ORIGINAL PAPER

A unified statistical framework for detecting trends in multi-timescale precipitation extremes: application to non-stationary intensity-duration-frequency curves

Guillaume Chagnaud[1] · Geremy Panthou[1] · Théo Vischel[1] · Juliette Blanchet[1] · Thierry Lebel[1]

Abstract

There is a large agreement that global warming induces changes of precipitation regimes of different nature and amplitude depending on the timescale considered. This question is of special concern regarding extreme rainfall that might have critical socio-environmental consequences. A unified framework is proposed here for detecting trends in extreme rainfall. It is based on the GEV distribution, whose parameters depend both on a simple scaling formulation to account for multiple time durations of rainfall and on time to account for the non-stationarity deriving from climatic trends. The implementation of the model is illustrated in the Sahel region by analyzing 30 in situ rainfall series of 28 years measured at time-steps from 2 to 24 h. While the separate analysis of the point series proves inconclusive for detecting trends at any of the time-steps considered, the inclusion of all the series and time-steps into the proposed unified model allows trends to be detected at a high level of confidence (p-value $< 1\%$). This trend essentially appears in the scale parameter of the regional GEV distribution, involving a 15 to 20% increase of the 10-year rainfall in 28 years, and a 23 to 30% increase of the 100-year rainfall. The main advantages of the proposed framework are (i) its parsimony, allowing for reducing the uncertainty associated with the model inference; (ii) its capacity for detecting trends either in the mean and/or in the variability of the extreme events; and (iii) its ability for producing non-stationary Intensity-Duration-Frequency curves that are coherent over a range of durations of accumulation.

Keywords Trend detection · Precipitation extremes · Non-stationarity · Intensity-duration-frequency curves

1 Introduction

As reviewed recently by Allan et al. (2020), a warmer Earth climate induces changes of precipitation regimes through different mechanisms. There is high confidence based on robust physics and idealised CO_2 forcing experiments that global mean precipitation increases at $\sim 2.5\%$ K^{-1} of global warming, termed hydrological sensitivity (Allen and Ingram 2002; Fläschner et al. 2016). However, a given value of mean annual precipitation change may hide different patterns from a rainfall regime perspective, depending on the timescale and region considered. In regions where the rainfall decadal variability is especially

strong, as is the case for monsoon climates heavily dependent on sea surface temperatures, the mean annual rainfall signal linked to the hydrological sensitivity may not even be detectable due to a counter-acting phase of the oceanic temperature oscillations. Yet, at the same time, hydro-climatic intensification (as termed by Giorgi et al. 2011) might take place, driven by fast adjustment effects linked to the Clausius-Clapeyron thermodynamics and with changes in the atmospheric circulation shaping regional discrepancies (see, e.g., Pfahl et al. 2017). Such intensification might entail longer dry spells and more intense rain events (Trenberth et al. 2003). As a consequence, it is widely expected that the regime of extremes will be more affected by climate warming than the central part of the rainfall distributions (Trenberth 2011; Westra et al. 2014), at least in the initial phase of the thermodynamical re-equilibrium. Such a trend has already been detected in global and regional observational records (see, e.g., Alpert et al. 2002; Donat et al. 2013; Panthou

✉ Guillaume Chagnaud
  guillaume.chagnaud@univ-grenoble-alpes.fr

[1] University Grenoble Alpes, CNRS, IRD, Grenoble INP, IGE, 38000 Grenoble, France

et al. 2014a). Since extremes are bearing critical stakes for populations, detecting changes in the extreme rainfall events distribution is both a challenge for scientists and a pressing issue for countries and population that will have to face the consequences. Moreover, the timescale to consider depends on different factors (e.g., the type of water-related natural hazard for which a protection is required, the nature of the infrastructure to be designed, the dynamics of the considered watershed). This paper thus proposes a unified statistical framework for analyzing the modification of extreme rainfall regimes, coherently across timescales as well as across regions. In addition to providing insights into climate change, this unified framework is of interest for the purpose of comparison between observations and extreme events simulated in climate models (on this issue, see, e.g., Kharin et al. 2007; Wehner et al. 2010). It can also meet the recurrent needs of water managers and decision-makers in charge of planning rainfall hazard protection systems, with the noticeable originality of accounting for temporal non-stationarity.

1.1 Accounting for trends in the regime of precipitation extremes

Given the all-time impact of droughts and floods on populations and even on civilizations, it is not astonishing that hydrologists and climatologists alike has long focused attention on extreme rainfall. Characterizing the upper tail of rainfall distributions is the statistical key to predicting the probability of occurrence of critically extreme rainfall, a challenge that was taken up by many researchers in different ways until the so-called Extreme Value (EV) Theory emerged. This framework was initially conceived for analyzing single point series, but further developments led to dealing with rainfall extremes in a regional context and for an array of timescales, with two issues in mind: (i) reducing the uncertainty of the estimated parameters and (ii) maintaining some coherency between the EV distributions inferred for different timescales. In that respect, the simple scaling (SS) approach provides a nice framework for pooling together samples of rainfall data aggregated over various durations, allowing for both a more robust inference of the distribution parameters and the coherency between timescales (Koutsoyiannis et al. 1998; Menabde et al. 1999). This simple scaling framework can be applied either to a single rainfall record (see, e.g., Blanchet et al. 2016; Sane et al. 2018) or in a regional context (Panthou et al. 2014b). While time stationarity is usually assumed when applying the EV theory to a particular dataset, framing it in a non-stationary context allows for:

– Detecting whether there is a significant trend in the intensity of extreme rainfall over a range of time-steps

(i.e., durations of accumulation), and whether these trends preserve the simple scaling relationship,
– Allowing for a more robust inference of the time-varying parameters of the EV model.
– Providing a more realistic estimation of the return period of observations.
– Building a more relevant framework for designing infrastructures through the computation of non-stationary IDF curves.

1.2 Methodological issues related to trend detection and model inference

The methodological issues to be solved for adjusting a non-stationary EV model on a given dataset pertain to two categories: one is related to detectability of potential rainfall trends and the other to parameter inference. Detecting a trend in a series is therefore a matter of signal-to-noise ratio. As regarding the inference methodological issues, they are especially sensitive given the number of parameters of any non-stationary EV model. Accounting for non-stationarity involves adding to a standard stationary model a number of parameters (as few as possible) deemed sufficient to represent this non-stationarity to represent this non-stationarity. The main consequence is an obvious loss of robustness of the inference process. A compromise has thus to be found so as to test which parameters seem to be time-varying—if any—while limiting their number to a minimum, keeping in mind that quality checked series of daily and sub-daily rainfall extremes are not so numerous in many regions of the world, especially when looking for long and complete series in order to increase the signal-to-noise ratio. Detecting which parameters are significantly non-stationary in such a context is thus challenging and imposes using sophisticated methodologies.

1.3 Implication of non-stationarity for the estimation of return levels and IDF curves

A return level (i.e., the value of a quantile for a given frequency of non-exceedance or return period) is a statistical concept that assumes the temporal stationarity of the sampled population. In case of temporal drifting, return levels should be provided with a sticker precising at which temporal horizon they apply (Salas and Obeysekera 2014): a 100-year return level of a 24-h rainfall might have been 100 mm at the end of the twentieth century and 110 mm 20 years later. The same remark applies to IDF curves commonly used for designing infrastructures and operating flood warning systems, since they are basically a compilation of return levels—corresponding to a given return period—over an array of durations of accumulation. In many (if not in most) regions, stationary IDF curves

established under a stationarity assumption are still in use (e.g., Svensson and Jones 2010). Yet, several studies have recently revealed the potential non-stationarity of rainfall extremes. For instance, Cheng and AghaKouchak (2015) explore the impact of ignoring non-stationarity of rainfall extremes at 5 stations in the USA exhibiting increasing trends of annual maxima at several durations. They conclude that "the shorter the duration the larger the differences between the non-stationary and stationary extremes." With the same purpose in mind, Sarhadi and Soulis (2017) redefine the notion of return level in a changing climate, thus yielding a time-varying risk of failure. Using an hourly rainfall dataset located in the Great Lakes area in a Bayesian framework, they show that an EV model with time-varying location and scale parameters best describes the extreme rainfall behavior at most sites. They similarly conclude that basing IDF curves on a stationarity assumption leads to an under-estimation of the frequency of occurrence of extreme rainfall events, more prominent for small durations (1–2 h) and large return periods (50 years). In the two previously mentioned studies, a non-stationary EV distribution is adjusted to each sample of specific duration of accumulation. To our knowledge, the study of Ouarda et al. (2019) is the only one that builds a unified non-stationary IDF (NS-IDF) model. Applying the NS-IDF model to two stations in Ontario and California, they found that accounting for the time- and/or climate index-dependence of extreme rainfall produces more accurate IDF models.

1.4 Objectives of the paper

Based on the various issues discussed above, the paper has three main goals:

- Deriving a global and coherent framework for modelling the distribution of rainfall extreme values for a set of durations of accumulation, in a possibly non-stationary context (Section 2).
- Providing an efficient methodology for identifying the parameters of the model that best account for temporal non-stationarity, if it does prove significant. The two matters—detecting a trend and inferring the best parameters to represent it—are narrowly linked and jointly explored through a bootstrap Generalized Linear Model (GLM) approach (Section 3). Its practical implementation (i.e., choosing a proper class of model to deal with data in a given region) is illustrated in Section 4 through a case study located in the West African Sahel, where rainfall variability is high at both the interannual and decadal scales.
- Testing the added value of the proposed approach at detecting trends and providing more realistic return

levels and IDF curves, by implementing it in a step by step procedure whereby the complexity of the model is progressively increased from an initial point and stationary class of models to the most complex class of non-stationary regional multi-timescale models (Section 5).

Section 6 is then discussing some remaining methodological issues while Section 7 summarizes the main advances of the present paper.

2 Theoretical background of the non-stationary GEV simple scaling (NS-GEV-SS) framework

2.1 The stationary GEV-SS framework

2.1.1 GEV framework and block maxima sampling

The Extreme Value Theory (EVT) framework, initially developed by Fisher and Tippett (1928), completed by Gnedenko (1943) and thoroughly described in Coles (2001), proposes two ways of sampling extreme rainfalls: the Block Maxima (BM) consists of defining blocks of equal length and extracting the maximum of each block, and the Peak-Over-Threshold (POT) wherein values exceeding a given threshold are extracted. BM sampling is often preferred to POT sampling, for two reasons: first, it avoids dealing with the choice of a threshold value and, secondly, it ensures the independence of the members of the resulting sample (for a given duration), the drawback being the possibly small size of the sample. Let I denote the random variable associated with the annual maximum of rainfall accumulation; i then denotes a value taken by I in its range of variation. Within the BM analysis framework, the EVT asserts that for large enough blocks, a sample of independent and identically distributed maxima (see Leadbetter 1974, for a generalization to short-term–dependent random variables) converges in distribution towards one of the Generalized Extreme Value (GEV) distributions, whose cumulative distribution function F is:

$$F(i; \mu, \sigma, \xi) = Pr(I \leq i)$$
$$= \exp\left\{-\left[1 + \xi\left(\frac{i-\mu}{\sigma}\right)\right]^{-1/\xi}\right\}, \quad (1)$$
$$\text{for } \xi \neq 0 \text{ and } 1 + \xi\left(\frac{i-\mu}{\sigma}\right) > 0$$

where μ, $\sigma > 0$, and ξ are the location, scale, and shape parameters, respectively. The two first parameters govern the location and the deviation of the distribution about the center, respectively, while the shape parameter defines the

behavior of the distribution tails, leading to distinguish three type of distributions:

– The heavy-tailed Frechet distribution for $\xi > 0$,
– The bounded reversed Weibull distribution for $\xi < 0$,
– The light-tailed Gumbel distribution for $\xi = 0$.

The associated probability density function f reads:

$$
f(i; \mu, \sigma, \xi) = \frac{1}{\sigma} \left[1 + \xi \left(\frac{i - \mu}{\sigma} \right) \right]^{\frac{-1}{\xi} - 1}
$$
$$
\times \exp \left\{ - \left[1 + \xi \left(\frac{i - \mu}{\sigma} \right) \right]^{-1/\xi} \right\} \quad (2)
$$

2.1.2 Temporal scale invariance

The second aspect upon which relies the unified statistical framework is the temporal scale invariance of rainfall (e.g., Schertzer and Lovejoy 1987), i.e., the scaling of rainfall statistical properties over a given range of durations of accumulation. A variety of mathematical models have been developed to account for this fractal property (Langousis and Veneziano 2007; Langousis et al. 2009; Veneziano and Yoon 2013). Here, the strict sense simple scaling (Gupta and Waymire 1990, see also Menabde et al. 1999) is used to model the temporal scale invariance of the annual maximum of rainfall intensities:

$$
I(D) \stackrel{d}{=} I(D_0) \left(\frac{D}{D_0} \right)^{\eta} \quad (3)
$$

where $I(D)$ is the annual maximum rainfall intensity at duration D, D_0 a reference duration, and η a scalar comprised between -1 and 0 to be inferred from the data (the negative sign results from the decrease in intensity with increasing duration of aggregation); $\stackrel{d}{=}$ stands for the equality in distribution, with one consequence being that the normalized random variable at duration D, $I(D) \times (D/D_0)^{-\eta}$, has the same distribution as the random variable at the reference duration $I(D_0)$. In terms of moments, this leads to (Gupta and Waymire 1990):

$$
E[I(D)] = E[I(D_0)] \left(\frac{D}{D_0} \right)^{\eta} \quad (4)
$$

And to a more general extent, for any moment of order q:

$$
E^{(q)}[I(D)] = E^{(q)}[I(D_0)] \left(\frac{D}{D_0} \right)^{\eta q}
$$
$$
\Rightarrow \log\{E^{(q)}[I(D)]\} = \eta q \log \left(\frac{D}{D_0} \right) + \log\{E^{(q)}[I(D_0)]\} \quad (5)
$$

It can thus be empirically verified whether the moments of order q have a linear relationship (in log-log space) with durations, indicating that the simple scaling model is an adequate representation of the data over the considered range of durations (Bougadis and Adamowski 2006). In such a case, several samples collected at the same station for various durations of accumulation can be scaled so as to be pooled into a larger, unique sample.

2.1.3 GEV-SS model

Combining the extreme value and the simple scaling theories leads to a GEV-Simple Scaling (GEV-SS) framework in which the statistical parameters of the GEV distribution for any duration D are straightforwardly deduced from a reference distribution at D_0 through the following equations (e.g., Blanchet et al. 2016; Yeo et al. 2020):

$$
\mu_D = \mu_0 \left(\frac{D}{D_0} \right)^{\eta}, \ \sigma_D = \sigma_0 \left(\frac{D}{D_0} \right)^{\eta} \ \text{and} \ \xi_D = \xi \quad (6)
$$

Note that to remain in a simple scaling model, ξ must be considered invariant with duration. Compared to a discrete duration approach, the number of parameters to be estimated is reduced, thus increasing the robustness of the model inference. To summarize, the stationary GEV and GEV-SS models are fully described by a set of three and four parameters, respectively. Let θ be this set of parameters; we have:

– $\theta = (\mu, \sigma, \xi)$ for the GEV model,
– $\theta = (\eta, \mu, \sigma, \xi)$ for the GEV-SS model.

2.2 Non-stationary GEV and GEV-SS models

The GEV-based framework previously described can be used in a non-stationary context in view of detecting trends in extremes of precipitation. The GEV and GEV-SS models are readily made non-stationary by expressing one or more of their parameters as a function of time or any other physically based, time-dependent covariate. For instance, the global mean surface temperature can be used when one is interested in detecting the possible effect of global warming on the precipitation regime (e.g., Westra et al. 2013). The dependency function may be regular and monotonous (linear, polynomial, etc.), periodic (sinusoidal), or non-regular (existence of break points). For the sake of limiting the number of parameters to be inferred—and thus the robustness of the overall inference process—it is however clear that in most cases, simple linear functions are preferred in case of regularity or simple trigonometric functions in case of periodicity. Choosing a type of function and a relevant covariate is a matter of prior knowledge, theoretical understanding, and/or preliminary empirical assessment (e.g., with a moving-window approach). Once the time-dependency formulation is set, the non-stationary GEV (NS-GEV) and non-stationary GEV-SS (NS-GEV-SS)

models are described in their most general form by the following set of time-dependent parameters θ_t:

- $\theta_t = (\mu_t, \sigma_t, \xi_t)$ for the NS-GEV model,
- $\theta_t = (\eta_t, \mu_t, \sigma_t, \xi_t)$ for the NS-GEV-SS model.

2.3 Increasing the signal-to-noise ratio through a regional approach

GEV and GEV-SS models can be applied to either point-wise (PGEV, PGEV-SS) or regional (RGEV, RGEV-SS) data. In this latter case, the regional information is used to enlarge the sample size and thus to increase the signal-to-noise ratio and the probability of detecting trends more robustly. The regional information can be accounted for by including spatial covariates within the modelling framework (see, e.g., Blanchet and Davison 2011; Panthou et al. 2012) or simply by pooling the series together in case where regional independence is assumed.

2.4 Return levels and IDF curves

The NS-GEV model is expressed as a function of time-varying parameters, so do all the quantiles, whose general formulation derives from Eq. 1 as:

$$i(T; \mu, \sigma, \xi) = \mu + \frac{\sigma}{\xi} \left\{ \left[-\log\left(1 - \frac{1}{T}\right) \right]^{-\xi} - 1 \right\} \quad (7)$$

where T stands for the return period, i.e., corresponding to the probability of a value being exceeded on average once over a time period of length T. Using the expression of the scaled parameters (Eq. 6) and accounting for non-stationarity gives the time-dependent return level expression for any duration D for which the model is valid and for any return period T. Insofar as ξ is often considered as invariant with t because of its large sampling variance for samples of usual size (as is the case in this study), this reads:

$$i_t(D, T) = \left(\frac{D}{D_0}\right)^{\eta_t} \left\{ \mu_{0,t} + \frac{\sigma_{0,t}}{\xi} \left\{ \left[-\log\left(1 - \frac{1}{T}\right) \right]^{-\xi} - 1 \right\} \right\} \quad (8)$$

This equation is used to compute the non-stationary IDF (NS-IDF) curves, synthesizing the relationship existing between EV rainfall intensities for various durations and probabilities of occurrence (or similarly, return periods), thus providing a time-varying integrated view of the structure of the annual maxima for a range of durations. When η is further considered as time-invariant and μ and σ are represented by linear functions of time, Eq. 8 provides $i_t(D, T)$ as a linear function of time, yielding a linear displacement of the IDF curves. This is shown in the
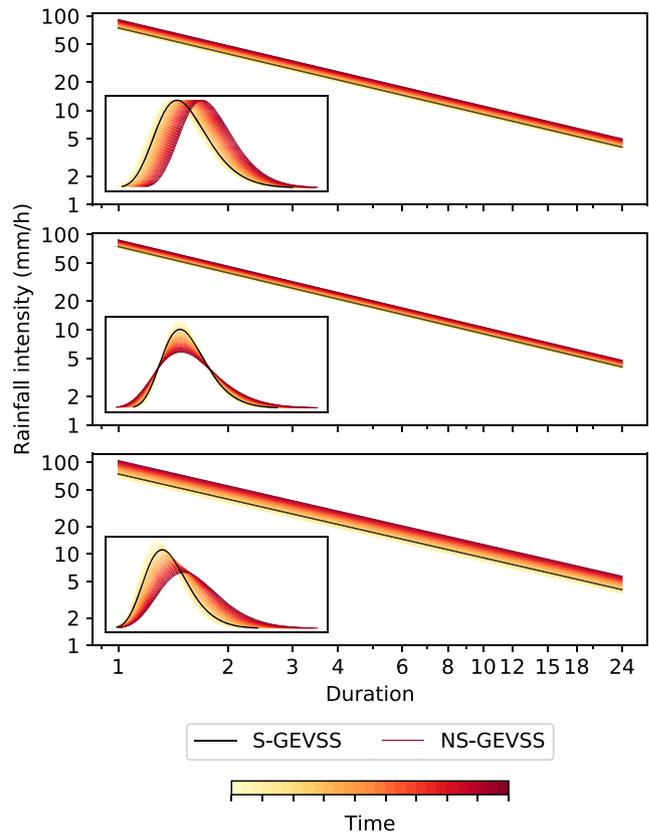
Fig. 1 Schematic view of a 20%/decade increase in either the location (top), scale (center), or both location and scale (bottom) GEV distribution parameters, assuming η and ξ to be stationary in time. The considered durations are plotted on the horizontal axis while the ordinate axis represents the rainfall intensity (mm/h), both in logarithmic coordinates. The IDF curves correspond to the 10-year rainfall for one station (Sandidey) of the case study analyzed in Sections 4 and 5. The corresponding probability density functions (PDF) are shown in boxes. Black lines are the stationary IDF curves and PDF, while the color shading represents the yearly evolution of the non-stationary IDF curves and PDF

illustrative example given in Fig. 1, where a 20%/decade increase in one or two of the GEV parameters is prescribed, yielding an upward displacement of the 10-year rainfall IDF curves.

3 Implementation process: parameter inference, trend detection, and uncertainty assessment

3.1 Inference

The model parameters are estimated with the maximum likelihood (ML) method. Despite convergence and bias issues, particularly for samples of size < 100 (Hosking et al. 1985), it is readily made non-stationary (Katz et al. 2002). In a stationary context, the likelihood function L of a

GEV distribution with probability density function f (Eq. 2) reads:

$$L(\theta) = \prod_{k=1}^{N} f(i_k|\theta)$$

$$= \prod_{k=1}^{N} \frac{1}{\sigma} \left[1 + \xi \left(\frac{i_k - \mu}{\sigma} \right) \right]^{-1-\frac{1}{\xi}} \quad (9)$$

$$\times \exp \left\{ - \left[1 + \xi \left(\frac{i_k - \mu}{\sigma} \right) \right]^{-1/\xi} \right\}$$

where i_k is the realization of the random variable I at year k, $\theta = (\mu, \sigma, \xi)$ is the set of the GEV distribution parameters, and N is the number of years with observations. It is more convenient, for optimization purpose, to work with the log-likelihood function l:

$$l(\theta) = \log[L(\theta)] = \sum_{k=1}^{N} \log[f(i_k|\theta)] \quad (10)$$

In the case of a GEV-SS model, the log-likelihood function of the GEV distribution at the duration D reads:

$$l(\theta_D) = \sum_{k=1}^{N} \log[f(i_{D,k}|\theta_D)]$$

$$= -\sum_{k=1}^{N} \log \sigma_D \quad (11)$$

$$- \sum_{k=1}^{N} \left(1 + \frac{1}{\xi} \right) \log \left[1 + \xi \left(\frac{i_{D,k} - \mu_D}{\sigma_D} \right) \right]$$

$$- \sum_{k=1}^{N} \left\{ \left[1 + \xi \left(\frac{i_{D,k} - \mu_D}{\sigma_D} \right) \right]^{-1/\xi} \right\}$$

To infer the set of parameters of the NS-GEV-SS model using the whole dataset, the log-likelihood functions corresponding to each duration are added, with the parameters being scaled to the reference duration according to Eq. 6 while non-stationarity is accounted for including time-dependent parameters. The function to be optimized then reads:

$$l(\theta) = -m \sum_{k=1}^{N} \log \sigma_{0,k} - \sum_{D} \log \left(\frac{D}{D_0} \right) \sum_{k=1}^{N} \eta_k$$

$$- \sum_{D} \sum_{k=1}^{N} \left(\frac{1+\xi}{\xi} \right) \log \left[1 + \xi \left(\frac{i_{D,k} \left(\frac{D}{D_0} \right)^{-\eta_k} - \mu_{0,k}}{\sigma_{0,k}} \right) \right]$$

$$- \sum_{D} \sum_{k=1}^{N} \left[1 + \xi \left(\frac{i_{D,k} \left(\frac{D}{D_0} \right)^{-\eta_k} - \mu_{0,k}}{\sigma_{0,k}} \right) \right]^{-1/\xi} \quad (12)$$

where m is the number of durations. This definition of the log-likelihood function is likely misspecified as it

implicitly assumes independence of the annual maxima across years and durations. This latter assumption might not hold as maxima at consecutive durations can originate from the same event. However, taking into account this dependence is tricky and would not lead to improvement, particularly when dealing with the distribution margins, as is the case here (Blanchet et al. 2016, see also Sebille et al. 2017). An IDF model that accounts for the dependence between samples of several durations is provided by Tyralis and Langousis (2019). Here, we use the definition of Eq. 12 and minimize its opposite (the negative log-likelihood, NLLH hereafter) with the Nelder-Mead optimization method (Nelder and Mead 1965) implemented in the Python package Scipy.

3.2 Model selection

3.2.1 Visual diagnosis

The model fit to the data can be visually checked with a quantile-quantile plot. The data must be transformed to stationary residuals ϵ, e.g., to stationary Gumbel residuals as follows (Katz 2013):

$$\epsilon_{D,k} = \frac{1}{\xi} \log \left[1 + \xi \left(\frac{i_{D,k} - \mu_{D,k}}{\sigma_{D,k}} \right) \right], \ k \in 1, 2, \ldots, N \quad (13)$$

where $\mu_{D,k}$, $\sigma_{D,k}$, and ξ are the NLLH estimates of the non-stationary model parameters at duration D and year k.

3.2.2 Trend significance

Assessing whether the additional parameter(s) representing non-stationarity improves the model fit to the data is achieved by testing for the significance of the non-stationary models against the null hypothesis of (i) no trend or (ii) a smaller number of time-varying parameters. Because of the dependence between durations, the model significance is assessed empirically, using a semi-parametric bootstrap resampling method. The test statistic is defined as:

$$\Delta NLLH_{M_{null} - M_{alter}} = l(\theta_{alter}) - l(\theta_{null}) \quad (14)$$

where l is the time-dependent log-likelihood function defined in Eq. 12 and $\theta_{alter} = (\eta^a, \mu^a, \sigma^a, \xi^a)$ and $\theta_{null} = (\eta^n, \mu^n, \sigma^n, \xi^n)$ are the sets of parameters of the alternative model M_{alter} and the null model M_{null}, respectively. Since M_{null} is nested in M_{alter}, i.e., M_{null} can be obtained by setting one or more parameters of M_{alter} to 0, $\Delta NLLH_{M_{null} - M_{alter}} \geq 0$. Following Katz et al. (2002), the bootstrap has to be performed on independent and identically distributed data (see also Kharin and Zwiers 2005; Cannon 2010). The hypothesis testing procedure goes as follows:

(i) The alternative model (M_{alter}) is fitted on the original sample, giving the alternative model parameters θ_{alter}.

(ii) The residuals of M_{alter} are computed for each duration D and year k with the alternative model parameters according to:

$$
\begin{aligned}
\epsilon_{D,k} &= \frac{1}{\xi^a} \log\left[1 + \xi^a \left(\frac{i_{D,k} - \mu_{D,k}^a}{\sigma_{D,k}^a} \right) \right] \\
&= \frac{1}{\xi^a} \log\left[1 + \xi^a \left(\frac{i_{D,k} \left(\frac{D}{D_0} \right)^{-\eta_k^a} - \mu_{0,k}^a}{\sigma_{0,k}^a} \right) \right], \\
&\quad k \in 1, 2, \ldots, N
\end{aligned}
\tag{15}
$$

the $\epsilon_{D,t}$ thus follow a stationary Gumbel distribution.

(iii) The rows of the original $N \times m$ matrix of residuals (where N is the number of observed years and m the number of considered durations) are drawn with replacement to form a bootstrapped sample of residuals ($\tilde{\epsilon}_{D,k}$), thus preserving the dependence between durations. This resampling procedure removes any trend in the original sample of residuals. Any remaining trend in the bootstrapped sample is therefore solely due to randomness.

(iv) The bootstrapped residuals $\tilde{\epsilon}_{D,k}$ are transformed back into rainfall intensities using the parameter estimates of the null model:

$$
\begin{aligned}
\tilde{i}_{D,k} &= \mu_{D,k}^n + \frac{\sigma_{D,k}^n}{\xi^n} \left[\exp(\tilde{\epsilon}_{D,k} \, \xi^n) - 1 \right] \\
&= \left(\frac{D}{D_0} \right)^{\eta_k^n} \left\{ \mu_{0,k}^n + \frac{\sigma_{0,k}^n}{\xi^n} \left[\exp\left(\tilde{\epsilon}_{D,k} \, \xi^n \right) - 1 \right] \right\}, \\
&\quad k \in 1, 2, \ldots, N
\end{aligned}
\tag{16}
$$

The $\left\{ \tilde{i}_{D,k} \right\}$ thus follow M_{null}.

(v) Both the null and the alternative models are fitted on the bootstrapped sample to obtain M'_{alter} and M'_{null} with parameters $\theta'_{alter} = (\eta'^a, \mu'^a, \sigma'^a, \xi'^a)$ and $\theta'_{null} = (\eta'^n, \mu'^n, \sigma'^n, \xi'^n)$, respectively.

(vi) The negative log-likelihood difference between models M'_{null} and M'_{alter}:

$$
\Delta NLLH_{M'_{null} - M'_{alter}} = l\left(\theta'_{alter} \right) - l\left(\theta'_{null} \right)
\tag{17}
$$

is computed and stored.

(vii) Operations (iii) to (vi) are repeated 200 times to obtain the empirical cumulative distribution of the $\Delta NLLH_{M'_{null} - M'_{alter}}$.

(viii) The p-value is the probability of having a value at least as extreme as the original $\Delta NLLH_{M_{null} - M_{alter}}$ according to the empirical distribution.

3.3 Uncertainty

The uncertainty on the parameters (and the quantiles) is estimated using the same semi-parametric bootstrap approach as in Section 3.2.2, with the only difference that in step (iv) the bootstrapped sample of residuals is transformed back into annual maximum of rainfall intensities with the parameter estimates of the alternative model M_{alter} ($\eta^a, \mu^a, \sigma^a, \xi^a$). The M'_{alter} parameter estimates are stored and the 90% confidence intervals of each parameter are given by the 5th–95th percentile range of these bootstrap parameter estimates. Mélèse et al. (2018) show the ability of this bootstrap resampling method to provide reliable uncertainty assessment of the GEV-SS model parameters (and derived quantiles) in a stationary context.

4 Case study description

4.1 Data

The AMMA-CATCH Niger (ACN) rainfall observatory (Lebel et al. 2009), located in southwestern Niger (1.6°–3° E, 13°–14° N), has been documenting the evolution of the Sahelian rainfall regime over the past 28 years, at high space-time resolution. This dataset has no equivalent in sub-Saharan Africa and provides a unique opportunity to test the robustness of the proposed modelling framework for a region of high space-time variability of rainfall (Balme et al. 2006), which makes non-stationarity particularly difficult to detect (Panthou et al. 2013).

The observing network is made of 30 tipping-bucket rain gauges covering a $\sim$16,000-km^2 area and providing 5-min rainfall records for the 1990–2017 period (see Galle et al. (2018) for a description of the data collection and pre-processing). Panthou et al. (2014b) have shown that the scale invariance property holds for the ACN dataset at durations ranging from 1 to 24 h, although a deviation appeared for $D = 1$ h, leading us to consider durations ranging from 2 to 24 h in this study (see Appendix 1 for the verification of the SS validity for the case study). Note also that a Koutsoyiannis formulation of the temporal scale invariance model (Koutsoyiannis et al. 1998) was tested on this range of durations and displayed very similar results (not shown). Time series of annual maxima are constructed by aggregating the 5-min rainfall amounts with a rolling window of length ranging from 2 to 24 h, and averaging over this window. The rolling mean ensures that rainfall events are not split. The annual maxima are then extracted for each duration, thus yielding for each station m vectors of length N of sliding maxima (van Montfort 1990) with m the number of durations D ($D \in \{2, 4, 6, 12, 18, 24h\}$) and N the number of observed years. Sampling annual maxima

ensures the independence of data within each duration sample.

Figure 2 shows the mean and standard deviation of the series of annual maximum rainfall intensities of the 30 ACN stations. For $D = 2h$ (Fig. 2a), the means range from 23 to 30 mm/h and the standard deviations range from 5.3 to 11.1 mm/h, showing the large sampling effect over the study area, with stations spaced by a few tens of kilometers having means of annual maximum at the two sides of the overall range of values (23 mm/h and 30mm/h). This high variability among stations is a clue that a regional approach will better sample the overall variability of the annual maxima than a single station. The same conclusion holds when considering other timescales, as can be seen for $D = 24h$ in Fig. 2b. Also worth noticing in Fig. 2 is the lack of spatial organization, which leads to the use of the regional model described in Section 4.2.3.

A first attempt at estimating the presence of trends in the series of annual maxima is illustrated in Fig. 3. The time series of annual maxima are displayed for Niamey Airport at all selected durations D (Fig. 3a). Linear trend estimates are computed with a Mann-Kendall test applied to each time series. Trends range from 1.44 to 6.13%/decade, yet none of them is significant at the 5% level. Figure 3b shows the trend estimates at the 30 ACN stations for $D = 12h$. Trends range from -11.6 to 16.6%/decade with only one station displaying a significant trend (circled in yellow). Twenty stations out of thirty have a positive trend and the median across the thirty stations is 5.29%/decade, indicating a possible overall increase of the annual maxima over the past 28 years (1990–2017). However, since only one station has a significant trend (furthermore a negative trend), it would be hard to draw any reliable conclusion from this non-parametric approach where point series are analyzed separately.

4.2 Workflow

To address the trend detection issue in this Sahelian case study, three models are gradually built within the GEV-based framework described in Section 2 and compared:

– The NS-PGEV model: It is a non-stationary (NS) GEV model fitted for each point series (P) and for each duration. It aims at evaluating how a GEV with time-varying parameters is able to detect trends on local series for different rainfall durations.
– The NS-RGEV: It differs from the previous NS-PGEV model in that the 30 stations now form a single regional (R) sample. Its aim is to evaluate whether a regional approach changes the ability to detect trends in terms of intensity and significance for the dataset under study.
– The NS-RGEV-SS: Its specificity compared to the previous model is to integrate all the rainfall durations through the Simple Scaling (SS) relationship. Its aim is to evaluate the added value of this scaling integration in trend detection abilities.

Note that in order to assess the significance of the trends (cf. Section 3.2), the stationary counterparts of these three non-stationary models are also fitted (S-PGEV, S-RGEV, and S-RGEV-SS) and referred to as M0 for each model. The experimental protocol and the associated model names are summarized in Fig. 4.

4.2.1 Implementation of non-stationarity

In the proposed parametric approach, the non-stationary GEV parameters are assumed to be linear functions of time. A non-stationary parameter $\nu(t)$ then reads $\nu(t) = \nu_0 + \nu_1 t$ where ν_0 and ν_1 are the intercept and the slope of the GLM, respectively, and t is the covariate, normalized so as to

Fig. 2 The ACN stations with mean (color shading) and standard deviation (circle size) of the annual maxima time series for (**a**) $D = 2h$ and (**b**) $D = 24h$
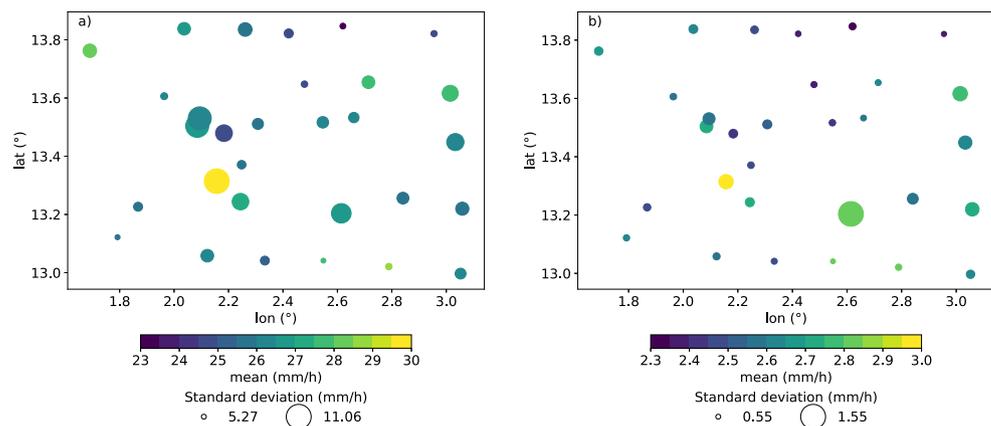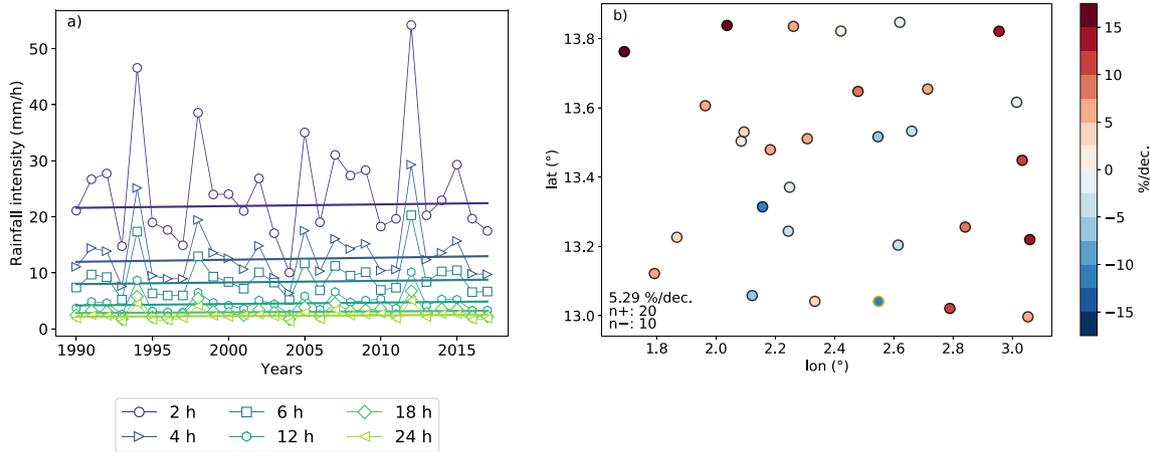
Fig. 3 (**a**) Time series of annual maxima for the range of considered time-steps at Niamey Airport with Mann-Kendall trend estimates. (**b**) Mann-Kendall relative trend (%/decade) estimates of the annual maximum rainfall intensities for $D = 12h$ at the 30 ACN stations. The only station displaying a significant trend at the 5% level is circled in yellow. The median is indicated in the bottom left corner along with the number of stations with positive (n+) and negative (n−) trends

vary between 0 and 1 for optimization purpose. For each of the approaches previously described (namely NS-PGEV, NS-RGEV, and NS-RGEV-SS), three parameterizations of non-stationarity are considered, leading to the three following classes of non-stationary model: M1 where μ only is time-dependent; M2 where σ only is time-dependent; and M3 where both μ and σ are time-dependent. The various combinations of models and non-stationarity parameterizations are summarized in Table 1.

4.2.2 Detection of trend at the point-wise scale: NS-PGEV models

Due to the limited length of the point series, it is assumed for the sake of robustness that in all classes of NS-PGEV models (M0, M1, M2, and M3), the shape parameter ξ of the GEV distribution is fixed to the mean of the stationary shape parameters estimated over all stations and durations. A likelihood ratio test is used to test for the alternative

Fig. 4 Experimental protocol (figure generated with LibreOffice Impress)
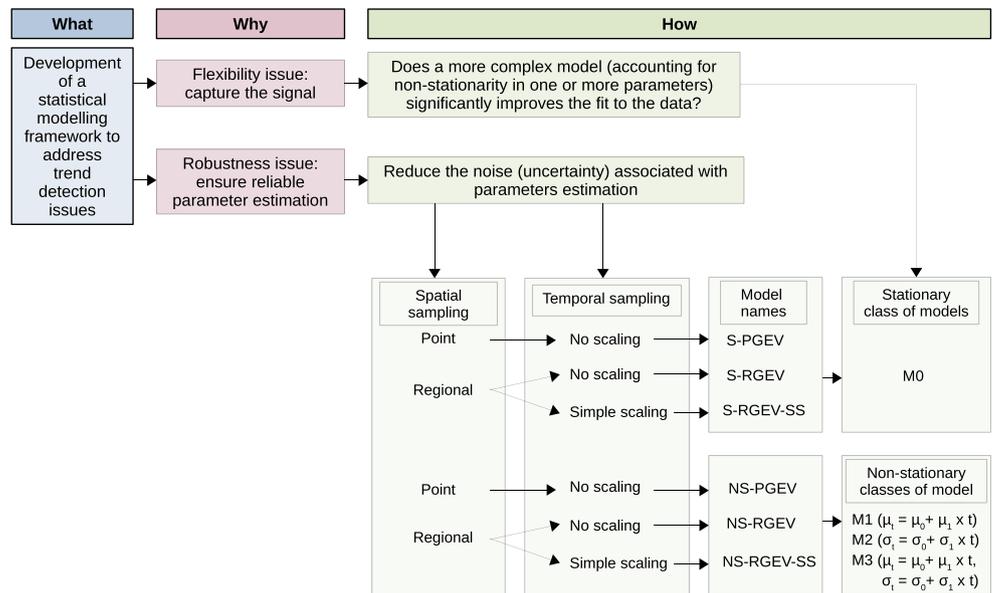
Table 1 Model names, classes, parameters, and degrees of freedom (DoF)

Name	Class	η	μ	σ	ξ	DoF
S-PGEV	M0	–	μ	σ	ξ	3
	M1	–	$\mu_0 + \mu_1 t$	σ	ξ	4
NS-PGEV	M2	–	μ	$\sigma_0 + \sigma_1 t$	ξ	4
	M3	–	$\mu_0 + \mu_1 t$	$\sigma_0 + \sigma_1 t$	ξ	5
S-RGEV	M0	-	μ	σ	ξ	3
	M1	–	$\mu_0 + \mu_1 t$	σ	ξ	4
NS-RGEV	M2	–	μ	$\sigma_0 + \sigma_1 t$	ξ	4
	M3	–	$\mu_0 + \mu_1 t$	$\sigma_0 + \sigma_1 t$	ξ	5
S-RGEV-SS	M0	η	μ	σ	ξ	4
	M1	η	$\mu_0 + \mu_1 t$	σ	ξ	5
NS-RGEV-SS	M2	η	μ	$\sigma_0 + \sigma_1 t$	ξ	5
	M3	η	$\mu_0 + \mu_1 t$	$\sigma_0 + \sigma_1 t$	ξ	6

model M_{alter} (either M1, M2, or M3) significance against the null model M_{null} (here, M0). The test statistics $S = 2 \times (l_{M_{alter}} - l_{M_{null}})$ with l the log-likelihood function, is compared to the α-quantile q_α of the χ^2 distribution with k degrees of freedom, where k is the number of additional degrees of freedom in M_{alter} compared to M_{null}. If S exceeds q_α the alternative model is accepted at level α.

The relative trend (in %/decade) of the non-stationary parameter ν, r_ν is computed from the inferred values of ν_0 and ν_1 as follows:

$$r_\nu = \frac{\nu(t=1) - \nu(t=0)}{\nu(t=0)} \times \frac{10}{N} \times 100$$
$$= \frac{\nu_0 + \nu_1 \times 1 - \nu_0}{\nu_0} \times \frac{10}{N} \times 100 = \frac{\nu_1}{\nu_0} \times \frac{1000}{N} \quad (18)$$

where N is the number of years. The corresponding trend in rainfall intensities can be derived similarly after using Eq. 7 to obtain the quantiles of interest. The total change (in %) of a specific return level value, Δi, over the study period is also computed as:

$$\Delta i = \frac{i(t=1) - i(t=0)}{i(t=0)} \times 100 \quad (19)$$

where i is the return level value obtained with Eq. 8.

4.2.3 Spatial aggregation: NS-RGEV models

Individual point series of 28 annual maxima have clearly a high sampling variance. Insofar as these 30 point series could be pooled to build an aggregate sample, this sampling variance would be significantly diminished. Obviously, given their low spacing (8 km for the nearest stations), the individual rainfall series cannot be considered as being strictly independent. However, rainfall extremes are

associated in this region with convective cells of limited life duration and spatial extent (Mathon and Laurent 2001); annual maxima series are thus more loosely correlated than the annual total series (the median Pearson coefficients of correlation across the stations of our case study are 0.08 and 0.12 for AMS at D=2 h and D=24 h, respectively, and 0.32 for the annual totals, see Appendix 2). Moreover, there is no external factor such as topography or vegetation that could produce some systematic spatial rainfall pattern over the considered domain, as confirmed by the lack of spatial consistency of the mean and standard deviation of the annual maxima (Fig. 2). Therefore, the 30 stations are considered to sample the same extreme rainfall population and the corresponding time series are pooled into one regional sample. The regional sample then becomes a matrix with m columns of EV variables (m being the number of considered durations) and $N \times q$ rows/observations (N being the number of years and q the number of stations). The three classes of NS-RGEV models (M1, M2, M3) are then fitted separately to each of the m series. The significance of each NS-RGEV class of models is tested against the class of stationary RGEV model (M0) using the bootstrap resampling method described in Section 3.2 to account for the spatial dependence. The sample size of each series being increased from N to $N \times q$, the adjustment of each model becomes more robust. Consequently, the shape parameter is relaxed to be estimated within the NLLH minimization.

4.2.4 Temporal aggregation on regional sample: NS-RGEV-SS models

The simple scaling approach provides a way of further enlarging the sample size by pooling multi-duration samples in a single scaled sample. This in turn allows for a more robust estimation of the GEV distribution parameters (Sane et al. 2018), by adjusting a scaled GEV on the scaled multi-duration samples, which is of major importance when one wants to estimate quantiles of large return period. In this procedure, ξ is still considered time-invariant and estimated within the NLLH minimization. In addition, it is also considered in a first step that η is stationary (the validity of this assumption is further examined in Section 6), leaving three classes of NS-RGEV-SS models to be tested and compared, as summarized in Table 1.

5 Results

5.1 Point-wise non-stationary GEV models

Following the workflow presented in Section 4.2, the first step is to fit a NS-PGEV model separately to each of the 180 EV series (30 stations × 6 durations). This yields the

values of the parameter trends corresponding to the best fit for each class of models as well as the associated *p*-value measuring the significance of using a non-stationary model as compared to using a stationary model. The results obtained for each of the 180 series are summarized in Fig. 5, where a color code is used to represent the trend value (from dark blue for large negative trends to dark red for large positive trends) and a black contouring is used to show models significant at the 5% level. The overall impression produced when looking at Fig. 5 is a global lack of coherency. When fitting a M1 model, some stations display a strong negative trend ($r_\mu < 0$), such as Guilahel, while others (such as Yillade) display a strong positive trend for all durations. Some stations display a mix of negative and positive trends depending on the considered duration. There are only 10 out of the 180 M1 models with significant improvement at the 5% level compared to the stationary model. The picture is more coherent for models M2 with a large overall predominance of positive trends ($r_\sigma > 0$). However, with only 17 M2 models displaying

a significant improvement, the number of significant non-stationary models is still low. The fitting of M3 models produces a somewhat smoother pattern for both μ and σ, with the exception of Torodi that displays a 213%/decade increase in the scale parameter. The M3 models bring a significant improvement in only 13 cases. There is only 1 series out of 180 (Kare, 4 h) displaying significance at the 5% level for all three models.

Figure 6 provides a spatial representation of these results for $D = 12$h. Here, also, no clear pattern of negative versus positive trends appears, except maybe for a predominance of positive trends in the north of the study domain, but none of them is significant.

Notwithstanding the fact that 7% only of the non-stationary models are significant, Table 2 provides the median across the 30 stations of μ, σ, and $\{\mu, \sigma\}$ trend estimates for each duration. Again, there is a lack of consistency in these numbers with r_σ ranging from 6.88%/decade for 12 h to 13.2%/decade for 2 h with the M2 models, while it ranges from 12.5%/decade for 6 h

Fig. 5 Trends (%/decade) in the classes of NS-PGEV model parameters (**a**) M1 (μ), (**b**) M2 (σ), (**c**) M3 (μ), and (**d**) M3 (σ) for the range of considered durations and the ACN stations. Stations with significant trends at the 5% level are outlined in black
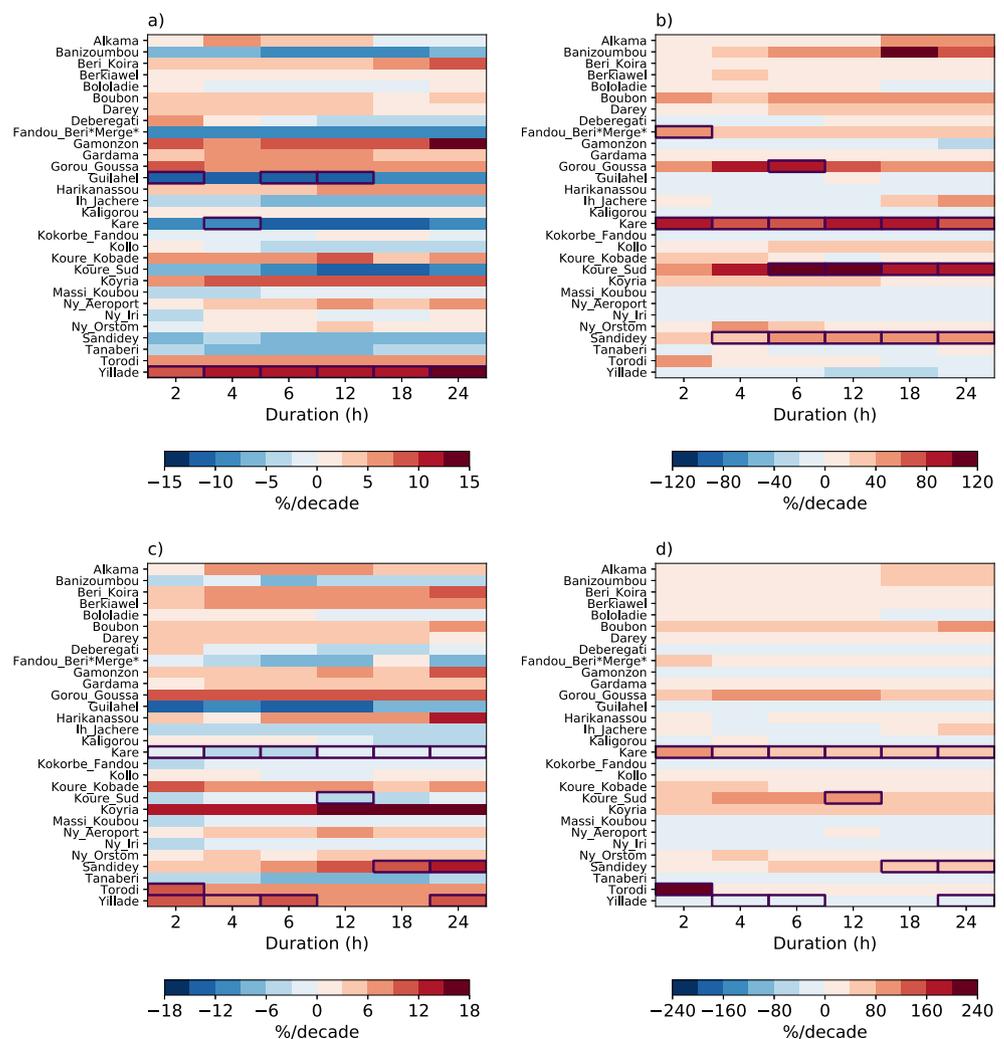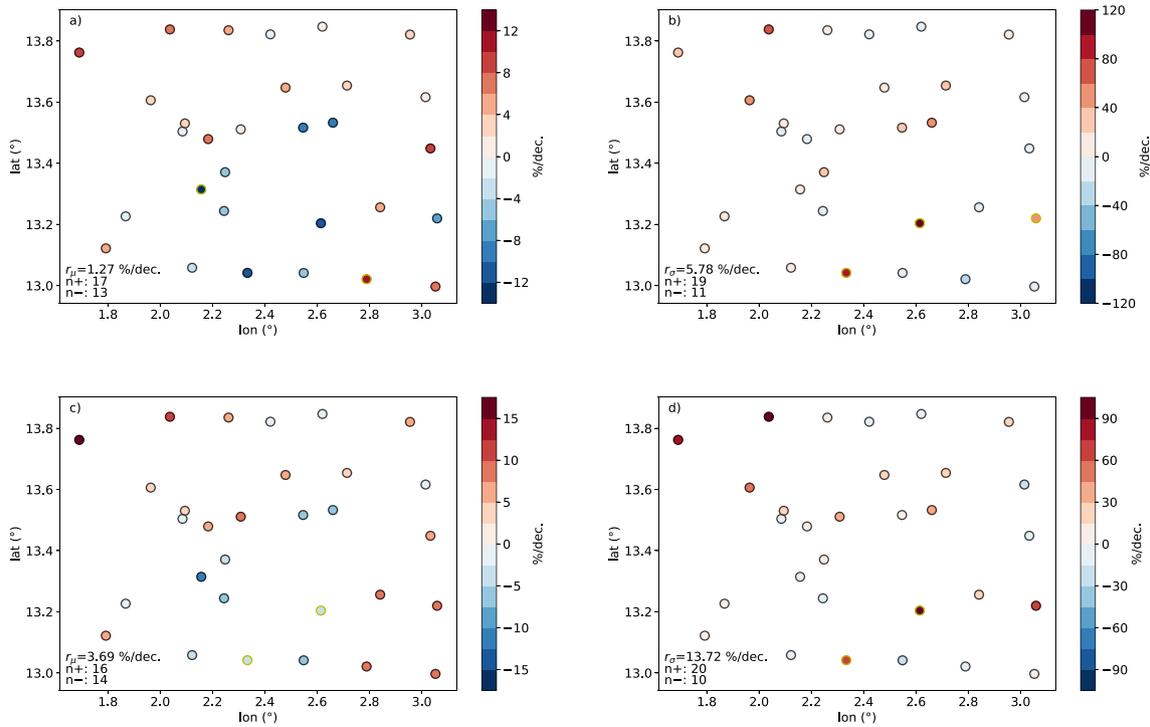
Fig. 6 Trends (in %/decade) in the classes of NS-PGEV model parameters (**a**) M1 (μ), (**b**) M2 (σ), and (**c, d**) M3 (μ, σ) for the duration $D = 12$h. Stations with significant trends at the 5% level are circled in yellow. Values in the bottom left corner are spatial median of the parameter trend (r_ν) and the number of stations with positive (n+) and negative (n–) trends

to 19.6%/decade for 24 h with the M3 models. The total changes of the 2-, 10-, and 100-year return levels (Δi_2, Δi_{10}, Δi_{100}) over the study period are also displayed to provide insights into the changes taking place at the regional scale on these extreme rainfalls: with M1, both the 10- and 100-year return level values underwent the same change over the study, ranging from 0.67 to 2.89 %, while the 2-year return levels underwent slightly larger increase, reaching 4.31% for $D = 4$h. With both M2 and M3 the 100-year return level has increased by a larger amount as compared to the 2- and 10-year return levels. This stronger sensitivity of the large return period values is due to the larger trend on the scale parameter of the GEV distribution (Katz and Brown 1992).

In conclusion, it clearly stems from these results that detecting a statistically significant trend in the EV rainfalls of this region from individual point series is challenging even though there is a fuzzy signal that this trend might be positive and mostly born by the σ parameter of the GEV model, more strongly affecting large return period events.

5.2 Regional non-stationary GEV models

In this section, we assess the expectation of a more powerful (in the statistical meaning) trend detection that could stem from the spatial aggregation. The results summarized in

Table 3 show the considerable improvement of the trend detection power that results from this regional approach, with the class of NS-RGEV models M2 displaying statistically significant trends for all durations against the stationary model (M0). The trends are also significant at the

Table 2 Medians across the ACN stations of the NS-PGEV models M1, M2, and M3 parameter trends r (in %/decade) and total change (%) over the study period in the 2-, 10-, and 100-year return levels

		2 h	4 h	6 h	12 h	18 h	24 h
M1	r_μ	0.65	1.71	0.91	1.27	0.7	0.59
	Δi_2	1.67	4.31	2.30	3.21	1.78	1.50
	Δi_{10}	1.16	2.89	1.48	2.10	1.15	1.0
	Δi_{100}	0.80	2.07	0.98	1.38	0.75	0.67
M2	r_σ	13.2	13.1	9.09	5.78	10.2	6.88
	Δi_2	2.82	2.37	1.91	1.22	2.12	1.47
	Δi_{10}	12.5	11.0	8.86	5.70	10.0	6.84
	Δi_{100}	19.6	17.2	13.8	8.95	15.8	10.7
M3	r_μ	2.2	2.56	2.86	3.69	3.48	3.08
	r_σ	13.8	14.9	12.5	13.7	18.8	19.6
	Δi_2	8.03	8.84	10.7	13.3	11.9	10.8
	Δi_{10}	18.9	17.5	16.9	20.3	23.3	27.0
	Δi_{100}	24.8	25.2	21.6	24.2	32.2	36.7

Table 3 Parameter trends r (in %/decade) and significances (p-value) in comparison with the stationary model M0 of the NS-RGEV models for each considered duration

		2 h	4 h	6 h	12 h	18 h	24 h
M1	r_μ	−0.49	−0.55	−0.85	−0.83	−0.74	−0.16
	p-value	0.76	0.72	0.66	0.65	0.71	0.90
	Δi_2	−1.32	−1.50	−2.27	−2.27	−2.04	−0.60
	Δi_{10}	−0.89	−1.01	−1.51	−1.50	−1.35	−0.39
	Δi_{100}	−0.64	−0.71	−1.05	−1.04	−0.92	−0.27
M2	r_σ	14.2	14.2	14.6	14.0	12.8	11.9
	p-value	**0**	**0**	**0**	**0**	**0**	**0.01**
	Δi_2	3.38	3.35	3.49	3.38	3.15	3.06
	Δi_{10}	14.8	14.8	15.5	15.0	14.0	13.5
	Δi_{100}	22.2	22.4	23.4	22.7	21.2	20.4
M3	r_μ	1.43	1.64	1.58	1.72	1.91	2.53
	r_σ	16.1	16.6	17.0	16.6	15.8	15.8
	p-value	**0.02**	**0.04**	**0.04**	**0.04**	0.06	0.06
	Δi_2	7.74	8.27	8.26	8.83	9.09	11.0
	Δi_{10}	19.8	20.5	21.1	21.5	21.2	23.1
	Δi_{100}	27.7	28.7	29.8	30.0	29.4	31.2

Models significant at the 5% level are indicated in bold. Δi_2, Δi_{10}, and Δi_{100} are the total changes (%) of the 2-, 10-, and 100-year return levels over the study period

5% level with the M3 class of NS-RGEV models for the durations 2 h, 4 h, 6 h, and 12 h. For M1, none of the trends is significant at the 5% level.

The trends in the two classes of NS-RGEV models that prove statistically significant are consistent with the regional vision highlighted in the previous section (Table 2), with (i) positive trends on μ and σ and (ii) larger trend estimates in the parameters of M3 as compared to their M1 and M2

counterparts. A feature worth noticing is the consistency of the trends over the range of durations, most prominent for the M2 and M3 classes of NS-RGEV models. Looking at the changes in return levels, results from the NS-RGEV models M2 and M3 are consistent with those highlighted in Table 2, considering both the orders of magnitude and the stronger increase of large return period events.

5.3 Regional non-stationary GEV-Simple Scaling models

Based on the results of Sections 5.1 and 5.2, we explore whether introducing a timescale invariance relationship makes the trend detection more robust. The evaluation is carried out based on the semi-parametric bootstrap resampling method presented in Section 3.2. Figure 7 shows the empirical CDF obtained for each Null/Alternative configuration that was tested, based on the 200 bootstrap simulations carried out for each configuration. M2 is significant versus M0 at the 1% level, while M3 is significant versus both M0 and M1 at the 2% level (Table 4). This confirms the results obtained in Section 5.2, albeit for a slight increase of the trend detection power.

The parameter trends given in Table 4 are in line with those of the NS-RGEV approach (Table 3). The simple scaling relationship involves that the relative values of the trends are the same for all time-steps for which the relationship holds. The total change, shown in Table 5, increases with the return levels (3.3 to 8.9 for 2-year and 22.6 to 30.1 for 100-year) as a result of the larger trend in the scale parameter. Another meaningful way of looking at the impact of non-stationarity on extreme events is to note that the 10-year return level in 1990 has become a nearly 6-year return level at the end of the study period with the NS-RGEV-SS model M2.

Fig. 7 Empirical cumulative distribution functions of the 200 bootstrapped samples significance test statistic ($\Delta NLLH_{M'_{null}-M'_{alter}}$) for the classes of NS-RGEV-SS models M1 (red), M2 (green), M3 (dark blue) with respect to the null model M0 and of M3 against M1 (light blue) and M2 (purple). Markers indicate the model significance test statistics obtained from the original sample ($\Delta NLLH_{M_{null}-M_{alter}}$, horizontal axis) and the corresponding empirical probability of non-exceedance (1 - p-value, vertical axis)
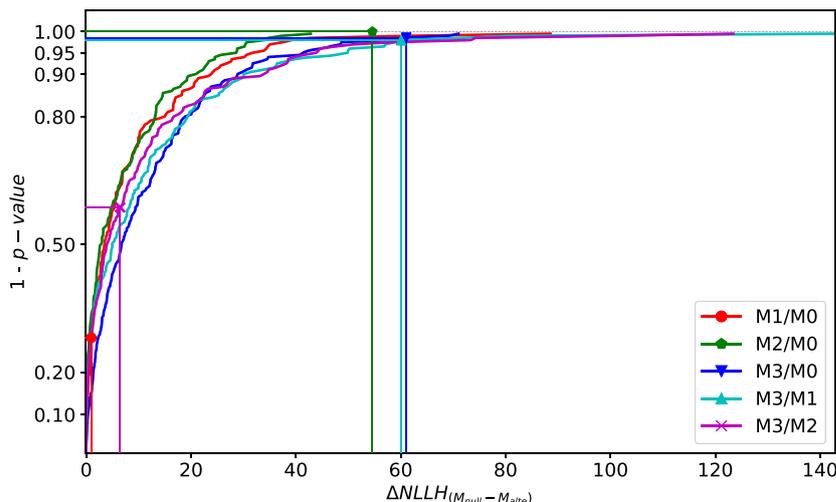
Table 4 Parameter trends r (%/decade) and p-values (significance at the 5% level is indicated in bold) of the two classes of NS-RGEV-SS models for which a significant non-stationarity is detected (M2 and M3)

Models	r_μ	r_σ	p-value	η	ξ
M2	–	14.6	**0.0**	−0.94	0.053
M3	1.9	18.0	**0.014/M0**;	−0.94	0.054
			0.02/M1		
			0.43/M2		

The values of the scaling exponent (η) and the shape parameter (ξ) of the NS-RGEV-SS model are also displayed

5.4 Non-stationary IDF curves

Non-stationary IDF curves are implemented at the regional scale (Fig. 8) as a consistent by-product of the NS-RGEV-SS approach. They provide a synthetic and end-user oriented view of the impact of trends in extreme rainfalls over the considered range of durations ($D \in \{2, 4, 6, 12, 18, 24h\}$) and for various return periods (2, 5, 10, and 30 years). Deriving NS-IDF curves from the NS-GEV-SS model has the advantage of preserving the consistency of annual maxima over time, since rainfall intensities for various durations of accumulation increase at the same rate. The non-stationarity is highlighted comparing the IDF curves at the start of the study period (Fig. 8a, b) with those at the end of the study period (Fig 8c, d). The two selected classes of NS-RGEV-SS model (M2 and M3) lead to an upward shift of IDF curves.

The 90% confidence intervals of the non-stationary return levels are also displayed. It can be noted that the start and the end of the study period approximately correspond to the time when the stationary return levels (thin black lines) become separate from the confidence interval of the non-stationary return levels (color shading). This feature is more prominent for M2 as a result of the narrower confidence intervals associated with this model's return level estimates (see Table 5). It is further investigated with Fig. 9, showing for each year with which level of confidence (ranging from 80 to 99%) the non-stationary return levels do not

Table 5 Total change (%) over the study period of the 2-, 10-, and 100-year return levels (Δi_2, Δi_{10}, and Δi_{100}) derived from the two NS-RGEV-SS models for which a significant non-stationarity is detected (M2 and M3)

Models	Δi_2	Δi_{10}	Δi_{100}
M2	3.3 [1.2; 5.5]	14.8 [4.9; 25.7]	22.6 [7.4; 39.7]
M3	8.8 [-5; 21]	21.4 [0.31; 42.4]	30.1 [1.9; 57.7]

Numbers in brackets are the 90% confidence intervals

overlap with their stationary counterpart: an upper bound of the confidence interval below the stationary return level is shaded in blue and a lower bound of the confidence interval above the stationary value is shaded in red. The larger the degree of separation, the darker the color.

Despite larger return level trend estimates, the larger confidence intervals of the NS-RGEV-SS model M3 make it less significantly different from the stationary return levels than M2. Also worth noticing in Fig. 9a is the dependence to the return period: as this latter increases, the non-stationary return levels become larger than the stationary ones earlier, though at the 80% level of significance only. Although the non-stationary return levels exceed their stationary counterpart at the 80% confidence level most of the time, return levels under stationary assumption may be underestimated compared with present-day non-stationary return levels. Provided the detected trends remain unchanged, future stationary return levels will be even more underestimated.

6 Discussions

While the results presented in Section 5 indicates that combining the regional and simple scaling approaches adds robustness both in the detection of temporal trends and in the estimation of the GEV model parameters for the Sahelian case study, three important points deserve further discussion. The first is related to the number of time-varying parameters to consider. The second takes us back to the signal-to-noise ratio issue introduced in Section 1.2. The third deals with the climatological implications of the trends detected on our specific case study.

6.1 Which parameters should be considered time-varying in the NS-RGEV-SS model?

In Section 4.2, a double assumption was made regarding the time-invariance of both the shape parameter ξ of the GEV and the parameter η of the simple scaling relationship. In order to examine to which extent such a choice is justified, let us first recall that a standard GEV-SS model has four parameters and that the inference of ξ is known to be ill-conditioned for usual sample sizes ranging from, say, 30 to 100. Allowing all four parameters (μ, σ, ξ and η) to vary in time with a simple linear function adds 4 parameters to the inference process, making it fairly difficult to handle from both a methodological and a robustness points of view. This is why in many regional studies (see, e.g., Panthou et al. 2012; Sarhadi and Soulis 2017), ξ is assumed to be invariant over the study area and also invariant in time in non-stationary models (Sarhadi and Soulis 2017). There is less background regarding η since, to our knowledge,

Fig. 8 Stationary (thin black lines) and non-stationary (thick lines) IDF curves derived from the NS-RGEV-SS models M2 (**a**, **c**) and M3 (**b**, **d**) representing the 2-, 5-, 10-, and 30-year return levels in 1990 (**a**, **b**) and 2017 (**c**, **d**). The shading corresponds to the 90% confidence intervals of the non-stationary return levels computed from 200 bootstrapped samples. The parameter values are shown in the bottom left corner with μ and σ expressed in mm/h for $D = 2$h, the first and second values of the time-varying parameters corresponding to the slope (ν_1) and the intercept (ν_0), respectively
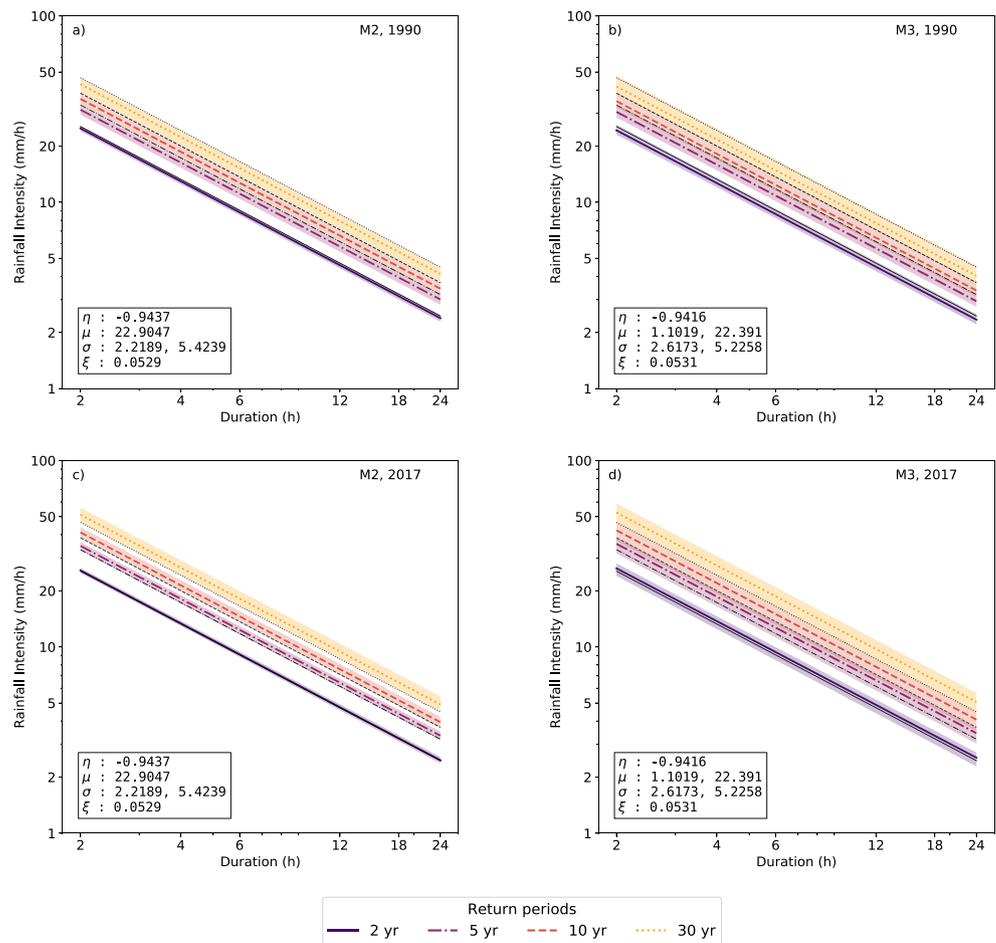
this is the first time that a complete NS-RGEV-SS model is implemented. We hence had no preconceived idea on whether η could really be considered time-invariant, as compared (in terms of significance) to μ and/or σ.

Therefore, a complementary study was launched to test the possible contribution of a time-varying η in accounting for the pattern of non-stationarity across durations of rainfall accumulation. This is achieved by expressing η as a linear
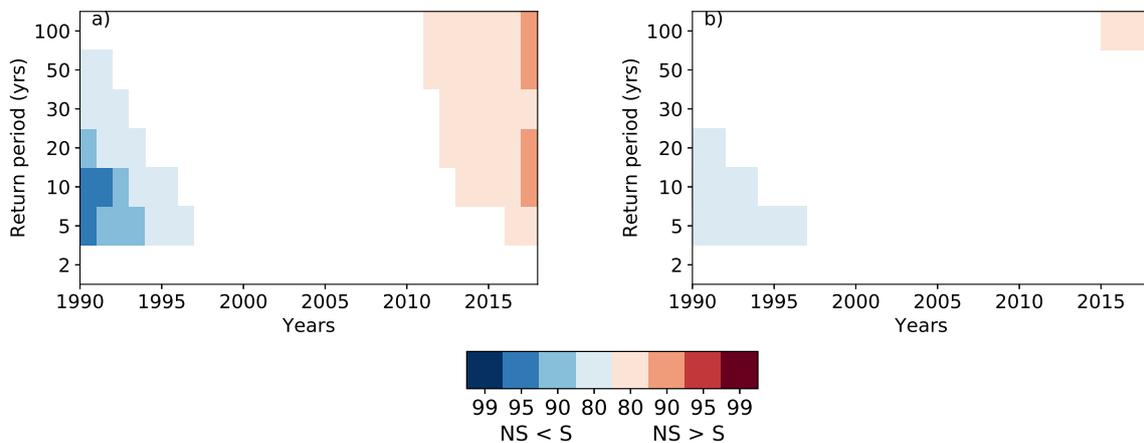
Fig. 9 Non-overlap significance of non-stationary return levels estimated from the NS-RGEV-SS models (**a**) M2 and (**b**) M3 for 2-, 5-, 10-, 20-, 30-, 50-, and 100-year return periods. The confidence intervals are obtained from 200 bootstrapped samples. Blue shading indicates that the upper bound of the non-stationary return level is smaller then the stationary return level while red shading indicates that the lower bound of the non-stationary return level confidence interval exceeds the stationary return levels. Result are shown for 80%, 90%, 95%, and 99% confidence levels (the wider the confidence interval the more distinct the stationary and non-stationary return levels)

Table 6 Additionnal NS-GEV-SS model parameters and degrees of freedom

Models	η	μ	σ	ξ	DoF
M0'	$\eta_0 + \eta_1 t$	μ	σ	ξ	5
M1'	$\eta_0 + \eta_1 t$	$\mu_0 + \mu_1 t$	σ	ξ	6
M2'	$\eta_0 + \eta_1 t$	μ	$\sigma_0 + \sigma_1 t$	ξ	6

Table 7 NS-RGEV-SS model parameter trends r (%/decade) and associated significances

Models	r_η	r_μ	r_σ	p-value	ξ
M0'	0.71	–	–	0.48/M0	0.022
M1'	3.61	−6.5	–	**0.08/M0**; **0.03/M1**	0.044
M2'	0.87	–	15.4	**0.03/M0**; 0.37/M2	0.055

Significance at the 5% level is indicated in bold

function of time, leading to test three additionnal NS-RGEV-SS models (Table 6):

The significance of these three NS-RGEV-SS models is also assessed with the bootstrap resampling method (Section 3.2). The p-values are shown in Table 7. The largest improvement from adding a temporal trend in η is clearly for the class of models where μ is time-varying while σ is time-invariant (M1'), with a p-value reaching 0.08 against M0 and 0.027 against M1 (itself not significant at the 5% level compared to the stationary model). The NS-RGEV-SS model M2' is significant compared to M0, but its additional flexibility is not significant compared to its M2 counterpart. Therefore, the significant additional flexibility of M2 and M2' as compared to the stationary model is carried out by the time-dependence release of σ, not by adding the time-dependence in η, further confirmed by the fact that M0' is not significantly better than M0.

For the sake of robustness of the estimation process and in the absence of any overwhelming evidence that such an hypothesis is not acceptable, it is probably justified to assume the stationarity of η. In cases where the non-stationarity of η could not be totally excluded, it is important to keep in mind some implications:

– It questions the relevance of studies assuming its stationarity when downscaling IDF curves from daily data, as suggested by Cannon and Innocenti (2019).
– It might entail a change in the typology of rainfall systems which could have become more stationary, and/or of larger horizontal extent. One might also think of a change in the relative contribution of the various types of rainy systems, e.g., localized vs organized convection, or a change in the share of convective events versus stratiform fronts.

As often when using statistical models to deal with natural phenomena, there is an ontological conflict between the simplification required for a robust parameter estimation and the complexity of the underlying physical processes. It is the context (amount of available data, prior physical knowledge allowing to hierarchize processes depending on the environment considered, goal of the statistical representation) that will lead the user in choosing a more or less complex model.

6.2 Is the signal-to-noise ratio improved when using a NS-RGEV-SS model?

Figure 10 compares the total increase of the 10- and 100-year return level (%) over the study period inferred from the NS-RGEV and NS-RGEV-SS models, along with the associated 90% confidence intervals. It clearly stems from Fig. 10a and c that the six NS-RGEV/M2 models (one for each time-step) and the NS-RGEV-SS/M2 model provide extremely coherent estimates of the total change for the 10-year and 100-year return levels—around 15% for the 10-year rainfall and in the range of 20 to 25% for the 100-year rainfall. The corresponding signal-to-noise ratio is in the range of 0.7–0.75 for the 10-year level and close to 1 for the 100-year level. The regional/simple scaling approach is thus capturing the non-stationarity of extreme rainfall intensities in a more parsimonious way with 5 parameters describing the full signal agaisnt 24 (4×6) for the unscaled regional approach. By comparison, the M3 class of models yields (i) a somewhat larger estimate of the total changes (20–25% for the 10-year rainfall and around 30% for the 100-year rainfall); (ii) a smaller signal-to-noise ratio ranging from 0.5 to 0.6. It should also be noted that the median of the NS-RGEV-SS/M3 bootstrap simulations is significantly lower than the value estimated from the fitting to the original sample. This traduces a larger instability and may be seen as a direct consequence of M3 models having an extra trend parameter, as compared to M2 models. With no claim to generality, it can thus be concluded that the NS-RGEV-SS/M2 class of models is, in this particular case, the most appropriate for properly describing the positive trend in the intensity of extreme rainfall. Finally, these signal-to-noise ratio considerations are valid under the hypothesis of weakly auto-correlated data. This latter assumption has been checked for the ACN dataset computing the auto-correlations at lag-1 of the 180 EV times series of our case study (mean value $r = -0.08$).

6.3 Are the strongest events becoming stronger?

A major outcome concerning our case study is the much larger temporal trend detected for the scale parameter σ of the GEV distribution as compared to the small trend
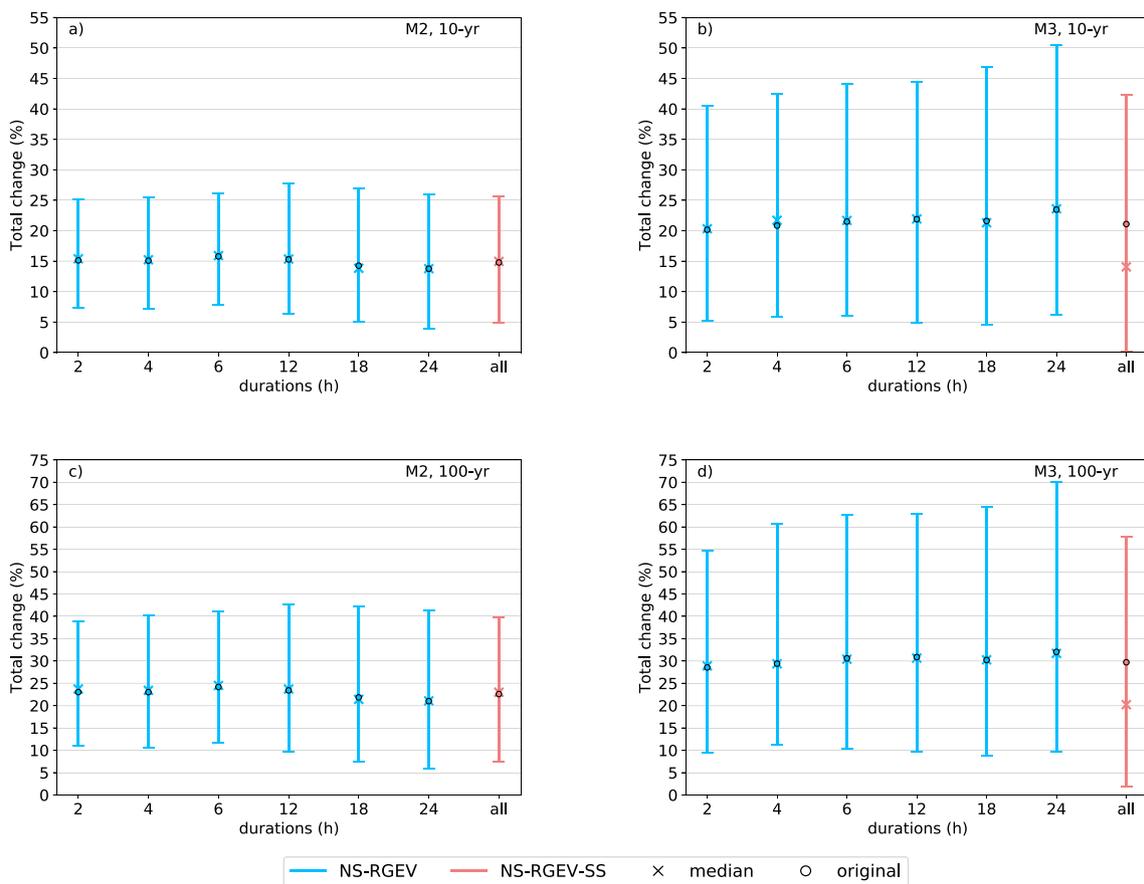
Fig. 10 Total change (%) over the study period in (**a, b**) 10- and (**c, d**) 100-year return levels inferred from the M2 (**a, c**) and M3 (**b, d**) models. Error bars represent the 90% confidence intervals with the cross indicating the median across the 200 bootstrapped samples. The value inferred from the models fitted to the original sample is indicated by a circle

observed for the location parameter μ. It is worth noting that, because μ is the GEV parameter with the smallest sampling standard deviation, it is of common use to first consider a trend in μ. In our case, it is clear that considering solely a time-varying location parameter could lead to a mistaking diagnosis of no trend (or even of a negative trend), thus rejecting the need for non-stationary IDF curves. Another side issue to consider is that the increase of σ for model M2 (assuming μ to be constant) implies a 5% increase of the mean and a 10% increase of the standard deviation of the GEV distribution, defined in Eqs. 20 and 21, respectively (Appendix 3). Such increased interannual variability of the annual maxima involves that the strongest events are becoming stronger, which concurs with our theoretical understanding of global warming imprint on the hydrological cycle (see, e.g., Trenberth et al. 2003). Reminding that this applies only to a small area, this

conclusion calls for extending this research to a larger regional scale.

7 Conclusion

Global warming is inducing changes in the regimes of precipitation worldwide, thus questioning the stationary hypothesis in which is rooted the classical statistical representation of rainfall distributions. The statistical framework presented here aims at providing a coherent tool for testing the presence of trends in time series of precipitation extremes. It combines a GEV distribution representation of the annual maxima with a simple scaling model, allowing to adjust a scaled GEV model of annual maxima valid across a range of durations. To this approach already proposed in the literature, we add a non-stationary

component by expressing the statistical model parameters as a function of time, leading to a non-stationary GEV-Simple Scaling (NS-GEV-SS) model, subsequently allowing for computing consistent non-stationary IDF curves. Because the time dependence increases the number of parameters to be estimated, a trade-off has to be found between model flexibility—a larger number of time-dependent parameters will increase the chance to capture a potential non-stationary signal—and parsimony, ensuring a more robust parameter inference. The model parameters are inferred with a time-dependent maximum likelihood estimator and the model significance is assessed with a semi-parametric bootstrap method to account for the dependence of the scaled samples.

The ability of this theoretical framework to detect non-stationarity is investigated by applying it to series of annual maxima of rainfall intensities for event durations ranging from 2 to 24 h in a region, the Sahel, where detecting trends in rainfall extremes is notoriously challenging, mainly because of the high space-time variability of rainfall. A linear formulation with time as covariate is used to account for non-stationarity, although a more complex parameterization is of course possible. Dealing first with individual point rainfall series, non-stationary models proved to be better than stationary models for only 7% of the 180 series at the 5% level. Even though this could lead to an overall diagnostic that there is no significant extreme precipitation trends in this region, it is worth noting that 63% of the series display a positive trend (albeit significant at the 5% level in 9% only of the cases) of the scale parameter of the GEV distribution. In a second step, the data of the 30 stations are pooled into a unique sample, thus creating 6 regional samples (one for each of the time-steps considered). This regional approach yields a substantial improvement in the detection of trends, with the class of models including a time-varying scale parameter displaying a positive trend significant at the 1% level for all 6 durations. Ultimately, a single regional time-scaled sample is built, using the temporal scale invariance property to pool the 6 regional samples. A more parsimonious model is thus obtained, allowing the production of consistent non-stationary IDF curves. It is worth noting that the parameter ruling the temporal scale invariance property seems stationary—or at least it does not display any statistically significant trend.

The combination of the regional approach with the simple scaling framework leads to infer that the area is undergoing a significant rainfall intensification with the 10-year rainfall increasing by 15 to 20% over the past 28 years and the 100-year rainfall by 23 to 30%. It thus appears that the ongoing intensification has a greater effect on the intensity of the strongest events, which is consistent with the theoretical expectation that "latent heat release strengthens the storm in proportion to their intensity" (Pendergrass 2018). As large as the increase of return levels might be on our case study region, the fact that this trend was not detected on individual stations means that it is just starting to emerge from the range of uncertainty linked to (i) the natural variability of rainfall, especially when it comes to extremes, and (ii) the statistical methods used to model their behavior; demonstrating what is gained from the proposed unified approach. Future work should be conducted in the following complementary directions:

–   Test the framework in other regions of the world, especially those with spatial gradients of rainfall. For instance, including spatial covariate(s) to account for heterogeneous spatial patterns of extreme rainfall statistics (conditioned by, e.g., topographical features) might preserve the acceptability of the unified framework.
–   Further develop the statistical framework, specifically the temporal scale invariance and/or regional models, to more faithfully represent the extreme rainfall behavior in space and across a range of durations.

Appendix 1: Check of the Simple Scaling model validity for the stations and durations considered in the study

The suitability of a simple scaling modeling of the temporal scale invariance of annual maxima is checked over the 30 stations of the ACN dataset in two steps:

–   First, we check the linearity of the $\log(E^{(q)}[I(D)])$ vs $\log(D)$ (5) for $q = \{1, 2, 3, 4\}$. The results are shown in Fig. 11, where each marker represents a station. Figure 12 shows the Pearson coefficient of the linear regression obtained at each station, for each moment order; they are all very close to -1.
–   Second, we check the shape of the relationship between the slopes of the linear regressions of step 1 for each moment order (referred to as $k(q)$) and the moment orders q. For the SS model to be considered a valid approximation, we need the linearity of $k(q)$ vs q. Otherwise, a multi-scaling model should be preferred. As shown in Fig. 13 (left), the $k(q)$ vs q linear fit is very satisfying for all the stations. Figure 13 (right) shows the slope (corresponding to the η parameter of the SS model) and the Pearson coefficient of these regressions.
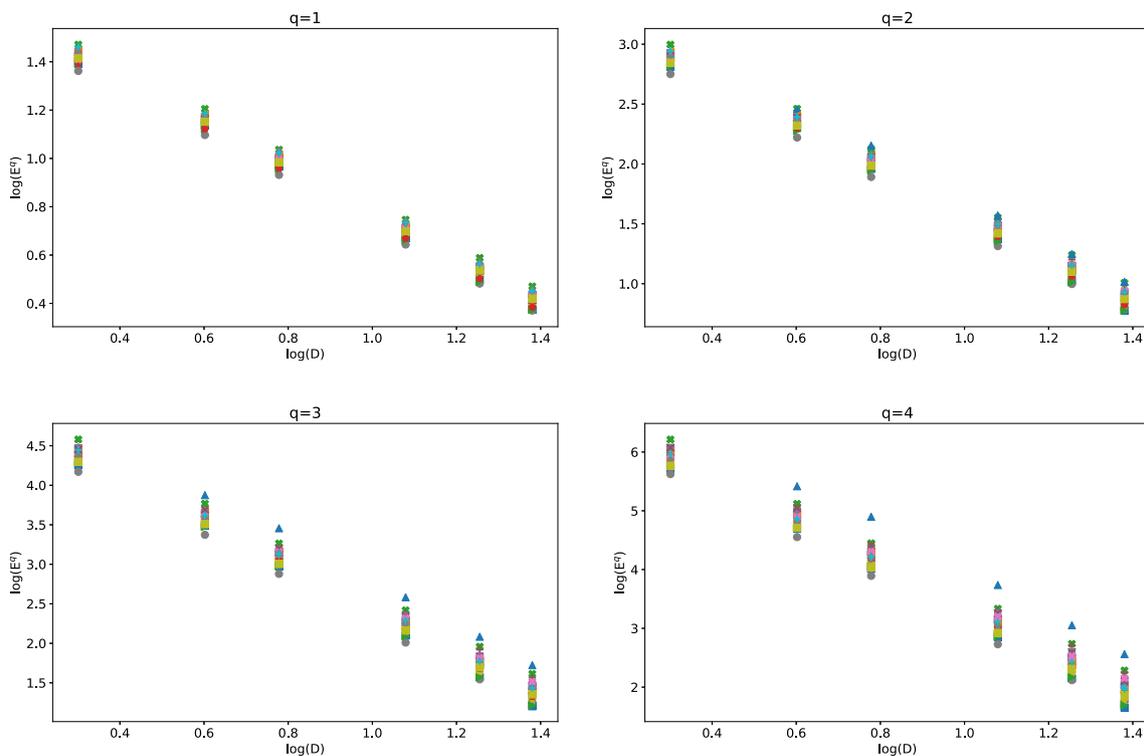
Fig. 11 Log of the moments of order $q = \{1, 2, 3, 4\}$ as a function of the log of durations $D = \{2, 4, 6, 12, 18, 24h\}$ for the 30 stations of the ACN dataset (markers)
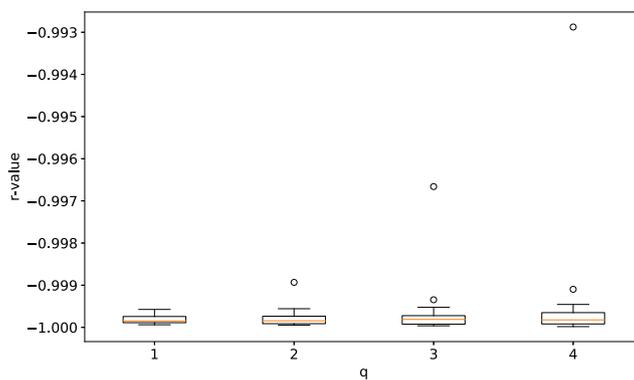
Fig. 12 Distribution among the 30 stations of the ACN dataset of the $\log(E^{(q)}[I(D)])$ vs $\log(D)$ linear regression Pearson coefficients (r-values)
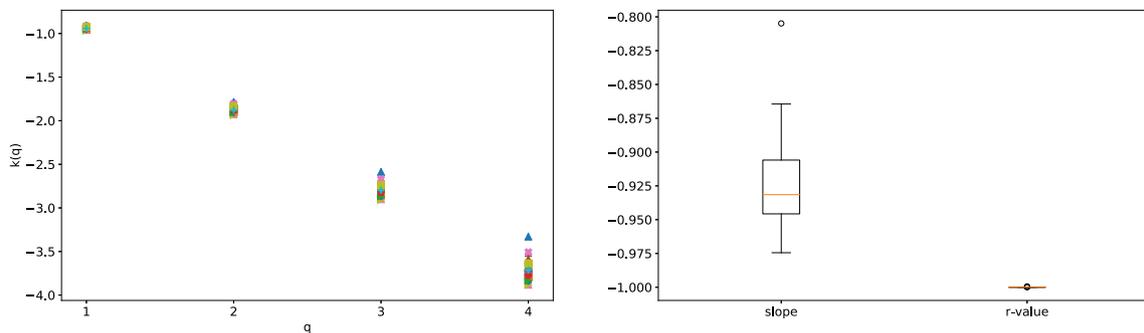
Fig. 13 (left) $k(q)$ as a function of $q = \{1, 2, 3, 4\}$ and (right) slopes and Pearson coefficient (r-values) of the $k(q)$ vs q linear regressions for the 30 stations of the case study. Markers are as in Fig.11

Appendix 2: Cross-correlation matrices of annual maxima for *D*=2h and *D*=24h and annual totals
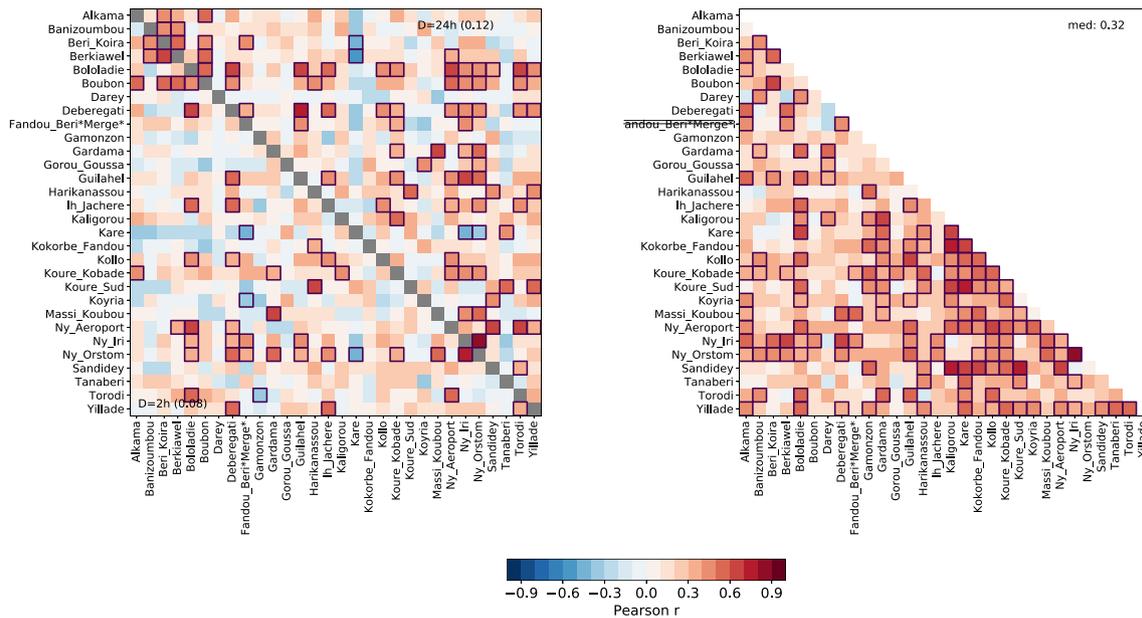
Fig. 14 Pearson coefficients of correlation (r) between the time series of (**a**) annual maximum of rainfall intensities for *D*=2h (lower left) and *D*=24h (upper right) and (**b**) annual totals at the 30 stations of the ACN dataset. Correlations significant at the 5% level are outlined in black. The medians across the stations are indicated in the corners

Figure 14

Appendix 3: Mean and variance of the GEV distribution

Equations 20 and 21 give the relationship between the mean (M) and variance (V), respectively, of the GEV distribution and its parameters.

$$M = \mu + \frac{\sigma}{\xi}(g_1 - 1) \text{ if } \xi \neq 0 \text{ and } \xi < 1 \qquad (20)$$

$$V = \frac{\sigma^2}{\xi^2}(g_2 - g_1^2) \text{ if } \xi \neq 0 \text{ and } \xi < 1/2 \qquad (21)$$

with $g_k = \Gamma(1 - k\xi)$ where Γ is the gamma function.

Availability of data and material Data are available on http://bd. amma-catch.org

Code availability The code is available on request.

Declarations

Conflict of interest The authors declare no competing interests.

References

Allan RP, Barlow M, Byrne MP, Cherchi A, Douville H, Fowler H, Gan TY, Pendergrass A, Rosenfeld D, Swann A, Wilcox LJ, Zolina O (2020) Advances in understanding large-scale responses of the water cycle to climate change. Annals of the New York Academy of Sciences. https://doi.org/10.1111/nyas.14337

Allen MR, Ingram WJ (2002) Constraints on future changes in climate and the hydrologic cycle. Nature 419(6903):228–232. https://doi.org/10.1038/nature01092

Alpert P, Ben-Gai T, Baharad A, Benjamini Y, Yekutieli D, Colacino M, Diodato L, Ramis C, Homar V, Romero R, Michaelides S, Manes A (2002) The paradoxical increase of Mediterranean extreme daily rainfall in spite of decrease in total values. Geophys Res Lett 29(11):1536. https://doi.org/10.1029/2001GL013554

Balme M, Vischel T, Lebel T, Peugeot C, Galle S (2006) Assessing the water balance in the Sahel: Impact of small scale rainfall variability on runoff. J Hydrol 331(1-2):336–348. https://doi.org/10.1016/j.jhydrol.2006.05.020

Blanchet J, Davison AC (2011) Spatial modeling of extreme snow depth. Ann Appl Stat 5(3):1699–1725. www.jstor.org/stable/23069351

Blanchet J, Ceresetti D, Molinié G, Creutin JD (2016) A regional GEV scale-invariant framework for Intensity–Duration–Frequency analysis. J Hydrol 540:82–95. https://doi.org/10.1016/j.jhydrol.2016.06.007 https://linkinghub.elsevier.com/retrieve/pii/S0022169416303584

Bougadis J, Adamowski K (2006) Scaling model of a rainfall intensity-duration-frequency relationship. Hydrol Process 20(17):3747–3757. https://doi.org/10.1002/hyp.6386

Cannon AJ (2010) A flexible nonlinear modelling framework for non-stationary generalized extreme value analysis in hydroclimatology. Hydrol Process 24(6):673–685. https://doi.org/10.1002/hyp.7506

Cannon AJ, Innocenti S (2019) Projected intensification of sub-daily and daily rainfall extremes in convection-permitting climate model simulations over North America: implications for future intensity-duration-frequency curves. Nat Hazards Earth Syst Sci 19(2):421–440. https://doi.org/10.5194/nhess-19-421-2019

Cheng L, AghaKouchak A (2015) Nonstationary precipitation intensity-duration-frequency curves for infrastructure design in a changing climate. Sci Rep 4(1):7093. https://doi.org/10.1038/srep07093

Coles S (2001) An introduction to statistical modeling of extreme values springer series in statistics. Springer, London, London. https://doi.org/10.1007/978-1-4471-3675-0

Donat MG, Alexander LV, Yang H, Durre I, Vose R, Dunn RJH, Willett KM, Aguilar E, Brunet M, Caesar J, Hewitson B, Jack C, Klein Tank AMG, Kruger AC, Marengo J, Peterson TC, Renom M, Oria Rojas C, Rusticucci M, Salinger J, Elrayah AS, Sekele SS, Srivastava AK, Trewin B, Villarroel C, Vincent LA, Zhai P, Zhang X, Kitching S (2013) Updated analyses of temperature and precipitation extreme indices since the beginning of the twentieth century: The HadEX2 dataset. J Geophys Res Atmos 118(5):2098–2118. https://doi.org/10.1002/jgrd.50150

Fisher RA, Tippett LHC (1928) Limiting forms of the frequency distribution of the largest or smallest member of a sample. Math Proc Camb Philos Soc 24(2):180–190. https://doi.org/10.1017/S0305004100015681

Fläschner D, Mauritsen T, Stevens B (2016) Understanding the intermodel spread in global-mean hydrological sensitivity. J Clim 29(2):801–817. https://doi.org/10.1175/JCLI-D-15-0351.1

Galle S, Grippa M, Peugeot C, Moussa IB, Cappelaere B, Demarty J, Mougin E, Panthou G, Adjomayi P, Agbossou E, Ba A, Boucher M, Cohard JM, Descloitres M, Descroix L, Diawara M, Dossou M, Favreau G, Gangneron F, Gosset M, Hector B, Hiernaux P, Issoufou BA, Kergoat L, Lawin E, Lebel T, Legchenko A, Abdou MM, Malam-Issa O, Mamadou Y, Nazoumou Y, Pellarin T, Quantin G, Sambou B, Seghieri J, Séguis L, Vandervaere JP, Vischel T, Vouillamoz JM, Zannou A, Afouda S, Alhassane A, Arjounin M, Barral H, Biron R, Cazenave F, Chaffard V, Chazarin JP, Guyard H, Koné A, Mainassara I, Mamane A, Oi M, Ouani T, Soumaguel N, Wubda M, Ago E, Alle I, Allies A, Arpin-Pont F, Awessou B, Cassé C, Charvet G, Dardel C, Depeyre A, Diallo F, Do T, Fatras C, Frappart F, Gal L, Gascon T,

Gibon F, Guiro I, Ingatan A, Kempf J, Kotchoni D, Lawson F, Leauthaud C, Louvet S, Mason E, Nguyen C, Perrimond B, Pierre C, Richard A, Robert E, Román-Cascón C, Velluet C, Wilcox C (2018) AMMA-CATCH, a critical zone observatory in west africa monitoring a region in transition. Vadose Zone J 17(1):0. https://doi.org/10.2136/vzj2018.03.0062

Giorgi F, Im ES, Coppola E, Diffenbaugh NS, Gao XJ, Mariotti L, Shi Y (2011) Higher hydroclimatic intensity with global warming. J Clim 24(20):5309–5324. https://doi.org/10.1175/2011JCLI3979.1

Gnedenko B (1943) Sur la distribution limite du terme maximum d'une serie aleatoire. Ann Math 44(3):423–453. https://doi.org/10.2307/1968974

Gupta VK, Waymire E (1990) Multiscaling properties of spatial rainfall and river flow distributions. J Geophys Res 95(D3):1999. https://doi.org/10.1029/JD095iD03p01999

Hosking JRM, Wallis JR, Wood EF (1985) Estimation of the generalized extreme-value distribution by the method of probability-weighted moments. Technometrics 27(3):251–261. https://doi.org/10.1080/00401706.1985.10488049

Katz RW (2013) Statistical methods for nonstationary extremes. In: AghaKouchak A, Easterling D, Hsu K, Schubert S, Sorooshian S (eds) Extremes in a changing climate: detection, analysis and uncertainty. Springer, Netherlands, Dordrecht, pp 15–37. https://doi.org/10.1007/978-94-007-4479-0_2

Katz RW, Brown BG (1992) Extreme events in a changing climate: Variability is more important than averages. Clim Chang 21(3):289–302. https://doi.org/10.1007/BF00139728

Katz RW, Parlange MB, Naveau P (2002) Statistics of extremes in hydrology. Adv Water Resour 25(8-12):1287–1304. https://doi.org/10.1016/S0309-1708(02)00056-8

Kharin V, Zwiers FW (2005) Estimating extremes in transient climate change simulations. J Clim 18(8):1156–1173. https://doi.org/10.1175/JCLI3320.1

Kharin V, Zwiers FW, Zhang X, Hegerl G (2007) Changes in temperature and precipitation extremes in the IPCC ensemble of global coupled model simulations. J Clim 20(8):1419–1444. https://doi.org/10.1175/JCLI4066.1

Koutsoyiannis D, Kozonis D, Manetas A (1998) A mathematical framework for studying rainfall intensity-duration-frequency relationships. J Hydrol 206(1-2):118–135. https://doi.org/10.1016/S0022-1694(98)00097-3

Langousis A, Veneziano D (2007) Intensity-duration-frequency curves from scaling representations of rainfall. Water Resour Res vol 43(2). https://doi.org/10.1029/2006WR005245

Langousis A, Veneziano D, Furcolo P, Lepore C (2009) Multifractal rainfall extremes: Theoretical analysis and practical estimation. Chaos, Solitons Fractals 39(3):1182–1194. https://doi.org/10.1016/j.chaos.2007.06.004 https://linkinghub.elsevier.com/retrieve/pii/S0960077907003700

Leadbetter MR (1974) On extreme values in stationary sequences. Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete 28(4):289–303. https://doi.org/10.1007/BF00532947

Lebel T, Cappelaere B, Galle S, Hanan N, Kergoat L, Levis S, Vieux B, Descroix L, Gosset M, Mougin E, Peugeot C, Seguis L (2009) AMMA-CATCH studies in the Sahelian region of West-Africa: An overview. J Hydrol 375(1-2):3–13. https://doi.org/10.1016/j.jhydrol.2009.03.020 https://linkinghub.elsevier.com/retrieve/pii/S0022169409001875

Mathon V, Laurent H (2001) Life cycle of Sahelian mesoscale convective cloud systems. Q J Roy Meteorol Soc 127(572):377–406. https://doi.org/10.1002/qj.49712757208

Menabde M, Seed A, Pegram G (1999) A simple scaling model for extreme rainfall. Water Resour Res 35(1):335–339. https://doi.org/10.1029/1998WR900012

Mélèse V, Blanchet J, Molinié G (2018) Uncertainty estimation of Intensity–Duration–Frequency relationships: A regional analysis. J Hydrol 558:579–591. https://doi.org/10.1016/j.jhydrol.2017.07.054 https://linkinghub.elsevier.com/retrieve/pii/S002216941730519X

van Montfort MA (1990) Sliding maxima. J Hydrol 118(1):77–85. https://doi.org/10.1016/0022-1694(90)90251-R

Nelder JA, Mead R (1965) A simplex method for function minimization. Comput J 7(4):308–313. https://doi.org/10.1093/comjnl/7.4.308

Ouarda TBMJ, Yousef LA, Charron C (2019) Non-stationary intensity-duration-frequency curves integrating information concerning teleconnections and climate change. Int J Climatol 39(4):2306–2323. https://doi.org/10.1002/joc.5953

Panthou G, Vischel T, Lebel T, Blanchet J, Quantin G, Ali A (2012) Extreme rainfall in West Africa: A regional modeling. Water Resour Res vol 48(8). https://doi.org/10.1029/2012WR012052

Panthou G, Vischel T, Lebel T, Quantin G, Pugin AC, Blanchet J, Ali A (2013) From pointwise testing to a regional vision: An integrated statistical approach to detect nonstationarity in extreme daily rainfall. Application to the Sahelian region. J Geophys Res Atmos 118(15):8222–8237. https://doi.org/10.1002/jgrd.50340

Panthou G, Vischel T, Lebel T (2014a) Recent trends in the regime of extreme rainfall in the Central Sahel. Int J Climatol 34(15):3998–4006. https://doi.org/10.1002/joc.3984

Panthou G, Vischel T, Lebel T, Quantin G, Molinié G (2014b) Characterising the space–time structure of rainfall in the Sahel with a view to estimating IDAF curves. Hydrol Earth Syst Sci 18(12):5093–5107. https://doi.org/10.5194/hess-18-5093-2014

Pendergrass AG (2018) What precipitation is extreme? Science 360(6393):1072–1073

Pfahl S, O'Gorman PA, Fischer EM (2017) Understanding the regional pattern of projected future changes in extreme precipitation. Nat Clim Change 7(6):423–427. https://doi.org/10.1038/nclimate3287

Salas JD, Obeysekera J (2014) Revisiting the concepts of return period and risk for nonstationary hydrologic extreme events. J Hydrol Eng 19(3):554–568. https://doi.org/10.1061/(ASCE)HE.1943-5584.0000820

Sane Y, Panthou G, Bodian A, Vischel T, Lebel T, Dacosta H, Quantin G, Wilcox C, Ndiaye O, Diongue-Niang A, Diop Kane M (2018) Intensity–duration–frequency (IDF) rainfall curves in Senegal. Nat Hazards Earth Syst Sci 18(7):1849–1866. https://doi.org/10.5194/nhess-18-1849-2018

Sarhadi A, Soulis ED (2017) Time-varying extreme rainfall intensity-duration-frequency curves in a changing climate. Geophys Res Lett 44(5):2454–2463. https://doi.org/10.1002/2016GL072201

Schertzer D, Lovejoy S (1987) Physical modeling and analysis of rain and clouds by anisotropic scaling multiplicative processes. J Geophys Res 92(D8):9693. https://doi.org/10.1029/JD092iD08p09693

Sebille Q, Fougères AL, Mercadier C (2017) Modeling extreme rainfall A comparative study of spatial extreme value models. Spat Stat 21:187–208. https://doi.org/10.1016/j.spasta.2017.06.009 https://linkinghub.elsevier.com/retrieve/pii/S2211675316300975

Svensson C, Jones D (2010) Review of rainfall frequency estimation methods. J Flood Risk Manag 3(4):296–313. https://doi.org/10.1111/j.1753-318X.2010.01079.x

Trenberth KE (2011) Changes in precipitation with climate change. Clim Res 47(1):123–138. https://doi.org/10.3354/cr00953 http://www.int-res.com/abstracts/cr/v47/n1-2/p123-138/

Trenberth KE, Dai A, Rasmussen RM, Parsons DB (2003) The changing character of precipitation. Bull Am Meteorol Soc 84(9):1205–1218. https://doi.org/10.1175/BAMS-84-9-1205

Tyralis H, Langousis A (2019) Estimation of intensity–duration–frequency curves using max-stable processes. Stoch Env Res Risk A 33(1):239–252. https://doi.org/10.1007/s00477-018-1577-2

Veneziano D, Yoon S (2013) Rainfall extremes, excesses, and intensity-duration-frequency curves: A unified asymptotic framework and new nonasymptotic results based on multifractal measures: Rainfall Extremes, Excesses, and IDF Curves. Water Resour Res 49(7):4320–4334. https://doi.org/10.1002/wrcr.20352

Wehner MF, Smith RL, Bala G, Duffy P (2010) The effect of horizontal resolution on simulation of very extreme US precipitation events in a global atmosphere model. Clim Dynam 34(2-3):241–247. https://doi.org/10.1007/s00382-009-0656-y

Westra S, Alexander LV, Zwiers FW (2013) Global increasing trends in annual maximum daily precipitation. J Clim 26(11):3904–3918. https://doi.org/10.1175/JCLI-D-12-00502.1

Westra S, Fowler HJ, Evans JP, Alexander LV, Berg P, Johnson F, Kendon EJ, Lenderink G, Roberts NM (2014) Future changes to the intensity and frequency of short-duration extreme rainfall. Rev Geophys 52(3):522–555. https://doi.org/10.1002/2014RG000464

Yeo MH, Nguyen VTV, Kpodonu T (2020) Characterizing extreme rainfalls and constructing confidence intervals for IDF curves using Scaling-GEV distribution model. Int J Climatol p joc. 6631. https://doi.org/10.1002/joc.6631