Informativeness of teleconnections in local and regional frequency analysis of rainfall extremes

Andrea Magnini[1], Valentina Pavan[2], and Attilio Castellarin[1]

[1]Department of civil, environmental, chemical and materials engineering (DICAM), University of Bologna, Italy
[2]ARPAE-SIMC Emilia Romagna, Bologna, Italy

Correspondence: Andrea Magnini (andrea.magnini@unibo.it)

Abstract. We propose an effective and reproducible framework to assess the informative content of teleconnections for representing and modeling the frequency regime of rainfall extremes at regional scale. Our dataset consists of 680 annual maximum series of rainfall depth, with 1 and 24 hours durations, located in northern Italy. We compute at-site time series of L-moments through sliding time windows; then we discretize the study region into tiles, where L-moments time series are averaged. We

5   observe that the Western Mediterranean Oscillation index (WeMOI) shows strong spatial correlation patterns with gridded L-moments. Finally, in a preliminary application of climate-informed regional frequency analysis, the L-moments are modelled as functions of WeMOI. We observe high variability of WeMOI-driven rainfall percentiles predictions, and an increase in overall goodness-of-fit of the regional model relative to the stationary framework. Overall, our research suggests promising pathways for climate-informed local and regional frequency analysis of rainfall extremes, and describes general methods, that

10   can be adapted to different environmental variables.

1   Introduction

There is strong evidence that large-scale climate oscillations, also called teleconnections, have a significant influence on a region's climate (e.g., Bardossy and Plate, 1992; Bonsal and Shabbar, 2008; Rasouli et al., 2020). Several authors have in-

15   vestigated the link between teleconnections and the seasonal regime of precipitation, and found strong influence for monthly (e.g., Das et al., 2020; Romano et al., 2022) or 3-month (i.e., seasonal, e.g., Belkhiri and Krakauer, 2023; González-Pérez et al., 2022) cumulate rainfall and number of wet days (e.g., Ouachani et al., 2013; Ríos-Cornejo et al., 2015). Other authors focused on rainfall extremes, and showed how to exploit teleconnections to model the non-stationarity of the frequency regime of annual maxima (Cheng and AghaKouchak, 2014; Fauer and Rust, 2023; Ouarda et al., 2019; Ragno et al., 2018). Not only

20   the parameters of the frequency distribution can be represented as a function of teleconnections (e.g., El Adlouni and Ouarda, 2009), but also the distribution itself may change (Ouarda et al., 2019). Investigations on the balance between increased complexity and better reliability from non-stationary frequency analysis of rainfall extremes point out that the improvement is

worth the effort (Ouarda et al., 2020). Overall, the last findings appear to suggest the need for non-stationary frequency analysis of rainfall extremes, depending on teleconnections (see Nerantzaki and Papalexiou, 2022). Most of the studies about this

25 topic show applications with a limited number of stations, where observations are abundant enough for fitting non-stationary local frequency analysis models. However, the dependence of the extreme rainfall regime on teleconnections may have regional patterns, instead of being specific for some isolated sites. In case this dependence could be described as a function of space, this would lead to non-stationary models for regional frequency analysis, similarly as in the case of local frequency analysis.

The presence of regional structures in the dependence between teleconnections and rainfall has been investigated by several

30 authors, focusing on monthly/seasonal or annual totals (Caroletti et al., 2021; Das et al., 2020; Ríos-Cornejo et al., 2015) or droughts (Romano et al., 2022). Differently, this field remains highly unexplored for rainfall extremes. This is a complex problem, as rainfall extremes have higher statistical and spatial complexity than seasonal/monthly rainfall totals or number of wet days. Thus, the correlation between teleconnections and extreme rainfall can vary significantly in sign, strength and significance within space (see e.g., Jayaweera et al., 2023). In fact, climate may have strong local variations due to orography (Marra

35 et al., 2021), which makes it difficult to understand which teleconnections are more relevant to a specific region. Moreover, the length and quality of the observed timeseries play an important role in the reliability of the obtained results (Martins and Stedinger, 2000; Nerantzaki and Papalexiou, 2022; Ouarda et al., 2020).

In this study, we propose a framework for assessing the informativeness of teleconnections in frequency analysis of hourly and daily rainfall extremes. In particular, we want to investigate (1) whether it is possible to delineate robust regional zonation of

40 the correlation with teleconnections, and (2) what is the effect and suitability of non-stationary frequency analysis in significantly correlated areas. Accordingly, the study and proposed framework are structured in two parts, a correlation analysis and a regional frequency analysis, respectively. Our study area is northern-central Italy, where 680 timeseries with at least 30 years of records are available. We focus on annual maxima of precipitation with duration of 1 and 24 hours. In the first part of the research, we consider six teleconnection patterns with proven influence on the rainfall regime in the study area (Caroletti et al.,

45 2021; Criado-Aldeanueva and Soto-Navarro, 2020; Krichak et al., 2014): North Atlantic Oscillation (NAO), Pacific Decadal Oscillation (PDO), East Atlantic – West Russia pattern (EA-WR), El Niño Southern Oscillation (ENSO), Mediterranean Oscillation Index (MOI), and Western Mediterranean Oscillation Index (WeMOI).

Differently from other studies, we do not perform our correlation analysis on the raw timeseries of the teleconnections and annual maxima. Here, two strategies are simultaneously adopted, to conveniently aggregate the data temporally and spatially.

50 First, we consider sliding time windows, which allows us to (a) account only for long-term variability components, and (b) consider the variation of the timeseries' statistics during the recorded period. In particular, we consider two linear moments (or L-moments, see Hosking and Wallis, 1997): the mean and L-coefficient of variation (L-CV) for each station. Second, we divide the study region into tiles: within each tile, we average the at-site estimation of the L-moments, in order to obtain timeseries of robust regional L-moments estimates. Finally, we compute tile-wise the correlation between the timeseries of the L-moments

55 and the rolling mean of the teleconnections and we define raster maps of the correlation with the mean and L-CV.

In the second part, we focus on areas (tiles) where significant correlations are present, and we fit a tile-wise polynomial relationship between the L-moments and the most influent teleconnection. By using a hierarchical approach (see e.g. Gabriele and
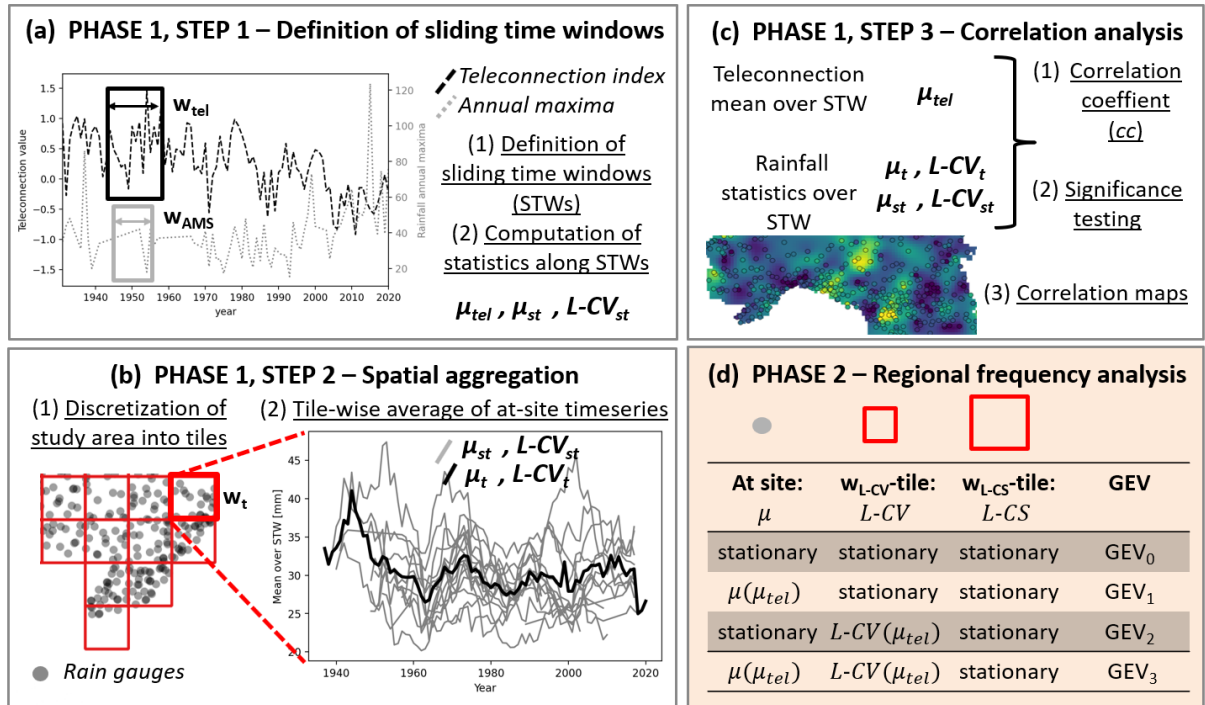
Figure 1. General methodological framework of the present study: first phase (white background, panels (a), (b) and (c)) and second phase (coloured background, panel (d)). Vectors are highlighted with bold, italic font, substeps are numbered and underlined.

Arnell, 1991; Castellarin et al., 2001), we define regional Generalized Extreme Value distributions (GEV, see Jenkinson, 1955) in a stationary and non-stationary framework. Finally, the research is enriched by a critical discussion on the generality and reproducibility of the proposed methodology. It is showed that beside the choice of the study area, our methods are innovative and universally applicable.

2 Methodological framework

We propose an innovative and structured methodological framework for assessing the effectiveness of teleconnections-informed frequency analysis of rainfall extremes. The general methodology is structured into two phases. Some elements of the procedure, which include numerical parameters and specific functions, need to be adjusted according to the specific study case; for the sake of brevity, all these elements will be referred to as "parameters". The general methodology is represented in Figure 1 and described in this Section, while the parametrization adopted for this study is detailed in Section 4.1.

2.1 Phase 1: correlation analysis

The first part aims at (1) evaluating the relation between teleconnection indexes and the local climate indexes (i.e., the statistics
of the rainfall regime), (2) investigating the spatial structure of the correlation, and (3) producing maps of the correlation
structure over the study area. These objectives are obtained through three fundamental steps which we describe below. The first
step is the definition of two sliding time windows (STWs). One STW is used for the teleconnections, with width w_{tel} years.
Over this STW, the mean of the teleconnection μ_{tel} is computed.

Another STW is used for the Annual Maximum Series (AMS), with width of w_{AMS} years. Over this STW, the at-site mean (μ)
and L-coefficient of variation (L-CV, see Hosking and Wallis, 1997) of the AMS of rainfall depths are computed (see Figure
1.a). Thus, for each gauging station (st) the time series of the mean ($\boldsymbol{\mu_{st}}$) and L-CV ($\boldsymbol{L\text{-}CV_{st}}$) are obtained; these have length
$n - w_{AMS} + 1$, where n is the number of years of observations for the considered site:

$$\boldsymbol{\mu_{st}} = \left\{ \mu_{1,st}, \mu_{2,st}, ..., \mu_{n-w_{AMS}+1,st} \right\} \tag{1}$$

$$\boldsymbol{L\text{-}CV_{st}} = \left\{ L\text{-}CV_{1,st}, L\text{-}CV_{2,st}, ...L\text{-}CV_{n-w_{AMS}+1,st} \right\} \tag{2}$$

where $\mu_{1,st}$, $\mu_{2,st}$ and $\mu_{n-w_{AMS}+1,st}$ represent the mean computed over the first, second and last time-steps defined by the
STW at site st; the same notation is used for the L-CV. Notably, at each time-step of the STW, the computed value (i.e., mean
or L-CV) is attributed to the last year included in that interval. Thus, each year is represented by a statistical moment (or
L-moment) that describes the w_{AMS} previous years, including itself.

The second step is the discretization of the spatial domain into single tiles (or cells, or pixels) that do not overlap with each
other. The spatial resolution is w_t. For each single tile (t), the timeseries $\boldsymbol{\mu_{st}}$ and $\boldsymbol{L\text{-}CV_{st}}$ of the gauged sites within the tile are
averaged yearly (see Figure 1.b). Thus, regional timeseries ($\boldsymbol{\mu_t}$ and $\boldsymbol{L\text{-}CV_t}$) for each tile are obtained:

$$\boldsymbol{\mu_t} = \left\{ \frac{\sum_{i=1}^{n_1} \mu_{1,i}}{n_1}, \frac{\sum_{i=1}^{n_2} \mu_{2,i}}{n_2}, ..., \frac{\sum_{i=1}^{n_{n-w_{AMS}+1}} \mu_{n-w_{AMS}+1,i}}{n_{n-w_{AMS}+1}} \right\} \tag{3}$$

$$\boldsymbol{L\text{-}CV_t} = \left\{ \frac{\sum_{i=1}^{n_1} L\text{-}CV_{1,i}}{n_1}, \frac{\sum_{i=1}^{n_2} L\text{-}CV_{2,i}}{n_2}, ..., \frac{\sum_{i=1}^{n_{n-w_{AMS}+1}} L\text{-}CV_{n-w_{AMS}+1,i}}{n_{n-w_{AMS}+1}} \right\} \tag{4}$$

where $n_1, n_2, ...n - w_{AMS} + 1$, are the numbers of stations with available rainfall statistics at steps 1,2, ..., $n - w_{AMS} + 1$
of the STW.

The third step is the correlation analysis (see Figure 1.c). The correlation coefficient (cc) is computed for each tile between
the averaged timeseries ($\boldsymbol{\mu_t}$ and $\boldsymbol{L\text{-}CV_t}$) and the rolling mean of the considered teleconnection ($\boldsymbol{\mu_{tel}}$), and the significance of
the resulting correlation is tested (with significance 5%). We propose to adopt the Spearman coefficient, as it accounts also

for non-linear correlations, yet preliminary experiments showed very similar results with the Pearson correlation coefficient.
Only stations with p-value≤0.05 are considered as significantly correlated. Only the teleconnections showing high numbers of
significantly correlated pixels are considered for the further steps.

Finally, an additional analysis of the correlation coefficients is needed to delineate eventual spatial structures. Since the definition of the parameters described above (i.e., w_{tel}, w_{AMS}, w_t) is necessarily affected by subjectivity and uncertainty, the
robustness, or reliability, of the detected correlation should be assessed: the method adopted for this assessment is detailed in
Section 4.1. Then, a map of the correlation field over the study area can be created.

2.2 Phase 2: regional frequency analysis

The second part of the present study aims at assessing (1) the effect of teleconnections on frequency analysis of extreme
rainfall, and (2) the benefit of considering teleconnections as covariates for non-stationary frequency analysis.

For the sake of generality, the Generalized Extreme Value distribution (GEV, see Jenkinson, 1955) is considered, given its
flexibility and representativeness of the frequency regime of hydrological extremes (e.g., Papalexiou and Koutsoyiannis, 2013;
Salinas et al., 2014). Nevertheless, the proposed framework is not limited to the selection of a given frequency distribution;
on the contrary, it could be easily extended and adapted to the case in which alternative theoretical frequency distributions are
tested against each other. The cumulate probability distribution of the GEV, $F_{GEV}(x)$, depending on the location, scale and
shape parameters (ξ,α,k), is defined as follows:

$$F_{GEV}(x) = e^{-e^{-y}} \text{ , where } y = \begin{cases} -k^{-1}log_e[1 - k(x-\xi)/\alpha], & k \neq 0 \\ (x-\xi)/\alpha, & k = 0 \end{cases} \tag{5}$$

The hierarchical method for regional frequency analysis is adopted (see Gabriele and Arnell, 1991). Accordingly, for a given
gauged site, the mean is computed from the at-site records, the L-CV exploits all the time series within a tile with resolution
w_{L-CV} where the target site is included, and the L-CS exploits all the time series within a tile with resolution w_{L-CS} (where
$w_{L-CV} \leq w_{L-CS}$). The values adopted in the present study for w_{L-CV} and w_{L-CS} are detailed in Section 4.1.
First, a stationary GEV is fit (GEV_0), where the mean over the whole time series is computed at-site, and the regional L-CV
and L-CS are obtained as described in Hosking and Wallis (1997) as a weighted average over their respective tiles by referring
to the complete sequences of annual maxima.

Second, three types of non-stationary GEV are set up. The first type, GEV_1, adopts the same L-CV and L-CS as the GEV_0,
whereas the mean varies as a function of the best teleconnection index (selected in phase 1). This function is fitted at-site,
according to the hierarchical regionalization framework. The choice for its shape, $f(x)$, is detailed in Section 4.1. The second
type GEV, GEV_2, adopts the same mean and L-CS as the GEV_0, while the L-CV varies as a function $f(x)$ of the best index.
This is fitted on the w_{L-CV}-averaged timeseries (**L-CV$_t$**), according to the hierarchical regionalization framework. The third
type GEV, GEV_3, adopts the same L-CS as the GEV_0, while the mean and L-CV are obtained with the same methods as for the
GEV_0 and GEV_2, respectively. Finally, the four models are compared by means of three commonly used metrics: the ratio of

125 maximized likelihood (RML), the Anderson-Darling statistic (AD), and the modified Shapiro-Wilk statistic (TN.SW). These three metrics, or discrimination statistics, as they are frequently defined, are selected based on the results of other authors. Ashkar and Aucoin (2012) and Ashkar and Ba (2017) found that the discrimination power of these three metrics is higher than other metrics after multiple and extensive testing. In particular, Laio et al. (2009) found that the Anderson-Darling statistic had better discrimination power than other considered metrics when the parent distribution is a GEV. The RML is defined as in

130 Ashkar and Ba (2017):

$$RML = log_e\Big(\frac{LH_{non-stationary}}{LH_{GEV_0}}\Big) \tag{6}$$

The AD statistic is defined as in Laio et al. (2009) and Laio (2004). For an ordered series of annual maxima, $\boldsymbol{x} = \{x_1 < x_2 < ... < x_i < ... < x_n\}$, the AD depends on the originally version proposed by Anderson and Darling (1952):

$$AD_{1952} = -n - \frac{1}{n}\sum_{i=1}^{n}[(2i-1)\cdot log_e(F(x_i)) + (2(n-i)+1)\cdot log_e(1-F(x_i))] \tag{7}$$

135 $$AD = \begin{cases} 0.0403 + 0.116\Big(\frac{AD_{1952}-\xi}{\beta}^{\eta/0.861}\Big), & 1.2\xi \leq AD_{1952} \\ \Big[0.0403 + 0.116\Big(\frac{0.2\xi}{\beta}^{\eta/0.861}\Big)\Big] \cdot \frac{AD_{1952}-0.2\xi}{\beta}, & 1.2\xi > AD_{1952} \end{cases} \tag{8}$$

where $F(x)$ is the cumulative density function of the probability distribution under exam, AD_{1952} is the orginal Anderson-Darling statistic (see also Das, 2022), and ξ, β and η are distribution-dependent coefficients that are tabled by Laio (2004) (Tables 3 and 5).

The TN.SW statistic, generated by the modified Shapiro-Wilk test (see the original test by Shapiro and Wilk (1965)), is

140 performed as in Ashkar and Ba (2017). First, the series of annual maxima is transformed into a standard normal sample by applying the cumulative density function of the distribution under exam, $F(x)$, and the percentile function of the standard normal distribution, $\Phi^{-1}(x)$:

$$z_i = \Phi^{-1}(F(x_i)) \tag{9}$$

Then, the TN.SW is computed:

145 $$TN.SW = \Big(\sum_{i=1}^{n} v_i z_i\Big)^2 \Big/ \sum_{i=1}^{n}(z_i - \mu_z)^2 \tag{10}$$

where the coefficients v_i are computed as in the approximation described by Royston (1992), and μ_z is the mean of the standardized vector $\boldsymbol{z}$.

These three metrics are computed for the four types of GEV at all the locations where the correlation is significant (i.e., p-value≤5%) and strong (i.e., Spearman coefficient≤-0.5 or ≥0.5). While the RML represents the comparison between the LH

150 of two distribution probabilities, the AD and TN.SW measure the goodness-of-fit of a single distribution to the data (i.e., larger values represent better fit). In order to compare the stationary with the non-stationary distributions, we consider the differences
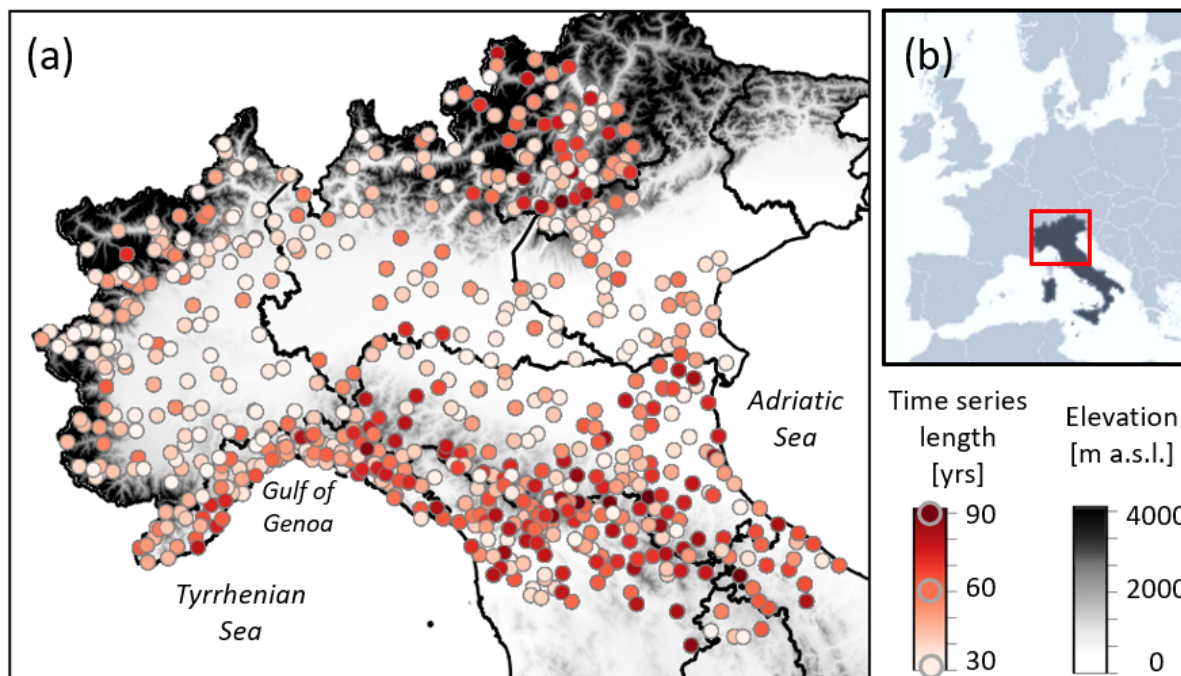
Figure 2. Elevation (a, in grey color scale) and location (b) of the study area. Length of the timeseries of sub-daily annual maximum rainfall depths (a, red color scale). In black: Italian administrative regions.

$\Delta_{AD} = AD_{non-stationary} - AD_{GEV_0}$ and $\Delta_{TN.SW} = TN.SW_{non-stationary} - TN.SW_{GEV_0}$ (Ashkar and Ba, 2017; Das, 2022). Thus, either for RML, Δ_{DA} and $\Delta_{TN.SW}$, positive values indicate that the considered non-stationary model provides a higher goodness-of-fit to the observed annual maxima relative to the stationary model.

3 Dataset

The study area includes most of northern and part of central Italy, a region characterized by great climate variability (see Figure 2.a). Two main mountain ranges are present: the Alps in the north, with a maximum elevation of 4000 m a.s.l., and the Apennines, crossing all along continental Italy, with a maximum elevation of ∼2100 m a.s.l. in the study area. The largest Italian plain, the Po plain, is located at the southern border of the Alps, following the course of the Po River from the northwest to the northeast, where low coasts are located.

We select 680 gauged stations (Figure 2.a) from the I2-RED dataset (Mazzoglio et al., 2020), filtering by a minimum of 30 years of data. Thus, all the selected timeseries should be long enough to show variations of the statistical regime during time, if these are present (see also Renard et al., 2008; Ouarda et al., 2019). For each station, we consider time series annual maximum cumulative rainfall over 1 and 24 consecutive hours, which represent distinct events: mainly convective and mainly synoptic, respectively. Data have been recorded between 1928 and 2020.

In the study, six teleconnections are considered; a detailed description of their nature is not reported here, since an interested reader can refer to the extensive literature cited in the text. Namely, these are the North Atlantic Oscillation (NAO, see Jones et al., 1997), the East Atlantic-West Russia (EA-WR) pattern (see Barnston and Livezey, 1987), the Pacific Decadal Oscillation (PDO, see Zhang et al., 1997), El Nino Southern Oscillation (ENSO, see Chen et al., 2019), Mediterranean Oscillation Index

170 (MOI, see Conte et al., 1991), and Western Mediterranean Oscillation Index (WeMOI, see Martin-Vide and Lopez-Bustins, 2006). All these climate indexes have significant influences on several locations in Europe and the Mediterranean (e.g., Caroletti et al., 2021; Krichak et al., 2014, 2002; Krichak and Alpert, 2005). The NAO, EA-WR, PDO and ENSO are freely accessible from the NOAA Physical Sciences Laboratory data base available at https://psl.noaa.gov/data/climateindices/list/. The MOI and WeMOI can be retrieved from the University of East Anglia's Climate Research Unit (CRU; https://crudata.uea.ac.uk/cru/

175 data/moi/).

4 Results

4.1 Parametrization of the procedure for the study area

Based on several preliminary experiments conducted, the parametrization of the general methodology to the study area is the following.

180 The first part of the study requires the definition of w_{tel}, w_{AMS}, w_t, and the methodology for the robustness analysis. First, w_{tel} is assumed as 30 years. Preliminary experiments with smaller w_{tel} provided similar results, yet not exactly the same. The final choice accounts to 30 years in order to smooth short interannual oscillations, considering instead a pluridecadal climate window. Shorter oscillations (i.e., <30 years) would not be of interest for the design of hydraulic structures, while longer ones (i.e., >30 years) would be highly uncertain to detect, due to limited length of the timeseries.

185 Second, the w_{AMS} parameter is assumed as 10yrs. This choice is also due to a balance between a minimum width for the computation of the L-CV and a minimum length of the timeseries of the rainfall statistics (i.e., $\boldsymbol{\mu}_{st}$ and $\boldsymbol{L\text{-}CV}_{st}$). In fact, on the one hand w_{AMS} <10 would lead to unacceptable inaccuracy for the definition of the L-CV. On the other hand, the $\boldsymbol{\mu}_{st}$ and $\boldsymbol{L\text{-}CV}_{\mathbf{st}}$ timeseries originated from an n-long annual maxima timeseries have length of $n - (w_{AMS} - 1)$. This means that the longer w_{AMS} would lead to shorter $\boldsymbol{\mu}_{\mathbf{st}}$ and $\boldsymbol{L\text{-}CV}_{\mathbf{st}}$, which in turn would lead to a smaller number of stations where

190 the correlation between teleconnections and rainfall statistics can be reliable. Indeed, to the aim of defining a zonation of the teleconnection-statistics correlation field, as in the present study, the number of sufficiently long timeseries needs to be as high as possible.

According to the adopted parametrization, equations 1 to 4 become the following:

$$\boldsymbol{\mu}_{\mathbf{st}} = \left\{\mu_{1,st}, \mu_{2,st}, ..., \mu_{n-10+1,st}\right\} \tag{11}$$

195 $$\boldsymbol{L\text{-}CV}_{\mathbf{st}} = \left\{L - CV_{1,st}, L - CV_{2,st}, ..., L - CV_{n-10+1,st}\right\} \tag{12}$$

$$\boldsymbol{\mu_t} = \left\{ \frac{\sum_{i=1}^{n_1} \mu_{1,i}}{n_1}, \frac{\sum_{i=1}^{n_2} \mu_{2,i}}{n_2}, ..., \frac{\sum_{i=1}^{n_{n-10+1}} \mu_{n-10+1,i}}{n_{n-10+1}} \right\} \tag{13}$$

$$\boldsymbol{L\text{-}CV_t} = \left\{ \frac{\sum_{i=1}^{n_1} L-CV_{1,i}}{n_1}, \frac{\sum_{i=1}^{n_2} L-CV_{2,i}}{n_2}, ..., \frac{\sum_{i=1}^{n_{n-10+1}} L-CV_{n-10+1,i}}{n_{n-10+1}} \right\} \tag{14}$$

For the spatial resolution of the tile size, w_t, we adopt four values: 0km (i.e., considering the single gauged stations, with no spatial discretization), 15km, 30km and 50km. This multiple choice comes from a balance. On the one hand, L-CV computed over a 10yrs STW may have low robustness, which can be addressed by averaging L-CVs from several stations within large tiles. On the other hand, larger tiles may be less statistically homogeneous. Moreover, averaging L-statistics over large tiles may smooth the variability of the rainfall regime, hiding local patterns where the morphology is complex. Since there is no universal rule for solving this balance, we decide to consider four different values for w_t. The suitability of these values is tested by means of the heterogeneity test described by Hosking and Wallis (1997) for the L-CV. The results show that most of the tiles are homogenous for all the resolutions. These analyses are not reported here for the sake of brevity.

Regarding the robustness of the correlation signal is defined through a reliability index. At each station st, it is defined as ri_{st}:

$$ri_{st} = sign(cc_{0km,st}) + sign(cc_{15km,st}) + sign(cc_{30km,st}) + sign(cc_{50km,st}) \tag{15}$$

Where $cc_{0km,st}$ is the correlation computed at station st, while $cc_{15km,st}$, $cc_{30km,st}$ and $cc_{50km,st}$ are the correlation coefficients relative to the tiles (with w_t 15km, 30km and 50km) where st is inserted. Non-significant correlations are considered as 0.

This is considered to be a measure of the spatial coherence of the correlation signal, which varies between -4 and 4. The absolute value is the coherence of the correlations at different tiles. The sign represents the sign of the prevailing correlation. For instance, considering a gauged station where the absolute value of ri is 4, the correlation is significant and with the same sign both when calculated at-site, and when it derives from a 15km aggregation, and at 30km, and at 50km. On the opposite, 0 represents areas with no significant correlation or where positive and negative correlations compensate with each other (e.g. positive correlation at-site and at a 15km tile, and negative correlation at 30km and 50km). In general, the ri can be rescaled between -1 and 1, in order be more similar to a correlation coefficient. In this study, the original scale is kept, as it is more easily interpretable.

Finally, the reliability index is interpolated by ordinary kriging (Hengl, 2007) to produce a robust map of the correlation field. Regarding the second part of the study, three parameters need to be adjusted: w_{L-CV}, w_{L-CS} and $f(x)$. The selected resolution w_{L-CV} is 30km, as it is a good balance between robustness (i.e., aggregating at least two L-CV at-site timeseries) and representing regional patterns and local variability. The selected resolution w_{L-CS} is 100km, based on low spatial variability of the skewness parameter (e.g., Gabriele and Arnell, Claps et al 2022) and improved robustness (see above).

The function $f(x)$ between teleconnections and L-statistics is shaped as a second-order polynomial function. This form has

225   been selected because of its simplicity (i.e., only two parameters are needed) and yet adaptability to the data, while the dependence observed of extreme rainfall statistics on teleconnections is clearly non-linear. Other choices are clearly possible. However, it should be underlined that the aims of the present study are mainly demonstrative of the potential of the proposed approach. Thus, the nature and selection of the best function is not discussed in detail.

4.2   Results of phase 1: correlation analysis

230   First, it is observed that the WeMOI shows a remarkably higher number of significant correlations in the study area than any other teleconnection considered in the study. When considering single stations, we found 387 significant correlations for hourly rainfall for WeMOI, while for the other indexes they range from 0 (ENSO) to 78 (PDO). Due to this huge difference, only the results obtained for WeMOI are presented below.

Figure 2 reports significant Spearman correlation coefficients with WeMOI at different resolutions for spatial aggregation for
235   the mean (panels (a)-(h)) and L-CV (panels (i)-(p)), respectively. Several stations present statistically significant correlation values with this index, with signs and amplitude changing depending on the site considered (panels (a) and (e) for mean, (i) and (m) for L-CV, Figure 3). Aggregating stations into tiles reduces spatial heterogeneity, and allows to describe the geographical pattern of the correlation field (see other panels).

Both for the mean and L-CV, the correlation fields of extreme rainfall with 1h duration present complex spatial patterns, with
240   small areas characterized by homogeneous values.

The reliability index (ri, see Equation 15) is showed for the mean and L-CV in Figure 4. As observed above, stable geographical structures in the correlation are larger (1) for the 24h duration than for the 1h (compare panels (a) and (b) with panels (c) and (d) in Figure 4), and (b) for the mean than for L-CV (compare panels (a) and (c) with panels (b) and (d) in Figure 4).

Concerning the mean of the AMS series, two major patterns of robust negative correlation are present in the Western coast of
245   the Tyrrhenian Sea (Gulf of Genoa) and in the north-eastern Alps. Less intense but significant positive correlation values are present along the northern flank of the northern Apennines. For the L-CV, the pattern in the Gulf of Genoa is still present, but less evident and less extended, while some positively correlated hotspots are present, mostly located in the Southern portion of the study area.

4.3   Results of phase 2: Frequency analysis

250   Following the adopted framework (see Section 3.2), regional frequency analysis is applied to all the locations where significant correlation is observed (i.e., p-value≤5%). Figure 4 illustrates the results obtained for a station (see panel (a) for its location) where the three non-stationary GEV can be fitted (i.e., GEV_1 in panel (c), GEV_2 in panel (d), and GEV_3, in panel (e)).

It is evident from the example provided in Figure 5 that the variation of the expected percentiles due to the dependence of the GEV parameters on the teleconnections may be very significant. The prediction of the 100-year 24-hour rainfall depth in
255   the selected location is equal to 370mm according to the stationary model (GEV_0), but may vary between 340 and 375 if the empirical relationship between WeMOI and L-CV is explicitly modelled (GEV_2), between 300 and 460 if the mean 24-hour annual maximum rainfall depth becomes a function of WeMOI (GEV_1), and may be as low as 275 and as high as 470 if both
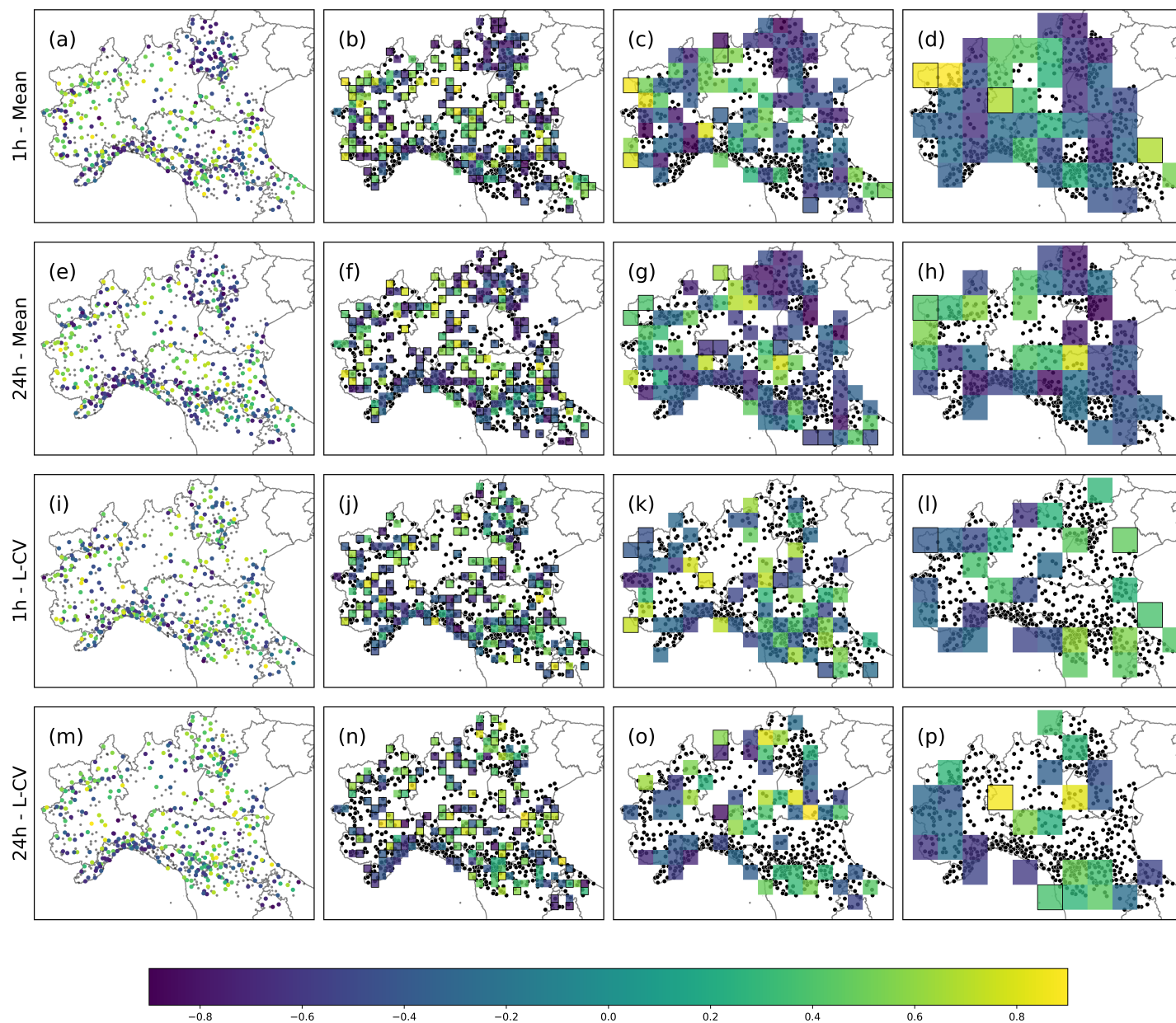
Figure 3. Correlation coefficient for the mean [L-CV] for AMS of rainfall depths with duration of 1h and 24h: at-site ((a)/(i) and (e)/(m)) and for tiles of size 15km ((b)/(j) and (f)/(n)), 30km ((c)/(k) and (g)/(o)), and 50km ((d)/(l) and (h)/(p)). Only statistically significant (at 5%) correlation coefficients are illustrated using a purple-green-yellow color scale. Black outlines highlight tiles where only one station is present

mean and L-CV are expressed as functions of WeMOI (GEV_3).

Figure 6 represents the RML, Δ_{AD}, and $\Delta_{TN.SW}$ (panels (a), (b) and (c), in this order). These measure the increase in
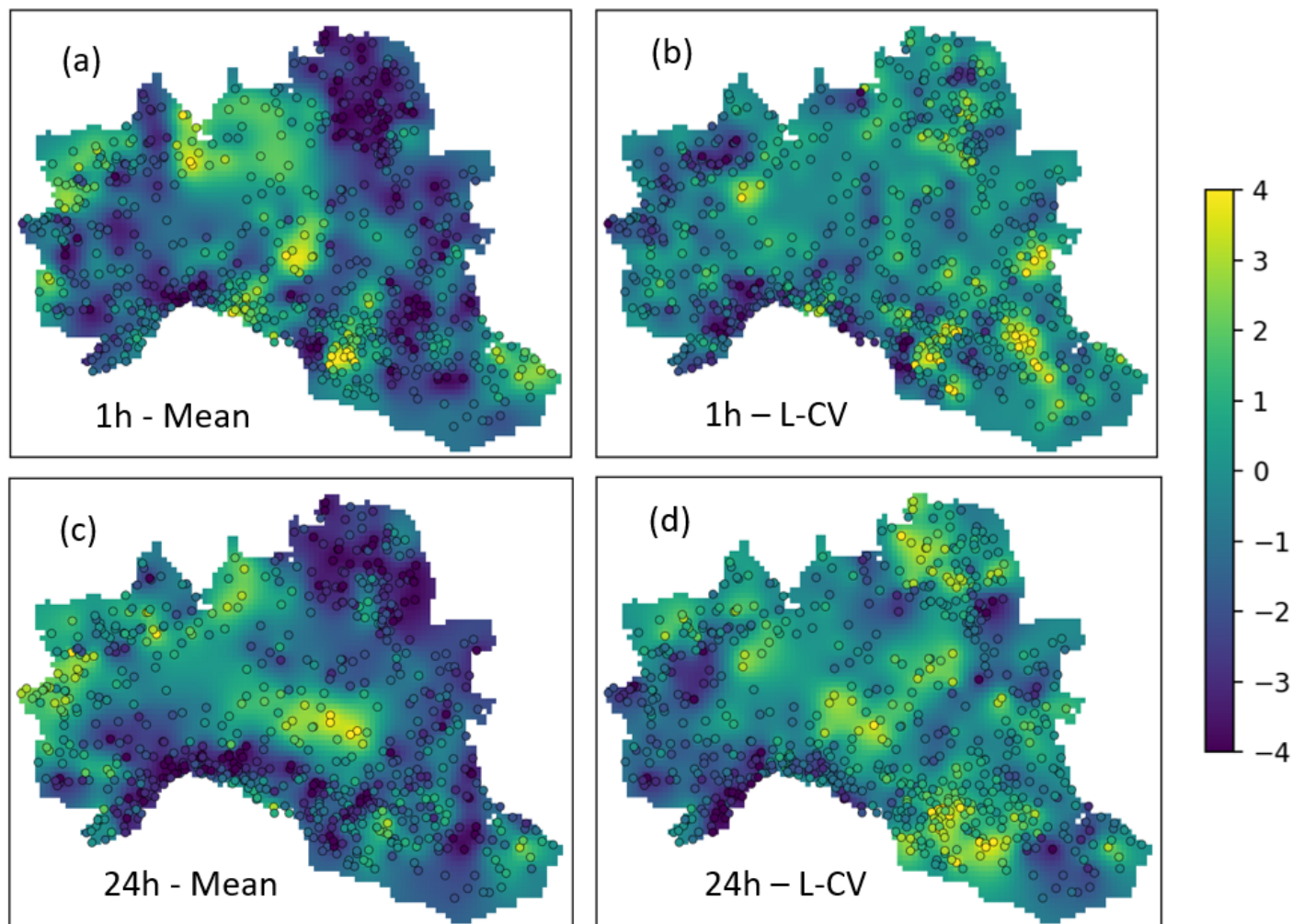
Figure 4. Raster maps of reliability index (ri) of the correlation between WeMOI and mean/L-CV of AMS of rainfall depths with duration of 1h (a/b) and 24h (c/d)

260   goodness-of-fit due to the non-stationary regional GEV models with respect to the stationary ones. They are computed for all locations in which statistically significant correlation between AMS statistics and WeMOI was detected (see Sections 2.1 and 2.2).

Values of RML (Figure 6.a) and Δ_{AD} are mainly positive for GEV_1 and GEV_3 (i.e., non-stationary models show a better representation relative to the stationary one), where negative values are present in less than 25% of the cases. For GEV_2, in

265   nearly half of the stations examined the precipitation regime is better represented by a stationary framework. Differently, values of $\Delta_{TN.SW}$ are almost equally divided into positive and negative for both the 1h and 24h durations and all the non-stationary models.
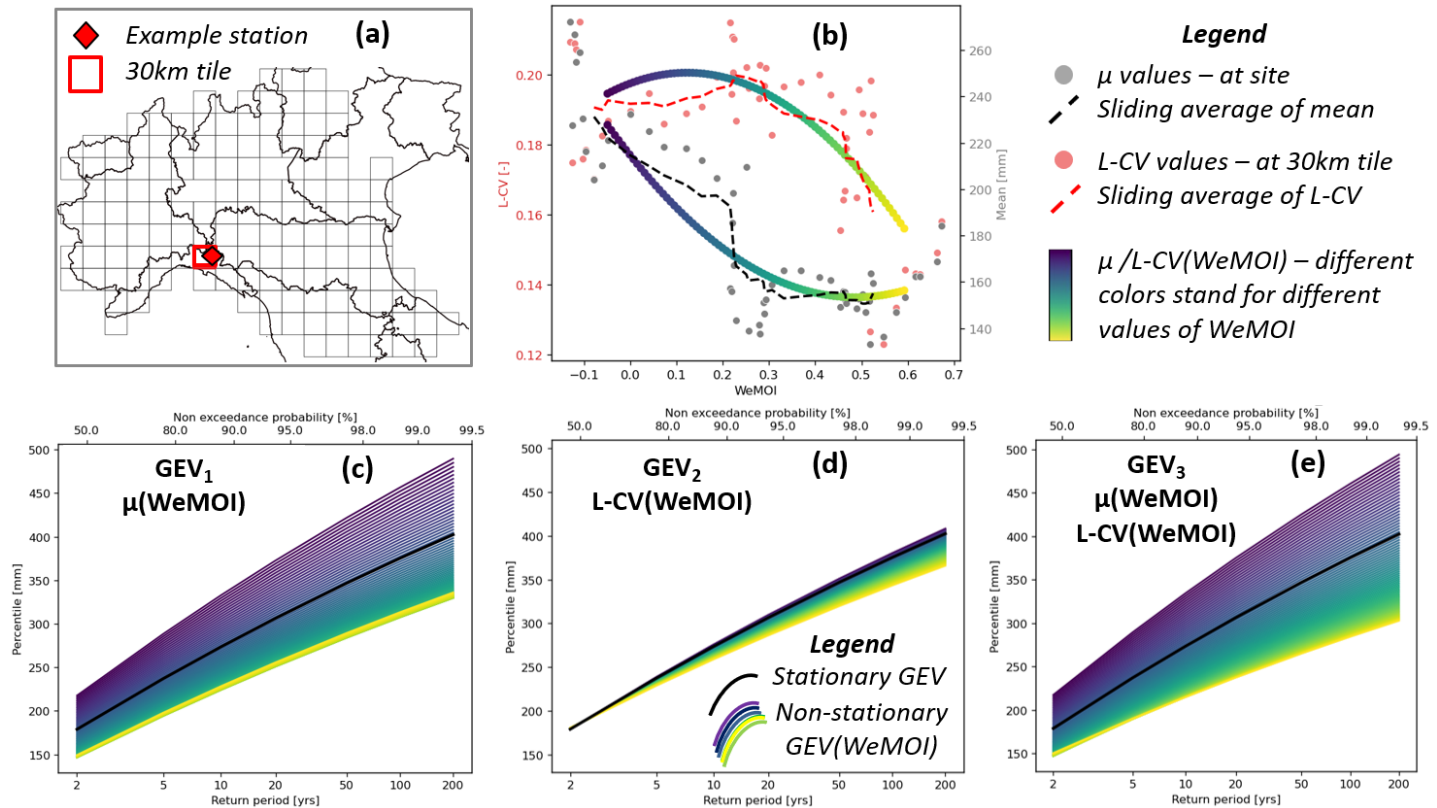
Figure 5. Example of location where regional frequency analysis for 24h duration is performed. Upper panels: location of example station (a), polynomial function of mean (μ) and L-CV depending on WeMOI (b). Lower panels: expected percentiles with given return periods in stationary (black line) and non-stationary framework (colored scale lines) when only the mean depends on the WeMOI (c), only the L-CV depends on the WeMOI (d), or both mean and L-CV depend on the WeMOI (e).

Overall, the GEV_2 model shows the lowest number number of cases (which is still above 50%) where it should be preferred to the stationary regime. For all the three metrics, light differences are observed between the results for 1h and 24h.

5 Discussion

5.1 Zonation of the correlation

The stronger influence of the WeMOI compared to the other teleconnections was expected. In fact, the WeMOI consists of the normalized difference between Cadiz, in the South of Spain, and Padua, in northern Italy, and thus, it describes the formation of precipitation systems over the Tyrrhenian Sea (Lopez-Bustins et al., 2020; Redolat et al., 2019). Strong significant correlations with the WeMOI were also found in other studies, as for Central Italy by Romano et al., 2021 and for Southern Italy by Caroletti et al., 2021. In particular, the presence of significant intense negative correlation values in the Gulf of Genoa and the
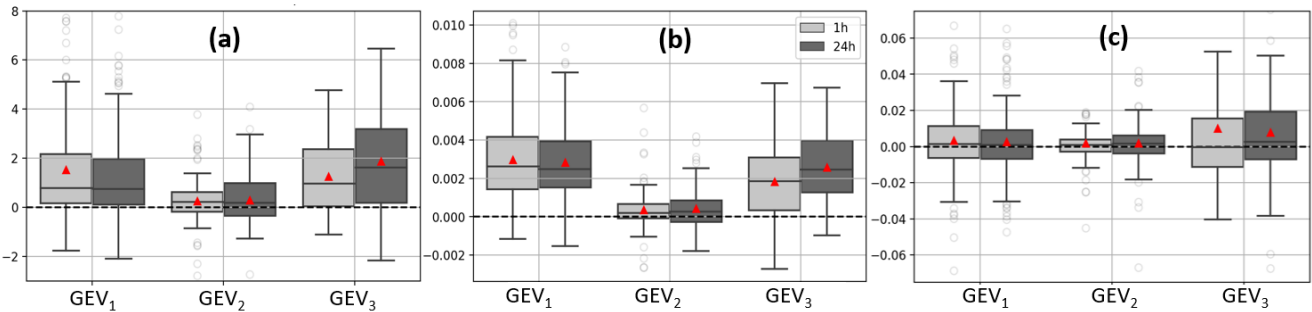
Figure 6. Evaluation metrics for stationary and non-stationary RFA framework for locations where at-site strong and significant correlation μ_{WeMOI}-μ_{st} is present (GEV_1), 30km strong and significant correlation μ_{WeMOI}-L-CV_t is present (GEV_2), and both correlations μ_{WeMOI}-μ_{st} and μ_{WeMOI}-L-CV_t are strong and significant (GEV_3). In particular, RML (a), Δ_{AD} (b), and $\Delta_{TN.SW}$ (c). For all the metrics, results larger than zero represent better performances of non-stationary over stationary framework.

North-East for the mean-WeMOI correlation (Figure 4.a and 4.c) is consistent with the known patterns of precipitation regimes over north-central Italy. In fact, intense daily precipitation values are expected over the Tyrrhenian Coast and north-eastern Alps in the presence of intense southwesterly flows from the Mediterranean, typical during the autumn season and favored by large

280   scale circulation anomalies associated with negative value of the WeMOI. On the other hand, in the presence of positive values of the index, precipitation is expected to be linked to cut-off flows mostly affecting the Adriatic section of the Apennines. In the remaining portion of the study area, precipitation systems are more complex, as influenced by the passage of cut-of flow favoring precipitation over the southern and eastern portion of the Apennines area. In this case, correlation patterns with the WeMOI are expected to be more fragmented (Figure 4.c).

285   For the 1h rainfall maxima, mostly linked to convective phenomena, which are often characterized by a very limited spatial scale, the correlation with WeMOI presents more complex geographical patterns, even if still present (Figure 4.a).

A physical interpretation of the L-CV-WeMOI dependence is more complex. It is evident that some hotspots of correlation are present (Figure 3.k-l for and 3.o-p): negative correlations are mostly located in the North-West, which partially confirms the results for the mean (compare Figure 3.f-g with 3.n-o). Elsewhere, small and fragmented, mainly positive patterns are visible.

290   Considering the ri, the robustness of patterns to spatial aggregation is lower than for the mean, which may be due to higher uncertainty in at-site computation of L-CV within a 10-year time window. For this reason, we believe that for L-CV, a spatial aggregation at 30km should be preferred for RFA.

Based on the results described in Section 4, the methods adopted generally allowed a reliable zonation of the teleconnection-extreme rainfall correlation. This is due to two elements: the temporal and spatial aggregation of the data. First, the temporal

295   aggregation through sliding time windows allow to consider the statistics of the extreme rainfall during time, instead of the rainfall depth themselves. In this way, it is possible to filter out inter-annual variability of the seasonality and magnitude of the annual maxima and focus on the decadal precipitation statistics. Second, the spatial aggregation into tiles allows to obtain more reliable values of the rainfall statistics. This produces a smoothing local effect, that could be due to data fragmentation and

noise and enhance geographical pattern recognition (see Figure 3). The choice of w_t for spatial aggregation should be carefully

300   conducted, as it mainly depends on the morphology of the study area, and the density and location of the rain gauges. Generally, considering a number of different values for w_t is useful to analyze the reliability of the detected correlation patterns.

A different approach for spatial aggregation could be the one described in Castellarin et al. (2024), where overlapping tiles are used. However, including each gauged site in multiple tiles could result in eccessive smoothing effect of the orographic effect over the correlation field. Thus, this solution should be preferred only in case of scarce density of gauges network.

305   A key aspect of the proposed approach is its high adaptability. In fact, the same methodology with an appropriate parametrization could be used to study the influence of teleconnections on several environmental variables, such as AMS of floods or temperature or wind. Moreover, raster maps of the correlation field (as the ones in Figure 4) could be used as descriptors of the drivers of an environmental variable, and adopted as input of predictive raster-based models (e.g., for prediction of flood susceptibility, as in Magnini et al., 2023).

310   5.2   Non-stationary regional frequency analysis

Looking at the results of the second phase of our study (Figure 5), two main points are of general interest. First, the range of variability of the expected percentiles with non-stationary models depending on teleconnections is very significant, confirming what observed locally by other authors. Second, the regional dependency of rainfall statistics on teleconnections can be successfully exploited locally for frequency analysis. This is a useful improvement over present literature, as it allows to obtain

315   the μ_{st} and $L-CV_{st}$ even where observations are not locally available. Moreover, the observed variability of the expected maximum rainfall with given return periods could be used as teleconnection-informed uncertainty range for the design of hydraulic structures.

The results in Figure 6 show that an incontrovertible evaluation of which model is the most accurate is impossible. In fact, each metric, or test, evaluates differently the goodness-of-fit and may provide wrong indication of the best model (e.g., see

320   Laio et al., 2009; Ashkar and Aucoin, 2012; Ashkar and Ba, 2017). In general, it is positive that two metrics out of three (i.e., RML and Δ_{AD}) indicate that the non-stationary framework fits the data better for most of the stations analyzed. Nevertheless, it is not surprising that one of the metrics ($\Delta_{TN.SW}$) suggests to select the stationary approach in half of the cases. In fact, the framework we adopted for non-stationary RFA is based on the strong assumption that the same type of function (i.e., a polynomial function) can represent the teleconnection-statistic relationships within all the study area. Indeed, this approxima-

325   tion is sufficiently accurate in some stations, while being not adatp in others. This is probably also the reason why the GEV_1 models perform generally better than the GEV_2 ones. In fact, the polynomial approximation may fit the data better when these are collected locally (as for the $\boldsymbol{\mu}_{\mathbf{tel}}$-$\boldsymbol{\mu}_{\mathbf{st}}$ case) than when they result from tile-wise averaging (as for the $\boldsymbol{\mu}_{\mathbf{tel}}$-$\boldsymbol{L}$-$\boldsymbol{CV}_{\mathbf{st}}$ case). Accordingly, the GEV_3 models derive their goodness or badness from the sum of the GEV_1 and GEV_2 contributes. This leads them to be the best models for some stations, and the worst for others (see the extension of the GEV_3 boxplots in Figure 6.a

330   and 6.c), depending on the goodness of the polynomial approximation for μ and L-CV.

To summarize, the metrics showed in Figure 6 can be considered as encouraging, as they detect better goodness-of-fit of non-stationary RFA models than the stationary ones, despite the low-complexity of the function adopted for modelling the

dependence on teleconnections. It is important to underline that the aim of the present research is to investigate the potential of teleconnections as independent variables in RFA models, and not to propose a specific method for RFA. Indeed, the RFA results depend on a number of parameters, that are the widths of STWs for temporal aggregation of the teleconnection indexes (w_{tel}) and AMS (w_{AMS}), and the resolution for spatial aggregation (w_t). The choice of these parameters can be done by a careful sensitivity analysis for defining the spatial field of the teleconnection-statistic correlation (see Sections 2 and 4.2). The case of setting a framework for RFA is very different and much more complex, as it requires the formalization of a function of extreme rainfall statistics depending on teleconnection indexes. In this case, one should decide not only the shape of this function, but also the way its parameters vary in space and should be estimated. In our study, we adopted a simple framework, as this function has a limited number of parameters and the same shape (i.e., polynomial) in all the spatial domain. We showed a hierarchical RFA approach where the parameters of the polynomial functions are fitted at-site for the μ and at 30km-tiles for the L-CV. Our analyses overall suggest that even with a simple RFA framework, the use of teleconnections as dependent variables to describe the extreme rainfall regime may increase the accuracy in frequency modelling.

However, different approaches are indeed possible. First, the best resolution for spatial aggregation and the shape of the teleconnection-statistic function should be carefully evaluated for each specific case. Second, a more complex teleconnection-statistics function could be defined. A possible approach is the one proposed by Magnini et al. (2024), which leverages neural networks' capabilities to obtain functions whose parameters depend on the location of the considered site and other morphoclimatic descriptors. Indeed, the implementation and discussion of more sophisticated RFA methods to exploit teleconnections' informative content is complex, and should be addressed by future studies.

6 Conclusions

A growing number of recent studies show how large scale climatic indexes (or teleconnections) can be used as covariates to increase reliability of local frequency analysis of rainfall extremes across diverse geographical regions worldwide (e.g., Fauer and Rust, 2023; Ouarda et al., 2020; Ragno et al., 2018). It is theoretically possible to extend these methods to regional frequency analysis (RFA), but the teleconnection-extreme rainfall dependency at a regional scale should be first investigated. Beside its usefulness for correct estimation of the design rainfall for engineering applications, this topic is still not well addressed in the literature.

In the present study, we propose a framework to assess the link between teleconnections and the frequency regime of rainfall extremes at a regional scale, in order to perform climate-informed regional frequency analysis.

The approach is tested for a large and climatically diverse region in northern Italy. Our dataset consists of 680 annual maximum series (AMS) of hourly and daily (i.e., 1 and 24 hours durations) rainfall depth, recorded between 1921 and 2022. We select six global climate indexes, known to have significant correlation with local climate variability over the study area (Caroletti et al., 2021; Criado-Aldeanueva and Soto-Navarro, 2020; Romano et al., 2022): the North Atlantic Oscillation, Pacific Decadal Oscillation (PDO), East Atlantic – West Russia pattern, El Niño Southern Oscillation, Mediterranean Oscillation Index (MOI),

and Western Mediterranean Oscillation Index.

The main steps of the proposed framework can be summarized as follows. First, we define sliding time windows in order to obtain time series of teleconnections and statistics of annual maxima (i.e. L-moments of AMS, which in our case were used characterized the distribution of sub-daily rainfall extremes). Second, we discretize the study area into tiles where the

370 L-moments are averaged into regional predictions. Then, we evaluate the correlation of teleconnections with time series of spatially gridded L-moments. Finally, we show a preliminary application of climate-informed regional frequency analysis of rainfall extremes, where the L-moments are modelled as functions of WeMOI. We selected WeMOI as it shows a notably higher number of significant correlations with the statistics of sub-daily rainfall maxima than the other teleconnection indexes. Application to different regions and hydroclimatic contexts may result in a higher relevance of other teleconnections. More-

375 over, the relationship between WeMOI and L-moments of extreme rainfall shows clear spatial patterns across the study area, whose robustness is confirmed by their limited sensitivity to the chosen grid resolution and the partial agreement with previous studies (Caroletti et al., 2021; Romano et al., 2022). As well, this is coherent with the known spatial variability of precipitation regimes over the region.

The proposed approach is simple and easily reproducible, yet it is new with respect to the existing literature. In fact, while

380 most authors investigated the correlation between the teleconnections and the raw AMS, we consider the L-moments. This, in combination with spatial discretization of the domain, allows us to focus on the relationship between the teleconnections and the extreme rainfall regime, instead of the extreme values themselves, whose seasonality and interannual variability can affect the correlation analysis. Beside the preliminary nature of the RFA application, three commonly used metrics (e.g., Ashkar and Ba, 2017) detect overall an increase in goodness-of-fit with respect to a stationary approach, in line with previous studies

385 (Nerantzaki and Papalexiou, 2022), shows that teleconnections may be useful covariates in a regional a framework.

Overall, our research suggests promising pathways for climate-informed local and regional frequency analysis of rainfall extremes, and our methodology is highly adaptable to different environmental variables, such as floods and temperature.

Data availability. The rainfall data used in the present study are part of the dataset I2-RED (Mazzoglio et al., 2020), available under authorization at the website https://doi.org/10.5281/zenodo.4269509. Teleconnections indexes data are available at https://psl.noaa.gov/data/

390 climateindices/list/ and https://crudata.uea.ac.uk/cru/data/.

Author contributions. AM: conceptualization, investigation, methodology, software, writing – original draft, writing – review and editing. VP: methodology, writing – original draft, writing – review and editing, validation. AC: conceptualization, methodology, writing – original draft, writing – review and editing, validation, supervision.

Competing interests. The author declare they have no known competing financial interest or personal relationship that could have appeared

395 to influence the work reported in this paper

Acknowledgements. The work was supported by European Climate, Infrastructure and Environment Executive Agency (CINEA), under grant number 101069928 — LIFE21-IPC-IT-LIFE CLIMAX PO. The Authors thankfully acknowledge the use of free and open-source software, in particular Python (Van Rossum and Drake, 2003) and R (R Core Team, 2024). Also, sincere gratitude is due to Paola Mazzoglio and Pierluigi Claps, for giving access to the dataset I2-RED (Mazzoglio et al., 2020).

References

Anderson, T.W. and Darling, D.A., 1952. Asymptotic theory of certain "goodness-of-fit" criteria based on stochastic processes. The Annals of Mathematical Statistics, 23, 193–212. doi:10.1214/aoms/1177729437

Ashkar, F., Aucoin, F., 2012. Choice between competitive pairs of frequency models for use in hydrology: a review and some new results. Hydrological Sciences Journal 57, 1092–1106. https://doi.org/10.1080/02626667.2012.701746

Ashkar, F., Ba, I., 2017. Selection between the generalized Pareto and kappa distributions in peaks-over-threshold hydrological frequency modelling. Hydrological Sciences Journal 62, 1167–1180. https://doi.org/10.1080/02626667.2017.1302089

Bardossy, A., Plate, E.J., 1992. Space-time model for daily rainfall using atmospheric circulation patterns. Water Resources Research 28, 1247–1259. https://doi.org/10.1029/91WR02589

Barnston, A.G., Livezey, R.E., 1987. Classification, Seasonality and Persistence of Low-Frequency Atmospheric Circulation Patterns. Mon. Wea. Rev. 115, 1083–1126. https://doi.org/10.1175/1520-0493(1987)115<1083:CSAPOL>2.0.CO;2

Belkhiri, L., Krakauer, N., 2023. Quantifying the effect of climate variability on seasonal precipitation using Bayesian clustering approach in Kebir Rhumel Basin, Algeria. Stoch Environ Res Risk Assess 37, 3929–3943. https://doi.org/10.1007/s00477-023-02488-z

Bonsal, B., Shabbar, A., 2008. Impacts of Large-Scale Circulation Variability on Low Streamflows over Canada: A Review. Canadian Water Resources Journal 33, 137–154. https://doi.org/10.4296/cwrj3302137

Caroletti, G.N., Coscarelli, R., Caloiero, T., 2021. A sub-regional approach to the influence analysis of teleconnection patterns on precipitation in Calabria (southern Italy). Intl Journal of Climatology 41, 4574–4586. https://doi.org/10.1002/joc.7087

Castellarin, A., Magnini, A., Kyaw, K.K., Ciavaglia, F., Bertola, M., Blöschl, G., Volpi, E., Claps, P., Viglione, A., Marinelli, A., Vogel, R.M., 2024. Frequency of Italian Record-Breaking Floods over the Last Century (1911–2020). Atmosphere 15, 865. https://doi.org/10.3390/atmos15070865

Castellarin, A., Burn, D.H., Brath, A., 2001. Assessing the effectiveness of hydrological similarity measures for flood frequency analysis. Journal of Hydrology 241, 270–285. https://doi.org/10.1016/S0022-1694(00)00383-8

Chen, N., Thual, S., Hu, S., 2019. El Niño and the Southern Oscillation: Observation, in: Reference Module in Earth Systems and Environmental Sciences. Elsevier, p. B978012409548911766X. https://doi.org/10.1016/B978-0-12-409548-9.11766-X

Cheng, L., AghaKouchak, A., 2014. Nonstationary Precipitation Intensity-Duration-Frequency Curves for Infrastructure Design in a Changing Climate. Sci Rep 4, 7093. https://doi.org/10.1038/srep07093

Conte, M., Giuffrida, A., Tedesco, S., 1991. The Mediterranean Oscillation [L'Oscillazione Mediterranea]. Memorie - Societa Geografica Italiana 46, 115–124.

Criado-Aldeanueva, F., Soto-Navarro, J., 2020. Climatic Indices over the Mediterranean Sea: A Review. Applied Sciences 10, 5790. https://doi.org/10.3390/app10175790

Das, S., 2022. Performance of a multi-parameter distribution in the estimation of extreme rainfall in tropical monsoon climate conditions. Nat Hazards 110, 191–205. https://doi.org/10.1007/s11069-021-04942-z

Das, J., Jha, S., Goyal, M.K., 2020. On the relationship of climatic and monsoon teleconnections with monthly precipitation over meteorologically homogenous regions in India: Wavelet & global coherence approaches. Atmospheric Research 238, 104889. https://doi.org/10.1016/j.atmosres.2020.104889

435 El Adlouni, S., Ouarda, T.B.M.J., 2009. Joint Bayesian model selection and parameter estimation of the generalized extreme value model with covariates using birth-death Markov chain Monte Carlo. Water Resources Research 45, 2007WR006427. https://doi.org/10.1029/2007WR006427

Fauer, F.S., Rust, H.W., 2023. Non-stationary large-scale statistics of precipitation extremes in central Europe. Stoch Environ Res Risk Assess 37, 4417–4429. https://doi.org/10.1007/s00477-023-02515-z

440 Gabriele, S., Arnell, N., 1991. A hierarchical approach to regional flood frequency analysis. Water Resources Research 27, 1281–1289. https://doi.org/10.1029/91WR00238

González-Pérez, A., Álvarez-Esteban, R., Penas, A., Del Río, S., 2022. Analysis of recent rainfall trends and links to teleconnection patterns in California (U.S.). Journal of Hydrology 612, 128211. https://doi.org/10.1016/j.jhydrol.2022.128211

Hengl, T., 2007. A Practical Guide to Geostatistical Mapping of Environmental Variables, EUR 22904 EN. Office for Official Publications
445 of the European Communities, Luxembourg (Luxembourg.

Hosking, J.R.M., Wallis, J.R., 1997. Regional Frequency Analysis: An Approach Based on L-Moments, 1st ed. Cambridge University Press. https://doi.org/10.1017/CBO9780511529443

Jayaweera, L., Wasko, C., Nathan, R., Johnson, F., 2023. Non-stationarity in extreme rainfalls across Australia. Journal of Hydrology 624, 129872. https://doi.org/10.1016/j.jhydrol.2023.129872

450 Jenkinson, A.F., 1955. The frequency distribution of the annual maximum (or minimum) values of meteorological elements. Q.J Royal Met. Soc. 81, 158–171. https://doi.org/10.1002/qj.49708134804

Jones, P.D., Jonsson, T., Wheeler, D., 1997. Extension to the North Atlantic oscillation using early instrumental pressure observations from Gibraltar and south-west Iceland. Int. J. Climatol. 17, 1433–1450. https://doi.org/10.1002/(SICI)1097-0088(19971115)17:13<1433::AID-JOC203>3.0.CO;2-P

455 Krichak, S.O., Alpert, P., 2005. Decadal trends in the east Atlantic-west Russia pattern and Mediterranean precipitation. Int. J. Climatol. 25, 183–192. https://doi.org/10.1002/joc.1124

Krichak, S.O., Breitgand, J.S., Gualdi, S., Feldstein, S.B., 2014. Teleconnection–extreme precipitation relationships over the Mediterranean region. Theor Appl Climatol 117, 679–692. https://doi.org/10.1007/s00704-013-1036-4

Krichak, S.O., Kishcha, P., Alpert, P., 2002. Decadal trends of main Eurasian oscillations and the Eastern Mediterranean precipitation.
460 Theoretical and Applied Climatology 72, 209–220. https://doi.org/10.1007/s007040200021

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

Laio, F., 2004. Cramer–von Mises and Anderson-Darling goodness of fit tests for extreme value distributions with unknown parameters. Water Resources Research 40, 2004WR003204. https://doi.org/10.1029/2004WR003204

465 Lopez-Bustins, J.A., Arbiol-Roca, L., Martin-Vide, J., Barrera-Escoda, A., Prohom, M., 2020. Intra-annual variability of the Western Mediterranean Oscillation (WeMO) and occurrence of extreme torrential precipitation in Catalonia (NE Iberia). Nat. Hazards Earth Syst. Sci. 20, 2483–2501. https://doi.org/10.5194/nhess-20-2483-2020

Magnini, A., Lombardi, M., Bujari, A., Mattivi, P., Shustikova, I., Persiano, S., Patella, M., Bitelli, G., Bellavista, P., Lo Conti, F., Tirri, A., Bagli, S., Mazzoli, P., Castellarin, A., 2023. Geomorphic flood hazard mapping: from floodplain delineation to flood hazard characteriza-
470 tion. Hydrological Sciences Journal 68, 2388–2403. https://doi.org/10.1080/02626667.2023.2269909

Magnini, A., Lombardi, M., Ouarda, T.B.M.J., Castellarin, A., 2024. AI-driven morphoclimatic regional frequency modelling of sub-daily rainfall-extremes. Journal of Hydrology 631, 130808. https://doi.org/10.1016/j.jhydrol.2024.130808

Marra, F., Armon, M., Borga, M., Morin, E., 2021. Orographic Effect on Extreme Precipitation Statistics Peaks at Hourly Time Scales. Geophysical Research Letters 48, e2020GL091498. https://doi.org/10.1029/2020GL091498

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

Martin-Vide, J., Lopez-Bustins, J., 2006. The Western Mediterranean Oscillation and rainfall in the Iberian Peninsula. Intl Journal of Climatology 26, 1455–1475. https://doi.org/10.1002/joc.1388

Mazzoglio, P., Butera, I., Claps, P., 2020. I2-RED: A Massive Update and Quality Control of the Italian Annual Extreme Rainfall Dataset.
480 Water 12, 3308. https://doi.org/10.3390/w12123308

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

Ouachani, R., Bargaoui, Z., Ouarda, T., 2013. Power of teleconnection patterns on precipitation and streamflow variability of upper Medjerda Basin. Intl Journal of Climatology 33, 58–76. https://doi.org/10.1002/joc.3407

485 Ouarda, T.B.M.J., Charron, C., 2019. Changes in the distribution of hydro-climatic extremes in a non-stationary framework. Sci Rep 9, 8104. https://doi.org/10.1038/s41598-019-44603-7

Ouarda, T.B.M.J., Charron, C., St-Hilaire, A., 2020. Uncertainty of stationary and nonstationary models for rainfall frequency analysis. Intl Journal of Climatology 40, 2373–2392. https://doi.org/10.1002/joc.6339

Ouarda, T.B.M.J., Yousef, L.A., Charron, C., 2019. Non-stationary intensity-duration-frequency curves integrating information concerning
490 teleconnections and climate change. Intl Journal of Climatology 39, 2306–2323. https://doi.org/10.1002/joc.5953

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

R Core Team, 2024. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Ragno, E., AghaKouchak, A., Love, C.A., Cheng, L., Vahedifard, F., Lima, C.H.R., 2018. Quantifying Changes in Fu-
495 ture Intensity-Duration-Frequency Curves Using Multimodel Ensemble Simulations. Water Resources Research 54, 1751–1764. https://doi.org/10.1002/2017WR021975

Rasouli, K., Scharold, K., Mahmood, T.H., Glenn, N.F., Marks, D., 2020. Linking hydrological variations at local scales to regional climate teleconnection patterns. Hydrological Processes 34, 5624–5641. https://doi.org/10.1002/hyp.13982

Redolat, D., Monjo, R., Lopez-Bustins, J.A., Martin-Vide, J., 2019. Upper-Level Mediterranean Oscillation index and seasonal variability of
500 rainfall and temperature. Theor Appl Climatol 135, 1059–1077. https://doi.org/10.1007/s00704-018-2424-6

Renard, B., Lang, M., Bois, P., Dupeyrat, A., Mestre, O., Niel, H., Sauquet, E., Prudhomme, C., Parey, S., Paquet, E., Neppel, L., Gailhard, J., 2008. Regional methods for trend detection: Assessing field significance and regional consistency. Water Resources Research 44, 2007WR006268. https://doi.org/10.1029/2007WR006268

Ríos-Cornejo, D., Penas, Á., Álvarez-Esteban, R., Del Río, S., 2015. Links between teleconnection patterns and precipitation in Spain.
505 Atmospheric Research 156, 14–28. https://doi.org/10.1016/j.atmosres.2014.12.012

Romano, E., Petrangeli, A.B., Salerno, F., Guyennon, N., 2022. Do recent meteorological drought events in central Italy result from long-term trend or increasing variability? Intl Journal of Climatology 42, 4111–4128. https://doi.org/10.1002/joc.7487

Royston, P. Approximating the Shapiro-Wilk W-test for non-normality. Stat Comput 2, 117–119 (1992). https://doi.org/10.1007/BF01891203

Salinas, J.L., Castellarin, A., Kohnová, S., Kjeldsen, T.R., 2014. Regional parent flood frequency distributions in Europe – Part 2: Climate
510 and scale controls. Hydrol. Earth Syst. Sci. 18, 4391–4401. https://doi.org/10.5194/hess-18-4391-2014

Shapiro, S.S. and Wilk, M.B., 1965. An analysis of variance test for normality (complete samples). Biometrika, 52 (3–4), 591–611.
doi:10.1093/biomet/52.3-4.591

Van Rossum, G., Drake, F.L., 2003. Python Language Reference Manual, A Python manual. Network Theory Limited.

Zhang, Y., Wallace, J.M., Battisti, D.S., 1997. ENSO-like Interdecadal Variability: 1900–93. J. Climate 10, 1004–1020.
515   https://doi.org/10.1175/1520-0442(1997)010<1004:ELIV>2.0.CO;2