screening probability (expressed in terms of $\alpha$ -level probability), instead, eliminates tree-ring chronologies from the candidates that are poorly correlated with drought. After several tests, Cook et al. (1999) found maximum reconstruction fidelity using 450 km search radius and $\alpha$-level probability of 0.10 reconstructing gridded drought over the United States. However, for the OWDA domain (as for the MADA domain), the originally proposed PPR needed to be changed due to the uneven distribution of tree ring data over the domain and its sparse coverage, especially in the eastern portion of the domain (see Figure 4.1). Therefore, an ensemble of reconstructions at each grid point was introduced to improve accuracy in the

reconstruction. To do this, some arrangements were necessary. Firstly, the screening probability was deleted due to the high level of noise in the tree rings. Secondly, a dynamic search radius has replaced the fixed one and has become part of the ensemble approach used for the reconstruction.

As a result, an ensemble point-by-point regression was used to produce the OWDA. It is a generalization of the original PPR method (Cook et al., 1999) whereby each tree-ring chronology found within a given search radius around each grid point is weighted by some power of its correlation with the climate variable being reconstructed. See more details in (Cook et al., 1999) and (Cook et al., 2010a). The correlation-weighted chronologies are then used in the principal components regression (PCR) model at each grid point for reconstructing climate from tree rings. This is done in lieu of selecting a subset of chronologies based on a fixed correlation screening probability. For the OWDA, Cook et al. (2015) set the initial search radius to 1000 km. It is interesting to note that in this reconstruction a "dynamic search radius" is used to find at least 20 chronologies within the radius of a given grid point. According to this request, the search radius at the grid point is enlarged by 50 km increments until at least 20 chronologies are found. Here, PPR sequentially reconstructs each grid point of climate variables (i.e. 5414 target grid points of sc-PDSI calculated from version TS3.21 of the high-resolution CRU climate variables (Harris et al., 2014) from the network of annual tree-ring chronologies (i.e. 106 chronologies partly extracted from the International Tree-Ring Data Bank, ITRDB, and partly derived from contributions by European dendrochronologists) using principal components regression (PCR). This resulted in a 8-member ensemble, which was then robustly averaged for further use.

### Climate data

Long time series of rainfall and temperature are here used to validate the tree ring reconstructed JJA sc-PDSI with historical records. In particular, two gauge stations in Italy were chosen as testing cases: Linguaglossa (CT, Sicily Region, 37.84 N, 15.14 E; 530 m a.s.l.) and Padova (PD, Veneto Region, 45.41 N, 11.88 E; 12 m a.s.l.). The Padova time series includes 281 years of daily rainfall observations from 1725 to 2019, with only 14 missing years Marani and Zanetti (2015) and 223 years of daily air temperatures from 1774 to 1994 Camuffo (2002a,b,c); Cocheo and Camuffo (2002). Since temperature measurements for years 1725-1764 were taken indoor, they are excluded from the present analysis. For the Linguaglossa site, monthly rainfall series (1921–2008) and maximum and minimum monthly air temperature (1926–1996), available through the Water Observatory of Sicily Region whose data are collected on the website of the Servizio Informativo Agrometeorologico Siciliano (SIAS; `http://www.sias.regione.sicilia.it/frameset_dati.htm`), are used. In particular, the precipitation data used for Linguaglossa are the same used in the work by

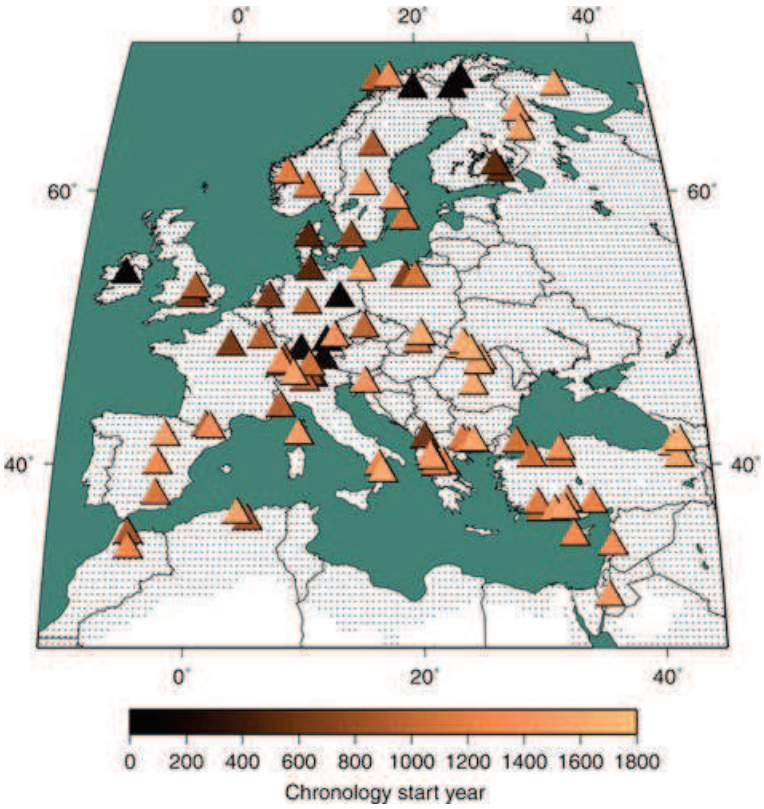


**Figure 4.1:** Map of the JJA sc-PDSI target field (small black grid points) and the 106 chronology tree-ring network (filled triangles) used for reconstruction. Image credit: Cook et al. (2015).

Bonaccorso et al. (2015), where the authors reconstructed missing years by applying the Thiessen polygons method. These raw data were pre-processed to (1) extract the monthly rainfall and precipitation time series, and (2) eliminate years with less than 12 months of rainfall and precipitation observations. Finally, the obtained monthly rainfall and temperature time series are merged. At the end of the process, 71 years of observations are available for Linguaglossa (1926–1996 without missing years) and 207 years are available for Padova (1774–1994 with 14 missing years). Figure 4.2 shows the data used for the subsequent analyses.

### 4.2.3 Method

#### Climate signal in the Old World Drought Atlas

Dendrochronologists have long found that ring widths from tree rings on arid or extremely cold sites reflect past variations in climate. Several studies have demonstrated the linkages between tree ring-derived indexes and the local hydrologic balance (e.g., Fritts et al., 1971; Briffa et al., 2002; Dobbertin, 2005). These linkages are particularly tight for the warm-season temperature, as the dominant growth-limiting factor of tree ring widths. As a first step, however, the response of tree ring growth to instrumental-based data of temperature and rainfall is here assessed. This analysis

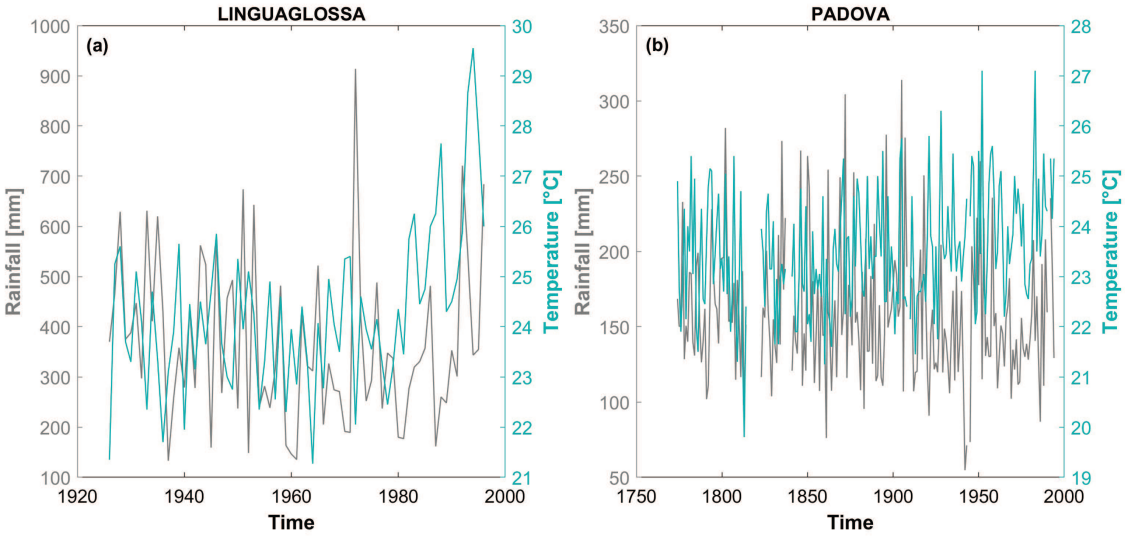


**Figure 4.2:** Rainfall and temperature data for the two selected locations: (**a**) Linguaglossa, CT, Italy, and (**b**) Padova, PD, Italy.

is based on simple correlation, i.e. Pearson correlation coefficient calculated between the total ring width chronologies (annual indices) and the climate records. Since soil moisture retains some memory of previous months, the correlation is computed considering 12 time lags. To be more specific, the climate signal to compare with the tree-ring chronologies was defined as the average monthly values from 1 to 12 previous months. This analysis was assessed for the Linguaglossa site since its annual ring-width indices are available from the International Tree-Ring Data Bank (ITRDB, `https://www.ncei.noaa.gov/products/paleoclimatology/tree-ring`). After this preliminary step, focusing on the two case studies (i.e., Linguaglossa and Padova) and prior to any drought occurrences analysis, the JJA sc-PDSI data from OWDA and those estimated from instrumental data – computed using the Thornthwaite formulation (Thornthwaite, 1948) to estimate the evapotranspiration – were compared based on (1) visual inspection of the time series, (2) the Pearson and Spearman's rank correlation coefficients, (3) scatter plots, and (4) contingency matrices.

### Definition of drought

The choice of the drought characteristic and the type of drought considered for a specific study are critical steps to take in defining events characterized by below-normal conditions.

As previously discussed, the drought metric used for event definition and analysis is the summer sc-PDSI, which reflects spring-summer soil moisture conditions. It is widely used as a paleoclimate reconstruction indicator (Cook et al., 1999, 2004, 2010a; Smerdon et al., 2015). The choice of the PDSI as a valuable drought indicator is e.g., justified in Cook et al. (2016), where they conclude that it represents a more complete picture of

the surface moisture balance due to a long-term view of moisture deficits and surpluses compared to indices derived from monthly precipitation anomalies alone. In fact, the PDSI integrates changes in the surface moisture balance over a time scale of 12 months (Guttman, 1998) and compares well with tree-growth models.

In the present analyses, the grid points in the OWDA domain corresponding to the selected study sites are identified and the corresponding sc-PDSI time series are extracted. Drought events are then defined as those periods of time in which the values of the sc-PDSI remains below a threshold, or truncation level, $thr$. This approach is known as run theory (Yevjevich, 1967), and allows the definition of different drought properties: (i) drought duration is the length of time during which the x-value (i.e. sc-PDSI here) remains below $thr$, (ii) the sum of the deviations between the truncation level and the trace of the sc-PDSI time series is the deficit amount, and (iii) the drought intensity, or drought severity, is obtained by dividing the deficit by duration. The truncation level may be a constant, a stochastic variable, a deterministic function, or a combination of these. In general, the literature expresses the truncation level in terms of some central tendency index of the drought sample, e.g., the mean, median, or mode.

In this application, the threshold level is varied to explore the sensitivity of results on its choice and to verify that the event independence condition is met. After exploring different threshold values, a value of –1 was adopted here. Moreover, the choice of this suitable threshold was also confirmed by the study of the correlation structure of the events (Figure 4.8). Figures 4.3 and 4.4 show the wet and dry periods identified in the time series using the threshold-based approach. In particular, this application focuses on the statistical characterization of the deficit amounts which constitute the sample used for the following statistical inference.

### Extreme value statistics

Due to the random nature of contributing factors, stochastic and probabilistic theories are proper approaches to describing the drought characteristics. In general, the most extreme events have the greatest impact but, by definition, are "rare" and therefore difficult to characterize empirically. While this issue is inherent in the definition of "extreme event", it is particularly critical for drought conditions due to their long characteristic time scales. To overcome this problem, here the metastatistical framework – (MEVD; Marani and Ignaccolo, 2015; Zorzetto and Marani, 2019) and (SMEV; Marra et al., 2019) – is applied. The MEVD formulation is based on the concept of "ordinary values", which are independent realizations of events without regard to their magnitude, for which a more extensive sample can be identified in the observations (see more details in Section 2.2). In the present case, the ordinary values are defined as the magnitudes (of duration, deficit, or intensity) of all the events for which the value of

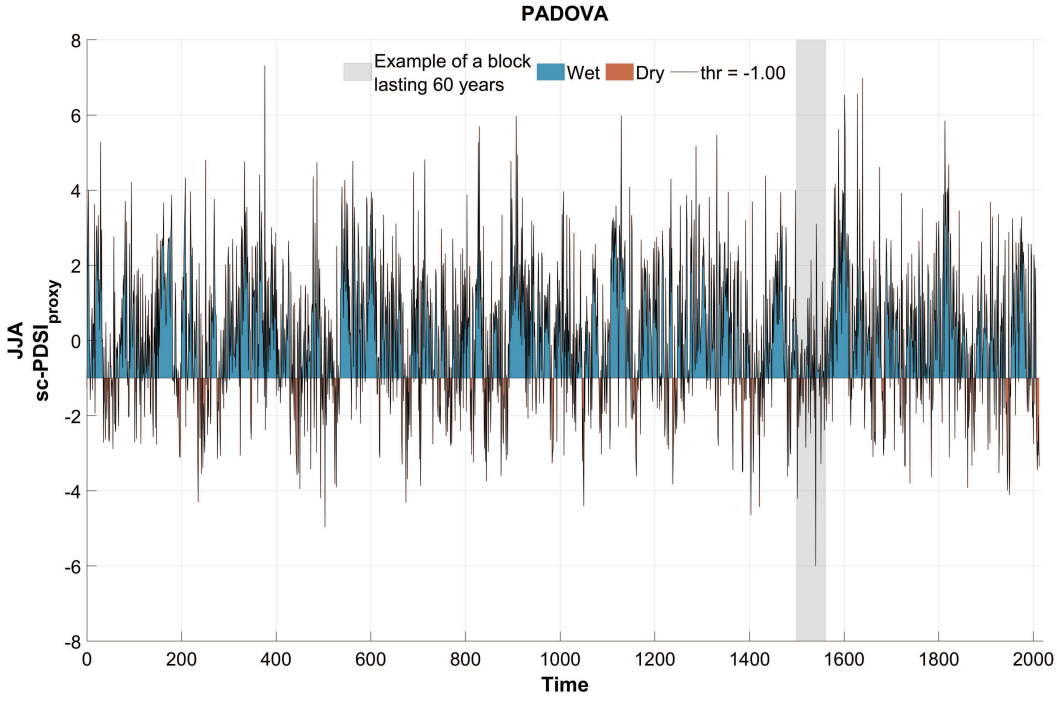


**Figure 4.3:** Run theory applied to summer (JJA) sc-PDSI time series for Padova, PD, Italy.

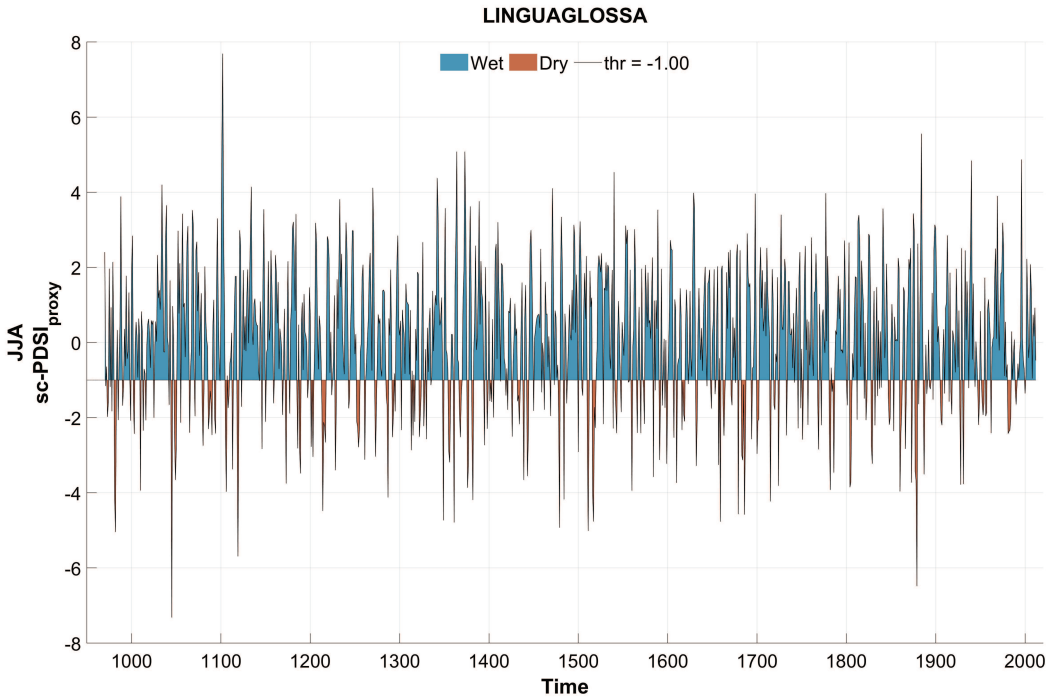


**Figure 4.4:** Run theory applied to summer (JJA) sc-PDSI time series for Linguaglossa, CT, Italy.

the PDSI remains below a set threshold. This definition is similar to that adopted in much work in the existing drought literature (see e.g., Heudorfer and Stahl, 2017; Stahl et al., 2020).

To our knowledge, only one precedent exists of application of the MEVD framework to extreme drought occurrence. Araujo et al. (2022) used the SMEV formulation to model changes in future drought risk assessments in Australia by using different climate projections. In that application, the drought events were identified as a continuous negative SPI/SPEI anomaly exceeding 1 standard deviation. Following Marra et al. (2020), the SMEV distribution was applied to the largest portion of ordinary events and a stretched-exponential decay describes the tail of the distribution. However, the work by Araujo et al. (2022) does not include any comparison with approaches based on the asymptotic results of the EV theory which is instead a pivotal analysis of this application.

In the present application, several candidate probability distributions – i.e., the Gamma, Weibull, 2-parameter log-normal, and generalized Pareto distributions – are tested to describe the ordinary values distribution (i.e., $F(x; \overrightarrow{\theta_j})$ or $F(x; \overrightarrow{\theta})$ in Eqs. 2.7 and 2.8, respectively). The most appropriate distribution is identified using diagnostic quantile–quantile scatter plots comparing the MEVD (SMEV)-estimated quantiles with the observed ones as part of the cross-validation procedure (the reader can refer to the next Section for more details). Based on this visual inspection of the outcome from several cross-validation randomizations, the Gamma distribution was selected. In order to identify the possible benefits of different assumptions, the MEVD is here applied according to three formulations: (1) the full *MEVD* formulation, in which both the ordinary-value distribution and the number of events in each block are allowed to vary from one block to another, (2) a formulation denoted as *MEVD**, in which the distribution of the ordinary events is constant across different blocks (i.e. $F(x; \overrightarrow{\theta}) =$ constant), while the number of events per block is allowed to vary, and (3) a *MEVD*** formulation, in which the probability distribution of ordinary events is allowed to vary, while the number of events per block is taken to be constant and equal to the overall average across all blocks. The MEVD-based estimates from these formulations are then compared with those obtained with the GEV distribution applied on the basis of the block maxima model.

As discussed earlier (see Section 2.2.1), the metastatistical formulations as well as the traditional EV theory require the specification of a "block" size within which the extreme value distribution is defined. We have experimented with different choices of block length to find that suitable choices depend on ensuring that a sufficient average number of events fall within each block, and this is dependent on climatic conditions at each location. A careful choice of block size is thus an important step in developing reliable probability estimates are possible. Here, a block length of 40 (Linguaglossa) and 60 (Padova) years is used. In the cross-validation experiments performed we typically

found that sensible results are obtained when the average number of events per block is greater or equal to about 8–10. Clearly, it is desirable to define the shortest possible block length, at least in metastatistical formulations, because this allows to capture inter-block variability at a finer temporal scale.

#### Cross-validation procedure

In this particular application to extreme meteorological droughts by using paleo-climatic archives, the CV procedure can be summarized as follows:

1. the time series of the drought periods is divided into blocks of fixed length, where the starting event to define the block is randomly chosen $N_{s,y}$ different times ($N_{s,y} = 100$ here). This step ensures that blocks are not always composed by the same set of events;

2. in every random starting point selection from step (1) $S$ of the $M$ blocks into which the record is divided are randomly selected as the calibration sub-sample, on which parameters are estimated. This randomization removes any possible non-stationarity and serial correlation. Note also that the internal event composition of each block remains unchanged to preserve both the ordinary value frequency distribution in each block and the distribution of the number of events per block. The $V = M - S$ blocks not used in parameter estimation are used for testing and constitute the "validation sample" or "test sample". The validation is performed by computing for each block maximum in the validation sample the corresponding Weibull empirical frequency estimator, $F_i = i/(V + 1)$ – where $i$ is the rank of the i-th block maximum in a list sorted in ascending order, is computed and equated to the cumulative probability from the fitted distribution. The empirical return period $T_r$ associated with each block maximum is then simply $T_{r,i} = 1/(1 - F_i)$. For each realization of calibration-validation partitioning (repeated $N_r = 100$ times here) a nondimensional estimation error (NDE) – i.e. the relative error between the estimated and observed quantiles – is computed as follows: $NDE_p(S, T_r) = [h_{(est,p)}(S, T_r) - h_{(obs,p)}(S, T_r)]/h_{(obs,p)}(S, T_r)$;

3. the CV scheme is repeated $N_{s,p}$ x $N_r$ times.

This procedure is performed for different calibration sample sizes ($S = 3$, and 5 blocks).

## 4.3 Results and discussion

### 4.3.1 Climate sensitivity of the OWDA tree-ring chronologies

The responses of annual tree-ring radial growth to climate are identified by using a Pearson correlation analysis between the tree-ring chronology and climate variables.

**Table 4.1:** Pearson and Spearman's rank coefficient for the two case studies

| | **Linguaglossa** | | **Padova** | |
|---|---|---|---|---|
| | Precipitation | Temperature | Precipitation | Temperature |
| Pearson | 0.04 | -0.08 | 0.16 | -0.04 |
| Sperman | -0.03 | -0.13 | 0.14 | -0.06 |

Since tree growth has a tendency to lag time (Fritts, 1976), as described in Section 4.2.3, the correlation is computed on backward windows of length from 1 to 12 months. The results are not shown here. Looking at the Linguaglossa site, positive correlations between plant growth and spring-summer temperature (i.e. May–August) were found. Instead, an undefined pattern and a weaker association were found between tree-ring width and precipitation variability.

To further identify the signature of the climate signal in the reconstructed sc-PDSI from the OWDA, Figure 4.5 shows dry and wet periods identified in the drought index time series reconstructed using proxy records (black line) or instrumental temperature and precipitation data (red line) for the two case studies. On the basis of the visual inspection of the two time series in overlapping years (i.e., 1926–1996 for Linguaglossa and 1774–1994 for Padova), with the exception of just a few periods, a similar trend is found, especially for the Linguaglossa location. Correlation analysis using the Pearson and Spearman's rank coefficients is also performed. For each case study, the obtained values do not show a strong correlation between the reconstructed sc-PDSI and climate data (i.e., annual mean precipitation and temperature). The two correlation metrics both yield a negative correlation (even if small) between temperature and proxy. When precipitation is considered, the Pearson and Spearman coefficients suggest a positive correlation for Padova and an opposite correlation sign for Linguaglossa. To be more specific, the values are reported in Table 4.1, while the corresponding scatter plots are reported in Figure 4.6. The JJA sc-PDSI values, from proxy and estimated from instrumental observations, are plotted against each other and a Total Least Squares method – TLS – was applied. TLS is a regression method that minimizes the sum of the squared distances between the points representing observations and the regression line (in the direction orthogonal to the line). Therefore, this model assumes that measurements of both variables are affected by "errors". In addition, since the error is present in both directions (i.e., x and y-axis), the scatter plots are discretized into 6 classes defined with respect to both the axes (x and y-axis). The average of the values within the class is represented with black triangles (red stars). However, the results of this comparative analysis do not seem to provide a precise pattern, and cloud of observations is quite dispersed.

The lack of a clear correlation signal in previous analyses may be due to a low signal to noise ratio and to non-linearities inherent to the computation of the sc-PDSI. A further

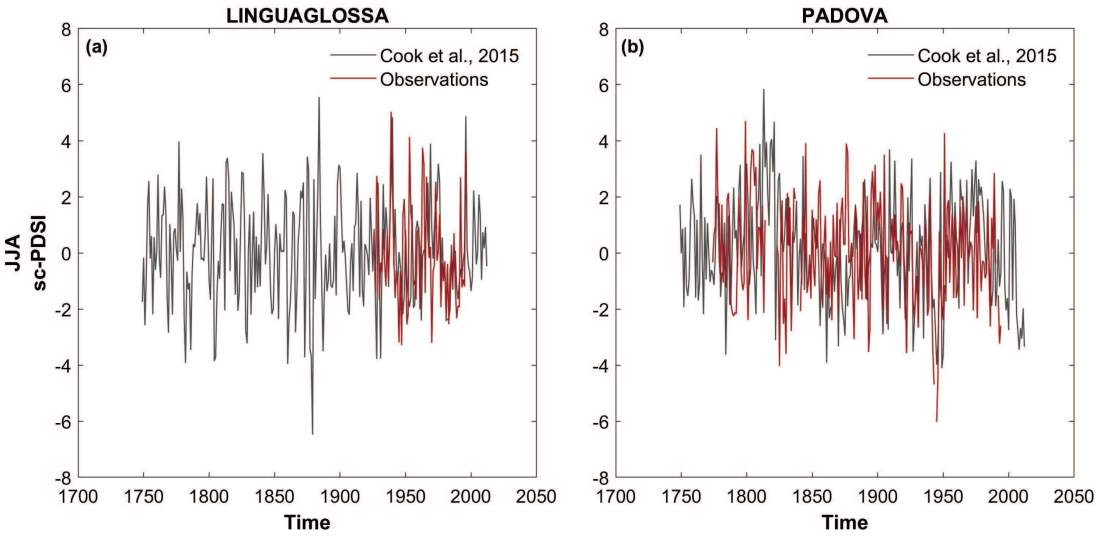


**Figure 4.5:** Self-calibrating Palmer Drought Severity Index, sc-PDSI, for (**a**) Linguaglossa (CT, Italy) and (**b**) Padova (PD, Italy) sites. The black line is the sc-PDSI extracted by the grid point of the OWDA domain corresponding to the study sites. Red line is the time series of the sc-PDSI estimated using observed temperature and precipitation data at each site.

analysis thus involved exploring the possible presence of a relation between corresponding percentiles of the sc-PDSI values from tree-ring information and from observations. Figure 4.7 shows contingency (or confusion) matrices: the matrices are obtained by dividing the samples into tertiles of sc-PDSI along both axes, and by computing the number of values from the sample falling in each cell (the number indicated in each matrix cell). Overall, results for both sites suggest that the matrices are diagonally dominant, except for the matrix element in position (3,3) in Figure 4.7(b). Despite the noise, these analyses show that there indeed is a detectable covariation between the two variables. In other words, when the tree-ring sc-PDSI value remains below its average, also the corresponding observation-based sc-PDSI remains below its own average. These considerations suggest that the use of tree-ring based sc-PDSI analyses may provide insight into the statistical properties of droughts that are similar to those that may be obtained from the observation-based PDSI (if they were available for the entire proxy historical record).

### 4.3.2 Extreme value analysis

After demonstrating that the summer drought index reconstructed from proxy paleoclimatic information indeed mirrors, to some extent, the variability of observed indexes, the analysis proceeds to the application of the extreme value models to gather information on the frequency of extreme drought events. As discussed in Chapter 2, the presence of serial correlation in time series can produce some issues in performing a frequency analysis because extreme value modeling requires the independence between two consecutive events. In general, as discussed by Dracup et al. (1980) in their work

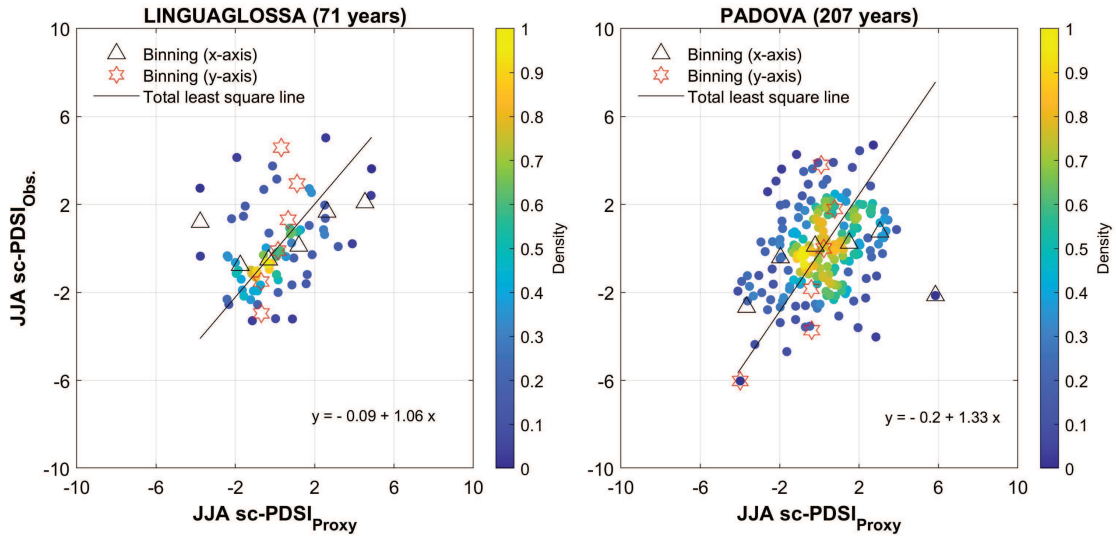


**Figure 4.6:** Scatter plots between JJA sc-PDSI extracted from the OWDA domain (x-axis) vs. those computed using rainfall and temperature measurements (y-axis) for (**a**) Linguaglossa (CT, Italy) and (**b**) Padova (PD, Italy) sites.

about the definition of droughts, a shorter period tends to result in greater serial correlation in the time series. Therefore, daily or monthly events usually show more serial correlation with respect to the yearly ones. According to this idea, the use of annual data resolution (as in this application) ensures that serial correlation is not an issue. The existing literature on threshold-based approaches (see e.g., Tallaksen et al., 1997) avoids splitting major droughts into several minor droughts and pools together events that are separated by less than 2–3 days. For example, to overcome this problem, Brunner et al. (2019), following the approach proposed by Tallaksen et al. (1997) and Van Loon (2015), apply a time series smoothing (over a window of 60 d) prior to event selection. In addition, they prescribe a minimum event duration of 20 d to ensure independence in daily discharge time series.

Therefore, prior to further analysis the autocorrelation structure of the drought series (i.e. the JJA sc-PDSI time series below the chosen threshold) is studied. For each case study, the autocorrelation function (ACF, Figure 4.8) is computed using time lags of fixed length (5 years, here). At the Padova site (Figure 4.6(b)), excluding the longest two lags, where few pairs of values from the series exist (see Figure 4.8(d)), the ACF decreases for increasing lag. Even though Linguaglossa displays a more slowly decreasing correlation, (Figure 4.8(a)), overall the correlation is always less than 0.4, which suggests a weak temporal correlation structure.

As in other applications of MEVD-based formulations (e.g., rainfall or extreme coastal water level), the parameters of the ordinary value distribution (in this case a Gamma distribution) are here estimated using the Probability Weighted Moments (PWMs; Greenwood et al., 1979). The predictive estimation uncertainty associated with the MEVD, MEVD*, MEVD**, SMEV, and GEV (fitted on block maxima) methods is

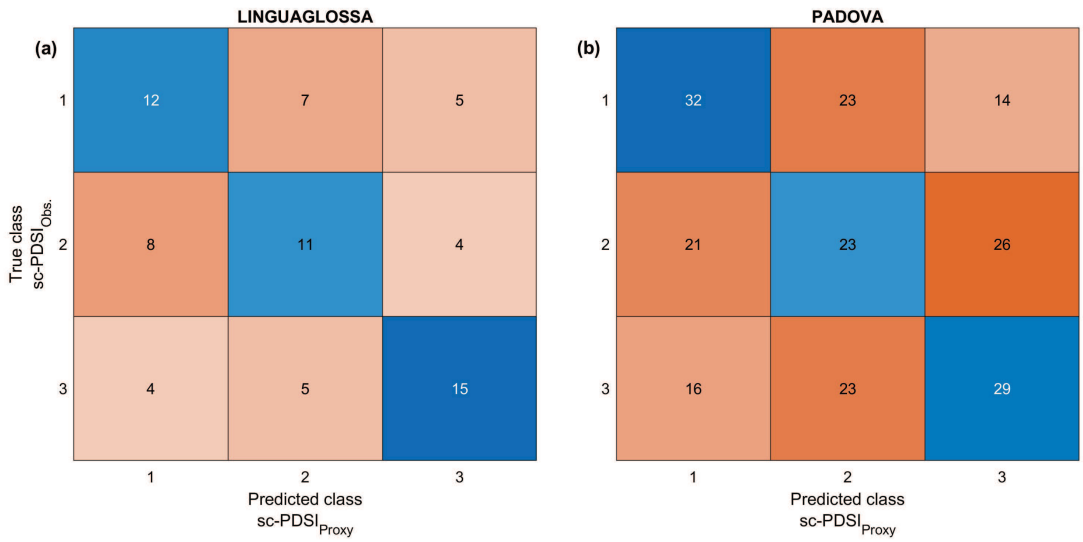


**Figure 4.7:** Contingency matrices for (**a**) Linguaglossa (CT, Italy) and (**b**) Padova (PD, Italy) sites.

evaluated here using the NDE (defined in Section 4.2.3) as a function of return period within the range of value of the latter accessible through the validation sub-sample, i.e. for $T_r = 1, \ldots, M - S + 1$. Additionally, the effect of different choices of calibration sample sizes between 3 and 5 blocks is explored. NDE values are summarized by computing their average across the cross-validation randomization procedure, along with their variability across different realizations represented by means of error bars (Figures 4.9 and 4.11) and violin plots (Figures 4.10 and 4.12). When we consider a calibration length of 5 blocks, we find somewhat different outcomes at the two study sites. When inter-block variability in the distribution of the ordinary events is accounted for (MEVD in green and MEVD** in purple colors) an underestimation of the empirical quantile is produced. On the other hand, when no variability in the ordinary-value distribution is allowed – i.e. a constant $F(x, \overrightarrow{\theta})$ is assumed – an overestimation tends to be produced. For the Padova site, and considering return periods less than 300 years (dashed line in Figure 4.9), MEVD-based estimation errors are close to zero and the small error standard deviation indicates a lower error variability (shorter "whiskers") with respect to the other approaches. Violin plots in Figure 4.10 suggest that: (1) the estimates obtained using all the MEVD-based approaches tested show a lower error variability (smaller uncertainty), and (2) the median error in MEVD estimates is close to zero, corresponding to approximately an unbiased estimates. For the lower return times, of most applicative interest in the present case, results suggest that the performance of the full-fledged MEVD is better than for the remaining approaches explored.

Results for the Linguaglossa sites (Figure 4.11, 4.12) show that also in this case the application of MEVD-based methods is advantageous when compared with the GEV formulation. However, in this case study the use of a formulation that neglects inter-

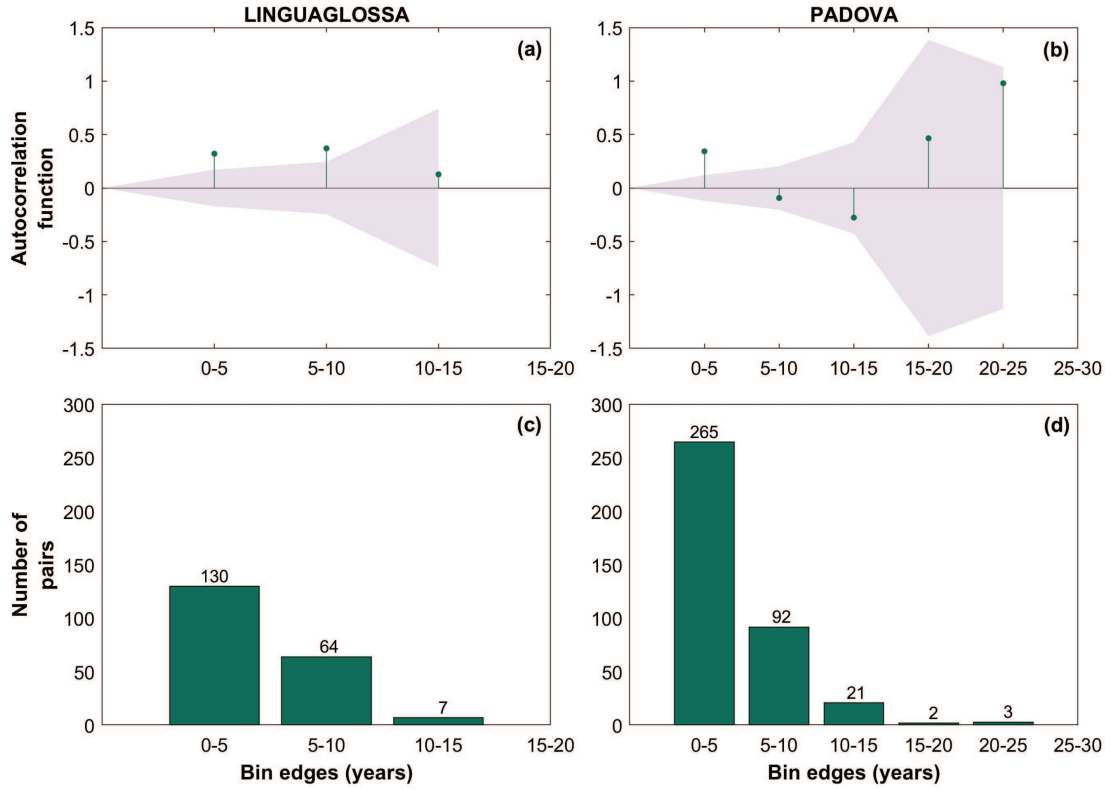


**Figure 4.8:** (**a**), (**b**) Autocorrelation function calculated by grouping the values into classes of fixed width (5 years). (**c**), (**d**) Histogram of the number of pairs in each class. Results are shown for the two case studies: (**a**),(**c**) Linguaglossa and (**b**),(**d**) Padova.

block variability (e.g., MEVD* and SMEV, namely in blue and yellow colors) seems to be more appropriate when $T_r$ is less than approximately 500 years. This can be explained by considering that, in this case, the probability distribution of ordinary events does not display a large inter-block variability. In other words, hydrometeorological regimes at Linguaglossa seem to display properties that are consistent over time, thus favoring methods that pool together data from all blocks when estimating parameter values.

Finally, a comparative analysis of the predictive estimations uncertainty considering a small number of blocks is performed, to explore the relevant case in which only a short time series is available. In particular, the situation in which just 3 blocks can be used for parameter estimation for both the MEVD-based approaches and the traditional GEV methods (Figures B.1–B.4). As expected, all formulations show a decreased performance compared to the case of $S = 5$ blocks. However, it may be seen that the GEV approach produces the largest estimation errors. For example, when Padova and the smallest value of $T_r$ are considered, GEV yields estimation errors of approximately –40 % (it was –0.06 % for $S = 5$ blocks). Stressing even more methods, e.g., considering $S = 2$ blocks, metastatistical formulations can be applied unlike the three-parameter GEV distribution is not applicable. The results of this limiting case

are reported in Appendix B (Figures B.5–B.8). Even though the estimates display high uncertainty, again approaches that neglect the inter-block variability are preferable for Linguaglossa.

Overall, the results suggest that no single approach is clearly superior for the entire range of return periods accessible through the validation sub-sample. However, MEVD-based estimates of extreme drought probability outperform, in terms of bias and uncertainty, traditional GEV estimates, particularly when the observational length is too short to apply conventional methods.

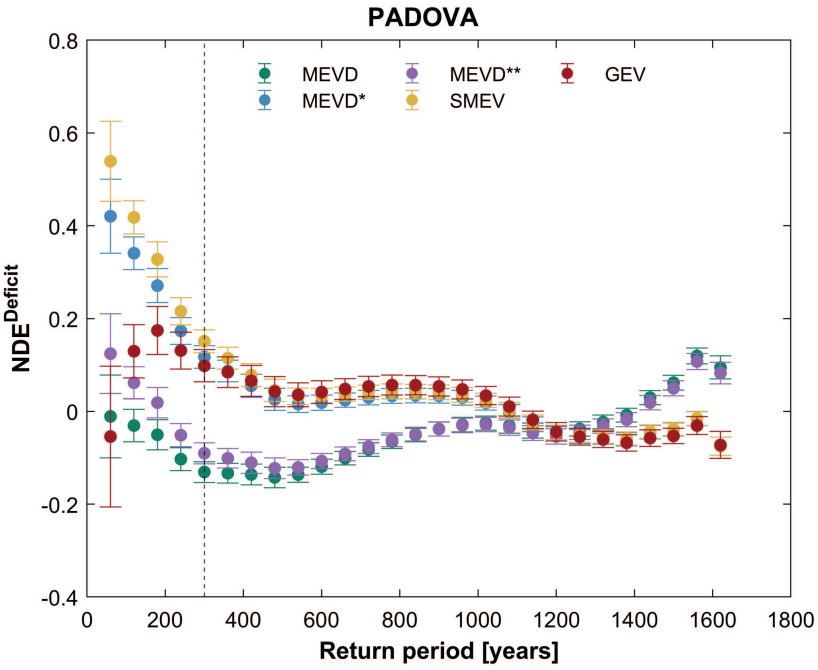


**Figure 4.9:** Error bars of the nondimensional estimation error (NDE) for the Padova site. Results are obtained for all the available return periods in the validation sub-sample and by using 5 blocks (each 60 years long) to estimate the parameters of the Gamma distribution. The symbols indicate the mean value (point) and the standard deviation (bar) of the NDE. Explanations of the acronyms are reported in the main text.

## 4.4 Conclusions

The results in this Chapter contribute to improving our ability to estimate drought return periods and provide estimates of the involved uncertainty. Notoriously, the analysis of the probabilistic structure of droughts is a non-trivial task, due to the limited number of drought events observed in the historical records of the relevant hydro-meteorological variables. Frequently, in practical applications, a standard inferential approach is not viable to suitably model drought characteristics.

To overcome these issues, this Chapter presents an at-site analysis of extreme meteorological droughts in two case studies in Italy by using the sc-PDSI reconstructed from

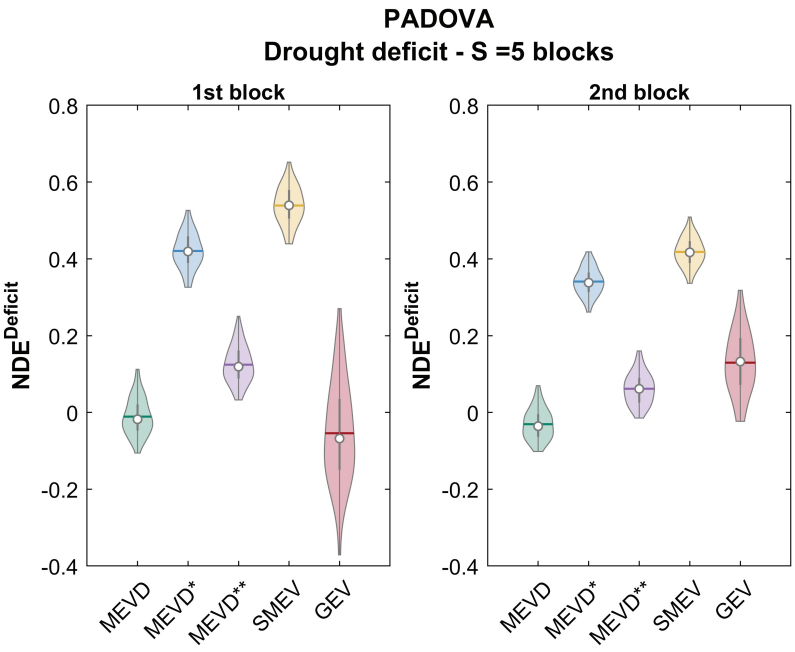


**Figure 4.10:** Violin plots of the nondimensional estimation error (NDE) for the Padova site. Panels report the results obtained for the first two return periods (indicated as 1st and 2nd block in fig.) in the validation sub-sample and by using 5 blocks (each 60 years long) to estimate the parameters of the Gamma distribution. Explanations of the acronyms are reported in the main text.

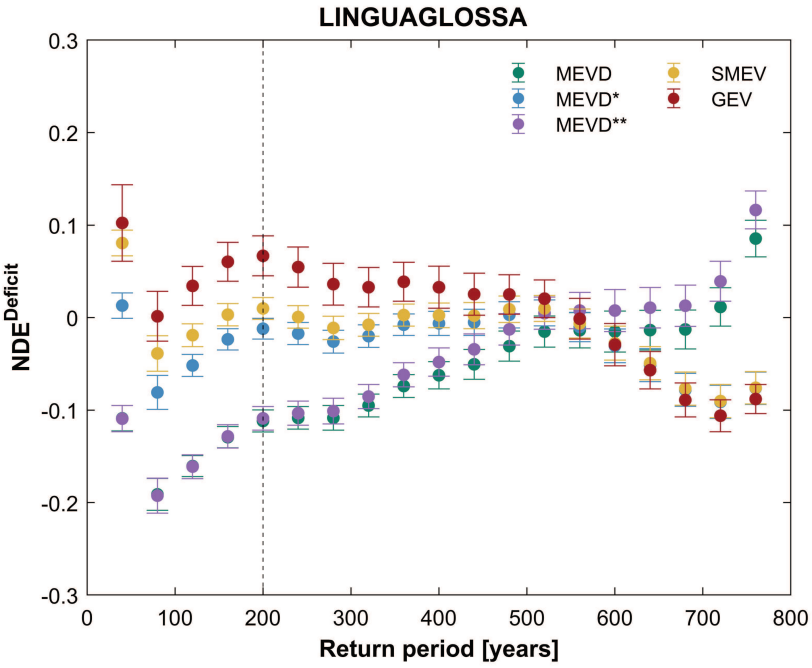


**Figure 4.11:** Error bars of the nondimensional estimation error (NDE) for the Linguaglossa site. Results are obtained for all the available return periods in the validation sub-sample and by using 5 blocks (each 40 years long) to estimate the parameters of the Gamma distribution. The symbols indicate the mean value (point) and the standard deviation (bar) of the NDE. Explanations of the acronyms are reported in the main text.

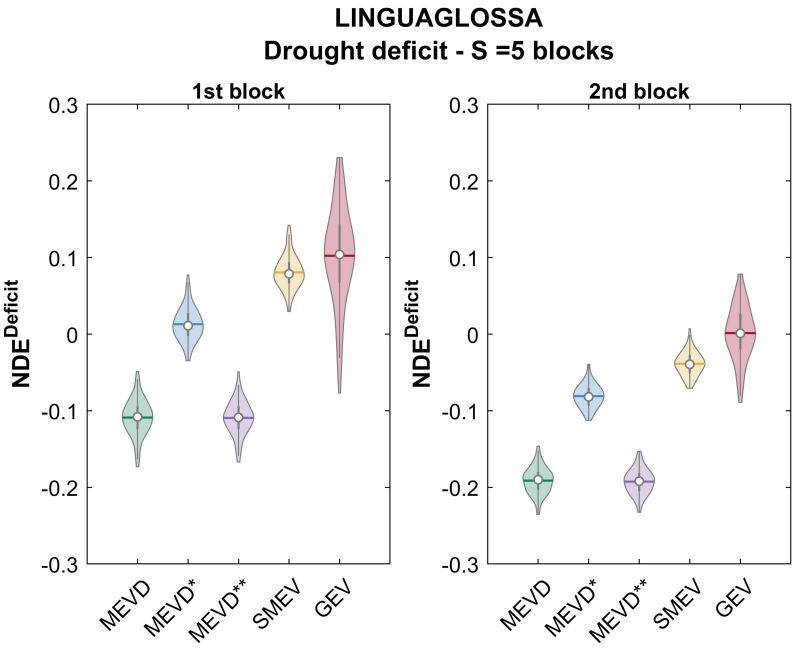


**Figure 4.12:** Violin plots of the nondimensional estimation error (NDE) for the Linguaglossa site. Panels report the results obtained for the first two return periods (indicated as 1st and 2nd block in fig.) in the validation sub-sample and by using 5 blocks (each 40 years long) to estimate the parameters of the Gamma distribution. Explanations of the acronyms are reported in the main text.

tree ring archives. To infer the probability of drought occurrences, non-asymptotic extreme-value formulations are used that allow for more efficient use of the available data. In particular, the MEVD, SMEV, and GEV-based performance estimates and their uncertainty are compared based on a cross-validation scheme. As recognized in several studies worldwide (e.g. Anchukaitis, 2017; O'Donnell et al., 2018), the analyses confirm that drought index reconstructions based on tree-ring data provide an opportunity to evaluate predictive uncertainty in extreme drought estimations. This Chapter also explores the use of three different versions of the MEVD that account for different degrees of inter-block variability, in the ordinary distribution parameters and in the number of events per block. In general, results indicate that MEVD-based formulations are more robust approaches with respect to traditional ones in applications to droughts. When comparing the two case studies, different results are found. This fact can be explained by their different hydro-meteorological regimes and the different long term variability at each site.

The comparative analysis of the predictive estimation uncertainty in extreme drought occurrences suggests that: (1) large inter-block variability in the probability distribution of drought properties may occur (e.g., see the Padova site), such that approaches that account for such variability outperform methods that assume the drought statistical structure to be constant, (2) conversely, using formulations that neglect inter-block

variability is more advantageous when such variability is less pronounces, such as in Linguaglossa. In this case, approaches that use all blocks to estimate distributional parameters yield a reduced estimation uncertainty.
Overall, this Chapter represents a contribution to the analysis of the probabilistic structure of droughts using non-asymptotic formulations. This application confirms the advantages and flexibility of these novel extreme value distributions also in the case of drought phenomena.
Recently, there has been growing interest in the accurate estimation of joint probabilities of drought for water management issues, monitoring, and early warning systems. Given that the environmental extreme phenomena are generally multivariate in nature, future work could focus on performing a multivariate frequency analysis considering the possible occurrence of different properties or types of drought. In this context, the metastatistical formulation based on the concept of "ordinary event" may provide advantages in high quantile estimation.

**Data availability.** All data set used are publicly available from data sources cited throughout this Chapter.
The OWDA data set is publicly available from the NOAA paleoclimate data base (`https://www.ncei.noaa.gov/access/paleo-search/study/19419`). Padova (Italy) climate data are from Marani and Zanetti (2015); Camuffo (2002a,b,c); Cocheo and Camuffo (2002). Linguaglossa (Italy) climate time series are available from the Water Observatory of Sicily Region, formerly the Sicilian Regional Hydrographic Service, whose data are available at the website of the Servizio Informativo Agrometeorologico Siciliano (SIAS; `http://www.sias.regione.sicilia.it/frameset_dati.htm`. The annual total ring-width indices used in this application are available from the International Tree-Ring Data Bank (ITRDB, `https://www.ncei.noaa.gov/products/paleoclimatology/tree-ring`).

# Chapter 5

# Summary and conclusions

This dissertation is a contribution to the statistical description of extremes using non-asymptotic approaches based on the Metastatistical Extreme Value Distribution, MEVD, and its simplified versions, SMEV, or Simplified Metastatistical Extreme Value distribution. The objective is to apply the MEVD framework to practical extreme-value problems requiring an optimized use of the observational information to estimate large extremes based on small samples.

For the first time, MEVD models have been applied to characterize the extreme distribution of coastal flooding events and drought occurrence. Even though these phenomena are different from a physical process perspective, in both cases the results show the advantages and flexibility afforded by these novel extreme value distributions.

The MEVD by construction is capable of capturing the stochastic inter-annual variability of process dynamics by multiple block-based parametrizations. The analysis has shown that a key parameter to reliably estimate high quantiles is the block size through which the extreme value distribution is defined. In particular, this choice defines the sample length for parameter estimation and, hence, strongly controls estimation uncertainty. For this reason, block length must ensure that a sufficient average number of events fall within each block. Hence, the choice of block size must be dependent on the phenomenon under analysis and on local climatic conditions.

In coastal flooding problem the analyses were focused on extreme values of the total coastal water level, given by the sum between the water level setup induced by meteorological forcing and the astronomical tide. The results highlight that the MEVD approach provides reliable estimates of high quantiles for almost all the gauge stations explored, particularly when sufficiently long calibration sample sizes are considered. Differences in performance between the MEVD framework and GEV-based approaches are subtle, and a definitive conclusion on an optimal solution independent of the return period of interest remains elusive. If the focus is on high-return-period quantiles estimation, the analyses show that the MEVD approach provides reliable estimates for almost all the gauge stations explored. Data from the Marseille gauge station exhibit

a behavior that deviates from those from other sites, showing an inferior predictive performance of the MEVD with respect to GEV-based approaches. The cause is to be found in the small average number of sea-level peaks every year. In this case, the small sample of yearly ordinary events available prevents the MEVD from adding significant information with respect to GEV-based approaches. Conversely, when methods performance is evaluated for intermediate return period values, GEV based on the block maxima method displays an overall greater robustness, and MEVD exhibits a greater absolute value of the estimation error. Finally, the sensitivity of extreme-water-level frequency to sea-level rise was explored. The results are location-dependent and show that, at a given site and for a set value of the total water level extreme, the relative change in return time grows linearly with return time itself. In fact, changes in return levels are nonlinear way. The results suggest that relative changes are more significant for smaller extremes than for larger ones. By comparing the percentage changes associated with the two climate change scenarios and the two return periods considered, Venice and Marseille are seen to experience the greatest changes in extreme total water levels. All sites display greater percent changes for the lower 100-year return period in each scenario, i.e. "more frequent" extremes will be most impacted by sea-level changes in the near future.

Focusing on drought occurrence, it is widely recognized that the study of the probabilistic structure of droughts suffers from the lack of sufficiently long observations and consolidated methods of analysis. Since direct hydrological observations do not provide a sufficient amount of data, the analysis was here focused on the use of reconstructions of Palmer's drought index based on proxy paleoclimate data. An important point addressed here regards the identification of a relationship between tree ring-derived index and local hydrologic balance. Using several correlation measures, we conclude that the use of tree-ring based sc-PDSI analyses may provide insight into the statistical properties of droughts that are similar to those that may be obtained from PDSI values computed using direct observations. Drought events were identified by using Run Theory and the effect of different choices of calibration sample sizes was explored. The MEVD was applied according to three formulations: (1) the full *MEVD* formulation, in which both the ordinary-value distribution and the number of events in each block are allowed to vary from one block to another, (2) a formulation denoted as *MEVD**, in which the distribution of the ordinary events is constant across different blocks (i.e. $F(x; \overrightarrow{\theta}) = \text{constant}$), while the number of events per block is allowed to vary, and (3) a *MEVD*** formulation, in which the probability distribution of ordinary events is allowed to vary, while the number of events per block is taken to be constant and equal to the overall average across all blocks. The comparative analysis of the predictive estimation uncertainty in extreme drought occurrences suggests that: (1) large inter-block variability in the probability distribution of drought properties may occur, such that approaches that account for this variability outperform methods that

assume the drought statistical structure to be constant, (2) conversely, using formulations that neglect inter-block variability is more advantageous when such variability is less pronounces. In other words, when the hydrometeorological regimes display properties that are consistent over time, thus favoring methods that pool together data from all blocks when estimating parameter values. In this case, approaches that use all blocks to estimate distributional parameters yield a reduced estimation uncertainty. Finally, the predictive uncertainty associated with MEVD-based approaches compares favorably with that characterizing traditional GEV-based methods. Importantly, for realistic cases in which only relatively short calibration samples were used, MEVD-based approaches allowed the estimation of high quantiles, even though with large (but quantified) uncertainty. Block-maxima GEV methods could not in these cases be applied because of a lack of sufficient data for parameter estimation.

Clearly, further improvements and developments to the present dissertation may be envisioned and can be summarized as follows.

(1) For coastal flooding, the size of the available data set does not allow the exploration of model predictive uncertainty for large values of the return period. Future work could investigate whether the estimation error can be further reduced by using alternative assumptions, e.g. by assuming "time-invariant" parameters in the ordinary-event distribution (SMEV). With such an approach distributional parameters can be estimated based on the entire calibration data set, rather than on relatively short sliding windows.

(2) For drought analysis, future work may explore a mixed-distribution approach, in which different populations of ordinary drought events are modeled. The MEVD could then be built by studying the rate of occurrence of the different components of the mixture.

# Appendix A

# Extreme-coastal-water levels

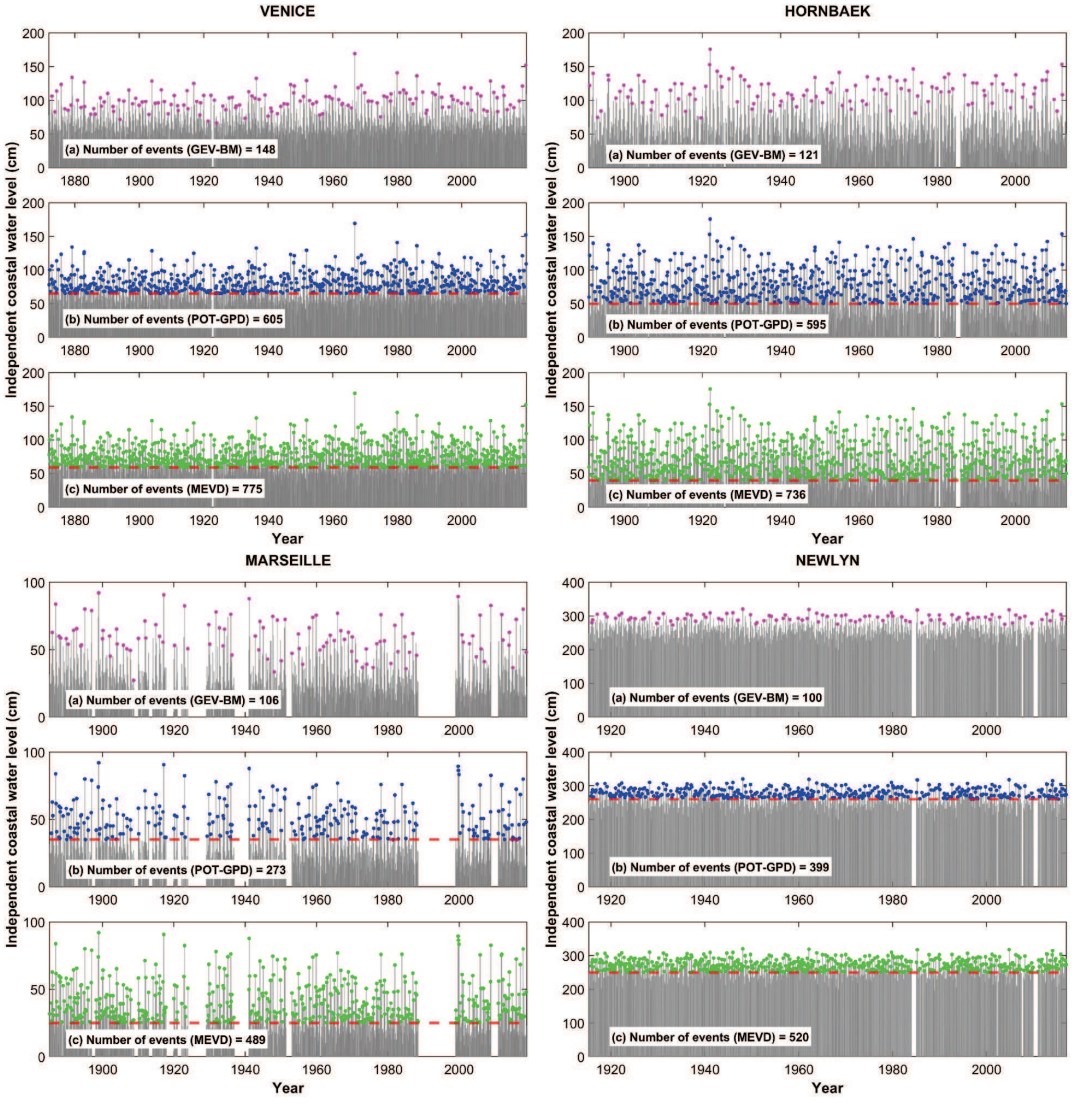


**Figure A.1:** Independent coastal water levels (gray line) for the Venice (IT), Hornbæk (DK), Marseille (FR) and Newlyn (UK) sites, and events on which the three approaches are fitted: (a) magenta dots show the annual maxima used for the GEV–BM method, (b) blue dots represent the exceedances over a threshold in the case of the POT–GPD approach, and (c) green dots display the ordinary values for the MEVD framework.

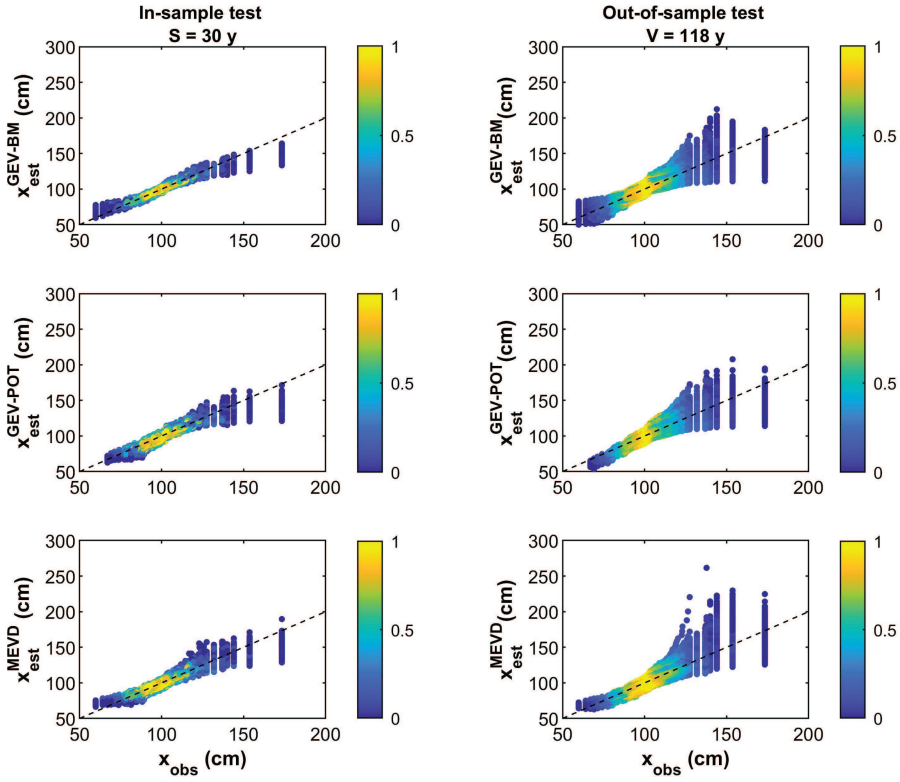


**Figure A.2: VENICE (IT)** – QQ plots of extreme coastal water level quantiles computed for the GEV–based approaches (BM and POT) and MEVD for the Venice station. The plots are obtained as a result of the cross-validation method used to test the global performance of the models and are estimated for 1000 random realizations and for sample size: a) $S = 30$ years (in-sample-test in the left column); b) $V = M - S$ years (out-of-sample test in the right column). The colors represent the points density around the 45° line (black dashed line) corresponding to the best fit.

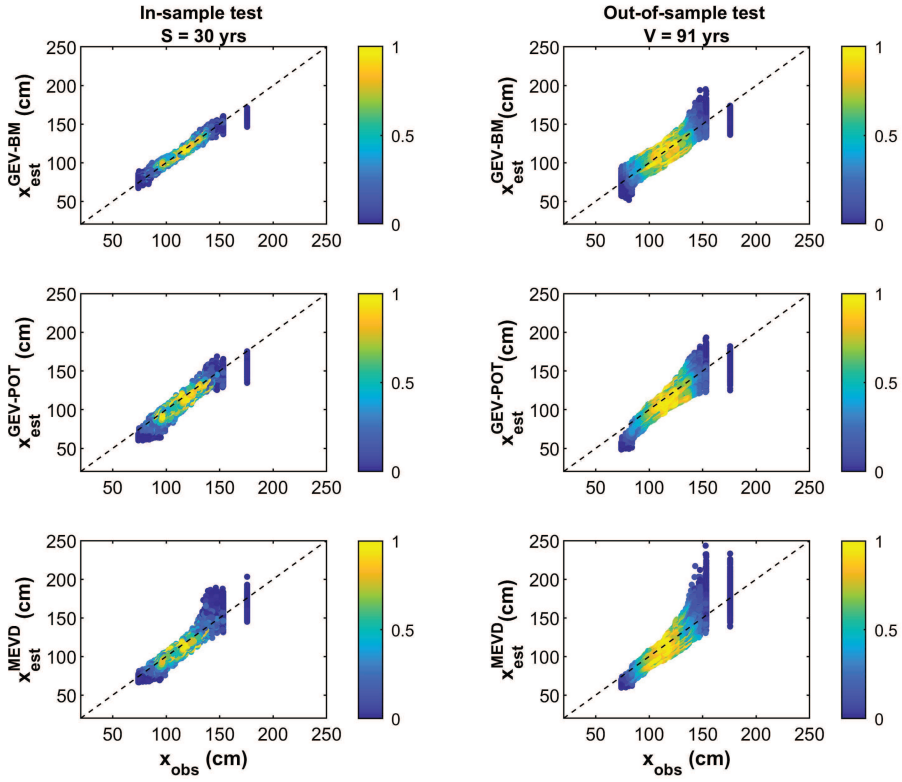


**Figure A.3: HORNBÆK (DK)** – QQ plots of extreme coastal water level quantiles computed for the GEV–based approaches (BM and POT) and MEVD for the Venice station. The plots are obtained as a result of the cross-validation method used to test the global performance of the models and are estimated for 1000 random realizations and for sample size: a) $S = 30$ years (in-sample-test in the left column); b) $V = M - S$ years (out-of-sample test in the right column). The colors represent the point density around the 45° line (black dashed line) corresponding to the best fit.

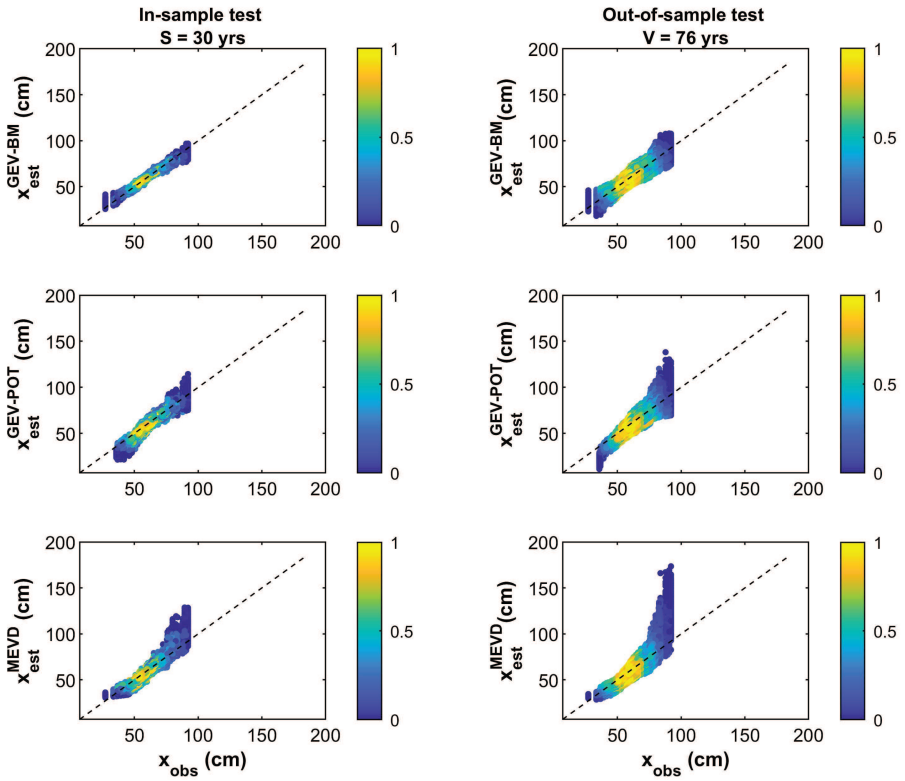


**Figure A.4: MARSEILLE (FR)** – QQ plots of extreme coastal water level quantiles computed for the GEV–based approaches (BM and POT) and MEVD (with parameter estimation *on non-overlapping sub-samples of fixed size (5 years)*) for the Marseille station. The plots are obtained as a result of the cross-validation method used to test the global performance of the models and are estimated for 1000 random realizations and for sample size: a) $S = 30$ years (in-sample-test in the left column); b) $V = M - S$ years (out-of-sample test in the right column). The colors represent the point density around the 45° line (black dashed line) corresponding to the best fit.

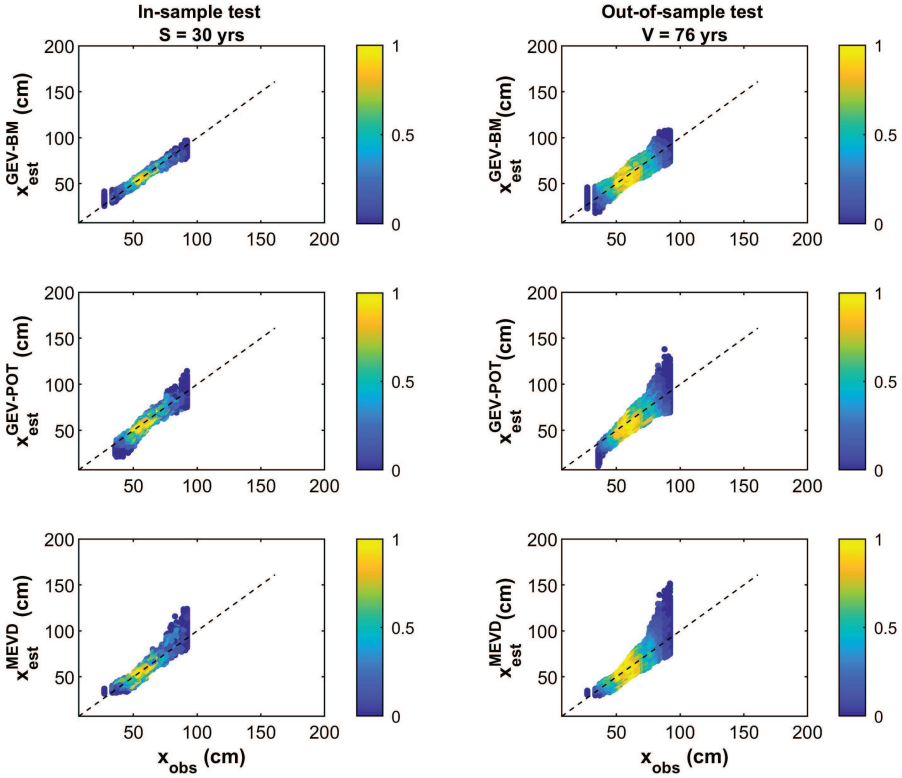


**Figure A.5: MARSEILLE (FR)** – QQ plots of extreme coastal water level quantiles computed for the GEV–based approaches (BM and POT) and MEVD (with parameter estimation *on data from the whole calibration sample*) for the Marseille station. The plots are obtained as a result of the cross-validation method used to test the global performance of the models and are estimated for 1000 random realizations and for sample size: a) $S = 30$ years (in-sample-test in the left column); b) $V = M - S$ years (out-of-sample test in the right column). The colors represent the point density around the 45° line (black dashed line) corresponding to the best fit.

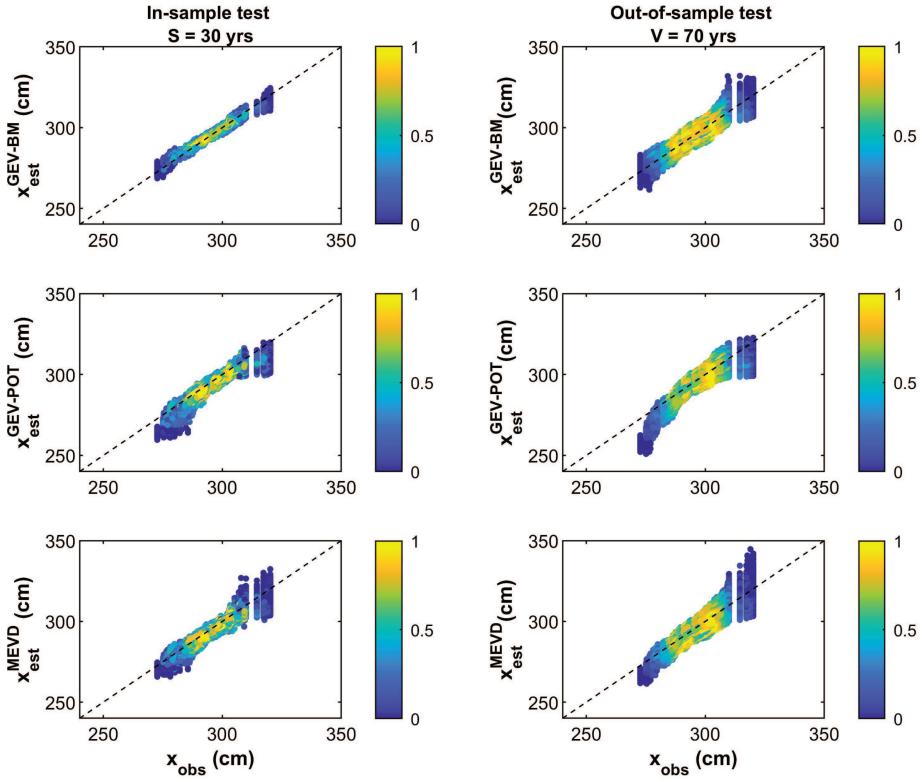


**Figure A.6: NEWLYN (UK)** – QQ plots of extreme coastal water level quantiles computed for the GEV–based approaches (BM and POT) and MEVD for the Venice station. The plots are obtained as a result of the cross validation method used to test the global performance of the models and are estimated for 1000 random realizations and for sample size: a) $S = 30$ years (in-sample-test in the left column); b) $V = M - S$ years (out-of-sample test in the right column). The colors represent the point density around the 45° line (black dashed line) corresponding to the best fit.

## Appendix B

# Extreme meteorological droughts

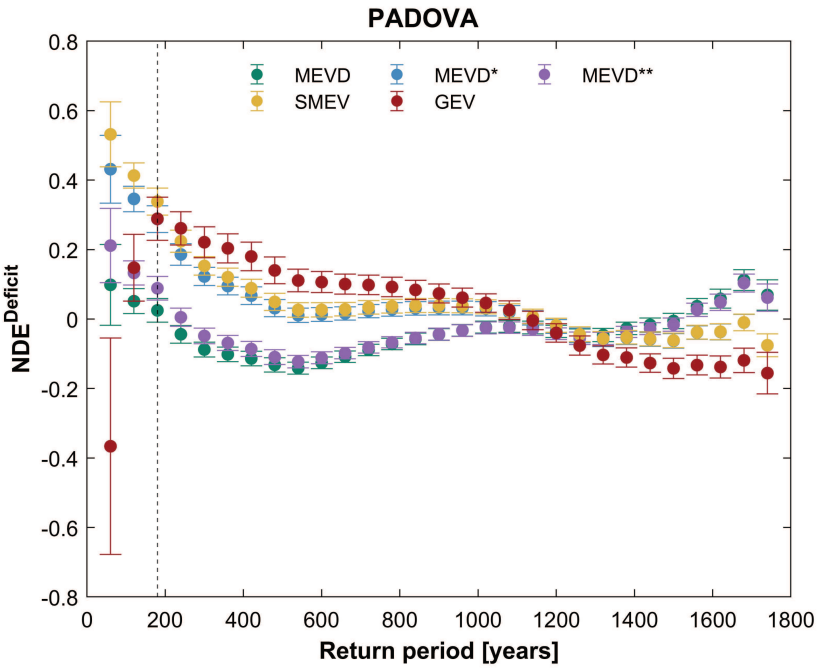


**Figure B.1:** Error bars of the nondimensional estimation error (NDE) for the Padova site. Results are obtained for all the available return periods in the validation sub-sample and by using **3** blocks (each 60 years long) to estimate the parameters of the Gamma distribution. The symbols indicate the mean value (point) and the standard deviation (bar) of the NDE. Explanations of the acronyms are reported in the main text.

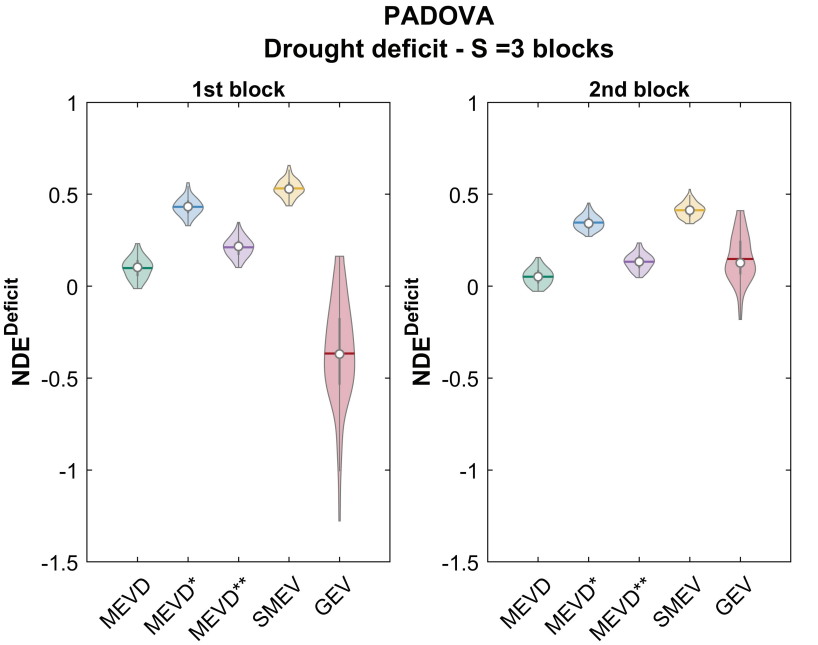


**Figure B.2:** Violin plots of the nondimensional estimation error (NDE) for the Padova site. Panels report the results obtained for the first two return periods (indicated as 1st and 2nd block in fig.) into the validation sub-sample and by using **3** blocks (each 60 years long) to estimate the parameters of the Gamma distribution. Explanations of the acronyms are reported in the main text.

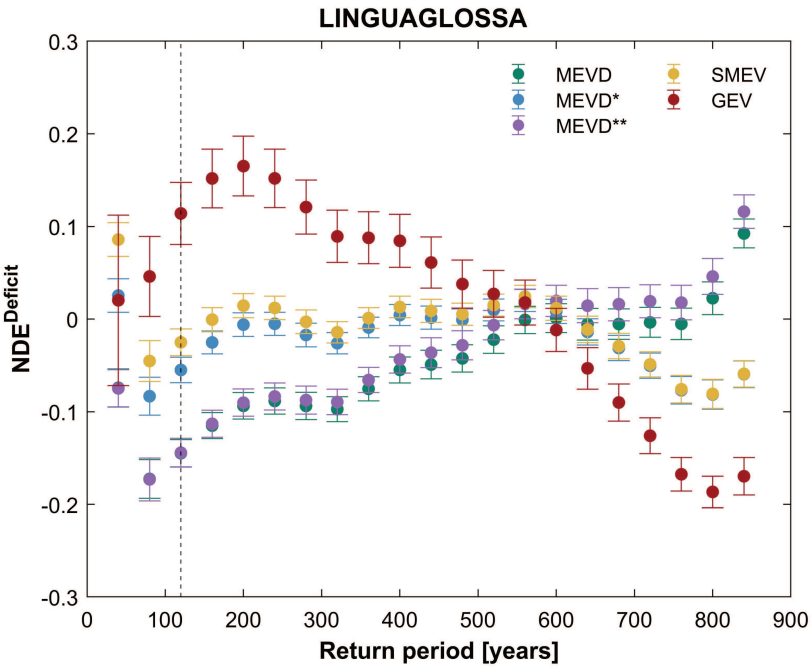


**Figure B.3:** Error bars of the nondimensional estimation error (NDE) for the Linguaglossa site. Results are obtained for all the available return periods in the validation sub-sample and by using **3** blocks (each 40 years long) to estimate the parameters of the Gamma distribution. The symbols indicate the mean value (point) and the standard deviation (bar) of the NDE. Explanations of the acronyms are reported in the main text.

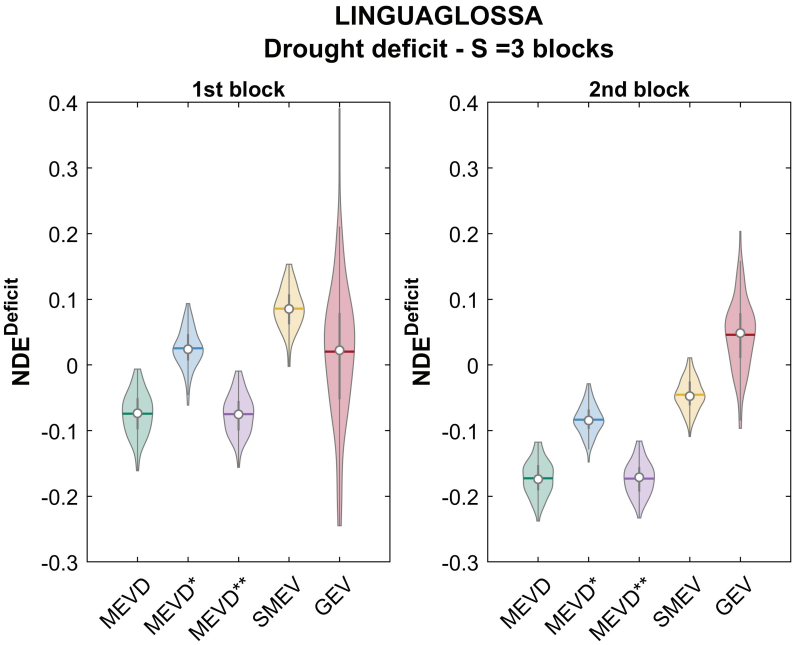


**Figure B.4:** Violin plots of the nondimensional estimation error (NDE) for the Linguaglossa site. Panels report the results obtained for the first two return periods (indicated as 1st and 2nd block in fig.) into the validation sub-sample and by using **3** blocks (each 40 years long) to estimate the parameters of the Gamma distribution. Explanations of the acronyms are reported in the main text.

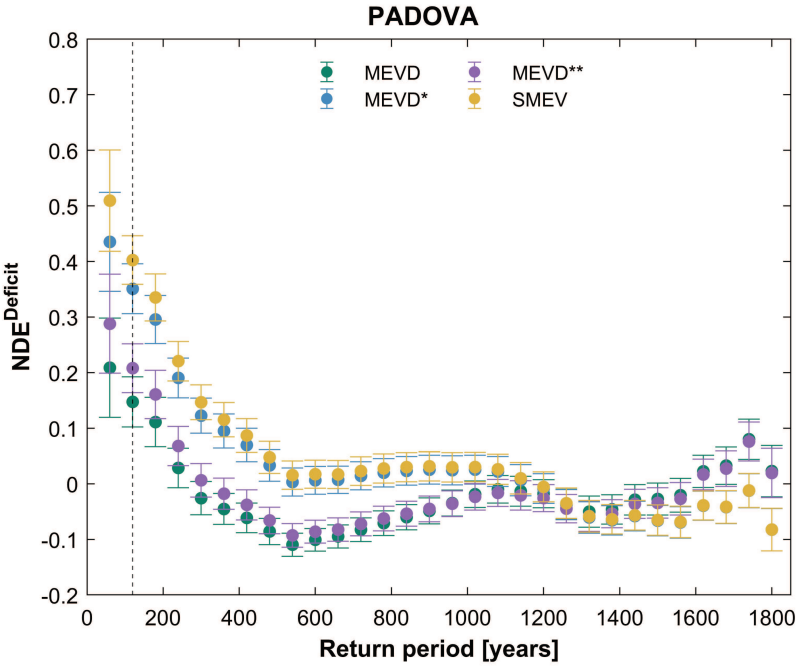


**Figure B.5:** Error bars of the nondimensional estimation error (NDE) for the Padova site. Results are obtained for all the available return periods in the validation sub-sample and by using **2** blocks (each 60 years long) to estimate the parameters of the Gamma distribution. The symbols indicate the mean value (point) and the standard deviation (bar) of the NDE. Explanations of the acronyms are reported in the main text.

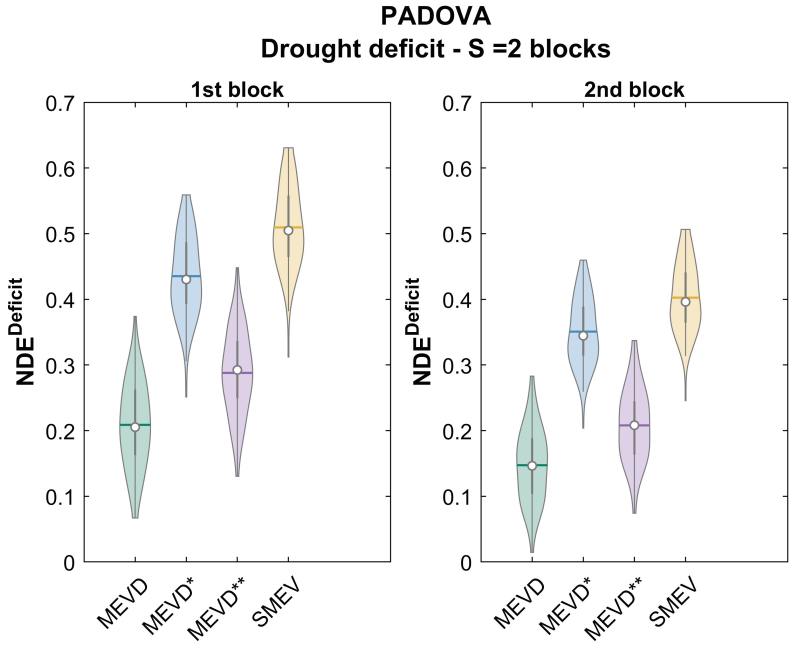


**Figure B.6:** Violin plots of the nondimensional estimation error (NDE) for the Padova site. Panels report the results obtained for the first two return periods (indicated as 1st and 2nd block in fig.) into the validation sub-sample and by using **2** blocks (each 60 years long) to estimate the parameters of the Gamma distribution. Explanations of the acronyms are reported in the main text.

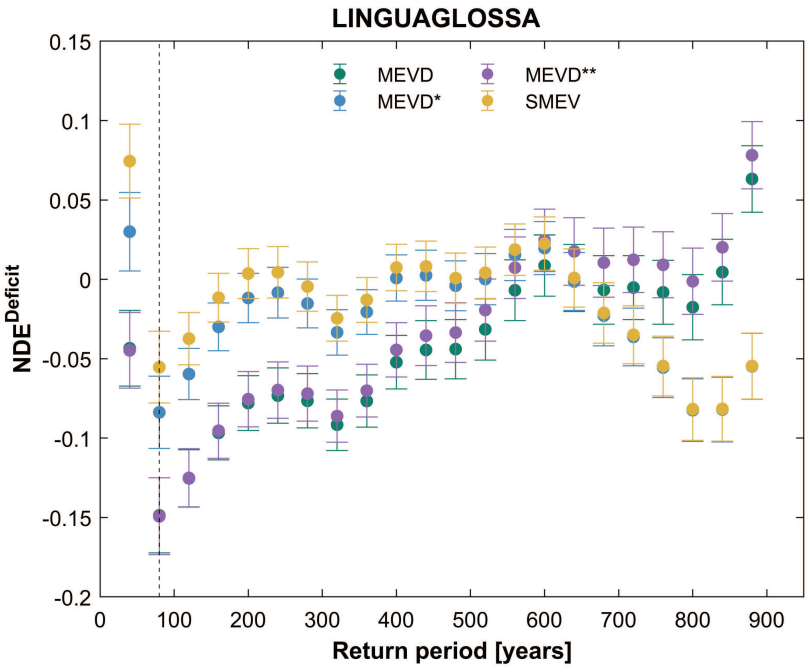


**Figure B.7:** Error bars of the nondimensional estimation error (NDE) for the Linguaglossa site. Results are obtained for all the available return periods in the validation sub-sample and by using **2** blocks (each 40 years long) to estimate the parameters of the Gamma distribution. The symbols indicate the mean value (point) and the standard deviation (bar) of the NDE. Explanations of the acronyms are reported in the main text.

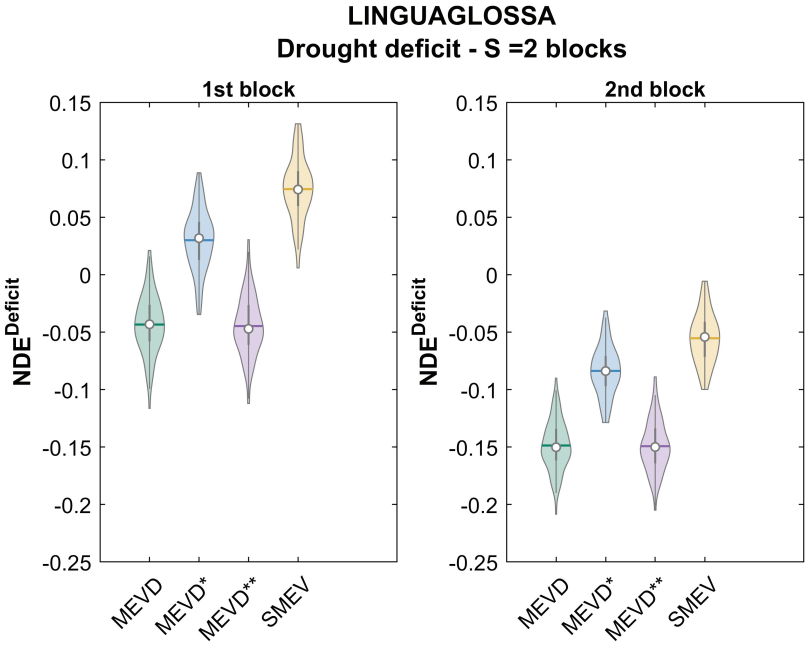


**Figure B.8:** Violin plots of the nondimensional estimation error (NDE) for the Linguaglossa site. Panels report the results obtained for the first two return periods (indicated as 1st and 2nd block in fig.) into the validation sub-sample and by using **2** blocks (each 40 years long) to estimate the parameters of the Gamma distribution. Explanations of the acronyms are reported in the main text.

# References


Aghakouchak, A., Feldman, D., Stewardson, M. J., Saphores, J.-D., Grant, S., and Sanders, B.: Australia's Drought: Lessons for California, Science, 343, 1430–1431, https://doi.org/10.1126/science.343.6178.1430, 2014.

Allan, R. P., Barlow, M., Byrne, M. P., Cherchi, A., Douville, H., Fowler, H. J., Gan, T. Y., Pendergrass, A. G., Rosenfeld, D., Swann, A. L. S., Wilcox, L. J., and Zolina, O.: Advances in understanding large-scale responses of the water cycle to climate change, Annals of the New York Academy of Sciences, 1472, 49–75, https://doi.org/10.1111/nyas.14337, 2020.

Alley, W. M.: The Palmer Drought Severity Index: Limitations and Assumptions, Journal of Applied Meteorology and Climatology, 23, 1100 – 1109, https://doi.org/10.1175/1520-0450(1984)023⟨1100:TPDSIL⟩2.0.CO;2, 1984.

Almar, R., Ranasinghe, R., Bergsma, E. W. J., Diaz, H., Melet, A., Papa, F., Vousdoukas, M., Athanasiou, P., Dada, O., Almeida, L. P., and Kestenare, E.: A global analysis of extreme coastal water levels with implications for potential coastal overtopping, Nature Communications, 12, 3775, https://doi.org/10.1038/s41467-021-24008-9, 2021.

Amponsah, W., Dallan, E., Nikolopoulos, E. I., and Marra, F.: Climatic and altitudinal controls on rainfall extremes and their temporal changes in data-sparse tropical regions, Journal of Hydrology, 612, 128 090, https://doi.org/10.1016/j.jhydrol.2022.128090, 2022.

Anchukaitis, K. J.: Tree Rings Reveal Climate Change Past, Present, and Future, Proceedings of the American Philosophical Society, 161, 244–263, URL `http://www.jstor.org/stable/45211559`, 2017.

Araujo, D. S. A., Marra, F., Merow, C., and Nikolopoulos, E. I.: Today's 100 year droughts in Australia may become the norm by the end of the century, Environ. Res. Lett., 17, 044 034, https://doi.org/10.1088/1748-9326/ac58ac, 2022.

Araújo, I. B. and Pugh, D. T.: Sea levels at Newlyn, 1915–2005: Analysis of trends for

future flooding risks, J. Coastal Res., 24, 03–212, https://doi.org/10.2112/060785.1, 2008.

Bakke, S. J., Ionita, M., and Tallaksen, L. M.: The 2018 northern European hydrological drought and its drivers in a historical perspective, Hydrology and Earth System Sciences, 24, 5621–5653, https://doi.org/10.5194/hess-24-5621-2020, 2020.

Balkema, A. A. and de Haan, L.: Residual life time at great age, Ann. Probab., 2, 792–804, https://doi.org/10.1214/aop/1176996548, 1974.

Barbariol, F., Bidlot, J.-R., Cavaleri, L., Sclavo, M., Thomson, J., and Benetazzo, A.: Maximum wave heights from global model reanalysis, Prog. Oceanogr., 175, 139–160, https://doi.org/10.1016/j.pocean.2019.03.009, 2019.

Bateni, M. M., Martina, M. L. V., and Arosio, M.: Multivariate return period for different types of flooding in city of Monza, Italy, Natural Hazards, 114, 811–823, https://doi.org/10.1007/s11069-022-05413-9, 2022.

Beck, C. and Cohen, E. G. D.: Superstatistics, Physica A: Statistical Mechanics and its Applications, 322, 267–275, https://doi.org/10.1016/S0378-4371(03)00019-0, 2003.

Beirlant, J., Goegebeur, Y., Segers, J. J. J., and Teugels, J.: Statistics of Extremes: Theory and Applications, John Wiley & Sons, Chichester, UK, 2004.

Benetazzo, A., Ardhuin, F., Bergamasco, F., Cavaleri, L., Guimarães, P. V., Schwendeman, M., Sclavo, M., Thomson, J., and Torsello, A.: On the shape and likelihood of oceanic rogue waves, Sci. Rep, 7, 1–11, https://doi.org/10.1038/s41598-017-07704-9, 2017.

Bernardara, P., Andreewsky, M., and Benoit, M.: Application of regional frequency analysis to the estimation of extreme storm surges, J. Geophys. Res.-Oceans, 116, https://doi.org/10.1029/2010JC006229, 2011.

Bernardara, P., Mazas, F., Kergadallan, X., and Hamm, L.: A two-step framework for over-threshold modelling of environmental extremes, Nat. Hazard Earth Sys., 14, 635–647, https://doi.org/10.5194/nhess-14-635-2014, 2014.

Bernier, N. B. and Thompson, K. R.: Tide-surge interaction off the east coast of Canada and northeastern United States, J. Geophys. Res.-Oceans, 112, https://doi.org/10.1029/2006JC003793, 2007.

Bommier, E.: Peaks-over-threshold modelling of environmental data, Master's thesis, Department of Mathematics, Uppsala University, 2014.

Bonaccorso, B., Cancelliere, A., and Rossi, G.: An analytical formulation of return period of drought severity, Stochastic Environmental Research and Risk Assessment, 17, 157–174, https://doi.org/10.1007/s00477-003-0127-7, 2003.

Bonaccorso, B., Cancelliere, A., and Rossi, G.: Probabilistic forecasting of drought class transitions in Sicily (Italy) using Standardized Precipitation Index and North Atlantic Oscillation Index, Journal of Hydrology, 526, 136–150, https://doi.org/10.1016/j.jhydrol.2015.01.070, 2015.

Bouras, E., Jarlan, L., Khabba, S., Er-Raki, S., Dezetter, A., Sghir, F., and Tramblay, Y.: Assessing the impact of global climate changes on irrigated wheat yields and water requirements in a semi-arid environment of Morocco, Scientific Reports, 9, 19 142, https://doi.org/10.1038/s41598-019-55251-2, 2019.

Briffa, K. R., Jones, P. D., Wigley, T. M. L., Pilcher, J. R., and Baillie, M. G. L.: Climate reconstruction from tree rings: Part 2, spatial reconstruction of summer mean sea-level pressure patterns over Great Britain, Journal of Climatology, 6, 1–15, https://doi.org/10.1002/joc.3370060102, 1986.

Briffa, K. R., Osborn, T. J., Schweingruber, F. H., Jones, P. D., Shiyatov, S. G., and Vaganov, E. A.: Tree-ring width and density data around the Northern Hemisphere: Part 1, local and regional climate signals, The Holocene, 12, 737–757, https://doi.org/10.1191/0959683602hl587rp, 2002.

Brunner, M. I., Liechti, K., and Zappa, M.: Extremeness of recent drought events in Switzerland: dependence on variable and return period choice, Natural Hazards and Earth System Sciences, 19, 2311–2323, https://doi.org/10.5194/nhess-19-2311-2019, 2019.

Cai, W., Purich, A., Cowan, T., van Rensch, P., and Weller, E.: Did Climate Change–Induced Rainfall Trends Contribute to the Australian Millennium Drought?, Journal of Climate, 27, 3145–3168, 2014.

Caldwell, P. C., Merrifield, M. A., and Thompson, P. R.: Sea level measured by tide gauges from global oceans – the Joint Archive for Sea Level holdings (NCEI Accession 0019568), Version 5.5, NOAA National Centers for Environmental Information [data set], URL `https://doi.org/10.7289/V5V40S7W`, 2015.

Camuffo, D.: History of the Long Series of Daily Air Temperature in Padova (1725–1998), In: Camuffo, D., Jones, P. D. (eds), Improved Understanding of Past Climatic Variability from Early Daily European Instrumental Sources, pp. 7–75, Springer Netherlands, Dordrecht, https://doi.org/10.1007/978-94-010-0371-1_2, 2002a.

Camuffo, D.: Calibration and Instrumental Errors in Early Measurements of Air Temperature, In: Camuffo, D., Jones, P. D. (eds), Improved Understanding of Past Climatic Variability from Early Daily European Instrumental Sources, pp. 297–329, Springer Netherlands, Dordrecht, https://doi.org/10.1007/978-94-010-0371-1_11, 2002b.

Camuffo, D.: Errors in Early Temperature Series Arising from Changes in Style of Measuring Time, Sampling Schedule and Number of Observations, Climatic Change, 53, 331–352, https://doi.org/10.1023/A:1014962623762, 2002c.

Cancelliere, A.: Non Stationary Analysis of Extreme Events, Water Resources Management, 31, 3097–3110, https://doi.org/10.1007/s11269-017-1724-4, 2017.

Caruso, M. F. and Marani, M.: Extreme-coastal-water-level estimation and projection: a comparison of statistical methods, Natural Hazards and Earth System Sciences, 22, 1109–1128, https://doi.org/10.5194/nhess-22-1109-2022, 2022.

Castillo, E., Hadi, A. S., Balakrishnan, N., and Sarabia, J. M.: Extreme Value and Related Models in Engineering and Science Applications, John Wiley & Sons, New York, 2005.

Chan, S., Chu, J., Zhang, Y., and Nadarajah, S.: An extreme value analysis of the tail relationships between returns and volumes for high frequency cryptocurrencies, Research in International Business and Finance, 59, 101 541, https://doi.org/10.1016/j.ribaf.2021.101541, 2022.

Chiu, Y., Chebana, F., Abdous, B., Bélanger, D., and Gosselin, P.: Mortality and morbidity peaks modeling: an extreme value theory approach, Stat. Methods Med. Res., 27, 1498–1512, https://doi.org/10.1177/0962280216662494, 2018.

Church, J. A. and White, N. J.: A 20th century acceleration in global sea-level rise, Surv. Geophys., 33, https://doi.org/10.1029/2005GL024826, 2006.

Church, J. A. and White, N. J.: Sea-Level Rise from the Late 19th to the Early 21st Century, Surv. Geophys., 32, https://doi.org/10.1007/s10712-011-9119-1, 2011.

Church, J. A., Clark, P. U., Cazenave, A., Gregory, J. M., Jevrejeva, S., Levermann, A., Merrifield, M. A., Milne, G. A., Nerem, R. S., Nunn, P. D., Payne, A. J., Pfeffer, W. T., Stammer, D., and Unnikrishnan, A. S.: Sea Level Change, in: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, edited by Stocker, T., Qin, D., Plattner, G.-K., Tignor, M., Allen, S., Boschung, J., Nauels, A., Xia, Y., Bex, V., and Midgley, P., chap. 13, pp. 1137–1216, Cambridge University Press, URL `https://www.ipcc.ch/report/ar5/wg1/`, 2013.

Cid, A., Menéndez, M., Castanedo, S., Abascal, A. J., Méndez, F. J., and Medina, R.: Long-term changes in the frequency, intensity and duration of extreme storm surge events in southern Europe, Clim. Dynam., 46, 1503–1516, https://doi.org/10.1007/s00382-015-2659-1, 2015.

Cipollini, S., Fiori, A., and Volpi, E.: Structure-based framework for the design and risk assessment of hydraulic structures, with application to offline flood detention basins, Journal of Hydrology, 600, 126 527, https://doi.org/10.1016/j.jhydrol.2021.126527, 2021.

Città di Venezia: Centro Previsioni e Segnalazioni Maree, URL `https://www.comune.venezia.it/it/content/centro-previsioni-e-segnalazioni-maree`, 2020.

Cocheo, C. and Camuffo, D.: Corrections of Systematic Errors and Data Homogenisation in the Daily Temperature Padova Series (1725–1998), Climatic Change, 53, 77–100, https://doi.org/10.1023/A:1014950306015, 2002.

Coles, S.: An Introduction to Statistical Modeling of Extreme Values, Springer, London, https://doi.org/10.1007/978-1-4471-3675-0, 2001.

Coles, S. and Tawn, J.: Bayesian modelling of extreme surges on the UK east coast, Philos. T. Roy. Soc. A, 363, 1387–1406, https://doi.org/10.1098/rsta.2005.1574, 2005.

Commission, E. and Centre, J. R.: Drought in Europe : July 2022 : GDO analytical report, Publications Office of the European Union, https://doi.org/doi/10.2760/014884, 2022.

Cook, B. I., Smerdon, J. E., Seager, R., and Coats, S.: Global warming and 21st century drying, Climate Dynamics, 43, 2607–2627, https://doi.org/10.1007/s00382-014-2075-y, 2014.

Cook, B. I., Palmer, J. G., Cook, E. R., Turney, C. S. M., Allen, K., Fenwick, P., O'Donnell, A., Lough, J. M., Grierson, P. F., Ho, M., and Baker, P. J.: The paleoclimate context and future trajectory of extreme summer hydroclimate in eastern Australia, Journal of Geophysical Research: Atmospheres, 121, 12,820–12,838, https://doi.org/10.1002/2016JD024892, 2016.

Cook, E. R., Briffa, K. R., and Jones, P. D.: Spatial regression methods in dendroclimatology: A review and comparison of two techniques, International Journal of Climatology, 14, 379–402, https://doi.org/10.1002/joc.3370140404, 1994.

Cook, E. R., Meko, D. M., Stahle, D. W., and Cleaveland, M. K.: Drought Reconstructions for the Continental United States, Journal of Climate, 12, 1145–1162, https://doi.org/10.1175/1520-0442(1999)012⟨1145:DRFTCU⟩2.0.CO;2, 1999.

Cook, E. R., Woodhouse, C. A., Eakin, C. M., Meko, D. M., and Stahle, D. W.: Long-Term Aridity Changes in the Western United States, Science, 306, 1015–1018, https://doi.org/10.1126/science.1102586, 2004.

Cook, E. R., Anchukaitis, K. J., Buckley, B. M., D'Arrigo, R. D., Jacoby, G. C., and Wright, W. E.: Asian Monsoon Failure and Megadrought During the Last Millennium, Science, 328, 486–489, https://doi.org/10.1126/science.1185188, 2010a.

Cook, E. R., Seager, R., Heim Jr, R. R., Vose, R. S., Herweijer, C., and Woodhouse, C.: Megadroughts in North America: placing IPCC projections of hydroclimatic change in a long-term palaeoclimate context, Journal of Quaternary Science, 25, 48–61, https://doi.org/10.1002/jqs.1303, 2010b.

Cook, E. R., Krusic, P. J., Anchukaitis, K. J., Buckley, B. M., Nakatsuka, T., Sano, M., and Members, P. A.: Tree-ring reconstructed summer temperature anomalies for temperate East Asia since 800 C.E., Climate Dynamics, 41, 2957–2972, https://doi.org/10.1007/s00382-012-1611-x, 2013.

Cook, E. R., Seager, R., Kushnir, Y., Briffa, K. R., Büntgen, U., Frank, D., Krusic, P. J., Tegel, W., van der Schrier, G., Andreu-Hayles, L., Baillie, M., Baittinger, C., Bleicher, N., Bonde, N., Brown, D., Carrer, M., Cooper, R., Čufar, K., Dittmar, C., Esper, J., Griggs, C., Gunnarson, B., Günther, B., Gutierrez, E., Haneca, K., Helama, S., Herzig, F., Heussner, K.-U., Hofmann, J., Janda, P., Kontic, R., Köse, N., Kyncl, T., Levanič, T., Linderholm, H., Manning, S., Melvin, T. M., Miles, D., Neuwirth, B., Nicolussi, K., Nola, P., Panayotov, M., Popa, I., Rothe, A., Seftigen, K., Seim, A., Svarva, H., Svoboda, M., Thun, T., Timonen, M., Touchan, R., Trotsiuk, V., Trouet, V., Walder, F., Ważny, T., Wilson, R., and Zang, C.: Old World megadroughts and pluvials during the Common Era, Science Advances, 1, e1500 561, https://doi.org/10.1126/sciadv.1500561, 2015.

Cox, D. and Isham, V.: Point Processes, Routledge, New York, https://doi.org/10.1201/9780203743034, 1980.

CRED - Centre for Research on the Epidemiology of Disasters: 2021 - Disasters in numbers, URL `https://cred.be/sites/default/files/2021_EMDAT_report.pdf`, 2022.

Dai, A., Trenberth, K. E., and Qian, T.: A Global Dataset of Palmer Drought Severity Index for 1870–2002: Relationship with Soil Moisture and Effects of Surface Warming, Journal of Hydrometeorology, 5, 1117–1130, https://doi.org/10.1175/JHM-386.1, 2004.

Dallan, E., Borga, M., Zaramella, M., and Marra, F.: Enhanced Summer Convection Explains Observed Trends in Extreme Subdaily Precipitation in the Eastern Italian

Alps, Geophysical Research Letters, 49, e2021GL096 727, https://doi.org/10.1029/2021GL096727, 2022a.

Dallan, E., Marra, F., Fosser, G., Marani, M., Formetta, G., Schär, C., and Borga, M.: How well does a convection-permitting climate model represent the reverse orographic effect of extreme hourly precipitation? (published online as a preprint), EGUsphere, 2022, 1–27, https://doi.org/10.5194/egusphere-2022-1037, 2022b.

Dalrymple, T.: Flood frequency Analysis, Manual of hydrology: Part 3. Flood-flow techniques 1543, Geological Survey Water Supply Paper 1543-A, https://doi.org/10.3133/wsp1543A, 1960.

Davison, A. C. and Smith, R. L.: Models for Exceedances over High Thresholds, J. Roy. Stat. Soc. B-Met., 52, 393–442, URL `http://www.jstor.org/stable/2345667`, 1990.

de Haan, L.: On regular variation and its application to the weak convergence of sample extremes, Math. Centre Tract, Math. Centrum, Amsterdam, 1970.

De Zea Bermudez, P. and Mendes, Z.: Extreme value theory in medical sciences: modeling total high cholesterol levels, Journal of Statistical Theory and Practice volume, 6, 468–491, https://doi.org/10.1080/15598608.2012.695673, 2012.

Dixon, M. J. and Tawn, J. A.: The effect of non-stationarity on extreme sea-level estimation, J. Roy. Stat. Soc. C-App., 48, 135–151, 1999.

Dobbertin, M.: Tree growth as indicator of tree vitality and of tree reaction to environmental stress: a review, European Journal of Forest Research, 124, 319–333, https://doi.org/10.1007/s10342-005-0085-3, 2005.

Douville, H., Raghavan, K., Renwick, J., Allan, R., Arias, P., Barlow, M., Cerezo-Mota, R., Cherchi, A., Gan, T., Gergis, J., Jiang, D., Khan, A., Pokam Mba, W., Rosenfeld, D., Tierney, J., and Zolina, O.: Water Cycle Changes, in: Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change, edited by Masson-Delmotte, V., Zhai, P., Pirani, A., Connors, S., Péan, C., Berger, S., Caud, N., Chen, Y., Goldfarb, L., Gomis, M., Huang, M., Leitzell, K., Lonnoy, E., Matthews, J., Maycock, T., Waterfield, T., Yelekçi, O., Yu, R., and Zhou, B., chap. 8, pp. 1055–1210, Cambridge University Press, URL `https://www.ipcc.ch/report/sixth-assessment-report-working-group-i/`, 2022.

Dracup, J. A., Lee, K. S., and Paulson Jr., E. G.: On the definition of droughts, Water Resources Research, 16, 297–302, https://doi.org/10.1029/WR016i002p00297, 1980.

Dubey, S. D.: A compound weibull distribution, Naval Research Logistics Quarterly, 15, 179–188, https://doi.org/10.1002/nav.3800150205, 1968.

Dubey, S. D.: Compound Gamma, Beta and F Distribution, Metrika, 16, 27–31, https://doi.org/10.1007/BF02613934, 1970.

Eliot, M.: Influence of interannual tidal modulation on coastal flooding along the Western Australian coast, J. Geophys. Res.-Oceans, 115, https://doi.org/10.1029/2010JC006306, 2010.

Elvidge, S. and Angling, M. J.: Using Extreme Value Theory for Determining the Probability of Carrington-Like Solar Flares, Space Weather, 16, 417–421, https://doi.org/10.1002/2017SW001727, 2018.

Emanuel, K. and Jagger, T.: On Estimating Hurricane Return Periods, Journal of Applied Meteorology and Climatology, 49, 837–844, 2010.

Embrechts, P., Klüppelberg, C., and Mikosch, T.: Modelling Extremal Events for insurance and finance, Springer, New York, 1997.

Ferrarin, C., Bajo, M., Benetazzo, A., Cavaleri, L., Chiggiato, J., Davison, S., Davolio, S., Lionello, P., Orlić, M., and Umgiesser, G.: Local and large-scale controls of the exceptional Venice floods of November 2019, Progress in Oceanography, 197, 102 628, https://doi.org/10.1016/j.pocean.2021.102628, 2021.

Ferro, C. A. T. and Segers, J.: Inference for clusters of extreme values, J. Roy Stat. Soc. B-Met., 65, 545–556, https://doi.org/10.1111/1467-9868.00401, 2003.

Finkenstädt, B. and Rootzén, H.: Extreme Values in Finance, Telecommunications and the Environment, Taylor & Francis Inc, Chapman & Hall/CRC Monographs on Statistics & Applied Probability, Boca Raton, Florida, 2004.

Fisher, R. A. and Tippett, L. H. C.: Limiting forms of the frequency distribution of the largest or smallest member of a sample, Math. Proc. Cambridge, 24, 180–190, https://doi.org/10.1017/S0305004100015681, 1928.

Fortunato, A. B., Li, K., Bertin, X., Rodrigues, M., and Miguez, B. M.: Determination of extreme sea levels along the Iberian Atlantic coast, Ocean Eng., 111, 471–482, https://doi.org/https://doi.org/10.1016/j.oceaneng.2015.11.031, 2016.

Forzieri, G., Feyen, L., Rojas, R., Flörke, M., Wimmer, F., and Bianchi, A.: Ensemble projections of future streamflow droughts in Europe, Hydrology and Earth System Sciences, 18, 85–108, https://doi.org/10.5194/hess-18-85-2014, 2014.

Fritts, H. C.: Tree Rings and Climate, Academic Press, New York, 1976.

Fritts, H. C., Blasing, T. J., Hayden, B. P., and Kutzbach, J. E.: Multivariate Techniques for Specifying Tree-Growth and Climate Relationships and for Reconstructing Anomalies in Paleoclimate, Journal of Applied Meteorology (1962-1982), 10, 845–864, 1971.

Fréchet, M. R.: Sur la loi de probabilité de l'écart maximum, Ann. Soc. Polon. Math., 6, 93–116, 1927.

Gibbs, W. J. and Maher, J. V.: Rainfall Deciles as Drought Indicators. Bureau of Meteorology Bull. 48. Commonwealth of Australia, Melbourne, Australia, 1967.

Gillette, H. P.: A creeping drought under way, Water and sewage works, 104, 1950.

Gnedenko, B. V.: Sur la distribution limite du terme maximum d'une serie aleatoire, Ann. Math., 44, 423–453, 1943.

Goring, D. G., Stephens, S. A., Bell, R. G., and P., P. C.: Estimation of Extreme Sea Levels in a Tide-Dominated Environment Using Short Data Records, Journal of Waterway, Port, Coastal, and Ocean Engineering, 137, 150–159, https://doi.org/10.1061/(ASCE)WW.1943-5460.0000071, 2011.

Greenwood, J. A., Landwehr, J. M., Matalas, N. C., and Wallis, J. R.: Probability weighted moments: definition and relation to parameters of several distributions expressable in inverse form, Water Resour. Res., 15, 1049–1054, https://doi.org/10.1029/WR015i005p01049, 1979.

Gründemann, G. J., Zorzetto, E., Beck, H. E., Schleiss, M., Van de Giesen, N., Marani, M., and van der Ent, R. J.: Extreme Precipitation Return Levels for Multiple Durations on a Global Scale, Earth and Space Science Open Archive, p. 51, https://doi.org/10.1002/essoar.10503814.3, 2020.

Gumbel, E. J.: Statistics of Extremes, Columbia University Press, New York, 1958.

Guttman, N. B.: Comparing the Palmer drought severity index and the standardized precipitation index, Journal of the American Water Resources Association, 34, 113–121, https://doi.org/10.1111/j.1752-1688.1998.tb05964.x, 1998.

Haigh, I. D., Nicholls, R., and Wells, N.: Assessing changes in extreme sea levels: Application to the English Channel, 1900–2006, Cont. Shelf. Res., 30, 1042–1055, https://doi.org/10.1016/j.csr.2010.02.002, 2010.

Haigh, I. D., Eliot, M., and Pattiaratchi, C.: Global influences of the 18.61-year nodal cycle and 8.85-year cycle of lunar perigee on high tidal levels, J. Geophys. Res.-Oceans, 116, 25 249, https://doi.org/10.1029/2010JC006645, 2011.

Haigh, I. D., MacPherson, L. R., Mason, M. S., Wijeratne, E. M. S., Pattiaratchi, C. B., Crompton, R. P., and George, S.: Estimating present day extreme water level exceedance probabilities around the coastline of Australia: tropical cyclone-induced storm surges, Clim. Dynam., 42, 139–157, https://doi.org/10.1007/s00382-012-1653-0, 2014a.

Haigh, I. D., Wahl, T., Rohling, E. J., Price, R. M., Pattiaratchi, C., Calafat, F. M., and Dangendorf, S.: Timescales for detecting a significant acceleration in sea level rise, Nat. Commun., 5, https://doi.org/10.1038/ncomms4635, 2014b.

Hall, T. M. and Sobel, A. H.: On the impact angle of Hurricane Sandy's New Jersey landfall, Geophys. Res. Lett., 40, 2312–2315, https://doi.org/10.1002/grl.50395, 2013.

Hamdi, Y., Bardet, L., Duluc, C.-M., and Rebour, V.: Extreme storm surges: a comparative study of frequency analysis approaches, Nat. Hazard Earth Sys., 14, 2053–2067, https://doi.org/10.5194/nhess-14-2053-2014, 2014.

Hamdi, Y., Bardet, L., Duluc, C.-M., and Rebour, V.: Use of historical information in extreme-surge frequency estimation: the case of marine flooding on the La Rochelle site in France, Nat. Hazard Earth Sys., 15, 1515–1531, https://doi.org/10.5194/nhess-15-1515-2015, 2015.

Hamdi, Y., Garnier, E., Giloy, N., Duluc, C.-M., and Rebour, V.: Analysis of the risk associated with coastal flooding hazards: a new historical extreme storm surges dataset for Dunkirk, France, Nat. Hazard Earth Sys., 18, 3383–3402, https://doi.org/10.5194/nhess-18-3383-2018, 2018.

Hamon, B. V. and Middleton, J. F.: Return periods of extreme sea levels: the exceedance probability method, International Hydrographic Review, 66, 165–177, 1989.

Harris, I., Jones, P., Osborn, T., and Lister, D.: Updated high-resolution grids of monthly climatic observations – the CRU TS3.10 Dataset, International Journal of Climatology, 34, 623–642, https://doi.org/h10.1002/joc.3711, 2014.

Hay, C., Morrow, E., Kopp, R., and Mitrovica, J. X.: Probabilistic reanalysis of twentieth-century sea-level rise, Nature, 517, 481–484, https://doi.org/10.1038/nature14093, 2015.

Heudorfer, B. and Stahl, K.: Comparison of different threshold level methods for drought propagation analysis in Germany, Hydrology Research, 48, 1311–1326, 2017.

Hosking, J. R. M.: L-moments: Analysis and estimation of distributions using linear combinations of order statistics, J. Roy. Stat. Soc. B-Met., 52, 105–124, 1990.

Hosking, J. R. M. and Wallis, J. R.: Parameter and Quantile Estimation for the Generalized Pareto Distribution, Technometrics, 29, 339–349, URL http://www.jstor.org/stable/1269343, 1987.

Hosking, J. R. M., Wallis, J. R., and Wood, E. F.: Estimation of the generalized extreme-value distribution by the method of probability-weighted moments, Technometrics, 27, 251–261, https://doi.org/10.1080/00401706.1985.10488049, 1985.

Hosseini, S. R., Scaioni, M., and Marani, M.: Extreme Atlantic hurricane probability of occurrence through the Metastatistical Extreme Value Distribution, Geophys. Res. Lett., 47, 2019GL086 138, https://doi.org/10.1029/2019GL086138, 2020.

Hu, L., Nikolopoulos, E. I., Marra, F., Morin, E., Marani, M., and Anagnostou, E. N.: Evaluation of MEVD-based precipitation frequency analyses from quasi-global precipitation datasets against dense rain gauge networks, Journal of Hydrology, 590, 125 564, https://doi.org/10.1016/j.jhydrol.2020.125564, 2020.

Huntington, T.: Climate warming-induced intensification of the hydrologic cycle: A review of the published record and assessment of the potential impacts on agriculture, Advances in Agronomy, 109, 1–53, https://doi.org/10.1016/B978-0-12-385040-9.00001-3, 2010.

Ionita, M., Tallaksen, L. M., Kingston, D. G., Stagge, J. H., Laaha, G., Van Lanen, H. A., Scholz, P., Chelcea, S. M., and Haslinger, K.: The European 2015 drought from a climatological perspective, Hydrology and Earth System Sciences, 21, 1397–1419, 2017.

Jenkinson, A.: The frequency distribution of the annual maximum (or minimum) values of meteorological elements, Q. J. Roy. Meteor. Soc., 81, 158–171, 1955.

Jevrejeva, S., Moore, J. C., Grinsted, A., and Woodworth, P. L.: Recent global sea level acceleration started over 200 years ago?, Geophys. Res. Lett., 35, https://doi.org/10.1029/2008GL033611, 2008.

Ji, L. and Peters, A. J.: Assessing vegetation response to drought in the northern Great Plains using vegetation and drought indices, Remote Sensing of Environment, 87, 85–98, https://doi.org/https://doi.org/10.1016/S0034-4257(03)00174-3, 2003.

Johns, B. and Ali, M. A.: The numerical modeling of storm surges in the Bay of Bengal, Q. J. Roy. Meteor. Soc., 106, 1–18, https://doi.org/10.1002/qj.49710644702, 1980.

Karl, T. R.: The Sensitivity of the Palmer Drought Severity Index and Palmer's Z-Index to their Calibration Coefficients Including Potential Evapotranspiration, Journal of Climate and Applied Meteorology, 25, 77–86, 1986.

Katz, R. W., Parlange, M. B., and Naveau, P.: Statistics of extremes in hydrology, Adv. Water Resour., 25, 1287–1304, https://doi.org/10.1016/S0309-1708(02)00056-8, 2002.

Katz, R. W., Brush, G. S., and Parlange, M. B.: Statistics of extremes: modeling ecological disturbances, Ecology, Ecology, 86, 1124–1134, https://doi.org/10.1890/04-0606, 2005.

Laaha, G., Gauster, T., Tallaksen, L. M., Vidal, J.-P., Stahl, K., Prudhomme, C., Heudorfer, B., Vlnas, R., Ionita, M., Van Lanen, H. A. J., Adler, M.-J., Caillouet, L., Delus, C., Fendekova, M., Gailliez, S., Hannaford, J., Kingston, D., Van Loon, A. F., Mediero, L., Osuch, M., Romanowicz, R., Sauquet, E., Stagge, J. H., and Wong, W. K.: The European 2015 drought from a hydrological perspective, Hydrology and Earth System Sciences, 21, 3001–3024, https://doi.org/10.5194/hess-21-3001-2017, 2017.

Li, G., Zhang, X., Zwiers, F., and Wen, Q. H.: Quantification of uncertainty in high-resolution temperature scenarios for North America, J. Climate, 25, 3373–3389, https://doi.org/10.1175/JCLI-D-11-00217.1, 2012.

Lloyd-Hughes, B.: The impracticality of a universal drought definition, Theoretical and Applied Climatology, 117, 607–611, https://doi.org/10.1007/s00704-013-1025-7, 2014.

Loaiciga, H. A., Valdes, J. B., Vogel, R., Garvey, J., and Schwarz, H.: Global warming and the hydrologic cycle, Journal of Hydrology, 174, 83–127, https://doi.org/10.1016/0022-1694(95)02753-X, 1996.

Lowe, J. A., Woodworth, P. L., Knutson, T., McDonald, R. E., McInnes, K., Woth, K., Von Storch, H., Wolf, J., Swail, V., Bernier, N., Gulev, S., Horsburgh, K., Unnikrishnan, A. S., Hunter, J., and Weisse, R.: Past and future changes in extreme sea levels and waves, in: Understanding Sea-Level Rise and Variability, edited by Church, J. A., Woodworth, P. L., Aarup, T., and Wilson, W. S., chap. 11, pp. 326–375, Wiley-Blackwell, https://doi.org/10.1002/9781444323276, 2010.

Mann, H. B.: Nonparametric Tests Against Trend, Econometrica, 13, 245–259, 1945.

Marani, M. and Ignaccolo, M.: A metastatistical approach to rainfall extremes, Adv. Water Resour., 79, 121–126, https://doi.org/10.1016/j.advwatres.2015.03.001, 2015.

Marani, M. and Zanetti, S.: Long-term oscillations in rainfall extremes in a 268 year daily time series, Water Resources Research, 51, 639–647, https://doi.org/10.1002/2014WR015885, 2015.

Marani, M. and Zorzetto, E.: Doubly stochastic distributions of extreme events, arXiv, https://doi.org/10.48550/ARXIV.1902.09862, 2019.

Marani, M., Katul, G. G., Pan, W. K., and Parolari, A. J.: Intensity and frequency of extreme novel epidemics, Proceedings of the National Academy of Sciences, 118, e2105482 118, https://doi.org/10.1073/pnas.2105482118, 2021.

Marra, F., Nikolopoulos, E. I., Anagnostou, E. N., and Morin, E.: Metastatistical extreme value analysis of hourly rainfall from short records: Estimation of high quantiles and impact of measurement errors, Adv. Water Resour., 117, 27–39, https://doi.org/10.1016/j.advwatres.2018.05.001, 2018.

Marra, F., Zoccatelli, D., Armon, M., and Morin, E.: A simplified MEV formulation to model extremes emerging from multiple nonstationary underlying processes, Advances in Water Resources, 127, 280–290, https://doi.org/10.1016/j.advwatres.2019.04.002, 2019.

Marra, F., Borga, M., and Morin, E.: A Unified Framework for Extreme Subdaily Precipitation Frequency Analyses Based on Ordinary Events, Geophysical Research Letters, 47, e2020GL090 209, https://doi.org/10.1029/2020GL090209, 2020.

Marra, F., Armon, M., and Morin, E.: Coastal and orographic effects on extreme precipitation revealed by weather radar observations, Hydrology and Earth System Sciences, 26, 1439–1458, https://doi.org/10.5194/hess-26-1439-2022, 2022.

Martins, E. S. and Stedinger, J. R.: Generalized Maximum Likelihood Pareto-Poisson estimators for partial duration series, Water Resources Research, 37, 2551–2557, https://doi.org/10.1029/2001WR000367, 2001.

McInnes, K. L., Hubbert, G., Macadam, I., and O'Grady, J.: An assessment of current and future vulnerability to coastal inundation due to sea-level extremes in Victoria, southeast Australia, Int. J. Climatol., 33, 33–47, https://doi.org/10.1002/JOC.3405, 2013.

McKee, T. B., Doesken, N. J., and Kleist, J.: The relationship of drought frequency and duration to time scales, in: Proceedings of the 8th Conference on Applied Climatology, vol. 17, pp. 179–183, Boston, MA, USA, 1993.

McPhillips, L. E., Chang, H., Chester, M. V., Depietri, Y., Friedman, E., Grimm, N. B., Kominoski, J. S., McPhearson, T., Méndez-Lázaro, P., Rosi, E. J., and Shafiei Shiva, J.: Defining Extreme Events: A Cross-Disciplinary Review, Earth's Future, 6, 441–455, https://doi.org/10.1002/2017EF000686, 2018.

Meehl, G., Stocker, T., Collins, W. D., Friedlingstein, P., Gaye, A., Gregory, J. M., Kitoh, A., Knutti, R., Murphy, J. M., Noda, A., Raper, S. C. B., Watterson, I. G.,

Weaver, A. J., and Zhao, Z. C.: Global Climate Projections, in: Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change, edited by Solomon, S., Qin, D., Manning, M., Chen, Z., Marquis, M., Averyt, K., Tignor, M., and Miller, H., chap. 10, pp. 747–846, Cambridge University Press, URL `http://www.ipcc.ch/publications_and_data/publications_ipcc_fourth_assessment_report_wg1_report_the_physical_science_basis.htm`, 2007.

Mekonnen, K., Melesse, A. M., and Woldesenbet, T. A.: Effect of temporal sampling mismatches between satellite rainfall estimates and rain gauge observations on modelling extreme rainfall in the Upper Awash Basin, Ethiopia, Journal of Hydrology, 598, 126 467, https://doi.org/10.1016/j.jhydrol.2021.126467, 2021.

Menéndez, M. and Woodworth, P. L.: Changes in extreme high water levels based on a quasi-global tide-gauge data set, J. Geophys. Res., 115, 1124–1134, https://doi.org/10.1029/2009JC005997, 2010.

Middleton, J. F. and Thompson, K. R.: Return periods of extreme sea levels from short records, Journal of Geophysical Research, 91, 11 707–11 716, https://doi.org/10.1029/JC091iC10p11707, 1986.

Miniussi, A. and Marani, M.: Estimation of Daily Rainfall Extremes Through the Metastatistical Extreme Value Distribution: Uncertainty Minimization and Implications for Trend Detection, Water Resour. Res., 56, e2019WR026 535, https://doi.org/0.1029/2019WR026535, 2020.

Miniussi, A. and Marra, F.: Estimation of extreme daily precipitation return levels at-site and in ungauged locations using the simplified MEV approach, Journal of Hydrology, 603, 126 946, https://doi.org/10.1016/j.jhydrol.2021.126946, 2021.

Miniussi, A., Marani, M., and Villarini, G.: Metastatistical Extreme Value Distribution applied to floods across the continental United States, Adv. Water Resour., 136, 103 498, https://doi.org/10.1016/j.advwatres.2019.103498, 2020a.

Miniussi, A., Villarini, G., and Marani, M.: Analyses through the metastatistical extreme value distribution identify contributions of tropical cyclones to rainfall extremes in the eastern United States, Geophys. Res. Lett., 47, e2020GL087 238, https://doi.org/10.1029/2020GL087238, 2020b.

Mishra, A. K. and Singh, V. P.: A review of drought concepts, Journal of Hydrology, 391, 202–216, https://doi.org/10.1016/j.jhydrol.2010.07.012, 2010.

Monteleone, B., Bonaccorso, B., and Martina, M.: A joint probabilistic index for objective drought identification: the case study of Haiti, Natural Hazards and Earth System Sciences, 20, 471–487, https://doi.org/10.5194/nhess-20-471-2020, 2020.

Nerantzaki, S. D. and Papalexiou, S. M.: Assessing extremes in hydroclimatology: A review on probabilistic methods, Journal of Hydrology, 605, 127 302, https://doi.org/h10.1016/j.jhydrol.2021.127302, 2022.

Oppenheimer, M., Glavovic, B., Hinkel, J., van de Wal, R., Magnan, A., Abd-Elgawad, A., Cai, R., Cifuentes-Jara, M., DeConto, R., Ghosh, T., Hay, J., Isla, F., Marzeion, B., Meyssignac, B., and Sebesvari, Z.: Sea Level Rise and Implications for Low-Lying Islands, Coasts and Communities, in: IPCC Special Report on the Ocean and Cryosphere in a Changing Climate, edited by Pörtner, H.-O., Roberts, D., Masson-Delmotte, V., Zhai, P., Tignor, M., Poloczanska, E., Mintenbeck, K., Alegría, A., Nicolai, M., Okem, A., Petzold, J., Rama, B., and Weyer, N. M., chap. 4, pp. 321–445, Cambridge University Press, URL `https://www.ipcc.ch/srocc/download-report/`, 2019.

O'Donnell, A. J., Cook, E. R., Palmer, J. G., Turney, C. S., and Grierson, P. F.: Potential for tree rings to reveal spatial patterns of past drought variability across western Australia, Environmental Research Letters, 13, 024 020, https://doi.org/10.1088/1748-9326/aaa204, 2018.

Palazzolo, N., Peres, D. J., Creaco, E., and Cancelliere, A.: Potential improvements of landslide prediction by hydro-meteorological thresholds: an investigation based on reanalysis soil moisture data and principal component analysis, Natural Hazards and Earth System Sciences Discussions, 2022, 1–22, https://doi.org/10.5194/nhess-2022-175, 2022.

Palmer, W. C.: Meteorological Drought, Office of Climatology. US Weather Bureau, Research Paper No. 45, 58 p., 1965.

Palmer, W. C.: Keeping Track of Crop Moisture Conditions, Nationwide: The New Crop Moisture Index, Weatherwise, 21, 156–161, https://doi.org/10.1080/00431672.1968.9932814, 1968.

Papalexiou, S. M. and Koutsoyiannis, D.: Battle of extreme value distributions: A global survey on extreme daily rainfall, Water Resources Research, 49, 187–201, https://doi.org/10.1029/2012WR012557, 2013.

Peng, D., Hill, E. M., and Meltzner, A. J.and Switzer, A. D.: Tide gauge records show that the 18.61-year nodal tidal cycle can change high water levels by up to 30 cm, J. Geophys. Res.-Oceans, 124, 736–749, https://doi.org/10.1029/2018JC014695, 2019.

Peña-Gallardo, M., Vicente-Serrano, S. M., Hannaford, J., Lorenzo-Lacruz, J., Svoboda, M., Domínguez-Castro, F., Maneta, M., Tomas-Burguera, M., and El Kenawy, A.: Complex influences of meteorological drought time-scales on hydrological

droughts in natural basins of the contiguous Unites States, Journal of Hydrology, 568, 611–625, https://doi.org/https://doi.org/10.1016/j.jhydrol.2018.11.026, 2019.

Pickands, III, J.: Statistical inference using extreme order statistics, Ann. Stat., 3, 119–131, https://doi.org/10.1214/aos/1176343003, 1975.

Pisarenko, V. F., Rodkin, M. V., and Rukavishnikova, T. A.: Estimation of the probability of strongest seismic disasters based on the extreme value theory, Izvestiya, Physics of the Solid Earth, 50, 311–324, https://doi.org/10.1134/S1069351314030070, 2014a.

Pisarenko, V. F., Sornette, A., Sornette, D., and Rodkin, M. V.: 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, 1599–1624, https://doi.org/10.1007/s00024-014-0882-z, 2014b.

Pugh, D. T. and Vassie, J. M.: Extreme sea levels from tide and surge probability, in: Proc. 16th International Conference on Coastal Engineering 1978, Hamburg, Germany, chap. 52, pp. 911–930, American Society of Civil Engineers, 1979.

Pugh, D. T. and Vassie, J. M.: Applications of the joint probability method for extreme sea level computations, in: Proc. Inst. Civ. Eng., Part 2, pp. 959–975, 1980.

Pugh, D. T. and Woodworth, P. L.: Sea-Level Science: Understanding Tides, Surges, Tsunamis and Mean Sea-Level Changes, Cambridge University Press, 2014.

Rashid, M. M. and Wahl, T.: How climate affects extreme events and hence ecological population models, Environ. Res. Commun., 4, 071 001, https://doi.org/0.1088/2515-7620/ac77de, 2022.

Rossi, G., Benedini, M., Tsakiris, G., and Giakoumakis, S.: On regional drought estimation and analysis, Water Resources Management, 6, 249–277, https://doi.org/10.1007/BF00872280, 1992.

Rouault, M. and Richard, Y.: Intensity and spatial extent of droughts in southern Africa, Geophysical Research Letters, 32, L15 702, https://doi.org/https://doi.org/10.1029/2005GL022436, 2005.

Rueda, A., Camus, P., Méndez, F. J., Tomás, A., and Luceño, A.: An extreme value model for maximum wave heights based on weather types, J. Geophys. Res.-Oceans, 121, 1262–1273, https://doi.org/10.1002/2015JC010952, 2016.

Rypkema, D. C., Horvitz, C. C., and Tuljapurkar, S.: How climate affects extreme events and hence ecological population models, Ecology, 100, e02 684, https://doi.org/10.1002/ecy.2684, 2019.

Scheffner, N. W., Borgman, L. E., and Mark, D. J.: Empirical simulation technique based storm surge frequency analyses, J. Waterway, Port, Coastal, Ocean Eng., 122, 93–101, https://doi.org/10.1061/(ASCE)0733-950X(1996)122:2(93), 1996.

Schellander, H., Lieb, A., and Hell, T.: Error structure of metastatistical and generalized extreme value distributions for modeling extreme rainfall in Austria, Earth and Space Science, 6, 1616–1632, https://doi.org/10.1029/2019EA000557, 2019.

Serinaldi, F., Lombardo, F., and Kilsby, C. G.: All in order: Distribution of serially correlated order statistics with applications to hydrological extremes, Advances in Water Resources, 144, 103 686, https://doi.org/10.1016/j.advwatres.2020.103686, 2020.

Smerdon, J. E., Cook, B. I., Cook, E. R., and Seager, R.: Bridging Past and Future Climate across Paleoclimatic Reconstructions, Observations, and Models: A Hydroclimate Case Study, Journal of Climate, 28, 3212–3231, https://doi.org/10.1175/JCLI-D-14-00417.1, 2015.

Smith, R. L.: Extreme value theory based on the r largest annual events, J. Hydrol., 86, 27–43, https://doi.org/10.1016/0022-1694(86)90004-1, 1986.

Solari, S., Egüen, M., Polo, M. J., and Losada, M. A.: Peaks Over Threshold (POT): A methodology for automatic threshold estimation using goodness of fit p-value, Water Resources Research, 53, 2833–2849, https://doi.org/10.1002/2016WR019426, 2017.

Songchitruksa, P. and Tarko, A. P.: The extreme value theory approach to safety estimation, Accident Anal. Prev., 38, 811–822, https://doi.org/10.1016/j.aap.2006.02.003, 2006.

Stahl, K., Vidal, J.-P., Hannaford, J., Tijdeman, E., Laaha, G., Gauster, T., and Tallaksen, L. M.: The challenges of hydrological drought definition, quantification and communication: an interdisciplinary perspective, Proceedings of the International Association of Hydrological Sciences, 383, 291–295, https://doi.org/10.5194/piahs-383-291-2020, 2020.

Staudinger, M., Stahl, K., and Seibert, J.: A drought index accounting for snow, Water Resources Research, 50, 7861–7872, https://doi.org/10.1002/2013WR015143, 2014.

Stedinger, J. R., Vogel, R. M., and Georgiou, E. F.: Frequency analysis of extreme events, In: Handbook of Hydrology, Chapt. 18, editor in chief D. J. Maidment, McGraw Hill, New York, 1993.

Tallaksen, L. M. and van Lanen, H. A. J.: Hydrological drought: processes and estimation methods for streamflow and groundwater, In: Developments in water science, vol. 48, Amsterdam, theNetherlands: Elsevier Science B.V., 2004.

Tallaksen, L. M., Madsen, H., and Clausen, B.: On the definition and modelling of streamflow drought duration and deficit volume, Hydrological Sciences Journal, 42, 15–33, https://doi.org/10.1080/02626669709492003, 1997.

Tawn, J. A.: An extreme-value theory model for dependent observations, J. Hydrol., 101, 227–250, https://doi.org/10.1016/0022-1694(88)90037-6, 1988.

Tawn, J. A. and Vassie, J. M.: Extreme sea levels: The joint probability method revisited and revised, in: Proc. Inst. Civ. Eng., Part 2, pp. 429–442, 1989.

Tebaldi, C., Strauss, B. H., and Zervas, C. E.: Modelling sea level rise impacts on storm surges along US coasts, Environ. Res. Lett., 7, 014–032, https://doi.org/10.1088/1748-9326/7/1/014032, 2012.

Thornthwaite, C. W.: An Approach toward a Rational Classification of Climate, Geographical Review, 38, 55–94, 1948.

Touchan, R., Anchukaitis, K. J., Meko, D. M., Sabir, M., Attalah, S., and Aloui, A.: Spatiotemporal drought variability in northwestern Africa over the last nine centuries, Climate Dynamics, 37, 237—-252, https://doi.org/10.1007/s00382-010-0804-4, 2011.

Tramblay, Y., Koutroulis, A., Samaniego, L., Vicente-Serrano, S. M., Volaire, F., Boone, A., Le Page, M., Llasat, M. C., Albergel, C., Burak, S., Cailleret, M., Kalin, K. C., Davi, H., Dupuy, J.-L., Greve, P., Grillakis, M., Hanich, L., Jarlan, L., Martin-StPaul, N., Martínez-Vilalta, J., Mouillot, F., Pulido-Velazquez, D., Quintana-Seguí, P., Renard, D., Turco, M., Türkeş, M., Trigo, R., Vidal, J.-P., Vilagrosa, A., Zribi, M., and Polcher, J.: Challenges for drought assessment in the Mediterranean region under future climate scenarios, Earth-Science Reviews, 210, 103 348, https://doi.org/10.1016/j.earscirev.2020.103348, 2020.

Trenberth, K. E.: Atmospheric moisture residence times and cycling: Implications for rainfall rates and climate change, Climatic change, 39, 667–694, 1998.

Ullrich, P. A., Xu, Z., Rhoades, A., Dettinger, M., Mount, J., Jones, A., and Vahmani, P.: California's Drought of the Future: A Midcentury Recreation of the Exceptional Conditions of 2012–2017, Earth's Future, 6, 1568–1587, https://doi.org/https://doi.org/10.1029/2018EF001007, 2018.

Ummenhofer, C. C., England, M. H., McIntosh, P. C., Meyers, G. A., Pook, M. J., Risbey, J. S., Gupta, A. S., and Taschetto, A. S.: What causes southeast Australia's worst droughts?, Geophysical Research Letters, 36, https://doi.org/https://doi.org/10.1029/2008GL036801, 2009.

USGSC - United States Geological Survey: The water cycle, URL https://www.usgs.gov/special-topics/water-science-school/science/water-cycle, 2022.

Valle-Levinson, A., Marani, M., Carniello, L., D'Alpaos, A., and Lanzoni, S.: Astronomic link to anomalously high mean sea level in the northern Adriatic Sea, Estuar. Coast. Shelf. S., 257, 107 418, https://doi.org/10.1016/j.ecss.2021.107418, 2021.

van der Schrier, G., Barichivich, J., Briffa, K. R., and Jones, P. D.: A scPDSI-based global data set of dry and wet spells for 1901–2009, Journal of Geophysical Research: Atmospheres, 118, 4025–4048, https://doi.org/10.1002/jgrd.50355, 2013.

Van Loon, A. F.: Hydrological drought explained, WIREs Water, 2, 359–392, https://doi.org/10.1002/wat2.1085, 2015.

van Rooy, M. P.: A rainfall anomaly index independent of time and space. Notos, 1965.

Vicente-Serrano, S. M., Beguería, S., and López-Moreno, J. I.: A multiscalar drought index sensitive to global warming: the standardized precipitation evapotranspiration index, Journal of climate, 23, 1696–1718, 2010.

Villarini, G., Smith, J. A., Ntelekos, A. A., and Schwarz, U.: Annual maximum and peaks-over-threshold analyses of daily rainfall accumulations for Austria, Journal of Geophysical Research: Atmospheres, 116, D05 103, https://doi.org/10.1029/2010JD015038, 2011.

Vogel, R. and Castellarin, A.: Risk, reliability, and return periods and hydrologic design. Handbook of Applied Hydrology, Singh, VP, Ed.; McGraw-Hill Book Company, New York, NY, USA, 2017.

Volpi, E., Fiori, A., Grimaldi, S., Lombardo, F., and Koutsoyiannis, D.: One hundred years of return period: Strengths and limitations, Water Resources Research, 51, 8570–8585, https://doi.org/10.1002/2015WR017820, 2015.

Volpi, E., Fiori, A., Grimaldi, S., Lombardo, F., and Koutsoyiannis, D.: Save hydrological observations! Return period estimation without data decimation, J. Hydrol., 571, 782–792, https://doi.org/10.1016/j.jhydrol.2019.02.017, 2019.

Von Mises, R.: La distribution de la plus grande de n valuers, Rev. math. Union interbalcanique, 1, 141–160, 1936.

Vousdoukas, M. I., Voukouvalas, E., Annunziato, A., Giardino, A., and Feyen, L.: Projections of extreme storm surge levels along Europe, Clim. Dynam., 47, 3171–3190, https://doi.org/10.1007/s00382-016-3019-5, 2016.

Wahl, T., Haigh, I. D., Nicholls, R. J., Arns, A., Dangendorf, S., Hinkel, J., and Slangen, A. B.: Understanding extreme sea levels for broad-scale coastal impact

and adaptation analysis, Nature communications, 8, 1–12, https://doi.org/10.1038/ncomms16075, 2017.

Wells, N., Goddard, S., and Hayes, M. J.: A Self-Calibrating Palmer Drought Severity Index, Journal of climate, 17, 2335–2351, https://doi.org/10.1175/1520-0442(2004)017⟨2335:ASPDSI⟩2.0.CO;2, 2004.

Wilhite, D. A.: Drought. A global assessment for drought, 2000.

Wilhite, D. A. and Glantz, M. H.: Understanding: the Drought Phenomenon: The Role of Definitions, Water International, 10, 111–120, https://doi.org/10.1080/02508068508686328, 1985.

Williams, P. D., Guilyardi, E., Sutton, R., Gregory, J., and Madec, G.: A new feedback on climate change from the hydrological cycle, Geophysical Research Letters, 34, L08 706, https://doi.org/10.1029/2007GL029275, 2007.

Woodworth, P. L. and Blackman, D. L.: Changes in high waters at Liverpool since 1768, Int. J. Climatol., 22, 697–7147, https://doi.org/https://doi.org/10.1002/joc.761, 2002.

Woodworth, P. L. and Blackman, D. L.: Evidence for systematic changes in extreme high waters since the mid-1970s, J. Climate, 17, 1190–1197, https://doi.org/10.1175/1520-0442(2004)017⟨1190:EFSCIE⟩2.0.CO;2, 2004.

Woodworth, P. L., Menéndez, M., and Roland Gehrels, W.: Evidence for Century-Timescale Acceleration in Mean Sea Levels and for Recent Changes in Extreme Sea Levels, Surv. Geophys., 32, 603–618, https://doi.org/10.1007/s10712-011-9112-8, 2011.

Yevjevich, V. M.: Objective approach to definitions and investigations of continental hydrologic droughts, An, Ph.D. thesis, Colorado State University. Libraries, 1967.

Zanchettin, D., Bruni, S., Raicich, F., Lionello, P., Adloff, F., Androsov, A., Antonioli, F., Artale, V., Carminati, E., Ferrarin, C., Fofonova, V., Nicholls, R. J., Rubinetti, S., Rubino, A., Sannino, G., Spada, G., Thiéblemont, R., Tsimplis, M., Umgiesser, G., Vignudelli, S., Wöppelmann, G., and Zerbini, S.: Sea-level rise in Venice: historic and future trends (review article), Natural Hazards and Earth System Sciences, 21, 2643–2678, https://doi.org/10.5194/nhess-21-2643-2021, 2021.

Zhang, K., Douglas, B. C., and Leatherman, S. P.: Twentieth-century storm activity along the U.S. east coast, J. Climate, 13, 1748–1761, https://doi.org/10.1175/1520-0442(2000)013⟨1748:TCSAAT⟩2.0.CO;2, 2000.

Zhang, W. and Gillies, R.: The Role of Anthropogenic Forcing in Western United States Hydroclimate Extremes, Geophysical Research Letters, 49, e2022GL100 659, https://doi.org/10.1029/2022GL100659, 2022.

Zhang, W.-Z., Shi, F., Hong, H.-S., Shang, S.-P., and Kirby, J. T.: Tide–surge interaction intensified by the Taiwan Strait, J. Geophys. Res.-Oceans, 115, https://doi.org/10.1029/2009JC005762, 2010.

Zorzetto, E. and Marani, M.: Downscaling of rainfall extremes from satellite observations, Water Resour. Res., 55, 156–174, https://doi.org/10.1029/2018WR022950, 2019.

Zorzetto, E. and Marani, M.: Extreme value metastatistical analysis of remotely sensed rainfall in ungauged areas: Spatial downscaling and error modelling, Adv. Water Resour., 135, 103 483, https://doi.org/10.1016/j.advwatres.2019.103483, 2020.

Zorzetto, E., Botter, G., and Marani, M.: On the emergence of rainfall extremes from ordinary events, Geophys. Res. Lett., 43, 8076–8082, https://doi.org/10.1002/2016GL069445, 2016.

Önöz, B. and Bayazit, M.: Effect of the occurrence process of the peaks over threshold on the flood estimates, J. Hydrol., 244, 86–96, https://doi.org/10.1016/S0022-1694(01)00330-4, 2001.