



Extreme value modeling with errors-in-variables in detection and attribution of changes in climate extremes

Yuen Tsz Abby Lau^{1,3} · Tianying Wang² · Jun Yan³ · Xuebin Zhang⁴

Received: 16 December 2022 / Accepted: 19 August 2023 / Published online: 11 September 2023
© The Author(s) 2023

Abstract

The generalized extreme value (GEV) regression provides a framework for modeling extreme events across various fields by incorporating covariates into the location parameter of GEV distributions. When the covariates are subject to errors-in-variables (EIV) or measurement error, ignoring the EIVs leads to biased estimation and degraded inferences. This problem arises in detection and attribution analyses of changes in climate extremes because the covariates are estimated with uncertainty. It has not been studied even for the case of independent EIVs, let alone the case of dependent EIVs, due to the complex structure of GEV. Here we propose a general Monte Carlo corrected score method and extend it to address temporally correlated EIVs in GEV modeling with application to the detection and attribution analyses for climate extremes. Through extensive simulation studies, the proposed method provides an unbiased estimator and valid inference. In the application to the detection and attribution analyses of temperature extremes in central regions of China, with the proposed method, the combined anthropogenic and natural signal is detected in the change in the annual minimum of daily maximum and the annual minimum of daily minimum.

Keywords Corrected score · Estimating equation · Measurement error · Monte Carlo

1 Introduction

Extreme events arise in many fields and often have catastrophic impacts. Examples include extreme weather (Field et al. 2012), natural hazards (Hamdi et al. 2021), and market crashes (Nolde and Zhou 2021). Statistical methods for

prediction and inference in analyzing extreme events need to account for their unique distributional characteristics. The generalized extreme value (GEV) distribution is the natural choice as the limiting distribution of sample maximum (Fisher and Tippett 1928). Covariates can be incorporated into the location parameter to predict extreme events or understand the behavior of extreme observations (Coles et al. 2001). Such GEV regression models have been widely used in an enormous variety of applications (Smith 1989; Wang et al. 2004; Huerta and Sansó 2007; Hyun et al. 2019).

The validity of GEV regression analyses depends on a critical assumption that there is no errors-in-variables (EIV) or measurement error in the data. This assumption may be violated in many applications, such as the detection and attribution analyses in the context of climate extremes (see Sect. 3). One major tool for the detection and attribution analyses is fingerprinting, which was first introduced to study changes in mean climate conditions (e.g., mean temperature) (Hasselmann 1993; Hegerl et al. 1996; Allen and Stott 2003), and later extended to extreme climate conditions (Zwiers et al. 2011; Wang et al. 2017, 2021). The fingerprinting method fits a GEV distribution to the observed extremes and takes the fingerprints of external forcings of interest as covari-

✉ Tianying Wang
Tianying.Wang@colostate.edu

Yuen Tsz Abby Lau
yuentszabby.lau@penntestmed.upenn.edu

Jun Yan
jun.yan@uconn.edu

Xuebin Zhang
xuebin.zhang@ec.gc.ca

¹ Department of Biostatistics and Epidemiology, University of Pennsylvania, Philadelphia, PA, USA

² Department of Statistics, Colorado State University, Fort Collins, USA

³ Department of Statistics, University of Connecticut, Storrs, CT, USA

⁴ Climate Research Division, Environment and Climate Change, Toronto, ON, Canada

ates in the location parameters. The fingerprint, or signal, of a particular external forcing, is not observed but estimated from climate-model-simulated extremes, which introduces errors (i.e., EIV) and causes bias in the resulting estimator of the coefficients of the fingerprints (Wang et al. 2021).

Despite the voluminous literature on EIV issues in other regression settings, it has been rarely studied in GEV regressions even for independent data, let alone for time series or spatiotemporal data. The main challenge owes to the complex structure of GEV distribution, which prevents a straightforward extension of many methods. In the context of the fingerprinting, for example, though the EIV issue has been solved in the mean climate models under a linear regression setting by total least squares (Allen and Stott 2003), integrated likelihood (Hannart 2016), and Bayesian model averaging (Katzfuss et al. 2017), none of these approaches can be applied to the extremes setting in a relatively straightforward manner. In fact, the bias issue in GEV modeling has been explicitly raised but left unsolved (Wang et al. 2021). Additionally, the correlation in the measurement errors further complicates the problem. Among a wide range of applications of GEV modeling of spatiotemporal data, the EIVs are likely dependent. One example is again the fingerprinting with time series, where the signals are most likely to be temporally dependent as the parameters in the signals are often shared over time.

A promising approach to GEV regression with EIV is the corrected-score method (Nakamura 1990). In principle, this method constructs a corrected score equation whose expectation conditional on the contaminated covariates is the same as the original score equations evaluated at the true covariates. It is attractive because, as a functional approach, it does not require any distributional specification of the covariates as some approaches do. Given the complex expressions of the score functions under the GEV regression, which can be derived from the score functions of GEV distribution (Prescott and Walden 1980) using the chain rule, such construction is challenging because the expectations of the score functions in terms of the contaminated covariates have no closed forms, which prohibits closed-form corrections. The Monte Carlo corrected scores (MCCS) method, however, is well suited for the situation. It approximates the expectations of the score functions in terms of the contaminated covariates by Monte Carlo integration, which facilitates corrected scores (Novick and Stefanski 2002).

Our contributions are two-fold. First, as the first attempt to address the EIV issue in GEV modeling using a functional approach, we propose an MCCS-based approach and extend it to the settings where the EIVs are temporally dependent as in fingerprinting applications. We further develop a small-sample correction to the sandwich variance estimator. This is practically important because the record lengths of extremes are usually insufficient for the variance estimator to

work well. Simulation studies show that the proposed method corrects the bias in estimating all the model parameters, especially the regression coefficients in the location parameter. The coverage rates of the confidence intervals based on our corrected sandwich variance estimator are acceptable. Our second contribution is an application of the methods to the detection and attribution analyses of changes in temperature extremes. The method accounts for the fact the unobserved signals are estimated from climate model simulations under the corresponding forcings. In applications we demonstrate the validity of the method via simulation studies that mimic the fingerprinting setting and real data.

The rest of the paper is organized as follows. Section 2 provides the MCCS framework under the GEV setting for both independent and temporally dependent EIVs. Section 3 adapts the method to the specific setting of fingerprinting for climate extremes in detection and attribution analyses. The methods are validated in simulation studies with both independent extremes and fingerprinting settings in Sect. 4. Section 5 reports an application to detection and attribution analyses of changes in temperature extremes in central China. A discussion concludes in Sect. 6.

2 Methodology

2.1 Statistical model

Consider GEV modeling with a regression setting for the location parameter. Let $F(\cdot \mid \mu, \sigma, \xi)$ denote a GEV distribution function with location μ , scale σ , and shape ξ , and Y_t be observed extreme with p -dimensional error-prone covariates X_t and q -dimensional accurate covariates Z_t for $t = 1, \dots, n$. Suppose that Y_t follows a GEV distribution with a location parameter that incorporates covariate effects. In particular,

$$Y_t \sim F(\cdot \mid \mu_t, \sigma, \xi), \quad \mu_t = Z_t^\top \alpha + X_t^\top \beta, \quad (1)$$

where α and β are q - and p -dimensional regression coefficients, respectively. An intercept term is included in α by including a one in Z_t . The parameters of interest are $\theta = (\alpha, \beta, \sigma, \xi)$.

Assumption 1 When X_t is observed, there exists an unbiased estimating function $\psi(Y_t, X_t, Z_t, \theta)$ with parameters θ such that

$$E\{\psi(Y_t, X_t, Z_t, \theta) \mid X_t, Z_t\} = 0, \quad t = 1, \dots, n.$$

Then, the parameter θ can be estimated consistently by solving

$$\sum_{t=1}^n \psi(Y_t, X_t, Z_t, \theta) = 0.$$

By fulfilling Assumption 1, the score function $\psi(Y_t, X_t, Z_t, \theta)$ of GEV loglikelihood is an unbiased estimating function, where

$$\psi(Y_t, X_t, Z_t, \theta) = \begin{pmatrix} Z_t f_1(Y_t; \mu_t, \sigma, \xi) \\ X_t f_1(Y_t; \mu_t, \sigma, \xi) \\ f_2(Y_t; \mu_t, \sigma, \xi) \\ f_3(Y_t; \mu_t, \sigma, \xi) \end{pmatrix}, \tag{2}$$

and f_1, f_2 , and f_3 are the partial derivatives of the log density of $F(\cdot \mid \mu, \sigma, \xi)$ with respect to μ, σ , and ξ , respectively, with specific expressions given in Appendix A (Prescott and Walden 1980). The maximum likelihood estimator (MLE) for θ can be obtained by solving the score equation $\sum_{t=1}^n \psi(Y_t, X_t, Z_t, \theta) = 0$. For $\xi > -0.5$, $\psi(Y_t, X_t, Z_t, \theta)$ has expectation zero and finite second moments (Smith 1985; Bücher and Segers 2017).

In situations where instead of X_t , a surrogate version W_t is observed, substituting X_t 's directly with W_t in estimating function (2) leads to a biased estimator of θ . The magnitude of the bias depends on the measurement error specification. Assume that

$$W_t = X_t + e_t, \quad t = 1, \dots, n, \tag{3}$$

where the p -dimensional measurement error e_t , independent of X_t and Y_t , is normally distributed with mean zero and covariance matrix Σ_t .

Assumption 2 When X_t is not observed, there exists a corrected score function of the observed data such that for $t = 1, \dots, n$,

$$E\{\tilde{\psi}(Y_t, W_t, Z_t, \theta) \mid (Y_t, X_t, Z_t)\} = \psi(Y_t, X_t, Z_t, \theta). \tag{4}$$

Based on Assumptions 1–2, we seek a corrected score function $\tilde{\psi}(Y_t, W_t, Z_t, \theta)$, which is conditionally unbiased given the observed quantities and results in estimating equations

$$\sum_{t=1}^n \tilde{\psi}(Y_t, W_t, Z_t, \theta) = 0.$$

While the corrected scores for some common families like normal and Poisson are available in a closed-form (Nakamura 1990), it is hard to obtain that of the GEV distribution because of its nonlinear and complicated structure. Therefore, we resort to the MCCS method developed by Novick and Stefanski (2002), where the conditional expectation in (4) is approximated by Monte Carlo integration. This method is

relatively robust and can yield satisfying results in certain settings even when the normality assumption of e_t is violated.

A further complication in practice, however, is that the measurement errors e_t 's in GEV could be correlated in various ways. For instance, in fingerprinting applications, the fingerprint or signal of an external forcing cannot be observed but is generally estimated from climate model simulations. The measurement errors e_t 's could be temporally correlated over time and we denote the covariance matrix for $e^\top = (e_1^\top, \dots, e_n^\top)^\top$ as Σ . This brings an additional complexity compared to the classical MCCS approach in Novick and Stefanski (2002). In the sequel, we address this issue after dealing with independent EIVs.

2.2 MCCS for GEV regression with independent EIV

First, we illustrate the MCCS method for GEV regression under the assumption that measurement error e_t 's are *mutually independent* for $t = 1, \dots, n$. Let $\epsilon_{b,t}$ be the b th Monte Carlo copy generated from $N(0, \Sigma_t)$, where Σ_t is the variance matrix for $e_t, t = 1, \dots, n$. The complex variates are constructed as $\tilde{W}_{b,t} = W_t + \iota \epsilon_{b,t}$ for $b = 1, \dots, B$, where B is a large integer and $\iota = \sqrt{-1}$. Then, the following Lemma was established by Novick and Stefanski (2002).

Lemma 1 (Novick and Stefanski 2002) *When the measurement error e_t 's are mutually independent, under integrability conditions and the assumption that $\psi(Y_t, X_t, Z_t, \theta)$ is an entire function of X_t in the complex plane,*

$$\begin{aligned} E\{\psi(Y_t, \tilde{W}_{b,t}, Z_t, \theta) \mid (Y_t, X_t, Z_t)\} \\ &= E[\Re\{\psi(Y_t, \tilde{W}_{b,t}, Z_t, \theta)\} \\ &\quad \mid (Y_t, X_t, Z_t)] \\ &= \psi(Y_t, X_t, Z_t, \theta), \end{aligned}$$

where $\Re(\cdot)$ is the real part of a complex number.

Consequently, $\Re\{\psi(Y_t, \tilde{W}_{b,t}, Z_t, \theta)\}$ is a corrected score based on a randomly generated random vector $\epsilon_{b,t}$. Then, we use Monte Carlo integration to approximate the conditional expectation as

$$\tilde{\psi}_B(Y_t, W_t, Z_t, \theta, \Sigma_t) = \frac{1}{B} \sum_{b=1}^B \Re[\psi(Y_t, \tilde{W}_{b,t}, Z_t, \theta)], \tag{5}$$

with its specific form in GEV provided in Appendix A. As $B \rightarrow \infty$, the MCCS converges to the exact conditional expectation (Novick and Stefanski 2002). The MCCS estimator $\hat{\theta}_{n,B}$ solves the estimating equations

$$\sum_{t=1}^n \tilde{\psi}_B(Y_t, W_t, Z_t, \theta, \Sigma_t) = 0.$$

In practice, Σ_t is unknown and needs to be replaced with a consistent estimator in the above derivation. We provide more details on estimating Σ_t in Sect. 3. Let θ_0 be the true value of θ . When (Y_t, W_t) are mutually independent for $t = 1, \dots, n$, and $\tilde{\psi}_B(Y_t, W_t, Z_t, \theta, \Sigma_t)$ and $\psi(Y_t, W_t, Z_t, \theta, \Sigma_t)$ are differentiable, there exists a consistent estimator for θ by solving $\sum_{t=1}^n \tilde{\psi}_B(Y_t, W_t, Z_t, \theta, \Sigma_t) = 0$ (Nakamura 1990). The solution can be obtained via `nleqslv` (Hasselmann 2018), a nonlinear equation solver package in R.

Proposition 1 *Under certain regularity conditions, with $n \rightarrow \infty$,*

$$\sqrt{n}(\tilde{\theta}_{n,B} - \theta_0) \rightarrow N(0, V_{\tilde{\psi}}),$$

where $V_{\tilde{\psi}} = D^{-1}C(D^{-1})^\top$, $C = E[\tilde{\psi}_B(Y_t, W_t, Z_t, \theta, \Sigma) \tilde{\psi}_B^\top(Y_t, W_t, Z_t, \theta, \Sigma_t)]$ and $D = E[(\partial/\partial\theta^\top)\tilde{\psi}_B(Y_t, W_t, Z_t, \theta, \Sigma_t)]$. The asymptotic variance $V_{\tilde{\psi}}$ is consistently estimated by a sandwich formula estimator $\tilde{V}_{n,\tilde{\psi}} = \tilde{D}_{n,\tilde{\psi}}^{-1}\tilde{C}_{n,\tilde{\psi}}(\tilde{D}_{n,\tilde{\psi}}^{-1})^\top$, where

$$\tilde{D}_{n,\tilde{\psi}} = \frac{1}{n} \sum_{t=1}^n \frac{\tilde{\psi}_B(Y_t, W_t, Z_t, \theta, \Sigma_t)}{\partial\theta^\top} \Big|_{\theta=\tilde{\theta}_{n,B}}$$

and

$$\tilde{C}_{n,\tilde{\psi}} = \frac{1}{n} \sum_{t=1}^n \tilde{\psi}_B(Y_t, W_t, Z_t, \tilde{\theta}_{n,B}, \Sigma_t) \tilde{\psi}_B^\top(Y_t, W_t, Z_t, \tilde{\theta}_{n,B}, \Sigma_t).$$

The proof based on standard estimating equation theory is outlined in Appendix B.1.

2.3 MCCS for GEV regression with temporally dependent EIV

In many GEV applications, such as the fingerprinting setup, measurement error e_t 's could be *temporally correlated* (Wang et al. 2021). Here we explain how to adapt the aforementioned MCCS approach to time-dependent EIV. To construct Monte Carlo replicates $\tilde{W}_{b,t}^* = W_t + \iota\epsilon_{b,t}^*$, we propose to generate $\epsilon_{b,t}^*$ as part of $\epsilon_b^{*\top} = (\epsilon_{b,1}^*, \dots, \epsilon_{b,n}^*)^\top$ from $N(0, \Sigma)$. Ideally, one would expect to construct corrected scores based on $Y^\top = (Y_1^\top, \dots, Y_n^\top)^\top$, $W^\top = (W_1^\top, \dots, W_n^\top)^\top$ and $Z^\top = (Z_1^\top, \dots, Z_n^\top)^\top$ to incorporate the temporal correlations. Such practice can be difficult or undesired for GEV modeling as the temporal dependence may be hard to specify. As θ is assumed to be shared across n years, we again consider the corrected score based on each year t as in Eq. (5) and obtain the unbiased estimator $\tilde{\theta}_{n,B}$. Though the estimating equation is constructed at each time point, the validity of the approach is stated in Proposition 1.

Lemma 2 *Assume that $\psi(Y_t, X_t, Z_t, \theta)$ is an entire function of X_t in the complex plane, and that measurement error $e^\top = (e_1^\top, \dots, e_n^\top)^\top$ are temporally correlated with covariance matrix Σ . With $\epsilon_b^{*\top} \sim N(0, \Sigma)$, the following equations still hold:*

$$\begin{aligned} E\{\psi(Y_t, \tilde{W}_{b,t}^*, Z_t, \theta) \mid (Y_t, X_t, Z_t)\} \\ = E[\Re\{\psi(Y_t, \tilde{W}_{b,t}^*, Z_t, \theta)\} \\ \mid (Y_t, X_t, Z_t)] = \psi(Y_t, X_t, Z_t, \theta). \end{aligned}$$

This Lemma is based on representing ψ as a multivariate power series and then applying Lemma 4 from Stefanski and Cook (1995). More details can be found in the Appendix B.3.

Then, we define

$$\tilde{\psi}_B^*(Y_t, W_t, Z_t, \theta, \Sigma) = \frac{1}{B} \sum_{b=1}^B \Re[\psi(Y_t, \tilde{W}_{b,t}^*, Z_t, \theta)], \quad (6)$$

and the MCCS estimator $\tilde{\theta}_{n,B}^*$ solves the estimating equations

$$\sum_{t=1}^n \tilde{\psi}_B^*(Y_t, W_t, Z_t, \theta, \Sigma) = 0. \quad (7)$$

Again, the R package `nleqslv` can be used to solve Equation (7).

Proposition 2 *Under certain regularity conditions, with $n \rightarrow \infty$,*

$$\sqrt{n}(\tilde{\theta}_{n,B}^* - \theta_0) \rightarrow N(0, V_{\tilde{\psi}^*}),$$

where $V_{\tilde{\psi}^*} = D^{*-1}C^*(D^{*-1})^\top$, $C^* = \text{Var}\{n^{-1/2} \sum_{t=1}^n \tilde{\psi}_B^*(Y_t, W_t, Z_t, \theta, \Sigma)\}$ and $D^* = E[(\partial/\partial\theta^\top)\tilde{\psi}_B^*(Y_t, W_t, Z_t, \theta, \Sigma)]$. The asymptotic variance $V_{\tilde{\psi}^*}$ can be estimated consistently as $\hat{V}_{n,\tilde{\psi}^*} = \hat{D}_{n,\tilde{\psi}^*}^{-1}\hat{C}_{n,\tilde{\psi}^*}(\hat{D}_{n,\tilde{\psi}^*}^{-1})^\top$, where

$$\hat{D}_{n,\tilde{\psi}^*} = n^{-1} \sum_{t=1}^n \frac{\tilde{\psi}_B^*(Y_t, W_t, Z_t, \theta, \Sigma_t)}{\partial\theta^\top} \Big|_{\theta=\tilde{\theta}_{n,B}^*}$$

and

$$\hat{C}_{n,\tilde{\psi}^*} = \frac{1}{nB} \left[\sum_{t=1}^n \sum_{b=1}^B \Re\{\psi(Y_t, \tilde{W}_{b,t}^*, \theta_{n,B}^*)\psi^\top(Y_t, \tilde{W}_{b,t}^*, \theta_{n,B}^*)\} \right],$$

for $t = 1, \dots, n$.

The detailed proof is provided in Appendix B.2.

Remark 1 Our work is built upon standard estimating equation theory. There are some sufficient but not necessary conditions for our asymptotic theory to hold as below.

1. The parameter space is compact, and there is a unique θ in the parameter space satisfying $E\{\tilde{\psi}_B^*(Y_t, W_t, Z_t, \theta, \Sigma)\} = 0$ for all $t = 1, \dots, n$.
2. The estimating equation (7) has a unique solution.
3. The matrix D^* is of full rank within a neighborhood of the true value θ_0 . For sufficiently large n , $\widehat{D}_{n, \tilde{\psi}^*}$ is of full rank with eigenvalues bounded away from 0 and $\pm\infty$ within a neighborhood of the true value θ_0 .

Sandwich variance estimators are known to under-estimate the variance for small samples (Mancl and DeRouen 2001; Westgate 2012). A major reason is that the middle matrix C^* is under-estimated. To mitigate the potential underestimation of $\widehat{C}_{n, \tilde{\psi}^*}$ when the sample size is small, and deal with the time dependency in extreme observations if that exists in the data, we propose a block bootstrap to account for the time dependency in $\tilde{\psi}_B^*$, rendering an estimate for the variance matrix of $\tilde{\psi}_B^*$. To create M bootstrap samples, we use the following bootstrap procedure for each $m = 1, \dots, M$ and $t = 1, \dots, n$:

1. Calculate residuals by subtracting $Z_t^\top \alpha + W_t^\top \beta$ from Y_t .
2. Group residuals into non-overlapping blocks of samples in specified length with replacement.
3. Obtain bootstrapped data $\mathcal{Y}_{m,t}$ by adding the grouped residuals to the $Z_t^\top \alpha + W_t^\top \beta$.
4. Calculate the $\tilde{\psi}_{m,B} = n^{-1} \sum_{t=1}^n B^{-1} \sum_{b=1}^B \mathfrak{R}[\psi(\mathcal{Y}_{m,t}, \tilde{W}_{m,b,t}^*, Z_t, \tilde{\theta}_{n,B}^*)]$ for $m = 1, \dots, M$, where $\tilde{W}_{m,b,t}^*$ is generated for the m th bootstrap sample.
5. Determine the covariance matrix of $\tilde{\psi}_{m,B}$ using M samples and multiplying the covariance matrix with n to get $\widehat{C}_{n, \tilde{\psi}^*}$.

This bootstrap procedure is computationally efficient as it only requires evaluating the outer-products of individual estimating functions instead of solving any estimating equations. Of note, the bootstrap is only used for estimating the middle part of the sandwich estimator. As it does not bootstrap extremes directly, this procedure also avoids some potential issues in sampling extremes (Gilleland 2020), and the confidence intervals of the estimates are still built based on the asymptotic normality.

3 Application in fingerprinting with time series

Fingerprinting in the context of extremes is a direct application of the proposed method. Fingerprinting aims to quantify the impact of external forcings on the climate variable of interest, in which the expected responses of the climate system to external forcings, also called fingerprints, are treated

as predictors (Li et al. 2021, 2023). These external forcings can be classified into two types: anthropogenic and natural, which are factors that impact the climate system while not being part of the climate system itself, such as greenhouse gas emissions and volcanic eruptions. The fingerprints are not observable but can be inferred from numerical climate model outputs. Climate models typically contain multiple initial condition ensembles, which, most of the time, can be treated as random independent and identically distributed samples from one multivariate spatial-temporal process due to their sensitivity to the initial conditions (Stein 2020). The estimated signals from these outputs are naturally temporally dependent. Here we limit our scope to a setting with time series data instead of spatiotemporal data; handling spatiotemporal EIVs requires a separate investigation. We investigate this application setting as follows.

Suppose the observed climate extremes Y_t specified in Sect. 2 are available at a particular grid box for year $t = 1, \dots, n$. The climate extremes of interest can be station annual maximum of daily maximum (TXx) and minimum (TNx) temperatures or the annual minimum of daily maximum (TXn) and minimum (TNn) temperatures. The error-prone covariates X_t in Equation (1) are signals of external forcings, and β are the corresponding coefficients of the signals. We also assume a constant shape and scale parameters throughout the year at the site, and an intercept is included in the location parameter, which makes Z_t a constant of 1 and α a scalar (Wang et al. 2021; Zwiers et al. 2011).

Since the signal vector X_t is unknown, it needs to be estimated from the climate models simulations. Following Zwiers et al. (2011) and Wang et al. (2021), we assume another GEV model for the simulated climate extremes. Let L_j denote the total number of available ensemble runs under the j th forcing from the climate models, and $U_{ij}^{(l)}$ extreme observations for the l th ensemble run under the j th forcing in year t , where $j = 1, \dots, p, t = 1, \dots, n$, and $l = 1, \dots, L_j$. The annual extremes $U_{ij}^{(l)}$ can be modeled as follows,

$$U_{ij}^{(l)} \sim F(\cdot | X_{tj}, \sigma'_j, \xi'_j), \quad X_{tj} = \mu(t; \gamma_j), \quad (8)$$

where σ'_j and ξ'_j are forcing-specific scale and shape parameters, respectively, and X_{tj} is the fingerprint or the signal of forcing j characterized by function μ with parameter vector γ_j . To account for the temporal dynamics, a flexible formulation for μ with B-splines is

$$\mu(t; \gamma_j) = S_t^\top \gamma_j$$

where $S_t = (S_1(t), \dots, S_D(t))^\top$, $S_d(t)$'s are a set of B-splines basis with D degree of freedom, and $\gamma_j = (\gamma_{j,1}, \dots, \gamma_{j,D})^\top$ is the coefficient vector for the basis. The B-spline representation is assumed to be flexible enough to

approximate the temporal pattern in signals. With all the ensemble runs under forcing j as independent replicates, we can obtain the MLE $\hat{\gamma}_j$ of γ_j and its covariance matrix Σ_{γ_j} . Therefore, $W_{tj} = \sum_{d=1}^D \hat{\gamma}_{j,d} S_d(t)$ is an MLE of X_{tj} with covariance matrix $S_t^\top \Sigma_{\gamma_j} S_t$.

Given the B-spline model is correctly specified, using $W_t = (W_{t1}, \dots, W_{tp})^\top$ rather than the true X_t introduces additional variability to Equation (1), resulting bias when estimating β . It has the same effect as EIV and needs to be recognized. This is exactly the unsolved issue raised by Wang et al. (2021), which results in biased estimation. Here the relationship between X_t and W_t is the same as in Equation (3), that is, $W_t = X_t + e_t$. In general, e_t and $e_{t'}$ for any $t \neq t'$ are not independent as the variance of e_{tj} and $e_{t'j}$ both depend on $\hat{\gamma}_j$. The covariance of $e_j = (e_{1j}^\top, \dots, e_{nj}^\top)^\top$, can be estimated by $S^\top \Sigma_{\gamma_j} S$ with $S = (S_1^\top, \dots, S_n^\top)^\top$. This fits the situation in Sect. 2.3. For each $j = 1, \dots, p$, we can generate $\tilde{W}_{b,j}^\top = (\tilde{W}_{b,1j}^\top, \dots, \tilde{W}_{b,nj}^\top)^\top$ from a multivariate normal distribution with mean $W^\top = (W_{1j}^\top, \dots, W_{nj}^\top)^\top$ and covariance matrix $S^\top \Sigma_{\gamma_j} S$. Since the runs under different forcings are independent, we can generate the Monte Carlo copies separately for each j . Then, $\tilde{W}_{b,tj}$ for any $t = 1, \dots, n$ and $j = 1, \dots, p$ can be used to construct the corrected score in Equation (6), yielding the unbiased estimates of θ as explained in Sect. 2.

The coefficients of the fingerprints β , also known as scaling factors, are the primary target of the inference. To account for the time dependency in the measurement error, we rely on the block bootstrap method with 5-year blocks detailed in Sect. 2.3 to construct the 90% confidence interval for β , which is frequently used in attribution and detection analyses. If the 90% corresponding confidence interval of a scaling factor is above zero, then the fingerprint is claimed to be detected. Additionally, given that a fingerprint is detected, a 90% confidence interval covering one is the necessary evidence that the observed changes can be attributed to the corresponding external forcing.

4 Simulation studies

4.1 Independence setting

We first considered the case with independent measurement errors. For simplicity, X_t is a univariate normal random variable with mean 0 and standard deviation 2, and $Z_t = 1$. The observed extremes Y_t was subsequently generated from a generalized extreme value distribution with $\theta = (1, 1, 4, -0.2)$ by using the R package `evd` (Stephenson 2002). The observed surrogate W_t was then generated from Equation (3) where Σ reduces to $\sigma_e^2 I$ with $\sigma_e^2 = 1$. To carry out the MCCS method, $\tilde{W}_{b,t}$ was obtained by adding $\epsilon_{b,t}$ to

W_t , where $\epsilon_{b,t}$ is also a normal variable with mean 0 and variance $\sigma_e^2 = 1$, and $b = 1, \dots, B$. Here we treat σ_e as known and focus on estimating θ . In the simulation, we consider the Monte Carlo sampling $B \in \{200, 400\}$ and the sample size $n \in \{100, 200\}$. For each configuration, 1000 data sets have been generated. Also, for each dataset, we obtained MLE for θ without correcting for W_t and $\tilde{\theta}$ by implementing MCCS with $\tilde{W}_{b,t}$. The MLE estimation was carried out by the package `ismev` (Heffernan and Stephenson 2018), and the MCCS was solved by using `nleqslv`.

Table 1 summarizes the results in estimating θ . The MLEs of θ are the naive estimates with measurement error ignored. The naive estimates of α , σ , and ξ are unbiased under all configurations. The naive estimate of parameter β , however, is severely biased, and the coverage rate is low across different sample sizes. On the other hand, the proposed method provides an unbiased estimate of β , and the coverage rate is close to the nominal level 90% for all the settings, including the case that $B = 200$. The other parameters are consistently estimated, and their coverage rates are close to the nominal level as well. In general, as n increases, the convergence rate improves, and the bias, RMSE, ESE, and ASE decreases as n increases. Regarding the performance of the algorithm, the percentage of successfully finding a root using `nleqslv` is around 95% when $n = 100$ and improves to 99.5% when n increases to 200 in our simulation. Given a fixed n , however, the performance of MCCS estimators is very similar across different B s in our setting. Nonetheless, the proposed method works well in correcting bias in β .

4.2 Fingerprinting setting

We also carried out simulation studies with fingerprinting setting described in Sect. 3 with one forcing ($p = 1$). In order to mimic the real detection and attribution process, we first simulated the climate model extremes $U_t^{(l)}$ from a generalized extreme model with $\sigma' = 2$ and $\xi' = -0.2$. The location parameter X_t is determined by a cubic spline where γ is uniformly distributed in $(38, 40)$. The number of ensemble runs L is assumed to be 50, and for each run, we have 100 years of data. Then the observed extreme Y_t was generated from a generalized extreme value distribution with $\{\alpha, \beta, \sigma, \xi\} = \{40, 1, 1, -0.2\}$. Again, two levels of the Monte Carlo sampling size, $\{200, 400\}$ are considered here.

Table 2 summarizes the results in estimating θ in the fingerprinting setup. Both methods provide unbiased estimates to σ and ξ . Unlike the previous results, the bias is observed in both α and β in the naive estimates, resulting in extremely poor coverage rates. This could lead to misleading conclusions in the detection and attribution analysis which mainly relies on the statistical inferences of β . In contrast, MCCS estimates of α and β are nearly unbiased. Even though the bias of $\hat{\alpha}$ estimated by MCCS is around -1.2 in Table 2, it

Table 1 Summary of simulation results based on 1000 replications

<i>n</i>	Parameter	Method	Bias	RMSE	ESE	ASE	CR
100	α	Naive method	0.036	0.450	0.449	0.451	0.895
		MCCS($B = 200$)	0.056	0.464	0.461	0.538	0.906
		MCCS($B = 400$)	0.055	0.461	0.458	0.492	0.907
	β	Naive method	-0.191	0.280	0.204	0.188	0.710
		MCCS($B = 200$)	0.024	0.266	0.265	0.264	0.874
		MCCS($B = 400$)	0.027	0.271	0.270	0.256	0.875
	σ	Naive method	0.038	0.318	0.315	0.319	0.907
		MCCS($B = 200$)	-0.092	0.357	0.345	0.399	0.867
		MCCS($B = 400$)	-0.091	0.359	0.348	0.350	0.867
	ξ	Naive method	-0.017	0.074	0.072	0.068	0.869
		MCCS($B = 200$)	-0.006	0.084	0.084	0.133	0.880
		MCCS($B = 400$)	-0.006	0.086	0.086	0.087	0.878
200	α	Naive method	0.002	0.321	0.321	0.320	0.904
		MCCS($B = 200$)	0.035	0.332	0.330	0.333	0.902
		MCCS($B = 400$)	0.035	0.332	0.330	0.333	0.903
	β	Naive method	-0.197	0.240	0.137	0.133	0.576
		MCCS($B = 200$)	0.014	0.176	0.176	0.171	0.892
		MCCS($B = 400$)	0.015	0.176	0.176	0.171	0.894
	σ	Naive method	0.085	0.240	0.224	0.224	0.883
		MCCS($B = 200$)	-0.025	0.241	0.240	0.238	0.892
		MCCS($B = 400$)	-0.025	0.242	0.240	0.238	0.893
	ξ	Naive method	-0.012	0.049	0.047	0.045	0.876
		MCCS($B = 200$)	-0.010	0.056	0.055	0.049	0.862
		MCCS($B = 400$)	-0.010	0.056	0.055	0.050	0.861

Bias is the bias of point estimates; RMSE is the root mean square error; ESE is the empirical standard error; ASE is the average of the standard errors of the estimator; and CR is the coverage rate of 90% confidence interval

is relatively unbiased given that the true value of α is 40. More importantly, the coverage rate of the scaling factor β in MCCS is much closer to 90%, which implies a more reliable decision on the detection and attribution of a signal. Even though the coverage rates of σ and ξ are lower than the nominal level using the MCCS method, they are still around 85%, and their MCCS point estimates are also unbiased. In simulations, the convergence rate of the MCCS method is around 99%, indicating high success in finding a root.

5 Temperature extremes in central China

The observed annual temperature extremes in the East Asia region over the period of 1951–2010 were extracted from the HadEX2 data (Donat et al. 2013). For each temperature extreme described in Sect. 3, grid boxes of resolution $5^\circ \times 5^\circ$ (latitude \times longitude) are available. After excluding the grid boxes with missing values, we have 47, 48, 41, and 43 grid boxes for TXx, TXn, TNx, and TNn, respectively. The longitude value covered by these grid boxes ranges from

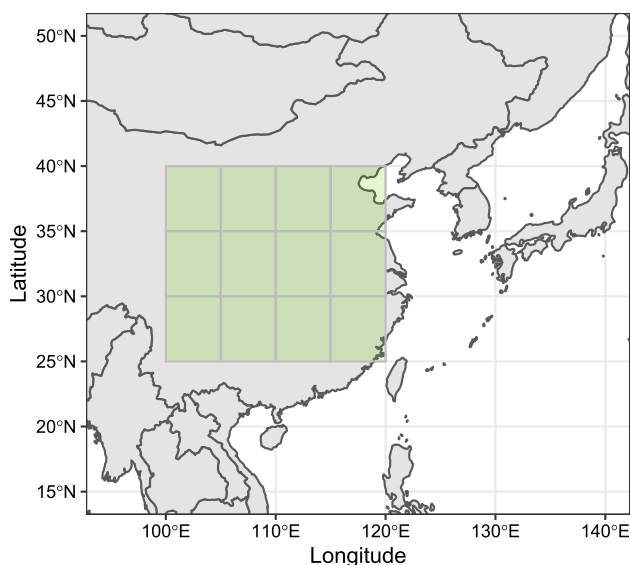
100E to 150E, and latitude value ranges from 25N to 50N. Since observations on land are more reliable than those in the ocean, we focused on twelve land grid boxes with latitude ranging from 30N to 40N and longitude covering from 100E to 115E, which covers a large part of inland China (Fig. 1).

The ALL (combination of anthropogenic and natural external) forcing is the single external forcing being considered. To avoid making any assumptions about the exchangeability of different climate models, we only used the second-generation Canadian Earth System Model (CanESM2) (Chylek et al. 2011) to estimate signals, which contains 50 runs under the ALL forcing. In order to match the spatial resolution of the HadEx2 data, the CanESM2 data also has been regridded to $5^\circ \times 5^\circ$ grid boxes. The 50 runs are independent replicates simulated under both anthropogenic factors (greenhouse gas and aerosol emissions) and natural-only factors (solar radiation and volcano eruptions). The signals under the ALL forcing at each grid box were estimated by fitting (8) to the 50 runs. A cubic spline with a 5-year knots was used to estimate the ALL signal. In regional detection and attribution analyses, it is often assumed that the scaling

Table 2 Summary of simulation results based on 1000 replications

Parameter	Method	Bias	RMSE	ESE	ASE	CR
α	Naive method	27.739	33.077	18.026	11.637	0.251
	MCCS($B = 200$)	-1.294	13.476	13.421	12.790	0.866
	MCCS($B = 400$)	-1.247	13.420	13.369	12.759	0.873
β	Naive method	-0.712	0.849	0.463	0.299	0.251
	MCCS($B = 200$)	0.033	0.346	0.344	0.328	0.867
	MCCS($B = 400$)	0.032	0.344	0.343	0.327	0.872
σ	Naive method	0.042	0.100	0.091	0.089	0.876
	MCCS($B = 200$)	-0.019	0.085	0.083	0.078	0.837
	MCCS($B = 400$)	-0.019	0.085	0.083	0.077	0.836
ξ	Naive method	-0.021	0.071	0.068	0.068	0.886
	MCCS($B = 200$)	-0.007	0.075	0.075	0.067	0.851
	MCCS($B = 400$)	-0.006	0.074	0.074	0.067	0.851

Bias is the bias of point estimates; RMSE is the root mean square error; ESE is the empirical standard error; ASE is the average of the standard errors of the estimator; and CR is the coverage rate of 90% confidence interval

**Fig. 1** Twelve grid boxes in central China were selected for analysis

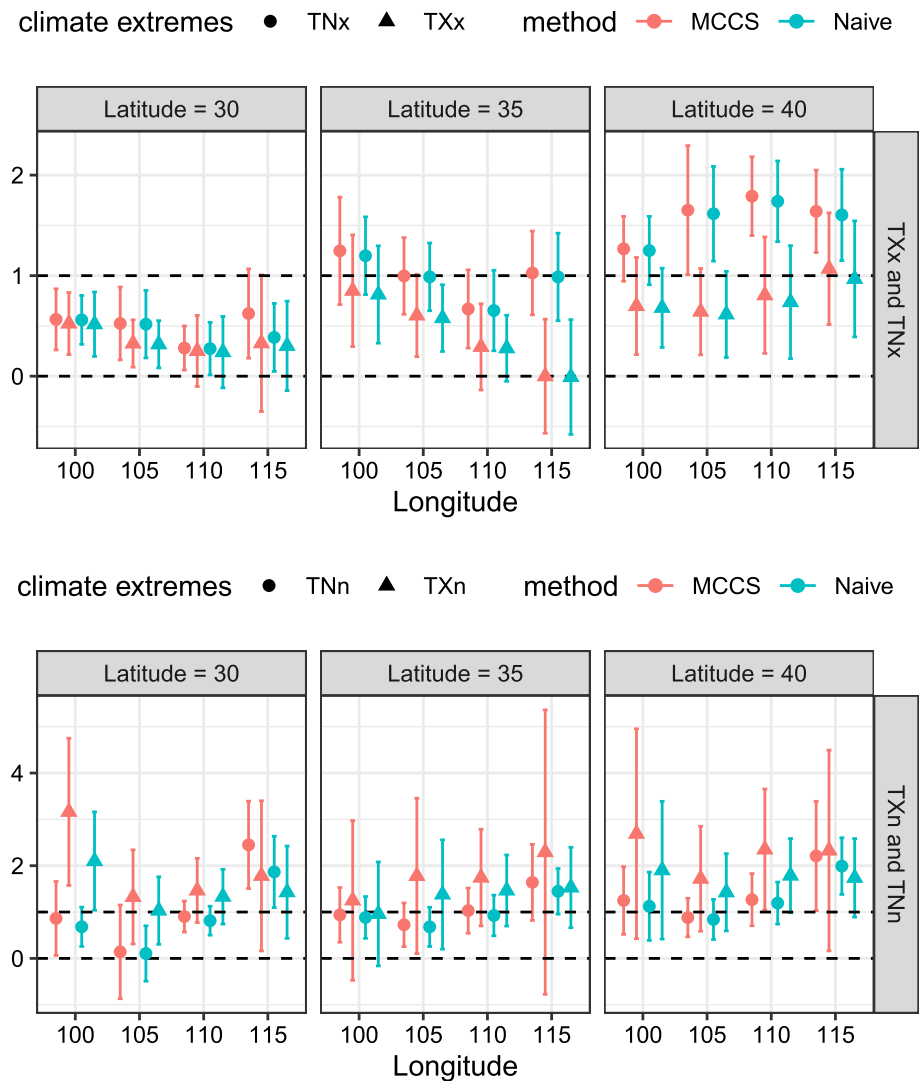
factors across all the grid boxes are the same (Zwiers et al. 2011; Wang et al. 2017, 2021), in which case, the spatial dependence among the grid boxes needs to be addressed. Since a spatial extension of the MCCS method needs a full investigation, here we apply the method to each grid box separately.

At each grid box, the fingerprinting method in Sect. 3 was applied to each of the four temperature extremes. The minimum extremes TXn and TNn were negated before the analyses. The size of Monte Carlo sampling was fixed at $B = 1,000$ for each analysis. For comparison, we also included the naive estimators, which ignored the measurement errors in the estimated fingerprint and fitted the data

using the MLE method. To ensure $\xi > -0.5$, we use a Beta(9, 6) prior to constraint ξ within a range of $(-0.5, 0.5)$ (Martins and Stedinger 2000) in both methods as done by Wang et al. (2021).

Figure 2 displays the 90% confidence intervals of the scaling factor for each temperature extreme in each of the 12 grid boxes. Additional figures of other parameters are presented in the Appendix 1. The estimates of ξ are negative in most grid boxes, which is consistent with the findings in Huang et al. (2016). The influence of the ALL forcing is detected for all the temperature extremes in most of the 12 grid boxes. In particular, the most detectable influence is on TNx, the annual maximum of daily minimum temperature, and the least detectable influence is on TXx, the annual maximum of daily maximum. In most cases, both naive and proposed methods reach consistent conclusions, although the point estimates of the scaling factor can be quite distinct from both methods in temperature minimums. For the maximums TNx and TXx, for example, the estimates of β are quite close except for the (115E, 30N) case. The analysis on the minimums, especially on the TXn, displays a much larger discrepancy in the $\hat{\beta}$. In the (110E, 40N) case, for instance, the difference in the $\hat{\beta}$ s bring about different conclusions in the attribution decision. Moreover, the proposed method tends to produce a larger standard error in most cases, which might lead to different conclusions even though the estimated values of β are close for both approaches. For example, both methods produce similar values in $\hat{\beta}$ in the (105E, 40N) grid box, but the lower bound of the 90% confidence interval is much closer to 1 using the proposed method while the counterpart for the naive method is way above 1. Overall, both methods agree on the detection and attribution

Fig. 2 The top and bottom graphs present the scaling factor estimation for all the temperature extremes and the corresponding 90% bootstrap confidence intervals for the selected grid boxes



results on TNx and TXx for all the grid boxes. ALL forcings are always detected on TNx using both methods, but its evidence of attribution is obvious in the grid boxes with latitude 40N. Different from the results on TNx, the influence of ALL forcings are not detected in grid boxes (110E, 30N), (115E, 30N), (110E, 35N), and (115E, 35N). There is also no evidence that the change in TXx is attributable to ALL forcings. For TNn and TXn, there are a few cases that the ALL forcings are not detectable. For the (115E, 30N) grid box, the influence of the ALL forcing is not detectable using the proposed method, while the naive method reach a different decision. In terms of attribution, only the (115E, 30N) and (115E, 40N) grid boxes show necessary evidence of attributing the ALL influence to the change in TNn. Similar to the results of TNn, the change in TXn can only be attributed to ALL forcings in two grid boxes. Interestingly, both methods agree on the attribution results for the (100E, 30N) grid box, but disagree on the attribution results for (110E, 40N) grid box.

For illustration purposes, we choose the (110E, 40N) grid box and report the point estimates (Est), the corresponding standard errors (SE), and the 90% confidence intervals (CI) of θ for each temperature extreme based on the two methods (Table 3). The estimates of ξ and σ between the two methods are quite similar, which is consistent with the simulation results in Sect. 4. On the other hand, the location parameters α and β can be very different between the naive estimates and the MCCA estimates in some cases. For instance, the estimates of α and β are very different between the two methods on TXn, directly resulting in a different conclusion. The naive method detects the ALL signal at all on TXn, but there is no evidence that the change on TXn is attributable to the influence of ALL forcings. On the opposite side, ALL forcings are not only detected by the proposed method but also demonstrate necessary evidence of attribution on TXn, with the corresponding 90% confidence interval being significantly greater than zero. For extremes other than TXn, both methods detect the ALL signal to the temperature change.

Table 3 Summaries of estimated parameters of the ALL signals for four temperature extremes in grid box indexed by longitude 110E and latitude 40N

Method	Param	TXx Est	SE	90% CI	TNx Est	SE	90% CI
α	Naive	5.733	13.286	(− 16.121, 27.587)	− 17.020	5.389	(− 25.884, − 8.156)
	MCCS	3.070	13.696	(− 19.458, 25.598)	− 18.152	5.262	(− 26.808, − 9.497)
β	Naive	0.737	0.341	(0.175, 1.298)	1.739	0.244	(1.338, 2.140)
	MCCS	0.805	0.352	(0.227, 1.384)	1.791	0.238	(1.399, 2.182)
σ	Naive	1.037	0.100	(0.873, 1.201)	0.686	0.062	(0.584, 0.788)
	MCCS	1.033	0.114	(0.844, 1.221)	0.677	0.070	(0.562, 0.792)
ξ	Naive	− 0.077	0.091	(− 0.227, 0.073)	− 0.112	0.069	(− 0.226, 0.002)
	MCCS	− 0.076	0.069	(− 0.189, 0.037)	− 0.113	0.061	(− 0.213, − 0.013)
Method	Param	TXn Est	SE	90% CI	TNn Est	SE	90% CI
α	Naive	19.567	8.333	(5.860, 33.273)	11.558	7.786	(− 1.249, 24.366)
	MCCS	29.222	13.503	(7.012, 51.432)	13.641	9.652	(− 2.234, 29.517)
β	Naive	1.786	0.486	(0.986, 2.587)	1.194	0.276	(0.740, 1.647)
	MCCS	2.350	0.793	(1.046, 3.654)	1.267	0.342	(0.705, 1.829)
σ	Naive	1.525	0.148	(1.281, 1.768)	1.462	0.133	(1.243, 1.680)
	MCCS	1.453	0.145	(1.214, 1.692)	1.448	0.127	(1.239, 1.656)
ξ	Naive	− 0.082	0.101	(− 0.248, 0.084)	− 0.147	0.084	(− 0.286, − 0.009)
	MCCS	− 0.052	0.107	(− 0.227, 0.124)	− 0.148	0.056	(− 0.239, − 0.056)

6 Discussion

Relatively little attention has been paid to understanding and correcting for the EIVs in the location parameters of GEV regressions. Motivated by the optimal fingerprinting method in the detection and attribution analysis of changes in climate extremes, we investigated the impact of measurement error in both the independent and time-correlated measurement error settings through simulation studies. The bias caused by the measurement error can be severe. Our MCCS method takes advantage of the existing score functions of the GEV distribution and the flexibility of Monte Carlo copies to correct for the bias of estimators. For inferences, the confidence intervals based on our block bootstrap approach appear to give acceptable coverage rates. Though our method was applied to single grid boxes, it is able to correct the bias and provide valid inference effectively. Given the fact that the signals at each grid box are much lower than those on a bigger spatial scale, our results are encouraging, indicating that our method can be served as a building block for further research on pooling boxes together to estimate β (Wang et al. 2021). The corrected score functions also allow further extension of the fingerprinting by incorporating covariates in the scale parameter to account for non-stationarity and other possible influencing factors (Huang et al. 2016). Overall, the MCCS

approach can be applied to GEV regressions with independent and temporally dependent EIVs that arise in many fields.

A few future directions are worth pursuing. Computationally, one concern is the numerical instability in estimation when the sample size is small. As many GEV regressions have a limited amount of observations, it would be useful to develop regularized estimation along the line of Firth (1993). For sandwich variance estimation, the “bread” part, or the inverse of the derivative of the corrected scores with respect to the parameters, could be poorly conditioned when the sample size is small. The regularized estimation might be helpful with this issue as well. Due to these computational issues, in our application to central China temperature extremes, we performed only one-signal analysis with the ALL signal. It is desirable to do two-signal analyses with both anthropogenic forcing and natural forcing at the same time. One possibility is to use a longer record of data, such as HadEX3 (Dunn et al. 2020) and CanESM5 (Swart et al. 2019). Methodologically, however, it would be interesting to extend the MCCS method to the spatial setting and restrict the scaling factors across grid boxes to be the same in one regional analysis (Zwiers et al. 2011). Incorporating spatial dependency to improve the effi-

ciency of the MCCS method similar to the combined score equations (Wang et al. 2021) merits a full investigation.

Declarations

Conflict of interest All authors declare that they have no conflicts of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

Appendix A Derivation of MCCS for GEV

The derivatives of the log density of $F(\cdot \mid \mu, \sigma, \xi)$ with respect to the parameter are (Prescott and Walden 1980)

$$\begin{aligned}
 f_1(Y; \mu, \sigma, \xi) &= \frac{1 + \xi - v}{\sigma w}, \\
 f_2(Y; \mu, \sigma, \xi) &= \frac{1}{\sigma} \left\{ -1 + \frac{(1 + \xi - v)(Y - \mu)}{\sigma w} \right\}, \\
 f_3(Y; \mu, \sigma, \xi) &= \begin{cases} \frac{\log(w)(1-v)}{\xi^2} - \frac{(1+\xi-v)(Y-\mu)}{\xi\sigma w} & \xi \neq 0 \\ \frac{1}{2} \left(1 - \exp\left(\frac{Y-\mu}{\sigma}\right) \right) \left(\frac{Y-\mu}{\sigma}\right)^2 - \frac{Y-\mu}{\sigma w} & \xi = 0 \end{cases},
 \end{aligned}$$

where $w = 1 + \xi(Y - \mu)/\sigma$, $v = w^{-1/\xi} \mathbb{1}(\xi \neq 0) + e^{-(Y-\mu)/\sigma} \mathbb{1}(\xi = 0)$, $\mu = Z^\top \alpha + X^\top \beta$, and $\mathbb{1}(\cdot)$ is the indicator function.

To derive the real part of the $\tilde{\psi}_B$, we first replace X with $\tilde{W}_{b,t}$ in $f_1(\cdot)$, $f_2(\cdot)$, and $f_3(\cdot)$. Then we can rewrite $Y - \mu$ as following,

$$Y - \mu = Y - W^\top \beta - Z^\top \alpha - \iota \epsilon_{b,t}^\top \beta = u - \iota v, \tag{A1}$$

where $u = Y - W^\top \beta - Z^\top \alpha$ and $v = \epsilon_{b,t}^\top \beta$. In addition, using polar representation, we have

$$\sigma + \xi(u - \iota v) = r e^{i\omega}, \tag{A2}$$

where $r = \sqrt{(u\xi + \sigma)^2 + (v\xi)^2}$ and $\omega = \text{atan2}\{-u\xi, (v\xi + \sigma)\}$. Note that $\text{atan2}\{-u\xi, (v\xi + \sigma)\}$ returns a single value ω such that $-\pi < \theta \leq \pi$ and, for some $r > 0$, $(a\xi + \sigma) = r \cos \omega$ and $-b\xi = r \sin \omega$. From Euler’s formula, we can

also write $r e^{i\omega}$ as

$$r e^{i\omega} = r \cos \omega + \iota r \sin \omega, \tag{A3}$$

or equivalently, $r e^{-i\omega} = r \cos \omega - \iota r \sin \omega$. Given these expressions, we can first start to rewrite $f_1(\cdot)$ in the following way, `allowdisplaybreaks`

$$\begin{aligned}
 f_1(Y; \mu, \sigma, \xi) &= \frac{1 + \xi - v}{\sigma w} \\
 &= \frac{1 + \xi - w^{-1/\xi}}{\sigma \{1 + \xi(Y - \mu)/\sigma\}} \\
 &= \frac{1 + \xi - \{1 + \xi(Y - \mu)/\sigma\}^{-1/\xi}}{\sigma \{1 + \xi(Y - \mu)/\sigma\}} \\
 &= \frac{1 + \xi - \{1 + \xi(u - \iota v)/\sigma\}^{-1/\xi}}{\sigma \{1 + \xi(u - \iota v)/\sigma\}} \text{ (use (A1))} \\
 &= (r e^{i\omega})^{-1} \left\{ 1 + \xi - (r e^{i\omega}/\sigma)^{-1/\xi} \right\} \text{ (use (A2))} \\
 &= \frac{\cos \omega - \iota \sin \omega}{r} \\
 &= \left[1 + \xi - \frac{r^{-1/\xi} \{\cos(\omega/\xi) - \iota \sin(\omega/\xi)\}}{\sigma^{-1/\xi}} \right] \text{ (use (A3))} \\
 &= c + \iota d,
 \end{aligned}$$

where

$$\begin{aligned}
 c &= \frac{\cos \omega}{r} \left\{ 1 + \xi - (\sigma/r)^{1/\xi} \cos(\omega/\xi) \right\} \\
 &\quad + \frac{\sin \omega}{r} (\sigma/r)^{1/\xi} \sin(\omega/\xi), \\
 d &= \frac{\cos \omega}{r} (\sigma/r)^{1/\xi} \sin(\omega/\xi) \\
 &\quad - \frac{\sin \omega}{r} \left\{ 1 + \xi - (\sigma/r)^{1/\xi} \cos(\omega/\xi) \right\}.
 \end{aligned}$$

In a similar way, we can also rearrange $f_2(\cdot)$ such that,

$$\begin{aligned}
 f_2(Y; \mu, \sigma, \xi) &= \frac{1}{\sigma} \{-1 + f_1 \cdot (u - \iota v)\} \\
 &= \frac{-1 + (c + \iota d)(u - \iota v)}{\sigma} \\
 &= \frac{-1 + cu + dv}{\sigma} + \iota \frac{du - cv}{\sigma}.
 \end{aligned}$$

We then focus on the case when $\xi \neq 0$ to separate the real and imaginary part of $f_3(\cdot)$, and we have,

$$f_3(Y; \mu, \sigma, \xi)$$

$$\begin{aligned}
 &= \frac{\log(w)(1 - \nu)}{\xi^2} - \frac{(\sigma f_2 + 1)}{\xi} \\
 &= \frac{1}{\xi^2} \log \left\{ \frac{\sigma + \xi(u - \nu)}{\sigma} \right\} \left[1 - \left\{ \frac{\sigma + \xi(u - \nu)}{\sigma} \right\}^{-1/\xi} \right] \\
 &\quad - \frac{(\sigma f_2 + 1)}{\xi} \\
 &= \frac{1}{\xi^2} \log \left(\frac{re^{i\omega}}{\sigma} \right) \left(1 - \frac{(re^{i\omega})^{-1/\xi}}{\sigma^{-1/\xi}} \right) - \frac{(\sigma f_2 + 1)}{\xi} \quad (\text{use (A2)}) \\
 &= \frac{\log r - \log \sigma + i\omega}{\xi^2} \times [1 - (\sigma/r)^{1/\xi} \{\cos(\omega/\xi) - i \sin(\omega/\xi)\}] \\
 &\quad - \frac{(\sigma f_2 + 1)}{\xi} \\
 &= \frac{\log(r/\sigma) \{1 - (\sigma/r)^{1/\xi} \cos(\omega/\xi)\}}{\xi} - \frac{\omega(\sigma/r)^{1/\xi} \sin(\omega/\xi)}{\xi^2} \\
 &\quad - \frac{(cu + dv)}{\xi} \\
 &\quad + i \frac{\omega \{1 - (\sigma/r)^{1/\xi} \cos(\omega/\xi)\}}{\xi^2} - i \frac{\log(r/\sigma) \sin(\omega/\xi)}{\xi^2} \\
 &\quad - i \frac{(du - cv)}{\xi}.
 \end{aligned}$$

The real part of the three functions can be readily obtained by,

$$\begin{aligned}
 \Re(f_1(Y; \mu, \sigma, \xi)) &= c, \\
 \Re(f_2(Y; \mu, \sigma, \xi)) &= \frac{-1 + cu + dv}{\sigma}, \\
 \Re(f_3(Y; \mu, \sigma, \xi)) &= \frac{\log(r/\sigma) \{1 - (\sigma/r)^{1/\xi} \cos(\omega/\xi)\}}{\xi^2} \\
 &\quad - \frac{\omega(\sigma/r)^{1/\xi} \sin(\omega/\xi)}{\xi^2} - \frac{(cu + dv)}{\xi}.
 \end{aligned}$$

Furthermore, $\Re(Wf_1) = \Re((W + \iota \epsilon_{b,t})(c + \iota d)) = cW - d\epsilon_{b,t}$, and $\Re(Zf_1) = Zc$. The derivation of the case when $\xi = 0$ can be done in a similar way.

Appendix B Sketch of technical arguments

B.1 Argument for Proposition 1

Denote $\tilde{\psi}_{B,t}(\theta) = \tilde{\psi}_B(Y_t, W_t, Z_t, \theta, \Sigma_t)$. By Taylor series,

$$n^{-1/2} \sum_{t=1}^n \left\{ \tilde{\psi}_{B,t}(\tilde{\theta}_{n,B}) - \tilde{\psi}_{B,t}(\theta_0) \right\}$$

$$\begin{aligned}
 &= n^{-1/2} \sum_{t=1}^n \left\{ \frac{\partial \tilde{\psi}_{B,t}(\theta_0)}{\partial \theta_0^\top} (\tilde{\theta}_{n,B} - \theta_0) + o_p \|\tilde{\theta}_{n,B} - \theta_0\| \right\} \\
 &= \left\{ n^{-1} \sum_{t=1}^n \frac{\partial \tilde{\psi}_{B,t}(\theta_0)}{\partial \theta_0^\top} \right\} \sqrt{n}(\tilde{\theta}_{n,B} - \theta_0) + o_p(1).
 \end{aligned}$$

Thus,

$$\begin{aligned}
 \sqrt{n}(\tilde{\theta}_{n,B} - \theta_0) &= - \left\{ n^{-1} \sum_{t=1}^n \frac{\partial \tilde{\psi}_{B,t}(\theta_0)}{\partial \theta_0^\top} \right\}^{-1} \\
 &\quad n^{-1/2} \sum_{t=1}^n \left\{ \tilde{\psi}_{B,t}(\tilde{\theta}_{n,B}) - \tilde{\psi}_{B,t}(\theta_0) \right\} + o_p(1).
 \end{aligned}$$

Subsequently, we have

$$n^{-1} \sum_{t=1}^n \frac{\partial \tilde{\psi}_{B,t}(\theta_0)}{\partial \theta_0^\top} = D + o_p(1),$$

and

$$\sqrt{n}(\tilde{\theta}_{n,B} - \theta_0) \rightarrow N(0, C),$$

where D and C are defined as in Proposition 1.

B.2 Argument for Proposition 2

Denote $\tilde{\psi}_{B,t}^*(\theta) = \tilde{\psi}_B^*(Y_t, W_t, Z_t, \theta, \Sigma)$. By Taylor series,

$$\begin{aligned}
 &n^{-1/2} \sum_{t=1}^n \left\{ \tilde{\psi}_{B,t}^*(\tilde{\theta}_{n,B}^*) - \tilde{\psi}_{B,t}^*(\theta_0) \right\} \\
 &= n^{-1/2} \sum_{t=1}^n \left\{ \frac{\partial \tilde{\psi}_{B,t}^*(\theta_0)}{\partial \theta_0^\top} (\tilde{\theta}_{n,B}^* - \theta_0) + o_p \|\tilde{\theta}_{n,B}^* - \theta_0\| \right\} \\
 &= \left\{ n^{-1} \sum_{t=1}^n \frac{\partial \tilde{\psi}_{B,t}^*(\theta_0)}{\partial \theta_0^\top} \right\} \sqrt{n}(\tilde{\theta}_{n,B}^* - \theta_0) + o_p(1).
 \end{aligned}$$

Thus,

$$\begin{aligned}
 \sqrt{n}(\tilde{\theta}_{n,B}^* - \theta_0) &= - \left\{ n^{-1} \sum_{t=1}^n \frac{\partial \tilde{\psi}_{B,t}^*(\theta_0)}{\partial \theta_0^\top} \right\}^{-1} \\
 &\quad n^{-1/2} \sum_{t=1}^n \left\{ \tilde{\psi}_{B,t}^*(\tilde{\theta}_{n,B}^*) - \tilde{\psi}_{B,t}^*(\theta_0) \right\} + o_p(1).
 \end{aligned}$$

Let $D^* = E \left\{ \frac{\partial \tilde{\psi}_B^*(Y_t, W_t, Z_t, \theta)}{\partial \theta^\top} \right\}$, then we have

$$D^* = E \left\{ \frac{\partial \tilde{\psi}_B^*(Y_t, W_t, Z_t, \theta)}{\partial \theta^\top} \right\}$$

$$\begin{aligned}
 &= \frac{\partial E\{\tilde{\psi}_B^*(Y_t, W_t, Z_t, \theta)\}}{\partial \theta^\top} \\
 &= \frac{\partial E[E\{\tilde{\psi}_B^*(Y_t, W_t, Z_t, \theta) \mid (Y_t, X_t, Z_t)\}]}{\partial \theta^\top} \\
 &= \frac{\partial E\{\psi^*(Y_t, X_t, Z_t, \theta)\}}{\partial \theta^\top} \\
 &= E\left\{\frac{\partial \psi^*(Y_t, X_t, Z_t, \theta)}{\partial \theta^\top}\right\} \\
 &= E\left[E\left\{\frac{\partial \psi^*(Y_t, X_t, Z_t, \theta)}{\partial \theta^\top} \mid X_t\right\}\right] \\
 &= E\left(\operatorname{Re}\left[E\left\{\frac{\partial \psi^*(Y_t, \tilde{W}_{t,b}^*, Z_t, \theta)}{\partial \theta^\top} \mid X_t\right\}\right]\right).
 \end{aligned}$$

Thus, we have

$$n^{-1} \sum_{t=1}^n \frac{\partial \tilde{\psi}_{B,t}^*(\theta_0)}{\partial \theta_0^\top} = D^* + o_p(1),$$

and

$$\sqrt{n}(\tilde{\theta}_{n,B}^* - \theta_0) \rightarrow N(0, C^*),$$

where $C^* = \operatorname{Var}\{n^{-1/2} \sum_{t=1}^n \tilde{\psi}_B^*(Y_t, W_t, Z_t, \theta, \Sigma)\}$. Considering the correlations among $\tilde{\psi}_{B,t}^*$'s caused by correlated measurement errors, C^* can be decomposed as $\operatorname{Var}\{\tilde{\psi}_{B,t}^*(\theta)\}$ and $\operatorname{Cov}\{\tilde{\psi}_{B,t}^*(\theta)\tilde{\psi}_{B,t'}^*(\theta)\}$.

For the b th Monte Carlo copy of year t , we generate \tilde{Z}_{b1} and \tilde{Z}_{b2} independently from $N(0, I_{np \times np})$ and construct $\tilde{W}_{b,t}^* = W_t + \iota \Delta_t \tilde{Z}_{b1}$ and $\tilde{W}_{b,t'}^* = W_{t'} + \iota \Delta_{t'} \tilde{Z}_{b2}$, where $\Delta = \Sigma^{1/2}$ and $\{\Delta_t\}_{p \times np}$ is the top t th chunk of Δ . We have

$$\begin{aligned}
 &E\left[\operatorname{Re}\left\{\psi(Y_t, \tilde{W}_{b,t}^*, Z_t, \theta)\psi^\top(Y_t, \tilde{W}_{b,t}^*, Z_t, \theta)\right\} \mid (Y_t, X_t, Z_t)\right] \\
 &= E\left[\operatorname{Re}\left\{\psi(Y_t, X_t + e_t + \iota \Delta_t \tilde{Z}_{b1}, Z_t, \theta)\right.\right. \\
 &\quad \left.\left.\psi^\top(Y_t, X_t + e_t + \iota \Delta_t \tilde{Z}_{b1}, Z_t, \theta) \mid (Y_t, X_t, Z_t)\right\}\right] \\
 &(\text{Expand around } X_t) \approx \psi(Y_t, X_t, Z_t, \theta)\psi^\top(Y_t, X_t, Z_t, \theta),
 \end{aligned}$$

and similarly,

$$\begin{aligned}
 &E\left[\operatorname{Re}\left\{\psi(Y_t, \tilde{W}_{b,t}^*, Z_t, \theta)\psi^\top(Y_{t'}, \tilde{W}_{b,t'}^*, Z_{t'}, \theta)\right\}\right. \\
 &\quad \left.\mid (Y_t, Y_{t'}, X_t, X_{t'}, Z_t, Z_{t'})\right] \\
 &\approx \psi(Y_t, X_t, Z_t, \theta)\psi^\top(Y_{t'}, X_{t'}, Z_{t'}, \theta).
 \end{aligned}$$

Thus, for $t = 1, \dots, n$ and $t' = 1, \dots, n$, a consistent estimator for C^* can be constructed as below,

$$\begin{aligned}
 \widehat{C}_{\tilde{\psi}^*} &= \frac{1}{nB} \left[\sum_{t,t'=1}^n \sum_{b=1}^B \Re\left\{\psi(Y_t, \tilde{W}_{b,t}^*, Z_t, \theta_{n,B}^*)\right.\right. \\
 &\quad \left.\left.\psi^\top(Y_{t'}, \tilde{W}_{b,t'}^*, Z_{t'}, \theta_{n,B}^*)\right\}\right].
 \end{aligned}$$

B.3 Argument for Lemma 2

We denote $\tilde{Z}_b^\top = (\tilde{Z}_{b,1}^\top, \dots, \tilde{Z}_{b,t}^\top, \dots, \tilde{Z}_{b,n}^\top)^\top$ from $N(0, I_{np \times np})$ and generate $\tilde{W}_b^* = W + \iota \Sigma^{1/2} \tilde{Z}_b$ for the b th Monte Carlo copy. The first equality is based on Lemma 1–2 in Stefanski and Cook (1995). Then, we denote $\{\Delta_t\}_{p \times np}$ be the top t th chunk of $\Sigma^{1/2}$ and have

$$\begin{aligned}
 E\left[\operatorname{Re}\left\{\psi(Y_t, \tilde{W}_{b,t}^*, \theta)\right\} \mid (Y_t, X_t)\right] &= E\left[\operatorname{Re}\left\{\psi(Y_t, X_t + e_t\right.\right.\right. \\
 &\quad \left.\left.\left.+ \iota \Delta_t \tilde{Z}_{b,t}, \theta)\right\}\right] = \psi(Y_t, X_t, \theta).
 \end{aligned}$$

As the measurement error e_t and $\Delta_t \tilde{Z}_{b,t}$ are independent and identically distributed random vectors generated from $N(0, \Sigma)$, the last equality is a result of Lemma 4 in Stefanski and Cook (1995). Hence, we finish the proof for Proposition 2.

Appendix C Additional figures

This section contains plots for parameters, α , σ , and ξ , in the application of modeling temperature extremes in central China. See Figs. 3, 4, and 5, respectively. Each plot includes the point estimates and corresponding 90% bootstrap confidence intervals. In most grid boxes, the estimates of shape parameter ξ are negative, which aligns with the findings in Huang et al. (2016).

Fig. 3 The top and bottom graphs present the location parameter α estimation for all the temperature extremes and the corresponding 90% bootstrap confidence intervals for the selected grid boxes

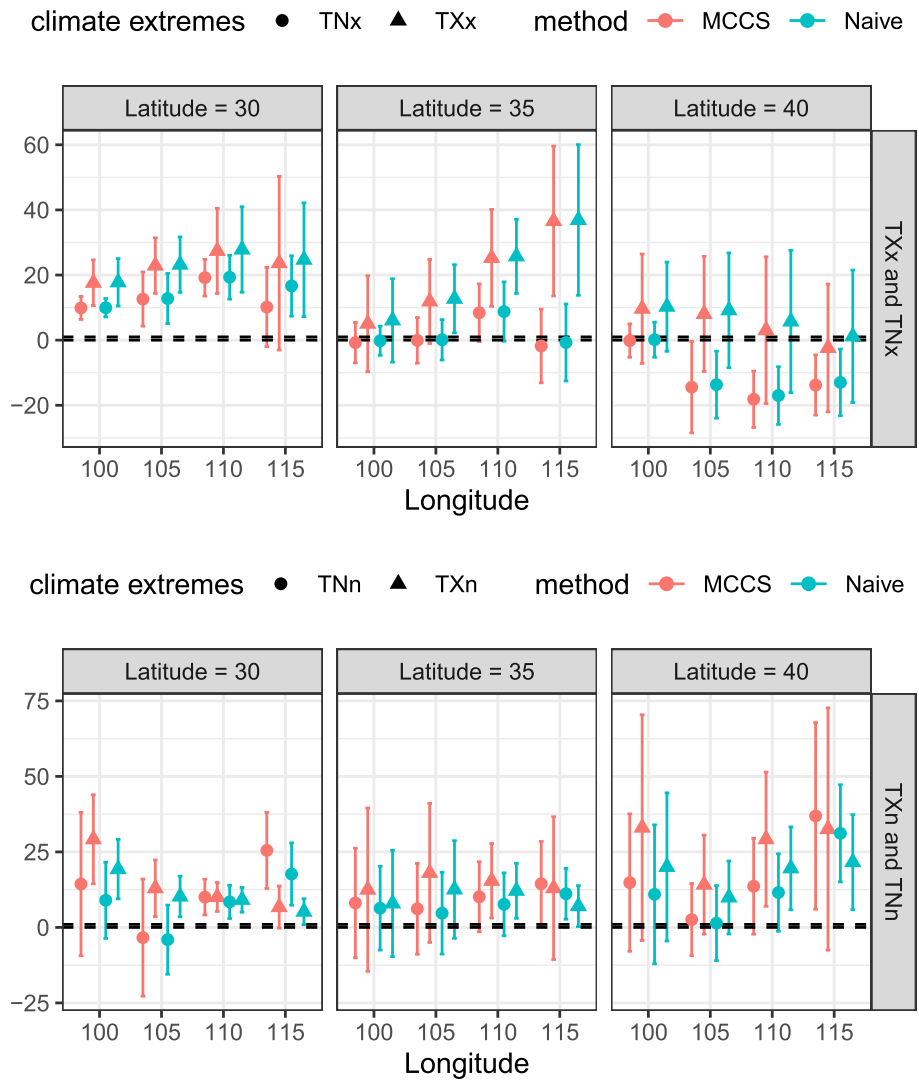


Fig. 4 The top and bottom graphs present the scale parameter σ estimation for all the temperature extremes and the corresponding 90% bootstrap confidence intervals for the selected grid boxes

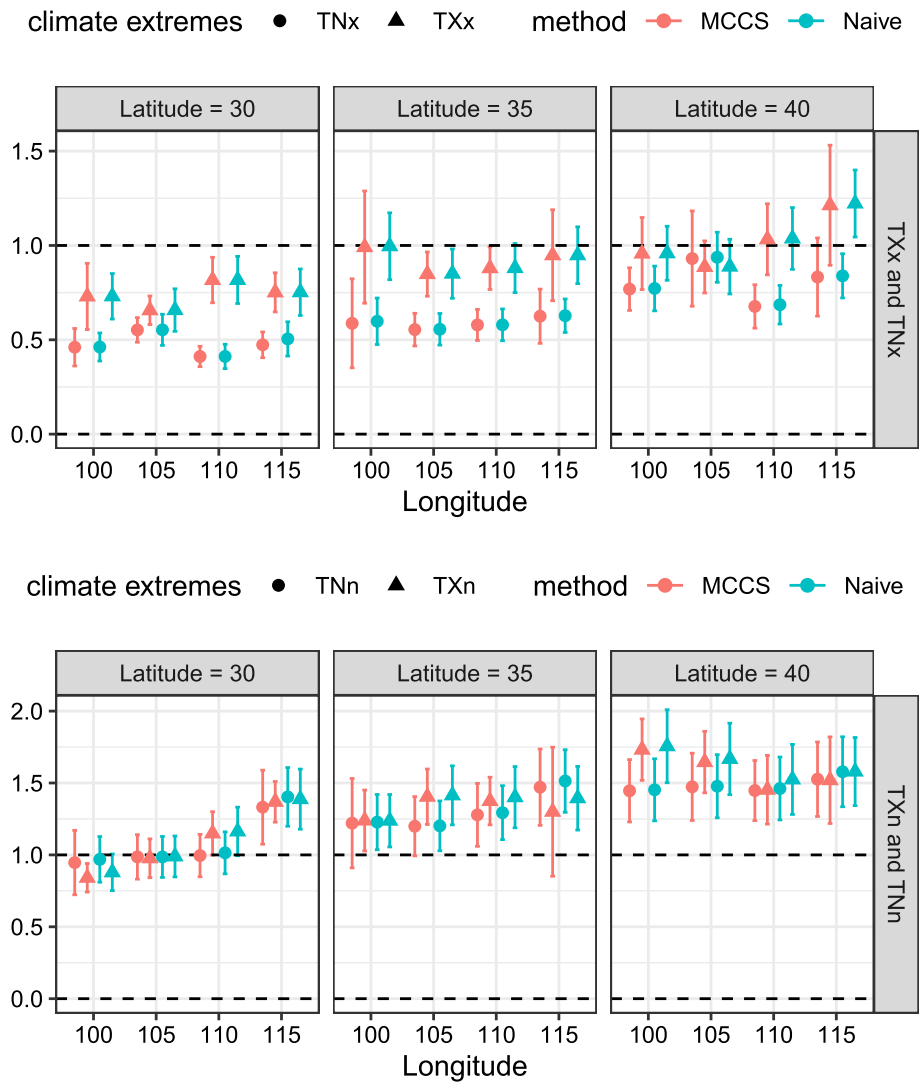
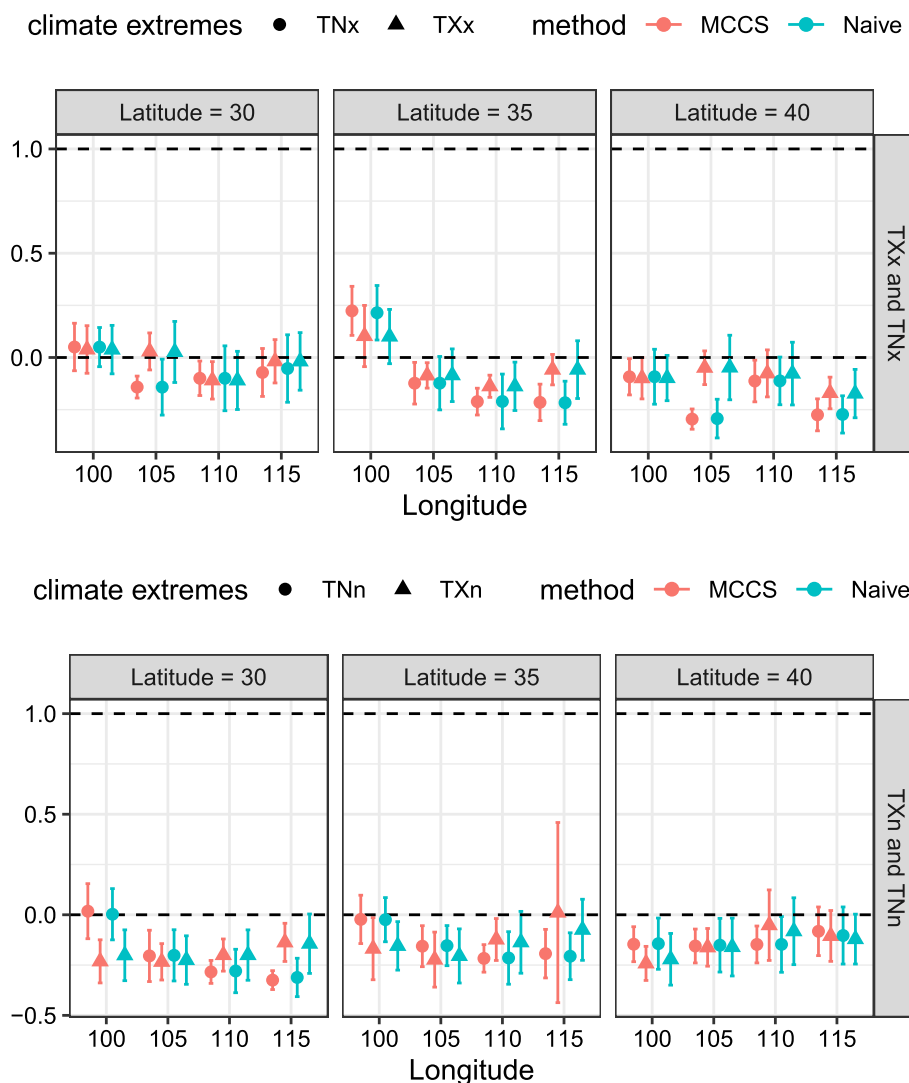


Fig. 5 The top and bottom graphs present the scale parameter ξ estimation for all the temperature extremes and the corresponding 90% bootstrap confidence intervals for the selected grid boxes



References

- Allen, M.R., Stott, P.A.: Estimating signal amplitudes in optimal fingerprinting, part I theory. *Clim. Dyn.* **21**(5–6), 477–491 (2003)
- Bücher, A., Segers, J.: On the maximum likelihood estimator for the generalized extreme-value distribution. *Extremes* **20**(4), 839–872 (2017)
- Chylek, P., Li, J., Dubey, M., Wang, M., Lesins, G.: Observed and model simulated 20th century arctic temperature variability: Canadian earth system model CanESM2. *Atmos. Chem. Phys. Discuss.* **11**(8), 22893–22907 (2011)
- Coles, S., Bawa, J., Trenner, L., Dorazio, P.: *An Introduction to Statistical Modeling of Extreme Values*, vol. 208. Springer (2001)
- Donat, M., Alexander, L., Yang, H., Durre, I., Vose, R., Dunn, R., Willett, K., Aguilar, E., Brunet, M., Caesar, J., et al.: Updated analyses of temperature and precipitation extreme indices since the beginning of the twentieth century: the hadex2 dataset. *J. Geophys. Res. Atmos.* **118**(5), 2098–2118 (2013)
- Dunn, R.J., Alexander, L.V., Donat, M.G., Zhang, X., Bador, M., Herold, N., Lippmann, T., Allan, R., Aguilar, E., Barry, A.A., et al.: Development of an updated global land in situ-based data set of temperature and precipitation extremes: Hadex3. *J. Geophys. Res. Atmos.* **125**(16), e2019JD032263 (2020)
- Field, C.B., Barros, V., Stocker, T.F., Dahe, Q.: *Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation: Special Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press (2012)
- Firth, D.: Bias reduction of maximum likelihood estimates. *Biometrika* **80**(1), 27–38 (1993)
- Fisher, R.A., Tippett, L.H.C.: Limiting forms of the frequency distribution of the largest or smallest member of a sample. *Math. Proc. Camb. Philos. Soc.* **24**(2), 180–190 (1928)
- Gilleland, E.: Bootstrap methods for statistical inference. Part II: extreme-value analysis. *J. Atmos. Ocean. Technol.* **37**(11), 2135–2144 (2020)
- Hamdi, Y., Haigh, I.D., Parey, S., Wahl, T.: Preface: advances in extreme value analysis and application to natural hazards. *Nat. Hazards Earth Syst. Sci.* **21**(5), 1461–1465 (2021)
- Hannart, A.: Integrated optimal fingerprinting: method description and illustration. *J. Clim.* **29**(6), 1977–1998 (2016)
- Hasselmann, B.: nleqslv: solve systems of nonlinear equations (2018). <https://CRAN.R-project.org/package=nleqslv>. R package version 3.3.2
- Hasselmann, K.: Optimal fingerprints for the detection of time-dependent climate change. *J. Clim.* **6**(10), 1957–1971 (1993)

- Heffernan, J.E., Stephenson, A.G.: *ismev: an introduction to statistical modeling of extreme values* (2018). <https://CRAN.R-project.org/package=ismev>
- Hegerl, G.C., von Storch, H., Hasselmann, K., Santer, B.D., Cubasch, U., Jones, P.D.: Detecting greenhouse-gas-induced climate change with an optimal fingerprint method. *J. Clim.* **9**(10), 2281–2306 (1996)
- Huang, W.K., Stein, M.L., McInerney, D.J., Sun, S., Moyer, E.J.: Estimating changes in temperature extremes from millennial-scale climate simulations using generalized extreme value (GEV) distributions. *Adv. Stat. Climatol. Meteorol. Oceanogr.* **2**(1), 79–103 (2016)
- Huerta, G., Sansó, B.: Time-varying models for extreme values. *Environ. Ecol. Stat.* **14**(3), 285–299 (2007)
- Hyun, N., Couper, D.J., Zeng, D.: Gumbel regression models for a monotone increasing continuous biomarker subject to measurement error. *J. Stat. Plan. Inference* **203**, 160–168 (2019)
- Katzfuss, M., Hammerling, D., Smith, R.L.: A Bayesian hierarchical model for climate change detection and attribution. *Geophys. Res. Lett.* **44**(11), 5720–5728 (2017)
- Li, Y., Chen, K., Yan, J., Zhang, X.: Uncertainty in optimal fingerprinting is underestimated. *Environ. Res. Lett.* **16**(8), 084043 (2021)
- Li, Y., Chen, K., Yan, J., Zhang, X.: Regularized fingerprinting in detection and attribution of climate change with weight matrix optimizing the efficiency in scaling factor estimation. *Ann. Appl. Stat.* **17**(1), 225–239 (2023)
- Mancl, L.A., DeRouen, T.A.: A covariance estimator for gee with improved small-sample properties. *Biometrics* **57**(1), 126–134 (2001)
- Martins, E.S., Stedinger, J.R.: Generalized maximum-likelihood generalized extreme-value quantile estimators for hydrologic data. *Water Resour. Res.* **36**(3), 737–744 (2000)
- Nakamura, T.: Corrected score function for errors-in-variables models: methodology and application to generalized linear models. *Biometrika* **77**(1), 127–137 (1990)
- Nolde, N., Zhou, C.: Extreme value analysis for financial risk management. *Annu. Rev. Stat. Appl.* **8**, 217–240 (2021)
- Novick, S.J., Stefanski, L.A.: Corrected score estimation via complex variable simulation extrapolation. *J. Am. Stat. Assoc.* **97**(458), 472–481 (2002)
- Prescott, P., Walden, A.: Maximum likelihood estimation of the parameters of the generalized extreme-value distribution. *Biometrika* **67**(3), 723–724 (1980)
- Smith, R.L.: Maximum likelihood estimation in a class of nonregular cases. *Biometrika* **72**(1), 67–90 (1985)
- Smith, R.L.: Extreme value analysis of environmental time series: an application to trend detection in ground-level ozone. *Stat. Sci.* **4**, 367–377 (1989)
- Stefanski, L.A., Cook, J.R.: Simulation-extrapolation: the measurement error jackknife. *J. Am. Stat. Assoc.* **90**(432), 1247–1256 (1995)
- Stein, M.L.: Some statistical issues in climate science. *Stat. Sci.* **35**(1), 31–41 (2020)
- Stephenson, A.G.: *evd: extreme value distributions*. *R News* **2**(2), 31–32 (2002)
- Swart, N.C., Cole, J.N., Kharin, V.V., Lazare, M., Scinocca, J.F., Gillett, N.P., Anstey, J., Arora, V., Christian, J.R., Hanna, S., et al.: The Canadian earth system model version 5 (CanESM5. 0.3). *Geosci. Model Dev.* **12**(11), 4823–4873 (2019)
- Wang, X.L., Zwiers, F.W., Swail, V.R.: North Atlantic ocean wave climate change scenarios for the twenty-first century. *J. Clim.* **17**(12), 2368–2383 (2004)
- Wang, Z., Jiang, Y., Wan, H., Yan, J., Zhang, X.: Detection and attribution of changes in extreme temperatures at regional scale. *J. Clim.* **30**(17), 7035–7047 (2017)
- Wang, Z., Jiang, Y., Wan, H., Yan, J., Zhang, X.: Toward optimal fingerprinting in detection and attribution of changes in climate extremes. *J. Am. Stat. Assoc.* **116**(533), 1–13 (2021)
- Westgate, P.M.: A bias-corrected covariance estimate for improved inference with quadratic inference functions. *Stat. Med.* **31**(29), 4003–4022 (2012)
- Zwiers, F.W., Zhang, X., Feng, Y.: Anthropogenic influence on long return period daily temperature extremes at regional scales. *J. Clim.* **24**(3), 881–892 (2011)

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