Frequency analysis of precipitation extremes under a changing climate: a case study in Heihe River basin, China
Qingyun Tian, Zhanling Li and Xueli Sun

ABSTRACT

The stationary assumption for the traditional frequency analysis of precipitation extremes has been challenged due to natural climate variability or human intervention. To overcome this challenge, this paper, taking Heihe River basin as the case study, performed the frequency analysis by developing a nonstationary GEV model for those seasonal maximum daily precipitation (SMP) time series with nonstationary characteristics by employing the GEV conditional density estimation network. In addition, the confidence intervals (CIs) of estimated return levels were also investigated by using the residual bootstrap technique. Results showed that, 7 of 12 SMP series were nonstationary. The parameters in the nonstationary model were specified as functions of time varying or correlated climate indices varying covariates. The frequency analysis showed that the return levels varied linearly or nonlinearly with covariates. Precipitation extremes with the same magnitude in the study area were found to be occurring more frequently in the future. The CIs of such return levels increased with time passing, especially those from the more complex GEV11 model, embedding a nonlinear increasing trend in model scale parameters. It implied that the increase of model complexity is likely to result in the increase of uncertainty in estimates.

Key words | confidence interval, GEV, nonstationary, precipitation extremes

INTRODUCTION

The Intergovernmental Panel on Climate Change (IPCC) in the Fifth Assessment Report of Working Group I clearly pointed out that global climate system warming is unquestionable (IPCC 2013). This may be reflected in the more intense process of the water cycle and the increases in frequency and magnitude of precipitation (Wobus et al. 2014; Wasko et al. 2016; Fotovatikhah et al. 2018; Li et al. 2018). Studies on precipitation extremes have caused widespread concern due to their considerable impacts on agriculture, industry, ecosystem, humans, etc. (You et al. 2011; Zhang et al. 2011; Derbile & Kasei 2012; Wang et al. 2012; Cheng et al. 2015; Lu et al. 2017).

Extreme value theory (EVT), one of the important branches of statistics, has been widely used in meteorological and hydrological extremes (e.g., AghaKouchak & Nasrollahi 2010; Shinjie et al. 2013; Vasiliades et al. 2015; Agilan & Umamahesh 2017). However, the stationary assumption in using the EVT in hydrology has been gradually challenged in the context of climate change (e.g., Coles 2001; Fischer et al. 2012; Soukissian & Tsalis 2015; Lu et al. 2017). Studies in the last few decades have indicated that the hydroclimatic records exhibit some types of nonstationary features, such as trends or shifts on time scales whether due to the climate variability or the effects of human activities (Milly et al. 2008; Peterson et al. 2008; Zolina et al. 2008; DeGaetano et al. 2011; Derbile & Kasei 2012; Wang et al. 2012; Cheng et al. 2015; Lu et al. 2017).

This is an Open Access article distributed under the terms of the Creative Commons Attribution Licence (CC BY 4.0), which permits copying, adaptation and redistribution, provided the original work is properly cited (http://creativecommons.org/licenses/by/4.0/).

doi: 10.2166/wcc.2020.170

772
© 2021 The Authors
Journal of Water and Climate Change | 12.3 | 2021
models need to be proposed and identified (Galiatsatou & Prinos 2011; Katz 2013; Salas & Obeysekera 2014; Gao et al. 2016; Šraj et al. 2016; Hamdi et al. 2018).

The nonstationary features of hydroclimatic extremes can be described by introducing the covariates into EVT (e.g., Cheng et al. 2014; Lu et al. 2017; Um et al. 2017). Specifically, the parameters in GEV, including location parameter, scale parameter, and shape parameter, could be expressed as regression structures allowing linear or quadratic dependence on time (e.g., Katz 2013; Kim et al. 2017). Besides the variable of time, appropriate explanatory climatic variables, such as Western Pacific index (WPI), East Asian summer monsoon index (EASMI), etc., also can be selected as covariates in some cases (e.g., Villarini et al. 2010; Cannon 2015; Vasilides et al. 2015; Gao et al. 2016; Šraj et al. 2016; Hao et al. 2019).

Frequency analyses of extremes are of great importance and great use for risk management and engineering design from a practical point of view (Katz 2013; Salas & Obeysekera 2014). Return levels at different return periods usually need to be estimated in frequency analysis. Katz et al. (2002) extended the traditional stationary return level into the nonstationary framework, in which the return level varies over time to keep the occurrence probability of an extreme event constant. In estimating the return levels of extremes, large uncertainties exist due to the short time series available (Serinaldi & Kilsby 2015; Kim et al. 2017; Hamdi et al. 2018). Furthermore, only point estimates of return levels are indeed of little use for decision-makers from a practical point of view. It is also a common belief today that a good estimation of uncertainty in extreme levels can be as important as the estimate of the level itself. Therefore, the confidence intervals (CIs), indicating the reliability and uncertainty of return level estimates, are required (Cooley 2013; Obeysekera & Salas 2014; Serinaldi & Kilsby 2015). As suggested by Serinaldi & Kilsby (2015), the bootstrap method will be selected to quantify the CIs of return level estimates in this study, since it relies on the available information without any asymptotic hypotheses, does not depend on a particular estimation method, and can easily be implemented regardless of the model complexity.

Overall, this study mainly focuses on conducting frequency analysis of precipitation extremes in a nonstationary context by developing a nonstationary GEV model with GEV-CDN. The second largest inland river basin in northwest China, the Heihe River basin, is selected as the case study. The seasonal maximum daily precipitation (SMP) series over the basin are analyzed. Both time and climate indices are considered as covariates in the nonstationary modeling. With the best fitting model, return levels at different return periods are estimated, and their 95% CIs, indicating the reliability of return level estimates, are finally quantified and discussed.

STUDY AREA AND DATASETS

Heihe River basin (Figure 1) is the second largest inland basin in northwest China with an area of 142,900 km². It is located in the climatic intersection between the Westerlies and the East Asian summer monsoon (Qin et al. 2010), which belongs to a typical continental arid climate. The basin has an important strategic position in northwest China. The middle reach of the basin, on the ancient ‘Silk Road’ and the current Asia–Europe Continental Bridge, is one of the top ten commodity grain bases in China, with a long history of agriculture; the Ejina Oasis in the lower reach is an important ecological security barrier. However, water resources in the basin are scarce. Understating the precipitation extremes in the basin is of great importance not only for the local water allocation but also for the ecological restoration.

The basin is divided into three reaches, the upper, middle, and lower reaches, which correspond to mountainous areas, oasis areas, and arid Gobi Desert areas, respectively. The upper reach of the basin, ranging from Qilian Mountain to Yingluoxia gorge, is the area where most of the runoff is generated. The annual average temperature is less than 2 °C and the annual precipitation is 200–500 mm, but reaches 700 mm in some mountain areas. The middle reach, ranging from Yingluo gorge to Zhengyi gorge, is the farmland oasis where most water resources are consumed by agricultural irrigation. The annual average temperature is 6–8 °C and the annual precipitation is 50–150 mm. The lower reach in Ejina Banner is characterized by very few annual precipitations, with even less than 50 mm. However, Ejina Oasis in the lower reach, as an ecological security barrier, plays an important role in preventing the expansion of desertification areas. Scarce water resources in the lower reach lead to it being difficult
to maintain a healthy and sustainable oasis system. Severe environmental deterioration has been experienced in the past several decades (Zhang et al. 2017).

The precipitation data in the basin used here are the gridded dataset provided by Cold and Arid Regions Science Data Center at Lanzhou (http://westdc.westgis.ac.cn) by merging gauge observations and regional climate model simulations. The data span from 1960 to 2014 with 3 km spatial resolution and daily temporal scale. The dataset has been cross-validated and verified (Yang et al. 2018). With the aim of better understanding the spatial and seasonal behaviors of precipitation extremes in the basin, the precipitation extremes during four seasons (spring, summer, autumn, and winter) from three reaches (upper, middle, and lower) are investigated. The areal precipitation over each reach is calculated by averaging all gridded precipitation data over the reach. The daily maximum precipitation series in each season for each reach are finally extracted and then used in the following analysis.

In developing a nonstationary GEV model, besides the time $t$, the EASMI and the Western Pacific index (WPI) are also considered as covariates, since Cheng et al. (2015) and Chen et al. (2017) have concluded that precipitation extremes over the study area vary with these two climate indices. The annual EASMI time series are then collected from http://ljp.gcess.cn/dct/page/65577. The WPI data (monthly/annual) are collected from http://www.cpc.ncep.noaa.gov/data/ provided by the U.S. National Centers for Environmental Prediction (NCEP).

**METHODS AND MATERIALS**

**Trend and stationary tests**

The Mann–Kendall (M-K) test is one of the commonly used trend tests, which is often used to analyze the trend of time series of precipitation, runoff, air temperature, and other variables. It is a rank-based nonparametric trend detection method that is less sensitive to outliers than parametric statistics, such as Pearson's correlation coefficient (Kendall 1958; Mann 1945). The null hypothesis in the M-K test states that there is no trend in the time series and observations which are randomly ordered. On the contrary, the alternative hypothesis means that there are increasing or decreasing monotonic trends. The resultant M-K test statistic (tau)
indicates how strong the trend is and whether it is increasing or decreasing. More details about the test can be found in studies of Libiseller & Grimvall (2002) and Gao et al. (2016). The M-K test is done by using ‘trend’ package in R.

The stationarity test is carried out using the augmented Dickey–Fuller test (ADF), which is an augmented version of the original Dickey–Fuller test, to verify the stationarity of a larger and more complicated set of time series (Said & Dickey 1984). The null hypothesis in the ADF test is nonstationary, that is, a unit root is present in the time series. The statistic Dickey–Fuller used in the test is a negative number. The more negative it is, the stronger the rejection of the null hypothesis that there is a unit root at some level of confidence. More details about the ADF test can be found in the studies of Chang & Park (2002) and Li et al. (2017). The ADF test is done by using ‘tseries’ package in R.

**GEV-CDN model**

The generalized extreme value conditional density estimation network (GEV-CDN) is employed for the evaluation of GEV modeling (Cannon 2010, 2011). Parameters in a GEV model are specified as functions of covariates using a probabilistic variant of the multilayer perceptron (MLP) neural network. Due to the flexibility of neural network architecture, the GEV model is capable of representing a wide range of nonstationary relationships, including those involving interactions between covariates. Taking time $t$ as the covariate as an example, the cumulative density function (cdf) of a random variable $Y$ drawn from a nonstationary GEV distribution is given by:

\[
F(Y; \mu(t), \alpha(t), \kappa(t)) = \begin{cases} 
\exp \left[ -\left( 1 - \kappa(t) \frac{Y - \mu(t)}{\alpha(t)} \right)^{\frac{1}{\kappa(t)}} \right] & \kappa \neq 0, \quad 1 - \kappa(t) \frac{Y - \mu(t)}{\alpha(t)} > 0 \\
\exp \left[ -\exp \left( 1 - \frac{Y - \mu(t)}{\alpha(t)} \right) \right] & \kappa = 0
\end{cases}
\]

(1)

where $\mu(t)$, $\alpha(t)$, and $\kappa(t)$ are the location, scale, and shape parameters, respectively. The location parameter specifies the center of the distribution, the scale parameter gives an indication of the size of deviations around the location, and the shape parameter governs the tail behavior of GEV distribution. These parameters are the functions of the covariate $t$.

A flowchart of applied GEV-CDN model is shown in Figure 2. Given covariates at time $t$, $x(t) = \{x_i(t), i = 1, ..., I \}$ ($I$ means the numbers of covariates). The output from the hidden-layer node $h_j$ is given by applying the hidden-layer...
activation function $m(\cdot)$ to the inner product between the covariates and the input-hidden layer weights $w_{ji}^{(1)}$ plus the bias $b_i^{(1)}$. The $m(\cdot)$ is taken to be a sigmoidal function, e.g., the hyperbolic tangent function $\tanh (\cdot)$ for the nonlinear GEV-CDN network. The identity function is adopted if the GEV-CDN mapping is to be strictly linear. The value of the $k$th output from the network $o_k$ is then given by the hidden-output layer weights $w_{kj}^{(2)}$ and hidden-output layer biases $b_k^{(2)}$. In the present study, the value of $\kappa^+$ is set to be 0.5 according to the approach of generalized maximum likelihood (GML) in Martins & Stedinger (2000). More details on the GEV-CDN architecture can be found in Cannon (2010).

**Model selection**

Four models with different complexity are defined within GEV-CDN model architecture, GEV0, GEV1, GEV11, and GEV2, respectively. GEV0 is the stationary case with all parameters independent of time; the latter three are nonstationary cases. GEV1 assumes a linear change in the location parameter, GEV2 assumes a quadratic change in the location parameter, and GEV11 assumes a linear change in the location and a log-linear change in the scale parameter. The shape parameter keeps constant with time changing in all models since it is not easy to be estimated precisely even for the stationary case (Coles 2001; El Adlouni et al. 2007) and thus usually suggested to be assumed as a constant (Hamdi et al. 2018). On the other hand, the variations of this parameter surely increase the complexity of the model, resulting in larger uncertainties in the frequency analysis.

The Akaike information criterion with small sample size correction (AICc) (Hurvich & Tsai 1989) are used to select the best performance model in GEV-CDN. The AICc are cost-complexity model selection criteria that penalize the negative log-likelihood as a function of the number of model parameters. The AICc are penalty functions estimated using the formula as follows:

$$AICc = -2 \log (L) + 2P + \frac{2P(P + 1)}{N - P - 1}$$  \hspace{1cm} (2)

where $L$ is the likelihood function, $P$ is the number of model parameters, and $N$ represents the sample size. The model that minimizes AICc is deemed to give the best trade-off between maximizing model fit and minimizing model complexity and is selected for the final use.

**Return level estimation**

Constant parameters correspond to a stationary GEV model. From Equation (1), the probabilistic quantile $q_p$ can be obtained from the following equation:

$$q_p = \begin{cases} \mu - \frac{\alpha}{\kappa} \left[1 - (\log p)^{-\kappa}\right], & \kappa \neq 0 \\ \mu - \alpha \log \left(-\log p\right), & \kappa = 0 \end{cases}$$  \hspace{1cm} (3)

where $p$ is the non-exceedance probability ($0 < p < 1$), and the annual maximum (or minimum) $q_p$ corresponds to the return period $T = 1/(1 - p)$.

For a nonstationary GEV model, estimates of extreme quantiles with non-exceedance probability $p$ can be derived by (Katz et al. 2002; Vasiliades et al. 2013):

$$q_{pt} = \begin{cases} \mu(t) - \frac{\alpha(t)}{\kappa(t)} \left[1 - (\log p)^{-\kappa(t)}\right], & \kappa(t) \neq 0 \\ \mu(t) - \alpha(t) \log \left(-\log p\right), & \kappa(t) = 0 \end{cases}$$  \hspace{1cm} (4)

where $\mu(t)$, $\alpha(t)$ and $\kappa(t)$ are time-dependent GEV parameters.

**Confidence interval estimation**

A bootstrap-based technique is used to estimate the CIs for return levels of a nonstationary GEV distribution. Following the suggestions of Kharin & Zwiers (2004) and Khalid et al. (2006), the residual bootstrap technique is used here. It proceeds by:

1. fitting a nonstationary GEV model to the observed data;
2. transforming residuals from the fitted model to be identically distributed:

$$\epsilon(t) = \left[1 - \kappa \frac{(y(t) - \mu(t))}{\alpha(t)^{1/\kappa}}\right]^{1/\kappa}$$  \hspace{1cm} (5)

where $\epsilon(t)$ is the $t$th transformed residual, and $\mu(t)$, $\alpha(t)$ and $\kappa$ are GEV parameters from the fitted model;
3. resampling the transformed residuals with replacement to form bootstrapped residuals \( \{e^{(b)}(t), t = 1, \ldots, N\} \);
4. rescaling the bootstrapped residuals by inverting the transformation:

\[
y^{(b)}(t) = \mu(t) - \alpha(t) \frac{(e^{(b)}(t) - 1)}{\kappa}
\]  

(6)

5. fitting a nonstationary GEV model to the bootstrapped samples;
6. estimating GEV parameters and return levels from the fitted model; and
7. repeating steps (i) to (vi) for a large number of times.

In this study, the above sections are done by using ‘GEVcdn’ package in R. See the Appendix for the R code to apply packages of ADF test, M-K test, and GEV-CDN model.

RESULTS AND DISCUSSION

Trend and stationary tests

The SMP time series for all seasons over the basin are shown in Figure 3. Visually, the series in the upper reach for all seasons have upward trends, while those in the middle and lower reaches for only winter season (December–February) have upward trends.

Table 1 shows the results of the M-K trend test and ADF stationary test for all the 12 SMP time series. The trend test results are quite close to those of visual observations.

![Figure 3](https://example.com/fi.png)

Figure 3 | The 12 SMP time series in Heihe River basin. From top to bottom, upper, middle, and lower reaches, are represented respectively. The solid line means the polynomial regression curves.
Significant positive trends are detected for SMP time series in all seasons for the upper reach at either 0.1 or 0.05 significance level. It is probably associated with the fact that the western and eastern parts of the upper reach are located on the mountain slopes of the westerlies and southeast monsoon. Qin et al. (2010) and Cheng et al. (2015) reported that the Asian monsoon and westerly winds play an important role in precipitation changes over the basin. Therefore, under global warming, these can lead to more precipitation and extreme events occurring. For SMP time series in winter, significant positive trends are detected in the middle and lower reaches, and the series in other seasons show no remarkable trends.

According to the \( p \) values of the ADF test in Table 1, all the series with significant trends are tested to be nonstationary \((p > 0.05)\). In addition, the SMP series in summer and autumn for the middle reach are also tested to be nonstationary although no significant trends are found from the M-K method. This can be explained since nonstationary behaviors can be either trends, mutations, cycles, random walks (with or without drift), or their combinations. Further analysis was then made by using moving \( t \)-test to examine whether these two series have significant change points. Results show that, the absolute values of statistic \( T \) of both sequences, 2.27 and 2.22, are greater than the critical value of 1.64 at 0.05 significance level, meaning that both sequences are nonstationary owing to the existence of significant change points.

### Table 1 | The results of trend and stationary tests for 12 SMP time series

| Region  | Season | M-K test \( p \) | tau | ADF test \( p \) | Dickey-Fuller |
|---------|--------|----------------|-----|----------------|--------------|
| Upper   | Spring | 0.00**       | 0.32| 0.07*          | -3.40        |
|         | Summer | 0.09*        | 0.16| 0.12*          | -3.10        |
|         | Autumn | 0.06*        | 0.17| 0.05*          | -3.50        |
|         | Winter | 0.00**       | 0.39| 0.24*          | -2.80        |
| Middle  | Spring | 0.22         | 0.11| 0.01           | -4.10        |
|         | Summer | 0.59         | 0.05| 0.17*          | -3.00        |
|         | Autumn | 0.71         | 0.04| 0.17*          | -3.00        |
|         | Winter | 0.00**       | 0.38| 0.25*          | -2.80        |
| Lower   | Spring | 0.67         | -0.04| 0.01       | -4.00        |
|         | Summer | 0.69         | -0.04| 0.02       | -4.00        |
|         | Autumn | 0.38         | 0.08| 0.02          | -3.90        |
|         | Winter | 0.00**       | 0.48| 0.16*          | -3.00        |

Note: * and ** represent the significance level of 0.1 and 0.05.

### Table 2 | The AICc values for 12 SMP time series by using GEV-CDN with time as covariate

| Region | Season | GEV0 | GEV1 | GEV11 | GEV2 |
|--------|--------|------|------|-------|------|
| Upper  | Spring | -11.9| -28.9| -26.6 | -21.8|
|        | Summer | -26.5| -30.3| -28.5 | -27.8|
|        | Autumn | 1.6  | 1.3  | 2.5   | 6.1  |
|        | Winter | -20.5| -40.8| -47.5 | -50.5|
| Middle | Spring | -106 | -113 | -113  | -108 |
|        | Summer | -18.5| -17  | -14.9 | -15.1|
|        | Autumn | -35  | -33.3| -31   | -28.3|
|        | Winter | 12.3 | 2.6  | -6.5  | 1    |
| Lower  | Spring | -59.4| -57.3| -54.9 | -53.1|
|        | Summer | -22.5| -20.8| -18.4 | -17.8|
|        | Autumn | -97.5| -95.2| -94.6 | -90.2|
|        | Winter | -24.1| -30  | -61.5 | -34.9|

Note: The smallest AICc values are represented in bold font.

### Stationary and nonstationary GEV modeling

GEV-CDN is then used for stationary and nonstationary extreme value analysis for all the 12 SMP series. The performance of four candidate GEV-CDN models (GEV0, GEV1, GEV2, and GEV11) are tested by using AICc. A smaller AICc value indicates a better fitting.

The AICc values are summarized in Table 2 for each model. It can be seen that among the 12 SMP series over the study area, seven of them are suitable to be fitted by nonstationary models due to the smaller AICc, and the other five series are suitable to be fitted by stationary models. Among the seven series, four of them are presenting better performance with the GEV1 model, and the other three, all belonging to winter series, are better fitted by the more complicated GEV11 or GEV2 models. It is worth noting that the selection of optimal GEV model is largely supported by the results of M-K and ADF tests. That is, most nonstationary time series show better performance with nonstationary models, and stationary time series show better performance with stationary models. Although nonstationary models (GEV1 and GEV11) show better fitting for the stationary middle-spring SMP series, the estimated AICc values are quite close from the stationary and nonstationary models, and thus the suitable models are finally determined by the results of goodness of fit test.

Besides the variable of time \( t \), the climate indices EASMI and WPI are also taken into account as the covariate in the nonstationary GEV models. The correlations between SMP
time series and the two indices were first analyzed by using Pearson’s correlation coefficients (Moazenzadeh et al. 2018). Results are shown in Table 3. Note that, EASMI has data only from June to August since it is defined as an area-averaged seasonally (June, July, and August) dynamical normalized seasonality (DNS) at 850 hPa within the East Asian monsoon domain (10°–40N, 110°–140E) (Li & Zeng 2005, 2005), and so the correlations between EASMI and SMP series merely in summer are tested. According to Table 3, only middle-summer SMP time series is correlated with EASMI at 0.05 significance level. Five of the 12 SMP series are significantly associated with WPI. EASMI is negatively correlated with the SMP series, whereas WPI is positively correlated.

Table 4 illustrates the corresponding AICc values by using the nonstationary models with climate indices as the covariate. Compared to Table 2, only middle-spring and middle-summer SMP series are better fitted with the climate indices as the covariate, and the other series are all better fitted with the time as the covariate.

**Frequency analysis of SMP series**

Return levels are usually used to describe and quantify the risk and thus of great importance in design engineering and risk management for water resources. Under an assumption of stationarity, the return level is the same for all years. While in the nonstationary case, the return level will vary over time with the probability of occurrence remaining constant (Chen & Chu 2014). Herein, the 2-year, 20-year, and 100-year return levels corresponding to 0.5, 0.95, and 0.99 quantiles, respectively, for those nonstationary SMP series are mainly focused on. Figure 4 illustrates the return levels estimated from the best nonstationary models with the time as covariate for both the current period of 1960–2014 and the future period of 2015–2060. The curves represent the return levels of 2-year, 20-year, and 100-year, respectively. The historical observations are presented in the figure in circles. The majority of the observations can be described by the 0.5 and 0.95 quantile curves, indicating that nonstationary models with time as covariate are able to describe the change properties of different components of precipitation extremes.

For comparison, the return levels estimated from the stationary models are also displayed in Figure 4. As expected, the stationary models produce stable return levels without any changes with time, while the nonstationary ones produce the changing return levels varying linearly or nonlinearly with time. For example, the 100-year return level of upper-spring SMP series is 22.3 mm from the stationary GEV0 model, while it varies from 19.9 mm in 1960 to 25.7 mm in 2014 from the nonstationary GEV1 model. The 20-year return level is 16.8 mm from the GEV0 model, and varies from 13.4 mm in 1960 to 19.1 mm in 2014 from the GEV1 model; in other words, an event with a 100-year recurrence-interval threshold value in 1960 has occurred, on average, once every 20–25 years by 2014. By 2060, it will happen every 5–10 years. This indicates that extreme precipitation events with the same magnitude

---

**Table 3** Pearson’s correlation coefficients between SMP time series and the climate indices

| Region   | Season | EASMI | WPI |
|----------|--------|-------|-----|
|          |        |       |     |
| Upper    | Spring | –     | –   | –   |
|          | Summer | 0.09  | –0.23| 0.51| –0.09|
|          | Autumn | –     | –   | 0.33| 0.13|
|          | Winter | –     | –   | 0.05*| 0.26|
| Middle   | Spring | –     | –   | 0.01*| 0.33|
|          | Summer | 0.01*| –0.32| 0.48| 0.09|
|          | Autumn | –     | –   | 0.08| 0.24|
|          | Winter | –     | –   | 0.009*| 0.33|
| Lower    | Spring | –     | –   | 0.09| 0.23|
|          | Summer | 0.74  | 0.04| 0.62| –0.07|
|          | Autumn | –     | –   | 0.63| –0.06|
|          | Winter | –     | –   | 0.002*| 0.41|

Note: Bold font indicates a significant correlation between SMP and the climate index.

**Table 4** The AICc values for SMP time series by using GEV-CDN with climate indices as covariate

| SMP       | Covariate | GEV1  | GEV11 | GEV2  |
|-----------|-----------|-------|-------|-------|
| Upper-spring| WPI       | –12.89| –12.03| –8.06|
| Upper-winter| WPI       | –24.83| –23.62| –25.25|
| Middle-spring| WPI      | –115.76| –114.91| –110.82|
| Middle-summer| EASMI    | –27.29| –24.97| –23.6|
| Middle-winter| WPI      | 11.61 | 9.96  | 12.45 |
| Lower-winter| WPI       | –24.75| –33   | –28.2|

Note: The smallest AICc values are represented in bold font.
occur more frequently in the future. From Figure 4, it is also clear that three return levels from the nonstationary models vary linearly with time, which are related with the forms of GEV1 model. This model is linear time-dependent of location parameter, leading to the associated return level estimates showing linear trends with time. GEV11 and GEV2 models are more complicated, with linear time dependence of location parameter and log-linear time dependence of scale parameter for GEV11, and quadratic time dependence of location parameter for GEV2, resulting in the associated return level estimates for the other three SMP series varying nonlinearly with time. It is observed that the return level estimates from GEV11 models show small divergences for lower levels (e.g., at two-year), while large divergence for higher levels (e.g., at 20-year or higher), indicating that the effect of the scale parameter on the high return period is greater than the low return period (Chen & Chu 2014). It is also observed that for the upper-winter SMP series, the return level estimates increase sharply around 1980, which is associated with its best fitted GEV2 model with quadratic time dependence of location for capturing the stepwise pattern of those extreme data which occurred at that period.

Since only the middle-spring and middle-summer SMP series are better fitted by the nonstationary GEV1 model with climate variables as covariates, Figure 5 just shows the relationship between the return level estimates of these two SMP series and WPI and EASMI. As a result, the return level estimates increase with WPI values increasing and EASMI values decreasing. Previous studies showed that the East Asian summer monsoon significantly weakened after 1978 over the study area, and resulted in some
changes in the amount, frequency, and intensity of precipitation in Heihe River basin (Ding et al. 2008, 2009; Cheng et al. 2015; Chen et al. 2017).

Confidence intervals of frequency analysis

The CIs of return level estimates are also examined. They are usually important and required for hydrological and climatological decision-makers from an operational point of view, since the CI estimates are more informative than the point estimates. The concept of CIs is usually used to indicate the reliability of the estimates. Bootstrap technique, mentioned in the section ‘Confidence interval estimation’, is used to compute the CIs of return levels.

Figure 6 illustrates the two-year return level for the nonstationary SMP series estimated from their respective best nonstationary models, along with their 95% CIs obtained by the bootstrap technique (1,000 times). Those estimations from the stationary models and the corresponding 95% CIs are also displayed in the same figure. Figures 7 and 8 present the results for the 20-year and 100-year return levels, respectively. It is clear that the width of CIs increases as the predicted return levels increase no matter whether stationary or nonstationary model. This implies that the higher the estimated return levels, the greater the uncertainty of its estimation. On the other hand, the uncertainty of future return level predictions increases as time goes on for the nonstationary models.

As also shown, comparatively, GEV1 (embedding a linear increasing trend) yields the width of 95% CIs without significant large differences between stationary and nonstationary models, especially for 20-year and 100-year return periods, meaning that both models can provide the approaching reliability of the estimated return levels. By considering the fact that the nonstationary GEV1 is capable of better describing the variations of precipitation extremes, it is regarded as a satisfactory model for the SMP fitting in this study within the nonstationary framework. With respect to the more complex GEV11 model (embedding a nonlinear increasing trend), the corresponding 95% CIs become larger over time, much wider than those from the stationary models as time increases, indicating that this nonstationary model produces less reliable predicted return levels, particularly with time passing. For instance, the GEV11 model predicts a 14.8 mm middle-winter SMP at 20-year return levels in 2060, four times the maximum value recorded during the past 50 years, and the width of 95% CIs ranges approximately from 6.9 mm to 30.8 mm. For precipitation extremes at higher return levels, such change presents even more dramatically and the 95% CIs even wider. Without questioning the physical meaning of these values, the huge widths of the CIs indicate that this model is not very informative and lacks reliability and practical applicability. Thus, the most complex GEV11 model is not recommended here due to the largest CIs width and the poor reliability it delivers. It further implies that the increase of model complexity is likely to result in the increase of uncertainty and the decrease of reliability in estimates. Thus, from a practical point of view, if no more information is available, it seems unnecessary to switch from a simple but clear framework to something that is more complex and does not guarantee better future predictions.
CONCLUSIONS

In this study, the frequency analysis of seasonal maximum 1-day precipitation (SMP) in the Heihe River basin is conducted by using the flexible, efficient, and robust GEV-CDN statistical model. Furthermore, the confidence interval estimates of return levels derived from the residual bootstrap method are also considered.

Results show that the upper reaches of the Heihe River basin are mountainous and more sensitive to climate change, and most of the SMP series show increasing trends at a significant level of 0.05 or 0.1 based on the Mann–Kendall test. These findings are consistent with some previous studies (e.g., Zhang et al. 2011; Wang et al. 2013).

The trend test results support the optimal model selection from the GEV-CDN. The nonstationary model performs better than the stationary model for the SMP time series with significant trend. For those SMP series without significant trend but correlated to the climate index, either WPI or EASMI, the nonstationary model with the climate index as covariate can effectively improve the quality of fitting according to the goodness of fit test.

The best nonstationary GEV model is used to estimate the 2-year, 20-year and 100-year return levels of precipitation extremes. As a result, precipitation extremes increase in frequency over time linearly or nonlinearly, and this increase will continue in the future, which means that the occurrence of precipitation extremes in the study area will increase in the future.

Figure 6 | Two-year return level estimates and the corresponding 95% CIs from the stationary and nonstationary GEV (with time as covariate) models.
The bootstrap approach demonstrates that the higher the return period, the greater the uncertainty of return level estimates. The uncertainties derived from the nonstationary GEV1 and GEV2 (single-parameter changes) are close to those from the stationary model in a short period, while becoming larger with time going on. The more complex GEV11 model (both location and scale parameters depend on time) presents the highest uncertainty due to more parameters involved in variations. The greater the uncertainty of return level estimates, the lower the reliability is. Thus, it is not necessary to switch from a simple but clear framework to something that is more complex while producing large uncertainties in the future predictions. In addition, when estimating return levels with a nonstationary model, the estimation results from the stationary model should be referred to as well to avoid causing large deviations.

This study developed nonstationary GEV models for fitting SMP with nonstationary characteristics in the study area by introducing time $t$ or climate indices of WPI and EASMI as covariates. Cannon (2010) and Vasiliades et al. (2015) considered that a single covariate may be not enough to explain the changes of precipitation. Thus, more covariates together with their interactions need to be considered in developing nonstationary models in future studies. In addition, using just one index of maximum daily precipitation to indicate the precipitation extremes in this study also has a certain limitation since precipitation extremes can be characterized in many different ways, not only the intensity of heavy precipitation, but also its duration and frequency, the accumulated amount of precipitation for

Figure 7 | Twenty-year return level estimates and the corresponding 95% CIs from the stationary and nonstationary GEV (with time as covariate) models.
a given period, etc. Indices such as maximum consecutive 5-day precipitation and number of heavy precipitation days (Cheng et al. 2015) can be selected for reference to discuss the characteristics of precipitation extremes in future research.

### ACKNOWLEDGEMENTS

This study is supported by the Fundamental Research Funds for the Central Universities (No. 35832015028) and NSFC (No. 411101038). The authors are grateful for the financial support by the China Scholarship Council and China University of Geosciences (Beijing). The authors also thank the anonymous referees for their comments and suggestions that have led to the quality of the paper being improved.

### SUPPLEMENTARY MATERIAL

The Supplementary Material for this paper is available online at https://dx.doi.org/10.2166/wcc.2020.170.

### REFERENCES

AghaKouchak, A. & Nasrollahi, N. 2010 Semi-parametric and parametric inference of extreme value models for rainfall data. Water Resour. Manage. 24, 1229–1249.
Agilan, V. & Umamahesh, N. V. 2017 Modelling nonlinear trend for developing nonstationary rainfall intensity-duration-frequency curve. *Int. Climatol.* 37, 1265–1281.

Cannon, A. J. 2020 A flexible nonlinear modelling framework for nonstationary generalized extreme value analysis in hydroclimatology. *Hydrol. Processes* 24, 673–685.

Cannon, A. J. 2021 GEVcdn: an R package for nonstationary extreme value analysis by generalized extreme value conditional density estimation network. *Comput. Geosci.* 37, 1532–1533.

Cannon, A. J. 2015 Revisiting the nonlinear relationship between ENSO and winter extreme station precipitation in North America. *Int. J. Climatol.* 35, 4001–4014.

Chang, Y. & Park, J. Y. 2002 On the asymptotics of adf tests for unit roots. *Econom. Rev.* 21 (4), 431–447.

Chen, Y. R. & Chu, P.-S. 2014 Trends in precipitation extremes and return levels in the Hawaiian Islands under a changing climate. *Int. J. Climatol.* 34, 3913–3925. doi:10.1002/joc.3950.

Chen, Y. N., Wang, H. J., Wang, Z. C. & Zhang, H. 2017 Characteristics of extreme climatic/hydrological events in the arid region of northwestern China. *Arid Land Geogr.* 40 (01), 1–9 (in Chinese).

Cheng, L. Y., AghaKouchak, A., Gilleland, E. & Katz, R. W. 2014 Non-stationary extreme value analysis in a changing climate. *Clim. Change* 127, 353–369.

Cheng, A. F., Feng, Q., Fu, G. B., Zhang, J. K., Li, Z. X., Hu, M. & Wang, G. 2015 Recent changes in precipitation extremes in the Heihe River basin, Northwest China. *Adv. Atmos. Sci.* 32 (10), 1391–1406. doi:10.1007/s00376-015-4199-3.

Coles, S. 2001 *An Introduction to Statistical Modelling of Extreme Values.* Springer, London, UK.

Cooley, D. 2013 Return periods and return levels under climate change. *Extremes Changing Clim.* 65, 97–114.

DeGaetano, A. T. 2009 Time-dependent changes in extreme-precipitation return-period amounts in the Continental United States. *J. Appl. Meteorol. Climatol.* 48 (10), 2086–2099. doi:10.1175/2009JAMC2179.

Derble, E. K. & Kasei, R. A. 2012 Vulnerability of crop production to heavy precipitation in North-Eastern Ghana. *Int. J. Clim. Change Strategies Manage.* 4, 36–53.

Ding, Y. H., Wang, Z. Y. & Sun, Y. 2008 Inter-decadal variation of the summer precipitation in East China and its association with decreasing Asian summer monsoon. Part I: observed evidences. *Int. J. Climatol.* 28, 1139–1161. doi:10.1002/joc.1615.

Ding, Y. H., Sun, Y., Wang, Z. Y., Zhu, Y. X. & Song, Y. F. 2009 Inter-decadal variation of the summer precipitation in China and its association with decreasing Asian summer monsoon part II: possible causes. *Int. J. Climatol.* 29, 1926–1944. doi:10.1002/joc.1759.

El Adlouni, S., Ouarda, T. B. M. J., Zhang, X., Roy, R. & Bobée, B. 2007 Generalized maximum likelihood estimators for the nonstationary generalized extreme value model. *Water Resour. Res.* 43, W03410. doi:10.1029/2005WR004545.

Fischer, T., Su, B. D., Luo, Y. & Scholten, T. 2012 Probability distribution of precipitation extremes for weather-index based insurance in the Zhujiang River Basin, South China. *J. Hydrol. Meteorol.* 13, 1023–1037.

Fotovatikhab, F., Herrera, M., Shamshirband, S., Chau, K., Ardabili, S. F. & Piran, M. 2018 Survey of computational intelligence as basis to big flow management: challenges, research directions and future work. *Eng. Appl. Comput. Fluid Mech.* 12 (1), 411–437.

Gallatsatou, G. & Prinos, P. 2011 Modeling non-stationary extreme waves using a point process and wavelets. *Stochastic Environ. Risk Assess.* 25 (2), 165–183.

Gao, M., Mo, D. Y. & Wu, X. Q. 2016 Nonstationary modeling of extreme precipitation in China. *Atmos. Res.* 182, 1–9.

Gu, X. H., Zhang, Q., Li, J. F., Singh, V. P. & Sun, P. 2019 Impact of urbanization on nonstationarity of annual and seasonal precipitation extremes in China. *J. Hydrol.* 575, 638–655.

Hamdi, Y., Duluc, C. M. & Rebour, V. 2018 Temperature extremes estimation of non-stationary return levels and associated uncertainties. *Atmosphere* 9 (4), 129. https://doi.org/10.3390/atmos9040129.

Hao, W. L., Shao, Q. X., Hao, Z. C., Ju, Q., Baima, W. D. & Zhang, D. 2019 Non-stationary modelling of extreme precipitation by climate indices during rainy season in Hanjiang River Basin, China. *Int. J. Climatol.* 39, 4154–4169.

Hurvich, C. M. & Tsai, C. 1989 Regression and time series model selection in small samples. *Biometrika* 76, 297–307.

IPCC 2013 *Climate Change 2013: The Physical Science Basis.* Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, Cambridge University Press, Cambridge, UK, p. 1535.

Katz, R. W. 2015 Statistical methods for nonstationary extremes. In *Extremes in a Changing Climate.* Springer, Dordrecht, pp. 15–27.

Katz, R. W., Parlang, M. B. & Naveau, P. 2002 Statistics of extremes in hydrology. *Adv. Water Resour.* 25, 1287–1304.

Kendall, M. G. 1958 A new measure of rank correlation. *Biometrika* 50, 81–93.

Khaliq, M., Ouarda, T., Ondo, J.-C., Gachon, P. & Bobée, B. 2006 Frequency analysis of a sequence of dependent and/or nonstationary hydrometeorological observations: a review. *J. Hydrol.* 329, 534–552.

Kharin, V. V. & Zwiers, F. W. 2004 Estimating extremes in transient climate change simulations. *J. Clim.* 18, 1156–1173.

Kim, H., Kim, S., Shin, H. & Heo, J.-H. 2017 Appropriate model selection methods for nonstationary generalized extreme value models. *J. Hydrol.* 547, 557–574.

Li, J. P. & Zeng, Q. C. 2005 A new monsoon index and the geographical distribution of the global monsoons. *Adv. Atmos. Sci.* 20, 299–302.

Li, J. P. & Zeng, Q. C. 2005 A new monsoon index, its interannual variability and relation with monsoon precipitation. *Clim. Environ. Res.* 10 (3), 351–365.

Li, W., Liu, Y. & Yang, J. 2017 Time series stationarity test and empirical analysis based on R language – taking Chinese macroeconomic variables as research objects. *China Bus. Trade* 16, 150–153 (in Chinese).
Li, X., Wang, X. & Babovic, V. 2008 Analysis of variability and trends of precipitation extremes in Singapore during 1980–2013. *Int. J. Climatol.* 38, 125–141.

Libiseller, C. & Grimvall, A. 2002 Performance of partial Mann-Kendall tests for trend detection in the presence of covariates. *Environmetrics* 13, 71–84.

Lu, F., Xiao, W. H., Yan, D. H. & Wang, H. 2017 Progresses on statistical modeling of non-stationary extreme sequences and its application in climate and hydrological change. *SHUILI XUEBAO* 48 (4), 379–389. 0559-9350(2017)04-0379-11 (in Chinese).

Mann, H. B. 1945 Nonparametric test against trend. *Econometrica* 13, 243–259.

Martins, E. S. & Stedinger, J. R. 2000 Generalized maximum-likelihood generalized extreme-value quantile estimators for hydrologic data. *Water Resour. Res.* 36, 737–744.

Milly, P. C. D., Betancourt, J., Falkenmark, M., Hirsch, R. M., Kundzewicz, Z. W., Lettenmaier, D. P. & Stouffer, R. J. 2008 Stationarity is dead: whither water management? *Science* 319 (5863), 573–574.

Moazenzadeh, R., Mohammadi, B., Shamshirband, S. & Chau, K. 2018 Coupling a firefly algorithm with support vector regression to predict evaporation in northern Iran. *Eng. Appl. Comput. Fluid Mech.* 12 (1), 584–597.

Obeysekera, J. & Salas, J. D. 2014 Quantifying the uncertainty of design floods under nonstationary conditions. *J. Hydrol. Eng.* 19 (7), 1438–1446.

Peterson, T. C., Zhang, X., Brunet-India, M. & Vazquez-Aguirre, J. L. 2008 Changes in North American extremes derived from daily weather data. *J. Geophys. Res.* 113, D07113. doi:10.1029/2007JD009433.

Qin, C., Yang, B., Burchardt, I., Hu, X. L. & Kang, X. C. 2010 Intensified pluvial conditions during the twentieth century in the inland Heihe River basin in arid northwestern China over the past millennium. *Global Planet. Change* 72, 192–200. doi:10.1016/j.gloplacha.2010.04.005.

Said, S. E. & Dickey, D. A. 1984 Testing for unit roots in autoregressive-moving average models of unknown order. *Biometrika* 71, 599–607.

Salas, J. D. & Obeysekera, J. 2014 Revisiting the concepts of return period and risk for nonstationary hydrologic extreme events. *J. Hydrol. Eng.* 19, 554–568.

Serinaldi, F. & Kilsby, C. G. 2005 Stationarity is undead: Uncertainty dominates the distribution of extremes. *Adv. Water Resour.* 77, 17–36.

Shinyie, W. L., Ismail, N. & Jemain, A. A. 2013 Semi-parametric estimation for selecting optimal threshold of extreme rainfall events. *Water Resour. Manage.* 27 (7), 2325–2352.

Soukissian, T. H. & Tsalis, C. 2013 The effect of the generalized extreme value distribution parameter estimation methods in extreme wind speed prediction. *Nat. Hazard.* 78, 1777–1809.

Sraj, M., Viglione, A., Parajika, J. & Blöschl, G. 2016 The influence of non-stationarity in extreme hydrological events on flood frequency estimation. *J. Hydrol. Hydromech.* 64, 426–437. doi:10.1515/johh-2016-0032.

Um, M.-J., Kim, Y., Markus, M. & Wuebbles, D. J. 2017 Modeling nonstationary extreme value distributions with nonlinear functions: an application using multiple precipitation projections for U.S. cities. *J. Hydrol.* 552, 396–406.

Vasilias, L., Galiatsatou, P. & Loukas, A. 2015 Nonstationary frequency analysis of annual maximum rainfall using climate covariates. *Water Resour. Manage.* 29, 339–358.

Villarini, G., Smith, J. A. & Napolitano, F. 2010 Nonstationary modelling of a long record of rainfall and temperature over Rome. *Adv. Water Resour.* 33, 1256–1267.

Wang, H. J., Chen, Y. N. & Chen, Z. S. 2012 Spatial distribution and temporal trends of mean precipitation and extremes in the arid region, northwest of China, during 1960–2010. *Hydrol. Process.* 27, 1807–1818. doi:10.1002/hyp.9339.

Wang, H. J., Chen, Y. N., Xun, S., Lai, D. M., Fan, Y. T. & Li, Z. 2013 Changes in daily climate extremes in the arid area of northwestern China. *Theor. Appl. Climatol.* 112, 15–28. doi:10.1007/s00704-012-0697-8.

Wasko, C., Sharma, A. & Westra, S. 2016 Reduced spatial extent of extreme storms at higher temperatures. *Geophys. Res. Lett.* 43 (8), 4026–4032. http://dx.doi.org/10.1002/2016GL068509.

Wobus, C., Lawson, M., Jones, R., Smith, J. & Martinich, J. 2014 Estimating monetary damages from flooding in the United States under a changing climate. *J. Flood Risk Manage.* 7, 217–229.

Yang, D.W., Wang, Y. H., Gao, B. & Qin, Y. 2015 Spatial interpolation of gauge precipitation using regional climate model simulation in the Heihe basin (1960–2014). *Cold and Arid Reg. Sci. Data Cent. Lanzhou*. doi:10.3972/heihe.127.2014.db.

You, Q., Kang, S., Aguilar, E., Pepin, N., Flugel, W.-A., Yan, Y., Xu, Y., Zhang, Y. & Huang, J. 2011 Changes in daily climate extremes in China and their connection to the large scale atmospheric circulation during 1961–2003. *Clim. Dyn.* 36 (11), 2399–2417.

Zhang, Q., Singh, V. P., Li, J. F. & Chen, X. H. 2011 Analysis of the periods of maximum consecutive wet days in China. *Geophys. Res. Lett.* 116, 1–18. doi:10.1029/2011JD016088.

Zhang, S. H., Ye, Z. X., Chen, Y. N. & Xu, Y. F. 2017 Vegetation responses to an ecological water conveyance project in the lower reaches of the Heihe River Basin. *Ecology Hydrology* 10 (6), e1866.

Zolina, O., Simmer, C., Kapala, A., Bachner, S., Gulev, S. & Maechel, H. 2008 Seasonally dependent changes of precipitation extremes over Germany since 1950 from a very dense observational network. *J. Geophys. Res. Atmos.* 113, D06110. doi:10.1029/2007JD008393.