Hydrological Sciences Journal

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/thsj20

The legacy of STAHY: milestones, achievements, challenges, and open problems in statistical hydrology

Elena Volpi, Salvatore Grimaldi, Amir Aghakouchak, Attilio Castellarin, Fateh Chebana, Simon Michael Papalexiou, Hafzullah Aksoy, András Bárdossy, Antonino Cancelliere, Yuanfang Chen, Roberto Deidda, Uwe Haberlandt, Ebru Eris, Svenja Fischer, Félix Francés, Dmitri Kavetski, Thomas Rodding Kjeldsen, Krzysztof Kochanek, Andreas Langousis, Luis Mediero Orduña, Alberto Montanari, Sofia D. Nerantzaki, Taha B. M. J. Ouarda, Ilaria Prosdocimi, Elisa Ragno, Chandra R. Rajulapati, Ana Isabel Requena, Elena Ridolfi, Mojtaba Sadegh, Andreas Schumann & Ashish Sharma

To cite this article: Elena Volpi, Salvatore Grimaldi, Amir Aghakouchak, Attilio Castellarin, Fateh Chebana, Simon Michael Papalexiou, Hafzullah Aksoy, András Bárdossy, Antonino Cancelliere, Yuanfang Chen, Roberto Deidda, Uwe Haberlandt, Ebru Eris, Svenja Fischer, Félix Francés, Dmitri Kavetski, Thomas Rodding Kjeldsen, Krzysztof Kochanek, Andreas Langousis, Luis Mediero Orduña, Alberto Montanari, Sofia D. Nerantzaki, Taha B. M. J. Ouarda, Ilaria Prosdocimi, Elisa Ragno, Chandra R. Rajulapati, Ana Isabel Requena, Elena Ridolfi, Mojtaba Sadegh, Andreas Schumann & Ashish Sharma (2024) The legacy of STAHY: milestones, achievements, challenges, and open problems in statistical hydrology, Hydrological Sciences Journal, 69:14, 1913-1949, DOI: 10.1080/02626667.2024.2385686

To link to this article: https://doi.org/10.1080/02626667.2024.2385686

Published online: 25 Sep 2024.

HYDROLOGICAL SCIENCES JOURNAL
2024, VOL. 69, NO. 14, 1913–1949
https://doi.org/10.1080/02626667.2024.2385686

FEATURED ARTICLE

The legacy of STAHY: milestones, achievements, challenges, and open problems in statistical hydrology

Elena Volpi[a], Salvatore Grimaldi[b], Amir Aghakouchak[c], Attilio Castellarin[d], Fateh Chebana[e], Simon Michael Papalexiou[f], Hafzullah Aksoy[g], András Bárdossy[h], Antonino Cancelliere[i], Yuanfang Chen[j], Roberto Deidda[k], Uwe Haberlandt[l], Ebru Eris[m], Svenja Fischer[n], Félix Francés[o], Dmitri Kavetski[p,q], Thomas Rodding Kjeldsen[r], Krzysztof Kochanek[s], Andreas Langousis[t], Luis Mediero Orduña[u], Alberto Montanari[d], Sofia D. Nerantzaki[v], Taha B. M. J. Ouarda[e], Ilaria Prosdocimi[w], Elisa Ragno[x], Chandra R. Rajulapati[y], Ana Isabel Requena[z], Elena Ridolfi[aa], Mojtaba Sadegh[ab], Andreas Schumann[ac] and Ashish Sharma[ad]

[a]Department of Civil, Computer Science and Aeronautical Technologies Engineering, Roma Tre University, Roma, Italy; [b]DIBAF Department, Tuscia University, Viterbo, Italy; [c]Department of Civil and Environmental Engineering, Center for Hydrometeorology and Remote Sensing, University of California, Irvine, California, USA; [d]Department DICAM, University of Bologna, Bologna, Italy; [e]Eau Terre Environment Research Centre, National Institute for Scientific Research, Quebec, Canada; [f]Department of Civil Engineering, University of Calgary, Calgary, Canada; [g]Department of Civil Engineering, Istanbul Technical University, Turkey; [h]Institute for Modelling Hydraulic and Environmental Systems, University of Stuttgart, Stuttgart, Germany; [i]Department of Civil Engineering and Architecture, University of Catania, Catania, Italy; [j]College of Hydrology and Water Resources, Hohai University, Nanjing, China; [k]Department of Civil and Environmental Engineering and Architecture, University of Cagliari, Cagliari, Italy; [l]Institute of Hydrology and Water Resources Management, Leibniz University of Hannover, Hannover, Germany; [m]Department of Civil Engineering, Ege University, Izmir, Turkey; [n]Hydrology and Environmental Hydraulics, Wageningen University & Research, Wageningen, The Netherlands; [o]Research Institute of Water and Environmental Engineering, Universitat Politècnica de València, Valencia, Spain; [p]School of Civil, Environmental and Mining Engineering, University of Adelaide, Adelaide, Australia; [q]Institute of Environmental Hydraulics IHCantabria, University of Cantabria, Santander, Spain; [r]Department of Architecture and Civil Engineering, University of Bath, Bath, UK; [s]Faculty of Building Services, Hydro and Environmental Engineering, Warsaw University of Technology, Warszawa, Poland; [t]Department of Civil Engineering, University of Patras, Patras, Greece; [u]Department of Civil Engineering: Hydraulic, Energy and Environment, Universidad Politécnica de Madrid, Madrid, Spain; [v]Department of Civil, Geological and Environmental Engineering, University of Saskatchewan, Saskatoon, Canada; [w]Department of Environmental Sciences, Informatics and Statistics, Ca' Foscari University of Venice, Venice, Italy; [x]Faculty of Civil Engineering and Geosciences, Delft University of Technology, Delft, Netherlands; [y]Department of Civil Engineering, University of Manitoba, Winnipeg, Canada; [z]Centre of Hydrographic Studies of the Centre for Studies and Experimentation of Public Works, Spanish Ministry of Transport Mobility and Urban Agenda, Madrid, Spain; [aa]Department of Civil, Constructional and Environmental Engineering, Sapienza University of Rome, Rome, Italy; [ab]Department of Civil Engineering, Boise State University, Boise, Idaho, USA; [ac]Faculty for Civil and Environmental Engineering, Ruhr-University Bochum, Bochum, Germany; [ad]Civil & Environmental Engineering, University of New South Wales, Sydney, Australia

ABSTRACT
Statistical tools are crucial for a variety of hydrological applications, whether to model processes and enhance understanding and knowledge or to design infrastructure systems. Given the rapid evolution of statistical methods and the need for a solid theoretical foundation for their correct application, a multidisciplinary community STAtistics in HYdrology Working Group (STAHY-WG) aggregated under the International Association of Hydrological Sciences (IAHS) umbrella to contribute to this research field. Now, more than 15 years since its inception, this paper summarizes the main achievements of this productive community collaboration in four (of many) branches of statistical hydrology: extreme value analysis, multivariate analysis, time series analysis, and regionalization. The aim is to provide an overview of recent developments, offer practical suggestions (e.g. software packages), and outline future challenges to support scientists and practitioners in their endeavours within the realm of statistical hydrology studies.

ARTICLE HISTORY
Received 19 September 2023
Revised 24 April 2024
Accepted 1 July 2024

EDITOR
S. Archfield

ASSOCIATE EDITOR
(not assigned)

KEYWORDS
ICSH; STAHY; extreme value analysis; GEV; metastatistical distributions; copula function; time series analysis; FARMA model; changing point analysis; rainfall simulations; regionalization

1 Introduction

Statistical methods for analysis, synthesis, and modelling of hydrological data have a long history and continue to be a topic of intense research. Such tools have proved to be very effective and useful in numerous applications. The effectiveness of the statistical description of hydrological processes reflects the enormous complexity of hydrological systems, which often makes a purely deterministic description ineffective (Koutsoyiannis 2021); indeed, all hydrological processes reflect a combination of both deterministic and stochastic elements (Vogel 1999). A clear exemplification is provided by Farmer and Vogel

CONTACT Elena Volpi elena.volpi@uniroma3.it Department of Civil, Computer Science and Aeronautical Technologies Engineering, Roma Tre University, V. Vito Volterra. 62, Rome 00146, Italy

This article has been corrected with minor changes. These changes do not impact the academic content of the article.

(2016), who document systematic bias in the estimation of design events, droughts and floods, when a strictly deterministic approach based on watershed simulation models is employed.

In the last few decades, the number of available statistical tools, approaches, and procedures in several scientific fields has been increasing faster than before. The correct application of classical, updated, and new methods has always been fundamental for hydrological applications; moreover, for a single hydrological application, there are many potential statistical approaches available. As a consequence, in 2007 the "STAtistics in HYdrology" (STAHY) Working Group of the International Association of Hydrological Sciences (IAHS) was launched as a virtual common space aimed at synthesizing the enormous amount of information and number of resources present in the literature and beyond. The mission of this group was to coordinate, optimize and concentrate resources with the aims for statisticians to understand hydrological applications, for hydrologists to understand and appropriately apply statistical tools and understand what is the correct or best approach, and for statistical hydrologists (who have peculiar expertise in the application and development of statistical methods for hydrological data) to easily be updated on recent advances in their research field. Thus, the working group contributed to advancing and promoting hydrological sciences worldwide, in line with the IAHS mission.

The STAHY Working Group successfully promoted and organized several initiatives towards the above directions. Contributions included several sessions at IAHS or International Union of Geodesy and Geophysics (IUGG) General Assembly, annual STAHY Workshops, and several summer schools and short courses contributing to the development of new generations of statistical hydrologists; STAHY also collected recent scientific studies on statistical hydrology and awarded yearly the best ones (STAHY Best Paper Award). During the XXV IUGG General Assembly held in Melbourne Australia in 2011, the STAHY Working Group was transformed into the International Commission on Statistical Hydrology (ICSHIAHS) to permanently give its contribution to the scientific community. Ever since, ICSH has brought together people who wanted to actively collaborate by sharing knowledge, information, papers, data, and numerical routines. Up to the beginning of 2023, more than 200 researchers from about 60 countries, from six continents, have joined the commission.

The operational idea of ICSH was to focus attention on and gather people around some main topics, emerging from the interests of or explicitly suggested by the involved community. The most frequently discussed subjects in the ICSH-STAHY community from the very beginning are: (i) extreme value analysis, (ii) multivariate analysis, (iii) single or multiple time series analysis and modelling, and (iv) regionalization. Extreme events are important hydrological phenomena with serious societal consequences; their probability of being exceeded (or not) is determined in a univariate (i) or a multivariate (ii) framework, depending on the specific conditions. However, for general hydrological purposes, the statistical prediction of the magnitude and occurrence of extreme events may not be sufficient, and we need to model the temporal (iii) or spatial (iv) evolution of the entire hydrological process, considering its correlation structure. Obviously, these are only some of the topics of interest to the statistical hydrology community. Providing a comprehensive compendium of all or at least majority of the research areas related to or derived from those described here is far beyond the scope of this work.

Moreover, the ICSH-STAHY community actively contributes to many of the questions listed in the Unsolved Problems in Hydrology (UPH) community initiative (Blöschl *et al.* 2019a), and, among them, the following ones are directly linked to ICSH-STAHY activities (even if almost all questions require the application of statistical tools); note that the questions are reported here with reference to their number in the UPH list.

(1) *Time variability and change*: 1. Is the hydrological cycle regionally accelerating/decelerating under climate and environmental change, and are there tipping points (irreversible changes)?
(2) *Space variability and scaling*: 6. What are the hydrological laws at the catchment scale and how do they change with scale?
(3) *Variability of extremes*: 9. How do flood-rich and drought-rich periods arise, are they changing, and if so why?
(4) *Measurements and data*: 16. What is the relative value of traditional hydrological observations versus soft data (qualitative observations from lay persons, data mining etc.), and under what conditions can we substitute space for time?
(5) *Modelling methods*: 20. How can we disentangle and reduce model structural/parameter/input uncertainty in hydrological prediction?
(6) *Interfaces with society*: 21. How can the (un)certainty in hydrological predictions be communicated to decision makers and the general public?

Now, 15 years since its establishment, the ICSH-STAHY community, represented by the authors of this work, aims to summarize the current state of the art in the aforementioned topics. Hence, the objective of this paper is to offer a critical review of the progress in the past 15 years, present standardized methods and procedures for various applications, and propose insights for the future advancement of statistical hydrology. Note that the authors of this work are past and current officers of the ICSH-STAHY, as well as some members of the commission who were actively involved in the organization of recent initiatives. We are aware that the present work cannot be exhaustive, nor do we pretend to cover all relevant issues, of which there are many. However, we will mention, when relevant, some other topics or emerging areas of research that are more or less closely related to those discussed here (such as Bayesian methods, stochastic rainfall–runoff modelling, machine learning, etc.). Cross-cutting, emerging issues may be discussed in detail in a possible future development of this work (a series of manuscripts dedicated to specific issues).

The remainder of this paper is organized according to the four main topics mentioned above; the following sections are each dedicated to one of the four research areas. For each of them, we present a brief state of the art, the milestones for driving practical applications, and the main open problems following the vision of the ICSH-STAHY community, in the eyes of the group of people authoring this work. Finally, each section providesa table with some of the most popular Matlab® and R (R Core Team 2020) packages for statistical analysis and modelling. Such packages can be accessed and downloaded either from The Comprehensive R Archive Network (CRAN), i.e. the official Rproject package archive, or from other open-access repositories.

2 Extreme value analysis

When an extreme event occurs, key questions arise: Is it going to happen again? How often? Can information from past events help with the prediction of the chance of future extremes? Answering these questions has led to significant advances in probability theory over the past century (see e.g. Coles 2001, Koutsoyiannis 2021). Indeed, human history has been shaped by and evolved also in response to natural disasters. In addition, anthropogenically induced change is expected to amplify the magnitude, frequency and effects of extreme events in the future.

The probability of an event (here, extreme flood, rainfall, drought) with a given magnitude can be broadly defined as the chance that the process of interest can exceed a certain high value (see Koutsoyiannis 2021 and references therein). Contemporary frequency analysis in hydrology, as it is known today, started with Hazen in 1930 and was popularized by Gumbel (1958). The study on the concept of probability led to developing many distribution functions and answering several basic scientific and applied questions (Kelley 1994, Hald 2005). Hydrological frequency analysis broadly encompasses a set of statistical methods and techniques to link the magnitude of an extreme event to its frequency or chance of exceedance through a probability distribution function (see e.g. Stedinger *et al.* 1993, El Adlouni *et al.* 2010, Camuffo *et al.* 2020), including distributions with an upper limit of the population, which corresponds to the concept of the probable maximum precipitation (PMP) or the probable maximum flood (PMF) (e.g. the recent Salas *et al.* 2020).

Extreme value analysis is well established in the literature and has a long history in hydrological theory and practice, as recalled above. Notwithstanding this, continued advances of hydrological knowledge incessantly open new questions, requiring innovative statistical tools to address them. In the following sections we briefly recall the most common approaches to extreme value analysis and focus on the latest open problems.

2.1 Common methodological approaches for extreme events frequency analysis

Typically, extreme value analysis involves the following five steps (e.g. Rao and Hamed 2019, Katz *et al.* 2002):

- *Exploratory analysis*. Since hydrological data are generally asymmetric and the interest is in the distribution tail (where extremes occur), skewness and kurtosis are of particular interest in exploratory data analyses. For example, being a good indicator of a distribution's shape, the skewness is useful to guide the selection of a representative distribution. It is relevant to mention that it is generally more convenient to use L-moments, L-skewness and L-kurtosis, instead of ordinary moments, for this scope, as demonstrated in the seminal paper by Hosking (1990), as well as the comparative assessment in Vogel and Fennessey (1993). Another interesting approach is that recently proposed by Koutsoyiannis (2021) based on the so-called knowable moments. On the other hand, outliers can negatively affect the selection of an appropriate distribution (see e.g. Lamontagne *et al.* 2016). For a summary of both parametric and nonparametric methods for exploratory data analysis, the reader is referred to Helsel and Hirsch (1992) and Helsel *et al.* (2020).
- *Testing assumptions*. This step is important to ensure the validity of the subsequent steps. A number of homogeneity and stationarity testing methods are available in the literature; readers are referred to Naghettini (2017) for a review of the methods and to Serinaldi *et al.* (2018), among others, for a discussion on trend tests (as briefly discussed in Section 4.3).
- *Modelling by probability distribution functions*. Fitting a probability distribution function to the observed data is perhaps the most developed step in hydrology and statistical theory literature; the remainder of this section is indeed dedicated to pointing out the most common modelling approaches.
- *Quantile and return period estimation*. The notions of quantile and return period (e.g. Fernández and Salas 1999, Veneziano *et al.* 2007, Langousis *et al.* 2009, Volpi 2019) for hydrological events are commonly used to describe the severity of extremes for a wide range of applications including hydraulic infrastructure design and risk assessment (e.g. Singh and Strupczewski 2002, Vogel and Castellarin 2017, among others).
- *Uncertainty estimation*. The evaluation of the uncertainty bounds (see e.g. Coles 2001), which depends on the length of the observed data and on modelling choices, is of paramount importance to understand and quantify the reliability for subsequent risk assessment and design. Note, however, that additional uncertainty components with respect to uncertainty bounds are generally needed for risk-based decision making (see e.g. Vogel and Castellarin 2017).

The fundamental theory of distributions for hydrological frequency analysis of extreme events relies on the early work of Maurice Fréchet who showed in 1927 the asymptotic distributions of sampled extremes, which are typically selected as the annual maxima. After Fisher's and Tippett's work in 1928, and Gnedenko's work in 1943, the Fréchet, Gumbel, and Weibull probability distributions emerged as limiting distributions of extremes (maxima sampled from fixed-length sequences of independent random variables, referred to also as block maxima). The combination of Fréchet, Gumbel, and Weibull resulted in the generalized extreme value (GEV) distribution (Coles 2001), which is characterized by the location, scale, and shape parameters; the latter determines not only the shape but also the tail behaviour of the GEV distribution. As previously mentioned, in environmental applications, the most commonly applied approach is composed of maxima extracted from each Elena Volpi-year-long block of data, simply named annual maxima (AM). This assumption is a commonly accepted compromise in order to retain a reasonable number of observations free of seasonal and dependence biases.

Under a different framework aiming to enlarge the sample dimension, but with some analogies to the asymptotic properties of GEV, it can be proved that the generalized Pareto (GP) is the expected asymptotic distribution of the rescaled exceedances above high enough thresholds, namely the peaks-over-threshold (POT), regardless of the distribution of the underlying process (Coles 2001). The more complex POT approach requires the proper definition of a high enough threshold to filter the data and calculate the frequency of the rescaled excesses (e.g. Deidda 2010, Langousis *et al.* 2016a). The POT approach makes use of the GP distribution, which is ruled by the threshold, scale parameter, and, similarly to GEV, shape parameter; again, the latter controls the tail behaviour of the distribution.

A large number of studies apply the GEV and GP distributions to analyse extremes (see e.g. Coles 2001, El Adlouni *et al.* 2007, Katz 2013, Tyralis *et al.* 2019, Emmanouil *et al.* 2020 among many others). In particular, Martins and Stedinger (2000) introduce the generalized maximum likelihood (ML) estimation for the GEV model; Hosking and Wallis (1987) provide seminal analyses and comparisons of both parameter estimation and quantile estimators for the GP model. Stedinger *et al.* (1993) (see also Madsen *et al.* 1997) summarize the various rigorous and classical studies, at that time, which led to the recommendation of POT over AM approaches when the arrival rate for the POT is large enough (1.65 events/year for Poisson arrivals with exponential exceedances). In such instances, POT analyses would yield more accurate estimates of extreme quantiles. O'Shea *et al.* (2023) argue that the POT series is better characterized using the four-parameter Kappa (or the combination of the GP and binomial) for the estimation of rare to veryrare design extremes. Along with GEV and GP, many other distribution functions are used for frequency analysis of extremes, like Log-Pearson Type III (e.g. Griff and Stedinger 2007), Burr type (e.g. Zaghloul *et al.* 2020) or kappa distributions (see e.g. Hosking and Wallis 1997 or Kjeldsen *et al.* 2017).

Frequency models can nevertheless differ not only in the choice of the shape of the distribution function, but also in the choice of the estimation method used to fit the model to the observed data. For a comprehensive review of the topic, including several examples of flood and rainfall extreme events, readers are referred to Nerantzaki and Papalexiou (2022).

It is noteworthy that there are no theoretical reasons that justify the apriori assumption of a single specific distribution under non-asymptotic conditions; thus, the adoption of a specific model is generally motivated by its ability to robustly represent the available observed data (see e.g. Laio 2004, El Adlouni *et al.* 2008, Calenda *et al.* 2009, orLaio *et al.* 2009for possible examples of model selection criteria). Note that even in the case of high accuracy in reproducing the observed data (which in fact hardly ever happens), any adopted model provides only uncertain, and possibly biased estimates of higher extremes (e.g. in extrapolation) because the limited length of the available samples provides only poor information about rare events (Klemeš 2000a, 2000b). Special attention should be paid to the tail behaviour (large or small extreme events) of the modelling distribution (El Adlouni *et al.* 2008, Kochanek *et al.* 2020; see also Merz *et al.* 2022 and references therein). In addition to parametric models, which can be easily applied in extrapolation, a wide range of distribution-free, non-parametric, semi-parametric or kernel-based methods have also been developed in the literature (see e.g. Lall 1995, Rao and Hamed 2019, Banfi *et al.* 2022).

Finally, Table 1 reports a selected list of packages already available for extreme events frequency analysis, mainly (but not only) R packages available from the R Core Team (2020).

2.2 Incorporating additional information: historical events, seasonality and time dependence of the underlying process

Besides the two classical approaches for considering extreme events in hydrology mentioned in the previous section (Section 2.1), namely AM and POT, several extensions exist that aim at increasing the information incorporated in the samples. Indeed, both these classical approaches lead to formulations neglecting a significant proportion of the observations, discarding the information contained in the bulk of the parent distribution along with most of the observations. The bulk of the distribution may add information on the extreme events assuming that they pertain to the same process (as in GEV and POT derivations). In addition, extreme events may be differentiated by the season they occur in (see e.g. Baratti *et al.* 2012, Kochanek *et al.* 2012, Strupczewski *et al.* 2012 for applications in flood frequency analysis, and Mascaro 2018 for extreme rainfall events). Another development is the use of flood-type-specific samples that are considered in a joint mixture model (as in Hirschboek 1987, Fischer 2018, Fischer and Schumann 2021). These subsamples are assumed to be identically distributed and more homogeneous than the annual or partial duration series since the events can be assumed to have a common genesis.

In general, any additional information about the process of interest notably improves return period estimates. In the last few decades, several researchers have evaluated the possibility of extending the data to past, non-systematically recorded events, leading to a significant increase in the length of the available datasets, with benefits towards the frequency analysis of hydrological extremes. This non-systematic data can have two different sources: historical and palaeoflood information. The use of historical information in the estimation of flood quantiles was first described by Benson (1950) and Leese (1973). Techniques for obtaining information from palaeofloods were first introduced by Costa (1978) and Kochel and Baker (1982).

Stedinger and Cohn (1986), Cohn and Stedinger (1987) and Frances *et al.* (1994) systematized the use of these types of information and demonstrated their enormous advantage in reducing the uncertainty of the estimated quantiles. This has led to the development of historical hydrology (see e.g. Benito *et al.* 2004, Macharo *et al.* 2015), that involves the use of

Table 1. Packages for extreme events frequency analysis and modelling (see also the CRAN task view: hydrological data and modelling https://cran.rproject.org/web/views/Hydrology.html).

Name	Brief description	Repo.	Link	Authors
ProNEVA	Process-informed nonstationary extreme value analysis	UCI.edu	https://amir.eng.uci.edu/software.php	AghaKouchak *et al.*
NEVA	Nonstationary extreme value analysis	UCI.edu	https://amir.eng.uci.edu/software.php	AghaKouchak *et al.*
MEV	Modelling of extreme values	CRAN	https://cran.r-project.org/package=mev	Belzile *et al.*
UKFL	UK flood estimation	CRAN	https://CRAN.R-project.org/package=UKFE	Hammond
PeakFQ	Flood frequency analysis based on US Bulletin 17C	USGS	http://water.usgs.gov/software/PeakFQ/	US Geological Survey (USGS)
Afins	Unbounded and bounded distributions with nonsystematic information	UPV	http://lluvia.dihma.upv.es/EN/software/software.html	Botero and Francés
GAMLSS	Generalized additive models for location, scale and shape	CRAN	https://cran.R-project.org/package=gamlss	Stasinopoulos *et al.*
extRemes	General functions for performing extreme value analysis	CRAN	https://cran.R-project.org/package=extRemes	Gilleland
lmom	Functions related to L-moments: computation of L-moments and trimmed L-moments of distributions and data samples; parameter estimation	CRAN	https://CRAN.R-project.org/package=lmom	Hosking
LMoFit	Advanced L-moment fitting of distributions	CRAN	https://CRAN.R-project.org/package=LMoFit	Zaghloul *et al.*
LMOMCO	L-moments, censored L-moments, trimmed L-moments, LCo-moments, and many distributions	CRAN	https://CRAN.R-project.org/package=lmomco	Asquith
USGSR/smwrStats	R functions to support statistical methods in water resources	CRAN	https://rdrr.io/github/USGS-R/smwrStats/	USGS

historical–archival methods, of hydrological modelling and stochastic frequency analysis (from Stedinger and Cohn 1986 to e.g. Francés *et al.* 1994, Naulet *et al.* 2005, Reis and Stedinger 2005, Calenda *et al.* 2009, Botero and Francés 2010, Blöschl *et al.* 2020, Saint Criq *et al.* 2022, Ostrowski *et al.* 2023, and many more).

A promising technique for improving the reliability of high return period quantiles is based on a better understanding of the emergence of extreme events from the bulk of the distributions, as suggested by Marani and Ignaccolo (2015) and Zorzetto *et al.* (2016) or differently by Volpi *et al.* (2019). This might require relaxing the asymptotic hypothesis or the underlying, common hypothesis of independence of sample data used for inference and projection. The extension of frequency analysis and return period estimate for time-dependent data was recently discussed by Volpi *et al.* (2015), Serinaldi and Kilsby (2016, 2018) and Serinaldi and Lombardo (2020) and references therein.

In particular, the metastatistical extreme value (MEV) approach by Marani and Ignaccolo (2015) relies on the assumption that the extreme events are block maxima among a finite and stochastically variable number of *ordinary events* from an underlying and possibly time-varying parent distribution. Then, the MEV approach relaxes the limiting assumptions of the classical extreme value theory by considering as random variables the parameters defining the number of ordinary events and the probability distribution of event magnitudes. This allows the use of the entire observational set to infer the distribution of extremes and significantly reduce the estimation uncertainty (see e.g. Marra *et al.* 2023).

Another way to increase the information acquired from samples of limited length is to apply statistical models that simultaneously incorporate information from a wide range of spatiotemporal scales. Under this setting, during the last four decades, scaling representations of hydrological processes have attracted much attention, with particular emphasis on extreme estimation under asymptotic (e.g. Veneziano and Langousis 2005a and the review in Veneziano *et al.* 2006) and pre-asymptotic (e.g. Langousis *et al.* 2013, Emmanouil *et al.* 2020, 2022, 2023, Grimaldi *et al.* 2022) conditions.

Stochastic models (see Section 4) have also been used to extract information from samples of limited lengths. The idea is to calibrate a model that is capable of generating replicas of observed time series, therefore obtaining multiple realizations of the considered stochastic process. These models were pioneered by the contributions of Andrey Markov (1856–1922) and then widely used by hydrologists since the second half of the 20th century (see e.g. Thomas and Fiering 1962). The use of stochastic models for inferring extreme values has been problematic for the complexity associated with the simulation of the distribution tails of non-Gaussian processes. However, recent scientific contributions opened promising doors for resolving such limitations (Papalexiou and Serinaldi 2020; see also Section 4). In particular, stochastic streamflow models are needed for risk-based hydrological management methods because such models can generate the ensembles needed for such planning exercises. Vogel (2017, section 2) gives a historical perspective on the application of such models in hydrology and provides arguments regarding the need for a new generation of stochastic watershed models for generating such streamflow ensembles, particularly when purely statistical/machine learning models are inadequate to the task, because they may not be able to incorporate explicitly the impact of anthropogenic influences.

2.3 Changing extremes: stationarity versus non-stationarity

A fundamental assumption for the extreme value analysis is that the random variable of interest should be independently and identically distributed (i.i.d.). Note that, as discussed in the previous section, while the independence assumption can be relaxed, the identical distribution implies that the statistical properties of the historical data, i.e. the relationship between the magnitudes and return periods based on frequency analysis of data, should be invariant and thus representative of future events. The longer the observed data series, the higher the chance that historical extremes represent a reasonable sample for what is expected to occur in the future (Klemeš 2000a, 2000b). However, if the statistics of future extremes are expected to significantly vary from the past, return periods and occurrence probabilities estimated based on historical observations may not be representative of the future scenarios.

While hydrological extremes are expected to show significant natural variability, anthropogenic activities including greenhouse gas emissions and land use–land cover changes are expected to alter the magnitude and severity of extreme events (e.g. Milly *et al.* 2008, 2015, Chiang *et al.* 2021b). Over the past two decades, numerous studies have pointed to more intense and frequent extreme rainfall events (e.g. Alexander *et al.* 2006, Westra *et al.* 2014, Fischer and Knutti 2016, Mallakpour and Villarini 2017, Ragno *et al.* 2018, Farris *et al.* 2021, Emmanouil *et al.* 2022, 2023), changes in the mean and variability of river flows (e.g. Blöschl *et al.* 2017, 2019b, Hodgkins *et al.* 2017), and sea level rise with implications for severe coastal flooding (e.g. Vermeer and Rahmstorf 2009, Wahl *et al.* 2015). As for floods, changes in both hydroclimatology and land use showed a strong impact; as an example, in a comparative national assessment on the magnification of floods, Vogel *et al.* (2011) found that urbanization impacts were at that time more severe (led to higher magnification factors) and obvious (common) than impacts due to climate change.

Current operational procedures, risk assessment methods and design guidelines are based on the so-called "stationarity" assumption, which implies invariant properties of hydrological extremes from historical records. In other words, a stationary approach assumes that the properties of the underlying stochastic model (i.e. all finite-dimensional distributions, including the statistics of extremes) do not change significantly relative to time or another physically-based variable. Indeed, change in the observed data does not necessarily imply a nonstationary underlying process; and stationarity is also related to ergodicity, which in turn is a prerequisite for making inferences from data (Cohn and Lins 2005, Koutsoyiannis and Montanari 2015). Furthermore, a nonstationary framework cannot be generally inferred from the observed data alone, i.e. without a broad a priori discussion about the physical reasons and the expected change (Serinaldi and Kilsby 2015, Koutsoyiannis 2016, Luke *et al.* 2017).

A non-stationary assumption corresponds to the changing properties of the hydrological extremes over time in response to a physically-based process. It is generally suggested to use as covariates for change the forcings of that change through climate, land use and/or reservoir indices instead of time (Villarini *et al.* 2009b, Katz 2013, López and Francés 2013, Ragno *et al.* 2019). Following a popular publication on the "death of the stationarity assumption" due to significant human activities (Milly *et al.* 2008), a lively debate emerged in the literature on the validity of the stationary and non-stationary approaches (Lins and Cohn 2011, Matalas 2012, Koutsoyiannis and Montanari 2015, Milly *et al.* 2015, Serinaldi and Kilsby 2015, Luke *et al.* 2017, Ragno *et al.* 2019). Some key questions are:

- Is the commonly used stationary method (or alternative versions still based on the same fundamental assumptions) sufficient for analysis of future extremes?
- Do we need a new paradigm for analysis of changing extremes?
- How can we confidently decide whether the statistics of extremes have changed or not?
- Is non-stationarity a property of the natural system (including human interactions) or simply a property of a numerical model?
- Do we have sufficient observations to test and develop nonstationary models for analysis of hydrological extremes?
- What is the predictability power of a nonstationary framework?

This is not an exhaustive list of all the questions; on the contrary, the list of relevant questions and concerns is long and still growing (e.g. Koutsoyiannis 2020). The purpose of this section is not to take a position regarding which approach is more justified in a changing environment. Instead, we highlight that this area of research still deserves more in-depth exploration and model development. Regarding the methods for change analysis of the observed data, we refer the reader to Section 4.3.

The non-stationary assumption is typically implemented through changing one or more parameters of the corresponding extreme value distribution with respect to an underlying physical process driving change. Time is often used as a surrogate for other drivers, but the use of appropriate physically-based covariates instead of time alone should be preferred, as previously mentioned. In the case of GEV, for example, one can allow the location, scale and/or shape parameters to vary as a (e.g. linear) function of one or more covariates. The shape parameter is the most sensitive one and it is difficult to estimate accurately when the availability of observations is limited (see e.g. Coles 2001, El Adlouni *et al.* 2007, Papalexiou and Koutsoyiannis 2013, Strupczewski *et al.* 2016, Deidda *et al.* 2021). Ouarda and Charron (2019) showed that the shape parameter can evolve as a function of time and low-frequency climate oscillation indices and that its evolution affects considerably the estimates of extreme events. The appropriateness of the non-linear function and its validity in the future are among the major sources of uncertainty (Koutsoyiannis 2020). As pointed out by Prosdocimi and Kjeldsen (2021), combining time-varying parameters can lead to counterintuitive behaviour in the extreme quantiles of interest for hydrological design. Hence, a broad discussion on what types of changes can be apriori expected under different conditions is necessary to clarify what parameter model structures are more apt to capture the expected changes (see e.g. Sharma *et al.* 2018, Hecht and Vogel 2020).

Regardless of the choice of parameter model structure, the outcome will be a distribution function that changes over time due to varying parameters. Ouarda *et al.* (2020) presented a comparison of the uncertainties in stationary and nonstationary extreme rainfall models and formulated words of caution about the use of nonstationary models with relatively small-size data samples. In recent studies, Prosdocimi *et al.* (2015) and Bertola *et al.* (2019) among many others showed how physical processes can be integrated into regional and at-site nonstationary analysis of extremes such as changes in flooding as a function of land use change, increases in temperature as a function of CO_2 concentration in the atmosphere, and changes in rainfall and snowmelt patterns, providing a deterministic justification for non-stationary statistics.

Despite progress in this area (summarized broadly bySlater *et al.* 2021b), in most places around the world, official codes and guidelines do not consider changing extremes. For the first time, the American Society of Civil Engineers (Committee on Adaptation to a Changing Climate 2018) published a manual of practice including a guideline for considering the observed and projected changes in extreme precipitation events assessment based on Ragno *et al.* (2018) for design and risk. To avoid assumptions associated with changing trends in the future (e.g. Koutsoyiannis 2020), the guideline recommends using a wide range of future projections to update historical precipitation intensity–duration–frequency (IDF) curves. This approach, presented here as an example, allows for quantifying the changes in the frequency of past events (or return period of a historical event), when historical and projected IDF curves are available. Figure 1 displays a historical IDF curve (blue line) and the projected IDF curves derived from a wide range of climate model simulations (red ensemble) in San Francisco, California (USA), as presented in Ragno *et al.* (2018). Given that this method is based on a wide range of models, one can derive the 5th and 95th percentiles as a measure of uncertainty (Fig. 1). Needless to say, when comparing the expected value from future projections and historical observations, the most conservative one should be considered for design and risk assessment. Given the uncertainties associated with future projections, the difference between future projections and historical observations can be very large.

Ouarda *et al.* (2019) proposed an IDF model in which the parameters depend on time and low-frequency climate oscillation indices and are estimated with the maximum composite likelihood method. Results indicated that the non-stationary IDF framework provides a better fit to the data and leads to more robust estimates. Other well-known approaches dealing with the assumption of non-stationarity are the effective return level (Katz *et al.* 2002) and the expected waiting time (Salas and Obeysekera 2014). The effective return level (or effective design value), proposed by Katz *et al.* (2002), is the quantile expressed as a function of a given covariate (i.e. time or physical). The expected waiting time, proposed by Salas and Obeysekera (2014), is the non-stationary return period of an event of interest derived as the expectation of a geometric distribution in which the probability of the first occurrence of the event of interest changes over time. Compared to other approaches, the expected waiting time is consistent with the definition of the return period in the stationary case.

To summarize, the extrapolation of non-stationary behaviours to the future should be done with caution, and only if the future can be predicted by using additional prior physical knowledge of the process (e.g. Prosdocimi *et al.* 2015, Serago

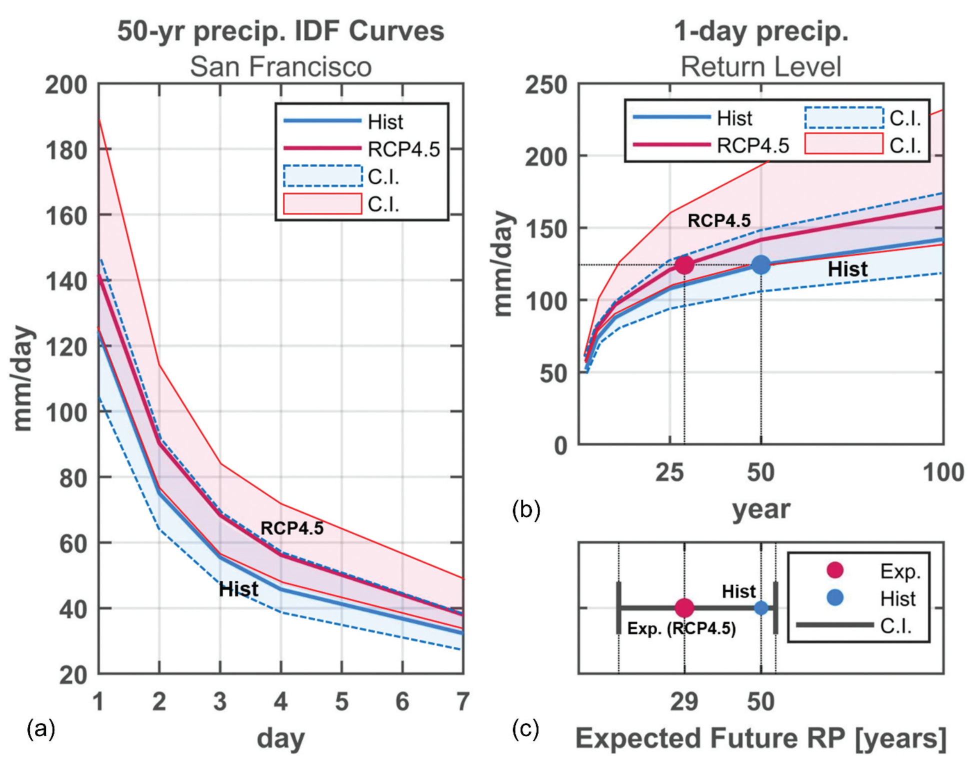

Figure 1. IDF curves together with their confidence intervals (C.I.) from multi-model climate simulations of the future (red) relative to the IDF curves from historical observations (blue): (a) intensity as function of the duration for $T = 50$ years; (b) 1-day precipitation as function of the return period; (c) expected return period according to climate simulation of the 1-day precipitation for $T = 50$ years (historical).

and Vogel 2018, Bertola *et al.* 2019, Ouarda *et al.* 2020). The additional information justifies the reduction of uncertainty that might result from explaining part of the variability in deterministic terms (Koutsoyiannis and Montanari 2015); otherwise, the effect would be that of introducing an additional source of uncertainty (Serinaldi and Kilsby 2015). Detailed discussion on the stationarity issue can be found in Vogel and Castellarin (2017) and Salas *et al.* (2018). Depending on the available information, different approaches can be applied, stationary and non-stationary; in any case, results should be evaluated in terms of both accuracy and uncertainty. In Table 1 some packages for non-stationary frequency analysis are listed.

3 Multivariate frequency analysis

Univariate frequency analysis deals with only one variable or feature of a hydrological phenomenon (e.g. drought intensity or flood peak); however, hydrological events are often characterized by several features that might be interrelated, e.g. flood peak, volume and duration, drought severity and duration, storm precipitation, intensity and duration, fluvial and oceanic floods, heatwave duration and severity, extreme discharge in two (or more) rivers in the same catchment. A multivariate frequency analysis framework may involve jointly modelling two or more features of the extreme events (e.g. flood volume and peak), or multisite analysis of extremes across space (e.g. flood peaks at several locations) or compound extreme events.

In the last two decades, multivariate frequency analysis has attracted increasing attention focusing on the joint treatment of different variables mainly using multivariate distributions (e.g. Yue *et al.* 1999) or copulas (e.g. Salvadori and De Michele 2004, Chebana 2022). In recent years, several studies, review papers and books have discussed multivariate hydrological frequency analysis methodologies and applications (e.g. Favre *et al.* 2004, Zhang and Singh 2006, 2019, Genest and Favre 2007, Salvadori *et al.* 2007, Ashkar and Aucoin 2011, Joe 2014, Genest and Chebana 2017). Some previous studies highlight the limitations of the univariate framework and justify the adoption of an alternative multivariate framework in hydrological applications (see Salvadori and De Michele 2004, Kao and Govindaraju 2007). This is particularly important for compound and cascading hazards and other interrelated hydrological extremes (e.g. Sadegh *et al.* 2018, Zscheischler *et al.* 2018, AghaKouchak *et al.* 2020).

The application of multivariate models in hydrology is not a new topic and numerous previous studies have employed multivariate normal, gamma and exponential distribution functions, among others, to understand and model the relationship among different variables. However, multivariate distributions have several limitations. For example, they require the marginal distributions of associated random variables to be within the same class, leading to only a limited number of the common distributions being extendable to a multivariate setting (e.g. Hao and Singh 2016). To overcome this drawback, in some cases a multivariate distribution can be derived analytically based on the nature of the process under investigation, e.g. for the case of droughts (Cancelliere and Salas 2010). An alternative and more general approach employs copula functions for hydrological frequency analysis to avoid the drawbacks of classical multivariate models (Nelsen 2006). In fact, copulas can model the dependence structure independently of the marginal distributions, making it possible to build multivariate distributions with different margins.

The literature on copulas and their application to hydrology is already substantial (for recent reviews see Hao and Singh 2016, Genest and Chebana 2017, Zhang and Singh 2019). Copulas are increasingly used in multivariate analysis, such as for precipitation (e.g. Grimaldi and Serinaldi 2006a, Sharma and Mujumdar 2019, Xu *et al.* 2020), droughts (e.g. Serinaldi *et al.* 2009, Vazifehkhah *et al.* 2019), floods (e.g. Grimaldi and Serinaldi 2006b, Durocher *et al.* 2016), river discharge estimates (e.g. Ragno *et al.* 2022), reservoir routing for dam design and safety assessment (Requena *et al.* 2013, Volpi and Fiori 2014), heatwaves (Mazdiyasni *et al.* 2019), storms and extremes (e.g. Corbella and Stretch 2013, Han *et al.* 2020) and multi-index drought assessment (e.g. Hao and AghaKouchak 2014). Copulas are also used for the joint modelling of extreme river temperature and low-flow conditions which can be harmful to aquatic life (Latif *et al.* 2023). They are also useful for multisite analysis (e.g. Serinaldi 2009).

In this section, we refer to the common methodological approaches used for multivariate frequency analysis based on copula functions. Considering multivariate hydrological frequency analysis also involves some challenges and issues (e.g. variable selection, choice of dimension, event selection, and dealing with more parameters), as it may require different definitions of a given statistical concept (e.g. return period) depending on the application, and often needs large data series compared to univariate analysis, as described in the last part of this section. Typically, multivariate frequency analysis involves the same four steps already described for the univariate setting in Section 2. Karahacane *et al.* (2020) is one of the rare papers dealing with most of these aspects simultaneously. We recall here some issues that are specific to multivariate analysis.

- *Exploratory analysis*. In a multivariate setting, Chebana and Ouarda (2011a) investigated this step and offered guidelines on how to explore the data prior to modelling. In addition, at this step, missing data and ties (if applicable) should also be treated (e.g. Ben Aissia *et al.* 2017), as well as sample dimension. There is no general recommendation concerning the minimum number of pairs or triplets for multivariate analysis; however, it is reasonable to avoid bivariate analysis with only 30–40 pairs.
- *Testing assumptions*. A number of multivariate homogeneity and stationarity testing methods are available in the literature. Specifically, multivariate trend tests can be found in Chebana *et al.* (2013), whereas multivariate shift testing (homogeneity) is discussed by Chebana *et al.* (2017) and Salvadori *et al.* (2018). A pivotal requirement, often not verified, is the serial independence condition that can easily be preliminarily checked on the autocorrelation structure of each sample (Chebana *et al.* 2013).
- *Modelling by probability distribution functions*. A large number of studies have focused on copula fitting and parameter estimation for a wide range of hydrological applications (e.g. Zhang and Singh 2006, Salvadori and

De Michele 2007, Kao and Govindaraju 2008, Requena *et al.* 2016). Given their importance, copulas are described in more detail in the following sections.

- *Multivariate quantile and return period analysis*. For multivariate return period analyses, one can refer, for instance, to Salvadori *et al.* (2007), Gräler *et al.* (2013), Serinaldi (2015), or Salvadori *et al.* (2016), whereas multivariate quantiles are investigated by Chebana and Ouarda (2011b). In summary, the multivariate return period is different from the univariate one, because in the multivariate setting the bijective relationship between return period and return level (or quantile) does not hold anymore. Volpi and Fiori (2012, 2014) treated design event selection, as a practical aspect, in the multivariate return period and quantile framework for risk assessment or infrastructure design. Note that the return period of structural failure does not necessarily match that of the hydrological load in a multivariate setting (see also Cipollini *et al.* 2021).
- *Uncertainty estimation*. As discussed below for the case of copula-based multivariate frequency analysis, uncertainty estimation in a multivariate setting still requires additional developments and implementations of practical applications (see e.g. Serinaldi 2013).

3.1 Copula functions

Copulas are an ingredient for constructing multivariate distributions with margins from different families. Basically, a copula is a multivariate distribution function with uniform margins. An attractive advantage of copulas is that the dependence between variables can be modelled separately from their marginal distributions (e.g. Nelsen 2006, Joe 2014). Another interesting feature is that multivariate analysis based on copulas can use all the common tools of univariate analysis.

In the following, we consider the bivariate case for simplicity even though the majority of the material presented below is available in higher dimensions. According to Sklar's theorem (Sklar 1959), the joint probability distribution of two random variables can be decomposed into two marginals as well as a copula to describe the dependence structure between the variables. The copula is unique when the marginals are continuous, which is a common assumption in hydrology. The copula captures the dependence structure between variables, and provides more information beyond descriptive dependence measures (e.g. Kendall's tau and Spearman's rho, defined below).

The Archimedean, meta-elliptical, and extreme value copula families are of particular interest in statistics as well as in hydrology. Lists and properties of different copulas can be found, for instance in Nelsen (2006), Salvadori *et al.* (2007), Joe (2014), and Zhang and Singh (2019); we recall in the following the main properties of some of the copulas.

- *Archimedean copulas*. This class is widely used in hydrology because: (i) its members are easy to construct and the parameter estimation is straightforward; (ii) this family includes a diverse set of copulas with different properties applicable to a wide range of data; and (iii) mathematically, they are elegant to treat. Archimedean copulas are based on a generator function, including Ali-Mikhail-Haq, Clayton, Frank, Joe and Gumbel-Hougaard, among others. Multi-parameter Archimedean copulas are also available and can be found in Joe (2014) in the general statistical context and e.g. in Sadegh *et al.* (2018) for hydrological applications. In hydrology and in multivariate analysis, the Archimedean copula family is probably the most widely used among different options (see, among many others, Kao and Govindaraju 2007, Chebana and Ouarda 2011b, Liu *et al.* 2020). Several studies have explored using Archimedean copulas at higher dimensions such as three and four (e.g. Grimaldi and Serinaldi 2006a, Zhang and Singh 2007a, 2007b); however, care should be taken in high-dimensional applications as the performance of most Archimedean copulas decreases as the dimension increases (Joe 2014). Finally, asymmetric (also called nested) versions of Archimedean copulas have been introduced and used in hydrology to model joint distributions of more than two random variables (e.g. Grimaldi and Serinaldi 2006b, Ma *et al.* 2013).
- *Extreme-value copulas*. This class is particularly attractive for block maxima data (i.e. annual maximum series of daily flows) – see e.g. Salvadori and De Michele (2011). Extreme-value copulas are defined on the basis of a dependence function, known as the Pickands dependence function, which plays a similar role for extreme-value copulas to generators for Archimedean copulas. Among the well-known copulas in this class, one can find Galambos, Hüsler-Resiss, Tawn and Gumbel (which is also Archimedean). Extreme-value copulas have been used in numerous multivariate hydrological studies, such as investigating floods and droughts (e.g. Zhang and Singh 2006, Salvadori and De Michele 2011, Sharma and Mujumdar 2019). See Genest and Chebana (2017) for the mathematical expressions in d-dimensions and Genest and Nešlehová (2012) for a review and more technical details.
- *Meta-elliptical copulas*. Meta-elliptical copulas are derived from elliptical distributions (e.g. Kotz and Nadarajah 2000). Convenience and flexibility are the key characteristics of meta-elliptical copulas (Genest *et al.* 2007). The normal and multivariate Student t copulas are among the most used copulas of this class in the hydrology literature including droughts, floods, and extreme rainfall analysis (e.g. Zhang and Singh 2019, table 7.1, Salvadori *et al.* 2007, Ma *et al.* 2013, Tosunoglu and Singh 2018). For more details and mathematical formulation, the interested reader is referred, for instance, to Genest and Chebana (2017) and Chebana (2022).
- *Other classes of copulas*. Even though the above classes of copulas are the most developed and used, a number of other classes are also available, such as the Farlie-Gumbel-Morgenstern (FGM), Plackett, entropy, Vine copula, flipped copula and others. Even though the FGM copula imposes moderate dependence, it is attractive because of its simplicity, and hence has been used in a number of hydrological studies (e.g. Papaioannou *et al.* 2016). The Plackett copula has also been used in

hydrology in a number of studies, such as Kao and Govindaraju (2008) and Papaioannou *et al.* (2016). Despite their significant potential, only limited hydrological applications have explored entropy copulas (e.g. Piantadosi *et al.* 2012). Given their flexibility, especially for high-dimensional analysis, vine copulas have recently received a great deal of attention in hydrology (e.g. Tosunoglu and Singh 2018). The nonparametric copula framework is also receiving increasing attention in the field of hydrology because of its flexibility and its capacity to adapt to any mutual dependence structure (e.g. Latif *et al.* 2024). Finally, it is worth mentioning the framework for the construction of multivariate non-Gaussian distributions presented by Bárdossy (2023); the framework can represent monotonic dependence but also dependencies with changing character, for example negative dependence for small values and positive dependence for high values.

3.2 Dependence measures, correlation and tail dependence

Dependence measure coefficients that summarize the degree of association between two or more variables are widely used in hydrology. The Pearson's ρ_P, Kendall's τ_K and Spearman's ρ_S correlation coefficients, and the upper–lower tail dependence coefficients, are the most commonly used measures of dependence (e.g. Nelsen 2006, Joe 2014). The Pearson correlation coefficient has several limitations, including linearity, and is not margin-free (Barber *et al.* 2020).

The population versions of τ_K and ρ_S can be obtained in terms of copula as explicit relations or through numerical approximations (for many Archimedean copulas see Zhang and Singh 2019). These relations are the basis of the method of moments for copula parameter estimation. Considering these coefficients is useful in preliminary copula selection as the values of τ_K and ρ_S for some copula families are restricted.

Tail dependence coefficients (upper λ_u^c and lower λ_l^c) can be used to detect and quantify the presence of extremal dependence. Tail dependence plays an important role in analysing dependent hydrological extremes (e.g. Poulin *et al.* 2007, Lee *et al.* 2013, Genest and Chebana 2017). Expressions of λ_u^c and λ_l^c can be obtained with respect to the parameters or generators of some common copulas (e.g. Salvadori *et al.* 2007, Joe 2014). Important differences between copulas can be revealed by investigating their λ_u^c and λ_l^c coefficients. As an example, the normal copula exhibits no tail dependence (null coefficients), whereas the Studentt copula offers substantially strong tail dependence (strictly positive coefficients). Hence, the former can potentially lead to an underestimation of joint extremes when considering multiple related hazards (e.g. McNeil *et al.* 2015). Estimators for tail dependence coefficients, developed by Capéraà *et al.* (1997) and Frahm (2006), have already been used in the hydrology literature (e.g. Requena *et al.* 2013). Lekina *et al.* (2015) considered different tail dependence measures in hydrology and recommended considering more than one, primarily because the upper tail dependence measure could fail to discriminate between the degrees of relative strength of dependence in asymptotically independent variables.

3.3 Multivariate inference with copulas

Selecting the most appropriate multivariate distribution for a given dataset is crucial. According to Sklar (1959), the selection of the joint distribution is equivalent to the selection of a copula and the margins. In this section, we focus on the key steps to select the copula since the process of choosing the margins aligns with the approach used in the univariate framework (e.g. Laio *et al.* 2009).

After transforming the margins into uniform margins, different copula models should be considered to find the best one for characterizing the dependence structure of the variables in hand. When selecting the appropriate copula, several issues need to be considered including the type of dependence structure, parameter estimation method, goodness-of-fit tests, selection criteria, and tail dependence (e.g. Genest and Chebana 2017).

- *Preliminary step*. A preliminary exploratory step can guide copula selection by, for instance, excluding some copulas that cannot describe the empirical dependence. This evaluation can start computing dependence measures (τ_K, ρ_S, and λ^c) to select potentially applicable copula candidates since not all copulas support all correlation coefficient and tail dependence values (e.g. Poulin *et al.* 2007, Michiels and Schepper 2008). Graphical tools (e.g. rank plots, rankit and K-plot, empirical copulas) can also provide interesting information about the dependence structure and hence help to select potential copula candidates through some characteristics such as departures from bivariate normality, presence of heavy tails, symmetry or asymmetry, strength of dependence, and extreme value.
- *Parameter estimation*. Several methods of copula parameter estimation are available in the literature, including the inference function of margins (IFM), maximum pseudo-likelihood (MPL), method of moments (MM), and Bayesian parameter estimation (see e.g. Hofert *et al.* 2012, Sadegh *et al.* 2018; for additional details the reader is referred to Genest and Chebana 2017). In the IFM method, the joint likelihood function is maximized in two steps (Joe 2014): first, the marginal log-likelihood is maximized, leading to an estimate of the margin parameters; then the latter is plugged into the joint likelihood function to obtain a log-likelihood function for the copula parameter. The obtained copula parameter estimator is consistent and asymptotically normal. By maximizing the pseudo-likelihood (MPL), the obtained estimator of the copula parameter is not affected by marginal misspecifications. The MPL estimator is consistent and asymptotically normal under reasonable conditions; however, it is generally less efficient than the full ML estimator when the margins are properly specified, except at independence. Overall, the MPL method is shown to have good performance, is considered the best

option, and is widely applied to one-parameter and multi-parameter copulas (e.g. Kim *et al.* 2007, Kojadinovic and Yan 2010). As indicated above, for some copulas, their parameter can be expressed as a function of Kendall's $_K$ or Spearman's $_S$; hence, a direct MM estimate of the copula parameter is obtained by estimating τ_K or ρ_S, respectively. Given their simplicity, MM estimators can provide reliable starting values for numerical (pseudo) likelihood maximization. Compared to IFM- or MPL-based estimators, the MM-based estimators are generally less efficient and may need to be adjusted to remain within the possibly limited range of the parameters. Further, moment-based methods have limitations in characterizing the underlying uncertainties; for this reason, Bayesian methods and global optimization approaches have gained attention for inferring copula parameters (Kwon and Lall 2016, Sadegh *et al.* 2018). For example, Sadegh *et al.* (2017) introduced a hybrid Markov chain Monte Carlo (MCMC) simulation within a Bayesian framework for estimating a wide range of copula families with one to three parameters. The MCMC simulations estimate the posterior distribution of each copula parameter value, which can be translated into uncertainty ranges for return periods and probability isolines.

- *Copula goodness-of-fit testing*. It is important and necessary to proceed with a formal goodness-of-fit testing of the preliminarily selected copula to ensure the model is representative. A number of goodness-of-fit tests for copulas have been proposed in the statistical literature (e.g. Zhang and Singh 2019). After a comprehensive assessment based on large simulations by Genest *et al.* (2009) and Berg (2009), it is deemed that goodness-of-fit testing based on empirical copula performs particularly well. Hence, a widely used goodness-of-fit test is based on the deviation between the empirical and the theoretical copula where the copula parameter is estimated on n pseudo-observations (such as the MPL estimator). One of the most powerful, easy to apply, general and widely used goodness-of-fit tests is based on the Cramér-von Mises statistics (Genest and Nešlehová 2012b).
    - Tests based on the Rosenblatt transformation or on nonparametric estimates of the copula density are among other goodness-of-fit tests available in the literature. These tests are general and conceptually valid for any copula. However, specific tests have been developed for specific dependence structures such as Clayton copula, Gaussian copula, and even extreme-value and Archimedean copula classes (see e.g. Genest and Nešlehová 2012b). Goodness-of-fit tests for a specific but large class of copulas are also emerging, for instance the class of bivariate exchangeable copulas, spatial copulas and multi-parameter copulas. Given the importance of the Archimedean and extreme-value copula classes in hydrological and other applications, specific goodness-of-fit tests are available (see Genest and Nešlehová 2012b or Genest and Chebana 2017 and the references therein). The p-value approximation of all the above tests (excluding Bayesian-based methods) can be obtained using a parametric bootstrap framework (Genest and Rémillard 2008). However, computational costs can be high, especially when the sample size increases. Based on the multiplier central limit theorems, Kojadinovic *et al.* (2011) proposed a faster and more efficient procedure for large samples; yet, for such samples, it is hard to find a copula that passes the above goodness-of-fit tests.
- *Selection criteria for copulas*. For a given dataset, the goodness-of-fit test can lead to more than one accepted model. Then, several selection criteria can be used to help identify the most appropriate copula from the set of acceptable ones (e.g. Zhang and Singh 2006, Requena *et al.* 2013, Genest and Chebana 2017). Among the popular criteria are the Akaike information criterion (AIC) and Bayesian information criterion (BIC). The general AIC formulation is based on the ML of a given model. However, for copula parameters, the MPL is preferred and the corresponding AIC becomes more precise than the one based on ML (similarly for BIC). MPL formulation is employed in several studies (see Joe 2014 and references therein). In addition to AIC and BIC, one can use the cross-validation copula information criterion. However, based on comparison in a bivariate simulation study, AIC and cross-validation are found to be overall very similar. For theoretical developments related to model selection for copulas, the interested reader is referred to Grønneberg and Hjort (2014).
- *Uncertainty assessment*. While uncertainty assessment is routinely implemented in univariate analysis, it is seldom addressed in the multivariate context. Uncertainty may affect copula parameter estimation, copula selection, and return period estimation. The relevance of dealing with uncertainty for reliable multivariate quantile estimation is analysed in several studies (e.g. Serinaldi 2013, Dung *et al.* 2015), where the usual lack of long data records for multivariate analysis is underlined. Lately, Bayesian approaches have been proposed to account for uncertainty in model parameters and return period estimations in the multivariate framework (e.g. Kwon and Lall 2016, Sadegh *et al.* 2018, Liu *et al.* 2020).

Several packages are available for the free software R (R Core Team 2020) for addressing copula-based frequency analysis; some of the most relevant ones are listed in Table 2.

3.4 Challenges and open problems

As described above, a significant contribution to multivariate hydrological frequency analysis using copulas has been produced in the last two decades. In this subsection, we highlight some additional aspects of multivariate analysis that are emerging or still require efforts from researchers.

A warning emerging from the literature is that analysts too often limit their attention to a few copula functions, typically

Table 2. Packages for multivariate, copula-based frequency analysis and modelling.

Name	Brief description	Repo.	Link	Authors
Copula	Multivariate dependence with copulas	CRAN	https://CRAN.R-project.org/package=copula	Hofert *et al.*
VineCopula	Statistical inference of vine copulas	CRAN	https://CRAN.R-project.org/package=VineCopula	Nagler *et al.*
MvCAT	Multivariate copula analysis toolbox	UCI.edu	https://amir.eng.uci.edu/MvCAT.php	AghaKouchak *et al.*
MhAST	Multihazard scenario analysis toolbox	UCI.edu	https://amir.eng.uci.edu/MhAST.php	AghaKouchak *et al.*
MSDI	Multivariate standardized drought index (MSDI)	CRAN	https://CRAN.R-project.org/package=drought	Zengchao
corTESTsrd	Significance testing of rank cros-correlations under SRD	CRAN	https://CRAN.R-project.org/package=corTESTsrd	Lun *et al.*

Archimedean, while, as shown here, the possible options are diverse, offering the opportunity to reach a better modelling outcome for the data sample. Furthermore, apart from the copula families mentioned above, it is important to mention multi-parameter copulas. Although multi-parameter copulas have not been fully explored, they are attracting more attention in recent years (e.g. Salvadori and De Michele 2010, Joe 2014, Requena *et al.* 2016, Sadegh *et al.* 2017, Ben Nasr and Chebana 2019). In principle, most estimation methods are valid for multi-parameter copulas, but most applications in the hydrology literature are limited to one-parameter copulas. Brahimi and Necir (2012) extended the MM-based method to multi-parameter copulas, whereas Brahimi *et al.* (2015) proposed a multivariate L-moment method. The latter has some interesting features for the case of small sample sizes such as in frequency analysis. Note, however, that using multi-parameter copulas when one-parameter copulas would suffice generally implies an increased estimation uncertainty. Bayesian approaches are also used to estimate the parameters of multi-parameter copulas and their underlying uncertainty.

By analogy with the univariate setting, the stationarity assumption could fail in the multivariate framework due to changes in urbanization, land use/cover or climate. Hence, to be more realistic and for an accurate risk estimation, multivariate non-stationarity modelling should be considered. This is a recent and emerging topic in statistical hydrology with a growing number of studies, including Chebana and Ouarda (2021) for floods, Jiang *et al.* (2015) for low-flow series, Kwon and Lall (2016) for droughts, and Feng *et al.* (2020) for flood coincidence risk. Conversely, the relaxation of the independence hypothesis that allows researchers to incorporate additional data for inference in the univariate setting still requires additional theoretical developments in the multivariate one. Indeed, as mentioned at the beginning of Section 3 (testing assumptions), the serial independence condition is a requirement for multivariate analysis.

4 Time series analysis and simulation

The rationale for time series modelling and simulation in hydrology is to resemble and reproduce the characteristics of a variable of interest for simulation and forecasting applications. This is typically performed by describing the historical evolution of the variable in time. Thus, records can be usefully extended using synthetic data generated by stochastic models, which is known as synthetic hydrology (Benson and Matalas 1967, Matalas 1967). Artificial datasets of adequate length are created from the characteristics of existing observations (which are insufficiently long for a reliable design or assessment of water systems), potentially providing a huge number of random sequences with the observed statistical characteristics; readers are referred to Table 3 for a list of R packages (R Core Team 2020) available for hydrological simulation. The stationarity hypothesis is the major assumption, which presumes a future with a non-dynamic behaviour in terms of statistical moments and correlation with the past.

Time series modelling of hydrological variables has a history of about a century and continues to evolve with intense research. Markov chains (MCs) were proposed at the beginning of the 20th century for streamflow simulation; then, with the introduction of the autoregressive (AR) models by Thomas and Fiering (1962) and Yevjevich (1963), the formal development of stochastic modelling has started. Literature reviews of these early studies clarify the historical development of time series modelling in hydrological studies; readers are referred to Mejia *et al.* (1972), Rodriguez-Iturbe *et al.* (1972), Lawrance and Kottegoda (1977), and references therein.

The historical focus on stochastic streamflow models (i.e. Thomas and Fiering 1962, Matalas 1967) was to enable hydrologists to evaluate the reliability, vulnerability and resilience of future water resource systems. Such computational tools and principles also enabled a more complete integration of uncertainty into water management decision making and have been

Table 3. Packages for time series analysis and modelling.

Name	Brief description	Repo.	Link	Authors
LPM	Linear parametric models applied to hydrological series	CRAN	https://CRAN.R-project.org/package=LPM	Tallerini and Grimaldi
CoSMoS	Complete stochastic modelling solution	CRAN	CoSMoS R | Complete Stochastic Modelling Solution (rproject.org)	Papalexiou *et al.*
MMM and MRD	Multisite Markov model (MMM) for rainfall generation, and its extension using exogenous predictor variables for downscaling (MRD)	UNSW	https://www.unsw.edu.au/research/hydrology-group/our-resources/multisite-rainfall-downscaling-mrd	Mehrotra *et al.*
SAMS2007	Stochastic analysis, Modelling and simulation	USBR	http://www.sams.colostate.edu/	Sveinsson *et al.*

in common use by several agencies worldwide for over 50 years (see Vogel 2017 for a brief review). Although the variable of interest is the discharge in a river cross-section, the continuous synthetic simulation methodology based on a stochastic model of precipitation and other meteorological variables, called a weather generator, is increasingly used. Many studies can be found in the literature that follow this approach, with different nuances. Among the more modern approaches, it is relevant to mention the stochastic weather generators by Steinschneider *et al.* (2019) for multivariate/multisite weather variables (i.e. rainfall and temperature) which can capture low-frequency climatic variability for use in water resource vulnerability assessments.

In a pioneering work, Cameron *et al.* (1999) combined a modification of the also pioneering weather generator developed by Eagleson (1972) with a semi-distributed hydrological model. This formed the basis for the development of stochastic watershed modelling. Stochastic watershed models (SWMs) are deterministic watershed models implemented using stochastic meteorological series, model parameters and model errors, to generate ensembles of streamflow traces that represent the variability in possible future streamflow. By combining deterministic watershed models, which are ideally suited to account for anthropogenic influences, with recent developments in uncertainty analysis and principles of stochastic simulation, SWMs are promising tools to accommodate climate, land use or other forms of change (see Vogel 2017 for a comprehensive discussion), and will certainly be the focus of future works. However, the remainder of this section is dedicated to purely stochastic modelling approaches, that were of main interest among the STAHY community in recent years.

Through this journey, there are several divisions in time series modelling. The basic classification is deterministic (like trends or jumps, periodicity, and seasonality) and stochastic (based on stationarity and ergodicity) components. Further, the modelling is extended to a single site and variable (univariate) or multiple sites and variables (multisite and multivariate). In addition, depending on the state of the variable, the modelling is divided into discrete, continuous and mixed types. Finally, the most basic classification in terms of stochastic modelling is into parametric and non-parametric approaches. The parametric approaches range from models such as AR to models based on machine learning algorithms. The nonparametric methods typically consist of kernel density and bootstrap techniques. Subsequent sections provide details only about the statistical parametric models, being aware of the recent impressive development and of the potentiality of machine learning and data science tools, which may be the subject of ongoing work in the ICSH.

The introduction of the Hurst phenomenon further advanced stochastic modelling (Hurst 1951), yet it was (and probably still is) controversial during the early stages (Klemeš 1974) among hydrologists. Though Hurst introduced these phenomena in the early 1950s, Kolmogorov introduced the same concept mathematically in the early 1940s. Long-term persistence and fractional Gaussian noise models were mainly popularized by Mandelbrot and Wallis (1968) and, later, by Beran (2017) in hydrology. In view of extensive arguments related to the interpretation of Hurst phenomena and uncertainties in the estimates of the Hurst exponent, simple models such as autoregressive moving average (ARMA) models capable of reproducing simple statistics were used by early hydrologists (Salas *et al.* 1980). Subsequently, several models have been proposed, such as the fractional Gaussian noise models, mixture models such as ARMA-Markov models, fractionally autoregressive moving average (FARMA) or autoregressive fractionally integrated moving average (ARFIMA) models, disaggregation models, models for intermittency, and general mixture models (Montanari *et al.* 1997).

4.1 Single-site modelling

In the following we provide a state-of-the-art review on parametric approaches, focusing on discharge and rainfall simulation.

4.1.1 Streamflow simulations

The first approach for modelling hourly and daily streamflow time series was the AR models. Thomas and Fiering (1962) introduced a streamflow generation model which is an AR model for generating monthly streamflow for the Clearwater River (USA). After that, several models have been developed for hydrological time series modelling. These models consider that a variable value at a specific time instant is related to the corresponding value at the previous instant(s). Streamflow at larger scales (week, month, season) can be described by stationary stochastic models after being seasonally standardized. The low-order AR and ARMA have been the most popular models for annual streamflow simulation (Box and Jenkins 1976). Hirsch (1979) compared six single-site data generation mechanisms and concluded that the ARMA model was superior to the AR model. Box and Jenkins (1976) proposed multiplicative models, in which the trend and seasonal components are multiplied and then added to the error component, to capture seasonal and annual statistics. However, these did not include parameters for the periodicity. When a periodic structure is present, the periodic autoregressive (PAR) or the periodic autoregressive moving average (PARMA) model (Salas and Obeysekara 1992) is more suitable. The latter has a more flexible correlation structure and preserves seasonal and annual statistics, at the expense of the number of parameters. The need to transform the time series into a normal one is a restraint of the PARMA and its multiplicative version. To overcome this limitation, a model with periodic correlation structure and periodic gamma marginal distribution, i.e. the PGAR(1) model, has been proposed (Fernandez and Salas 1986). Other versions, modifications, and developments of these models are the autoregressive integrated moving average (ARIMA) model and the ARFIMA or FARMA models (Oliveira and Maia 2018).

Other approaches were developed to account for short-term physical characteristics of flow (ascension and recession curve) together with the long-term statistical characteristics (mean, variance, lag-one, and higher lag correlation coefficients); see Claps *et al.* (2005). Among them, the shot noise models are based on overlapping pulses which represent the autocorrelation structure and are mostly used for variables with a strong and repetitive autocorrelation; the shot noise model has been used because of its ability to reproduce the physical behaviour and hydrological aspect of the streamflow process.

Streamflow processes depict an intermittent behaviour (similar to rainfall records) when the contribution of the hydrological basin to the river is significantly reduced due to lowered groundwater levels and no substantial snowmelt during the rainless season. Intermittency is effectively felt in arid and semi-arid regions (Salas and Fernandez 1993). Yevjevich (1984) summarizes the approaches used for modelling intermittent time series as (i) spell process, (ii) truncated process, and (iii) 1-0 approach. Another kind of model used for the daily streamflow is based on transition probabilities to understand the state of the stream (whether it has flow or not on a particular day, and whether it increases or decreases); a transfer function model using a wet/dry MC was proposed by Treiber and Plate (1977), where a pulse is assigned on a wet day by using a modification of the exponential distribution.

Annual, monthly, and daily multisite streamflow simulation is also achieved by parametric disaggregation models (Kumar *et al*. 2000) and nonparametric disaggregation models (Lee *et al*. 2010, Nowak *et al*. 2010). Other examples include a semi-parametric model for daily streamflow by Srinivas and Srinivasan (2005), a parametric multisite stochastic simulation framework for the generation of seasonal timeseries reproducing sub-annual statistics, short-term memory and long-term persistence (i.e. over year scaling of annual averages) by Langousis and Koutsoyiannis (2006), wavelet methods (Keylock 2012), entropy methods (Srivastav and Simonovic 2014), empirical decomposition methods (Lee and Ouarda 2012), copula methods (Chen *et al*. 2019), moving block bootstrap (Srinivas and Srinivasan 2005) and k-nearest neighbour bootstrap (Lall and Sharma 1996).

Simulation models grow in complexity as temporal resolution increases; i.e. a daily streamflow simulation model is expected to be more demanding than an annual model. This is not only because of the increasing amount of data involved when temporal resolution increases, but also because of the emergence of finer detail upon closer examination of the physical process. By delving into the time interval, a critical understanding of the temporal asymmetry (or time irreversibility; see Carsteanu and Langousis 2020, Koutsoyiannis 2021) in the rising and falling limbs of the daily streamflow hydrograph emerges. This is due to the physics behind the streamflow process which has different physical drivers for the ascension and recession curves of the hydrograph at a daily time step. It is therefore necessary to conserve the temporal asymmetry in the streamflow generation models (Serinaldi and Kilsby 2016). This is particularly important as a simulation not only replicates the statistical measures but also regenerates the physical structure of the daily streamflow process.

4.1.2 Rainfall simulation

Rainfall can be modelled by considering the process to be continuous (point process and cluster models), by using approaches of cumulative precipitation over non-overlapping time intervals aggregating rainfall at the desired time scale (i.e. hourly, daily, monthly), or by using mixed-type distributions to transform Gaussian time series to preserve marginal distribution and correlations. Typical examples are the MC and alternating renewal models. Multifractal simulation techniques are also a popular tool for temporal rainfall disaggregation of rainfall data (see e.g. Veneziano and Langousis 2010 for a review).

One of the first theories in modelling precipitation as a continuous process is the point process theory (e.g. Cox and Isham 1980). Based on this, the number of storms, the arrival rate and the rainfall amount for each storm are considered and simulated as independent random variables. This Poisson approach was modified to consider rainfall to have a random duration and intensity, independent from one another and usually exponentially distributed (Poisson rectangular pulse models); the storms may overlap so that the cumulative process is autocorrelated. The Poisson-based model, although flexible for a particular level of aggregation, presents limitations when studying a range of time scales (Rodriguez-Iturbe *et al*. 1987). The clustered point process-based models, where a cluster of activities starts at every point of a point process, offer more realistic representations of rainfall and can represent multiple timescales at once. Two such models are the Neyman-Scott and Bartlett-Lewis; the Poisson cluster process-based models are common structures for the generation of sub-daily rainfall time series, as rectangular profiles are flexible in approximating discrete rainfall which is aggregated over time intervals of 1 h or more (see e.g. Cowpertwait *et al*. 2007). Subsequently, several studies (e.g. Velghe *et al*. 1994, Kilsby *et al*. 2007, Kim and Onof 2020) investigated the performance of both models and provided modified versions better able to catch (for example) the autocorrelation, skewness, extreme values, etc., at different time scales. Among the available models, we recall the RainSim rainfall simulator based on the generalized Neyman-Scott rectangular pulses developed by Cowpertwait in 1995, that was used for single-and multisite applications, and later improved by Burton *et al*. (2008). Both rectangular pulse models, either in their original form or with modifications, were found to satisfactorily represent rainfall processes for a range of time scales for the UK, Scotland, Belgium, Switzerland, Spain, Australia, New Zealand, Ireland, South Africa, Greece, Italy, and the USA (Kossieris *et al*. 2018).

The MC model (or Markov process) describes a sequence of events whereby the present condition depends only on the antecedent state. The MC model is characterized by the transitions, which are the changes of state, and the transition probabilities. In the first-order MC, the probability of daily rainfall is conditioned on the status (wet or dry) of the previous day. The most substantial advantage of the model is the fact that it identifies the seasonality in daily rainfall occurrence, which can be described adequately by the process (Stern and Coe 1984), although the model has failed to fit the observed data in some cases. Some studies have noted that the method underestimates the dry spells (Wilks and Wilby 1999, Sharma and Mehrotra 2010). Usually, MC models of higher orders can overcome these issues (Wilks 1999, Hayhoe 2000). Higher-order MC models are based on wet-day probabilities of a few consecutive days, thus improving the model's "memory." The number of transition probabilities/parameters required increases exponentially as the order increases. Although these models improve the representation of inter-annual variance they still do not always manage to represent climatic variability

(Sharma and Mehrotra 2010). To improve the simulation of dry spells, the use of "hybrid-order" Markov models is proposed, where the "memory" extends further back in time but only for the dry spells (Wilks and Wilby 1999); to better represent rainfall clustering in time, the multi-state MC model is used to simulate both rainfall occurrence and the number of different precipitation bands (Haan *et al.* 1976). These models require a large number of parameters and long data records. Finally, non-stationary Markovian dependence is provided by the modified Markov model, which allows for the preservation of rainfall statistics up to the decennial scale, by conditioning the MC parameters on the number of past wet days (see e.g. Oriani *et al.* 2018 and references therein). Over the past few decades, several stochastic rainfall generators have been proposed based on MC models for the rainfall occurrence combined with a parametric probability distribution for the rainfall amount on a wet day; see, among others, Weather GENerator (WGEN) by Richardson (1981), and versatile stochastic daily weather generator (WeaGETS) by Chen *et al.* (2012).

Another category of rainfall time-series simulation models uses a simple renewal process to describe the alternation of wet and dry conditions. The term "renewal" implies independence between wet and dry period lengths. For the representation of wet and dry spells, logarithmic series, the truncated negative binomial distribution, the truncated geometric distribution, and other semi-empirical distributions have been proposed (Wilks and Wilby 1999). Another probability distribution describes the rainfall amount. The approach allows for the direct estimation of composite precipitation events but cannot depict the seasonality of the rainfall occurrence (Srikanthan and McMahon 2001).

As early as the mid-1960s, researchers identified the fractal behaviour of time series, according to which an object can be subdivided into reduced-size copies of the whole (Mandelbrot 1982). Based on the multifractal theory, fluctuations at a given scale can provide information on those at other scales via scale invariance. Statistical moments are associated with a scale parameter through a log–log linear relationship; thus, multifractal models are preferred for their ability to correctly reproduce the strongest events. The underlying idea of this framework is that these fields are the result of an underlying random multiplicative cascade process (Schertzer and Lovejoy 1987). Multiplicative cascade models can be used for temporal rainfall disaggregation of daily data to generate rainfall time series of high temporal resolution. First introduced by Yaglom (1966), they appeared to be promising and therefore have received significant attention ever since (e.g. Menabde *et al.* 1997, Deidda 2000, Veneziano and Langousis 2005b, 2010, Gaume *et al.* 2007, Langousis and Veneziano 2007). Multifractal approaches can be pulse-based, non-pulse-based using wavelet decompositions, and non-pulse-based using discrete or continuous multiplicative cascades (Flores 2004). Other works related to time series fractality in hydrology (Adarsh *et al.* 2020) include double trace moments (Tessier *et al.* 1996), wavelet transform modulus maxima (WTMM) (Kantelhardt *et al.* 2003), extended self-similarity (ESS) (Dahlstedt and Jensen 2005)and arbitrary order Hilbert spectral analysis (AOHSA) (Huang *et al.* 2009), to mention a few.

Recent approaches in univariate, multivariate, random field simulations of rainfall have focused on generating time series or random fields that explicitly preserve any desired marginal distribution at different locations and seasons, as well as any desired correlation structure (Papalexiou 2018, Papalexiou and Serinaldi 2020, Papalexiou *et al.* 2021). The scheme comprises five steps, as graphically demonstrated in Fig. 2; it is clearly generic and allows simulation of time series having any marginal distribution and autocorrelation and can be used for many different hydroclimatic processes such as relative humidity, wind speed, river streamflow, or any process having discrete and binary marginal distributions such as the number of extremes per year or wet–dry sequences (for more details see Papalexiou 2018).

Recently, the development of continuous hydrological modelling has made rainfall simulation models more important, as they constitute the input in this promising approach (Grimaldi *et al.* 2022). However, a unique approach that is easy to apply is still not available to the community. A challenge for the near future is to identify the best rainfall simulation approaches for specific hydrological applications providing user-friendly tools applicable without a restricting statistical background. This is a subject of current interest of the ICSH that will be addressed in future publications.

In addition to the parametric approach, it is worth mentioning that the nonparametric methods offer attractive alternatives since they can capture non-linear and non-normal data characteristics. Nonparametric regression refers to a group of approaches that are used to fit a curve when little a priori knowledge exists about its shape (Altman 1992). Running averages have been used since the late 19th century for determining time series trends by De Forest. Later, local location estimators like the kernel (e.g. Parzen 1962) and the nearest-neighbour regression estimators (e.g. Cover and Hart 1967) were introduced. The conditional kernel density estimation derives a probability density function from histogram data. The technique allows the creation of a smooth curve (empirical probability density estimate) given a set of data. In the kernel density estimation for discrete random variables such as rainfall (Lall 1995), wet and dry spell lengths are considered as an integer number of days and the sample relative frequencies are estimated. These relative frequencies are then smoothed with a kernel estimator. The kernel method is superior to the ML estimator which yields the relative frequency directly, as it allows the extrapolation of probabilities to spell lengths, and has higher mean square error efficiency (Rajagopalan and Lall 1995). Some kernel estimators are the geometric estimator, the maximum penalized likelihood estimator, the estimator by Hall and Titterington, and the discrete kernel estimator (see Rajagopalan *et al.* 1997 and references therein).

When a model is applied at a time interval shorter than a year or month; e.g. week or day, parameterization becomes computationally costly, because the seasonality emerges more effectively. This brings the problem that one single value of a parameter cannot be applicable over the year; i.e. monthly, weekly, or even daily parameterization is needed. This is a big burden for the modelling approach that can be overcome by fitting a seasonal function to the parameters such as the Fourier series (Aksoy and Bayazit 2000).

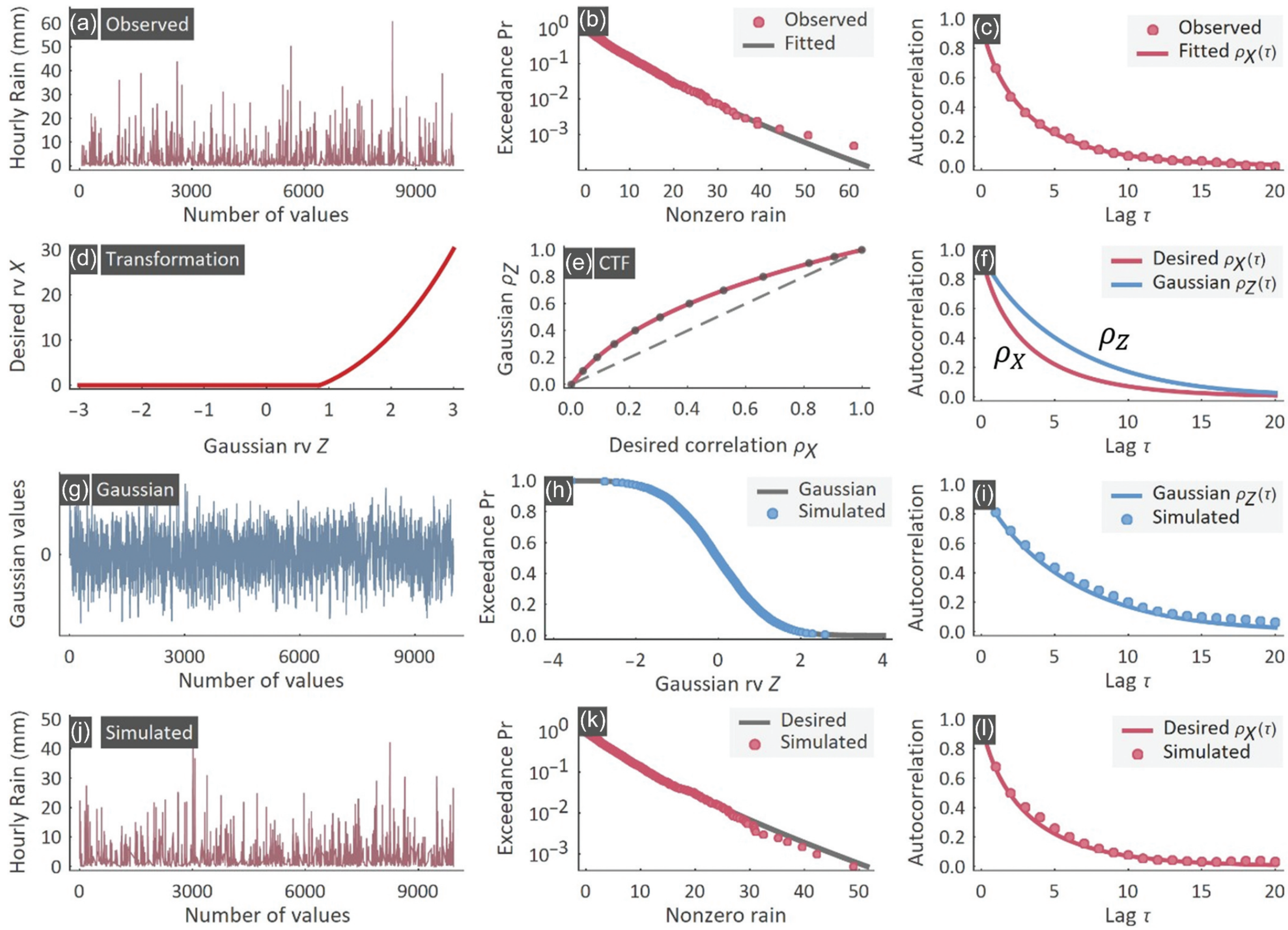

Figure 2. Step-by-step graphical demonstration of stochastic simulation that preserves explicitly a desired marginal distribution and correlation structure. (a, b, c) Observed precipitation time series and the fitted parametric distribution and autocorrelation structure; (d) transformation of Gaussian values to the desired mixed-type fitted distribution and (e, f) the corresponding correlation transformation function (CTF) and estimated autocorrelation structure of the Gaussian process; (g, h, i) generated Gaussian time series, its marginal and correlation structure; (j) final simulated time series that indeed have (k, l) the desired marginal distribution and autocorrelation structure.

4.2 Multisite modelling

4.2.1 Multisite streamflow generation

It is quite common for river flows in the same basin or nearby basins to exhibit significant cross-correlations among their tributaries since they receive runoff from the same parent rainfall. Therefore, multisite models which preserve cross-correlation between sites are logical (Matalas 1967). Some of the models mentioned above can be extended to multiple sites, i.e. the multivariate AR(1) and ARMA(1,1) are usually adequate for the annual time scale. Multivariate ARMA models often result in complex parameter estimation and model simplifications have been suggested such as the contemporaneous ARMA (CARMA) (Salas *et al.* 1980). ARIMA models have also been applied for the multisite case. The Markov cross-correlation pulse model has been used to extend single-site streamflow generation to multiple sites, preserving high daily cross-correlations (Xu *et al.* 2003). A hybrid seasonal MC model was also used at multiple sites by Szilagyi *et al.* (2006). The model used shot noise models in an MC-based framework, along with a conceptual model for flow recession, managing to generate long time series of daily flow which preserves long-term statistics. Medda and Bahr (2019) provide a list of models used for the multisite case, such as the hybrid stochastic model of Srinivas and Srinivasan (2005) with a parsimonious periodic parametric model and moving block bootstrap for resampling of the residuals, and the models found in Wang and Ding (2007), Hao and Singh (2013), and Srivastav and Simonovic (2014).

4.2.2 Multisite rainfall generation

Numerous studies have dealt with the spatial dependence of precipitation in space-time stochastic models under parametric (Papalexiou and Serinaldi 2020, Papalexiou *et al.* 2023), semiparametric, and non-parametric frameworks (Sharma and Mehrotra 2010). Multisite rainfall generation models can be broadly categorized into conditional models (hidden Markov models), extensions of MC models, and random cascade models (Srikanthan and McMahon 2001).

In the hidden Markov models, the multisite rainfall is simulated conditional on the weather states and/or atmospheric circulation patterns so that the effects of large-scale atmospheric circulation are incorporated into local weather (Zucchini and Guttorp 1991, Hughes and Guttorp 1994). In the hidden Markov model of Zucchini and Guttorp (1991) the climate process follows anMC. Other versions of the model are those of Bárdossy and Plate (1992), Wilson *et al.* (1992), and Charles *et al.* (1999). Wilks (1998) extended the first-order MC with the two states and a mixed exponential distribution for the simultaneous generation of rainfall at multiple locations by having individual models fitted to each site and using spatially correlated random numbers to introduce spatial dependence. Jothityangkoon *et al.* (2000) introduced a space-time model with a temporal first-order, four-state MC and a spatial model based on a non-homogeneous random cascade process, for the generation of daily rainfall. The model manages to reflect the spatial patterns of the long-term mean of all timescales, the spatial distribution of the wet fraction, the statistical characteristics of storm arrivals and interarrival times, and the exceedance probabilities of rainfall but underestimated the mean number of wet days and the mean wet spell lengths during winter months. To avoid the use of discrete weather states in continuous rainfall generation, Langousis and Kaleris (2014) and Langousis *et al.* (2016b) proposed and applied a scheme for stochastic simulation of daily rainfall conditional on upper-air predictor variables. While the scheme didnot involve any type of seasonal modelling, it was found to accurately reproduce several rainfall statistics at seasonal and annual time scales (wet day fractions, the alternation of wet and dry intervals, the distributions of dry and wet spell lengths, the distribution of rainfall intensities in wet days, short-range dependencies, the distribution of yearly rainfall maxima, among other), as well as the dependence of rainfall statics on the observation scale.

Other approaches used for multisite rainfall generation include generalized linear models (GLMs) (Chandler and Wheater 2002), reshuffling approach-based models (Clark *et al.* 2004), models that preserve exactly the marginal distributions and correlations (Papalexiou 2018, Papalexiou and Serinaldi 2020), nonparametric models like the k-nearest neighbour approach (Buishand and Brandsma 2001), and semi-parametric models that parameterize the rainfall occurrence generation process while using nonparametric methods to characterize rainfall amounts (Mehrotra and Sharma 2007). Fu *et al.* (2018) provide a comparison among four multisite weather generators: (1) the generalized linear model for daily climate time series, based on generalized GLMs, (2) the stochastic climate library which uses a first-order two-state MC, (3) the multisite precipitation generator, based on the approach of Wilks (1998), and (4) the multisite auto-regressive weather generator based on the theory of vector auto-regressive models. They note that all models reproduce the daily, monthly, and annual rainfall, extremes, and dry/wet spell lengths reasonably but also simulate a large range of variability.

4.3 Change detection in hydrological time series

As premised in Section 2.3, stationarity may be questionable, at least for some hydrological time series (Milly *et al.* 2008, 2015). However, non-stationarity is hard to detect, and often false assumptions on the underlying process can lead to falsely rejected hypotheses for non-stationarity tests (see e.g. Lins and Cohn 2011 and Serinaldi and Kilsby 2015). In the following we review some recent techniques for non-stationary analysis and simulation, emphasizing explicitly the high uncertainty when dealing with non-stationarity in short time series.

For change analysis, the detection of irregularities, jumps, changes, and trends is of the highest importance. For many stochastic models, non-stationarity – when detected – has to be removed before application, while recently developed stochastic models are able to include it in the model structure; either way, it is essential to know the time of occurrence of non-stationarities. To reliably differentiate non-stationarities from simple random effects in time series, statistical tests are applied that do or do not reject the hypothesis of stationarity (the null hypothesis) according to a given significance level, typically 5%. This corresponds to Type I error, which is the probability of detecting a change when it does not exist (societal over-preparedness). Type II error, related to statistical power, is the probability of missing the change when it exists; even though it informs us about the probability of under-preparedness, which is a fundamental issue for society, it is rarely considered in the analysis (Vogel *et al.* 2013). See also Prosdocimi *et al.* (2014) for a discussion on the importance of correctly specifying the null and alternative hypotheses in non-stationarity testing and how this relates to statistical power.

Common tests can be categorized according to some of their properties, such as whether they are parametric or non-parametric (i.e. if an assumption on the data distribution is made), and how conservative, powerful, efficient, or robust they are. The dependence structure for which the test was constructed must also be differentiated. Most tests can be applied for the case of i.i.d. data. However, if short- or long-range dependence occurs, the limit distributions and the test statistics have to be extended (see e.g. Serinaldi *et al.* 2018). Classical examples of short-range dependent (SRD) processes are mixing processes or Markovprocesses. Long-range dependence (LRD) is much harder to detect with respect to the short-range case, especially when the sample analysed spans a short observation period, as is often the case for hydro-meteorological records. However, many studies argue that flood and precipitation processes are characterized by long-range dependence in terms of long-term cyclic behaviour (Szolgayova *et al.* 2014, Koutsoyiannis 2021), yet other studies in long streamflow records do not verify the existence of long-range dependence (Markonis *et al.* 2018; see also the discussion in the introduction to Section 4).

Non-stationarities in time series are basically differentiated into two categories: change points and trends. The first category, change points, applies to the case where an abrupt change in the distribution of the random variables occurs at one point in time. The second category, trends, applies to continuous and monotone changes in the distribution. An in-depth introduction to the statistical methods for the detection of non-stationarity can be found in Helsel and Hirsch (1992).

Many distribution properties can change abruptly at any given point in time, e.g. due to the impact of some exogenous factors or a change in the measuring of the process: change point tests aim to detect these possible changes. Time points where changes take place are called *change points*; the interval included between two change points is a *segment*, and the procedure by which the segments of a time series are determined is called *segmentation*. Segmentation of a time series simply means dividing a given number of observations into subseries with statistical characteristics that are similar within each subseries and different between subseries (Salas *et al.* 1980, Helsel and Hirsch 1992). This is also called jump analysis and can be considered a change point detection problem for which statistical tests and Bayesian procedures are available (Pettitt 1979, Alexandersson 1986, Seidou *et al.* 2007). Many segmentation methods and a very extensive bibliography are presented in Basseville and Nikiforov (1993); additionally, recent segmentation procedures are available in the literature (Hubert 2000, Kehagias 2004, Kehagias *et al.* 2006, Gedikli *et al.* 2010). The simplest case is segmentation with regression-by-constant in which the aim is to determine the change points or boundaries where the average of the current segment is statistically different than the average of the next segment as well as that of the previous one. This shift or jump may be either positive or negative. Not only segmentation with regression-by-constant but also segmentation with regression-by-lines or higher-order polynomials can be used (Kehagias *et al.* 2006).

Common change-point tests focus on changes in mean, variance, and correlation. Corresponding change points can be identified at the time when the test statistic reaches its maximum. Change points in the mean refer to abrupt changes in the mean value of the underlying distribution. The simplest test statistic is based on the cumulative sums (CUSUM) of the series values before and after the change point. CUSUM tests for a change in mean exist for i.i.d., SRD, and LRD time series. Many CUSUM tests are non-parametric and efficient, but not robust. To overcome this problem, the robust Wilcoxon (Mann-Whitney) test was developed. Again, statistical theory for this test exists for i.i.d., SRD and LRD time series (see Dehling *et al.* 2013 and references therein). Another robust test for a change in the mean is the Pettitt (1979) test, usually applied for i.i.d. data. Changes in variance are less often investigated in hydrology. Common tests are the parametric CUSUM test (Inclan and Tiao 1994) which can be applied to i.i.d., SRD and LRD data and the robust, non-parametric test based on Gini's mean difference which can be applied to i.i.d. and SRD data (Gerstenberger *et al.* 2020). To test for a change in the correlation structure, tests based on correlation coefficients like Spearman's rho or the Pearson coefficient are applied. These tests apply to independent and weakly dependent time series (Wied and Galeano 2013, Dehling *et al.* 2017). Note that detecting a change in e.g. the mean does not imply that it is the only statistic that changes over time; as pointed out in Section 2.3, a thorough preliminary analysis is necessary to understand which structure is most appropriate to model change.

In contrast to change points, trends assume a monotonic, continuous change in the central tendency (often taken to be the mean) of the time series. The most common trend tests are the Mann-Kendall test and the Cox-Stuart test: both these tests are non-parametric and therefore can be applied to data series without assuming that they follow specific distributions. The Mann-Kendall test can be applied to independent data as well as weakly dependent data (Cabilio *et al.* 2013) or the special case of seasonally impacted data (Zhang *et al.* 2016). Note that the Mann-Kendall test corresponds to computing the Kendall correlation coefficient for the record under study and the time index. The Cox-Stuart test (e.g. Rutkowska 2013) typically requires longer observation records than the Mann-Kendall test and has less power. Moreover, the presence of autocorrelation limits the application of this test. Pre-whitening often reduces the power of trend tests or falsely raises their Type I error rates (Bayazit and Önöz 2007, Wang *et al.* 2015). Either ignoring the effect of autocorrelation or dealing with it without choosing adequate methods will result in inaccurate detection results. A robust trend detection strategy should involve the investigation of the autocorrelation structure of the data and the selection of the corresponding method that keeps a balance between maintaining a low Type I error and a relatively strong power of trend detection (e.g. O'Brien *et al.* 2020). Trend tests can be informative on whether a change in the mean of the time series has occurred, but cannot quantify the actual change. They are therefore often coupled with the Theil-Sen slope which provides a robust estimate to quantify the change which occurred over time in the mean of the time series (e.g. Yilmaz and Tosunoglu 2019). Note that a change in the mean can have an impact on the results of a trend test and vice versa.

At large scales, the detected trends identified at several sites could be statistically non-significant at a regional scale, where several distinct issues can be investigated. The most thoroughly studied is field significance: when a test is repeated with a given significance level on several locations, the aim is to determine the minimum number of locally significant trends to conclude, with a regional significance level, that the changes are not all due to chance. The second aspect involves the consistency of changes detected within a given region. Exploiting the concept of regionalization may be a step forward, as presented in Section 5. Furthermore, the spatial correlation between time series should be considered. See for example Renard *et al.* (2008) or Yue *et al.* (2003) for approaches to trend testing in a large region.

In addition, identifying the cause or physical underpinning of detected trends is often the main objective. There is a wide range of methods for the detection and attribution of changes designed for individual and compound hazards (Easterling *et al.* 2016, Chiang *et al.* 2021a, Slater *et al.* 2021b, Chiang *et al.* 2022). The partial Mann-Kendall test (PMK), for instance, can be used to test whether a detected trend in a hydrological time series is significant after removing its correlation with a covariate variable (Libiseller and Grimvall 2002).

If a distribution can be assumed for the data, it is possible to construct change point and trend tests within a parametric framework. If, for example, one wishes to test whether an abrupt change occurs in a given year in the location parameter, it is possible to compare the goodness of fit of two models: one in which the location parameter is assumed to be constant and another in which the location takes different values before and after a specified changeyear. The comparison might use goodness-of-fit criteria such as AIC or be based on a likelihood ratio test between the two nested models (Kundzewicz and Robson 2004). Katz *et al.* (2002) offer an introduction to this type of parametric models with applications for hydrological extremes. Similar approaches can be employed to construct trend tests and investigate whether changes can be detected in one or, as expected, more parameters of the parametric models, for example scale and location. Parametric approaches can also be further extended to allow for not only linear or monotonic changes in the parameters. For example, Villarini *et al.* (2009a) and Slater *et al.* (2021a) applied the generalized additive models for location, scale and shape (GAMLSS) to investigate possible changes in extremes. See also Debele *et al.* (2017) for a discussion on the usability of GAMLSS in the analysis of hydrological extremes.

Detection of changes in hydrology is a difficult task. The often skewed or zero-inflated data with unknown dependence structure and distribution require a detailed a priori analysis (Kundzwicz and Robson 2004) and cross-correlation has to be considered if spatial datasets are evaluated (Douglas *et al.* 2000). The test statistics with correct assumptions on dependence and distribution of the data have to be selected from a pool of available tests. When detecting a change point, it is necessary to determine its origin. Without attribution, i.e. the clarification of the deterministic causes, a change-point detection is of little value for hydrological purposes. Moreover, evidence supports the significance of the findings. For example, changes in mean in discharge data can be related to anthropogenic impacts like the building of dams (Seibert and McDonnel 2010). Often, changes in discharge time series can also be related to changes in climatological time series (Zhang *et al.* 2014), e.g. for changes in flood type frequencies (Fischer *et al.* 2019).

Trends and change points can be generated by different sources: (i) meteorological drivers, such as changes in extreme precipitation patterns; (ii) climatic drivers that can modify soil moisture contents in catchments, such as changes in temperature, annual precipitation or evapotranspiration; (iii) drivers that modify rainfall–runoff processes at the catchment scale, such as changes in land uses; and (iv) stream drivers that modify flood propagation processes, such as river training (Vorogushin *et al.* 2012). Therefore, time series recorded at catchments that are either natural or near-natural have to be used to identify climate-driven flood trends, avoiding the anthropogenic effects on catchment response. As an example, Bertola *et al.* (2021) propose a new framework for attributing flood changes to the potential drivers of extreme precipitation, antecedent soil moisture, and snowmelt, as a function of the return period, in a regional context. Timeseries collected by reference hydrological networks can be useful in hydrological change analyses (Burn *et al.* 2012, Whitfield *et al.* 2012).

For trend tests, such a relation between statistical results and hydrological evidence is even more important, since trends are often extrapolated into the future (see Iliopoulou and Koutsoyiannis (2020) on the possible issues connected to extrapolating trends). Changes in climate or anthropogenic changes like land-use can be the origin of such trends (Prosdocimi *et al.* 2015, Blöschl *et al.* 2019b, Bertola *et al.* 2021). However, cyclic behaviour or so-called flood-rich and flood-poor periods (Lun *et al.* 2020, Fischer *et al.* 2023) in the time series can have an impact on the significance of the results (Koutsoyiannis 2003). Depending on what observation period is considered, trends and change points can be significant or not (Kundzewicz *et al.* 2005, Serinaldi *et al.* 2018). As an example, Fig. 3 depicts varying significant linear trends of the mean (continuous coloured lines) detected for 30-year time windows of the annual maximum discharge series (AMS) at the Cologne/Rhine gauge (Germany).

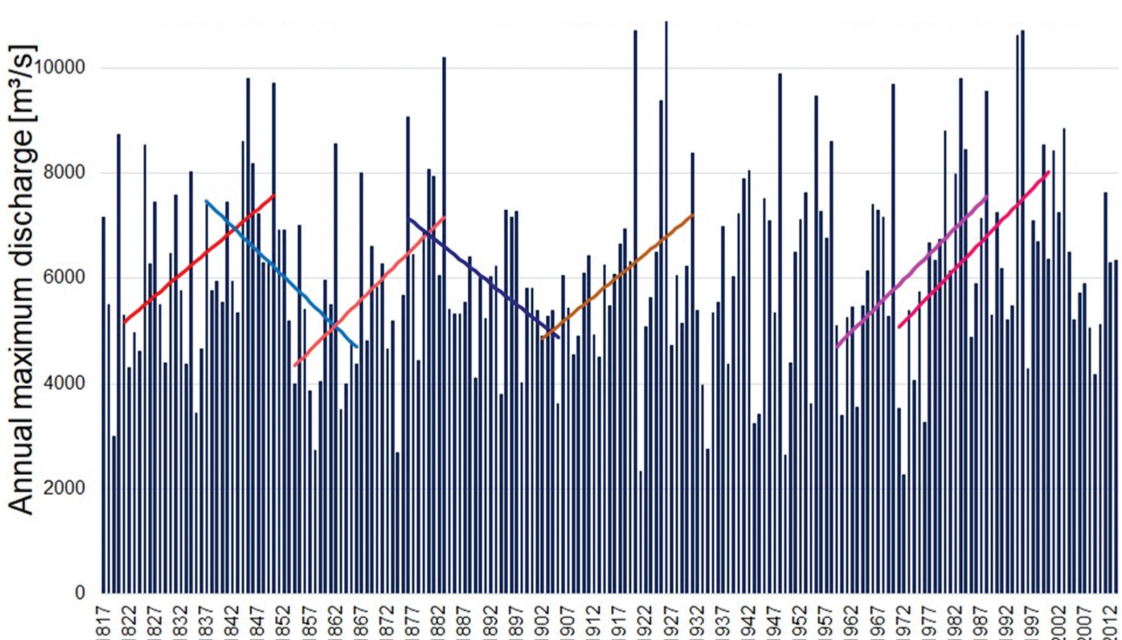

Figure 3. Example of varying significant trends detected for 30-year time windows of the annual maximum discharge series (AMS) at the Cologne/Rhine gauge; continuous coloured lines show the linear trends of the mean as emerging from the series.

Also, the time resolution of the data can have an impact on the results (Mangini *et al.* 2018). Several trends in hydrological time series around the world, like discharge or precipitation, have been detected in recent years, but not all of them can be confidently explained with evidence (Madsen *et al.* 2014). For example, precipitation time series may show a significant increasing trend while the corresponding flood time series do not (see also Sharma *et al.* 2018). Evidence of deterministic causes must be present for the incorporation of non-stationarities in hydrological models and especially for the extrapolation to the future.

5 Regionalization

The understanding and modelling of hydrological variability in space and the possibility of transferring hydrological knowledge and information from one region/catchment to another are central issues in hydrology for several scientific research areas and operational problems about water resources management and hydroclimatic risk assessment and mitigation; see e.g. the predictions in ungauged basins (PUB) initiative (Blöschl *et al.* 2013, Hrachowitz *et al.* 2013) and UPH Question 5 of 23 (Blöschl *et al.* 2019a).

This section addresses the state of the art, milestones, and future research avenues in the broad area of statistical regionalization of hydrological information by looking primarily at the prediction of rainfall and streamflow in ungauged locations. In particular, we look at the spatial interpolation of rainfall characteristics and frequency regimes, with a specific focus on rainfall extremes, and regionalization of streamflow regimes and flood frequency, looking primarily at (geo)statistical interpolation and regional frequency analysis. We also consider the related question of streamflow prediction in ungauged locations using rainfall–runoff models with regionalized parameters.

5.1 State of the art

5.1.1 Rainfall regionalization

Precipitation measurements are point observations usually obtained from gauging networks with sensors irregularly distributed in space. If rainfall is needed for ungauged points or areal rainfall is required, as for most hydrological applications, precipitation measurements have to be interpolated in space (e.g. Koutsoyiannis and Langousis 2011). There are many interpolation methods available, which can be mainly separated into deterministic/geometric approaches and geostatistical techniques. Deterministic approaches like Thiessen polygons or inverse distance weighting (IDW) are simple and fast, but do not consider the specific spatial persistence behaviour of the rainfall, nor can they use additional information sufficiently. Geometric spline interpolation is more advanced and can also use additional variables (Hutchinson 1998a, 1998b). Geostatistical techniques, relying on the concept of random functions providing manifold kriging versions, from simple stationary approaches with ordinary kriging over indicator kriging for precipitation occurrence interpolation (Berezowski *et al.* 2016) to non-stationary (in space) methods like co-kriging (Seo *et al.* 1990a, 1990b) and external drift kriging, are the state of the art. The latter especially allows easy incorporation of additional information into the interpolation, like elevation (Goovaerts 2000), satellite data (Thiemig *et al.* 2013), or weather radar measurements (Haberlandt 2007, Goudenhoofdt and Delobbe 2009).

An accurate estimation of extreme rainfall intensities is often needed for the design and assessment of urban infrastructure to minimize potential damage to society. This may be a challenge since records from rainfall gauges are often short and there is a need to also account for the effects of climate change. If only short records are available the estimation of the magnitude of events associated with long return periods is very uncertain. By combining information from different stations, i.e. by trading space for time, regionalization is a possible strategy to reduce the uncertainty in the estimation of design events for locations with short records or even ungauged locations.

Hence, a fundamental component of the process of setting up an interpolation procedure is the assessment of the accuracy of prediction in ungauged locations, which is usually done with cross-validation or split-sampling and enables the user to characterize spatial interpolation errors. Kriging methods are unbiased in space and regarded as standard tools for interpolating point rainfall. However, if time series are interpolated in space for each time step, the accumulated error for certain points can have a temporal bias, which is significant for hydrological applications (Bárdossy and Pegram 2013). A simple assessment of uncertainty is provided using estimates and estimation variance when the residuals are normally distributed. If this is not the case, indicator approaches or new interpolation methods based on copulas can be used (Bárdossy and Li 2008).

Regional frequency analysis of hydrological variables is a mature discipline, for which books and manuals exist detailing steps, procedures, statistical tests, and tools for setting up and testing regional models; a classic example is the textbook *Regional Frequency Analysis: An Approach Based on L-Moments* by Hosking and Wallis (1997). Concerning regional frequency analysis of rainfall extremes, geographic distance and mean annual precipitation (MAP) are generally recommended as similarity measures to be used to pool regional information from gauged sites due to the marked spatial nature of the rainfall extremes frequency regime (e.g. Ball *et al.* 2019) and its significant dependence on climate and orography, which can be effectively summarized by MAP (see e.g. Di Baldassarre *et al.* 2006, Persiano *et al.* 2020). Regional frequency analysis of rainstorms may be applied under different approaches, such as regional regression relationships (e.g. Brath *et al.* 2003), the index-event approach (e.g. Dalrymple 1960, Hosking and Wallis 1997, Burn 2014), or hierarchical regionalization (e.g. Alila 1999), among others. Some countries present official guidelines based on the use of a regional approach for rainfall frequency analysis, such as the *Flood Estimation Handbook* in the United Kingdom (Institute of Hydrology 1999). In other countries, the application of local approaches in which only the rainfall records at a given gauge station are used for estimating extreme rainfall events is still a common practice, instead (e.g. see Svensson and Jones 2010, where a review of nationwide rainfall frequency analysis procedures is provided considering nine countries).

The temporal resolution of precipitation matters for the selection of the optimal interpolation method and regarding interpolation performance. Usually, the interpolation error increases with increasing temporal resolution (see e.g. fig. 1 in Berndt and Haberlandt 2018), while topography is only of value as additional information for rainfall interpolation if the time steps are larger, becoming significant for data with about a weekly aggregation level (Bárdossy and Pegram 2013, Berndt and Haberlandt 2018).

Most relevant for short-time-step rainfall estimation in space are weather radar data. However, these usually have a large bias (Krajewski and Smith 2002, Berne and Krajewski 2013). Merging ground-based point rainfall measurements and radar-derived rainfall can provide corrected rainfall estimates with high resolution in space and time. Over the last few years several different merging methods have been applied, like external drift kriging (Haberlandt 2007, Goudenhoofdt and Delobbe 2009), conditional merging (Sinclair and Pegram 2005, Berndt *et al.* 2014, Kim *et al.* 2016), simple bias correction (Thorndahl *et al.* 2014, Rabiei and Haberlandt 2015) and Bayesian approaches (Todini 2001). An overview and comparison of merging methods especially targeting urban hydrological applications are provided by Ochoa-Rodriguez *et al.* (2019).

Finally, many recent developments in regional frequency analysis of hydrological variables aim at modelling or detecting climate signals and frequency alterations at a regional scale. Concerning rainfall events, evidence of the effect of climate change on extreme precipitation is found aroundthe world (e.g. Westra *et al.* 2014, Papalexiou and Montanari 2019, Persiano *et al.* 2020, Emmanouil *et al.* 2022, 2023), hence the need for it to be properly considered when estimating extreme rainfall intensities.

5.1.2 Streamflow regionalization

Analogously to rainfall, an accurate representation of streamflow regime and frequency, from low flows to floods, is of paramount importance for various water resources planning and management problems that are at the core of the safety and growth of societies. Achieving this objective is hampered by the limited density of stream-gauging networks, which are much sparser than rain-gauging ones even in more advanced countries (see e.g. Parajka *et al.* 2015), and often subject to a decline in the number of sensors in time due to high maintenance costs. For this reason, since the 1960s the hydrological scientific community has dedicated huge research attention and efforts to the development of tools and procedures for transferring hydrological information from streamgauges to ungauged river cross-sections. As summarized for instance by Blöschl *et al.* (2013; see e.g. chapters 7 and 9), classical approaches adopted for this transfer mainly consisted of statistical hydrological regionalization (see Fig. 4); these originally took the form of (log)linear regression of streamflow, low-flow, or flood quantiles against climate and geomorphological descriptors of gauged river basins, or the index-event approach (or "index-flood approach" when referring to flood flows) mentioned above.

Regional flood frequency analysis grew massively as a research topic from the mid-1980s to the revolutionary late 1990s and early 2000s when the paradigm of focused pooling was introduced (see e.g. Burn 1990, Reed 2002). Focused pooling consists of identifying poolinggroups of gauged sites from which to transfer the hydrological information by looking at hydrological similarities with the target site where the regional prediction is needed; hence, the approach dispenses with the delineation of geographically identifiable and hydrologically homogeneous regions (see cases a and b in Fig. 4), a hypothesis at the basis of the classical index-event approach, and adopts flexible pooling-groups of sites that depend on the target site (Ouarda *et al.* 2001; see Fig. 4, case c). A prominent example of focused-pooling is the so-called region of influence (ROI) approach (Burn 1990). Focusedpooling (e.g. the ROI approach) is generally considered the baseline regionalization approach and standard practice, to the point that e.g. the ROI approach is used in nationwide guidelines such as the Australian Rainfall and Runoff (ARR) guideline (Ball *et al.* 2019) and the *Flood Estimation Handbook* in the United Kingdom, as well as in nationwide studies (see e.g. Requena *et al.* 2019a in Canada). In some other areas, regression-based regional models are indicated as the national reference procedure. Examples are the Guidelines for Determining Flood Flow Frequency (Bulletin 17C) in the US, which recommend the Bayesian generalized least squares (B-GLS) approach for

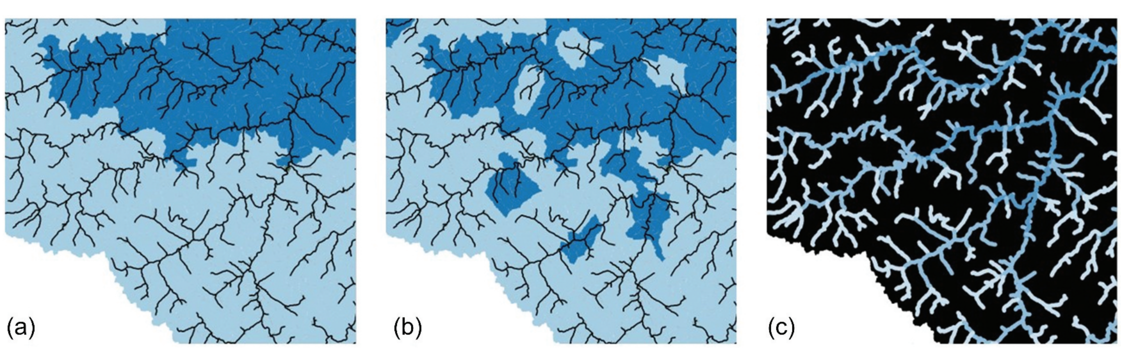

Figure 4. Spatial support for statistical regionalization: (a) classical interpretation (fixed and geographically contiguous homogeneous regions); (b) modern interpretation (non-contiguous homogeneous regions, resulting from e.g. cluster analysis of catchment descriptors); (c) current interpretation (stream network-based regionalization; typical of focused pooling, and more recently geostatistical or physiographic space-based interpolation).

setting up regional estimation equations (see e.g. Gruber and Stedinger 2008).

More recently, the paradigm of focusedpooling further evolved into procedures aiming at performing a seamless interpolation of streamflow descriptors (e.g. low-flow indices, flood quantiles, etc.) along streamnetworks (see Fig. 4, case c), therefore reproducing the hydrological continuity of streamflow regimes from our fragmented knowledge originating from the sparseness of sensors in stream gauging networks. Examples are methods that regress streamflow statistics, instead of quantiles, against catchment descriptors (Laio *et al.* 2011), or methods that contaminate hydrological regionalization principles with statistical interpolation such as kriging (see e.g. Merz and Blöschl 2005, Farmer 2016), either performed over the geographical space (geostatistics in the strict sense) or in an *n*-dimensional space whose coordinates are geomorphological and climatic catchment descriptors, which is also referred to as physiographic-space-based interpolation or canonical kriging in the scientific literature (see e.g. Shu and Ouarda 2007).

Topological kriging, or top-kriging, is a prominent example of the first category; similar to block-kriging in which the support area coincides with the drainage area, top-kriging was firstly introduced for regionalizing flood-quantiles (see Skøien *et al.* 2006) and then successfully applied in various hydro-climatic contexts for predicting a broad spectrum of point streamflow indices, such as lowflows (see e.g. Laaha *et al.* 2014), high flows and floods (Merz *et al.* 2008), flow duration curves (FDCs) (Pugliese *et al.* 2014, 2016, Castellarin *et al.* 2018), stream temperature (Laaha *et al.* 2013), habitat suitability indices (Ceola *et al.* 2018), and daily streamflow series (Skøien and Blöschl 2007, Vormoor *et al.* 2011, de Lavenne *et al.* 2016, Farmer 2016). Physiographic space-based interpolation, sometimes also referred to as canonicalcorrelation (Ouarda *et al.* 2001), proved to be a rather effective regionalization technique, which can be successfully applied for predicting flood quantiles (Shu and Ouarda 2007), lowflows (Castiglioni *et al.* 2011) and FDCs (Castellarin 2014). Recently, artificial intelligence and machine learning-based approaches have been gaining popularity due to the easy availability of computing power (e.g. Ouali *et al.* 2017, Desai and Ouarda 2021). These approaches need to be used with care to ensure that the physical concepts and the hydrological phenomena are adequately taken into consideration in the modelling effort.

Additionally, a promising regionalization approach involves the use of panel regression models, a methodology borrowed from the econometrics literature. Panel models represent the relationship between a time series of streamflow at one watershed and a time series of streamflow and/or time series of basin hydroclimatic characteristics at numerous other sites within a region. In other words, a panel model is a multivariate time series model especially designed for the types of regional time series problems encountered in hydrology. The economists Croissant and Millo (2008) provide R software for the implementation of panel regressions. Multicollinearity and omitted-variable bias are major limitations to developing multivariate (panel) regression models to estimate streamflow characteristics in ungauged watersheds (Farmer *et al.* 2015). Since the work of Brown *et al.* (2011) in documenting the impact of drought on economic growth, panel models have been increasingly used in hydrology for a number of applications. Panel models have been developed to evaluate the effect of urbanization on flood frequency (Blum *et al.* 2020), the impact of rainfall on low streamflow (Bassiouni *et al.* 2016), the prediction of groundwater levels (Izady *et al.* 2012), residential water demand modelling (Worthington *et al.* 2009), and determining the impact of urbanization on annual runoff coefficients (Steinschneider *et al.* 2013).

Another interesting approach for streamflow prediction in ungauged basins through statistical regionalization is the use of rainfall–runoff models with "regionalized" parameter values, which enables continuous simulation. This research area has been investigated using at least two distinct strategies: (i) "direct" regionalization of rainfall–runoff model parameter values, and (ii) "indirect" parameter regionalization through calibration to regionalized flow signatures. The first strategy can be implemented by spatial interpolation of model parameters calibrated at gauged locations (e.g. Merz and Blöschl 2004, Bárdossy and Li 2008, Vogel 2010). Earlier studies estimated rainfall–runoff model parameters at each site followed by attempts to relate model parameters to basin characteristics. The study by Fernandez *et al.* (2000) implemented both steps concurrently so that the multiple site models are calibrated simultaneously. Simultaneous estimation offers a better chance to reproduce observed streamflow behaviour across multiple sites and, importantly, to obtain more stable relationships between rainfall–runoff model parameters and basin characteristics. A significant advance along this research direction was achieved in the study by Samaniego *et al.* (2010), which developed detailed transfer functions that link catchment attributes to model parameters and calibrated the "hyper-parameters" of these transfer functions at multiple catchments across geographically large areas.

The second strategy, which normally yields smoother parameter variation in space, can be implemented by estimating relationships between flow signatures (e.g. mean and variance of flows, baseflow/flashiness indexes, etc.) and catchment attributes (e.g. catchment slope, geology, etc.). Traditionally, these relationships have been constructed using multi-linear regression techniques (e.g. Yadav *et al.* 2007) and more recently using machine learning techniques such as random forest (e.g. Snelder *et al.* 2013, Addor *et al.* 2018, Prieto *et al.* 2019). The identification of rainfall–runoff model structure under ungauged conditions and its impact on streamflow estimation has also received attention (Prieto *et al.* 2022). Regardless of the strategy, obvious challenges persist in terms of establishing spatial relationships for the pertinent quantities – which are essentially the same as the challenges in establishing spatial relationships in flood frequency analysis.

More recently, the main objective of regional frequency analysis may be properly accounting for non-stationarities and changes of frequency regimes at the regional scale (see e.g. Cunderlik and Burn 2003, Leclerc and Ouarda 2007, Bertola *et al.* 2020), which may imply using detected alterations and trends to form hydrologically similar regions (O'Brien and Burn 2014). Alternatively, a few studies aim at improving the robustness and reliability of trend detection in annual flood sequences by testing for trends at a regional scale (Kjeldsen and Prosdocimi 2021). Some non-stationary regional flood frequency analysis techniques proposed in the literature have been recently compared to each other, yet the body of studies in this area is still very limited and worth expanding soon (Kalai *et al.* 2020).

Together with textbooks and guidelines, several packages for implementing and testing regional statistical models, including geostatistical interpolation models such as top-kriging, have been developed under the free and open-source software R (R Core Team 2020). Table 4 reports some R packages developed and maintained by ICSH-STAHY members.

5.2 Challenges and open problems in regionalization

Geostatistical interpolation of point information produces a smoothed continuous representation of the point variable of interest in space (e.g. interpolation of point rainfall via ordinary kriging, but also interpolation of a given streamflow-index via top-kriging or physiographic-space-based interpolation). One of the main disadvantages of this procedure is the loss of variance, i.e. the underestimation of high values and the overestimation of low values (see e.g. Castellarin 2014). The preservation of variance should be considered when evaluating interpolation methods (Berndt *et al.* 2014). Usually, the assimilation of additional information, especially information deriving from weather radar for point rainfall, may improve the representation of the actual variance of the process. A noteworthy approach is the "maintenance of variance extension methods" (MOVE) introduced by Hirsch (1982). MOVE methods have been applied to rainfall, streamflow, and many other hydrological records, and extensions were developed to ensure that they perform as expected under either augmentation or extension (Vogel and Stedinger 1985). Augmentation

Table 4. Packages for regionalization analysis and modelling.

Name	Brief description	Repo.	Link	Authors
floodnetRfa	Package implementing FloodNet recommendations for flood frequency analysis	GitHub	https://github.com/floodnetProject16/floodnetRfa	Durocher *et al.*
geoFDC	Nonsupervised and geostatistically interpolated regional FDCs	GitHub	https://github.com/alessio-pugliese/geoFDC	Pugliese
lmomRFA	Functions for regional frequency analysis using the methods of Hosking and Wallis (1997)	CRAN	https://CRAN.R-project.org/package=lmomRFA	Hosking
ncdf4	Interface to Unidata netCDF (Version 4 or earlier) format data files	CRAN	https://CRAN.R-project.org/package=ncdf4	Pierce
nsRFA	A collection of statistical tools for objective (non-supervised) applications of the regional frequency analysis methods in hydrology	CRAN	https://CRAN.R-project.org/package=nsRFA	Viglione *et al.*
pREC	Package for quantifying regional information content for cross-correlated annual sequences, useful for assigning a frequency to regional envelope curves	GitHub	https://github.com/alessio-pugliese/pREC	Pugliese
RMWSPy	Python package for conditional spatial random field simulation and inverse modelling which can be used to simulate rainfall fields conditioned on commercial microwave link data	GitHub	https://github.com/SebastianHoerning/RMWSPy	Hörning
Rtop	Geostatistical interpolation of data with irregular spatial support such as runoff-related data or data from administrative units (to-kriging)	CRAN	https://cran.r-project.org/web/packages/rtop/index.html	Skøien *et al.*
TNDTK	Repository containing an application example for extracting period-of-record (FDCs) from daily streamflow series observed at gauged sites and computing FDCs at ungauged target sites using total negative deviation top-kriging	GitHub	https://github.com/SimonePersiano/TNDTK/tree/v1.0.0	Persiano
winfapReader	Interact with peak flow data in the United Kingdom	CRAN	https://cran.r-project.org/web/packages/winfapReader/index.html	Prosdocimi *et al.*
WREG	Develops regional estimation equations for streamflow characteristics that can be applied at ungauged basins (approaches allowed: ordinaryleastsquares, OLS; weightedleastsquares, WLS; and generalizedleastsquares, GLS)	GitHub	https://github.com/USGS-R/WREG	Farmer and USGS Team
statsmodels	Python module that provides classes and functions for the estimation of many different statistical models, as well as for conducting statistical tests, and statistical data exploration	statsmodels.org	https://www.statsmodels.org/stable/index.html	Seabold and Perktold
scipy.stats	Python module that contains a large number of probability distributions, summary and frequency statistics, correlation functions and statistical tests, masked statistics, kernel density estimation, quasi-Monte Carlo functionality, and more	scipy.org	https://docs.scipy.org/doc/scipy/reference/stats.html	Virtanen *et al.*
pyMannKendall	Python package for non-parametric Mann-Kendall family of trend tests	pypi.org	https://pypi.org/project/pymannkendall/	Hussain
pyHomogeneity	Python package for homogeneity test	pypi.org	https://pypi.org/project/pyhomogeneity/	Hussain

refers to the estimation of moments of hydrological records at short-record sites by transferring information from nearby sites with longer records. Extension refers to filling in the observations at the short-record site using the longer record site. MOVE methods are in common use by the US Geological Survey, as evidenced by their inclusion in the USGS R package already mentioned in Table 1. More recently, conditional simulation methods have been introduced to preserve the variance and provide uncertainty estimation (AghaKouchak *et al.* 2010, Seo *et al.* 2014, Hörning and Bárdossy 2018). It is relevant to mention also the patched-kriging approach proposed by Libertino *et al.* (2018), dealing with the loss of variance due to spatial interpolation of point maxima. Notwithstanding these advancements, the preservation of variance represents an interesting open problem for future analyses (see also Farmer and Vogel 2016). The relatively sparse precipitation networks very likely miss the true extremes of an event. This causes a serious problem for estimating point and areal extremes.

The quality of regionalization is often measured with correlation, Nash-Sutcliffe efficiency, or similar measures; the reader is referred to Barber *et al.* (2020), Lamontagne *et al.* (2020) and Clark *et al.* (2021)for a discussion on the variability of such measures when computed from daily streamflow. Unfortunately, in the case of a strong annual cycle, these numbers often indicate a much higher skill. Therefore, well-known annual cycles should be removed before the evaluation.

Focusing more specifically on point rainfall, it must be noted that precipitation is a space-time dynamic process, which is an issue also for multisite rainfall generation (Section 4.2.2). Therefore, a pure spatial interpolation is not optimal, as it does not consider the advection of rainfall fields and their temporal persistence, which can be especially relevant for a short-duration event. Special interpolation methods, which can account for this space-time behaviour, might improve the performance of rainfall interpolation (Fitzner and Sester 2015). Also, the interpolation of distribution functions, honoring the spatio-temporal aspects, might improve the interpolation performance (Lebrenz and Bárdossy 2019) and it is, therefore, worth investigating further. Another important problem is the large portion of zero rainfall. The number of dry stations increases with finer spatial resolution. This affects both the assessment of spatial variability and spatial/space-time interpolation.

Although it has been shown that blending weather radar data with raingauge observations generally results in smaller interpolation errors relative to using either type of data separately, the use of the data merger, e.g. as input for hydrological modelling, is not standard practice yet (Berne and Krajewski 2013). This might be due to the complexity of developing the merger products, or too little confidence in radar data products, but also to the specific result of specific scientific research projects (Price *et al.* 2014, Zhu *et al.* 2014). Research initiatives could advance our knowledge, providing valuable information for practical applications, including recent research activities considering the potential of alternative data sources to be used for spatial interpolation of rainfall, such as crowdsourced information (Haberlandt and Sester 2010, de Vos *et al.* 2019, Bárdossy *et al.* 2021) or data on the attenuation of electromagnetic signals in commercial microwave links (see e.g. Haese *et al.* 2017) or recordings of video cameras (Allamano *et al.* 2015). The combination of these alternative data sources with conventional raingauges shows great potential for improving the spatial interpolation of rainfall data and offers interesting opportunities for future research on merging techniques.

While, as noted above, the spatial interpolation of point rainfall may significantly benefit from information on spatial correlation of rainfall fields gathered by weather radar data (or commercial microwave links), we are currently lacking reliable information sources for improving our representation and modelling of spatial correlation (also termed intersite correlation, or cross-correlation) for regional frequency analyses of hydrological variables, and particularly so for extreme events such as rainstorms, floods, or lowflows.

Intersite correlation is a valuable indicator of hydrological similarity that can be effectively exploited in regional flood frequency analysis (see e.g. Archfield and Vogel 2010), and that is directly modelled and used by top-kriging and GLS for producing regional predictions of the hydrological variable of interest. Nevertheless, the effects of intersite correlation also need to be better understood and quantified as they control the actual information content of a regional sample (see e.g. Castellarin 2007 and references therein), and consequently the uncertainty and accuracy of regional predictions (Guo *et al.* 2021), they also impact statistical tests for assessing the regional homogeneity of a pooling group of sites (Castellarin *et al.* 2008, Lilienthal *et al.* 2018). Improving our understanding and representation of intersite cross-correlation is one of the key elements for enhancing the accuracy of regional estimates, bringing regional estimators closer to the unknown true theoretical values (Persiano *et al.* 2021).

Another serious problem is that precipitation and discharge are usually regionalized independently, which very often leads to inconsistent datasets (Kauffeldt *et al.* 2013). Therefore, methods to regionalize dependent variables simultaneously are required. Attempts to do so can be found in Grundmann *et al.* (2019) and Bárdossy *et al.* (2021).

The choice of distribution of the extreme values is a key decision in most regionalization studies. In the stationary case, most studies and procedures rely on well-known 2–4-parameter models such as the Log Pearson 3 (LP3), GEV or kappa distributions (see e.g. Hosking and Wallis 1997). Regionalization of, in particular, the shape parameter is often challenging due to a combination of high sampling variance of the third-order moment and the generally weak link to existing catchment descriptors (e.g. Lun *et al.* 2021). Another challenge is the need to select appropriate model structures when moving from stationary to non-stationary models. For example, Prosdocimi and Kjeldsen (2021) showed that extrapolation from popular non-stationary models can result in counter-intuitive estimates of future design floods when the location parameter is changing while the scale parameter is kept constant. Instead, they proposed a re-parametrized non-stationary version of the GEV model which preserves a constant coefficient of variation and thus ensures credible extrapolations. The challenge of selecting and regionalizing non-stationary models needs to be considered carefully to ensure the operational value of these models.

Finally, and more specifically on non-stationarity, quoting Faulkner *et al*. (2020), can we still predict the future from the past? Or equivalently, in the face of global environmental change (see Visessri and McIntyre 2016), is "trading space for time" still a reliable and viable working hypothesis? These are still very relevant and fundamental research questions for the PUB problem in the field of regional frequency analysis of hydrological variables. Our ever-growing computational capabilities combined with currently available cloud-computing possibilities and steadily increasing open accessibility to high-resolution global-coverage datasets seem to indicate regionalization of hydrological models (Guo *et al*. 2021) and continuous rainfall–runoff simulation as a promising way forward, one worth investing research efforts and resources in. In this context, regional modelling of extreme rainfall events, in terms of magnitude, spatial distribution, and frequency, as well as future climate scenarios, assume pivotal importance.

Concerning frequency analysis of rainfall extremes, further efforts could be dedicated to promoting the use of regional approaches for practical application. A global comparative assessment of the performance of regional rainfall approaches, in a similar way to that carried out for floods and lowflows by Salinas *et al*. (2013), would be beneficial for providing general recommendations. Regarding future climate scenarios, and focusing in particular on climate model simulations used and frequency analysis for estimating regional and continental future hydrological extreme events, global initiatives such as the Coordinated Regional Climate Downscaling Experiment (CORDEX, https://cordex.org) are currently in progress; they aim at gathering regional climate downscaled simulations from many climate models under similar scenarios, projections, resolutions, time scales and periods for facilitating their use in practice. Considering the wide variety of approaches to be used for estimating future extreme rainfall, some studies attempt to provide recommendations but the literature is still sparse. For instance, bias correction of climate model simulations for their use in extreme rainfall estimation is recommended to be applied to annual maxima under a regional approach (Li *et al*. 2017), using simulations and observations with a similar spatial and temporal resolution (Maraun 2013). So far, only a few studies deal with the estimation of future extreme rainfall by accounting for a regional approach (e.g. Ekström *et al*. 2005, DeGaetano and Castellano 2017, Li *et al*. 2017, Requena *et al*. 2019). Due to the large uncertainty involved in future extreme rainfall estimation (e.g. uncertainty in climate model simulations, spatial and temporal downscaling methods, re-gridding method, and bias correction methods, among others), the recurring general recommendation consists in considering as many methods as possible, and therefore this issue is still wide open for future research contributions.

6 Conclusions

In this paper, we provide a summary of the collaborative activities undertaken within the STAHY-WG (Statistical Hydrology Working Group) and later the ICSH (International Commission on Statistical Hydrology) of IAHS (International Association of Hydrological Sciences). The multidisciplinary nature of this community encompasses various research fields. However, in this paper, we specifically concentrate on four areas that have garnered significant attention over the past 15 years and we discuss open problems and challenges that persist within the domains of extreme value analysis, multivariate frequency analysis, time series analysis and simulation, and regionalization.

Regarding extreme value analysis, besides extensively studying the appropriate distributions such as the common GEV and GP functions, a recent promising development has emerged. It pertains to the potential to relax the constraint of the independence condition (i.e. metastatistical distribution), which would enable simplifying the frequency analysis and expanding and optimizing the available sample size. However, the issue of non-stationarity remains an unresolved challenge, despite considerable efforts devoted to its investigation. It needs to be adequately and comprehensively addressed in a general context.

Multivariate analysis, particularly involving the copula function, has gained recognition as a best practice. At the outset of the collaborative endeavours between STAHY-WG and the ICSH, the copula function was introduced in hydrology. While it showed great potential, there were several drawbacks in the inference procedure and limitations in its applications. However, through intensive activities such as short courses and workshops, significant progress has been made. The entire procedure has been reviewed, improved, and solidified, enhancing its applicability in hydrology. Currently, the multivariate inference procedure is wellestablished, similar to the univariate case. This contribution emphasizes certain considerations, such as the importance of conducting a preliminary analysis, carefully considering sample size, and exploring the full range of available copula functions. Looking ahead, a future challenge lies in eliminating the requirement for marginal autocorrelation, in alignment with the univariate case.

On the topic of time series analysis, in addition to strengthening established procedures for streamflow simulation, a substantial effort has been dedicated to rainfall simulation models, which are crucial for various hydrological analyses. One notable challenge is the existence of multiple approaches with different theoretical backgrounds, which can be confusing for practitioners. We acknowledge that it is impractical to identify a single best rainfall simulation model suitable for all applications. Therefore, the future challenge lies in two aspects. First, it involves establishing a proper classification system that links the different models to their specific applications. This would aid practitioners in selecting the most appropriate model for their needs. Second, there is a need to enhance the technological transfer process, making the rainfall simulation models more user-friendly for the endusers. By addressing these challenges, we can improve the accessibility and usability of rainfall simulation models, benefiting the broader community of users.

Regarding regional estimation techniques, there is a pressing need to consider the multivariate nature of the hydroclimatic variables under study (e.g. floods, lowflows, extreme rainfall events, droughts). This requires viewing them as intertwined elements of a single and complex unicum, and correctly representing their relationships and mutual constraints.

Undoubtedly, this contribution provides only a limited overview, acknowledging that numerous other intriguing and critical topics in statistical hydrology warrant similar analysis. This review aims to guide and support the future STAHY community in identifying the next research challenges in statistical hydrology, a field currently undergoing a revival driven in part by the rise of machine learning, big data, and artificial intelligence in hydrology. These are all sub-fields of, or heavily reliant on, the fundamental methods of statistical hydrology (as pointed out by R.M. Vogel in his review of this paper). Therefore, the interpretation of who should be part of the future STAHY community should be as broad and inclusive as possible.

Additionally, the ICSH recognizes the importance of stimulating the hydrological community to contribute solutions through the application of statistical hydrology tools. This aligns with the overarching goal of the IAHS to advance the field of hydrology and address key challenges through the HELPING – Science for Solution initiative in the coming decade. By embracing new technological advancements and fostering collaboration between data science and statistical methods, the hydrological community can strive towards innovative solutions and make significant progress in addressing complex water-related issues.

Acknowledgements

We thank the editor, Stacey Archfield, and two reviewers, Richard Vogel and Alberto Viglione, for their thoughtful comments, which helped to significantly improve the original version of the manuscript. The authors also thank the ICSH-STAHY members, who contributed to the discussions during the STAHY workshops and the IUGG and IAHS General Assemblies.

Disclosure statement

No potential conflict of interest was reported by the authors.

ORCID

Elena Volpi http://orcid.org/0000-0002-9511-1496
Attilio Castellarin http://orcid.org/0000-0002-6111-0612
Simon Michael Papalexiou http://orcid.org/0000-0001-5633-0154
Hafzullah Aksoy http://orcid.org/0000-0001-5807-5660
Ebru Eris http://orcid.org/0000-0003-0601-7666
Svenja Fischer http://orcid.org/0000-0002-6118-0223
Ilaria Prosdocimi http://orcid.org/0000-0001-8565-094X
Elena Ridolfi http://orcid.org/0000-0002-4714-2511

References

Adarsh, S., *et al.*, 2020. Multifractal description of daily rainfall fields over India. *Journal of Hydrology*, 586, 124913. doi:10.1016/j.jhydrol.2020.124913

Addor, N., *et al.*, 2018. A ranking of hydrological signatures based on their predictability in space. *Water Resources Research*, 54 (11), 8792–8812. doi:10.1029/2018WR022606.

AghaKouchak, A., *et al.*, 2020. Climate extremes and compound hazards in a warming world. *Annual Review of Earth and Planetary Sciences*, 48, 519–548. doi:10.1146/annurev-earth-071719-055228

AghaKouchak, A., Bárdossy, A., and Habib, E., 2010. Copula-based uncertainty modelling: application to multisensor precipitation estimates. *Hydrological Processes*, 24 (15), 2111–2124. doi:10.1002/hyp.7632.

Aksoy, H. and Bayazit, M., 2000. A model for daily flows of intermittent streams. *Hydrological Processes*, 14 (10), 1725–1744. doi:10.1002/1099-1085(200007)14:10<1725::AID-HYP108>3.0.CO;2-L.

Alexander, L.V., *et al.*, 2006. Global observed changes in daily climate extremes of temperature and precipitation. *Journal of Geophysical Research: Atmospheres*, 111 (D5). doi:10.1029/2005JD006290.

Alexandersson, H., 1986. A homogeneity test applied to precipitation data. *Journal of Climatology*, 6 (6), 661–675. doi:10.1002/joc.3370060607.

Alila, Y., 1999. A hierarchical approach for the regionalization of precipitation annual maxima in Canada. *Journal of Geophysical Research: Atmospheres*, 104 (D24), 31645–31655.

Allamano, P., Croci, A., and Laio, F., 2015. Toward the camera rain gauge. *Water Resources Research*, 51 (3), 1744–1757. doi:10.1002/2014WR016298.

Altman, N.S., 1992. An introduction to kernel and nearest-neighbor nonparametric regression. *The American Statistician*, 46 (3), 175–185. doi:10.1080/00031305.1992.10475879.

Archfield, S.A. and Vogel, R.M., 2010. Map correlation method: selection of a reference streamgage to estimate daily streamflow at ungaged catchments. *Water Resources Research*, 46 (10). doi:10.1029/2009WR008481.

Ashkar, F. and Aucoin, F., 2011. A broader look at bivariate distributions applicable in hydrology. *Journal of Hydrology*, 405 (3–4), 451–461. doi:10.1016/j.jhydrol.2011.05.043.

Ball, J., *et al.*, 2019. *Australian Rainfall and Runoff: A Guide to Flood Estimation*. © Commonwealth of Australia (Geoscience Australia), Version 4.2, 2019.

Banfi, F., Cazzaniga, G., and De Michele, C., 2022. Nonparametric extrapolation of extreme quantiles: a comparison study. *Stochastic Environmental Research and Risk Assessment*, 36 (6), 1579–1596. doi:10.1007/s00477-021-02102-0.

Baratti, E., *et al.*, 2012. Estimating the flood frequency distribution at seasonal and annual time scales. *Hydrology and Earth System Sciences*, 16 (12), 4651–4660. doi:10.5194/hess-16-4651-2012.

Barber, C., Lamontagne, J.R., and Vogel, R.M., 2020. Improved estimators of correlation and R^2 for skewed hydrologic data. *Hydrological Sciences Journal*, 65 (1), 87–101. doi:10.1080/02626667.2019.1686639.

Bárdossy, A., 2023. Changing correlations: a flexible definition of non-Gaussian multivariate dependence. *Stochastic Environmental Research and Risk Assessment*, 1–11.

Bárdossy, A. and Li, J., 2008. Geostatistical interpolation using copulas. *Water Resources Research*, 44 (7), W07412. doi:10.1029/2007WR006115.

Bárdossy, A. and Pegram, G., 2013. Interpolation of precipitation under topographic influence at different time scales. *Water Resources Research*, 49 (8), 4545–4565. doi:10.1002/wrcr.20307.

Bárdossy, A. and Plate, E.J., 1992. Space-time model for daily rainfall using atmospheric circulation patterns. *Water Resources Research*, 28 (5), 1247–1259. doi:10.1029/91WR02589.

Bárdossy, A., Seidel, J., and El Hachem, A., 2021. The use of personal weather station observations to improve precipitation estimation and interpolation. *Hydrology and Earth System Sciences*, 25 (2), 583–601. doi:10.5194/hess-25-583-2021.

Basseville, M. and Nikiforov, I.V., 1993. *Detection of abrupt changes: theory and application*. Vol. 104, Englewood Cliffs: prentice Hall.

Bassiouni, M., Vogel, R.M., and Archfield, S.A., 2016. Panel regressions to estimate low-flow response to rainfall variability in ungaged basins. *Water Resources Research*, 52 (12), 9470–9494. doi:10.1002/2016WR018718.

Bayazit, M. and Önöz, B.J.H.S.J., 2007. To prewhiten or not to prewhiten in trend analysis? *Hydrological Sciences Journal*, 52 (4), 611–624. doi:10.1623/hysj.52.4.611.

Ben Aissia, M.A., Chebana, F., and Ouarda, T.B., 2017. Multivariate missing data in hydrology–review and applications. *Advances in Water Resources*, 110, 299–309. doi:10.1016/j.advwatres.2017.10.002

Benito, G., *et al.*, 2004. Use of systematic, palaeoflood and historical data for the improvement of flood risk estimation. *Natural Hazards*, 31, 623–643.

Benson, M.A., 1950. Use of historical data in flood frequency analysis. *Eos, Transactions American Geophysical Union*, 31 (3), 419–424.

Benson, M.A. and Matalas, N.C., 1967. Synthetic hydrology based on regional statistical parameters. *Water Resources Research*, 3 (4), 931–935. doi:10.1029/WR003i004p00931.

Beran, J., 2017. *Statistics for long-memory processes*. Boca Raton, FL, US: Chapman and Hall, CRC.

Berezowski, T., *et al.*, 2016. CPLFD-GDPT5: high-resolution gridded daily precipitation and temperature data set for two largest polish river basins. *Earth System Science Data*, 8 (1), 127–139. doi:10.5194/essd-8-127-2016.

Berg, D., 2009. Copula goodness-of-fit testing: an overview and power comparison. *The European Journal of Finance*, 15 (7–8), 675–701. doi:10.1080/13518470802697428.

Berndt, C. and Haberlandt, U., 2018. Spatial interpolation of climate variables in Northern Germany—influence of temporal resolution and network density. *Journal of Hydrology: Regional Studies*, 15, 184–202.

Berndt, C., Rabiei, E., and Haberlandt, U., 2014. Geostatistical merging of rain gauge and radar data for high temporal resolutions and various station density scenarios. *Journal of Hydrology*, 508, 88–101. doi:10.1016/j.jhydrol.2013.10.028

Berne, A. and Krajewski, W.F., 2013. Radar for hydrology: unfulfilled promise or unrecognized potential? *Advances in Water Resources*, 51, 357–366. doi:10.1016/j.advwatres.2012.05.005

Bertola, M., *et al.*, 2020. Flood trends in Europe: are changes in small and big floods different? *Hydrology and Earth System Sciences*, 24 (4), 1805–1822. doi:10.5194/hess-24-1805-2020.

Bertola, M., *et al.*, 2021. Do small and large floods have the same drivers of change? A regional attribution analysis in Europe. *Hydrology and Earth System Sciences*, 25 (3), 1347–1364. doi:10.5194/hess-25-1347-2021.

Bertola, M., Viglione, A., and Blöschl, G., 2019. Informed attribution of flood changes to decadal variation of atmospheric, catchment and river drivers in Upper Austria. *Journal of Hydrology*, 577, 123919. doi:10.1016/j.jhydrol.2019.123919

Blöschl, G., *et al.*, Eds. 2013. *Runoff prediction in ungauged basins: synthesis across processes, places and scales*. Cambrige, UK: Cambridge University Press.

Blöschl, G., *et al.*, 2017. Changing climate shifts timing of European floods. *Science*, 357 (6351), 588–590. doi:10.1126/science.aan2506.

Blöschl, G., *et al.*, 2019a. Twenty-three unsolved problems in hydrology (UPH)–a community perspective. *Hydrological Sciences Journal*, 64 (10), 1141–1158. doi:10.1080/02626667.2019.1620507.

Blöschl, G., *et al.*, 2019b. Changing climate both increases and decreases European river floods. *Nature*, 573 (7772), 108–111. doi:10.1038/s41586-019-1495-6.

Blöschl, G., *et al.*, 2020. Current European flood-rich period exceptional compared with past 500 years. *Nature*, 583 (7817), 560–566. doi:10.1038/s41586-020-2478-3.

Blum, A.G., *et al.*, 2020. Causal effect of impervious cover on annual flood magnitude for the United States. *Geophysical Research Letters*, 47 (5). doi:10.1029/2019GL086480.

Botero, B. and Francés, F., 2010. Estimation of high return period flood quantiles using additional non-systematic information with upper bounded statistical models. *Hydrol Earth Syst Sci*, 14, 2617–2628. doi:10.5194/hess-14-2617-2010

Box, G.E. and Jenkins, G.M., 1976. *Time series analysis: forecasting and control. San Francisco*. Calif: Holden-Day.

Brahimi, B., Chebana, F., and Necir, A., 2015. Copula representation of bivariate L-moments: a new estimation method for multiparameter two-dimensional copula models. *Statistics*, 49 (3), 497–521. doi:10.1080/02331888.2014.932792.

Brahimi, B. and Necir, A., 2012. A semiparametric estimation of copula models based on the method of moments. *Statistical Methodology*, 9 (4), 467–477. doi:10.1016/j.stamet.2011.11.003.

Brath, A., Castellarin, A., and Montanari, A., 2003. Assessing the reliability of regional depth-duration-frequency equations for gaged and ungaged sites. *Water Resources Research*, 39 (12). doi:10.1029/2003WR002399.

Brown, C., *et al.*, 2011. Hydroclimate risk to economic growth in sub-Saharan Africa. *Climatic Change*, 106 (4), 621–647. doi:10.1007/s10584-010-9956-9.

Buishand, T.A. and Brandsma, T., 2001. Multisite simulation of daily precipitation and temperature in the Rhine basin by nearest□neighbor resampling. *Water Resources Research*, 37 (11), 2761–2776.

Burn, D.H., 1990. An appraisal of the “region of influence” approach to flood frequency analysis. *Hydrological Sciences Journal*, 35 (2), 149–165. doi:10.1080/02626669009492415.

Burn, D.H., *et al.*, 2012. Reference hydrologic networks II. Using reference hydrologic networks to assess climate-driven changes in streamflow. *Hydrological Sciences Journal*, 57 (8), 1580–1593. doi:10.1080/02626667.2012.728705.

Burn, D.H., 2014. A framework for regional estimation of intensity–duration–frequency (IDF) curves. *Hydrological Processes*, 28 (14), 4209–4218. doi:10.1002/hyp.10231.

Burton, A., *et al.*, 2008. RainSim: a spatial–temporal stochastic rainfall modelling system. *Environmental Modelling and Software*, 23 (12), 1356–1369. doi:10.1016/j.envsoft.2008.04.003.

Cabilio, P., Zhang, Y., and Chen, X., 2013. Bootstrap rank tests for trend in time series. *Environmetrics*, 24 (8), 537–549. doi:10.1002/env.2250.

Calenda, G., Mancini, C.P., and Volpi, E., 2009. Selection of the probabilistic model of extreme floods: the case of the River Tiber in Rome. *Journal of Hydrology*, 371 (1–4), 1–11. doi:10.1016/j.jhydrol.2009.03.010.

Cameron, D.S., *et al.*, 1999. Flood frequency estimation by continuous simulation for a gauged upland catchment (with uncertainty). *Journal of Hydrology*, 219, 169–187. doi:10.1016/S0022-1694(99)00057-8

Camuffo, D., Della Valle, A., and Becherini, F., 2020. A critical analysis of the definitions of climate and hydrological extreme events. *Quaternary International*, 538, 5–13. doi:10.1016/j.quaint.2018.10.008

Cancelliere, A. and Salas, J.D., 2010. Drought probabilities and return period for annual streamflows series. *Journal of Hydrology*, 391 (1–2), 77–89. doi:10.1016/j.jhydrol.2010.07.008.

Capéraà, P., Fougères, A.-L., and Genest, C., 1997. A nonparametric estimation procedure for bivariate extreme value copulas. *Biometrika*, 84 (3), 567–577. doi:10.1093/biomet/84.3.567.

Carsteanu, A.A. and Langousis, A., 2020. Break of temporal symmetry in a stationary markovian setting: evidencing an arrow of time, and parameterizing linear dependencies using fractional low-order joint moments. *Stochastic Environmental Research and Risk Assessment*, 34, 1–6. doi:10.1007/s00477-019-01749-0

Castellarin, A., 2007. Probabilistic envelope curves for design flood estimation at ungauged sites. *Water Resources Research*, 43 (4). doi:10.1029/2005WR004384.

Castellarin, A., 2014. Regional prediction of flow-duration curves using a three-dimensional kriging. *Journal of Hydrology*, 513, 179–191. doi:10.1016/j.jhydrol.2014.03.050

Castellarin, A., *et al.*, 2018. Prediction of streamflow regimes over large geographical areas: interpolated flow–duration curves for the Danube region. *Hydrological Sciences Journal*, 63 (6), 845–861. doi:10.1080/02626667.2018.1445855.

Castellarin, A., Burn, D.H., and Brath, A., 2008. Homogeneity testing: how homogeneous do heterogeneous cross-correlated regions seem? *Journal of Hydrology*, 360 (1–4), 67–76. doi:10.1016/j.jhydrol.2008.07.014.

Castiglioni, S., *et al.*, 2011. Smooth regional estimation of low-flow indices: physiographical space based interpolation and top-kriging. *Hydrology and Earth System Sciences*, 15 (3), 715–727. doi:10.5194/hess-15-715-2011.

Ceola, S., *et al.*, 2018. Hydro-power production and fish habitat suitability: assessing impact and effectiveness of ecological flows at regional scale. *Advances in Water Resources*, 116, 29–39. doi:10.1016/j.advwatres.2018.04.002

Chandler, R.E. and Wheater, H.S., 2002. Analysis of rainfall variability using generalized linear models: a case study from the west of Ireland. *Water Resources Research*, 38 (10), 10–11. doi:10.1029/2001WR000906.

Charles, S.P., Bates, B.C., and Hughes, J.P., 1999. A spatiotemporal model for downscaling precipitation occurrence and amounts. *Journal of Geophysical Research: Atmospheres*, 104 (D24), 31657–31669.

Chebana, F., 2022. *Multivariate frequency analysis of hydro-meteorological variables: a copula-based approach*. Amsterdam, Netherlands: Elsevier.

Chebana, F., Ben Aissia, M.A., and Ouarda, T.B.M.J., 2017. Multivariate shift testing for hydrological variables, review, comparison and application. *Journal of Hydrology*, 548, 88–103. doi:10.1016/j.jhydrol.2017.02.033

Chebana, F. and Ouarda, T.B., 2021. Multivariate non-stationary hydrological frequency analysis. *Journal of Hydrology*, 593, 125907. doi:10.1016/j.jhydrol.2020.125907

Chebana, F. and Ouarda, T.B.M.J., 2011a. Depth-based multivariate descriptive statistics with hydrological applications. *Journal Geophysical Research*, 116 (D10), D10120. doi:10.1029/2010JD015338.

Chebana, F. and Ouarda, T.B.M.J., 2011b. Multivariate quantiles in hydrological frequency analysis. *Environmetrics*, 22 (1), 63–78. doi:10.1002/env.1027.

Chebana, F., Ouarda, T.B.M.J., and Duong, T.C., 2013. Testing for multivariate trends in hydrologic frequency analysis. *Journal of Hydrology*, 486, 519–530. doi:10.1016/j.jhydrol.2013.01.007

Chen, J., Brissette, F.P., and Leconte, R., 2012. WeaGETS–a matlab-based daily scale weather generator for generating precipitation and temperature. *Procedia Environmental Sciences*, 13, 2222–2235. doi:10.1016/j.proenv.2012.01.211

Chen, L., *et al.*, 2019. Copula-based method for stochastic daily streamflow simulation considering lag-2 autocorrelation. *Journal of Hydrology*, 578, 123938. doi:10.1016/j.jhydrol.2019.123938

Chiang, F., *et al.*, 2021a. A multivariate conditional probability ratio framework for the detection and attribution of compound climate extremes. *Geophysical Research Letters*, 48 (15), e2021GL094361. doi:10.1029/2021GL094361.

Chiang, F., *et al.*, 2022. Intensified likelihood of concurrent warm and dry months attributed to anthropogenic climate change. *Water Resources Research*, 58 (6), e2021WR030411. doi:10.1029/2021WR030411.

Chiang, F., Mazdiyasni, O., and AghaKouchak, A., 2021b. Evidence of anthropogenic impacts on global drought frequency, duration, and intensity. *Nature Communications*, 12 (1), 1–10. doi:10.1038/s41467-021-22314-w.

Cipollini, S., Fiori, A., and Volpi, E., 2021. Structure-based framework for the design and risk assessment of hydraulic structures, with application to offline flood detention basins. *Journal of Hydrology*, 600, 126527. doi:10.1016/j.jhydrol.2021.126527

Claps, P., Giordano, A., and Laio, F., 2005. Advances in shot noise modeling of daily streamflows. *Advances in Water Resources*, 28 (9), 992–1000. doi:10.1016/j.advwatres.2005.03.008.

Clark, M., *et al.*, 2004. The Schaake shuffle: a method for reconstructing space–time variability in forecasted precipitation and temperature fields. *Journal of Hydrometeorology*, 5 (1), 243–262. doi:10.1175/1525-7541(2004)005<0243:TSSAMF>2.0.CO;2.

Clark, M.P., *et al.*, 2021. The abuse of popular performance metrics in hydrologic modeling. *Water Resources Research*, 57 (9), e2020WR029001. doi:10.1029/2020WR029001.

Cohn, T.A. and Lins, H.F., 2005. Nature's style: naturally trendy. *Geophysical Research Letters*, 32 (23). doi:10.1029/2005GL024476.

Cohn, T.A. and Stedinger, J.R., 1987. Use of historical information in a maximum likelihood framework. *Journal of Hydrol*, 96, 215–233. doi:10.1016/0022-1694(87)90154-5

Coles, S., 2001. Classical extreme value theory and models. *In*: *An introduction to statistical modeling of extreme values*. London, UK: Springer, 45–73.

Committee on adaptation to a changing climate. 2018. *Climate-resilient infrastructure: adaptive design and risk management*. Reston, VA, US: ASCE.

Corbella, S. and Stretch, D.D., 2013. Simulating a multivariate sea storm using archimedean copulas. *Coastal Engineering*, 76, 68–78. doi:10.1016/j.coastaleng.2013.01.011

Costa, J.E., 1978. Holocene stratigraphy in flood frequency analysis. *Water Resources Research*, 14 (4), 626–632. doi:10.1029/WR014i004p00626.

Cover, T. and Hart, P., 1967. Nearest neighbor pattern classification. *IEEE Transactions on Information Theory*, 13 (1), 21–27. doi:10.1109/TIT.1967.1053964.

Cowpertwait, P., Isham, V., and Onof, C., 2007. Point process models of rainfall: developments for fine-scale structure. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 463 (2086), 2569–2587. doi:10.1098/rspa.2007.1889.

Cox, D.R. and Isham, V., 1980. *Point processes*. Vol. 12. Boca Raton, Florida, US: CRC Press.

Croissant, Y. and Millo, G., 2008. Panel data econometrics in R: the plm package. *Journal of Statistical Software*, 27 (2), 1–43. doi:10.18637/jss.v027.i02.

Cunderlik, J.M. and Burn, D.H., 2003. Non-stationary pooled flood frequency analysis. *Journal of Hydrology*, 276 (1–4), 210–223. doi:10.1016/S0022-1694(03)00062-3.

Dahlstedt, K. and Jensen, H.J., 2005. Fluctuation spectrum and size scaling of river flow and level. *Physica A Statistical Mechanics & Its Applications*, 348, 596–610. doi:10.1016/j.physa.2004.09.039

Dalrymple, T., 1960. *Flood-frequency analyses (No. 1543)*. Reston, VA, US: U.S. Geological Survey Water-Supply Paper.

Debele, S.E., Strupczewski, W.G., and Bogdanowicz, E., 2017. A comparison of three approaches to non-stationary flood frequency analysis. *Acta Geophysica*, 65, 863–883. doi:10.1007/s11600-017-0071-4

DeGaetano, A.T. and Castellano, C.M., 2017. Future projections of extreme precipitation intensity-duration-frequency curves for climate adaptation planning in New York State. *Climate Services*, 5, 23–35. doi:10.1016/j.cliser.2017.03.003

Dehling, H., *et al.*, 2017. Testing for changes in Kendall's tau. *Econometric Theory*, 33 (6), 1352–1386. doi:10.1017/S026646661600044X.

Dehling, H., Rooch, A., and Taqqu, M.S., 2013. Non-parametric change-point tests for long-range dependent data. *Scandinavian Journal of Statistics*, 40 (1), 153–173. doi:10.1111/j.1467-9469.2012.00799.x.

Deidda, R., 2000. Rainfall downscaling in a space-time multifractal framework. *Water Resources Research*, 36 (7), 1779–1794. doi:10.1029/2000WR900038.

Deidda, R., 2010. A multiple threshold method for fitting the generalized Pareto distribution to rainfall time series. *Hydrology and Earth System Sciences*, 14 (12), 2559–2575. doi:10.5194/hess-14-2559-2010.

Deidda, R., Hellies, M., and Langousis, A., 2021. A critical analysis of the shortcomings in spatial frequency analysis of rainfall extremes based on homogeneous regions and a comparison with a hierarchical boundaryless approach. *Stochastic Environmental Research and Risk Assessment*, 35 (12), 2605–2628. doi:10.1007/s00477-021-02008-x.

de Lavenne, A., *et al.*, 2016. Transferring measured discharge time series: large-scale comparison of top-kriging to geomorphology-based inverse modeling. *Water Resources Research*, 52 (7), 5555–5576. doi:10.1002/2016WR018716.

Desai, S. and Ouarda, T.B., 2021. Regional hydrological frequency analysis at ungauged sites with random forest regression. *Journal of Hydrology*, 594, 125861. doi:10.1016/j.jhydrol.2020.125861

de Vos, L.W., *et al.*, 2019. Rainfall estimation accuracy of a nationwide instantaneously sampling commercial microwave link network: error dependency on known characteristics. *Journal of Atmospheric and Oceanic Technology*, 36 (7), 1267–1283. doi:10.1175/JTECH-D-18-0197.1.

Di Baldassarre, G., Castellarin, A., and Brath, A., 2006. Relationships between statistics of rainfall extremes and mean annual precipitation: an application for design-storm estimation in northern central Italy. *Hydrology and Earth System Sciences*, 10 (4), 589–601. doi:10.5194/hess-10-589-2006.

Douglas, E.M., Vogel, R.M., and Kroll, C.N., 2000. Trends in floods and low flows in the United States: impact of spatial correlation. *Journal of Hydrology*, 240 (1–2), 90–105. doi:10.1016/S0022-1694(00)00336-X.

Dung, N.V., *et al.*, 2015. Handling uncertainty in bivariate quantile estimation–An application to flood hazard analysis in the Mekong Delta. *Journal of Hydrology*, 527, 704–717. doi:10.1016/j.jhydrol.2015.05.033

Durocher, M., Chebana, F., and Ouarda, T.B., 2016. On the prediction of extreme flood quantiles at ungauged locations with spatial copula. *Journal of Hydrology*, 533, 523–532. doi:10.1016/j.jhydrol.2015.12.029

Eagleson, P.S., 1972. Dynamics flood Frequency. *Water Resources Research*, 8, 878–898. doi:10.1029/WR008i004p00878

Easterling, D.R., *et al.*, 2016. Detection and attribution of climate extremes in the observed record. *Weather and Climate Extremes*, 11, 17–27. doi:10.1016/j.wace.2016.01.001

Ekström, M., *et al.*, 2005. New estimates of future changes in extreme rainfall across the UK using regional climate model integrations. 2. Future estimates and use in impact studies. *Journal of Hydrology*, 300 (1–4), 234–251. doi:10.1016/j.jhydrol.2004.06.019.

El Adlouni, S., *et al.*, 2007. Generalized maximum likelihood estimators for the nonstationary generalized extreme value model. *Water Resources Research*, 43 (3). doi:10.1029/2005WR004545.

El Adlouni, S., Bobée, B., and Ouarda, T.B., 2008. On the tails of extreme event distributions in hydrology. *Journal of Hydrology*, 355 (1–4), 16–33. doi:10.1016/j.jhydrol.2008.02.011.

El Adlouni, S. and Ouarda, T.B.M.J., 2010. Frequency analysis of extreme rainfall events. *Rainfall: State of the Science*, 191, 171–188.

Emmanouil, S., *et al.*, 2020. Quantitative assessment of annual maxima, peaks-over-threshold and multifractal parametric approaches in estimating intensity-duration-frequency curves from short rainfall records. *Journal of Hydrology*, 589, 125151. doi:10.1016/j.jhydrol.2020.125151

Emmanouil, S., *et al.*, 2022. The spatiotemporal evolution of rainfall extremes in a changing climate: a CONUS-wide assessment based on multifractal scaling arguments. *Earth's Future*, 10, e2021EF002539. doi:10.1029/2021EF002539

Emmanouil, S., *et al.*, 2023. Exploring the future of rainfall extremes over CONUS: the effects of high emission climate change trajectories on the intensity and frequency of rare precipitation events. *Earth's Future*, 11, e2022EF003039. doi:10.1029/2022EF003039

Farmer, W.H., 2016. Ordinary kriging as a tool to estimate historical daily streamflow records. *Hydrology and Earth System Sciences*, 20 (7), 2721–2735. doi:10.5194/hess-20-2721-2016.

Farmer, W.H., Over, T.M., and Vogel, R.M., 2015. Multiple regression and inverse moments improve the characterization of the spatial scaling behavior of daily streamflows in the Southeast United States. *Water Resources Research*, 51 (3), 1775–1796. doi:10.1002/2014WR015924.

Farmer, W.H. and Vogel, R.M., 2016. On the deterministic and stochastic use of hydrologic models. *Water Resources Research*, 52 (7), 5619–5633. doi:10.1002/2016WR019129.

Farris, S., *et al.*, 2021. On the role ofserial correlation and field significance in detecting changes in extreme precipitation frequency. *Water Resources Research*, 57, e2021WR030172. doi:10.1029/2021WR030172

Faulkner, D., *et al.*, 2020. Can we still predict the future from the past? Implementing non-stationary flood frequency analysis in the UK. *Journal of Flood Risk Management*, 13 (1), e12582. doi:10.1111/jfr3.12582.

Favre, A.C., *et al.*, 2004. Multivariate hydrological frequency analysis using copulas. *Water Resources Research*, 40 (1). doi:10.1029/2003WR002456.

Feng, Y., *et al.*, 2020. Nonstationary flood coincidence risk analysis using time-varying copula functions. *Scientific Reports*, 10 (1), 1–12. doi:10.1038/s41598-019-56847-4.

Fernandez, B. and Salas, J.D., 1986. Periodic gamma autoregressive processes for operational hydrology. *Water Resources Research*, 22 (10), 1385–1396. doi:10.1029/WR022i010p01385.

Fernández, B. and Salas, J.D., 1999. Return period and risk of hydrologic events. I: mathematical formulation. *Journal of Hydrological Engineering*, 4 (4), 297–307. doi:10.1061/(ASCE)1084-0699(1999)4:4(297).

Fernandez, W., Vogel, R.M., and Sankarasubramanian, A., 2000. Regional calibration of a watershed model. *Hydrological Sciences Journal*, 45 (5), 689–707. doi:10.1080/02626660009492371.

Fischer, E.M. and Knutti, R., 2016. Observed heavy precipitation increase confirms theory and early models. *Nature Climate Change*, 6 (11), 986–991. doi:10.1038/nclimate3110.

Fischer, S., 2018. A seasonal mixed-POT model to estimate high flood quantiles from different event types and seasons. *Journal of Applied Statistics*, 45 (15), 2831–2847. doi:10.1080/02664763.2018.1441385.

Fischer, S., *et al.*, 2023. Detecting flood-type-specific flood-rich and flood-poor periods in peaks-over-threshold series with application to Bavaria (Germany). *Stoch Environ Res Risk Assess*, 37, 1395–1413. doi:10.1007/s00477-022-02350-8

Fischer, S., Schumann, A., and Bühler, P., 2019. Timescale-based flood typing to estimate temporal changes in flood frequencies. *Hydrological Sciences Journal*, 64 (15), 1867–1892. doi:10.1080/02626667.2019.1679376.

Fischer, S. and Schumann, A.H., 2021. Regionalisation of flood frequencies based on flood type-specific mixture distributions. *Journal of Hydrology X*, 13, 100107. doi:10.1016/j.hydroa.2021.100107

Fitzner, D. and Sester, M., 2015. Estimation of precipitation fields from 1-minute rain gauge time series–comparison of spatial and spatio-temporal interpolation methods. *International Journal of Geographical Information Science*, 29 (9), 1668–1693. doi:10.1080/13658816.2015.1040022.

Flores, C., 2004. Multiplicative cascade models for rain in hydro-meteorological disasters risk management. *ASTIN-Kolloquium, Bergen, Norway*. 6–9.

Frahm, G., 2006. On the extremal dependence coefficient of multivariate distributions. *Statistics and Probability Letters*, 76 (14), 1470–1481. doi:10.1016/j.spl.2006.03.006.

Francés, F., Salas, J.D., and Boes, D., 1994. Flood frequency analysis with systematic and historical or paleoflood data based on the two-parameter GEV models. *Water Resources Research*, 30 (6), 1653–1664. doi:10.1029/94WR00154.

Fu, G., Chiew, F.H., and Shi, X., 2018. Generation of multi-site stochastic daily rainfall with four weather generators: a case study of gloucester catchment in Australia. *Theoretical and Applied Climatology*, 134 (3–4), 1027–1046. doi:10.1007/s00704-017-2306-3.

Gaume, E., Mouhous, N., and Andrieu, H., 2007. Rainfall stochastic disaggregation models: calibration and validation of a multiplicative cascade model. *Advances in Water Resources*, 30 (5), 1301–1319. doi:10.1016/j.advwatres.2006.11.007.

Gedikli, A., Aksoy, H., and Unal, N.E., 2010. AUG-Segmenter: a user-friendly tool for segmentation of long time series. *Journal of Hydroinformatics*, 12 (3), 318–328. doi:10.2166/hydro.2009.084.

Genest, C., *et al.*, 2007. Metaelliptical copulas and their use in frequency analysis of multivariate hydrological data. *Water Resources Research*, 43 (9). doi:10.1029/2006WR005275.

Genest, C. and Chebana, F., 2017. Copula modeling in hydrologic frequency analysis. Chapter 30. *Handbook of Applied Hydrology. McGraw-Hill Education, New York*, 30–31.

Genest, C. and Favre, A.C., 2007. Everything you always wanted to know about copula modeling but were afraid to ask. *Journal of Hydrologic Engineering*, 12 (4), 347–368. doi:10.1061/(ASCE)1084-0699(2007)12:4(347).

Genest, C. and Nešlehová, J., 2012. Copulas and copula models. *In*: *Encyclopedia of Environmetrics*. 2nd ed. Vol. 2, Chichester: Wiley, 541–553.

Genest, C. and Rémillard, B., 2008. Validity of the parametric bootstrap for goodness-of-fit testing in semiparametric models. *Annales de l'IHP Probabilités et statistiques*, 44 (6), 1096–1127.

Genest, C., Rémillard, B., and Beaudoin, D., 2009. Goodness-of-fit tests for copulas: a review and a power study. *Insurance Mathematics & Economics*, 44 (2), 199–213.

Gerstenberger, C., Vogel, D., and Wendler, M., 2020. Tests for scale changes based on pairwise differences. *Journal of the American*

Statistical Association, 115 (531), 1336–1348. doi:10.1080/01621459.2019.1629938.
Gilleland, E. and Katz, R.W., 2016. extRemes 2.0: an extreme value analysis package in R. *Journal of Statistical Software*, 72 (8), 1–39. doi:10.18637/jss.v072.i08.
Goovaerts, P., 2000. Geostatistical approaches for incorporating elevation into the spatial interpolation of rainfall. *Journal of Hydrology*, 228 (1–2), 113–129. doi:10.1016/S0022-1694(00)00144-X.
Goudenhoofdt, E. and Delobbe, L., 2009. Evaluation of radar-gauge merging methods for quantitative precipitation estimates. *Hydrology and Earth System Sciences*, 13 (2), 195–203. doi:10.5194/hess-13-195-2009 .
Gräler, B., *et al*., 2013. Multivariate return periods in hydrology: a critical and practical review focusing on synthetic design hydrograph estimation. *Hydrology and Earth System Sciences*, 17 (4), 1281–1296. doi:10.5194/hess-17-1281-2013.
Griffis, V.W. and Stedinger, J.R., 2007. Log-Pearson type 3 distribution and its application in flood frequency analysis. I: distribution characteristics. *Journal of Hydrologic Engineering*, 12 (5), 482–491. doi:10.1061/(ASCE)1084-0699(2007)12:5(482).
Grimaldi, S., *et al*., 2022. Continuous hydrologic modelling for small and ungauged basins: a comparison of eight rainfall models for sub-daily runoff simulations. *Journal of Hydrology*, 610, 127866.
Grimaldi, S. and Serinaldi, F., 2006a. Design hyetograph analysis with 3-copula function. *Hydrological Sciences Journal*, 51 (2), 223–238. doi:10.1623/hysj.51.2.223.
Grimaldi, S. and Serinaldi, F., 2006b. Asymmetric copula in multivariate flood frequency analysis. *Advances in Water Resources*, 29 (8), 1155–1167. doi:10.1016/j.advwatres.2005.09.005.
Grønneberg, S. and Hjort, N.L., 2014. The copula information criteria. *Scandinavian Journal of Statistics*, 41 (2), 436–459. doi:10.1111/sjos.12042.
Gruber, A.M. and Stedinger, J.R., 13–16 May, 2008. Models of LP3 regional skew, data selection, and Bayesian GLS regression. *EWRI World Water & Environmental Resources Congress*. Honolulu, HI:, American Society of Civil Engineers, 1–10.
Grundmann, J., Hörning, S., and Bárdossy, A., 2019. Stochastic reconstruction of spatio-temporal rainfall patterns by inverse hydrologic modelling. *Hydrology and Earth System Sciences*, 23 (1), 225–237. doi:10.5194/hess-23-225-2019.
Gumbel, E.J., 1958. *Statistics of extremes*. New York: Columbia University Press.
Guo, Y., *et al*., 2021. Regionalization of hydrological modeling for predicting streamflow in ungauged catchments: a comprehensive review. *Wiley Interdisciplinary Reviews: Water*, 8 (1), e1487. doi:10.1002/wat2.1487.
Haan, C.T., Allen, D.M., and Street, J.O., 1976. A markov chain model of daily rainfall. *Water Resources Research*, 12 (3), 443–449. doi:10.1029/WR012i003p00443.
Haberlandt, U., 2007. Geostatistical interpolation of hourly precipitation from rain gauges and radar for a large-scale extreme rainfall event. *Journal of Hydrology*, 332 (1–2), 144–157. doi:10.1016/j.jhydrol.2006.06.028.
Haberlandt, U. and Sester, M., 2010. Areal rainfall estimation using moving cars as rain gauges–a modelling study. *Hydrology and Earth System Sciences*, 14 (7), 1139–1151. doi:10.5194/hess-14-1139-2010.
Haese, B., *et al*., 2017. Stochastic reconstruction and interpolation of precipitation fields using combined information of commercial microwave links and rain gauges. *Water Resources Research*, 53 (12), 10740–10756. doi:10.1002/2017WR021015.
Hald, A., 2005. *A history of probability and statistics and their applications before 1750*. New York, US: John Wiley & Sons.
Han, J.C., *et al*., 2020. Risk assessment through multivariate analysis on the magnitude and occurrence date of daily storm events in the Shenzhen bay area. *Stochastic Environmental Research and Risk Assessment*, 34 (5), 669–689. doi:10.1007/s00477-020-01793-1.
Hao, Z. and AghaKouchak, A., 2014. A nonparametric multivariate multi-index drought monitoring framework. *Journal of Hydrometeorology*, 15 (1), 89–101. doi:10.1175/JHM-D-12-0160.1.
Hao, Z. and Singh, V.P., 2013. Modeling multisite streamflow dependence with maximum entropy copula. *Water Resources Research*, 49 (10), 7139–7143. doi:10.1002/wrcr.20523.
Hao, Z. and Singh, V.P., 2016. Review of dependence modeling in hydrology and water resources. *Progress in Physical Geography*, 40 (4), 549–578. doi:10.1177/0309133316632460.
Hayhoe, H.N., 2000. Improvements of stochastic weather data generators for diverse climates. *Climate Research*, 14 (2), 75–87. doi:10.3354/cr014075.
Hecht, J.S. and Vogel, R.M., 2020. Updating urban design floods for changes in central tendency and variability using regression. *Advances in Water Resources*, 136, 103484. doi:10.1016/j.advwatres.2019.103484
Helsel, D.R., *et al*., 2020. Statistical methods in water resources: US geological survey techniques and methods. *book*, 4, 458.
Helsel, D.R. and Hirsch, R.M., 1992. *Statistical methods in water resources*. Amsterdam, Netherlands: Elsevier.
Hirsch, R.M., 1979. Synthetic hydrology and water supply reliability. *Water Resources Research*, 15 (6), 1603–1615. doi:10.1029/WR015i006p01603.
Hirsch, R.M., 1982. A comparison of four streamflow record extension techniques. *Water Resources Research*, 18 (4), 1081–1088. doi:10.1029/WR018i004p01081.
Hirschboeck, K.K., 1987. Hydroclimatically-defined mixed distributions in partial duration flood series. *In*: *Hydrologic frequency modeling*. Dordrecht, Netherlands: Springer, 199–212.
Hodgkins, G.A., *et al*., 2017. Climate-driven variability in the occurrence of major floods across North America and Europe. *Journal of Hydrology*, 552, 704–717. doi:10.1016/j.jhydrol.2017.07.027
Hofert, M., Mächler, M., and Mcneil, A.J., 2012. Likelihood inference for archimedean copulas in high dimensions under known margins. *Journal of Multivariate Analysis*, 110, 133–150. doi:10.1016/j.jmva.2012.02.019
Hörning, S. and Bárdossy, A., 2018. Phase annealing for the conditional simulation of spatial random fields. *Computers & Geosciences*, 112, 101–111. doi:10.1016/j.cageo.2017.12.008
Hosking, J.R., 1990. L-moments: analysis and estimation of distributions using linear combinations of order statistics. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 52 (1), 105–124. doi:10.1111/j.2517-6161.1990.tb01775.x.
Hosking, J.R.M. and Wallis, J.R. (1997). *Regional frequency analysis* (p. 240).
Hosking, J.R. and Wallis, J.R., 1987. Parameter and quantile estimation for the generalized Pareto distribution. *Technometrics*, 29 (3), 339–349. doi:10.1080/00401706.1987.10488243.
Hrachowitz, M., *et al*., 2013. A decade of Predictions in Ungauged Basins (PUB)—a review. *Hydrological Sciences Journal*, 58 (6), 1198–1255. doi:10.1080/02626667.2013.803183.
Huang, Y., *et al*., 2009. Analysis of daily river flow fluctuations using empirical mode decomposition and arbitrary order Hilbert spectral analysis. *Journal of Hydrology*, 373 (1–2), 103–111. doi:10.1016/j.jhydrol.2009.04.015.
Hubert, P., 2000. The segmentation procedure as a tool for discrete modeling of hydrometeorological regimes. *Stochastic Environmental Research and Risk Assessment*, 14, 297–304. doi:10.1007/PL00013450
Hughes, J.P. and Guttorp, P., 1994. Incorporating spatial dependence and atmospheric data in a model of precipitation. *Journal of Applied Meteorology and Climatology*, 33 (12), 1503–1515. doi:10.1175/1520-0450(1994)033<1503:ISDAAD>2.0.CO;2.
Hurst, H.E., 1951. Long-term storage capacity of reservoirs. *Transactions of the American Society of Civil Engineers*, 116 (1), 770–799. doi:10.1061/TACEAT.0006518.
Hutchinson, M.F., 1998a. Interpolation of rainfall data with thin plate smoothing splines. Part I: two dimensional smoothing of data with short range correlation. *Journal of Geographic Information and Decision Analysis*, 2 (2), 139–151.
Hutchinson, M.F., 1998b. Interpolation of rainfall data with thin plate smoothing splines: II. Analysis of topographic dependence. *Journal of Geographic Information and Decision Analysis*, 2 (2), 168–185.

Iliopoulou, T. and Koutsoyiannis, D., 2020. Projecting the future of rainfall extremes: better classic than trendy. *Journal of Hydrology*, 588, 125005. doi:10.1016/j.jhydrol.2020.125005

Inclan, C. and Tiao, G.C., 1994. Use of cumulative sums of squares for retrospective detection of changes of variance. *Journal of the American Statistical Association*, 89 (427), 913–923.

Institute of Hydrology. 1999. *Flood Estimation Handbook*. Wallingford, UK: Centre for Ecology & Hydrology.

Izady, A., *et al.*, 2012. Application of "panel-data" modeling to predict groundwater levels in the Neishaboor plain, Iran. *Hydrogeology Journal*, 20 (3), 435. doi:10.1007/s10040-011-0814-2.

Jiang, C., *et al.*, 2015. Bivariate frequency analysis of nonstationary low-flow series based on the time-varying copula. *Hydrological Processes*, 29 (6), 1521–1534. doi:10.1002/hyp.10288.

Joe, H., 2014. *Dependence modeling with copulas*. Boca Raton, Florida, US: Chapman and Hall/CRC.

Jothityangkoon, C., Sivapalan, M., and Viney, N.R., 2000. Tests of a space-time model of daily rainfall in southwestern Australia based on nonhomogeneous random cascades. *Water Resources Research*, 36 (1), 267–284. doi:10.1029/1999WR900253.

Kalai, C., *et al.*, 2020. Comparison of nonstationary regional flood frequency analysis techniques based on the index-flood approach. *Journal of Hydrologic Engineering*, 25 (7), 06020003. doi:10.1061/(ASCE)HE.1943-5584.0001939.

Kantelhardt, J.W., *et al.*, 2003. Multifractality of river runoff and precipitation: comparison of fluctuation analysis and wavelet methods. *Physica A Statistical Mechanics & Its Applications*, 330 (1–2), 240–245. doi:10.1016/j.physa.2003.08.019.

Kao, S.C. and Govindaraju, R.S., 2007. A bivariate frequency analysis of extreme rainfall with implications for design. *Journal of Geophysical Research: Atmospheres*, 112 (D13). doi:10.1029/2007JD008522.

Kao, S.C. and Govindaraju, R.S., 2008. Trivariate statistical analysis of extreme rainfall events via the plackett family of copulas. *Water Resources Research*, 44 (2). doi:10.1029/2007WR006261.

Karahacane, H., *et al.*, 2020. Complete multivariate flood frequency analysis, applied to northern Algeria. *Journal of Flood Risk Management*, 13 (4), e12619. doi:10.1111/jfr3.12619.

Katz, R.W., 2013. Statistical methods for nonstationary extremes. *In*: *Extremes in a changing climate*. Dordrecht, Netherlands: Springer, 15–37.

Katz, R.W., Parlange, M.B., and Naveau, P., 2002. Statistics of extremes in hydrology. *Advances in Water Resources*, 25 (8–12), 1287–1304. doi:10.1016/S0309-1708(02)00056-8.

Kauffeldt, A., *et al.*, 2013. Disinformative data in large-scale hydrological modelling. *Hydrology and Earth System Sciences*, 17 (7), 2845–2857. doi:10.5194/hess-17-2845-2013.

Kehagias, A., 2004. A hidden Markov model segmentation procedure for hydrological and environmental time series. *Stochastic Environmental Research and Risk Assessment*, 18, 117–130. doi:10.1007/s00477-003-0145-5

Kehagias, A., Nidelkou, E., and Petridis, V., 2006. A dynamic programming segmentation procedure for hydrological and environmental time series. *Stochastic Environmental Research and Risk Assessment*, 20, 77–94. doi:10.1007/s00477-005-0013-6

Kelley, D., 1994. *Introduction to probability*. London: Macmillan Publishing Company.

Keylock, C.J., 2012. A resampling method for generating synthetic hydrological time series with preservation of cross-correlative structure and higher-order properties. *Water Resources Research*, 48 (12). doi:10.1029/2012WR011923.

Kilsby, C.G., *et al.*, 2007. A daily weather generator for use in climate change studies. *Environmental Modelling and Software*, 22 (12), 1705–1719. doi:10.1016/j.envsoft.2007.02.005.

Kim, D., *et al.*, 2016. Spatial composition of AMSR2 soil moisture products by conditional merging technique with ground soil moisture data. *Stochastic Environmental Research and Risk Assessment*, 30, 2109–2126. doi:10.1007/s00477-016-1300-0

Kim, D. and Onof, C., 2020. A stochastic rainfall model that can reproduce important rainfall properties across the timescales from several minutes to a decade. *Journal of Hydrology*, 589, 125150. doi:10.1016/j.jhydrol.2020.125150

Kim, G., Silvapulle, M.J., and Silvapulle, P., 2007. Comparison of semiparametric and parametric methods for estimating copulas. *Computational Statistics and Data Analysis*, 51 (6), 2836–2850. doi:10.1016/j.csda.2006.10.009.

Kjeldsen, T.R., Ahn, H., and Prosdocimi, I., 2017. On the use of a four-parameter kappa distribution in regional frequency analysis. *Hydrological Sciences Journal*, 62 (9), 1354–1363. doi:10.1080/02626667.2017.1335400.

Kjeldsen, T.R. and Prosdocimi, I., 2021. Assessment of trends in hydrological extremes using regional magnification factors. *Advances in Water Resources*, 149, 103852. doi:10.1016/j.advwatres.2021.103852

Klemeš, V., 1974. The Hurst phenomenon: a puzzle? *Water Resources Research*, 10 (4), 675–688. doi:10.1029/WR010i004p00675.

Klemeš, V., 2000a. Tall tales about tails of hydrological distributions. I. *Journal of Hydrologic Engineering*, 5 (3), 227–231. doi:10.1061/(ASCE)1084-0699(2000)5:3(227).

Klemeš, V., 2000b. Tall tales about tails of hydrological distributions. II. *Journal of Hydrologic Engineering*, 5 (3), 232–239. doi:10.1061/(ASCE)1084-0699(2000)5:3(232).

Kochanek, K., *et al.*, 2020. The bias of the maximum likelihood estimates of flood quantiles based solely on the largest historical records. *Journal of Hydrology*, 584, 124740. doi:10.1016/j.jhydrol.2020.124740

Kochanek, K., Strupczewski, W.G., and Bogdanowicz, E., 2012. On seasonal approach to flood frequency modelling. Part II: flood frequency analysis of Polish rivers. *Hydrological Processes*, 26 (5), 717–730. doi:10.1002/hyp.8178.

Kochel, R.C. and Baker, V.R., 1982. Paleoflood hydrology. *Science*, 215 (4531), 353–361. doi:10.1126/science.215.4531.353.

Kojadinovic, I. and Yan, J., 2010. Comparison of three semiparametric methods for estimating dependence parameters in copula models. *Insurance Mathematics & Economics*, 47 (1), 52–63.

Kojadinovic, I., Yan, J., and Holmes, M., 2011. Fast large-sample goodness-of-fit tests for copulas. *Statistica Sinica*, 21, 841–871. doi:10.5705/ss.2011.037a

Kossieris, P., *et al.*, 2018. A rainfall disaggregation scheme for sub-hourly time scales: coupling a bartlett-lewis based model with adjusting procedures. *Journal of Hydrology*, 556, 980–992. doi:10.1016/j.jhydrol.2016.07.015

Kotz, S. and Nadarajah, S., 2000. *Extreme value distributions*. London: Imperial College Press.

Koutsoyiannis, D., 2003. Climate change, the hurst phenomenon, and hydrological statistics. *Hydrological Sciences Journal*, 48 (1), 3–24. doi:10.1623/hysj.48.1.3.43481.

Koutsoyiannis, D., 2016. Generic and parsimonious stochastic modelling for hydrology and beyond. *Hydrological Sciences Journal*, 61 (2), 225–244. doi:10.1080/02626667.2015.1016950.

Koutsoyiannis, D., 2020. Revisiting the global hydrological cycle: is it intensifying? *Hydrology and Earth System Sciences*, 24 (8), 3899–3932. doi:10.5194/hess-24-3899-2020.

Koutsoyiannis, D., 2021. *Stochastics of hydroclimatic extremes - a cool look at risk*. Athens: Kallipos Open Academic Editions, 333.

Koutsoyiannis, D. and Langousis, A., 2011. Precipitation. *In*: P. Wilderer (in chief) and S. Uhlenbrook, eds. *Treaties on water sciences: hydrology*. Oxford: Academic Press, Vol. 2, 27–78.

Koutsoyiannis, D. and Montanari, A., 2015. Negligent killing of scientific concepts: the stationarity case. *Hydrological Sciences Journal*, 60 (7–8), 1174–1183. doi:10.1080/02626667.2014.959959.

Krajewski, W.F. and Smith, J.A., 2002. Radar hydrology: rainfall estimation. *Advances in Water Resources*, 25 (8–12), 1387–1394. doi:10.1016/S0309-1708(02)00062-3.

Kumar, D.N., Lall, U., and Petersen, M.R., 2000. Multisite disaggregation of monthly to daily streamflow. *Water Resources Research*, 36 (7), 1823–1833. doi:10.1029/2000WR900049.

Kundzewicz, Z.W., *et al.*, 2005. Trend detection in river flow series: 1. Annual maximum flow/Détection de tendance dans des séries de débit fluvial: 1. Débit maximum annuel. *Hydrological Sciences Journal*, 50 (5). doi:10.1623/hysj.2005.50.5.797.

Kundzewicz, Z.W. and Robson, A.J., 2004. Change detection in hydrological records—a review of the methodology/revue méthodologique de la détection de changements dans les chroniques hydrologiques. *Hydrological Sciences Journal*, 49 (1), 7–19. doi:10.1623/hysj.49.1.7.53993.

Kwon, -H.-H. and Lall, U., 2016. A copula-based nonstationary frequency analysis for the 2012–2015 drought in California. *Water Resources Research*, 52, 5662–5675. doi:10.1002/2016WR018959

Laaha, G., *et al*., 2013. Spatial prediction of stream temperatures using top-kriging with an external drift. *Environmental Modeling and Assessment*, 18, 671–683. doi:10.1007/s10666-013-9373-3

Laaha, G., Skøien, J.O., and Blöschl, G., 2014. Spatial prediction on river networks: comparison of top-kriging with regional regression. *Hydrological Processes*, 28 (2), 315–324. doi:10.1002/hyp.9578.

Laio, F., 2004. Cramer-von mises and anderson-darling goodness of fit tests for extreme value distributions with unknown parameters. *Water Resources Research*, 40, W09308. doi:10.1029/2004WR003204

Laio, F., *et al*., 2011. Spatially smooth regional estimation of the flood frequency curve (with uncertainty). *Journal of Hydrology*, 408 (1–2), 67–77. doi:10.1016/j.jhydrol.2011.07.022.

Laio, F., Di Baldassarre, G., and Montanari, A., 2009. Model selection techniques for the frequency analysis of hydrological extremes. *Water Resources Research*, 45 (W07416). doi:10.1029/2007WR006666.

Lall, U., 1995. Recent advances in nonparametric function estimation: hydrologic applications. *Reviews of Geophysics*, 33 (S2), 1093–1102. doi:10.1029/95RG00343.

Lall, U. and Sharma, A., 1996. A nearest neighbor bootstrap for resampling hydrologic time series. *Water Resources Research*, 32 (3), 679–693. doi:10.1029/95WR02966.

Lamontagne, J.R., *et al*., 2016. Robust flood frequency analysis: performance of EMA with multiple grubbs-beck outlier tests. *Water Resources Research*, 52 (4), 3068–3084. doi:10.1002/2015WR018093.

Lamontagne, J.R., Barber, C.A., and Vogel, R.M., 2020. Improved estimators of model performance efficiency for skewed hydrologic data. *Water Resources Research*, 56 (9), e2020WR027101. doi:10.1029/2020WR027101.

Langousis, A., *et al*., 2009. Multifractal rainfall extremes: theoretical analysis and practical estimation. *Chaos, Solitons & Fractals*, 39, 1182–1194. doi:10.1016/j.chaos.2007.06.004

Langousis, A., *et al*., 2016a. Threshold detection for the generalized pareto distribution: review of representative methods and application to the NOAA NCDC daily rainfall database. *Water Resources Research*, 52 (4), 2659–2681. doi:10.1002/2015WR018502.

Langousis, A., *et al*., 2016b. Assessing the relative effectiveness of statistical downscaling and distribution mapping in reproducing rainfall statistics based on climate model results. *Water Resources Research*, 52, 471–494. doi:10.1002/2015WR017556

Langousis, A., Carsteanu, A.A., and Deidda, R., 2013. A simple approximation to multifractal rainfall maxima using a generalized extreme value distribution model. *Stochastic Environmental Research and Risk Assessment*, 27 (6), 1525–1531. doi:10.1007/s00477-013-0687-0.

Langousis, A. and Kaleris, V., 2014. Statistical framework to simulate daily rainfall series conditional on upper-air predictor variables. *Water Resources Research*, 50 (5), 3907–3932. doi:10.1002/2013WR014936.

Langousis, A. and Koutsoyiannis, D., 2006. A stochastic methodology for generation of seasonal time series reproducing over-year scaling behavior. *Journal of Hydrology*, 322 (1–4), 138–154. doi:10.1016/j.jhydrol.2005.02.037.

Langousis, A. and Veneziano, D., 2007. Intensity-duration-frequency curves from scaling representations of rainfall. *Water Resources Research*, 43. doi:10.1029/2006WR005245.

Latif, S., *et al*., 2023. Copula-based joint modelling of extreme river temperature and low flow characteristics in the risk assessment of aquatic life. *Weather and Climate Extremes*, 41, 100586. doi:10.1016/j.wace.2023.100586

Latif, S., *et al*., 2024. A new nonparametric copula framework for the joint analysis of river water temperature and low flow characteristics for aquatic habitat risk assessment. *Journal of Hydrology*, 634, 131079. doi:10.1016/j.jhydrol.2024.131079

Lawrance, A.J. and Kottegoda, N.T., 1977. Stochastic modelling of riverflow time series. *Journal of the Royal Statistical Society: Series A (General)*, 140 (1), 1–31. doi:10.2307/2344516.

Lebrenz, H. and Bárdossy, A., 2019. Geostatistical interpolation by quantile kriging. *Hydrology and Earth System Sciences*, 23 (3), 1633–1648. doi:10.5194/hess-23-1633-2019.

Leclerc, M. and Ouarda, T.B., 2007. Non-stationary regional flood frequency analysis at ungauged sites. *Journal of Hydrology*, 343 (3–4), 254–265. doi:10.1016/j.jhydrol.2007.06.021.

Lee, T., Modarres, R., and Ouarda, T.B.M.J., 2013. Data-based analysis of bivariate copula tail dependence for drought duration and severity. *Hydrological Processes*, 27 (10), 1454–1463. doi:10.1002/hyp.9233.

Lee, T. and Ouarda, T.B., 2012. Stochastic simulation of nonstationary oscillation hydroclimatic processes using empirical mode decomposition. *Water Resources Research*, 48 (2). doi:10.1029/2011WR010660.

Lee, T., Salas, J.D., and Prairie, J., 2010. An enhanced nonparametric streamflow disaggregation model with genetic algorithm. *Water Resources Research*, 46 (8). doi:10.1029/2009WR007761.

Leese, M.N., 1973. Use of censored data in the estimation of gumbel distribution parameters for annual maximum flood series. *Water Resources Research*, 9, 1534–1542. doi:10.1029/WR009i006p01534

Lekina, A., Chebana, F., and Ouarda, T.B., 2015. On the tail dependence in bivariate hydrological frequency analysis. *Dependence Modeling*, 3 (1). doi:10.1515/demo-2015-0015.

Li, J., *et al*., 2017. A comparison of methods for estimating climate change impact on design rainfall using a high-resolution RCM. *Journal of Hydrology*, 547, 413–427. doi:10.1016/j.jhydrol.2017.02.019

Libertino, A., *et al*., 2018. Regional-scale analysis of extreme precipitation from short and fragmented records. *Advances in Water Resources*, 112, 147–159. doi:10.1016/j.advwatres.2017.12.015

Libiseller, C. and Grimvall, A., 2002. Performance of partial mann–kendall tests for trend detection in the presence of covariates. *Environmetrics: The Official Journal of the International Environmetrics Society*, 13 (1), 71–84. doi:10.1002/env.507.

Lilienthal, J., Fried, R., and Schumann, A., 2018. Homogeneity testing for skewed and cross-correlated data in regional flood frequency analysis. *Journal of Hydrology*, 556, 557–571. doi:10.1016/j.jhydrol.2017.10.056

Lins, H.F. and Cohn, T.A., 2011. Stationarity: wanted dead or alive? 1. *JAWRA Journal of the American Water Resources Association*, 47 (3), 475–480. doi:10.1111/j.1752-1688.2011.00542.x.

Liu, Y.R., *et al*., 2020. Development of a Bayesian-copula-based frequency analysis method for hydrological risk assessment–The Naryn River in Central Asia. *Journal of Hydrology*, 580, 124349. doi:10.1016/j.jhydrol.2019.124349

López, J. and Francés, F., 2013. Non-stationary flood frequency analysis in continental Spanish rivers, using climate and reservoir indices as external covariates. *Hydrology and Earth System Sciences*, 17, 3189–3203. doi:10.5194/hess-17-3189-2013

Luke, A., *et al*., 2017. Predicting nonstationary flood frequencies: evidence supports an updated stationarity thesis in the United States. *Water Resources Research*, 53, 5469–5494. doi:10.1002/2016WR019676

Lun, D., *et al*., 2020. Detecting flood-rich and flood-poor periods in annual peak discharges across Europe. *Water Resources Research*, 56 (7), e2019WR026575. doi:10.1029/2019WR026575.

Lun, D., *et al*., 2021. Characteristics and process controls of statistical flood moments in Europe–a data-based analysis. *Hydrology and Earth System Sciences*, 25 (10), 5535–5560. doi:10.5194/hess-25-5535-2021.

Ma, M., *et al*., 2013. Multivariate drought characteristics using trivariate gaussian and student copulas. *Hydrological Processes*, 27 (8), 1175–1190. doi:10.1002/hyp.8432.

Machado, M.J., *et al*., 2015. Flood frequency analysis of historical flood data under stationary and non-stationary modelling. *Hydrology and Earth System Sciences*, 19 (6), 2561–2576. doi:10.5194/hess-19-2561-2015.

Madsen, H., *et al*., 2014. Review of trend analysis and climate change projections of extreme precipitation and floods in Europe. *Journal of Hydrology*, 519, 3634–3650. doi:10.1016/j.jhydrol.2014.11.003

Madsen, H., Rasmussen, P.F., and Rosbjerg, D., 1997. Comparison of annual maximum series and partial duration series methods for modeling extreme hydrologic events: 1. At-site modeling. *Water Resources Research*, 33 (4), 747–757. doi:10.1029/96WR03848.

Mallakpour, I. and Villarini, G., 2017. Analysis of changes in the magnitude, frequency, and seasonality of heavy precipitation over the contiguous USA. *Theoretical and Applied Climatology*, 130 (1), 345–363. doi:10.1007/s00704-016-1881-z.

Mandelbrot, B.B. and Mandelbrot, B.B., 1982. *The fractal geometry of nature*. Vol. 1, New York: WH freeman.

Mandelbrot, B.B. and Wallis, J.R., 1968. Noah, Joseph, and operational hydrology. *Water Resources Research*, 4 (5), 909–918. doi:10.1029/WR004i005p00909.

Mangini, W., *et al*., 2018. Detection of trends in magnitude and frequency of flood peaks across Europe. *Hydrological Sciences Journal*, 63 (4), 493–512. doi:10.1080/02626667.2018.1444766.

Marani, M. and Ignaccolo, M., 2015. A metastatistical approach to rainfall extremes. *Advances in Water Resources*, 79, 121–126. doi:10.1016/j.advwatres.2015.03.001

Maraun, D., 2013. Bias correction, quantile mapping, and downscaling: revisiting the inflation issue. *Journal of Climate*, 26 (6), 2137–2143. doi:10.1175/JCLI-D-12-00821.1.

Markonis, Y., *et al*., 2018. Global estimation of long-term persistence in annual river runoff. *Advances in Water Resources*, 113, 1–12. doi:10.1016/j.advwatres.2018.01.003

Marra, F., Amponsah, W., and Papalexiou, S.M., 2023. Non-asymptotic Weibull tails explain the statistics of extreme daily precipitation. *Advances in Water Resources*, 173, 104388. doi:10.1016/j.advwatres.2023.104388

Martins, E.S. and Stedinger, J.R., 2000. Generalized maximum-likelihood generalized extreme-value quantile estimators for hydrologic data. *Water Resources Research*, 36 (3), 737–744. doi:10.1029/1999WR900330.

Mascaro, G., 2018. On the distributions of annual and seasonal daily rainfall extremes in central Arizona and their spatial variability. *Journal of Hydrology*, 559, 266–281. doi:10.1016/j.jhydrol.2018.02.011

Matalas, N.C., 1967. Mathematical assessment of synthetic hydrology. *Water Resources Research*, 3 (4), 937–945. doi:10.1029/WR003i004p00937.

Matalas, N.C., 2012. Comment on the announced death of stationarity. *Journal of Water Resources Planning and Management*, 138 (4), 311–312. doi:10.1061/(ASCE)WR.1943-5452.0000215.

Mazdiyasni, O., *et al*., 2019. Heat wave intensity duration frequency curve: a multivariate approach for hazard and attribution analysis. *Scientific Reports*, 9 (1), 1–8. doi:10.1038/s41598-019-50643-w.

McNeil, A.J., Frey, R., and Embrechts, P., 2015. *Quantitative risk management: concepts, techniques and tools*. 2nd. Princeton, NJ: Princeton University Press.

Medda, S. and Bhar, K.K., 2019. Comparison of single-site and multi-site stochastic models for Streamflow generation. *Applied Water Science*, 9, 1–14. doi:10.1007/s13201-019-0947-3

Mehrotra, R. and Sharma, A., 2007. A semi-parametric model for stochastic generation of multi-site daily rainfall exhibiting low-frequency variability. *Journal of Hydrology*, 335 (1–2), 180–193. doi:10.1016/j.jhydrol.2006.11.011.

Mejia, J.M., Rodriguez-Iturbe, I., and Dawdy, D.R., 1972. Streamflow simulation: 2. The broken line process as a potential model for hydrologic simulation. *Water Resources Research*, 8 (4), 931–941. doi:10.1029/WR008i004p00931.

Menabde, M., *et al*., 1997. Multiscaling properties of rainfall and bounded random cascades. *Water Resources Research*, 33 (12), 2823–2830. doi:10.1029/97WR02006.

Merz, B., *et al*., 2022. Understanding heavy tails of flood peak distributions. *Water Resources Research*, 58 (6), e2021WR030506. doi:10.1029/2021WR030506.

Merz, R. and Blöschl, G., 2004. Regionalisation of catchment model parameters. *Journal of Hydrology*, 287 (1–4), 95–123. doi:10.1016/j.jhydrol.2003.09.028.

Merz, R. and Blöschl, G., 2005. Flood frequency regionalisation—spatial proximity vs. catchment attributes. *Journal of Hydrology*, 302 (1–4), 283–306. doi:10.1016/j.jhydrol.2004.07.018.

Merz, R., Blöschl, G., and Humer, G., 2008. National flood discharge mapping in Austria. *Natural Hazards*, 46, 53–72. doi:10.1007/s11069-007-9181-7

Michiels, F. and De Schepper, A., 2008. A copula test space model how to avoid the wrong copula choice. *Kybernetika*, 44 (6), 864–878.

Milly, P.C., *et al*., 2008. Stationarity is dead: whither water management? *Science*, 319 (5863), 573–574. doi:10.1126/science.1151915.

Milly, P.C., *et al*., 2015. On critiques of "Stationarity is dead: whither water management?". *Water Resources Research*, 51 (9), 7785–7789. doi:10.1002/2015WR017408.

Montanari, A., Rosso, R., and Taqqu, M.S., 1997. Fractionally differenced ARIMA models applied to hydrologic time series: identification, estimation, and simulation. *Water Resources Research*, 33 (5), 1035–1044. doi:10.1029/97WR00043.

Naghettini, M., Ed. 2017. *Fundamentals of statistical hydrology*. Cham: Springer International Publishing.

Nasr, I.B. and Chebana, F., 2019. Multivariate L-moment based tests for copula selection, with hydrometeorological applications. *Journal of Hydrology*, 579, 124151.

Naulet, R., *et al*., 2005. Flood frequency analysis on the Ardèche river using French documentary sources from the last two centuries. *Journal of Hydrology*, 313 (1–2), 58–78. doi:10.1016/j.jhydrol.2005.02.011.

Nelsen, R.B., 2006. *An introduction to copulas*. Dordrecht, Netherlands: Springer.

Nerantzaki, S.D. and Papalexiou, S.M., 2022. Assessing extremes in hydroclimatology: a review on probabilistic methods. *Journal of Hydrology*, 605, 127302. doi:10.1016/j.jhydrol.2021.127302

Nowak, K., *et al*., 2010. A nonparametric stochastic approach for multisite disaggregation of annual to daily streamflow. *Water Resources Research*, 46 (8). doi:10.1029/2009WR008530.

O'Brien, N.L., *et al*., 2020. Trend detection in the presence of positive and negative serial correlation: a comparison of block maxima and peaks-over-threshold data. *Water Resources Research*, 57, e2020WR028886. doi:10.1029/2020WR028886

O'Brien, N.L. and Burn, D.H., 2014. A nonstationary index-flood technique for estimating extreme quantiles for annual maximum streamflow. *Journal of Hydrology*, 519, 2040–2048. doi:10.1016/j.jhydrol.2014.09.041

Ochoa-Rodriguez, S., *et al*., 2019. A review of radar-rain gauge data merging methods and their potential for urban hydrological applications. *Water Resources Research*, 55 (8), 6356–6391. doi:10.1029/2018WR023332.

Oliveira, B. and Maia, R., 2018. Stochastic generation of streamflow time series. *Journal of Hydrologic Engineering*, 23 (10), 04018043. doi:10.1061/(ASCE)HE.1943-5584.0001695.

Oriani, F., *et al*., 2018. Simulating rainfall time-series: how to account for statistical variability at multiple scales? *Stochastic Environmental Research and Risk Assessment*, 32 (2), 321–340. doi:10.1007/s00477-017-1414-z.

O'Shea, D., *et al*., 2023. Improved extreme rainfall frequency analysis using a rwo-step kappa approach. *Water Resources Research*, 59 (4), e2021WR031854. doi:10.1029/2021WR031854.

Ostrowski, P., Falkowski, T., and Kochanek, K., 2023. Reconstructing parameters of the holocene paleofloods in alluvial lowland river valleys–An example from the Bug valley (East Poland). *Journal of Hydrology*, 624, 129930. doi:10.1016/j.jhydrol.2023.129930

Ouali, D., Chebana, F., and Ouarda, T.B., 2017. Fully nonlinear statistical and machine-learning approaches for hydrological frequency estimation at ungauged sites. *Journal of Advances in Modeling Earth Systems*, 9 (2), 1292–1306. doi:10.1002/2016MS000830.

Ouarda, T.B., *et al*., 2001. Regional flood frequency estimation with canonical correlation analysis. *Journal of Hydrology*, 254 (1–4), 157–173. doi:10.1016/S0022-1694(01)00488-7.

Ouarda, T.B. and Charron, C., 2019. Changes in the distribution of hydro-climatic extremes in a non-stationary framework. *Scientific Reports*, 9 (1), 1–8. doi:10.1038/s41598-019-44603-7.

Ouarda, T.B., Charron, C., and St-Hilaire, A., 2020. Uncertainty of stationary and nonstationary models for rainfall frequency analysis. *International Journal of Climatology*, 40 (4), 2373–2392. doi:10.1002/joc.6339.

Ouarda, T.B., Yousef, L.A., and Charron, C., 2019. Non-stationary intensity-duration-frequency curves integrating information concerning teleconnections and climate change. *International Journal of Climatology*, 39 (4), 2306–2323. doi:10.1002/joc.5953.

Papaioannou, G., *et al*., 2016. Joint modelling of flood peaks and volumes: a copula application for the Danube River. *Journal of Hydrology and Hydromechanics*, 64 (4), 382–392. doi:10.1515/johh-2016-0049.

Papalexiou, S.M., 2018. Unified theory for stochastic modelling of hydroclimatic processes: preserving marginal distributions, correlation structures, and intermittency. *Advances in Water Resources*, 115, 234–252. doi:10.1016/j.advwatres.2018.02.013

Papalexiou, S.M., Koutsoyiannis, D., and Makropoulos, C., 2013. How extreme is extreme? An assessment of daily rainfall distribution tails. *Hydrology and Earth System Sciences*, 17 (2), 851–862. doi:10.5194/hess-17-851-2013.

Papalexiou, S.M. and Montanari, A., 2019. Global and regional increase of precipitation extremes under global warming. *Water Resources Research*, 55 (6), 4901–4914. doi:10.1029/2018WR024067.

Papalexiou, S.M. and Serinaldi, F., 2020. Random fields simplified: preserving marginal distributions, correlations, and intermittency, with applications from rainfall to humidity. *Water Resources Research*, 56 (2), e2019WR026331. doi:10.1029/2019WR026331.

Papalexiou, S.M., Serinaldi, F., and Clark, M.P., 2023. Large-Domain multisite precipitation generation: operational blueprint and demonstration for 1,000 sites. *Water Resources Research*, 59 (3), e2022WR034094. doi:10.1029/2022WR034094.

Papalexiou, S.M., Serinaldi, F., and Porcu, E., 2021. Advancing Space-time simulation of random fields: from storms to cyclones and beyond. *Water Resources Research*, 57 (8), e2020WR029466. doi:10.1029/2020WR029466.

Parajka, J., *et al*., 2015. The role of station density for predicting daily runoff by top-kriging interpolation in Austria. *Journal of Hydrology and Hydromechanics*, 63 (3), 228–234. doi:10.1515/johh-2015-0024.

Parzen, E., 1962. On estimation of a probability density function and mode. *The Annals of Mathematical Statistics*, 33 (3), 1065–1076. doi:10.1214/aoms/1177704472.

Persiano, S., *et al*., 2020. Changes in seasonality and magnitude of subdaily rainfall extremes in Emilia-Romagna (Italy) and potential influence on regional rainfall frequency estimation. *Journal of Hydrology: Regional Studies*, 32, 100751.

Persiano, S., *et al*., 2021. A comparison between generalized least squares regression and top-kriging for homogeneous cross-correlated flood regions. *Hydrological Sciences Journal*, 66 (4), 565–579. doi:10.1080/02626667.2021.1879389.

Pettitt, A.N., 1979. A non-parametric approach to the change-point problem. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 28 (2), 126–135.

Piantadosi, J., *et al*., 2012. Maximum entropy methods for generating simulated rainfall. *Numerical Algebra, Control & Optimization*, 2 (2), 233. doi:10.3934/naco.2012.2.233.

Poulin, A., *et al*., 2007. Importance of tail dependence in bivariate frequency analysis. *Journal of Hydrologic Engineering*, 12 (4), 394–403. doi:10.1061/(ASCE)1084-0699(2007)12:4(394).

Price, K., *et al*., 2014. Comparison of radar and gauge precipitation data in watershed models across varying spatial and temporal scales. *Hydrological Processes*, 28 (9), 3505–3520. doi:10.1002/hyp.9890.

Prieto, C., *et al*., 2019. Flow prediction in ungauged catchments using probabilistic random forests regionalization and new statistical adequacy tests. *Water Resources Research*, 55 (5), 4364–4392. doi:10.1029/2018WR023254.

Prieto, C., *et al*., 2022. An exploration of Bayesian identification of dominant hydrological mechanisms in ungauged catchments. *Water Resources Research*, 58 (3), e2021WR030705. doi:10.1029/2021WR030705.

Prosdocimi, I. and Kjeldsen, T., 2021. Parametrisation of change-permitting extreme value models and its impact on the description of change. *Stochastic Environmental Research and Risk Assessment*, 35 (2), 307–324. doi:10.1007/s00477-020-01940-8.

Prosdocimi, I., Kjeldsen, T.R., and Miller, J.D., 2015. Detection and attribution of urbanization effect on flood extremes using nonstationary flood-frequency models. *Water Resources Research*, 51 (6), 4244–4262. doi:10.1002/2015WR017065.

Prosdocimi, I., Kjeldsen, T.R., and Svensson, C., 2014. Non-stationarity in annual and seasonal series of peak flow and precipitation in the UK. *Natural Hazards and Earth System Sciences*, 14 (5), 1125–1144. doi:10.5194/nhess-14-1125-2014.

Pugliese, A., *et al*., 2016. Regional flow duration curves: geostatistical techniques versus multivariate regression. *Advances in Water Resources*, 96, 11–22. doi:10.1016/j.advwatres.2016.06.008

Pugliese, A., Castellarin, A., and Brath, A., 2014. Geostatistical prediction of flow–duration curves in an index-flow framework. *Hydrology and Earth System Sciences*, 18 (9), 3801–3816. doi:10.5194/hess-18-3801-2014.

Rabiei, E. and Haberlandt, U., 2015. Applying bias correction for merging rain gauge and radar data. *Journal of Hydrology*, 522, 544–557. doi:10.1016/j.jhydrol.2015.01.020

Ragno, E., *et al*., 2018. Quantifying changes in future intensity-duration-frequency curves using multimodel ensemble simulations. *Water Resources Research*, 54 (3), 1751–1764. doi:10.1002/2017WR021975.

Ragno, E., *et al*., 2019. A generalized framework for process-informed nonstationary extreme value analysis. *Advances in Water Resources*, 130, 270–282. doi:10.1016/j.advwatres.2019.06.007

Ragno, E., Hrachowitz, M., and Morales-Nápoles, O., 2022. Applying non-parametric Bayesian networks to estimate maximum daily river discharge: potential and challenges. *Hydrology and Earth System Sciences*, 26 (6), 1695–1711. doi:10.5194/hess-26-1695-2022.

Rajagopalan, B. and Lall, U., 1995. A kernel estimator for discrete distributions. *Journal of Nonparametric Statistics*, 4 (4), 409–426. doi:10.1080/10485259508832629.

Rajagopalan, B., Lall, U., and Tarboton, D.G., 1997. Evaluation of kernel density estimation methods for daily precipitation resampling. *Stochastic Hydrology and Hydraulics*, 11, 523–547. doi:10.1007/BF02428432

Rao, A.R. and Hamed, K.H., 2019. *Flood frequency analysis*. Boca Raton, Florida, US: CRC press.

R Core Team, 2020. *R Core Team R: a language and environment for statistical computing*. Vienna, Austria: Foundation for Statistical Computing.

Reed, D.W., 2002. Reinforcing flood–risk estimation. *Philosophical Transactions of the Royal Society of London*, 360 (1796), 1373–1387. doi:10.1098/rsta.2002.1005.

Reis, D.S., Jr and Stedinger, J.R., 2005. Bayesian MCMC flood frequency analysis with historical information. *Journal of Hydrology*, 313 (1–2), 97–116. doi:10.1016/j.jhydrol.2005.02.028.

Renard, B., *et al*., 2008. Regional methods for trend detection: assessing field significance and regional consistency. *Water Resources Research*, 44 (8). doi:10.1029/2007WR006268.

Requena, A.I., Burn, D.H., and Coulibaly, P., 2019. Estimates of gridded relative changes in 24-h extreme rainfall intensities based on pooled frequency analysis. *Journal of Hydrology*, 577, 123940. doi:10.1016/j.jhydrol.2019.123940

Requena, A.I., Chebana, F., and Mediero, L., 2016. A complete procedure for multivariate index-flood model application. *Journal of Hydrology*, 535, 559–580. doi:10.1016/j.jhydrol.2016.02.004

Requena, A.I., Mediero, L., and Garrote, L., 2013. A bivariate return period based on copulas for hydrologic dam design: accounting for reservoir routing in risk estimation. *Hydrology and Earth System Sciences*, 17 (8), 3023–3038. doi:10.5194/hess-17-3023-2013.

Richardson, C.W., 1981. Stochastic simulation of daily precipitation, temperature, and solar radiation. *Water Resources Research*, 17 (1), 182–190. doi:10.1029/WR017i001p00182.

Rodriguez-Iturbe, I., De Power, B.F., and Valdes, J.B., 1987. Rectangular pulses point process models for rainfall: analysis of empirical data. *Journal of Geophysical Research: Atmospheres*, 92 (D8), 9645–9656. doi:10.1029/JD092iD08p09645.

Rodriguez-Iturbe, I., Mejia, J.M., and Dawdy, D.R., 1972. Streamflow simulation: 1. A new look at markovian models, fractional gaussian noise, and crossing theory. *Water Resources Research*, 8 (4), 921–930. doi:10.1029/WR008i004p00921.

Rutkowska, A., 2013. Statistical methods for trend investigation in hydrological non-seasonal series. *Acta Scientiarum Polonorum-Formatio Circumiectus*, 12 (4), 85–94.

Sadegh, M., *et al.*, 2018. Multihazard scenarios for analysis of compound extreme events. *Geophysical Research Letters*, 45 (11), 5470–5480. doi:10.1029/2018GL077317.

Sadegh, M., Ragno, E., and AghaKouchak, A., 2017. Multivariate Copula Analysis Toolbox (MvCAT): describing dependence and underlying uncertainty using a B ayesian framework. *Water Resources Research*, 53 (6), 5166–5183.

Saint Criq, L., *et al.*, 2022. Extreme sea level estimation combining systematic observed skew surges and historical record sea levels. *Water Resources Research*, 58 (3), e2021WR030873. doi:10.1029/2021WR030873.

Salas, J.D., *et al.*, 1980. *Applied modeling of hydrologic time series*. Colorado, US: Water Resources Publication.

Salas, J.D., *et al.*, 2020. PMP and climate variability and change. *Journal of Hydrologic Engineering*, 25 (12), 03120002. doi:10.1061/(ASCE)HE.1943-5584.0002003.

Salas, J.D. and Fernandez, B., 1993. Models for data generation in hydrology: univariate techniques. *Stochastic Hydrology and Its Use in Water Resources Systems Simulation and Optimization*, 47–73.

Salas, J.D. and Obeysekera, J., 2014. Revisiting the concepts of return period and risk for nonstationary hydrologic extreme events. *Journal of Hydrologic Engineering*, 19 (3), 554–568. doi:10.1061/(ASCE)HE.1943-5584.0000820.

Salas, J.D. and Obeysekera, J.T.B., 1992. Conceptual basis of seasonal streamflow time series models. *Journal of Hydraulic Engineering*, 118 (8), 1186–1194. doi:10.1061/(ASCE)0733-9429(1992)118:8(1186).

Salas, J.D., Obeysekera, J., and Vogel, R.M., 2018. Techniques for assessing water infrastructure for nonstationary extreme events: a review. *Hydrological Sciences Journal*, 63 (3), 325–352. doi:10.1080/02626667.2018.1426858.

Salinas, J.L., *et al.*, 2013. Comparative assessment of predictions in ungauged basins–Part 2: flood and low flow studies. *Hydrology and Earth System Sciences*, 17 (7), 2637–2652. doi:10.5194/hess-17-2637-2013.

Salvadori, G., *et al.*, 2007. *Extremes in nature: an approach using copulas*. Vol. 56. Dordrecht, Netherlands: Springer.

Salvadori, G., *et al.*, 2016. A multivariate copula-based framework for dealing with hazard scenarios and failure probabilities. *Water Resources Research*, 52 (5), 3701–3721. doi:10.1002/2015WR017225.

Salvadori, G., *et al.*, 2018. Hazard assessment under multivariate distributional change-points: guidelines and a flood case study. *Water*, 10 (6), 751. doi:10.3390/w10060751.

Salvadori, G. and De Michele, C., 2004. Frequency analysis via copulas: theoretical aspects and applications to hydrological events. *Water Resources Research*, 40 (12). doi:10.1029/2004WR003133.

Salvadori, G. and De Michele, C., 2010. Multivariate multiparameter extreme value models and return periods: a copula approach. *Water Resources Research*, 46 (10). doi:10.1029/2009WR009040.

Salvadori, G. and De Michele, C., 2011. Estimating strategies for multiparameter multivariate extreme value copulas. *Hydrology and Earth System Sciences*, 15 (1), 141–150. doi:10.5194/hess-15-141-2011.

Samaniego, L., Bárdossy, A., and Kumar, R., 2010. Streamflow prediction in ungauged catchments using copula-based dissimilarity measures. *Water Resources Research*, 46 (2). doi:10.1029/2008WR007695.

Schertzer, D. and Lovejoy, S., 1987. Physical modeling and analysis of rain and clouds by anisotropic scaling multiplicative processes. *Journal of Geophysical Research: Atmospheres*, 92 (D8), 9693–9714. doi:10.1029/JD092iD08p09693.

Seibert, J. and McDonnell, J.J., 2010. Land-cover impacts on streamflow: a change-detection modelling approach that incorporates parameter uncertainty. *Hydrological Sciences Journal–Journal Des Sciences Hydrologiques*, 55 (3), 316–332. doi:10.1080/02626661003683264.

Seidou, O., Asselin, J.J., and Ouarda, T.B.M.J., 2007. Bayesian multivariate linear regression with application to changepoint models in hydrometeorological variables. *Water Resources Research*, 43, W08401. doi:10.1029/2005WR004835

Seo, D.J., *et al.*, 1990b. Stochastic interpolation of rainfall data from rain gages and radar using cokriging: 2. Results. *Water Resources Research*, 26 (5), 915–924.

Seo, D.J., *et al.*, 2014. Improving real-time estimation of heavy-to-extreme precipitation using rain gauge data via conditional bias-penalized optimal estimation. *Journal of Hydrology*, 519, 1824–1835. doi:10.1016/j.jhydrol.2014.09.055

Seo, D.J., Krajewski, W.F., and Bowles, D.S., 1990a. Stochastic interpolation of rainfall data from rain gages and radar using cokriging: 1. Design of experiments. *Water Resources Research*, 26 (3), 469–477.

Serago, J.M. and Vogel, R.M., 2018. Parsimonious nonstationary flood frequency analysis. *Advances in Water Resources*, 112, 1–16. doi:10.1016/j.advwatres.2017.11.026

Serinaldi, F., 2009. A multisite daily rainfall generator driven by bivariate copula-based mixed distributions. *Journal of Geophysical Research: Atmospheres*, 114 (D10). doi:10.1029/2008JD011258.

Serinaldi, F., *et al.*, 2009. Probabilistic characterization of drought properties through copulas. *Physics and Chemistry of the Earth*, 34 (10–12), 596–605. doi:10.1016/j.pce.2008.09.004.

Serinaldi, F., 2013. An uncertain journey around the tails of multivariate hydrological distributions. *Water Resources Research*, 49, 6527–6547. doi:10.1002/wrcr.20531

Serinaldi, F., 2015. Dismissing return periods! *Stochastic Environmental Research and Risk Assessment*, 29 (4), 1179–1189. doi:10.1007/s00477-014-0916-1.

Serinaldi, F. and Kilsby, C., 2015. Stationarity is undead: uncertainty dominates the distribution of extremes. *Advances in Water Resources*, 77, 17–36. doi:10.1016/j.advwatres.2014.12.013

Serinaldi, F. and Kilsby, C.G., 2016. The importance of prewhitening in change point analysis under persistence. *Stochastic Environmental Research and Risk Assessment*, 30 (2), 763–777. doi:10.1007/s00477-015-1041-5.

Serinaldi, F. and Kilsby, C.G., 2018. Unsurprising surprises: the frequency of record-breaking and overthreshold hydrological extremes under spatial and temporal dependence. *Water Resources Research*, 54 (9), 6460–6487. doi:10.1029/2018WR023055.

Serinaldi, F., Kilsby, C.G., and Lombardo, F., 2018. Untenable nonstationarity: an assessment of the fitness for purpose of trend tests in hydrology. *Advances in Water Resources*, 111, 132–155. doi:10.1016/j.advwatres.2017.10.015

Serinaldi, F. and Lombardo, F., 2020. Probability distribution of waiting time of the kth extreme event under serial dependence. *Journal of Hydrologic Engineering*, 25. doi:10.1061/(ASCE)HE.1943-5584.0001923.

Sharma, A. and Mehrotra, R., 2010. Rainfall generation. *Washington DC American Geophysical Union Geophysical Monograph Series*, 191, 215–246.

Sharma, A., Wasko, C., and Lettenmaier, D.P., 2018. If precipitation extremes are increasing, why aren't floods? *Water Resources Research*, 54 (11), 8545–8551. doi:10.1029/2018WR023749.

Sharma, S. and Mujumdar, P.P., 2019. On the relationship of daily rainfall extremes and local mean temperature. *Journal of Hydrology*, 572, 179–191. doi:10.1016/j.jhydrol.2019.02.048

Shu, C. and Ouarda, T.B., 2007. Flood frequency analysis at ungauged sites using artificial neural networks in canonical correlation analysis physiographic space. *Water Resources Research*, 43 (7). doi:10.1029/2006WR005142.

Sinclair, S. and Pegram, G., 2005. Combining radar and rain gauge rainfall estimates using conditional merging. *Atmospheric Science Letters*, 6 (1), 19–22. doi:10.1002/asl.85.

Singh, V.P. and Strupczewski, W.G., 2002. On the status of flood frequency analysis. *Hydrological Processes*, 16 (18), 3737–3740. doi:10.1002/hyp.5083.

Sklar, M. (1959), Fonctions de répartition à n dimensions et leurs marges. Publications de l'Institut de statistique de l'Université de Paris. 8: p. 229–231.

Skøien, J.O. and Blöschl, G., 2007. Spatiotemporal topological kriging of runoff time series. *Water Resources Research*, 43 (9). doi:10.1029/2006WR005760.

Slater, L., *et al.*, 2021a. Global changes in 20-year, 50-year, and 100-year river floods. *Geophysical Research Letters*, 48, e2020GL091824. doi:10.1029/2020GL091824

Slater, L.J., *et al.*, 2021b. Nonstationary weather and water extremes: a review of methods for their detection, attribution, and management. *Hydrology and Earth System Sciences*, 25 (7), 3897–3935. doi:10.5194/hess-25-3897-2021.

Snelder, T.H., *et al.*, 2013. Regionalization of patterns of flow intermittence from gauging station records. *Hydrology and Earth System Sciences*, 17 (7), 2685–2699. doi:10.5194/hess-17-2685-2013.

Srikanthan, R. and McMahon, T.A., 2001. Stochastic generation of annual, monthly and daily climate data: a review. *Hydrology and Earth System Sciences*, 5 (4), 653–670. doi:10.5194/hess-5-653-2001.

Srinivas, V.V. and Srinivasan, K., 2005. Hybrid moving block bootstrap for stochastic simulation of multi-site multi-season streamflows. *Journal of Hydrology*, 302 (1–4), 307–330. doi:10.1016/j.jhydrol.2004.07.011.

Srivastav, R.K. and Simonovic, S.P., 2014. An analytical procedure for multi-site, multi-season streamflow generation using maximum entropy bootstrapping. *Environmental Modelling and Software*, 59, 59–75. doi:10.1016/j.envsoft.2014.05.005

Stedinger, J.R. and Cohn, T.A., 1986. Flood frequency analysis with historical and paleoflood information. *Water Resources Research*, 22 (5), 785–793. doi:10.1029/WR022i005p00785.

Stedinger, J.R., Vogel, R.M., and Foufoula-Georgiou, E., 1993. *Frequency analysis of extreme events, chapter 18, handbook of hydrology*. New York, US: McGraw-Hill Book Company, David R. Maidment, Editor-in-Chief.

Steinschneider, S., *et al.*, 2019. A weather-regime-based stochastic weather generator for climate vulnerability assessments of water systems in the western United States. *Water Resources Research*, 55 (8), 6923–6945. doi:10.1029/2018WR024446.

Steinschneider, S., Yang, Y.C.E., and Brown, C., 2013. Panel regression techniques for identifying impacts of anthropogenic landscape change on hydrologic response. *Water Resources Research*, 49 (12), 7874–7886. doi:10.1002/2013WR013818.

Stern, R.D. and Coe, R., 1984. A model fitting analysis of daily rainfall data. *Journal of the Royal Statistical Society: Series A (General)*, 147 (1), 1–18. doi:10.2307/2981736.

Strupczewski, W.G., *et al.*, 2012. On seasonal approach to flood frequency modelling. Part I: two-component distribution revisited. *Hydrological Processes*, 26 (5), 705–716. doi:10.1002/hyp.8179.

Strupczewski, W.G., *et al.*, 2016. Comparison of two nonstationary flood frequency analysis methods within the context of the variable regime in the representative Polish rivers. *Acta Geophysica*, 64, 206–236. doi:10.1515/acgeo-2015-0070

Svensson, C. and Jones, D.A., 2010. Review of rainfall frequency estimation methods. *Journal of Flood Risk Management*, 3 (4), 296–313. doi:10.1111/j.1753-318X.2010.01079.x.

Szilagyi, J., Balint, G., and Csik, A., 2006. Hybrid, Markov chain-based model for daily streamflow generation at multiple catchment sites. *Journal of Hydrologic Engineering*, 11 (3), 245–256. doi:10.1061/(ASCE)1084-0699(2006)11:3(245).

Szolgayova, E., *et al.*, 2014. Factors influencing long range dependence in streamflow of European rivers. *Hydrological Processes*, 28 (4), 1573–1586. doi:10.1002/hyp.9694.

Tessier, Y., *et al.*, 1996. Multifractal analysis and modeling of rainfall and river flows and scaling, causal transfer functions. *Journal of Geophysical Research: Atmospheres*, 101 (D21), 26427–26440. doi:10.1029/96JD01799.

Thiemig, V., *et al.*, 2013. Hydrological evaluation of satellite-based rainfall estimates over the volta and baro-akobo basin. *Journal of Hydrology*, 499, 324–338. doi:10.1016/j.jhydrol.2013.07.012

Thomas, J. and Fiering, M., 1962. Mathematical synthesis of streamflow sequences for the analysis of River Basins by Simulation. *In*: *Design of water-resource systems: new techniques for relating economic objectives, engineering analysis, and governmental planning*. Harvard, MA, US: Harvard University Press, 459–493.

Thorndahl, S., Nielsen, J.E., and Rasmussen, M.R., 2014. Bias adjustment and advection interpolation of long-term high resolution radar rainfall series. *Journal of Hydrology*, 508, 214–226. doi:10.1016/j.jhydrol.2013.10.056

Todini, E., 2001. A Bayesian technique for conditioning radar precipitation estimates to rain-gauge measurements. *Hydrology and Earth System Sciences*, 5 (2), 187–199. doi:10.5194/hess-5-187-2001.

Tosunoglu, F. and Singh, V.P., 2018. Multivariate modeling of annual instantaneous maximum flows using copulas. *Journal of Hydrologic Engineering*, 23 (3), 04018003. doi:10.1061/(ASCE)HE.1943-5584.0001644.

Treiber, B. and Plate, E.J., 1977. A stochastic model for the simulation of daily flows/Modèle stochastique pour la simulation des débits journaliers. *Hydrological Sciences Journal*, 22 (1), 175–192. doi:10.1080/02626667709491703.

Tyralis, H., Papacharalampous, G., and Tantanee, S., 2019. How to explain and predict the shape parameter of the generalized extreme value distribution of streamflow extremes using a big dataset. *Journal of Hydrology*, 574, 628–645. doi:10.1016/j.jhydrol.2019.04.070

Vazifehkhah, S., Tosunoglu, F., and Kahya, E., 2019. Bivariate risk analysis of droughts using a nonparametric multivariate standardized drought index and copulas. *Journal of Hydrologic Engineering*, 24 (5), 05019006. doi:10.1061/(ASCE)HE.1943-5584.0001775.

Velghe, T., *et al.*, 1994. Evaluation of cluster-based rectangular pulses point process models for rainfall. *Water Resources Research*, 30 (10), 2847–2857. doi:10.1029/94WR01496.

Veneziano, D., *et al.*, 2007. Marginal methods of intensity-duration-frequency estimation in scaling and nonscaling rainfall. *Water Resources Research*, 43, W10418.

Veneziano, D. and Langousis, A., 2005a. The maximum of multifractal cascades: exact distribution and approximations. *Fractals*, 13 (4), 311–324. doi:10.1142/S0218348X0500291X.

Veneziano, D. and Langousis, A., 2005b. The areal reduction factor a multifractal analysis. *Water Resources Research*, 41 (7), W07008.

Veneziano, D., Langousis, A., and Furcolo, P., 2006. Multifractality and rainfall extremes: a review. *Water Resources Research*, 42 (6). doi:10.1029/2005WR004716.

Veneziano, D. and Langousis, A., 2010. Scaling and fractals in hydrology. *In*: B. Sivakumar and R. Berndtsson, eds. *Advances in data-based approaches for hydrologic modeling and forecasting*. Singapore: World Scientific, 145.

Vermeer, M. and Rahmstorf, S. (2009). Global sea level linked to global temperature. *Proceedings of the national academy of sciences*, *106*(51), 21527–21532.

Villarini, G. *et al.*, 2009a. On the stationarity of annual flood peaks in the continental United States during the 20th century. *Water Resources Research*, 45 (8),

Villarini, G. *et al.*, 2009b. Flood frequency analysis for nonstationary annual peak records in an urban drainage basin. *Advances in Water Resources*, 32 (8), 1255–1266.

Visessri, S. and McIntyre, N., 2016. Regionalisation of hydrological responses under land-use change and variable data quality. *Hydrological Sciences Journal*, 61 (2), 302–320. doi:10.1080/02626667.2015.1006226.

Vogel, R.M., 1999. Stochastic and deterministic world views. *Journal of Water Resources Planning and Management*, 125 (6), 311–313.

Vogel, R.M., 2010. Regional calibration of watershed models. *In*: *Watershed models*. Boca Raton, Florida, US: CRC Press, 71–96.

Vogel, R.M., 2017. Stochastic watershed models for hydrologic risk management. *Water Security*, 1, 28–35. doi:10.1016/j.wasec.2017.06.001

Vogel, R.M. and Castellarin, A., 2017. Chapter 78 - Risk, reliability, and return periods and hydrologic design. *In*, and V.P. Singh, ed. *Handbook of applied hydrology*. New York, NY, USA: McGraw-Hill Book Company.

Vogel, R.M. and Fennessey, N.M., 1993. L moment diagrams should replace product moment diagrams. *Water Resources Research*, 29 (6), 1745–1752. doi:10.1029/93WR00341.

Vogel, R.M., Rosner, A., and Kirshen, P.H., 2013. Likelihood of Societal preparedness for global change –trend detection. *Natural Hazards and Earth System Science, Brief Communication*, 13, 1–6.

Vogel, R.M. and Stedinger, J.R., 1985. Minimum variance streamflow record augmentation procedures. *Water Resources Research*, 21 (5), 715–723.

Vogel, R.M., Yaindl, C., and Walter, M., 2011. Nonstationarity: flood magnification and recurrence reduction factors in the United States 1. *JAWRA Journal of the American Water Resources Association*, 47 (3), 464–474.

Volpi, E., *et al.*, 2015. One hundred years of return period: strengths and limitations. *Water Resources Research*, 51 (10), 8570–8585. doi:10.1002/2015WR017820.

Volpi, E., 2019. On return period and probability of failure in hydrology. *Wiley Interdisciplinary Reviews: Water*, 6 (3), e1340. doi:10.1002/wat2.1340.

Volpi, E., *et al.*, 2019. Save hydrological observations! Return period estimation without data decimation. *Journal of Hydrology*, 571, 782–792. doi:10.1016/j.jhydrol.2019.02.017

Volpi, E. and Fiori, A., 2012. Design event selection in bivariate hydrological frequency analysis. *Hydrological Sciences Journal*, 57 (8), 1506–1515. doi:10.1080/02626667.2012.726357.

Volpi, E. and Fiori, A., 2014. Hydraulic structures subject to bivariate hydrological loads: return period, design, and risk assessment. *Water Resources Research*, 50 (2), 885–897. doi:10.1002/2013WR014214.

Vormoor, K., *et al.*, 2011. Geostatistical regionalization of daily runoff forecasts in Norway. *International Journal of River Basin Management*, 9 (1), 3–15. doi:10.1080/15715124.2010.543905.

Vorogushyn, S. and Merz, B., 2012. What drives flood trends along the Rhine River: climate or river training? *Hydrology & Earth System Sciences Discussions*, 9 (12), 13537–13567.

Wahl, T., *et al.*, 2015. Increasing risk of compound flooding from storm surge and rainfall for major US cities. *Nature Climate Change*, 5 (12), 1093–1097. doi:10.1038/nclimate2736.

Wang, W. and Ding, J., 2007. A multivariate non-parametric model for synthetic generation of daily streamflow. *Hydrological Processes: An International Journal*, 21 (13), 1764–1771. doi:10.1002/hyp.6340.

Wang, W.P., *et al.*, 2015. Variance correction prewhitening method for trend detection in autocorrelated data. *Journal of Hydrologic Engineering*, 20 (12), 04015033. doi:10.1061/(ASCE)HE.1943-5584.0001234.

Westra, S., *et al.*, 2014. Future changes to the intensity and frequency of short-duration extreme rainfall. *Reviews of Geophysics*, 52 (3), 522–555. doi:10.1002/2014RG000464.

Whitfield, P.H., *et al.*, 2012. Reference hydrologic networks I. The status and potential future directions of national reference hydrologic networks for detecting trends. *Hydrological Sciences Journal*, 57 (8), 1562–1579. doi:10.1080/02626667.2012.728706.

Wied, D. and Galeano, P., 2013. Monitoring correlation change in a sequence of random variables. *Journal of Statistical Planning and Inference*, 143 (1), 186–1.

Wilks, D.S., 1998. Multisite generalization of a daily stochastic precipitation generation model. *Journal of Hydrology*, 210 (1–4), 178–191. doi:10.1016/S0022-1694(98)00186-3.

Wilks, D.S., 1999. Interannual variability and extreme-value characteristics of several stochastic daily precipitation models. *Agricultural and Forest Meteorology*, 93 (3), 153–169. doi:10.1016/S0168-1923(98)00125-7.

Wilks, D.S. and Wilby, R.L., 1999. The weather generation game: a review of stochastic weather models. *Progress in Physical Geography*, 23 (3), 329–357. doi:10.1177/030913339902300302.

Wilson, L.L., Lettenmaier, D.P., and Skyllingstad, E., 1992. A hierarchical stochastic model of large-scale atmospheric circulation patterns and multiple station daily precipitation. *Journal of Geophysical Research: Atmospheres*, 97 (D3), 2791–2809. doi:10.1029/91JD02155.

Worthington, A.C., Higgs, H., and Hoffmann, M., 2009. Residential water demand modeling in Queensland, Australia: a comparative panel data approach. *Water Policy*, 11 (4), 427–441. doi:10.2166/wp.2009.063.

Xu, P., *et al.*, 2020. Copula-based seasonal rainfall simulation considering nonstationarity. *Journal of Hydrology*, 590, 125439. doi:10.1016/j.jhydrol.2020.125439

Xu, Z., Schumann, A., and Li, J., 2003. Markov cross-correlation pulse model for daily streamflow generation at multiple sites. *Advances in Water Resources*, 26 (3), 325–335. doi:10.1016/S0309-1708(02)00172-0.

Yadav, M., Wagener, T., and Gupta, H., 2007. Regionalization of constraints on expected watershed response behavior for improved predictions in ungauged basins. *Advances in Water Resources*, 30 (8), 1756–1774. doi:10.1016/j.advwatres.2007.01.005.

Yaglom, A.M., 1966. The influence of fluctuations in energy dissipation on the shape of turbulence characteristics in the inertial interval. *Soviet Physics-Doklady*, 11, 26–29.

Yevjevich, V., 1984. Extremes in hydrology. *In*: *Statistical Extremes and Applications*. Dordrecht, Netherlands: Springer, 197–220.

Yevjevich, V.M., 1963. Fluctuations of wet and dry years. Part 1. Research data assembly and mathematical models. *Hydrology Paper 1*, Colorado, US: Colorado State Univ., Fort Collins.

Yilmaz, M. and Tosunoglu, F., 2019. Trend assessment of annual instantaneous maximum flows in Turkey. *Hydrological Sciences Journal*, 64 (7), 820–834. doi:10.1080/02626667.2019.1608996.

Yue, S., *et al.*, 1999. The Gumbel mixed model for flood frequency analysis. *Journal of Hydrology*, 226 (1–2), 88–100. doi:10.1016/S0022-1694(99)00168-7.

Yue, S., Pilon, P., and Phinney, B.O.B., 2003. Canadian streamflow trend detection: impacts of serial and cross-correlation. *Hydrological Sciences Journal*, 48 (1), 51–63. doi:10.1623/hysj.48.1.51.43478.

Zaghloul, M., *et al.*, 2020. Revisiting flood peak distributions: a pan-Canadian investigation. *Advances in Water Resources*, 145, 103720. doi:10.1016/j.advwatres.2020.103720

Zhang, L. and Singh, V.P., 2007a. Gumbel-Hougaard copula for trivariate rainfall frequency analysis. *Journal of Hydrologic Engineering*, 12 (4), 409–419. doi:10.1061/(ASCE)1084-0699(2007)12:4(409).

Zhang, L. and Singh, V.P., 2007b. Trivariate flood frequency analysis using the Gumbel-Hougaard copula. *Journal of Hydrologic Engineering*, 12 (4), 431–439. doi:10.1061/(ASCE)1084-0699(2007)12:4(431).

Zhang, L. and Singh, V.P., 2019. *Copulas and their applications in water resources engineering*. Cambridge, UK: Cambridge University Press.

Zhang, L.S.V.P. and Singh, V.P., 2006. Bivariate flood frequency analysis using the copula method. *Journal of Hydrologic Engineering*, 11 (2), 150–164. doi:10.1061/(ASCE)1084-0699(2006)11:2(150).

Zhang, Q., *et al.*, 2014. Flood frequency analysis with consideration of hydrological alterations: changing properties, causes and implications. *Journal of Hydrology*, 519, 803–813. doi:10.1016/j.jhydrol.2014.08.011

Zhang, Y., Cabilio, P., and Nadeem, K., 2016. Improved seasonal Mann–Kendall tests for trend analysis in water resources time series. In: *Advances in time series methods and applications: the A. Ian McLeod Festschrift*, pp. 215–229.

Zhu, D., Xuan, Y., and Cluckie, I., 2014. Hydrological appraisal of operational weather radar rainfall estimates in the context of different modelling structures. *Hydrology and Earth System Sciences*, 18 (1), 257–272. doi:10.5194/hess-18-257-2014.

Zorzetto, E., Botter, G., and Marani, M., 2016. On the emergence of rainfall extremes from ordinary events. *Geophysical Research Letters*, 43 (15), 8076–8082. doi:10.1002/2016GL069445.

Zscheischler, J., *et al.*, 2018. Future climate risk from compound events. *Nature Climate Change*, 8 (6), 469–477. doi:10.1038/s41558-018-0156-3.

Zucchini, W. and Guttorp, P., 1991. A hidden Markov model for space-time precipitation. *Water Resources Research*, 27 (8), 1917–1923. doi:10.1029/91WR01403.