ORIGINAL PAPER

Parametrisation of change-permitting extreme value models and its impact on the description of change

Ilaria Prosdocimi[1] · Thomas Kjeldsen[2]

Abstract

The potential for changes in environmental extremes is routinely investigated by fitting change-permitting extreme value models to long-term observations, allowing one or more distribution parameters to change as a function of time or some other covariate. In most extreme value analyses, the main quantity of interest is typically the upper quantiles of the distribution, which are often needed for practical applications such as engineering design. This study focuses on the changes in quantile estimates under different change-permitting models. First, metrics which measure the impact of changes in parameters on changes in quantiles are introduced. The mathematical structure of these change metrics is investigated for several change-permitting models based on the Generalised Extreme Value (GEV) distribution. It is shown that for the most commonly used models, the predicted changes in the quantiles are a non-intuitive function of the distribution parameters, leading to results which are difficult to interpret. Next, it is posited that commonly used change-permitting GEV models do not preserve a constant coefficient of variation, a property that is typically assumed to hold for environmental extremes. To address these shortcomings a new (parsimonious) model is proposed: the model assumes a constant coefficient of variation, allowing the location and scale parameters to change simultaneously. The proposed model results in changes in the quantile function that are easier to interpret. Finally, the consequences of the different modelling choices on quantile estimates are exemplified using a dataset of extreme peak river flow measurements in Massachusetts, USA. It is argued that the decision on which model structure to adopt to describe change in extremes should also take into consideration any requirements on the behaviour of the quantiles of interest.

Keywords Environmental extremes · Parametric models · Return period · Quantile function

1 Introduction

There is widespread interest in quantifying the impacts of climate and other anthropogenic changes on the likelihood of very extreme natural hazards (IPCC 2012). When assessing the risk of a natural hazard, for example to design critical infrastructures, an estimate is needed of the magnitude of extremes expected to occur with a certain rarity, typically derived as quantiles of a statistical distribution. This estimation is carried out under the assumption that the past recordings of the variable under study would still be representative of the stochastic behaviour of the variable at present and in the future. This can be formalised as an assumption that the process under study is stationary (Coles 2001; François et al. 2019). With the increasing evidence of the impacts of climate and other anthropogenic changes, the validity of this assumption is increasingly being challenged. As a result, a large of number of studies have investigated the evidence for change in risk levels of several natural hazards in observational data (e.g. Coates et al. 2014; Guerreiro et al. 2018; François et al. 2019, and references therein), typically using statistical approaches based on Extreme Value Analysis (Cooley 2009; Katz et al. 2002). Often, a default distribution is assumed and one or more distribution parameters are modelled as a

✉ Ilaria Prosdocimi
ilaria.prosdocimi@unive.it

Thomas Kjeldsen
T.R.Kjeldsen@bath.ac.uk

[1] Università Ca' Foscari Venezia, Via Torino 155, Venezia Mestre, Italy

[2] University of Bath, Claverton Down, Bath BA2 7AY, UK

function of covariates, e.g. time. The hypothesis that the behaviour of extremes might have been changing is formulated as a functional relationship between one or more of the distribution parameters and the covariate(s). This type of change-permitting analysis is also referred to as non-stationary analysis, while the models with constant values for all parameters throughout time are often referred to as stationary analysis.

Table 1 shows an extremely non-exhaustive list of published articles based on such change-permitting analysis: the list is provided to showcase the variety of modelling choices and applications of change-permitting models. From the Table it can be seen that different combinations of distributions, model structures and covariates have been employed in the literature. All the differences in the model components eventually have an impact in the final understanding about changes in extremes. Specifically, most of the studies listed in Table 1 focus on the changes to the higher quantiles of the distribution.

In this study we discuss ways to describe the changes in frequencies and magnitude focussing on change in the quantile function, showing that different model structures necessarily result in different functional forms of changes in the quantile function. Further, we show that commonly adopted model structures can result in non-intuitive behaviour of the quantile function. We argue that a reflection on the type of change in higher quantiles which can be derived

for each model structure should be performed at the initial stages of any analysis to ensure that the analysis results can provide the most appropriate communication of how extremes might be changing.

In particular, we formalise the investigation of how the different model structures impact the description of change on effective quantile estimates across different exceedance probabilities using two sets of criteria. The first set of criteria measures the impact of the changes which depend on the covariates used in the change-permitting models. The second set measures the effect of switching from a no-change model to a change-permitting model. Some of these measures have already been partially proposed in the literature: Vogel et al. (2011) introduced the magnification factor as a way to describe relative changes in quantiles in time. Further, the R package `extRemes` Gilleland and Katz (2016) provides functions to compute the difference in effective quantile estimates under different values of explanatory variables. Moreover, several studies in the literature, including those listed in Table 1, carry out some investigation of the implied impact on the magnitude of estimated quantile of using change-permitting models against results obtained when assuming constant parameters.

The analytical study of the change criteria shows that a minority of model structures results in simple definitions of change in the quantile function. One of these is a model structure which maintains a constant coefficient of

Table 1 Examples of studies which use change-permitting models to assess changes in annual maxima records

Reference	Variable (Y)	Distribution	Model structure	Covariate (X)
Vogel et al. (2011)	River flow	Log-Normal $LN(\mu, \sigma)$	location: $\mu = \mu_0 + \mu_1 x$ scale: $\sigma = \sigma_0$ (constant)	Time
Villarini et al. (2009)	River flow	Gumbel(μ, σ) (among others)	Location: $\mu = g(x)$ (non-parametric) scale: $\sigma = g(x)$ (non-parametric)	Time, population, rainfall
Cheng and AghaKouchak (2014)	Rainfall	GEV(μ, σ, ξ)	Location: $\mu = \mu_0 + \mu_1 x$ scale: $\sigma = \sigma_0$ (constant) shape: $\xi = \xi_0$ (constant)	Time
Luo and Zhu (2018)	Wave height	GEV(μ, σ, ξ)	Location: $\mu = \mu_0 + \mu_1 x$ scale: $\sigma = \sigma_0$ (constant) shape: $\xi = \xi_0$ (constant)l	Time
Šraj et al. (2016)	River flow	GEV(μ, σ, ξ)	Location: $\mu = \mu_0 + \mu_1 x$ scale: $\sigma = \exp\{\sigma_0 + \sigma_1 x\}$ shape: $\xi = \xi_0$ (constant)	Time, rainfall
Lu et al. (2020)	River flow	GEV(μ, σ, ξ)	Location: $\mu = g(x)$ (several parametric) scale: $\sigma = g(x)$ (several parametric) shape: $\xi = \xi_0$ (constant)	Time, rainfall, reservoir index
Griffin et al. (2019)	River flow	Gen. Logistic GLO(ξ, α, κ)	Location: $\xi = \xi_0 + \xi_1 x$ scale: $\alpha = \exp\{\alpha_0 + \alpha_1 x\}$ shape: $\kappa = g(t)$	Time
Brown et al. (2014)	Rainfall and temperature (observed and modelled)	GEV(μ, σ, ξ)	Location: $\mu = \mu_0 + \mu_1 x$ scale: $\sigma = \exp\{\sigma_0 + \sigma_1 x\}$ $\xi = \xi_0 + \xi_1 x$	global temperature

GEV Generalised Extreme Value distribution

variation, thus enforcing a change in the scale of the distribution when the location is allowed to change. Such models are not commonly used in the investigation of changes in extremes, even though a constant coefficient of variation is a common characteristic of estimates based on environmental extremes records (Overeem et al. 2009; Menabde et al. 1999; Serago and Vogel 2018; Blanchet et al. 2009).

The manuscript is organised as follow: a brief introduction of extreme values analysis and the models used to assess changes in extremes is given in Sect. 2. The quantile-based measures of change are presented in Sect. 3 and are applied to a case study of peak flow records in Massachusetts, USA, in Sect. 4. Section 5 closes the paper offering a perspective on the importance of model specification when investigating changes in extremes. The maximum likelihood estimation framework is adopted throughout the manuscript, although some of the concepts discussed would easily apply to models estimated within a Bayesian framework.

2 Statistical models for changing extremes

2.1 Extreme value models and design events; the fixed-parameters case

Assessing the risk of a natural hazard involves an assessment about the frequency at which events of given magnitudes can be expected to be exceeded. Typically, this is done by estimating the *design event*, which is expected to be exceeded with a annual exceedance probability (AEP) of p in any given year. If the distribution of event magnitudes is constant in time and events from year to year can be deemed independent from one another, the AEP has a direct relationship with the return period T, the average period of time over which one event at-least as big as the design event would be recorded, where $T = 1/p$ (see Volpi 2019). Since the interest of the estimation focusses on the most extreme magnitudes which might be recorded, rather than the typical magnitudes, Extreme Value Analysis is used to model the distribution of the extreme values. In particular, extremes can be defined as the maximum value in a sample recorded over fixed periods of time (for example a year). It can be shown that the Generalised Extreme Value (GEV) distribution is the limiting distribution for maxima of stationary series (Coles 2001), although other distributions are sometimes used when estimating design events for engineering purposes (Castellarin et al. 2012).

In general, it is assumed that the variable of interest, denoted by Y, follows a certain distribution parametrised by $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_d)$ with probability density function

$f(y; \boldsymbol{\theta})$. Typically a two ($d = 2$) or three-parameter ($d = 3$) distribution is used (Castellarin et al. 2012) and, in a first instance, all parameters are assumed to be constant. In the remainder we will focus on the GEV distribution due to its central role in the analysis of extremes. The GEV is a three parameter distribution characterised by a location parameter μ (with $\mu \in (-\infty, \infty)$), a scale parameter σ (with $\sigma \in (0, \infty)$) and a shape parameter ξ (with $\xi \in (-\infty, \infty)$). When $\xi = 0$ the distribution reduces to a two parameter distribution: the Gumbel distribution. The distribution's domain depends on the sign of the shape parameter taking the following values: $-\infty < y \leq \mu - \sigma/\xi$ for $\xi < 0$, $\mu - \sigma/\xi \leq y < \infty$ for $\xi > 0$ and $-\infty < y < \infty$ for $\xi = 0$. For a given sample of extremes $\boldsymbol{y} = (y_i, \ldots, y_n)$ it is possible and typically straightforward to find an estimate of $\boldsymbol{\theta}$ denoted as $\hat{\boldsymbol{\theta}}$. Estimates for the design event Q_T (where T indicates the return period) can be derived from the quantile function corresponding to the fitted distribution $q(1 - p; \hat{\boldsymbol{\theta}})$ (where $p = 1/T$) so that $P(Y > Q_T) = p$. For the GEV distribution the quantile function takes the form:

$$q(1-p; \mu, \sigma, \xi) = \begin{cases} \mu + \dfrac{\sigma}{\xi}\left[(-\log(1-p))^{-\xi} - 1\right] & \text{for } \xi \neq 0 \\ \mu + \sigma\{-\log(-\log(1-p))\} & \text{for } \xi = 0 \end{cases}$$

(1)

where p indicates the exceedance probability. The quantile function for the GEV shown in Eq. (1) can be rewritten as:

$$q(1-p; \mu, \sigma, \xi) = \begin{cases} \mu + \dfrac{\sigma}{\xi}\left[y_p^{-\xi} - 1\right] & \text{for } \xi \neq 0 \\ \mu + \sigma(-\log(y_p)) & \text{for } \xi = 0 \end{cases}$$

(2)

taking $y_p = (-\log(1-p))$. Notice that for $\xi = 0$ the quantile function of the Gumbel distribution is a linear transformation of $-\log(y_p)$ with μ being the intercept and σ being the slope of the line.

2.2 Incorporating change in extreme value models; the change-permitting case

The hypothesis that the behaviour of extremes might have been changing is typically formulated by assuming a functional relationship between one or more of the distribution parameters, and one or more covariates. The validity of the hypothesis is assessed by comparing the goodness-of-fit of the models with varying parameters against models with fixed parameters. This approach is established for extreme value analysis (Coles 2001) and has some similarity to Generalized Linear Models and their extensions such as Vector Generalised Additive Models (Yee 2015) or Generalised Additive Models for Location Scale and Shape (GAMLSS, Rigby and Stasinopoulos 2005; Wood 2017), which are also referred to as distributional regression

models (Umlauf and Kneib 2018). In this work, to keep the presentation and computations as simple as possible, only linear models (with appropriate link functions) with one covariate are explored. The concepts can be extended to the case in which more than one covariate would enter the model and in which the relationship between the covariate and the distribution parameter (or their link transformation) would be better described by some non-linear function. Indeed, different covariates might enter the model in either a linear or non-linear fashion affecting one or more of the distribution parameters: see Yee and Stephenson (2007) for a thorough discussion on this within the extreme values framework.

As an example of the change-permitting approach the annual maxima series of the Aberjona River at Winchester, Massachusetts (USGS Gage 01102500) is used. Serago and Vogel (2018) reported that the watershed upstream of the station has undergone significant urban development which would typically increase flood magnitudes. The $n = 78$ annual maximum peak flow records are assumed to be independent and drawn from a GEV distribution, possibly with one or more parameters changing throughout the recording period. The peak flow records are shown in Fig. 1, and it seems clear that through time the observations have been increasing in magnitude and possibly also become more variable.

Four models with different change-permitting structures are fitted to the observations to investigate possible changes in the peak flow distribution. At first a model M_0 with all parameters kept constant is considered. It is assumed that each element in the sample is iid and GEV-distributed:

$$M_0 : Y_i \sim GEV(\mu, \sigma, \xi) \text{ for } i = 1, \ldots, n.$$

Next, two models in which the location of the distribution is allowed to change as a function of time with different link functions are used. As in Serago and Vogel (2018) time is used as a covariate in the model, as a surrogate for
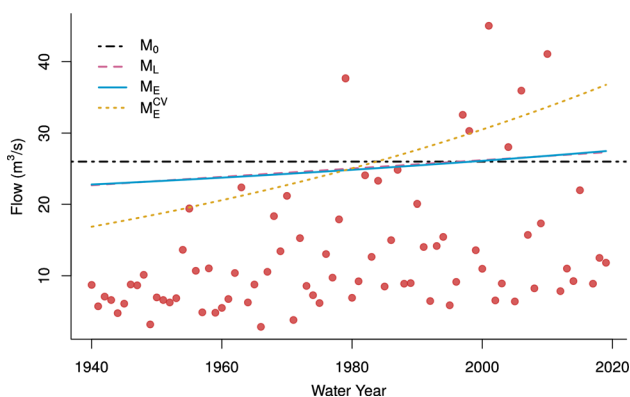
Fig. 1 Peak flow data for the Aberjona River together with the effective 30-year design event according to different change-permitting models presented in the main text

the increased urbanisation levels. Assuming that the variable describing river peak flow at time i follows a GEV distribution with a different location for each year, $Y_i \sim GEV(\mu_i, \sigma, \xi)$, the following two models are fitted: a model with a linear link between the location and time (M_L) and a model using an exponential function to link the location parameter and time (M_E):

$$M_L \quad : Y_i \sim GEV(\mu_i, \sigma_C, \xi_C), \text{ with } \mu_i = \mu_{C0} + \mu_{C1}x_i \qquad \text{for } i = 1, \ldots, n$$
$$M_E \quad : Y_i \sim GEV(\mu_i, \sigma_C, \xi_C), \text{ with } \mu_i = \exp\{\eta_{C0} + \eta_{C1}x_i\} \quad \text{for } i = 1, \ldots, n$$

In the last fitted model (M_E^{CV}) the location is modelled as a function of time via an exponential link function and the scale is allowed to change proportionally to the location:

$$M_E^{CV} : Y_i \sim GEV(\mu_i, \tau\mu_i, \xi_C), \text{ with } \mu_i = \exp\{\eta_{C0} + \eta_{C1}x_i\} \text{ for } i = 1, \ldots, n. \tag{3}$$

where τ is the ratio between the scale and the location of the distribution. The model structure is discussed further in Sect. 2.3. The C subscript is used to emphasise the parameters of change-permitting models.

In Fig. 1 the estimated 30 year event (i.e the design event with AEP = 1/30) under the four different models fitted to the data are shown. These are obtained by plugging the estimated parameter values in the quantile function in Eq. (1) and correspond to the effective quantiles discussed in Katz et al. (2002). Notice that for all four models shown in the Figure, different values of the scale and shape parameters will be estimated because different structures are used for the location function.

The maximum likelihood estimates of the model parameters in all four models discussed above are given in Table 2, together with their standard errors. In all models the coefficients describing the change in time for the location parameter are estimated to be positive and significantly different from 0 at the 5% significance level. The estimated value for the location parameter in 1940 and 2019 are also displayed in Table 2: the final estimation of the location at the beginning and at the end of the record is fairly similar for the three models in which the location is allowed to change: the choice of the link function has a minor effect on the estimated value of the location function within the recording period.

In Table 2 the estimated parameter values for an additional model (M_S) in which the scale is allowed to change with time using an exponential link function are also provided. The model structure is the following:

$$M_S : Y_i \sim GEV(\mu_C, \sigma_i, \xi_C), \text{ with } \sigma_i = \exp\{\gamma_{C0} + \gamma_{C1}x_i\}$$
$$\text{for } i = 1, \ldots, n.$$

This model should be used rarely and is presented here for comparative purposes: the scale parameter describes the variation of the distribution around its centre and typically

Table 2 Maximum likelihood parameter estimates for the models presented in main text for Aberjona River peak flow data

Model	Parameters				Estimate		log-lik
M_0	$\hat{\mu}$ (se)	–	$\hat{\sigma}$ (se)	$\hat{\xi}$ (se)	$\hat{\mu}_{1940}$	$\hat{\mu}_{2019}$	log-lik
	8.52 (0.57)	– (–)	4.46 (0.50)	0.34 (0.10)	8.52	8.52	−255.11
M_L	$\hat{\mu}_{C0}$ (se)	$\hat{\mu}_{C1}$ (se)	$\hat{\sigma}_C$ (se)	$\hat{\xi}_C$ (se)	$\hat{\mu}_{1940}$	$\hat{\mu}_{2019}$	log-lik
	4.06 (1.37)	6.98 (1.95)	4.09 (0.47)	0.38 (0.11)	6.41	11.04	−249.83
M_E	$\hat{\eta}_{C0}$ (se)	$\hat{\eta}_{C1}$ (se)	$\hat{\sigma}_C$ (se)	$\hat{\xi}_C$ (se)	$\hat{\mu}_{1940}$	$\hat{\mu}_{2019}$	log-lik
	1.61 (0.18)	0.81 (0.22)	4.06 (0.47)	0.39 (0.11)	6.56	11.25	−249.73
M_E^{CV}	$\hat{\eta}_{C0}$ (se)	$\hat{\eta}_{C1}$ (se)	$\hat{\tau}$ (se)	$\hat{\xi}_C$ (se)	$\hat{\mu}_{1940}$	$\hat{\mu}_{2019}$	log-lik
	1.39 (0.20)	1.17 (0.30)	0.48 (0.04)	0.30 (0.11)	5.94	12.93	−247.81
M_S	$\hat{\mu}_C$ (se)	$\hat{\gamma}_{C0}$ (se)	$\hat{\gamma}_{C1}$ (se)	$\hat{\xi}_C$ (se)	$\hat{\sigma}_{1940}$	$\hat{\sigma}_{2019}$	log-lik
	8.42 (0.79)	1.38 (0.70)	0.18 (1.07)	0.32 (0.14)	4.21	4.75	−255.09

one would first need to correctly model changes in the location parameter as a function of some predictors to correctly characterise changes in the scale. For this station there is little evidence that scale alone is changing when keeping the location fixed, although the variability of the record shown in Fig. 1 would appear to be higher in the later years. Indeed for the data in Fig. 1 there appear to be an increase in both the overall magnitude and the variability of flow extremes: this feature can be well described by the model M_E^{CV}, in which the location is allowed to change as a function of time and the scale changes proportionally to the location.

To assess the goodness of fit of the change-allowing models to the data, beside assessing whether the parameter describing the change is significantly different from zero, it is often appropriate to check that the distribution estimated under each model gives a good fit to the observed data. As discussed, among others, in Coles (2001) this can be done using graphical tools such as probability-probability plots (pp-plot) or quantile-quantile plots (qq-plot) which can provide a visual check for the validity of the distributional assumption for the response variable. Software such as `ismev` and `extRemes` makes the testing of model significance and the assessment of the goodness of fit fairly straightforward. For the Aberjona river case the goodness of fit plots (not shown) indicate that the GEV assumption seem to hold. Nevertheless, when assessing the significance of a change-permitting model or selecting the most suitable model by means of information criteria (e.g. AIC or BIC), it is also important to assess and take into account the goodness of fit of different models to the data.

2.3 A model preserving a constant coefficient of variation

The model in Eq. (3) has rarely been employed in extreme value analysis: it is mentioned in passing by Smith (1999) that such a model gave a better fit to the rainfall data under study, but it has not been employed in the investigation of changes in environmental extremes, to the best of the authors' knowledge.

The model is constructed so that when the location changes so does the scale, while the ratio between the scale and the location (τ) remains constant. Although τ is not exactly the coefficient of variation for a GEV distribution it is tightly related to its value. Most of the studies investigating changes in extremes in the literature (see for example Table 1), use GEV models with a varying location or scale (or location and scale which are allowed to vary independently from each other). These models have a varying coefficient of variation: a change in, say, the location parameter would involve also a change in the coefficient of variation. The model suggested in Eq. (3) keeps a constant coefficient of variation providing a parsimonious representation of changes in both location and scale. As mentioned in Serago and Vogel (2018), the estimates of the coefficient of variation are often found to be approximately constant across series measuring the same variable, while location and scale are well-known to vary greatly with some external variable. For example, in river flow measurements both the location and the scale typically vary as a function of watershed size and mean rainfall, while the estimates of the coefficient of variation are relatively constant across different watersheds. Similarly, rainfall accumulation across longer durations will tend to be larger and also exhibit more variability. These well-known properties of extremes are the basis for the widely adopted index-flood method and regional flood

frequency analysis (Hosking and Wallis 1997) and for the standard methods used to derive Intensity Duration Frequency (IDF) curves (Menabde et al. 1999). These properties are sometimes referred to as the scaling properties of environmental extremes (Gupta and Waymire 1990). Beside the case of rainfall and river flow extremes, simple and multi-scale properties have been found to hold for, for example, snow accumulation, which scales with altitude (Blanchet et al. 2009), and wind speed, which scales with duration (Diebold and Heller 2019). The model in Eq. (3) is inspired by these so-called scaling properties of extremes: by keeping a constant coefficient of variation changes in the location result in proportional changes of the scale. As discussed in the next Section this also provide a straightforward description of changes in the distribution quantiles.

3 Measuring the impact of change

Once a relevant change in the distribution parameters is identified, this implies that the overall distribution of the variable of interest is deemed to have changed. The left panels in Fig. 2 gives an illustration of how potential changes in the parameters correspond to changes in the distribution by displaying the probability density functions (pdf) of a GEV distribution under different parameter values. The right panels of the Figure show instead the return plots, which depict the quantile curve as a function of a transform of the AEP. The curves shown in the Figure could represent the past and present day pdf and quantile functions under different change-permitting models: the black solid line defines a baseline, for example the

estimate obtained for the beginning of the record when using time as covariate. The other three lines represent: (1) a change only in the location parameter (red dashed line), (2) a change only in the scale parameter (blue point and dash line), and (3) a change in both the location and scale parameter while maintaining the ratio between the scale and the location parameter constant (brown short and long-dashed line).The difference in the parameter values in the distributions results in quantile functions which all differ from the quantile function of the baseline model in a different way. The changes in the quantiles, which are related to the design events and are therefore the quantity of interest of extreme value analysis for environmental extremes, are the focus of the measures of change suggested in the next section. A number of metrics and approaches to redefine the concept of design event in the presence of change have been being proposed in the literature (see, among others, Rootzén and Katz 2013; Salas et al. 2018; Hu et al. 2018), but in this study we focus on metrics based on changes in quantiles.

3.1 Measuring change: some quantile-based metrics

Two quantile-based change metrics are introduced here: the difference, D, and the ratio, M, between two quantile functions. These two metrics can be computed to compare design events under different values of the covariate in the change-permitting models (indicated by superscript δ) or to compare design events under the fixed-parameters and change-permitting models. In the first case, one could compare the estimated quantiles for time point x and time point $x + \Delta x$, with Δx indicating the difference in time

Fig. 2 Probability distribution functions and quantile functions for a Gumbel distribution and a GEV distribution with parameters as those indicated in the legend
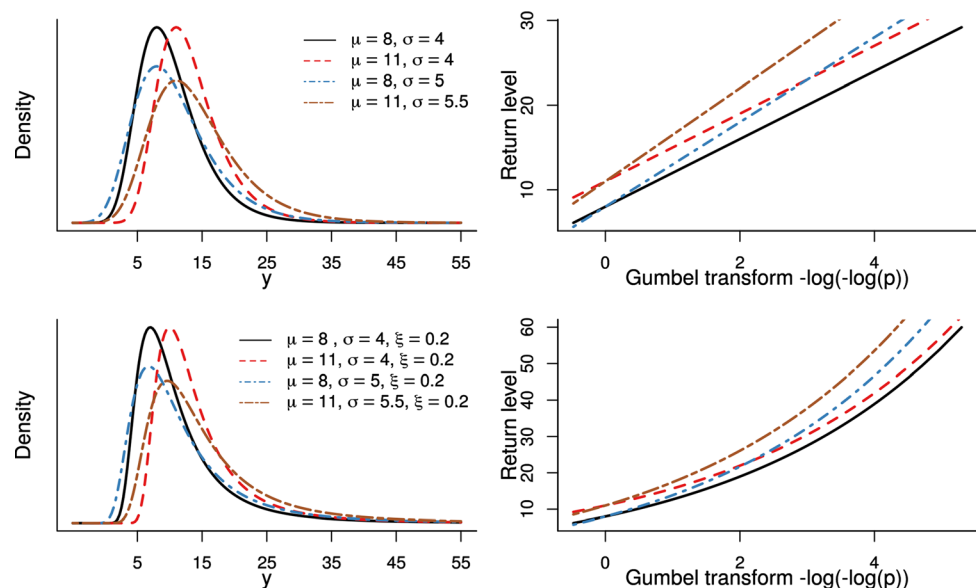
between the two time points under study (or the difference in any other covariate values). Denoting with $\boldsymbol{\theta}_C(x^*)$ the values of the parameters in the change-permitting model which are obtained evaluating the parameter functions at values x^* of the covariate, the following change metrics are defined:

$$D^\delta(p, \Delta x) = q(p, \boldsymbol{\theta}_C(x + \Delta x)) - q(p, \boldsymbol{\theta}_C(x)) \tag{4}$$

$$M^\delta(p, \Delta x) = \frac{q(p, \boldsymbol{\theta}_C(x + \Delta x)))}{q(p, \boldsymbol{\theta}_C(x))} \tag{5}$$

$$D(p, x^*) = q(p, \boldsymbol{\theta}_C(x^*)) - q(p, \boldsymbol{\theta}) \tag{6}$$

$$M(p, x^*) = \frac{q(p, \boldsymbol{\theta}_C(x^*))}{q(p, \boldsymbol{\theta})} \tag{7}$$

The metrics of change introduced above can be used to assess the impact of the changing behaviour of extremes as described by different model structures on the estimated quantiles of the distribution, i.e. the design events. Note that the quantity in Eq. (5) is a generalisation of the magnification factor introduced by Vogel et al. (2011), while values of the quantity in Eq. (4) can be derived within the R package extRemes (Gilleland and Katz 2016).

3.2 Measuring change within the change-permitting models

The types of changes in the parameter values in Fig. 2 reflect some of the changes observed in parameter values obtained from change-permitting models often used in environmental studies. The model defined by a linear change in the location parameter of a GEV distribution with a constant scale, in which it assumed that $Y_i \sim GEV(\mu_{C0} + \mu_{C1}x_i, \sigma, \xi)$, would corresponds to the red dashed line. This modelling assumption leads to a vertical shift of the quantile function. In this case the metric $D^\delta(p, \Delta x)$ remains constant across all values of p and takes value $D^\delta(p, \Delta x) = \mu_{C1}\Delta x$. In contrast, the ratio $M^\delta(p, \Delta x)$ depends on p, Δx and all the four parameter in the model: i.e. the ratio between quantiles would be different across exceedance probabilities (return periods) since

$$M^\delta(p, \Delta x) = \frac{\mu_{C1}\Delta x}{\mu_0 + \mu_{C1}x + \sigma(y_p^{-\xi} - 1)/\xi} + 1.$$

The values of $D^\delta(p, \Delta x)$ and $M^\delta(p, \Delta x)$ for a number of models commonly used when assessing the presence of change in environmental extremes are provided in the "Appendix". There are only a few change-permitting GEV models for which one of the metric of change does not depend on the exceedance probability. One of them is the case in which the location is allowed to change, either linearly, as in the case discussed above, or as an exponential function of the covariate: in this case the value of

$D^\delta(p, \Delta x)$ is constant across all exceedance probabilities. When using the model proposed in Eq. (3), it is the ratio of quantiles that is a constant which only depends on Δx and μ_{C1}:

$$M^\delta(p, \Delta x) = \exp\{\mu_{C1}\Delta x\}$$

Vogel et al. (2011) used time as a covariate and defined M^δ as a "decadal" magnification factor based on $\Delta x = 10$ years for the case of the two-parameter log-normal distribution, which is characterised by a constant coefficient of variation. This expression of change is independent of return period and has a simple interpretation. The model introduced here in Eq. (3) allows this concept to be applied with the more widely used GEV distribution and to provide a convenient way to communicate the outcome of change-permitting analysis to end-users.

3.3 Measuring change against the fixed-parameters model

For any model used to assess the existence of change in extremes, a natural question is also how estimates from a change-permitting model compare to the corresponding estimates obtained using fixed parameters, which are likely to be the basis of the current design event estimates. If the distribution is changing, then the probability of exceeding a pre-specified value would be different at the present time than it was at any time in the record and, more importantly, the estimation based on the assumption of a unique distribution for all data, would be biased.

The two metrics proposed in Eqs. (6) and (7), $D(p, x^*)$ and $M(p, x^*)$, provide a natural way to assess these impacts, but for almost no model can these two quantities reduce to simple metrics such as those available for $D^\delta(p, \Delta x)$ or $M^\delta(p, \Delta x)$. Nevertheless, it is possible for a fixed value of the covariate x^*, to investigate how $M(p, x^*)$ or $D(p, x^*)$ change as a function of p. Often, up to a certain value of $\tilde{p}$, the quantiles of the change-permitting model would be larger (respectively smaller) than the fixed-parameter model, while after $\tilde{p}$ the quantiles of the change-permitting model would be smaller (respectively larger). In practical applications, this can be a source of doubt for decision making, since events up to probability $\tilde{p}$ might be, say, underestimated when using fixed-parameters models, while events with probability of exceedance greater than $\tilde{p}$ would appear to be overestimated. This can happen, for example, when a linear trend is assumed for the location parameter and found to be, say, positive and some variability in the data can be explained by such trend, so that the estimate for the scale parameter diminishes in value: this is the case for the Aberjona river as shown in Table 2. When this happens one of the possible consequences is a shifting and

"flattening" the return curve so that there is an increase in the quantiles which are exceeded relatively frequently, while the very rare events are found to be generally smaller. Similarly, when including a linear trend in the location parameter, the estimate for the shape parameter might vary so that the return curve under the change-permitting model might increase at a faster or slower rate.

Figure 3 shows a comparison of the scale and shape parameter for a fixed-parameters model and for a change-permitting model with a linear trend in the location as a function of time applied to 40 annual maximum series of instantaneous peak river flow in Massachusetts each containing more than 65 years of data (see Sect. 4 for a complete description of the dataset). For 31 out of 40 stations, the scale parameter is smaller for the change-permitting model, reflecting that allowing the location parameter to change explains a portion of the variability in the records. The shape parameter estimates also change (middle figure), with both increases and decreases observed.The impacts on the quantile estimation when migrating from a fixed-parameters to a change-permitting model can be assessed using $M(p, x^*)$, i.e. the ratio between quantiles computed under the change-permitting and the fixed parameters models, for different value of p. In the right panel of Fig. 3 the scatterplot of $M(p, x^*)$ for the median (i.e. $p = 0.5$) and for the 100-year event (i.e. $p = 0.01$) is shown. For all stations in the datasets the change-permitting models are evaluated at the end of the record of the stations ($x^* = max(x)$). The sign of the change for the frequent ($p = 0.5$) and rare ($p = 0.01$) events is different at times (these are the red stars in the bottom right quadrant and top left quadrant). These instances represent cases where the introduction of a change-permitting model will result in changes in design flood estimates at high return periods which are different in direction than those for low return periods; a counter intuitive outcome for most practical applications. Moreover, the magnitude of the change can be different for the two AEP values, due to the difference in the scale and the shape parameter between the two models. For this dataset changes in the high quantile (AEP of 0.01) tend to be smaller than changes in the median, reflecting that the scale parameter estimates tend to be smaller for change-permitting model than for the fixed-parameters model.

3.3.1 Enforcing change structures

Although having different direction of changes for events of different rarity can be practically challenging, this property might be something that a modeller wishes to exploit when analysing extremes. For example, there might be some physical reasoning by which the magnitude of frequent events under some pre-specified value of the covariates is expected to be reduced while the rare events would become larger. For example, Sharma et al. (2018) discuss how different types of change of the distribution of high flows might be expected for different catchment types, while Guerreiro et al. (2018) show that changes in the magnitude of extreme rainfall are different for the relatively common and the most extreme events.

When fixing the shape parameter to be the same in the fixed-parameters and the change-permitting models, the point at which the two return curves cross can be found analytically for some of the commonly used models, namely the model with a linear change in the location and a model with an exponential change in the scale (see
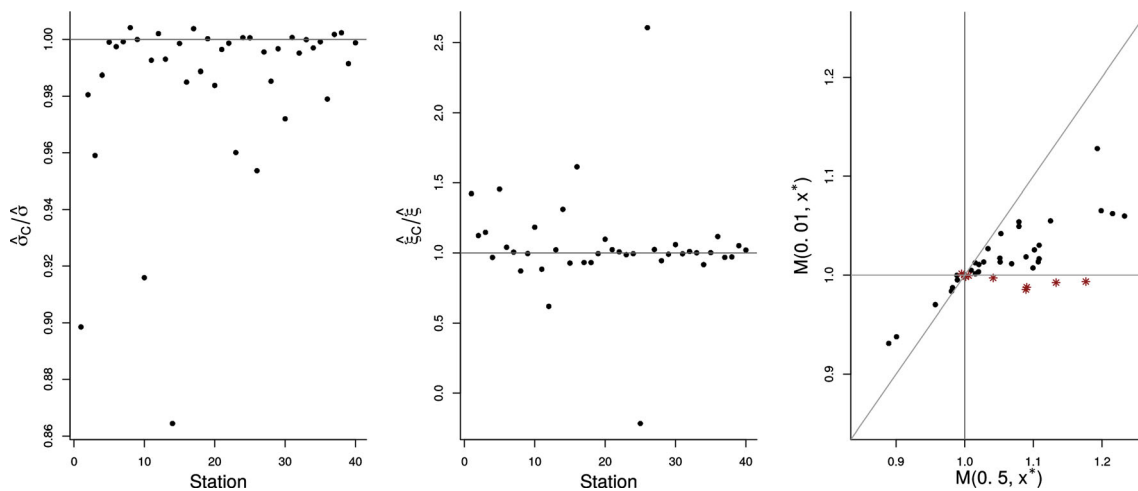
Fig. 3 The impact of fitting a linear trend in the location for the Massachusetts peak flow dataset. Left panel: ratio of the estimated scale parameter under the change-permitting model and the fixed-parameter model. Central panel: ratio of the estimated scale parameter under the change-permitting and the fixed-parameter model. Right panel: scatterplot of M(0.5, x^*) and M(0.01, x^*); red stars indicate stations for which the direction of change is different for $p = 0.5$ and $p = 0.01$

"Appendix"). Under the fixed parameters model the assumed distribution is $Y \sim GEV(\mu, \sigma, \xi)$. Under the linear change in the location model, the assumed distribution is $Y_i \sim GEV(\mu_{C0} + \mu_{C1}x_i, \sigma_C, \xi)$. Noticeably, the ξ parameter is kept to be the same under both models. The $D(p, x^*)$ function for this change-permitting model is found to be:

$$D(p, x^*) = (\mu_{C0} + \mu_{C1}x^* - \mu) + \frac{\sigma_C - \sigma}{\xi}(y_p^{-\xi} - 1)$$

The two quantile functions have the same value (i.e. they cross) at the point in which $D(p, x^*) = 0$, which occurs at:

$$y_p = \left(1 + \xi \frac{\mu - \mu_{C0} - \mu_{C1}x^*}{\sigma_C - \sigma}\right)^{-1/\xi}$$

provided that $\xi(\mu - \mu_{C0} - \mu_{C1}x^*)/(\sigma_C - \sigma) > -1$. Manipulating the formula for $D(p, x^*)$ further, one can define the value that σ_C should take to ensure that the two quantile curves cross at a fixed y_p value for a given value of (μ_{C0}, μ_{C1}) as:

$$\sigma_C = \sigma + \frac{\mu - \mu_{C0} - \mu_{C1}x^*}{y_p^{-\xi} - 1}\xi.$$

This could be enforced in the model estimation to ensure that the change in the sign of $D(p, x^*)$ happens exactly at a desired AEP $\tilde{p}$. Further, one can study the sign of $D(p, x^*)$ as a function of p. It can be found that quantiles from the change-permitting model exceed the quantiles from the fixed-parameters models ($D(p, x^*) > 0$) as long as:

$$\sigma_C > \sigma + \frac{\mu - \mu_{C0} - \mu_{C1}x^*}{y_p^{-\xi} - 1}\xi. \tag{8}$$

From this result it is possible to further investigate the properties a change-permitting model should have to ensure that effective return periods of a certain rarity are larger ($D(p, x^*) > 0$) or smaller ($D(p, x^*) < 0$) than the return periods under the fixed-parameters model. First, notice that

$$\begin{cases} y_p^{-\xi} - 1 > 0 & \text{for } \xi < 0 \text{ and } p > 1 - e^{-1} \\ y_p^{-\xi} - 1 > 0 & \text{for } \xi > 0 \text{ and } p < 1 - e^{-1} \end{cases}$$

where $p = 1 - e^{-1}$ corresponds approximately to an AEP of 0.63 and a return period of T = 1.58. Assuming that one would want the two curves to cross at a value of $p > 1 - e^{-1}$, we see that for $D(p, x^*)$ to be positive (i.e. to ensure that the dis-equality in Eq. (8) holds) one needs to find that the location parameter evaluated at x^* (i.e. the value $\mu_{C0} + \mu_{C1}x^*$) is smaller that the location parameter in the fixed parameters model (μ). This would mean, for example, that if time is the covariate in the model and x^* is taken to be the end of the record, to ensure that the design events under the change-permitting model are larger than the ones obtained under the fixed parameter model for event with AEP lower than a certain value $\tilde{p}$, one would

need to find a negative trend in the location and at the same time an increase in the scale parameter compared to the fixed-parameters model. Combining these findings it would be possible to define change-permitting models such that the effective return curve derived at a certain value of the covariate crosses the fixed-parameters return curve exactly at a pre-specified AEP $\tilde{p}$ and that has higher (or lower) estimated return levels than the fixed-parameters model for events with AEP smaller than $\tilde{p}$. This could be done by including constrains or using convenient re-parametrisation within the likelihood functions used to estimate both the fixed and the change-permitting models.

The constrains needed to ensure a positive $D(p, x^*)$ under different change-permitting models are more cumbersome than the ones presented above for the case of the linear change in location. Nevertheless, it is generally true that if one wishes to ensure that the estimated magnitudes of rare events (i.e. events with a small AEP) evaluated at a given value x^* in the change-permitting model are larger than the ones under the fixed-parameters model, the scale parameter under the change permitting needs to be larger, while the location parameter needs to be smaller (provided the shape is kept to be the same in both models).

4 Case study: the Massachusetts peak flow data

In this Section records of annual maxima of instantaneous peak river flow recorded from rivers in the state of Massachusetts, USA, are used to explore some of the practical consequences for the estimation of quantiles when imposing different model structures to allow for change in the distribution of peak flows. The data consists of the annual maximum series of instantaneous peak flow for 40 stations in the state with at least 65 years of valid records. The longest record is 115 year long and all records end after 2015. Some of these stations record flow at locations for which the upstream watershed has undergone changes in urbanisation similar to those experienced by the Aberjona watershed. In this study we do not wish to assess whether peak river flow has changed across the state, nor to identify the drivers of such change: we instead focus on the consequences of different parametrisation of change on the estimated design events. We also do not attempt to make an assessment of which of the model structures employed to characterise possible changes in the records is the most suitable one for the different river flow records in Massachusetts: more complete checks on the goodness of fit of each model fit for every record would be needed. Overall, finding the most suitable statistical model to characterise possible changes in the distribution of river flow peaks would require a more thorough statistical investigation of

the data and a deeper understanding of the possible causes driving the observed changes (e.g. changes in the land use in the watershed).

A number of different fixed-parameters and change-permitting models are fitted via standard maximum likelihood estimation to the data series of each station in the dataset. For all models time is used as a covariate and the following transformed logit function is used for the shape parameter:

$$\xi = \text{logit}(\zeta) - 0.5$$

to ensure that $\xi \in [-0.5, 0.5]$. This constraint (adapted from the `gevlss` function in the `mgcv` R package, Wood 2017) ensures that the mean and variance of the GEV are finite and the maximum likelihood estimates are consistent (see Smith 1985). At first, the standard model with fixed parameters, M_0, is estimated for each record, together with the different change-permitting models (M_L, M_E, M_S and M_E^{CV}) which were applied to the Aberjona river data as seen in Table 2.

4.1 From fixed- to change-permitting models

At first we compare the support in favour of a change in the behaviour of extremes across the different models. For each of the model different assumptions are made regarding which property of the distribution is changing and how this change is related to the covariate. Therefore, rather than comparing the raw estimated values for the parameters describing the estimated changes across different stations and different models, the parameter estimates are standardised by their estimated standard deviation and compared in Fig. 4. In the left panel the trend parameters for the model M_L and M_E are compared. For the model with a linear link function, M_L, the parameter of interest is μ_{C1}, while for M_E, in which an exponential link function is used, the parameter of interest corresponds to η_{C1}. For these two models, the detected changes in the location parameter are similar in sign and strength: most changes are positive, indicating an increase in the location parameter over time for the majority of stations. As in the Aberjona case the estimated location function within the recording period is similar when using a linear or an exponential link function.

In contrast, the sign and strength of the trend when modelling location and scale are very different. The scatterplot in central panel of Fig. 4 indicates that for sites with strong changes in the location parameter there is no corresponding strong evidence of change in the scale parameter when this is modelled as a changing function of time while keeping the location fixed (M_S). However, when the scale is allowed to change proportionally to the location (M_{CV}^E), the strength and direction of the trend is similar to those found when allowing only the location to change (right panel). At each site, using the model in Eq. (3) would results in identifying a changing behaviour of extremes similar in sign and strength to the one identified when only allowing the location parameter to change.

Importantly, the consequence of the identified change on the quantiles would be different for different models due to differences in model structures. These differences are shown for the Aberjona river in Fig. 5. The top panel and the bottom left panel in Fig. 5 show in detail the comparison of the return curves for the fixed-parameter model (M_0) and for different change-permitting models evaluated in year 2019 (the final year in the record). These curves are derived from the estimated parameters shown in Table 2: changes in the parameters in the change-permitting models
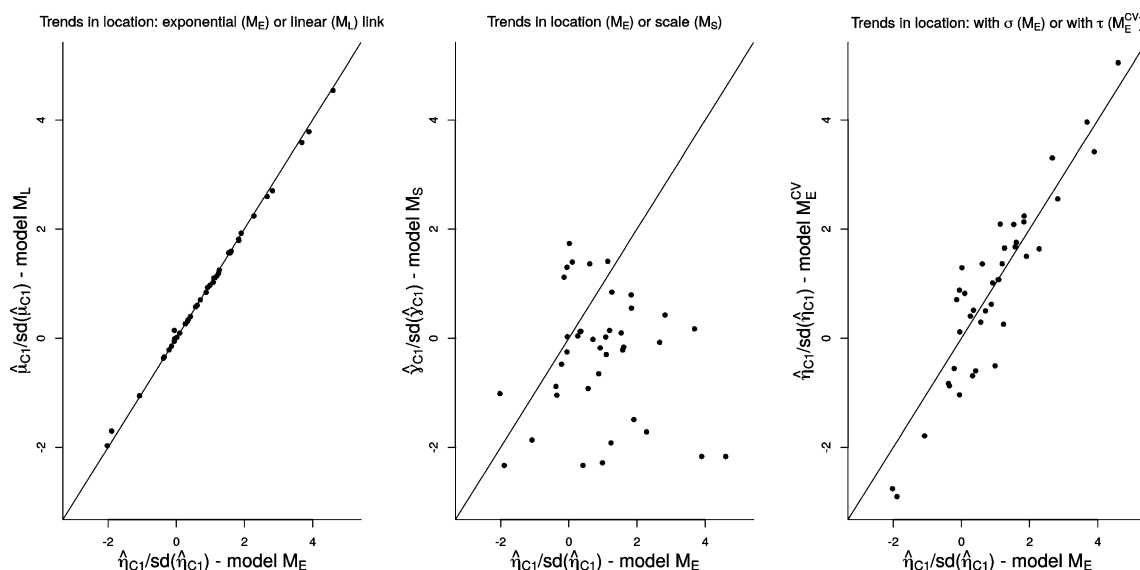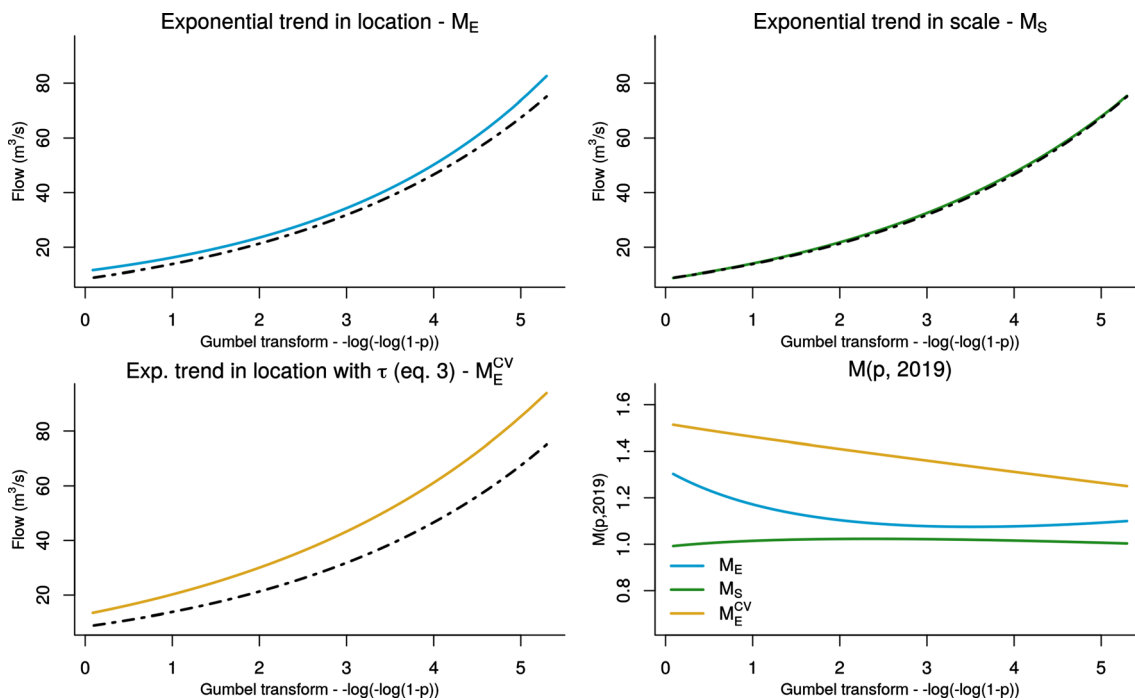

Fig. 4 Scatterplot of scaled trend coefficients under different modelling assumptions

Fig. 5 Comparison of return curves under the fixed-parameters model (dashed black line) and different change-permitting models for the Aberjona river. The bottom right panel shows the evolution of $M(p, 2019)$ for the change-permitting models across different AEP p

imply different relative changes across the AEP values. For models allowing changes in the location parameter it was found that the parameter value in year 2019 is generally larger than the one of the fixed parameter model and this implies an overall increase in the return curve. In contrast, the minimal difference between the parameter estimates for the fixed-parameter model M_0 and the model in which the scale is allowed to change in time M_S is reflected in almost identical return curves (top-right panel) and consequently an almost flat $M(p, 2019)$ line (lower right panel).The relative changes of the return curves for the different change-permitting models in all stations are shown in Fig. 6. For all stations in all models $M(p, x^*)$ is evaluated in the last year available in the record, thus Fig. 6 shows the ratio of the effective return curve of the different change-permitting model structures evaluated at the end of the record period against the estimated return curve for the fixed parameter model. The left panel shows the relative change in quantiles across AEP induced by assuming an exponential trend in the location parameter. In this case the relative changes implied in the quantile functions tend to be minimal for the events with AEP of approximately 0.1 and the relative changes in the frequent events (Gumbel transform smaller than .4) tend to be larger than the relative changes for rare events (Gumbel transform larger than 4.6) . The contrary is true when the scale alone is allowed to change (central panel): overall there are much larger relative changes for the rarer events. When the scale is allowed to change proportionally to the location (right panel) the

relative changes for the different AEP are fairly comparable for most stations. As in Fig. 3, for a handful of stations the direction of change for the median and the 100-year event are different when only the location is allowed to change; this is also true when the scale only is allowed to change. When using the model in Eq. (3) the changes have instead always the same direction in the range of AEP considered in Fig. 6, with most of them being fairly constant. In conclusion, the assumption made in terms of what parts of the distribution are allowed to change has a strong impact on the implied changes in higher quantiles which are typically of interest for engineering design.

4.2 Changes over the record-period

Figure 7 shows the return curves for the change-permitting models fitted to the Aberjona river series and evaluated for year 1965 and 2015. This allows the comparison of the possible changes in the quantiles of the peak flow distribution in 50 years. The bottom right panel shows the $M^\delta(p, 50)$ values: as expected when the parametrisation presented in Eq. (3) is used the relative change in the two quantiles is the same and is equal to $\exp\{\mu_{C1}\Delta t\}$.

Figure 8 compares the relative changes of the quantiles in 2015 and 1965 as a function of AEP ($M^\delta(p, 50)$) for the change-permitting models M_L, M_E and M_E^{CV} for all stations in the study. Thus, the left panel shows the relative changes over 50 years in the effective return levels when the model
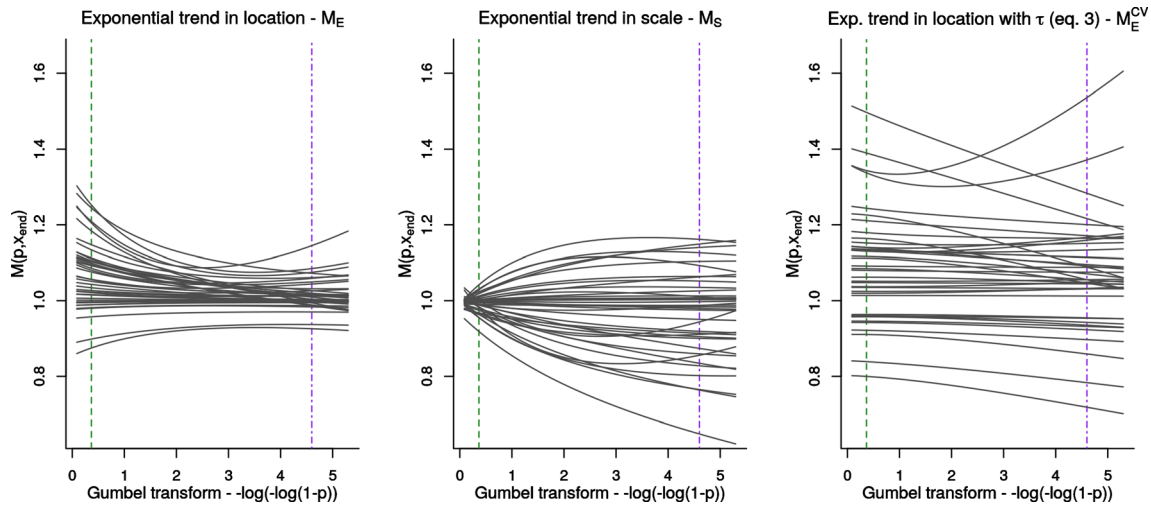
Fig. 6 $M(p, x_{end})$ across different AEP p for the different change-permitting models. The purple and green vertical lines indicate, respectively, the AEP = 0.5 and AEP = 0.01 values
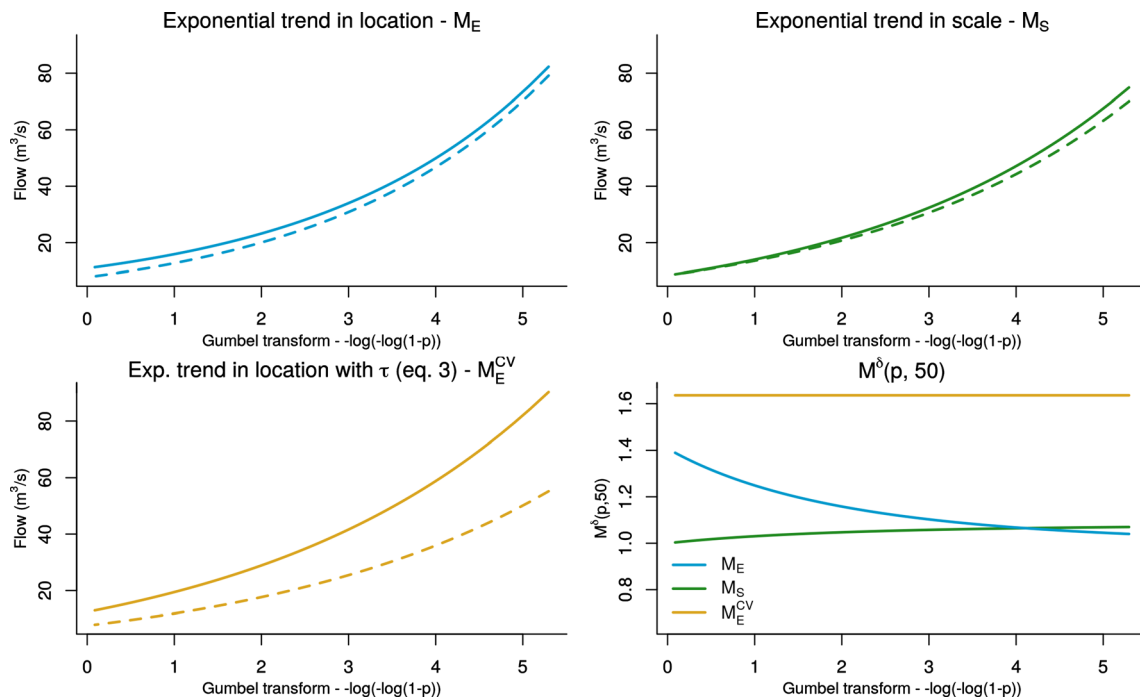

Fig. 7 Comparison of return curves for the change-permitting models in 1965 (dashed lines) and 2015 (solid lines) for the Aberjona river. The bottom right panel shows the evolution of $M^\delta(p, 50)$ across different AEP p

structure allow changes in the distribution as an exponential change in the location parameter. In this case, the changes in the frequent events are much larger than those in the rarer events: an increase in the location parameter would entail much larger estimates in the later years for the median, but only mildly larger estimates in the later years for the 100-year event. The behaviour is reversed when the scale only is allowed to change: frequent events are estimated to have relatively similar magnitude in 2015 and 1965, while

the estimated magnitude of rare events is very different with effective design events for low AEP being either much larger or much smaller in 2015 than in 1965. The consequence of using the parametrisation of change presented in Eq. (3) is evident in the right hand panel: changes are constant across AEP giving an easy way to assess how the quantiles of the distribution might have changed over the 50 year period.

No single model of change can be appropriate to study changes in all extremes series, and an evaluation of the
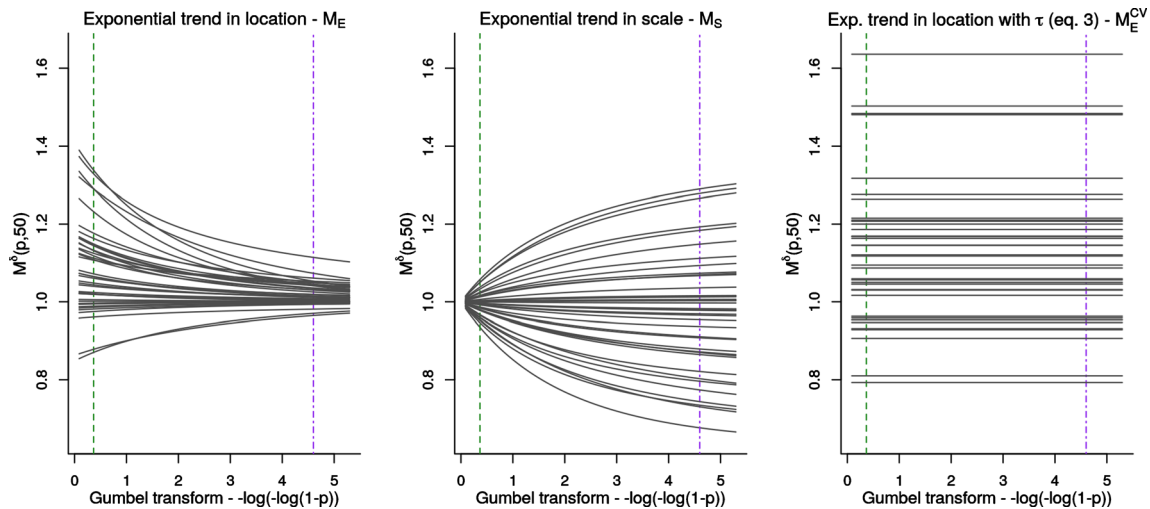
Fig. 8 $M^\delta(p, 50)$ across different AEP p for the change-permitting models. The purple and green vertical lines indicate, respectively, the AEP = 0.5 and AEP = 0.01 values

goodness of fit for the fitted models can be useful to assess whether any model can provide a better representation than others for the series under study. This might include both comparative measures such as AIC or BIC and graphical tools such as qq-plots which can indicate whether the distributional assumption made for the data is suitable. Furthermore, models fitted to the data should be reasonable and coherent with the physics principles underlying the process under study. Nevertheless, some decisions are typically made at an initial modelling stage of the parametric forms of change to consider as possible models. This decision could also be informed by considerations of whether any of the assumed forms can provide more usable metrics to describe changes in design events (when these are of interest) or any other quantity of interest derived from the fitted model.

### 4.3 Assessing uncertainty of quantile estimates

The assessment of the uncertainty in the quantities of interest is a key element when performing any statistical analysis. Different approaches have been employed in the literature to assess the variability of the estimated quantiles, the most common ones being the profile likelihood, the delta method, and the non-parametric and parametric bootstrap (Cooley 2013; Kyselỳ 2008). In particular, Kyselỳ (2008) shows that the latter method appears to provide good uncertainty estimates under several cases. The parametric bootstrap is therefore employed here to compare the variability of the estimated quantiles under different model structure for all stations in the study. A brief outline of the parametric bootstrap procedure is provided below.

For any model fitted to a river flow series of size n, an estimate for the vector of d parameters $\hat{\boldsymbol{\theta}}$ is obtained. Using the estimated parameters, B bootstrap samples of size n are randomly generated from the appropriate distribution parametrised by $\hat{\boldsymbol{\theta}}$. For each bootstrap sample the vector $\hat{\boldsymbol{\theta}}^*$ is estimated, together with the quantiles of interest possibly evaluated at a specific covariate value $\tilde{x}$, $q^* = q(1 - p; \hat{\boldsymbol{\theta}}^*(\tilde{x}))$. This results in a bootstrap sample of quantiles $(q_1^*, \ldots, q_B^*)$. Percentile based $(1 - \alpha) * 100\%$ confidence intervals for the quantile of interest are derived as the $\alpha/2$ and $1 - \alpha/2$ percentiles of the bootstrap sample.

The scatterplots in Fig. 9 compare the width of the 95% confidence intervals for the design event of AEP equal to 0.5 derived at all stations for the different change-permitting model structures against the width of the interval for the fixed-parameters model. For all station the quantiles for the change-permitting models are evaluated at the last available year, i.e the end of the record. Noticeably, for all models in which the location is allowed to change (M_L and M_E) confidence intervals are found to be wider. In contrast, the width of the confidence intervals for the frequent events under a model with a varying scale M_S (bottom right panel) appears to be comparable to the one derived when assuming fixed parameters in time (M_0). This is not the case though when the intervals under comparison refer to rarer events. Figure 10 is structured as Fig. 9, but showing the width of confidence intervals for rare design events (AEP = 0.01). The scatterplot in bottom right panel in this case shows that the intervals for the model in which scale is allowed to change (M_S) are generally wider than those obtained when no change is allowed (M_0). This is not the case for models in which the mean only is allowed to change (M_L and M_E): in the top two panels the interval widths for the M_L and M_E model are similar to those for the no-change (M_0) model. The width of the intervals obtained using the model which allows the scale to change proportionally to the location (M_{CV}^E) is also typically wider: the
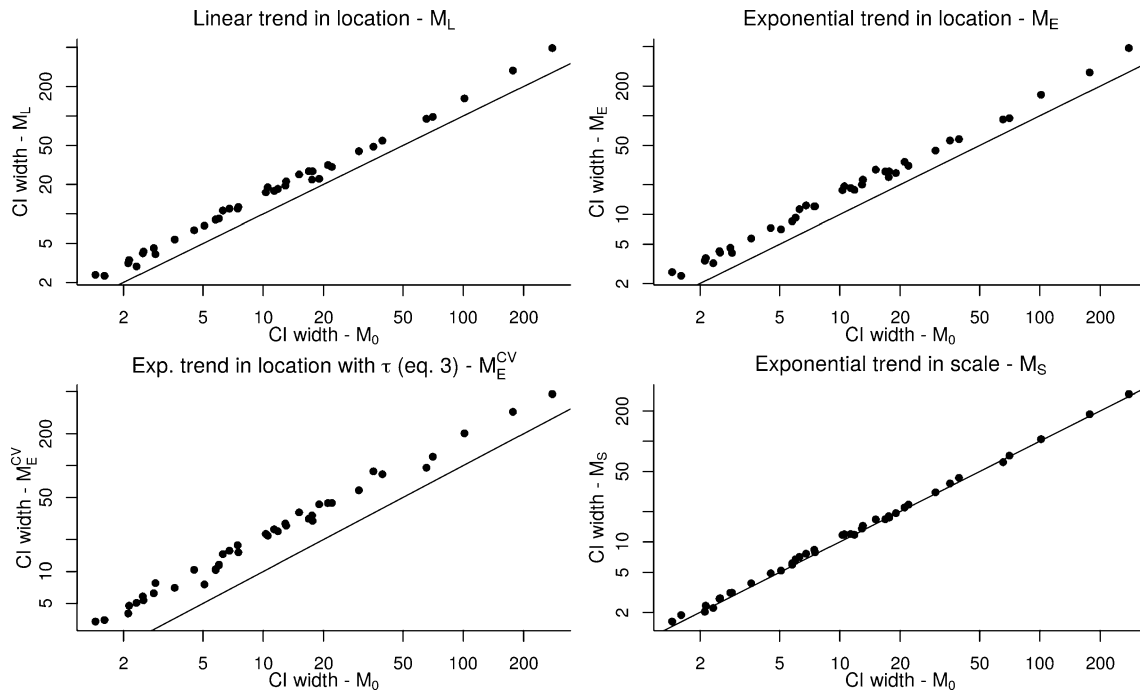
Fig. 9 Bootstrap 95% confidence interval width (log-scale) for the effective design event of AEP = 0.5 under different change-permitting model structures against the width of the interval for the fixed parameter model (log-scale)
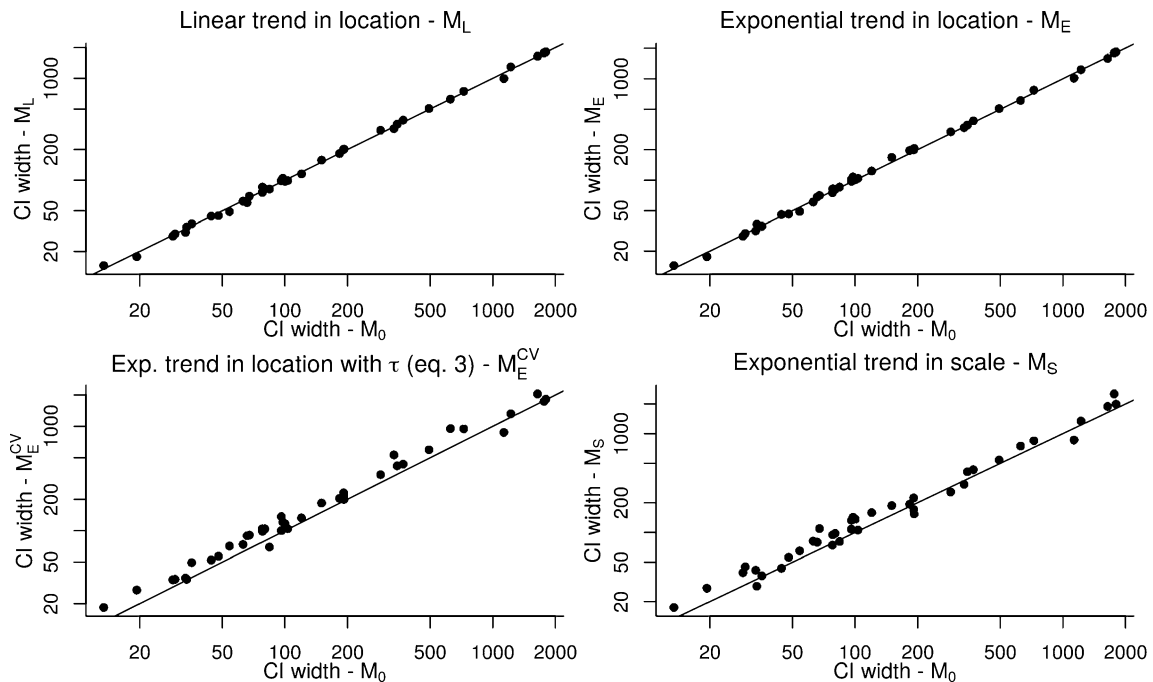
Fig. 10 Bootstrap 95% confidence interval width (log-scale) for the effective design event of AEP = 0.01 under different change-permitting model structures against the width of the interval for the fixed parameter model (log-scale)

variability of estimates of the magnitude of rare events is highly influenced by changes in the scale parameter.

Overall the variability of quantile estimates can be quite different under the various model structures investigated in this study. A better understanding of how the variability of the quantiles of interest might be affected by different parametrisation could also contribute to the decision of which model is adopted to describe change.

5 Discussion

In this paper the effect of different model structures used to describe changes in extremes is discussed by focusing on the impacts on quantiles as estimated under different change-permitting parametrisations. Although throughout the paper only univariate parametric models have been discussed, the results can be easily extended to the case of multivariate and non-parametric models. Further, although all analysis and calculations have been carried out assuming a GEV distribution, the findings could be useful for other distributions whose quantile function has the same structure as the one in Eq. (2), such as the Generalised Pareto distribution which is typically employed when analysing peaks-over-threshold records and has also been used to detect changes in environmental extremes (Silva et al. 2016, 2017). In that case, changes in the location parameter would correspond to changes in the threshold above which value are considered to be extremes. This has already been already proposed in the literature (see for example Roth et al. 2012; Eastoe and Tawn 2009), although with no specific consideration of the impacts of the modelling assumption on the quantile estimation.

The results from this study show that extreme value models routinely used in the applied literature can lead to very different assessments of change, but the consequences of the different modelling choices on the estimated design events are rarely discussed and explored. Moreover, many of the models typically employed entail a varying coefficient of variation, leading to changes in quantile estimates which are difficult to study under different values of the covariates and cumbersome to reconcile with the estimates derived under the fixed-parameters models. A wider discussion on what types of changes can be expected under different conditions (see for example Sharma et al. 2018), might help in clarifying what model structures are more apt to capture the expected changes. Possibly, the modelling options which are currently explored are not the ones which best describe the changes expected from a warming climate or an increase in population and urban areas. Using simple metrics which describe changes such as those proposed in Eqs. (4) to (7) can be useful to explore the consequences of different models structures on the estimation of quantities which are routinely extracted from any estimation procedure. By comparing the expected changes to the types of changes which can be described under different models, it would be possible to ensure that models of change employed the most suitable parametrisation to describe what is perceived to be the most likely direction and shape of change. As shown, sentences like "We expect all quantiles to have changed in time by a constant factor" or "We would expect events of AEP smaller than 0.02 to be larger in the future and events with AEP larger than 0.02

to be smaller" can be translated into models with specific parametrisation and possibly constrained values for certain parameters. It is important that the description of change of the quantities of interest derived from change-permitting models can be interpretable and therefore considered credible by the end users. We believe the model based on τ presented in Eq. (3), in which the scale is allowed to change proportionally to the location, is a first step towards such models, since it allows for the ratio between quantiles to be the same for all exceedance probabilities, and therefore provides a direct communication of how changes in the distributions affect the final main output of interest.

Data Availability The Massachusetts River peak flow data was downloaded from the United States Geological Survey by means of the `dataRetrieval` R package (De Cicco et al. 2018).

Software Availability Code which implements the model presented in Eq. (3) based on the `gev.fit` function of the `ismev` R package by Gilleland et al. (2018) is made available at https://gist.github.com/ilapros/8193d65a4c9426b27d7ad66a87aa16a5 and as supplementary material to this paper.

Appendix

Calculations for $D^\delta(p, \Delta x)$ and $M^\delta(p, \Delta x)$

- In the case in which $Y_i \sim GEV(\mu_{C0} + \mu_{C1} x_i, \sigma_C, \xi_C)$

$$D^\delta(p, \Delta x) = \mu_{C1} \Delta x \ .$$

The evolution of $D^\delta(p, \Delta x)$ changes as a function of μ_{C1} and of Δx: the difference between quantiles is constant across different return periods.

$$M^\delta(p, \Delta x)$$
$$= 1 + \frac{(\mu_{C1} \Delta x) \xi_C}{\xi_C(\mu_{C0} + \mu_{C1} x) + \sigma_C(y_p^{-\xi_C} - 1)} \ .$$

The evolution of $M^\delta(p, \Delta x)$ changes as a function of $(\mu_{C0}, \mu_{C1}, \sigma_C, \xi_C)$ and of p and Δx: the ratio between quantiles is different across different return periods.

- In the case in which $Y_i \sim GEV(\mu_C, \exp\{\gamma_{C0} + \gamma_{C1}x_i\}, \xi_C)$

$$D^\delta(p, \Delta x) = \frac{\exp\{\gamma_{C0} + \gamma_{C1}x\}}{\xi_C}(y_p^{-\xi_c} - 1)$$
$$(\exp\{\gamma_{C1}\Delta x\} - 1) .$$

The evolution of $D^\delta(p, \Delta x)$ changes as a function of $(\gamma_{C0}, \gamma_{C1}, \xi_C)$ and of p and Δx: the difference between quantiles varies across return periods.

$$M^\delta(p, \Delta x)$$
$$= 1 + \frac{\exp\{\gamma_{C0} + \gamma_{C1}x\}(y_p^{-\xi_c} - 1)(\exp\{\gamma_{C1}\Delta x\} - 1)}{\xi_C\mu_C + \exp\{\gamma_{C0} + \gamma_{C1}x\}(y_p^{-\xi_c} - 1)} .$$

The evolution of $M^\delta(p, \Delta x)$ changes as a function of $(\mu_C, \gamma_{C0}, \gamma_{C1}, \xi_C)$ and of p and Δx: the ratio between quantiles is different across different return periods.

- In the case in which $Y_i \sim GEV(\mu_{C0} + \mu_{C1}x, \exp\{\gamma_{C0} + \gamma_{C1}x_i\}, \xi_C)$:

$$D^\delta(p, \Delta x) = \mu_{C1}\Delta x + \frac{\exp\{\gamma_{C0} + \gamma_{C1}x\}}{\xi_C}(y_p^{-\xi_c} - 1)$$
$$(\exp\{\gamma_{C1}\Delta x\} - 1) .$$

The evolution of $D^\delta(p, \Delta x)$ changes as a function of $(\mu_{C1}, \gamma_{C0}, \gamma_C, \xi_C)$ and of p and Δx: the difference between quantiles varies across different return periods.

$$M^\delta(p, \Delta x)$$
$$= 1 + \frac{\xi_C\mu_{C1}\Delta x + \exp\{\gamma_{C0} + \gamma_{C1}x\}(y_p^{-\xi_c} - 1)(\exp\{\gamma_{C1}\Delta x\} - 1)}{\xi_C(\mu_{C0} + \mu_{C1}x) + \exp\{\gamma_{C0} + \gamma_{C1}x\}(y_p^{-\xi_c} - 1)} .$$

The evolution of $M^\delta(p, \Delta x)$ changes as a function of $(\mu_{C0}, \mu_{C1}, \gamma_{C0}, \gamma_{C1}, \xi_C)$ and of p and Δx: the ratio between quantiles is different across different return periods.

- In the case in which $Y_i \sim GEV(\exp\{\eta_{C0} + \eta_{C1}x_i\}, \sigma_C, \xi_C)$:

$$D^\delta(p, \Delta x) = \exp\{\eta_{C0} + \eta_{C1}x + \eta_{C1}\Delta x\} - \exp\{\eta_{C0} + \eta_{C1}x\} .$$

The evolution of $D^\delta(p, \Delta x)$ changes as a function of (η_{C0}, η_{C1}): the difference between quantiles is constant across different return periods.

$$M^\delta(p, \Delta x) = 1 + \frac{\xi_C \exp\{\eta_{C0} + \eta_{C1}x\}(\exp\{\eta_{C1}\Delta x\} - 1)}{\xi_C \exp\{\eta_{C0} + \eta_{C1}x\} + \sigma_C(y_p^{-\xi_c} - 1)} .$$

The evolution of $M^\delta(p, \Delta x)$ changes as a function of $(\eta_{C0}, \eta_{C1}, \sigma_C, \xi_C)$ and of p and Δx: the ratio between quantiles is different across different return periods.

- In the case in which $Y_i \sim GEV(\exp\{\eta_{C0} + \eta_{C1}x_i\}, \tau\exp\{\eta_{C0} + \eta_{C1}x\}, \xi_C)$:

$$D^\delta(p, \Delta x) = \exp\{\eta_{C0} + \eta_{C1}x\}(\exp\{\eta_{C1}\Delta x\} - 1)$$
$$\left(1 + \frac{\tau}{\xi_C}(y_p^{-\xi_c} - 1)\right) .$$

The evolution of $D^\delta(p, \Delta x)$ changes as a function of $(\eta_{C0}, \eta_{C1}, \tau, \xi_C)$ and of p and Δx: the difference between quantiles varies across different return periods.

$$M^\delta(p, \Delta x) = \exp\{\eta_{C1}\Delta x\} .$$

The evolution of $M^\delta(p, \Delta x)$ changes as a function of η_{C1} and of Δx: the ratio between quantiles is constant across different return periods.

- In the case in which $Y_i \sim GEV(\mu_{C0} + \mu_{C1}x_i, \tau(\mu_{C0} + \mu_{C1}x_i), \xi_C)$:

$$D^\delta(p, \Delta x) = \mu_{C1}\Delta x\left(1 + \frac{\tau}{\xi_C}(y_p^{-\xi_c} - 1)\right) .$$

The evolution of $D^\delta(p, \Delta x)$ changes as a function of (μ_{C1}, τ, ξ_C) and of p and Δx: the difference between quantiles varies across different return periods.

$$M^\delta(p, \Delta x) = 1 + \frac{\mu_{C1}\Delta x}{\mu_{C0} + \mu_{C1}x} .$$

The evolution of $M^\delta(p, \Delta x)$ changes as a function of (μ_{C0}, μ_{C1}) and of Δx: the ratio between quantiles is constant across different return periods.

Calculations for the crossing point between $q(1 - p; \theta)$ and $q(1 - p; \theta_c)$

In the fixed-parameters model it is assumed that $Y_i \sim GEV(\mu, \sigma, \xi)$. The value y_p at which the two quantile functions for the fixed-parameters and change-permitting model cross is derived for a generic change-permitting model below, assuming that the shape parameter under the change-permitting model has the same value ξ as in the fixed-parameters model. In general, under the change-permitting model it is assumed that $Y_i \sim GEV(\mu_C(x_i), \sigma_C(x_i), \xi)$ and the two quantiles function cross at the value p such that $D(p, x^*) = 0$, which is found to be:

$$y_p = \left(1 + \xi\frac{\mu - \mu_C(x^*)}{\sigma_C(x^*) - \sigma}\right)^{-1/\xi} .$$

This entails that in the case in which $Y_i \sim GEV(\mu_C, \exp\{\sigma_{C0} + \sigma_{C1}x_i\}, \xi)$ we find:

$$y_p = \left(1 + \xi\frac{\mu - \mu_C}{\exp\{\sigma_{C0} + \sigma_{C1}x^*\} - \sigma}\right)^{-1/\xi} .$$

R code to implement the model in Eq. (3)

The function below is built upon the `ismev::gev.fit` function in R. The output is an object of class `gev.fit`.

```r
gevcv.fit <- function (xdat, ydat = NULL,
        mul = NULL, taul = NULL, shl = NULL,
        mulink = identity, taulink = identity, shlink = identity,
        muinit = NULL, tauinit = NULL, shinit = NULL, show = TRUE,
        method = "Nelder-Mead", maxit = 10000, ...) {
  z <- list()
  npmu <- length(mul) + 1
  npcv <- length(taul) + 1
  npsh <- length(shl) + 1
  z$trans <- FALSE
  in2 <- sqrt(6 * var(xdat))/pi
  in1 <- mean(xdat) - 0.57722 * in2
  ins <- in2/in1
  if (is.null(mul)) {
    mumat <- as.matrix(rep(1, length(xdat)))
    if (is.null(muinit))
      muinit <- in1
  }
  else {
    z$trans <- TRUE
    mumat <- cbind(rep(1, length(xdat)), ydat[, mul])
    if (is.null(muinit))
      muinit <- c(in1, rep(0, length(mul)))
  }
  if (is.null(taul)) {
    taumat <- as.matrix(rep(1, length(xdat)))
    if (is.null(tauinit))
      tauinit <- in2
  }
  else {
    z$trans <- TRUE
    taumat <- cbind(rep(1, length(xdat)), ydat[, taul])
    if (is.null(tauinit))
      tauinit <- c(in2, rep(0, length(taul)))
  }
  if (is.null(shl)) {
    shmat <- as.matrix(rep(1, length(xdat)))
    if (is.null(shinit))
      shinit <- 0.1
  }
  else {
    z$trans <- TRUE
    shmat <- cbind(rep(1, length(xdat)), ydat[, shl])
    if (is.null(shinit))
      shinit <- c(0.1, rep(0, length(shl)))
  }
  z$model <- list(mul, taul, shl)
  z$link <- deparse(substitute(c(mulink, taulink, shlink)))
  init <- c(muinit, tauinit, shinit)
  gev.lik <- function(a) {
    mu <- mulink(mumat %*% (a[1:npmu]))
    cv <- taulink(taumat %*% (a[seq(npmu + 1, length = npcv)]))
    xi <- shlink(shmat %*% (a[seq(npmu + npcv + 1, length = npsh)]))
    sc <- cv*mu
    y <- (xdat - mu)/sc
    y <- 1 + xi * y
    if (any(y <= 0) || any(sc <= 0))
      return(10^6)
    sum(log(sc)) + sum(y^(-1/xi)) + sum(log(y) * (1/xi +
                                                    1))
  }
  x <- optim(init, gev.lik, hessian = TRUE, method = method,
          control = list(maxit = maxit, ...))
  z$conv <- x$convergence
  mu <- mulink(mumat %*% (x$par[1:npmu]))
  cv <- taulink(taumat %*% (x$par[seq(npmu + 1, length = npcv)]))
  xi <- shlink(shmat %*% (x$par[seq(npmu + npcv + 1, length = npsh)]))
  sc <- mu * cv
  z$nllh <- x$value
  z$data <- xdat
  if (z$trans) {
    z$data <- -log(as.vector((1 + (xi * (xdat - mu))/sc)^(-1/xi)))
  }
  z$mle <- x$par
  z$cov <- solve(x$hessian)
  z$se <- sqrt(diag(z$cov))
  z$vals <- cbind(mu, cv, xi)
  if (show) {
    if (z$trans)
      print(z[c(2, 3, 4)])
    else print(z[4])
    if (!z$conv)
      print(z[c(5, 7, 9)])
  }
  class(z) <- "gev.fit"
  invisible(z)
}
```

References

Blanchet J, Marty C, Lehning M (2009) Extreme value statistics of snowfall in the Swiss Alpine region. Water Resour Res. https://doi.org/10.1029/2009WR007916

Brown SJ, Murphy JM, Sexton DMH, Harris GR (2014) Climate projections of future extreme events accounting for modelling uncertainties and historical simulation biases. Clim Dyn. https://doi.org/10.1007/s00382-014-2080-1

Castellarin A, Kohnová S, Gaál L, Fleig A, Salinas J, Toumazis A, Kjeldsen T, Macdonald N (2012) Review of applied-statistical methods for flood-frequency analysis in Europe. NERC/Centre for Ecology & Hydrology http://nora.nerc.ac.uk/id/eprint/19286

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

Coates L, Haynes K, O'Brien J, McAneney J, de Oliveira FD (2014) Exploring 167 years of vulnerability: an examination of extreme heat events in Australia 1844–2010. Environ Sci Policy 42:33–44. https://doi.org/10.1016/j.envsci.2014.05.003

Coles SG (2001) An introduction to statistical modeling of extreme values. Springer, London

Cooley D (2009) Extreme value analysis and the study of climate change. Clim Change 97(1–2):77. https://doi.org/10.1007/s10584-009-9627-x

Cooley D (2013) Return periods and return levels under climate change. In: Extremes in a changing climate, Springer, pp 97–114

De Cicco LA, Lorenz D, Hirsch RM, Watkins W (2018) dataRetrieval: R packages for discovering and retrieving water data available from U.S. federal hydrologic web services. Reston, VA, https://doi.org/10.5066/P9X4L3GE, https://code.usgs.gov/water/dataRetrieval

Diebold M, Heller P (2019) Scaling wind fields to estimate extreme wave heights in mountainous lakes. J Appl Water Eng Res 7(1):1–9. https://doi.org/10.1080/23249676.2017.1317294

Eastoe EF, Tawn JA (2009) Modelling non-stationary extremes with application to surface level ozone. J R Stat Soc Ser C (Applied Statistics) 58(1):25–45. https://doi.org/10.1111/j.1467-9876.2008.00638.x

François B, Schlef K, Wi S, Brown C (2019) Design considerations for riverine floods in a changing climate—a review. J Hydrol 574:557–573. https://doi.org/10.1016/j.jhydrol.2019.04.068

Gilleland E, Katz RW (2016) extRemes 2.0: an extreme value analysis package in R. J Stat Softw 72(8):1–39. https://doi.org/10.18637/jss.v072.i08

Gilleland E, Heffernan J, Stephenson AG (2018) ismev: an introduction to statistical modeling of extreme values. https://CRAN.R-project.org/package=ismev, R package version 1.42

Griffin A, Vesuviano G, Stewart E (2019) Have trends changed over time? a study of uk peak flow data and sensitivity to observation period. Nat Hazards Earth Syst Sci 19(10):2157–2167. https://doi.org/10.5194/nhess-19-2157-2019

Guerreiro SB, Fowler HJ, Barbero R, Westra S, Lenderink G, Blenkinsop S, Lewis E, Li XF (2018) Detection of continental-scale intensification of hourly rainfall extremes. Nat Clim Change 8(9):803

Gupta VK, Waymire E (1990) Multiscaling properties of spatial rainfall and river flow distributions. J Geophys Res Atmos 95(D3):1999–2009

Hosking JRM, Wallis JR (1997) Regional frequency analysis: an approach based on L-moments. Cambridge University Press, Cambridge

Hu Y, Liang Z, Singh VP, Zhang X, Wang J, Li B, Wang H (2018) Concept of equivalent reliability for estimating the design flood

under non-stationary conditions. Water Resour Manag 32(3):997–1011. https://doi.org/10.1007/s11269-017-1851-y

IPCC (2012) Managing the risks of extreme events and disasters to advance climate change adaptation. A Special Report of Working Groups I and II of the Intergovernmental Panel on Climate Change. Tech. rep., Intergovernmental Panel on Climate Change

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

Kyselỳ J (2008) A cautionary note on the use of nonparametric bootstrap for estimating uncertainties in extreme-value models. J Appl Meteorol Climatol 47(12):3236–3251. https://doi.org/10.1175/2008JAMC1763.1

Lu F, Song X, Xiao W, Zhu K, Xie Z (2020) Detecting the impact of climate and reservoirs on extreme floods using nonstationary frequency models. Stoch Environ Res Risk Assess 34:169–182. https://doi.org/10.1007/s00477-019-01747-2

Luo Y, Zhu LS (2018) Investigation of trends in extreme significant wave heights in the south china sea. Aquatic Ecosyst Health Manag. https://doi.org/10.1080/14634988.2018.1467194

Menabde M, Seed A, Pegram G (1999) A simple scaling model for extreme rainfall. Water Resour Res 35(1):335–339

Overeem A, Buishand T, Holleman I (2009) Extreme rainfall analysis and estimation of depth-duration-frequency curves using weather radar. Water Resour Res. https://doi.org/10.1029/2009WR007869

Rigby RA, Stasinopoulos DM (2005) Generalized additive models for location, scale and shape,(with discussion). Appl Stat 54:507–554

Rootzén H, Katz RW (2013) Design life level: quantifying risk in a changing climate. Water Resour Res 49(9):5964–5972. https://doi.org/10.1002/wrcr.20425

Roth M, Buishand TA, Jongbloed G, Klein Tank AMG, van Zanten JH (2012) A regional peaks-over-threshold model in a nonstationary climate. Water Resour Res 48(11):W11533. https://doi.org/10.1029/2012WR012214

Salas JD, Obeysekera J, Vogel RM (2018) Techniques for assessing water infrastructure for nonstationary extreme events: a review. Hydrol Sci J 63(3):325–352. https://doi.org/10.1080/02626667.2018.1426858

Serago JM, Vogel RM (2018) Parsimonious nonstationary flood frequency analysis. Adv Water Resour 112:1–16. https://doi.org/10.1016/j.advwatres.2017.11.026

Sharma A, Wasko C, Lettenmaier DP (2018) If precipitation extremes are increasing, why aren't floods? Water Resour Res 54(11):8545–8551. https://doi.org/10.1029/2018WR023749

Silva AT, Naghettini M, Portela MM (2016) On some aspects of peaks-over-threshold modeling of floods under nonstationarity using climate covariates. Stoch Environ Res Risk Assess 30(1):207–224. https://doi.org/10.1007/s00477-015-1072-y

Silva AT, Portela MM, Naghettini M, Fernandes W (2017) A Bayesian peaks-over-threshold analysis of floods in the itajaí-açu river under stationarity and nonstationarity. Stoch Environ Res Risk Assess 31(1):185–204. https://doi.org/10.1007/s00477-015-1184-4

Smith RL (1985) Maximum likelihood estimation in a class of nonregular cases. Biometrika 72(1):67–90

Smith RL (1999) Trends in rainfall extremes. Unpublished manuscript, available online https://rls.sites.oasis.unc.edu/postscript/rs/rx.pdf

Šraj M, Viglione A, Parajka J, Blöschl G (2016) The influence of non-stationarity in extreme hydrological events on flood frequency estimation. J Hydrol Hydromech 64(4):426–437

Umlauf N, Kneib T (2018) A primer on Bayesian distributional regression. Stat Model 18(3–4):219–247

Villarini G, Smith JA, Serinaldi F, Bales J, Bates PD, Krajewski WF (2009) Flood frequency analysis for nonstationary annual peak records in an urban drainage basin. Adv Water Resour 32(8):1255–1266. https://doi.org/10.1016/j.advwatres.2009.05.003

Vogel RM, Yaindl C, Walter M (2011) Nonstationarity: flood magnification and recurrence reduction factors in the United States. JAWRA J Am Water Resour Assoc 47(3):464–474. https://doi.org/10.1111/j.1752-1688.2011.00541.x

Volpi E (2019) On return period and probability of failure in hydrology. Wiley Interdisc Rev Water 6(3):e1340. https://doi.org/10.1002/wat2.1340

Wood S (2017) Generalized: additive models an introduction with R, 2nd edn. Chapman and Hall/CRC, London

Yee TW (2015) Vector generalized linear and additive models: with an implementation in R. Springer, New York, USA

Yee TW, Stephenson AG (2007) Vector generalized linear and additive extreme value models. Extremes 10(1–2):1–19

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.