



**HAL**  
open science

## Revising Return Periods for Record Events in a Climate Event Attribution Context

Philippe Naveau, Aurélien Ribes, Francis Zwiers, Alexis Hannart, Alexandre Tuel, Pascal Yiou

► **To cite this version:**

Philippe Naveau, Aurélien Ribes, Francis Zwiers, Alexis Hannart, Alexandre Tuel, et al.. Revising Return Periods for Record Events in a Climate Event Attribution Context. *Journal of Climate*, 2018, 31 (9), pp.3411 - 3422. 10.1175/JCLI-D-16-0752.1 . hal-01806833

**HAL Id: hal-01806833**

**<https://hal.science/hal-01806833>**

Submitted on 9 Jun 2021

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

## Revising Return Periods for Record Events in a Climate Event Attribution Context

PHILIPPE NAVEAU, AURÉLIEN RIBES,<sup>a</sup> FRANCIS ZWIERS,<sup>b</sup> ALEXIS HANNART,<sup>c</sup>  
ALEXANDRE TUEL,<sup>d</sup> AND PASCAL YIOU

*LSCE/IPSL, CEA-CNRS-UVSQ, Université Paris-Saclay, Gif-sur-Yvette, France*

(Manuscript received 19 October 2016, in final form 6 December 2017)

### ABSTRACT

Both climate and statistical models play an essential role in the process of demonstrating that the distribution of some atmospheric variable has changed over time and in establishing the most likely causes for the detected change. One statistical difficulty in the research field of detection and attribution resides in defining events that can be easily compared and accurately inferred from reasonable sample sizes. As many impacts studies focus on extreme events, the inference of small probabilities and the computation of their associated uncertainties quickly become challenging. In the particular context of event attribution, the authors address the question of how to compare records between the counterfactual “world as it might have been” without anthropogenic forcings and the factual “world that is.” Records are often the most important events in terms of impact and get much media attention. The authors will show how to efficiently estimate the ratio of two small probabilities of records. The inferential gain is particularly substantial when a simple hypothesis-testing procedure is implemented. The theoretical justification of such a proposed scheme can be found in extreme value theory. To illustrate this study’s approach, classical indicators in event attribution studies, like the risk ratio or the fraction of attributable risk, are modified and tailored to handle records. The authors illustrate the advantages of their method through theoretical results, simulation studies, temperature records in Paris, and outputs from a numerical climate model.

### 1. Introduction

Civil engineers, flood planners, and other risk managers have to take into account the nonstationarity of our climate system when computing return levels<sup>1</sup> to assess, for example, the quality of existing dams or heights of new dikes and other structures. It is natural to

wonder how a 100-yr return level for a dike built 50 years ago, under the assumption of stationarity, could be different from today’s value. As emphasized by [Rootzén and Katz \(2013\)](#), a delicate issue is the interpretation of return periods and return levels within a nonstationary world. A 100-yr return level computed 50 years ago, today, or in 50 years may differ and become a relative quantity. Another challenge is to discriminate between a set of potential causes in order to explain detected changes (see, e.g., [Zwiers et al. 2013](#); [Hegerl and Zwiers 2011](#); [IPCC 2013](#)).

To deal with relative probability changes and causal questions, researchers in the event attribution community (see, e.g., [Uhe et al. 2016](#)) frequently work with the risk ratio  $RR = p_1/p_0$  of two probabilities of the same event,  $p_0$  and  $p_1$ , under two different situations: a “world that might have been” (a counterfactual world without anthropogenic forcings) and the “world that is” (a factual world). The archetypical example of a common event is that the random variable of interest, say annual maximum temperature, exceeds a high threshold. The so-called fraction of attributable risk  $FAR = 1 - 1/RR$

<sup>1</sup> Given a fixed number of years  $r$ , the return level, say  $x_r$ , is the value that should be crossed on average once during the period  $r$  under the assumption of stationarity. The number  $r$  is called the return period, and, hence, it is associated with an event of probability  $1/r$ . The return level  $x_r$ , associated with period  $r$  can also be defined as the value that is exceeded with probability  $1/r$  in any given year.

<sup>a</sup> Current affiliation: CNRM, Météo France/CNRS, Toulouse, France.

<sup>b</sup> Current affiliation: Pacific Climate Impacts Consortium, University of Victoria, Victoria, British Columbia, Canada.

<sup>c</sup> Current affiliation: Ouranos, Montreal, Québec, Canada.

<sup>d</sup> Current affiliation: École Polytechnique, Palaiseau, France.

*Corresponding author:* Philippe Naveau, naveau@lsce.ipsl.fr

DOI: 10.1175/JCLI-D-16-0752.1

© 2018 American Meteorological Society. For information regarding reuse of this content and general copyright information, consult the [AMS Copyright Policy \(www.ametsoc.org/PUBSReuseLicenses\)](#).

has also been used (see, e.g., Stott et al. 2004, 2016), and it can be viewed as the probability of necessary causation within the mathematical paradigm of causality theory (see Hannart et al. 2016).

Instead of focusing on return period changes, which is a frequent approach in event attributions studies, our goal in this paper is to determine how record event frequencies differ between the factual and counterfactual worlds. Many reasons can be invoked to motivate this change of viewpoint regarding the event of interest. Practically, records have often large economic and human impacts. In terms of media communications, the news anchor typically asks, whenever a record of temperatures or rainfall is broken, if such a record is due to climate change. It is important to have the appropriate probability tool to precisely address this question. Conceptually, record events (see, e.g., King 2017; Meehl et al. 2009; Resnick 1987) are different from maxima and excesses above a high threshold. Reformulating extreme event attribution in terms of records rather than threshold excesses simplifies the statistical issues because only one probability has to be estimated instead of two. Under some additional assumptions based on extreme value theory (EVT; see, e.g., Embrechts et al. 1997; Coles 2001), we will show that the computation of FAR for records can be simplified dramatically and leads to some simple formulas when evaluated through Pearl's causality theory (see Hannart et al. 2016). The results yield interesting conclusions when applied to two datasets of annual maxima of daily temperature maxima: observations recorded in Paris and outputs from the Météo-France CNRM-CM5 climate model. Still, our objective is not to replace the classical FAR and RR, but to offer a complementary instrument for climatologists, hydrologists, and risk managers. Although strongly connected, the interpretation of our new FAR and RR for records will differ from conventional approaches.

#### a. Main assumptions and notations

The probabilistic framework in event attribution studies often starts with the simple and reasonable hypothesis that two independent samples are available, say  $\mathbf{X} = (X_1, \dots, X_m)^T$  with a cumulative distribution function<sup>2</sup> (CDF) of  $G(x) = \mathbb{P}(X \leq x)$ , and  $\mathbf{Z} = (Z_1, \dots, Z_n)^T$  with a CDF of  $F(z) = \mathbb{P}(Z \leq z)$ . The random vectors  $\mathbf{X}$  and  $\mathbf{Z}$  represent ensemble members from numerical models run under two different scenarios. In a causality investigation, the vector  $\mathbf{X}$  is said to provide the relevant

time series in the so-called counterfactual world, while its counterpart  $\mathbf{Z}$  describes the same information but in the factual world.

As relative probability ratios of the type  $p_1/p_0$  or  $(p_1 - p_0)/p_1$  are often the main objects of interest in event attribution, we will always assume that such ratios are well defined and finite. Such a hypothesis is classical within the event attribution community and simplifies mathematical expressions and interpretations. Assuming a finite FAR implies that the right-hand end point of  $G(\cdot)$  is at least as large as that of  $F(\cdot)$ .

#### b. Records and return periods

Breaking a record simply means that the current observation exceeds all past measurements (see Resnick 1987). This type of event is often reported in the media and, in such instances, climatologists are often asked if the frequency of record breaking has increased. To define a record mathematically, suppose that a sample of annual<sup>3</sup> data, say  $\mathbf{Y} = (Y_1, \dots, Y_k)^T$ , is given. The year  $r$  is a record year if

$$Y_r > \max(Y_1, \dots, Y_{r-1}).$$

To relate this mathematical inequality with the media question about the likelihood of this event, one can compute the probability of observing such an event:

$$\mathbb{P}[Y_r > \max(Y_1, \dots, Y_{r-1})] = \frac{1}{r}. \quad (1)$$

This equality holds if the random variables  $Y_i$  are independently and identically distributed (IID). Basically, Eq. (1) emphasizes two points. First, records are expected to occur less frequently as the period length  $r$  increases. Second, given a fixed time period  $r$ , the chance of the record occurring in a given year is uniform,  $1/r$ . Equation (1) will be the backbone of our conceptual and inferential strategies. In particular, it does not depend on the marginal distribution of the  $Y$ , that is, on  $\mathbb{P}(Y \leq y)$ .

#### c. Interpreting records in an event attribution context

Given any return period  $r$ , different types of records can be defined by comparing data from our different worlds. For example, the event  $[Z_r > \max(X_1, \dots, X_{r-1})]$

<sup>2</sup> All CDFs in this article are assumed to be continuous.

<sup>3</sup> For ease of interpretation, the temporal increment will always be a year in this work, but this is arbitrary and other increments could be used instead.

characterizes the fact that  $Z_r$  is a record, relative to the counterfactual world.

Equation (1) can be applied to the counterfactual sample<sup>4</sup>  $\mathbf{X}$ :

$$p_{0,r} = \mathbb{P}[X_r > \max(X_1, \dots, X_{r-1})] = \frac{1}{r}.$$

To assess the probability that a factual observation like  $Z_r$  could be a record in the counterfactual world, it is convenient to introduce

$$p_{1,r} = \mathbb{P}[Z_r > \max(X_1, \dots, X_{r-1})].$$

Comparing  $p_{0,r}$  and  $p_{1,r}$  answers the question of whether the frequency of today’s records was different in the past. The mirrored option would be to compute the chance that a counterfactual realization like  $X_r$  becomes a record in the factual world, that is,  $\mathbb{P}[X_r > \max(Z_1, \dots, Z_{r-1})]$ . Although possible, we do not pursue this here because the “world that might have been” provides a more likely stationary referential than the “world that is.”

To make the link with past event attribution studies, we use the notation  $p_{0,r}$  and  $p_{1,r}$ ; they differ in their interpretation and inference properties from the classical survival functions  $p_0(u) = \mathbb{P}(X > u)$  and  $p_1(u) = \mathbb{P}(Z > u)$ , where  $u$  is any number on the real line, such as  $p_1(u) > 0$ .

Given the return period  $r$ , we have to find the solution of the system defined by  $p_{0,r}$  and  $p_{1,r}$ . By construction,  $p_{0,r}$  is always equal to  $1/r$ , and consequently, there is only one quantity to estimate:  $p_{1,r}$ . The consequence of fixing  $p_{0,r} = 1/r$  is particularly relevant when dealing with intercomparison studies like CMIP5 or the upcoming CMIP6 (see, e.g., Kharin et al. 2013; Taylor et al. 2012). To compare the values of  $p_1$  from different climate models, it would be best to fix  $p_0$  to the same value for all climate models. Otherwise, the interpretation of the ratio  $p_1/p_0$  will only be relative, an important shortcoming for assessments like CMIP5. For example, the ratio  $p_{1,\text{IPSL}}(u)/p_{0,\text{IPSL}}(u)$  for the French IPSL climate model is not directly comparable to that from, say, the American NCAR model because  $p_{0,\text{IPSL}}(u)$  and  $p_{0,\text{NCAR}}(u)$  differ due to model error. This inter-model discrepancy can be large. This issue disappears in our framework because  $p_{0,r,\text{IPSL}} = p_{0,r,\text{NCAR}} = 1/r$  for all  $r \geq 2$ . Having the same reference frequency, it is now possible to compare different relative ratios, that is, to perform intermodel comparisons with respect to the same yardstick.

The remaining task is to estimate  $p_{1,r}$ . To implement it, the rest of the paper will be organized as follows. Section 2 presents our statistical model and its corresponding inferential strategy. Examples will be given in section 3. In particular, the analysis of yearly temperature maxima from a climate model will illustrate the applicability of our approach. Section 4 will conclude by highlighting potential improvements and further directions.

## 2. Statistical model for $p_{1,r}$

A simple approach for deriving explicit formulas for  $p_{1,r}$  is to make the connection<sup>5</sup> with moments and moment generating functions (MGFs) by writing

$$p_{1,r} = \mathbb{E}[G(Z)^{r-1}] = \mathbb{E}\{\exp[-(r-1)W]\}, \quad (2)$$

where the new random variable

$$W = -\log G(Z)$$

takes its values in the interval  $[0, \infty)$  since  $G(Z) \in (0, 1]$ . Equation (2) will play a fundamental role because it will enable us to compute the probability that a factual observation like  $Z_r$  could be a record in the counterfactual world. In particular, we do not need to observe a record at year  $r$  to infer  $p_{1,r} = \mathbb{P}[Z_r > \max(X_1, \dots, X_{r-1})]$ . Instead, we simply need to compute the expectation  $\mathbb{E}[G(Z)^{r-1}]$ . Hence,  $p_{1,r}$  can also be viewed and interpreted as an MGF with respect to the random variable  $W$ .

In the remainder of this work, we will always assume that the random variable  $W$  follows an exponential distribution with mean  $\theta$ . The MGF for an exponential is well known (see, e.g., Henze and Meintanis 2005) and gives a simple form for

$$p_{1,r} = \frac{1}{1 + (r-1)\theta}. \quad (3)$$

A natural question is to wonder which cases obey this exponential constraint for  $W$ .

Table 1 provides three examples from extreme value theory, a well-developed probability theory for distributions of maxima and related upper-tail behaviors (see, e.g., Embrechts et al. 1997; Coles 2001). From this table, one can see that a shifted Gumbel random variable, that is,  $Z = X + \mu$ , produces a random variable  $W = -\log G(Z)$  that is exponentially distributed. The same result can

<sup>4</sup>The IID assumption appears reasonable for yearly maxima in the counterfactual world, the topic of interest in our examples.

<sup>5</sup>The proofs of all our mathematical statements can be found in the appendix.

TABLE 1. Three examples for  $X$  and  $Z$  (see, e.g., Fig. 1). Note that  $X$  and  $Z$  have to have the same lower and upper end points. For the Weibull type, this imposes  $x < 1/(-\xi)$  with  $\sigma > 0$  and  $\mu = (1 - \sigma)/(-\xi)$ .

$X$ Gumbel distributed with CDF
$G(x) = \exp[-\exp(-x)]$ with $x$ real, $\mu \geq 0$ , and
$Z = X + \mu$
$X$ Fréchet distributed with CDF
$G(x) = \exp(-x^{-1/\xi})$ with $x > 0$ , $\xi > 0$
$Z = \sigma X$ with $\sigma > 0$
$X$ Weibull distributed with CDF
$G(x) = \exp[-(1 + \xi x)_+^{-1/\xi}]$ with $\xi < 0$
$Z = \mu + \sigma X$

be obtained from an inflated Fréchet, that is,  $Z = \sigma X$  with  $\sigma > 0$ , or from a Weibull-EVT type,<sup>6</sup> that is,  $Z = \mu + \sigma X$ . This hints that block maxima, a block being a season or a year, can produce factual and counterfactual records that are in compliance with an exponentially distributed  $W$ . In practice, we do not have to fit a generalized extreme value distribution (GEV) to the factual and counterfactual data. We need to assume that  $W$  is exponentially distributed.

a. Testing the null hypothesis,  $H_0: W \sim \exp(\theta)$

The random variable  $W = -\log G(Z)$  can be approximated by its empirical version

$$\hat{W}_i = -\log \hat{G}_m(Z_i),$$

where  $i = 1, \dots, n$ , and  $\hat{G}_m(z)$  represents a slightly modified<sup>7</sup> version of the empirical CDF of the sample  $\mathbf{X}$ . To assess the goodness of fit with respect to the null hypothesis,  $H_0: W \sim \exp(\theta)$  with an unknown  $\theta$  against general alternatives, we refer to Henze and Meintanis (2005) and Ascher (1990). Both authors indicated that the test of Cox and Oates (1984) performs well for a wide range of testing alternatives. Its statistic is given by  $S_n = n + \sum_{i=1}^n [1 - (\hat{W}_i/\bar{W}_i) \log(\hat{W}_i/\bar{W}_i)]$  with  $\bar{W}_i = (\hat{W}_1 + \dots + \hat{W}_n)/n$ . It can be shown that  $\sqrt{6/n} S_n/\pi$  is asymptotically standard normal under  $H_0$  (see, e.g., Henze and Meintanis 2005).

<sup>6</sup> With the constraint that  $X$  and  $Z$  share the same support, that is,  $x < 1/(-\xi)$  with  $\sigma > 0$  and  $\mu = (1 - \sigma)/(-\xi)$ .

<sup>7</sup> As  $\mathbb{G}_m(Z_i)$ , with  $\mathbb{G}_m$  representing the empirical CDF, could be equal to zero for any  $i$  such  $Z_i < \min(X_1, \dots, X_n)$ , the expression  $-\log \mathbb{G}_m(Z_i)$  could provide  $-\log 0 = \infty$ . This explains why, to compute  $\hat{W}_i = -\log \hat{G}_m(Z_i)$ , we use a kernel estimate  $\hat{G}_m$  with infinite support, and not  $\mathbb{G}_m$ , to approximate the CDF  $G$  (see the np package of Racine and Hayfield in R). This approximation is only used for the test  $H_0$ , but not to estimate  $p_{1,r}$ , and consequently  $\theta$ . See Eq. (4).

b. Inference of  $\theta$  under  $H_0: W \sim \exp(\theta)$

The constant  $\theta$ , which plays an essential role, can be interpreted in different ways. Since  $p_{1,2} = \mathbb{P}(Z > X) = 1/(1 + \theta)$ , the constant  $\theta$  can be viewed as the ratio  $\theta = \mathbb{P}(Z \leq X)/\mathbb{P}(Z > X)$  that measures a type of fatness between  $X$  and  $Z$ . If  $X$  and  $Z$  have the same distribution, then  $\theta$  equals one. If  $\theta$  is smaller than one, then  $Z$  is more likely than  $X$  to take large values, and vice versa. The constant  $\theta$  is also the limit of the ratio  $p_{0,r}/p_{1,r}$  as  $r$  goes to infinity.

Under  $H_0: W \sim \exp(\theta)$ , we need to estimate the single parameter  $\theta$ . As  $p_{1,2} = 1/(1 + \theta)$  with  $p_{1,2} = \mathbb{E}[G(Z)]$ , we define

$$\hat{\theta} = \frac{1}{\hat{p}_{1,2}} - 1 \quad \text{with} \quad \hat{p}_{1,2} = \frac{1}{n} \sum_{i=1}^n \mathbb{G}_m(Z_i),$$

where  $\mathbb{G}_m(\cdot)$  corresponds to the empirical CDF associated with the sample  $\mathbf{X}$ . The asymptotic properties<sup>8</sup> of  $\hat{\theta}$  can be deduced from

$$\sqrt{n}(\hat{p}_{1,2} - p_{1,2}) \rightarrow N(0, \lambda_2^2), \tag{4}$$

where

$$\lambda_2^2 = \frac{1}{1 + 2\theta} - \frac{2}{(1 + \theta)^2} + \frac{2}{(1 + \theta)(2 + \theta)}.$$

The delta method (see, e.g., Proposition 6.1.5 in Brockwell and Davis 1991) gives

$$\sqrt{n}(\hat{\theta} - \theta) \rightarrow N(0, \sigma_\theta^2), \tag{5}$$

with  $\sigma_\theta^2 = \lambda_2^2/p_{1,2}^2 = [(1 + \theta)^2/(1 + 2\theta)] - 2 + [2(1 + \theta)/(2 + \theta)]$ . This leads to the  $(1 - \alpha) \times 100\%$  asymptotic confidence interval for  $\theta$ ;  $\hat{\theta} \pm z_{1-\alpha/2}(\hat{\sigma}_\theta/\sqrt{n})$ , where  $z_{1-\alpha/2}$  represents the standard normal quantile at  $1 - \alpha/2$ .

c. Inference of  $\text{far}(r)$  and risk ratio under

$$H_0: W \sim \exp(\theta)$$

From the expression for  $p_{1,r}$  in Eq. (3), the FAR for records defined by  $\text{far}(r) = 1 - p_{0,r}/p_{1,r}$  can be expressed as

$$\text{far}(r) = 1 - \frac{1}{r p_{1,r}} = (1 - \theta) \left(1 - \frac{1}{r}\right). \tag{6}$$

Table 2 provides  $\text{far}(r)$  and  $\text{FAR}(u)$  expressions for our three GEV examples defined in Table 1. Figure 1 compares  $\text{far}(r)$  with  $\text{FAR}(x_r)$  for the same three distributions, where  $x_r = G^{-1}(1 - 1/r)$  is the classical return

<sup>8</sup> See the appendix for the proof.

TABLE 2. Explicit FAR( $u$ ) and far( $r$ ) for the three examples in Table 1. The “limit” row indicates the asymptotic value of far( $r$ ) as  $r \uparrow \infty$  and, equivalently, the FAR( $u$ ) asymptote as  $u$  moves toward the upper support bound. See also Fig. 1.

$X$ Gumbel distributed and $Z = X + \mu$	
FAR( $u$ )	$1 - \frac{1 - \exp[-\exp(-u)]}{1 - \exp\{-\exp[-(u - \mu)]\}}$
far( $r$ )	$[1 - \exp(-\mu)](1 - 1/r)$
Limit	$1 - \exp(-\mu)$
$X$ Fréchet distributed and $Z = \sigma X$	
FAR( $u$ )	$1 - \frac{1 - \exp(-u^{-1/\xi})}{1 - \exp[-(u/\sigma)^{-1/\xi}]}$
far( $r$ )	$(1 - \sigma^{-1/\xi})(1 - 1/r)$
Limit	$1 - \sigma^{-1/\xi}$
$X$ Weibull distributed and $Z = \mu + \sigma X$	
FAR( $u$ )	$1 - \frac{1 - \exp[-(1 + \xi u)^{-1/\xi}]}{1 - \exp[-\sigma^{1/\xi}(1 + \xi u)^{-1/\xi}]}$
far( $r$ )	$(1 - \sigma^{-1/\xi})(1 - 1/r)$
Limit	$1 - \sigma^{-1/\xi}$

level. The limits of far( $r$ ) and FAR( $x_r$ ) as  $r$  becomes large are equal (horizontal dashed lines; see also the “Limit” rows in Table 2). In other words, far( $r$ ) and FAR( $u$ ) are asymptotically equivalent for extremes.

Equation (6) naturally suggests the estimator

$$\widehat{\text{far}}(r) = (1 - \hat{\theta}) \left(1 - \frac{1}{r}\right). \tag{7}$$

The delta method applied to Eq. (5) gives

$$\sqrt{n}[\widehat{\text{far}}(r) - \text{far}(r)] \rightarrow N \left[0, \left(1 - \frac{1}{r}\right)^2 \sigma_\theta^2\right],$$

and the asymptotic confidence interval for far( $r$ ) is  $\widehat{\text{far}}(r) \pm z_{1-\alpha/2}(1/\sqrt{n})(1 - 1/r)\hat{\sigma}_\theta$ . The relative error of  $\widehat{\text{far}}(r)$  defined as the renormalized standard deviation does not depend on  $r$ . For  $\theta \neq 1$ , it is equal to

$$\frac{\sqrt{\text{var}[\widehat{\text{far}}(r)]}}{\widehat{\text{far}}(r)} = \frac{\sqrt{\text{var}(\hat{\theta})}}{1 - \theta} = \frac{\sigma_\theta}{1 - \theta}.$$

Also note that this ratio is infinite when  $\theta = 1$  (no difference between the factual and counterfactual worlds). Hence, it is counterproductive to compute far( $r$ ) in this special case.

The record risk ratio  $\text{rr}(r) = p_{1,r}/p_{0,r}$  under model (6) does not have this issue. For the estimator

$$\widehat{\text{rr}}(r) = \frac{r}{1 + (r - 1)\hat{\theta}},$$

we have

$$\sqrt{n}[\widehat{\text{rr}}(r) - \text{rr}(r)] \rightarrow N \left\{0, \frac{r^2(r - 1)^2}{[1 + (r - 1)\theta]^4} \sigma_\theta^2\right\},$$

and consequently, this relative error of the risk ratio satisfies

$$\frac{\sqrt{\text{var}[\widehat{\text{rr}}(r)]}}{\widehat{\text{rr}}(r)} \underset{r \uparrow \infty}{\sim} \frac{\sigma_\theta}{\theta^2},$$

and it is well defined for  $\theta = 1$ .

### 3. Examples

#### a. Annual temperature maxima in Paris

To illustrate how our approach can be applied to climatological time series, we study one of the longest records of annual maxima of daily temperature maxima in Paris (weather station of Montsouris). To represent the factual and counterfactual worlds, this Paris time series is divided into two parts: 1900–30 and 1990–2015. Although the influence of anthropogenic forcing may not be completely absent in the first part of the record, Fig. 2 still shows that the hypothesis of exponentially distributed  $W_i$  seems reasonable. The histogram and the q-q plot do not contradict the exponential hypothesis. The top-right panel compares a decreasing convex dotted line (our fitted exponential model) with the observational histogram that is also decreasing and convex. The lower-left panel shows the ranked observations versus the ones expected from our model. Being close to the diagonal (the solid line) indicates a good fit.

To complement these visual devices, our exponential fit is compared, via the Akaike information criterion (AIC) and Bayesian information criterion (BIC), to a Weibull fit, a natural but more complex alternative. The lower scores for the exponential pdf (AIC = 11.83 and BIC = 13.05 vs AIC = 13.2 and BIC = 15.7) confirms our preference for the simpler model. Concerning the test of Cox and Oates (1984), the inferred  $p$  value is 0.51, indicating that there is no reason to reject the exponential model for our Parisian data. The upper-left panel of Fig. 2 displays far( $r$ ) with its associated 90% confidence interval. Even from a single location like Paris, one can deduce that far( $r$ ) is not equal to zero, but is rather around 0.5.

#### b. Analysis of yearly temperature maxima from CNRM-CM5

To illustrate our method with numerical climate model outputs, we study yearly maxima of daily maxima of near-surface air temperature generated by the Météo-France CNRM-CM5 climate model. We focus on a

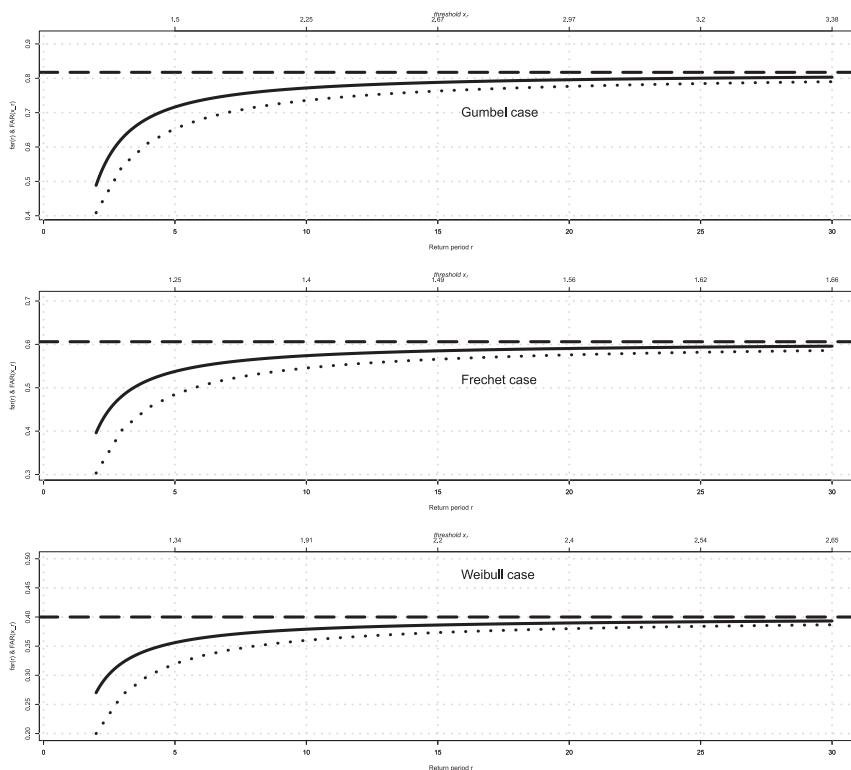


FIG. 1. Comparison of  $\text{far}(r)$  (dotted lines) defined by Eq. (6) with  $\text{FAR}(x_r)$  (solid lines). The horizontal dashed lines correspond to the limits shown in Table 2.

single climate simulation with “all forcings”<sup>9</sup> for the period 1975–2005. This should mimic the current climatology, that is, the counterfactual world, at least according to this model. During this 30-yr climatological period, we assume that the climate system is approximately stationary with respect to annual maxima. Concerning the “world that might have been” without anthropogenic forcings, we consider a second run from the same numerical model, but with only natural forcings. For this counterfactual world, we assume stationarity over the full length of the simulation, that is, for the period 1850–2012.

Our first inquiry is to see if our null hypothesis on  $W$  is rejected or not. The white areas in Fig. 3 correspond to the locations where the exponential hypothesis has been rejected by the Cox and Oates test at the 5% level. For the other points, the estimated value of  $\theta$  inferred from Eq. (5) ranges from 0.2 to 1.2. A  $\theta$  near one (orange) indicates identical factual and counterfactual worlds. A  $\theta$  close to zero (blue) implies that the factual and counterfactual worlds strongly differ with an increase

in temperature records. Note that the estimated  $\theta$  can be larger than one, meaning that  $\hat{p}_1 < \hat{p}_0$ . Consequently, the risk ratio should be used instead of the  $\text{far}$  in such cases.

Figure 4 displays the largest values of  $p_{1,r} - p_{0,r}$  over all values of  $r \geq 2$ . In causality theory, this difference is called the probability of sufficient and necessary causations, and under  $H_0$ , it is equal to  $p_{1,r\theta} - p_{0,r\theta} = (1 - \sqrt{\theta})/(1 + \sqrt{\theta})$  with  $r_\theta = 1 + 1/\sqrt{\theta}$ . These probabilities summarize evidence of both sufficient and necessary causation (see, e.g., Hannart et al. 2016). The subset of locations such that the Cox and Oates  $p$  values are greater than 5% and  $\hat{\theta} < 1$  are displayed in Fig. 4. For example, this number for Iceland is 35%.

Interpreting the two maps jointly highlights the African continent, Brazil, and western Australia, which seem to represent regions with low  $\theta$  and large  $p$  values, that is, with a strong contrast between the all forcing and only natural forcing runs. Northern Brazil, most of Africa, and western Australia appear to show evidence of such causations most strongly. Variability in surface air temperature is low in tropical regions and over many ocean areas, and it is strongly reduced in areas of sea ice loss, so we expect large relative differences between the counterfactual and factual worlds in these cases. Simulated variability is larger in other places, relative to the simulated change in mean climate. As a consequence, both PN and PS remain

<sup>9</sup> That is, the model is driven by specifying the history of greenhouse gas concentrations, aerosols, land use, volcanic aerosol, and solar output over the period of simulation from preindustrial times up to 2005.

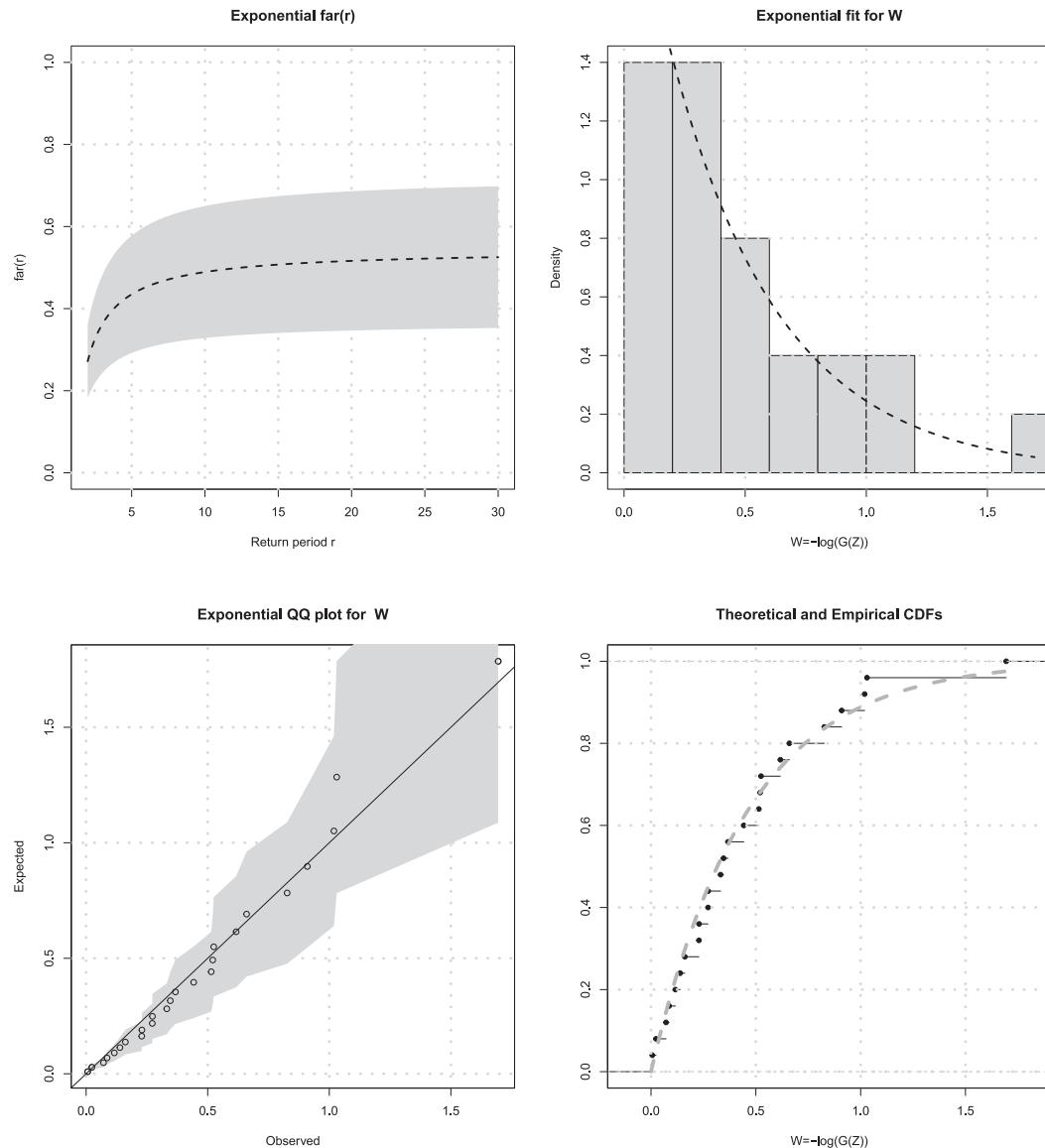


FIG. 2. Analysis of annual maxima of daily temperature maxima recorded at the weather station of Montsouris in Paris. To compute  $\text{far}(r)$ , the period 1900–30 plays the role of the counterfactual world, and the recent climatology 1990–2015 represents the factual world. The 90% confidence intervals are based on Eq. (5).

small locally in these regions. This includes the Mediterranean basin, which is sometimes referred to as a “hot spot.” Note that given the period of time considered for the “factual world,” non-GHG forcings like anthropogenic aerosols can contribute to this result (offsetting part of the GHG-induced warming).

#### 4. Discussions and future work

##### a. Event definition

Our work is closely related to the fundamental and sometimes overlooked question of how to define the

event itself. Ideally, the event of interest should be easily interpretable, its probability should be simple to compute in both worlds, and it should belong to an invariant reference framework in order to compare different studies. These desiderata were fulfilled by slightly moving away from the idea of defining the event as an excess over a fixed threshold and instead setting the event as a record.

A potential worry could have been that treating record events instead of excesses above the threshold  $u$  could drastically move us away from  $\text{FAR}(u)$ . Figure 1 showed that this concern was unfounded. In other words, the practitioner interested in assessing the likelihood of

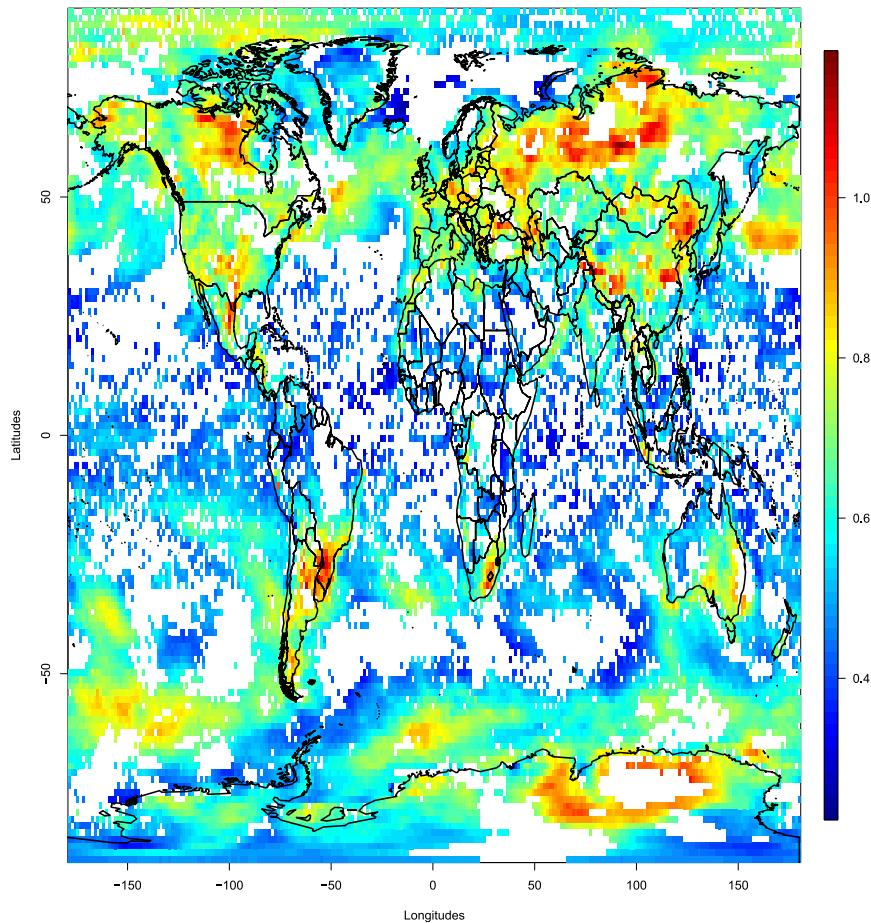


FIG. 3. Comparing yearly temperature maxima from two runs of the Météo-France CNRM-CM5 climate model. The factual run has all forcings (1975–2005), and the counterfactual has only natural forcings (1850–2012). This figure displays  $\hat{\theta}$  for the locations such that the  $p$  value of the Cox and Oates test ( $H_0$ :  $W$  is exponential) is greater than 5%. The white areas correspond to locations where the null hypothesis can be rejected at the 5% level.

extreme events relative to a counterfactual scenario has the choice of using  $\text{far}(r)$  or  $\text{FAR}(x_r)$  for large  $r$  or, equivalently,  $\text{rr}(r)$  or  $\text{RR}(x_r)$ .

#### b. Null hypothesis

Statistically, inferences of the probabilities of interest are simple, fast, and efficient under  $H_0$ :  $W \sim \exp(\theta)$ . In particular, such an assumption is satisfied for simple GEV setups (see Table 1). Theoretically, it is rooted on the variable  $W = -\log G(Z)$ . This transformed variable shows that the interest is neither in only  $Z$  nor only  $X$  (throughout the CDF  $G$ ) but in how they combine. In other words, our main message is that it is counterproductive to minutely study the two samples  $\mathbf{X}$  and  $\mathbf{Z}$  independently. As the FAR and risk ratios are relative quantities, working with the relative statistic  $W$  greatly simplifies their inferences.

Concerning the applications, our analysis of a long temperature record in Paris and of a climate model indicates that the exponential assumption for  $W$  appears to be reasonable for temperature maxima over land areas. Still, a comprehensive study with a large number of both weather stations and climate models is needed to confirm this hypothesis. Future work should also focus on precipitation maxima. In this situation, we conjecture that the Weibull MGF, with its extra parameter  $k$ , could be a better fit than the exponential MGF because extreme precipitation has a stronger variability. The Weibull parameter  $k$  should capture the changes between the factual and counterfactual tails. Another option when  $H_0$  is rejected is to estimate nonparametrically  $p_{1,r}$  [see Eq. (A1)], which provides the necessary information to compute confidence intervals in this case. The main advantage of this approach is that goodness-of-fit tests are

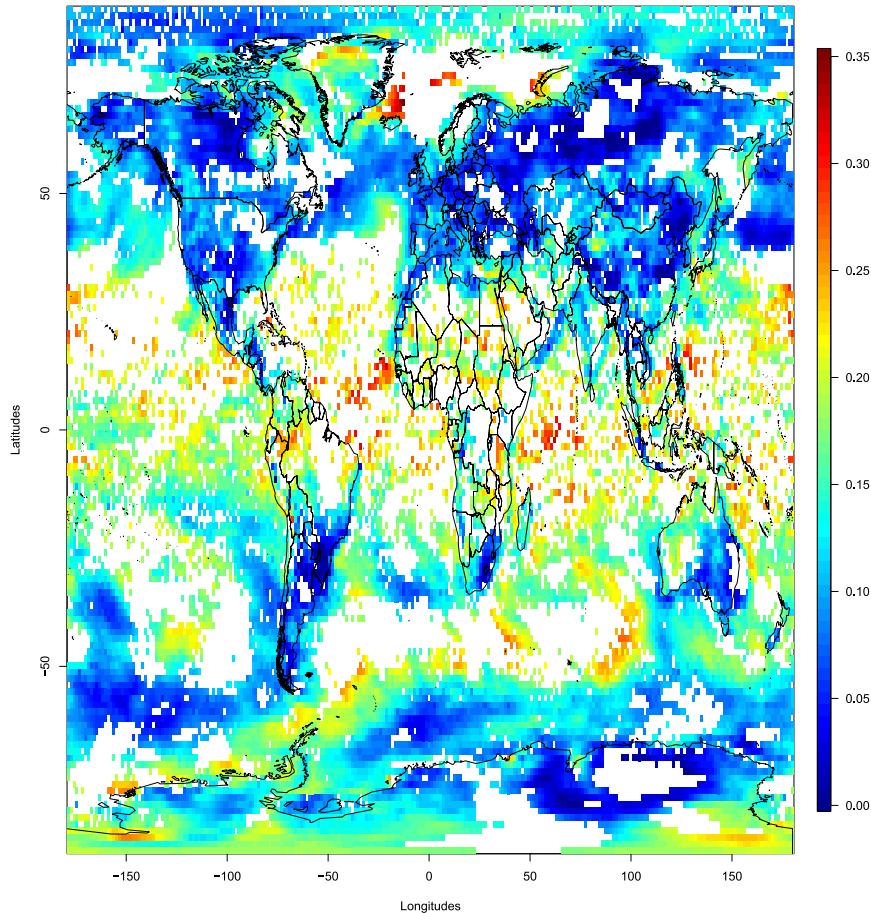


FIG. 4. As in Fig. 3, but for probabilities of sufficient and necessary causations at locations such that  $\theta \leq 1$ , and the Cox and Oates  $p$  value is greater than 5%. This probability is defined as  $p_{1,r_\theta} - p_{0,r_\theta} = (1 - \sqrt{\theta})(1 + \sqrt{\theta})$ , with  $r_\theta = 1 + 1/\sqrt{\theta}$ .

not required, but its main drawback is that the sample size  $n$  should always be much greater than the return period of interest so  $\sqrt{r}/n$  goes to zero.

To summarize our findings about  $H_0$ , it is important to emphasize that this hypothesis allows us to work with small sample sizes; see our Paris example with  $n = 30$ . This is a huge difference with nonparametric methods that need thousands of factual and counterfactual simulations. Hence, our methodology is fairly general. One possible application could be conditional attributions, when the event itself can be linked to another covariate, like the atmospheric circulation. Further work on  $\text{far}(r)$  would be to estimate this value's conditional to atmospheric circulation properties, in order to separate the roles of changes in circulation and temperature increases (see, e.g., [Yiou et al. 2017](#)).

Finally, our goal in this paper is not to discard the classical  $\text{FAR}(u)$  and  $\text{RR}(u)$  in favor of  $\text{far}(r)$  and  $\text{rr}(r)$ . The two have different interpretations, and they complement each other. Here, the new  $p_{0,r}$  and  $p_{1,r}$  should be viewed as

an additional risk management tool for practitioners interested in causality, especially those who base their decision-making process on records. Conceptually, our approach has the advantage of providing a fixed reference framework, which is a necessity for intercomparison studies and could help future intercomparison studies like CMIP6.

*Acknowledgments.* This work is supported by ERC Grant 338965-A2C2 and P. Naveau's work was partially financed by the EUPHEME project (ERA4CS ERA-NET).

## APPENDIX

### Mathematical Computation

#### a. Proof of Eq. (2)

Since  $F(z) = \mathbb{P}(Z \leq z)$ , we can write that

$$\begin{aligned}
 p_{1,r} &= \mathbb{P}[Z_r > \max(X_1, \dots, X_{r-1})] \\
 &= \int \mathbb{P}[z > \max(X_1, \dots, X_{r-1})] dF(z), \\
 &= \int \mathbb{P}^{r-1}(X < z) dF(z), \quad \text{since } X_i \text{ iid,} \\
 &= \mathbb{E}[G(Z)^{r-1}], \quad \text{since } G(z) = \mathbb{P}(X \leq z).
 \end{aligned}$$

b. Proof of Eq. (4)

In this section, we prove a more general result than Eq. (4); the latter can be deduced by taking  $r = 2$  in the following proposition.

Let

$$\hat{p}_{1,r} = \frac{1}{n} \sum_{i=1}^n \mathbb{G}_n^{r-1}(Z_i),$$

where  $\mathbb{G}_n(z) = \frac{1}{m} \sum_{j=1}^m \mathbb{I}(X_j \leq z)$ . Then, we have

$$\sqrt{n}(\hat{p}_{1,r} - p_{1,r}) \rightarrow N(0, \lambda_r^2), \tag{A1}$$

with

$$\begin{aligned}
 \lambda_r^2 &= p_{1,2r-1} - p_{1,r}^2 \\
 &+ (r-1)^2 (\mathbb{E}\{G^{r-2}(Z_1)G^{r-2}(Z_2)[G(Z_1) \wedge G(Z_2)]\} \\
 &- p_{1,r}^2).
 \end{aligned}$$

Note that the estimator  $\hat{p}_{1,r}$  and its asymptotic properties do not depend on the null hypothesis:  $H_0: W \sim \exp(\theta)$ . In other words,  $\hat{p}_{1,r}$  is a very general nonparametric estimator of  $p_{1,r}$  for any integer  $r \geq 2$ .

The key point for the proof of Eq. (A1) is the following result about the empirical process  $\mathbb{G}_m$  (see Mason and Van Zwet 1987):

$$\sup_z |\sqrt{m}[\mathbb{G}_m(z) - G(z)] - B[G(z)]| = O\left(\frac{\log m}{\sqrt{m}}\right), \tag{A2}$$

where  $B(u)$  represents a classical Brownian bridge on  $[0, 1]$  with  $\mathbb{E}[B(u)] = 0$  and covariance

$$\text{cov}[B(u), B(v)] = (u \wedge v) - uv.$$

It follows that for all  $i = 1, \dots, n$ , we can write uniformly for any  $z_i$

$$\sqrt{m}[\mathbb{G}_m(z_i) - G(z_i)] = B[G(z_i)] + O\left(\frac{\log m}{\sqrt{m}}\right).$$

Letting  $h(u) = u^{r-1}$ , we can write that

$$\begin{aligned}
 \sqrt{n}(\hat{p}_{1,r} - p_{1,r}) &= \sqrt{n} \left\{ \frac{1}{n} \sum_{i=1}^n h[\mathbb{G}_n(Z_i)] - h[G(Z_i)] \right\} \\
 &+ \sqrt{n} \left\{ \frac{1}{n} \sum_{i=1}^n h[G(Z_i)] - p_{1,r} \right\} = N_1 + N_2,
 \end{aligned}$$

with obvious notation for  $N_1$  and  $N_2$ . Concerning the latter, the central limit theorem tells us that  $\sqrt{n} \left\{ (1/n) \sum_{i=1}^n h[G(Z_i)] - p_{1,r} \right\}$  converges to a zero-mean Gaussian variable with variance  $p_{1,2r-1} - p_{1,r}^2$ . Concerning  $N_1$ , the function  $h(u) = u^{r-1}$  being a differentiable function on  $[0, 1]$ , Proposition 6.1.5 in Brockwell and Davis (1991) allows us to focus on the approximation of  $N_1$  (the so-called delta-s method):

$$\begin{aligned}
 \sqrt{n} \left\{ \frac{1}{n} \sum_{i=1}^n h'[\mathbb{G}_n(Z_i)] [\mathbb{G}_n(Z_i) - G(Z_i)] \right\} \\
 = \frac{1}{n} \sum_{i=1}^n h'[G(Z_i)] B[G(Z_i)] + O\left(\frac{\log n}{\sqrt{n}}\right).
 \end{aligned}$$

As  $\mathbb{E}[B(u)] = 0$ , we have

$$\begin{aligned}
 \mathbb{E}\{h'[G(Z_i)] B[G(Z_i)]\} &= \mathbb{E}\{h'[G(Z_i)] \mathbb{E}_{Z_i}\{B[G(Z_i)] | Z_i\}\} \\
 &= 0.
 \end{aligned}$$

Using  $\text{cov}[B(u), B(v)] = (u \wedge v) - uv$ , we have

$$\begin{aligned}
 \mathbb{E}\{h'[G(Z_i)] B[G(Z_i)] h'[G(Z_j)] B[G(Z_j)]\} &= \mathbb{E}\{h'[G(Z_i)] h'[G(Z_j)] \{[G(Z_i) \wedge G(Z_j)] - G(Z_i)G(Z_j)\}\} \\
 &= (r-1)^2 \mathbb{E}\{G^{r-2}(Z_i) G^{r-2}(Z_j) \{[G(Z_i) \wedge G(Z_j)] - G(Z_i)G(Z_j)\}\}.
 \end{aligned}$$

It follows that the covariance of  $(1/n) \sum_{i=1}^n h'[G(Z_i)] B[G(Z_i)]$  is equal to

$$\begin{aligned}
 &\frac{1}{n} (\mathbb{E}\{G^{2r-4}(Z)[G(Z) - G^2(Z)]\} \\
 &+ (n-1)(r-1)^2 (\mathbb{E}\{G^{r-2}(Z_1)G^{r-2}(Z_2)[G(Z_1) \wedge G(Z_2)]\} \\
 &- \mathbb{E}^2[G^{r-1}(Z)]).
 \end{aligned}$$

For large  $n$ , this leads to

$$\begin{aligned}
 &\text{var} \left\{ \frac{1}{n} \sum_{i=1}^n h'[G(Z_i)] B[G(Z_i)] \right\} \\
 &\sim (r-1)^2 (\mathbb{E}\{G^{r-2}(Z_1)G^{r-2}(Z_2)[G(Z_1) \wedge G(Z_2)]\} \\
 &- \mathbb{E}^2[G^{r-1}(Z)]).
 \end{aligned}$$

In other words, the random variable  $\sqrt{n}\{(1/n)\sum_{i=1}^n h'[G(Z_i)][G_n(Z_i) - G(Z_i)]\}$  asymptotically follows a zero-mean Gaussian distribution with variance  $(r - 1)^2(\mathbb{E}\{G^{r-2}(Z_1)G^{r-2}(Z_2)[G(Z_1) \wedge G(Z_2)]\} - p_{1,r}^2)$ . In addition, the covariance

$$\mathbb{E}\{h'[G(Z_i)]B[G(Z_i)]\} = \mathbb{E}\{h'[G(Z_i)]\mathbb{E}\{B[G(Z_i)]|Z_i\}\}$$

is equal to zero because  $\mathbb{E}\{B(u)\} = 0$ . In summary, the sum  $N_1 + N_2$  can be viewed (asymptotically in a distributional sense) as the sum of two zero-mean Gaussian random variables, with total variance

$$\lambda_r^2 = p_{1,2r-1} - [1 + (r - 1)^2]p_{1,r}^2 + (r - 1)^2\mathbb{E}\{G^{r-2}(Z_1)G^{r-2}(Z_2)[G(Z_1) \wedge G(Z_2)]\}.$$

For the special case defined by Eq. (3), we can write that

$$\lambda_r^2 = \frac{1}{1 + (2r - 2)\theta} - \frac{1}{[1 + (r - 1)\theta]^2} + (r - 1)^2 \left\{ \frac{2}{[1 + \theta(r - 1)][2 + \theta(2r - 3)]} - \frac{1}{[1 + (r - 1)\theta]^2} \right\}.$$

As  $\lambda_r^2 = 1/(2r\theta)[1 + (1/\theta)] + O(1/r)$  and  $\sigma_r = \lambda_r/(rp_{1,r}^2)$ , we can write

$$\sigma_r \sim \sqrt{r} \sqrt{\frac{1 + \theta}{2}}, \text{ for large } r.$$

This indicates that this nonparametric technique should be used with caution for large  $r$  as the standard deviation increases with the rate  $\sqrt{r}$ . The sample size  $n$  should be always much greater than the return period of interest so  $\sqrt{r/n}$  goes to zero.

*c. Computation of  $p_{1,r}$  in the GEV case*

For  $\xi_i \neq 0$ , we denote

$$T_i(x) = \left( 1 + \xi_i \frac{x - \mu_i}{\sigma_i} \right)_+^{1/\xi_i},$$

for  $i \in \{0, 1\}$ , and we have  $T_i^{-1}(x) = \mu_i + (\sigma_i/\xi_i)(x^{\xi_i} - 1)$ . For any GEV distributions in our tables, we have  $F(x) =$

$\exp[-1/T_1(x)]$  and  $G(x) = \exp[-1/T_0(x)]$ . In this context, we can write

$$\begin{aligned} T_1[T_0^{-1}(x)] &= \left[ 1 + \xi_1 \frac{\mu_0 + \frac{\sigma_0}{\xi_0}(x^{\xi_0} - 1) - \mu_1}{\sigma_1} \right]^{1/\xi_1}, \\ &= \left[ 1 + \frac{\xi_1}{\sigma_1}(\mu_0 - \mu_1) - \frac{\sigma_0}{\sigma_1} \frac{\xi_1}{\xi_0} + \frac{\sigma_0}{\sigma_1} \frac{\xi_1}{\xi_0} x^{\xi_0} \right]^{1/\xi_1}, \\ &= \left( \frac{\xi_1 \sigma_0}{\xi_0 \sigma_1} \right)^{1/\xi_1} x^{\xi_0/\xi_1}, \end{aligned} \tag{A3}$$

whenever  $\mu_0 - \mu_1 = \sigma_0/\xi_0 - \sigma_1/\xi_1$  (same support condition in the Weibull and Fréchet cases).

For  $\xi_0 = \xi_1 = 0$ , we set for  $i \in \{0, 1\}$ ,  $T_i(x) = \exp[(x - \mu_i)/\sigma_i]$  and  $T_i^{-1}(x) = \mu_i + \sigma_i \ln x$ . Then, we have

$$T_1[T_0^{-1}(x)] = x^{\sigma_0/\sigma_1} \exp\left(\frac{\mu_0 - \mu_1}{\sigma_1}\right). \tag{A4}$$

As  $T_1(\cdot)$  is a nondecreasing function, we can write that

$$A = \mathbb{P}\{T_1(Z_r) \geq \max[T_1(X_1), \dots, T_1(X_{r-1})]\}.$$

To continue, we need to derive the pdf of  $T_1(Z_r)$ . If  $Z$  follows a GEV( $\mu_1, \sigma_1, \xi_1$ ), then

$$\begin{aligned} \mathbb{P}\{T_1(Z_r) \leq x\} &= \mathbb{P}\{Z_r \leq T_1^{-1}(x)\}, \\ &= \exp\left\{-\frac{1}{T_1[T_1^{-1}(x)]}\right\} = \exp\left(-\frac{1}{x}\right). \end{aligned}$$

Hence,  $T_1(Z_r)$  is unit-Fréchet. Going back to the computation of  $p_{1,r}$ , we can use the independence assumption within the sample  $\mathbf{X}$  to write that

$$\begin{aligned} p_{1,r} &= \int \mathbb{P}\{\max[T_1(X_1), \dots, T_1(X_{r-1}) \leq x]\} dF_{T_1(Z_r)}(x), \\ &= \int [\mathbb{P}\{T_1(X) \leq x\}]^{r-1} dF_{T_1(Z_r)}(x). \end{aligned}$$

We need to compute

$$\mathbb{P}\{T_1(X) \leq x\} = \mathbb{P}\{X \leq T_1^{-1}(x)\} = \exp\left(-\frac{1}{T_0[T_1^{-1}(x)]}\right).$$

To finish the computation, we study separately the Gumbel and Fréchet cases. If  $\xi_0 = \xi_1 = 0$ , then  $T_0[T_1^{-1}(x)] = x^{\sigma_1/\sigma_0} \exp[(\mu_1 - \mu_0)/\sigma_0]$ . Hence,  $p_{1,r} = \int_0^\infty \exp(-c_r x^{-\sigma_1/\sigma_0}) d[\exp(-1/x)]$ , with  $c_r = (r - 1)\exp[(\mu_1 -$

$\mu_0)/\sigma_0]$ . In other words, we have  $p_{1,r} = \int_0^1 \exp[-c_r(-\log u)^{\sigma_1/\sigma_0}] du$ . When  $\sigma_0 = \sigma_1$ , this becomes  $p_{1,r} = 1/(1 + c_r)$ .

If the shape parameters are nonnull and of the same sign, then we have

$$\begin{aligned} T_0[T_1^{-1}(x)] &= \left[ 1 + \frac{\xi_0}{\sigma_0}(\mu_1 - \mu_0) - \frac{\sigma_1}{\sigma_0} \frac{\xi_0}{\xi_1} + \frac{\sigma_1}{\sigma_0} \frac{\xi_0}{\xi_1} x^{\xi_1} \right]^{1/\xi_0}, \\ &= \left( \frac{\xi_0 \sigma_1}{\xi_1 \sigma_0} \right)^{1/\xi_0} x^{\xi_1/\xi_0}, \end{aligned}$$

whenever  $\mu_1 - \mu_0 = \sigma_1/\xi_1 - \sigma_0/\xi_0$  (same support condition in the Weibull and Fréchet cases). It follows that

$$p_{1,r} = \int_0^\infty \exp(-d_r x^{-\xi_1/\xi_0}) d[\exp(-1/x)],$$

with  $c_r = (r - 1)[(\xi_0 \sigma_1)/(\xi_1 \sigma_0)]^{-1/\xi_0}$ . In other words, we have  $p_{1,r} = \int_0^1 \exp[-c_r(-\log u)^{\xi_1/\xi_0}] du$ .

For the special case when  $\xi_0 = \xi_1$ , then this simplifies to  $p_{1,r} = 1/(1 + c_r)$ .

#### REFERENCES

- Ascher, S., 1990: A survey of tests for exponentiality. *Commun. Stat. Theory Methods*, **19**, 1811–1825, <https://doi.org/10.1080/03610929008830292>.
- Brockwell, P. J., and R. A. Davis, 1991: *Time Series: Theory and Methods*, 2nd ed. Springer Series in Statistics, Springer, 577 pp.
- Coles, S. G., 2001: *An Introduction to Statistical Modeling of Extreme Values*. Springer Series in Statistics, Springer, 208 pp.
- Cox, D. R., and D. Oakes, 1984: *Analysis of Survival Data*, Chapman & Hall/CRC Monogr. on Statistics and Applied Probability, No. 21, Chapman and Hall, 212 pp.
- Embrechts, P., C. Klüppelberg, and T. Mikosch, 1997: *Modelling Extremal Events for Insurance and Finance*. Stochastic Modelling and Applied Probability Series, Vol. 33, Springer, 648 pp.
- Hannart, A., J. Pearl, F. E. L. Otto, P. Naveau, and M. Ghil, 2016: Causal counterfactual theory for the attribution of weather and climate-related events. *Bull. Amer. Meteor. Soc.*, **97**, 99–110, <https://doi.org/10.1175/BAMS-D-14-00034.1>.
- Hegerl, G., and F. Zwiers, 2011: Use of models in detection and attribution of climate change. *Wiley Interdiscip. Rev.: Climate Change*, **2**, 570–591, <https://doi.org/10.1002/wcc.121>.
- Henze, N., and S. G. Meintanis, 2005: Recent and classical tests for exponentially: A partial review with comparisons. *Metrika*, **61**, 29–45, <https://doi.org/10.1007/s001840400322>.
- IPCC, 2013: *Climate Change 2013: The Physical Science Basis*. Cambridge University Press, 1535 pp., <https://doi.org/10.1017/CBO9781107415324>.
- Khariin, V. V., F. W. Zwiers, X. Zhang, and M. Wehner, 2013: Changes in temperature and precipitation extremes in the CMIP5 ensemble. *Climatic Change*, **119**, 345–357, <https://doi.org/10.1007/s10584-013-0705-8>.
- King, A. D., 2017: Attributing changing rates of temperature record breaking to anthropogenic influences. *Earth's Future*, **5**, 1156–1168, <https://doi.org/10.1002/2017EF000611>.
- Mason, D. M., and W. R. Van Zwet, 1987: A refinement of the KMT inequality for the uniform empirical process. *Ann. Probab.*, **15**, 871–884, <https://doi.org/10.1214/aop/1176992070>.
- Meehl, G. A., C. Tebaldi, G. Walton, D. Easterling, and L. McDaniel, 2009: Relative increase of record high maximum temperatures compared to record low minimum temperatures in the U.S. *Geophys. Res. Lett.*, **36**, L23701, <https://doi.org/10.1029/2009GL040736>.
- Resnick, S., 1987: *Extreme Values, Regular Variation, and Point Processes*. Springer Series in Operations Research and Financial Engineering, Springer, 320 pp.
- Rootzén, H., and R. W. Katz, 2013: Design Life Level: Quantifying risk in a changing climate. *Water Resour. Res.*, **49**, 5964–5972, <https://doi.org/10.1002/wrcr.20425>.
- Stott, P. A., D. A. Stone, and M. R. Allen, 2004: Human contribution to the European heatwave of 2003. *Nature*, **432**, 610–614, <https://doi.org/10.1038/nature03089>.
- , and Coauthors, 2016: Attribution of extreme weather and climate-related events. *Wiley Interdiscip. Rev.: Climate Change*, **7**, 23–41, <https://doi.org/10.1002/wcc.380>.
- Taylor, K., R. Stouffer, and G. Meehl, 2012: An overview of CMIP5 and the experiment design. *Bull. Amer. Meteor. Soc.*, **93**, 485–498, <https://doi.org/10.1175/BAMS-D-11-00094.1>.
- Uhe, P., F. E. L. Otto, K. Haustein, G. J. van Oldenborgh, A. D. King, D. C. H. Wallom, M. R. Allen, and H. Cullen, 2016: Comparison of methods: Attributing the 2014 record European temperatures to human influences. *Geophys. Res. Lett.*, **43**, 8685–8693, <https://doi.org/10.1002/2016GL069568>.
- Yiou, P., A. Jézéquel, P. Naveau, F. E. L. Otto, R. Vautard, and M. Vrac, 2017: A statistical framework for conditional extreme event attribution. *Adv. Stat. Climate Meteor. Oceanogr.*, **3**, 17–31, <https://doi.org/10.5194/ascmo-3-17-2017>.
- Zwiers, F., and Coauthors, 2013: Climate extremes: Challenges in estimating and understanding recent changes in the frequency and intensity of extreme climate and weather events. *Climate Science for Serving Society*, G. R. Asrar and J. W. Hurrell, Eds., Springer, 339–389.