Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-619, 2016
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Published: 9 December 2016

Hydrology and Earth System Sciences Discussions
Open Access

Comparative study of flood projections under the climate scenarios: links with sampling schemes, probability distribution models, and return level concepts

Lingqi Li[1], Lihua Xiong[1,*], Chong-Yu Xu[1,2], Shenglian Guo[1], Pan Liu[1]

[1]State Key Laboratory of Water Resources and Hydropower Engineering Science, Wuhan University, Wuhan 430072, China
[2]Department of Geosciences, University of Oslo, PO Box 1047 Blindern, N-0316 Oslo, Norway

E-mail addresses:
L. Li (lqli@whu.edu.cn),
L. Xiong (xionglh@whu.edu.cn),
C.-Y. Xu (c.y.xu@geo.uio.no),
S. Guo (slguo@whu.edu.cn),
P. Liu (liupan@whu.edu.cn)

Correspondence to:

Lihua Xiong, PhD, Professor
State Key Laboratory of Water Resources and Hydropower Engineering Science
Wuhan University
Wuhan 430072, China
E-mail: xionglh@whu.edu.cn
Telephone: +86-13871078660
Fax: +86-27-68773568

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-619, 2016
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Published: 9 December 2016

Hydrology and Earth System Sciences
Discussions

Abstract

Traditional stationarity strategy for extrapolating future design floods requires renovation in response to the possible nonstationarity caused by changing climate. Capable of tackling such problem, the expected-number-of-events (ENE) method is employed with both Annual Maximum (AM) and Peaks over Threshold (POT) sampling schemes expatiated. The existing paradigms of the ENE method are extended focusing on the over-dispersion emerged in POT arrival rate, for which by virtue of the ability to account, the Negative Binomial (NB) distribution is proposed as an alternative since the common assumption of homogeneous Poisson process would likely be invalid under nonstationarity. Flood return levels are estimated and compared under future climate scenarios (embodied by the two covariates of precipitation and air temperature) using the ENE method for both sampling schemes in the Weihe basin, China. To further understand how flood estimation responds to climate change, a global sensitivity analysis is performed. It is found that design floods dependent on nonstationarity are usually but not necessarily more different from those analyzed by stationarity strategy due to the interaction between air temperature and precipitation. In general, a large decrease in flood projection could be induced under nonstationarity if air temperature presents dramatically increasing trend or reduction occurs in precipitation, and vice versa. AM-based flood projections are mostly smaller than POT estimations (unless a low threshold is assumed) and more sensitive to changing climate. The outcome of the biased flood estimates resulting from an unrestricted use of the Poisson assumption suggests a priority to the NB

distribution when fitting POT arrival rate with significantly larger variance than the mean. The study supplements the knowledge of future design floods under changing climate and makes an effort to improve guidance of choices in flood inference.

Keywords: Nonstationarity; Flood return level; Peaks over Threshold; Annual Maximum; Negative Binomial distribution; Climate change

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-619, 2016
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Published: 9 December 2016

1. Introduction

Flood frequency analysis, one of the most widely used tools in hydrology, is of great significance for theoretical research and practical application in flood projection and risk management. Reliable flood return level estimation requires careful considerations basically from three aspects, i.e., sampling schemes, probability distribution models, and return level concept.

Primarily, two kinds of sampling schemes are used in common for the flood-related studies (Coles, 2001), i.e., the Annual Maximum (AM) (block defined as year scale in block maxima sampling) and Peaks over Threshold (POT) (also known as partial duration series). The AM sampling, extracting the annual maximum peak flows from the observed discharge series, is simpler than the POT sampling that collects the discharges above a fixed high threshold. Hence, the AM realizes a wider use in hydrology than the POT, but losing 'real flood' information is inevitable because small discharge included in a dry year could be misleading (Lang et al., 1999). The POT, free from the sampling restriction of the AM that picks only one event per year, seems to be rational as it substantially contains two flood characteristics to be portrayed separately: the magnitude and the arrival rate (annual number of exceedances above the threshold) (Önöz and Bayazit, 2001).

No matter for AM or POT floods, flood frequency analysis has undoubtedly, for a long time, indulged in such a prevailing approach that flood events, subject to the underlying assumptions of being independent and identically distributed (*i.i.d.*), share the same probability distribution. This description can be epitomized as the stationarity strategy used in traditional flood frequency analysis (Coles, 2001). Since the impacts of climate change on hydrological system have been reported repeatedly (IPCC, 2013), nonstationarity, as a special concept in contrast to stationarity, has generally enjoyed popular supports in academia. A number of researchers have been absorbed in, for instance, revealing the invalidation of stationarity strategy (Khaliq et al., 2006; Milly et al., 2008), describing the temporal variability of hydrological characteristics (Villarini et al., 2009a; Machado et al., 2015; Xiong et al., 2015a), and exploring the reasons behind the changes (Ishak et al., 2013; López and Francés, 2013; Jiang et al., 2015; Xiong et al., 2015b).

Important as it is, questions around "stationarity is still alive or wanted dead" (Lins and Cohn, 2011; Koutsoyiannis, 2011) have been subsequently pointed out sharply, remaining more or less as a controversial puzzle. In an attempt to clarify these issues, so far there have appeared various arguments. For example, Lins and Cohn (2011) admitted the existence of nonstationarity but simultaneously suggested the use of stationarity to elude the potentially high uncertainty of nonstationary influences on hydrologic studies. Montanari and Koutsoyiannis (2014) asserted that stationarity is immortal for the need of mitigating natural hazards. Koutsoyiannis and Montanari (2015) stated, persuasively, that the misunderstanding of stationarity has let "changes" be mistakenly labeled as "nonstationarity." Likewise, Serinaldi and Kilsby (2015) deliberately titled their main topic with "stationarity is undead" to alert of the uncertainty related to nonstationary flood frequency analysis. In response to the

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-619, 2016
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Published: 9 December 2016

thoughtful literatures with the opposing opinions mentioned above, Milly et al. (2015) reiterated the viewpoints of Milly et al. (2008) who claimed that “stationarity is dead” by using “Policy Forum” to communicate the necessity of considering nonstationarity in hydrology in the 21st century. Stedinger and Griffis (2011) explained conservatively that formulating nonstationary models with finite flood records can be defensible when physical-causal basis for multi-decadal projections is known. Indeed, stationarity, as the solid cornerstone laid for hydrologic frequency analysis, does deserve to be active (Koutsoyiannis, 2011). Nevertheless, there is reason to afford an opportunity to nonstationarity for advancing hydrologic research (Milly et al., 2015). Advocating nonstationarity at present is intended to arouse the consciousness in the scientific community due to the on-going climate changes yet without smothering stationarity.

In the presence of nonstationarity, a good few of studies (also this paper) materialize nonstationary hydrologic variables with resorting to the time-variant characters of the variable moments (Khaliq et al., 2006), i.e., nonstationary flood-frequency distribution model is constructed by the theoretical probability distribution whose statistical parameters are assumed to be no longer fixed over time (Stedinger and Griffis, 2011; Milly et al., 2015). To addressing the causes of nonstationarity, researchers attribute the changes, qualitatively, by nonparametric cross-correlation analyses (Ishak et al., 2013), and quantitatively, by linking the time-varying distribution parameters to the exploratory variables, e.g., time and potential driving forces (Prosdocimi et al., 2015; Serinaldi and Kilsby, 2015; Silva et al., 2015; Xiong et al., 2015b), among many others.

Analyses for flood frequency or return level have been accomplished with the AM, POT, or both worldwide with either stationary or nonstationary hypotheses (e.g., Villarini et al., 2012; López and Francés, 2013; Machado et al., 2015; Xiong et al., 2015a). However, attentions paid to the comparison of AM and POT flood series in flood frequency analysis are relatively limited in the previous research endeavors. Rosbjerg (1985) whose research was completed on a stationary background deemed that the POT series modeled with heavy-tailed distributions should yield more advisable flood estimates than AM. Madsen et al. (1997) suggested that the POT series was generally preferable to AM series for at-site flood estimation under stationarity. More recently, Bezak et al. (2014) found that the POT series gave higher flood estimates than the AM series for larger return periods on stationary conditions. Prosdocimi et al. (2015) concluded that POT models outperformed AM models in respect of detecting the external causes of nonstationary floods.

It is worth noting that nonstationarity in the flood series caused by the changing environments has made stationarity strategy for return level estimation problematic (Khaliq et al., 2006; Sivapalan and Samuel, 2009; Villarini et al., 2009b; López and Francés, 2013). There have been numerous studies on return level inference associated with hydro-climatic extreme events that consider nonstationary conditions. For example, return level was proposed, with the corresponding return period as the expected waiting time until an exceedance occurs (Olsen et al., 1998; Wigley, 2009; Salas and Obeysekera, 2014), or as the quantile over which the expected number of

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-619, 2016
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Published: 9 December 2016

Hydrology and Earth System Sciences
Discussions
Open Access
EGU

events (ENE) during a given return period is one (Parey et al., 2007, 2010). Comparative analyses on such two methods were performed in Cooley (2013), Du et al. (2015), etc. Besides, risk-oriented approaches for deriving flood estimators considering nonstationarity have also been devised in some literatures with diversities in their scope (Sivapalan and Samuel, 2009).

These profound studies, with considerable efforts made on the extrapolation of hydro-climatic extremes under nonstationarity, have presented capacity and depth in theory, and most focus on the block (e.g., AM) sampling. Other sampling, i.e., the POT, seems not to receive much attention in estimating return levels by the method adapted to the context of nonstationarity in the literatures except Parey et al. (2010) who set an example with the application of the ENE method to the POT case yet without much more discussions on mathematical treatment, and Silva et al. (2015) who estimated the flood hazards based on the POT framework by making the engineering design life period equal to the past observation periods. Additionally, exploration on future design floods in nonstationarity context is still limited as well as the analyses on how climate change could influence flood projections.

This paper is aimed to achieve multi-decadal flood projections under the future climate scenarios and investigate the effect of climate changes on design floods. Essentially, the study can serve as a complement of the available ENE method from the following aspects. First, design floods are estimated with two sampling schemes of AM and POT and compared on not only stationary but also nonstationary conditions. Second, the ENE method is extended for the POT sampling with an

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-619, 2016
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Published: 9 December 2016

emphasis on describing the POT arrival rates. The POT arrival rates have in fact been chronically accepted on faith to follow a homogeneous Poisson process under stationarity (Shane and Lynn, 1964). However, such assumption has been reported to be invalid due to two-type sources of nonstationarity which will be addressed herein: (i) heterogeneity of Poisson process intensity (Cunnane, 1979; Villarini et al., 2012; Silva et al., 2015), for which the Poisson distribution is retained no longer with invariant Poisson process intensity, or rather, parameterized as functions of climatic covariates; (ii) over-dispersion of observations. Theoretically, the Poisson distribution holds identical variance and mean of population, whereas it is often the case that the variance is rarely equal to, and even significantly higher than, the mean (Cunnane, 1979). Therefore, the Negative Binomial (NB) distribution is recruited as an alternative to the Poisson distribution following the findings from Ben-Zvi (1991) and Önöz and Bayazit (2001). Finally, the sensitivity of flood estimations to changing climate is analyzed for reference to future inference.

2. Methodology

Analysis of flood return levels is undertaken briefly following: preliminary diagnosis for nonstationarity evidence, modeling of both AM and POT samplings under stationarity and nonstationarity, respectively, (i.e., using the assumed probability distributions with parameters as functions of constant or climatic covariates), extrapolation of flood by applying the ENE method to these models, and investigation on how climatic effect affects flood estimations.

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-619, 2016
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Published: 9 December 2016

2.1. Diagnostics for nonstationarity

Justifying the presence of nonstationarity is of great importance for the investigation of hydro-climatic events in a changing world (Montanari and Koutsoyiannis, 2014; Serinaldi and Kilsby, 2015; Milly et al., 2015; Xiong et al., 2015b). Importance attached to the gradual evolution of observation time series, is emphasized, for which the preliminary detection is implemented by three nonparametric trend tests: the Mann-Kendall (MK) (Mann, 1945; Kendall, 1975), the pre-whitening (PW) (von Storch, 1995), and the trend-free pre-whitening (TFPW) (Yue et al., 2002). The latter two tests are proposed initially to mitigate the adverse influence of lag-1 serial correlation r_1 on the robustness of the MK method. Instead of testing the MK statistics $Z_{MK}(\cdot)$ of the original observation series $\{X_t, t=1,2,...,N\}$, they use the new independent series of $X_t' = X_t - r_1 X_{t-1}$ and Y_t'' from Eq. (1), respectively.

$$\begin{aligned}
&S = \underset{\forall t_1<t_2}{\mathrm{median}}\left(\frac{X_{t_2} - X_{t_1}}{t_2 - t_1}\right) \\
&Y_t = X_t - St \\
&Y_t' = Y_t - r_1 Y_{t-1} \\
&Y_t'' = Y_t' + St
\end{aligned} \tag{1}$$

where S is the Sen's slope (Sen, 1968). The partial MK test (Libiseller and Grimvall, 2002) is then employed to identify the attribution of the detected significant trend via the statistics as

$$Z_{PMK} = \frac{Z_{MK}(X) - \hat{\rho} Z_{MK}(E)}{\sqrt{(1-\hat{\rho}^2)N(N-1)(2N+5)/18}} \tag{2}$$

where $\hat{\rho}$ denotes the correlation coefficient between $Z_{MK}(X)$ of dependent variable X and $Z_{MK}(E)$ of a physical covariate E. This test can be thought of testing the significance of trend in the modified dependent variable after removing the linear dependence on a covariate. It is inferred that dependent variable may co-vary with the physical covariate if the *p*-value of Z_{PMK} becomes larger than the given significance level (0.05). The closer the *p*-value is to one, the greater the extent to which the dependent variable relates to the physical covariate.

To verify the conjecture if the homogeneous Poisson process assumption is valid under changing circumstances, the Bohning (1994) test is applied to the observed series of POT arrival rates for testing against the alternative hypothesis that the variance of population S^2 is larger than the mean $\bar{X}$. The test statistic $\sqrt{\left(\frac{n-1}{2}\right)\left(\frac{S^2}{\bar{X}}-1\right)}$ asymptotically converges to the normal distribution for a large population. Given the finite sample size, a bootstrap simulation is performed to generate randomly 10000 replications from original series and for each replication calculate the statistic values. According to the given significance level (0.05), the Poisson assumption would be rejected if the *p*-value of the attained empirical distribution for test statistic is less than 0.05.

2.2. Probability distribution modeling

Modelling of flood series was undertaken for recruiting the theoretical probability distribution as potential candidates. In this paper, the distribution to be considered is selected based on the successful applications in previous studies (e.g., Madsen et al., 1997; Lang et al., 1999; Du et al., 2015) but for the purpose at current stage not including all of them. The AM floods are assumed to follow three different types of

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-619, 2016
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Published: 9 December 2016

probability distributions of the lognormal 3 (LNO3), Log-Pearson type 3 (LP3), and Generalized Extreme Value (GEV) (Jenkinson, 1955). They all include three parameters $\boldsymbol{\theta}_t(\mu_t, \sigma_t, \xi_t)$: the location parameter μ_t associated with the magnitude of the series; the scale parameter σ_t related to the variability of the series; and the shape parameter ξ_t that reflects the skewness and also the tail behavior of the probability distribution, which is thought to be enough for a good description of flood characteristics.

The POT floods are in fact portrayed separately by the magnitude of POT exceedances over a fixed threshold u and the attached arrival rates $\{M_t, t = 1, 2, ..., \tau\}$. The former series is modeled by the Generalized Pareto (GP) distribution (Pickands, 1975). This distribution is bound to the threshold u on the left with two-dimensional parameters (σ_t, ξ_t). If $\xi_t = 0$, it will be transferred to an exponential distribution with a single parameter $1/\sigma_t$. For fitting POT arrival rate, both Poisson and Negative Binomial (NB) (Anscombe, 1950) distributions are employed. The traditional use of the Poisson distribution only contains one parameter, i.e., the location parameter that is also termed the Poisson process intensity. This limitation makes it difficult to better adapt to the application under changing climate as explicated in the introduction. However, the alternative proposal of the NB distribution is competent in this regards owing to the scale parameter involved to represent the over-dispersion that may exist in POT arrival rate.

Table 1 summarizes the basic information for these distribution candidates. The parametric link function $g(\cdot)$ is a fairly general specification used to transform the distribution parameters of concern, for example, as natural logarithms (to ensure the positive value), or as identities. The Generalized Additive Models in Location, Scale, and Shape (GAMLSS) (Rigby and Stasinopoulos, 2005) is adopted for modeling the selected distribution, since it has been proven beneficial in providing a higher degree of flexibility to describe different hydro-meteorological variables through various families of distribution (López and Francés, 2013). For each candidate distribution, the transformed parameters (μ_t, σ_t, ξ_t) are modeled under stationarity (as constant) and nonstationarity (as linear functions of climatic covariates), respectively.

$$g(\cdot \mid \mu_t, \sigma_t, \xi_t) = \begin{cases} \text{constant} & \text{if stationary} \\ \mathbf{E}\boldsymbol{\beta} & \text{if nonstationary} \end{cases} \tag{3}$$

where a multidimensional vector of physical covariates candidates, $\mathbf{E}(1, E_1, E_2, \ldots)$, has a value of one in the first location for the intercept term. $\boldsymbol{\beta}\,(\beta_0, \beta_1, \beta_2, \ldots)$ is the vector of parametric coefficients to be numerically estimated by maximum likelihood technique. The computation can be easily finished by the iteration algorithms for optimization available in GAMLSS package on R software and determine the effective number of covariates (n_β). The assumption of a linear dependence on physical covariates should be regarded as a tradeoff between the diversity of covariates and the suspicion of over-fitting, which can be practicable in consideration of the referential experience (e.g., Villarini et al., 2009a, b; Xiong et al., 2015a, b). It must be noted that the mathematical expectation $E(M_t) = \mu_t = m$ is satisfied under stationarity whether using the Poisson or NB distribution (Coles, 2001). The combination of GP model for the magnitudes and (Poisson/NB) model for the arrival rates constitutes together a complete POT model. If time-varying parameters in the

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-619, 2016
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Published: 9 December 2016

Hydrology and Earth System Sciences Discussions

POT model exist, the nonstationary POT model will then be constructed.

2.3. Model selection and assessment

The model selection follows a generalized Akaike information criterion, i.e., $-2\ell + \# \cdot n_{\beta}$, to balance the considerations between structure complexity and goodness-of-fit, in which the penalty factor $\# = 2$ refers to the original AIC (Akaike, 1974) and $\# = \ln(N)$ to the Bayesian information Criterion (BIC) (Schwarz, 1978). ℓ is the value of the likelihood function. The priority choice is the model with minimum AIC and/or BIC values that tends to best capture the variation of observation with the simplest model structure. The model adequacy is diagnosed with a focus on the normality and independence of theoretical residuals r_t. Exempt from the influence of variability in the estimated parameters for a nonstationary model, the theoretical residuals r_t can be produced by inverting the fitted distribution function and finding the equivalent standard normalized quantiles (Dunn and Smyth, 1996), i.e., $r_t = \Phi^{-1}(\mathrm{Prob}_t)$, where Φ^{-1} is the inverse function of standard normal distribution, Prob_t is an abstraction for the theoretical probability at time t, having separate forms equal to, $F(x_t|\boldsymbol{\theta}_t)$ for the AM, $H(x_t|\boldsymbol{\theta}_t, u)$ for the magnitudes of POT exceedance, and a randomized value on the interval $[\Pr(M_t < m_t - 1|\theta_t), \Pr(M_t < m_t|\theta_t)]$ for a discrete integer response from the POT arrival rates (Rigby and Stasinopoulos, 2005). The following tests for r_t are utilized (at the 5% significance level):

(i) The normal Q-Q plot and its detrended version called the worm plot (Buuren and Fredriks, 2001). Given an observation series x_t rearranged in the

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-619, 2016
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Published: 9 December 2016

Hydrology and Earth System Sciences
Open Access
Discussions

descending order with the rank of $n(x_t)$, the empirical probability Prob_t^* is defined by $\dfrac{n(x_t)-0.44}{N+0.12}$ (Gringorten, 1963). The normal Q-Q plot of $\Phi^{-1}(\mathrm{Prob}_t^*)$ against r_t indicate a reasonable model if all the point-pairs lie around the unit diagonal (1:1 line). Instead of the vertical axis in Q-Q plot, the worm plot shows the differences between $\Phi^{-1}(\mathrm{Prob}_t^*)$ and r_t. A preferable model-fitting can be demonstrated if the distribution of data resembles a flat worm-like string within the 95% confidence interval.

(ii) The coefficients of determination for probability (R_{PP}^2) in Eq. (4) and for quantile (R_{QQ}^2) in Eq. (5), respectively. The higher the values of them, the better the model performs.

$$R_{PP}^2 = 1 - \frac{\sum_{t=1}^{N}[\mathrm{Prob}_t^* - \mathrm{Prob}_t]^2}{\sum_{t=1}^{N}[\mathrm{Prob}_t^* - \overline{\mathrm{P}}\mathrm{rob}_t]^2} \tag{4}$$

$$R_{QQ}^2 = 1 - \frac{\sum_{t=1}^{N}[\Phi^{-1}(\mathrm{Prob}_t^*) - r_t]^2}{\sum_{t=1}^{N}[\Phi^{-1}(\mathrm{Prob}_t^*) - \overline{r}_t]^2} \tag{5}$$

2.4. Return level formulations

To begin with, stationarity strategy of flood return level estimation in the classical extreme value theory is revisited in perspective of both the AM and POT. Under the stationarity assumption, the T-year return level x_T, subject to AM observations, is defined as the quantile for which the exceedance probability $\Pr(X > x_T)$ is $1/T$ for any particular year.

$$\Pr(X > x_T) = 1 - F(x_T|\boldsymbol{\theta}) = 1/T \tag{6}$$

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-619, 2016
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Published: 9 December 2016

Hydrology and Earth System Sciences Open Access Discussions

In view of the POT series with the threshold u, the exceedance probability of Eq. (7) has a similar form to that of Eq. (6) but with an additional constant item m, i.e., average annual arrival rates of the POT.

$$\Pr(X > x_T \,|X > u) = 1 - H(x_T \,|\boldsymbol{\theta}, u) = \frac{1}{m \cdot T} \tag{7}$$

As Coles (2001) warns, overreliance on the stationarity strategy of flood return level estimation presented above is risky unless the use of stationary assumption is pertinent. There is thereof a growing interest to understand how flood return levels could be when the possible nonstationarity has been accounted for.

The method of expected number of events (ENE) is therefore employed that facilitates the presentation of design flood in both stationarity and nonstationarity contexts. It defines return level x_T being a unique value such that the expected number of events over x_T in the next *T*-year return period will be one (Parey et al. 2007, 2010). This advantage makes the method able to provide unique design value for reference even though the flood behaviors observe nonstationarity, which is beyond the capacity of traditional stationarity strategy. For instance, dramatic (or pointless) *T*-year return levels of AM floods that change along the time axis will be obtained when applying Eq. (6) to the nonstationary models with time-varying parameters (López and Francés, 2013). The general formulation for any hypothetic probability distribution models can be expressed by

$$\begin{aligned} &\sum_{t=t_0+1}^{t_0+T} [1 - F(x_T \,|\boldsymbol{\theta}_t)] = 1 && \text{for AM} \\ &\sum_{t=t_0+1}^{t_0+T} \left\{[1 - H(x_T \,|\boldsymbol{\theta}_t, u)] \cdot E(M_t)\right\} = 1 && \text{for POT} \end{aligned} \tag{8}$$

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-619, 2016
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Published: 9 December 2016

where t_0+1 is the starting year of the flood return period. Here t_0 is set to be the end year of observation record for illustrating the method with future application. It can be noted that the magnitude of each POT exceedance in one year t is assumed to follow the same distribution $H(x_T|\boldsymbol{\theta}_t,u)$. The mathematical expectation $E(M_t)$ can be substituted by either Poisson or NB model with a given assumption for arrival rates, which on stationarity conditions, however, can transform Eq. (8) as

$$\begin{aligned} &1-F(x_T|\boldsymbol{\theta})=\frac{1}{T},\quad t=t_0+1,t_0+2,\ldots,t_0+T &&\text{for AM} \\ &1-H(x_T|\boldsymbol{\theta},u)=1/(m\cdot T),\quad t=t_0+1,t_0+2,\ldots,t_0+T &&\text{for POT} \end{aligned} \tag{9}$$

that coincides with the inferences in Eqs. (6) and (7), respectively.

In this study, the return level inferences are executed under both stationarity and nonstationarity. Taking account of the contradiction between the limited sample size and reliability of flood estimation, the return level formulations are intended to engage the study of design flood coupled with the nonparametric bootstrap resampling technique, as recommended to enhance the representativeness of sample (Salas and Obeysekera, 2014; Serinaldi and Kilsby, 2015). The *i.i.d.* assumption for stationarity strategy leads to a direct resampling of observation series for calculating the 95% confidence interval of return levels, while under nonstationarity the original data should be transformed into a standardized variable $\tilde{x}_t$ to follow an identical standard distribution before bootstrapping. This standard distribution is subjectively selected and naturally based on the distributional family that the fitting model belongs to, e.g., the standard Gumbel distribution used for the GEV model, the standard exponential distribution for the GP model. For the sake of convenience, the standard normal distribution is used for all constructed models. The nonstationary flood inference

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-619, 2016
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Published: 9 December 2016

Hydrology and Earth System Sciences Discussions

comprises several steps:

(i) Calculate *T*-year flood return levels x_T by applying Eq. (8) to the distribution models built in section 2.2 for AM and POT series, respectively.

(ii) Transform flood observation values x_t into $\tilde{x}_t$ using Eq. (10) (Coles, 2001) with the known model parameters obtained in step (1).

(iii) Resample $\tilde{x}_t$ with replacement for a large replication (i.e., 5000) and generate new observation samples by the inverse solution of Eq. (10).

(iv) Refit the same distribution for each new observation sample and calculate the return levels following step (1). The 95% confidence intervals for x_T are obtained.

$$\tilde{x}_t = \begin{cases} \ln\left[1+\xi_t\left(x_t-\mu_t\right)/\sigma_t\right]/\xi_t & \xi_t \neq 0 \\ \left(x_t-\mu_t\right)/\sigma_t & \xi_t = 0 \end{cases} \tag{10}$$

2.5. Global sensitivity analysis

A variance-based global sensitivity analysis is carried out with the Sobol' method (Sobol', 1993) to help understand the influence level of changing climate on return level estimations, which is important for future flood inference with due caution to the uncertainty originating from climate scenarios. This method is independent of model structure enabling an effective identification of both single and interactive parameter sensitivities and has been reported to outperform other methods (Tang et al., 2007). However, it is out of the scope to consider its own advantages/disadvantages.

Designate Ψ being all the parameters to be studied by the Sobol' method (i.e., climatic covariates) and flood return level x_T as a response variable that can be

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-619, 2016
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Published: 9 December 2016

Hydrology and Earth System Sciences
Open Access
Discussions

expressed according to the target function of Eq. (8). The total variance of x_T can be decomposed into

$$V = \sum_i V_i + \sum_i \sum_{i<j} V_{ij} + \ldots + V_{1,2,\ldots k} \tag{11}$$

where V_i is the first-order variance of the i-th parameter, indicating the contribution of single parameter to overall model uncertainty; V_{ij} is the second-order variance explained by the interactions between paired parameters of index i and j, and so on. The first-order sensitivity indices for the i-th parameter quantify the average proportion of V_i on the total variance (without any interactions with other parameters)

$$S_i = \frac{V_i}{V} \tag{12}$$

and the total-order sensitivity indices are

$$S_{T_i} = \frac{V_i + V_{ij} + \ldots + V_{ij\ldots k}}{V} = \frac{V - V_{\sim i}}{V} = 1 - \frac{V_{\sim i}}{V} \tag{13}$$

where $V_{\sim i}$ defines the average variance without any effect from the i-th parameter. The difference between the first-order and total-order sensitivity indices is the interaction between the i-th parameter and others ($i \neq j \neq k$). Due to the complexity of analytical solutions, V, V_i, and $V_{\sim i}$ are approximately estimated by Monte Carlo numerical integration (Saltelli, 2002) using Eqs. (14-18), to which the Sensitivity package on R provides easy access with the high computing power.

$$\tilde{\Psi}_s^A = \begin{bmatrix} \psi_{A1}^1 & \cdots & \psi_{Ai}^1 & \cdots & \psi_{Ak}^1 \\ \psi_{A1}^2 & \cdots & \psi_{Ai}^2 & \cdots & \psi_{Ak}^2 \\ \vdots & \cdots & \vdots & \cdots & \vdots \\ \psi_{A1}^n & \cdots & \psi_{Ai}^n & \cdots & \psi_{Ak}^n \end{bmatrix}; \quad \tilde{\Psi}_s^B = \begin{bmatrix} \psi_{B1}^1 & \cdots & \psi_{Bi}^1 & \cdots & \psi_{Bk}^1 \\ \psi_{B1}^2 & \cdots & \psi_{Bi}^2 & \cdots & \psi_{Bk}^2 \\ \vdots & \cdots & \vdots & \cdots & \vdots \\ \psi_{B1}^n & \cdots & \psi_{Bi}^n & \cdots & \psi_{Bk}^n \end{bmatrix} \tag{14}$$

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-619, 2016
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Published: 9 December 2016

$$\hat{h}_0 = \frac{1}{n}\sum_{s=1}^{n} h(\tilde{\Psi}_s^A) \tag{15}$$

$$\hat{V} = \frac{1}{n}\sum_{s=1}^{n} h^2(\tilde{\Psi}_s^A) - \hat{h}_0^2 \tag{16}$$

$$\hat{V}_i = \frac{1}{n}\sum_{s=1}^{n} h(\tilde{\Psi}_s^A) h(\tilde{\Psi}_{\sim is}^B, \tilde{\Psi}_{is}^A) - \hat{h}_0^2 \tag{17}$$

$$\hat{V}_{\sim i} = \frac{1}{n}\sum_{s=1}^{n} h(\tilde{\Psi}_s^A) h(\tilde{\Psi}_{\sim is}^A, \tilde{\Psi}_{is}^B) - \hat{h}_0^2 \tag{18}$$

where $\tilde{\Psi}_s^A$ and $\tilde{\Psi}_s^B$ are two different sample matrices by Monte Carlo simulation that each column shows the sample vector for each parameter. The sample vector is randomly selected from a uniform distribution for the given parameter ranges. $s = 1, 2, \ldots, n$ specifies the row number with the total simulation sample size of n (set to 1000 herein). $\tilde{\Psi}_{is}^A$ (or $\tilde{\Psi}_{is}^B$) represents the sample vector in the i-th column of $\tilde{\Psi}_s^A$ ($\tilde{\Psi}_s^B$) highlighted with the box in Eq. (14), while $\tilde{\Psi}_{\sim is}^A$ ($\tilde{\Psi}_{\sim is}^{(B)}$) denotes all the sample vectors except that of the i-th parameter. The function $h(\tilde{\Psi}_{\sim is}^A, \tilde{\Psi}_{is}^B)$ can be perceived as the calculation with $\tilde{\Psi}_s^A$ of which the i-th sample vector has been replaced by that from $\tilde{\Psi}_s^B$, which is similar to understand $h(\tilde{\Psi}_{\sim is}^B, \tilde{\Psi}_{is}^A)$.

3. Study area and data

3.1. Study area description

The Weihe is the biggest tributary of the Yellow River with a length of 818 km. It originates from the Niaoshu Mountain at Weiyuan County, mainly flows through Gansu and Shaanxi Provinces and Ningxia Hui Autonomous Region, and joins the Yellow River at Tongguan County. The Weihe basin, located in Northern China, has

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-619, 2016
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Published: 9 December 2016

an approximate drainage area of 134,800 km^2 (Xiong et al., 2014). It has a temperate continental monsoon climate, naturally showing semi-humid and semi-arid characteristics (Zuo et al., 2012). Average annual total precipitation of the Weihe basin is unevenly distributed with more (800-1000 mm) in the southern region and less (400-700 mm) precipitation in the northern region. The annual mean air temperature in the whole basin is about 6-14℃, and average annual mean runoff depth is about 100 mm (Du et al., 2015). The catchment downstream of Huaxian gauging station is used as the study region. This region covers 80% of the Weihe basin with a 106,498 km^2 drainage area. Figure 1 presents the geographical information of the Weihe basin and the study region.

3.2. Meteorological data

The observations of daily total precipitation and daily mean air temperature from 22 meteorological stations over the period 1960-2009 were provided by the China Meteorological Administration. The weighted areal precipitation and air temperature series were generated by the Thiessen polygon (e.g., Du et al., 2015), using 10 and 12 stations in and around the Huaxian catchment, respectively. Five physical factors were chosen as proxies to represent physical covaraites (E) based on the previous research (Xiong et al., 2015a): annual total precipitation (P_{total}); annual maximum precipitation on consecutive one, three, and seven days (denoted as P_{max_1d}, P_{max_3d}, and P_{max_7d}, respectively); and annual mean air temperature (T_{mean}). These factors can be viewed as the pertinent surrogate on behalf of basin climatic characteristics that may exert important effects on river runoff generation, rainfall-runoff process, etc.

Having access to the future meteorological data (i.e., projected precipitation and air temperature) is of great significance for extrapolating future design floods when the return level formulations of Eq. (8) are used with physically-based models. In this paper the General Circulation Models (GCMs) data sets that have been used worldwide are employed to obtain insight into the unknown future climate and the statistical downscaling model (Wilby et al., 2002), a model combining multiple linear regression and stochastic weather generator, to deal with the mismatched spatial resolution between the large-scale GCMs outputs and local-scale climate information. The statistical downscaling model (SDSM) is selected for use due to its merits such as the convenient operation of producing spatially and temporally continuous fine-scale precipitation and air temperature information at a basin scale (Raff et al., 2009). Its technical procedure mainly includes: analyzing the correlation between the NOAA National Centres for Environmental Prediction (NCEP) reanalysis predictors and historical precipitation record by a multiple linear regression; running weather generator in the SDSM to simulate precipitation based on the constructed multiple linear regressions; calibrating the SDSM by assessing the predictive performance; and projecting precipitation scenarios from the GCMs data in the calibrated SDSM. The details of SDSM have been interpreted in the recent publication (Du et al., 2015) thus not being covered here for brevity. Interested readers can find more information in Wilby et al. (2002). The data of 26 NCEP reanalysis predictors for the period of 1960-2009 were available from the NOAA Earth System Research Laboratory (ESRL) (http://www.esrl.noaa.gov). The latest version of GCMs from the Coupled Model

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-619, 2016
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Published: 9 December 2016

Hydrology and Earth System Sciences
Open Access
Discussions

Intercomparison Project Phase 5 (CMIP5) have projected new generation scenarios of greenhouse gas emissions, i.e., the Representative Concentration Pathways (RCPs), as recommended by the Fifth Assessment Report of the Intergovernmental Panel on Climate Change (IPCC, 2013). The RCP8.5, a scenario representative of considerable greenhouse gas concentration levels, was chosen as motivated by Peters et al. (2013) who considered that the RCP8.5 reflects the recent trends of global carbon dioxide emissions reasonably. The same 26 predictors of seven different GCMs (CanESM2, CCSM4, CNRM-CM5, GFDL-ESM2M, MIROC-ESM, MIROC-ESM-CHEM, and NorESM1-M) under the RCP8.5 scenario for the future period of 2010-2099 were downloaded from the CMIP5 website (http://cmip-pcmdi.llnl.gov/cmip5). Gridded daily data of both NCEP and GCMs were first interpolated to each of 22 meteorological stations by the Inverse Distance Weighting method (Atkinson and Tate, 2000), and then processed into weighted areal series for the Huaxian catchment by the Thiessen polygon.

3.3. Flood data

Daily flow records of the Huaxian station were collected from 1960 to 2009 by the Yangtze River Waterway Bureau. Two sampling schemes of AM and POT were utilized to describe the flood events. The threshold of POT sampling is determined according to the preselected annual number of peaks per year on average (Lang et al., 1999). The peaks defined as the highest values on a centered 17-day window, are restrained by the criteria of Eq. (19), proposed by USWRC (1982).

$$TI > 5 + \log(Area)$$
$$Q_{\text{int}} < 0.75 \min(Q_1, Q_2) \quad (19)$$

where TI is the time interval between two consecutive peaks in days, Area is the basin area (km^2) and Q_{int} represents the intermediate flows between two consecutive peaks (Q_1 and Q_2). Assuming that average POT arrival rate per year is two, three, and four, respectively, the POT series are screened out, i.e., three POT magnitude series denoted by POT2 (u =1060 m^3/s), POT3 (u =780 m^3/s), and POT4 (u =530 m^3/s); and their corresponding arrival rate series by POT_AR2, POT_AR3, and POT_AR4, respectively.

4. Results

4.1. Data analyses: stationary or nonstationary?

Based on the diagnostics in section 2.1, changes of temporal trends were explored over the observation period of 1960-2009 for all flood-feature series (Table 2) including AM, POT magnitudes (POT2, POT3, and POT4), and POT arrival rates (POT_AR2, POT_AR3, and POT_AR4), as well as the physical covariates (P_{total}, P_{max_1d}, P_{max_3d}, P_{max_7d}, and T_{mean}). A significantly decreasing trend was detected in the AM series regardless of whether the MK, PW, or TFPW method was used. Similarly, no differences occurred among the results of the MK, PW, and TFPW tests for the POT series. The POT magnitudes showed non-significantly declined tendencies in POT2, POT3, and POT4 series. However, their corresponding arrival rates (POT_AR2, POT_AR3, and POT_AR4) exhibited dramatically negative trends. The statistics Z_{MK} (-4.12, -4.00, and -3.35), Z_{PW} (-3.19,-2.72, and -2.52), and

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-619, 2016
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Published: 9 December 2016

Z_{TFPW} (-4.26, -3.80, and -3.73) whose values in brackets were presented orderly for POT_AR2, POT_AR3, and POT_AR4 indicated that these downward trends in the POT arrival rates became more significant with the generally increasing threshold. Moreover, the results of Bohning (1994) test statistic demonstrated that the assumption of homogeneous Poisson process would not be applicable to describe POT arrival rates under current environments for the three POT arrival rates. Physical covariates regarding the precipitation-related variables (P_{total}, P_{max_1d}, P_{max_3d}, P_{max_7d}) presented no significant trends according to the MK tests, but declining trends were detected in both P_{total} and P_{max_7d} by their respective PW and TFPW statistics. T_{mean} showed a clear uptrend using all the trend tests (not shown in Table 2).

How could the detected trends in floods be when linking with the potential influencing factor (climatic covariates)? The PMK test was applied to investigate whether the trends can still be significant after removing the dependence on each of the physical covariates (P_{total}, P_{max_1d}, P_{max_3d}, P_{max_7d}, and T_{mean}) and the associate extent. The *p*-values of the PMK test showed that the detected trends in AM and POT_AR2 would disappear once associated with either T_{mean} or P_{total}. POT_AR3 and POT_AR4 series also had a dependence on P_{total} and T_{mean}, respectively. However, the detected trends were less affected by $P_{max\text{-}1d}$, $P_{max\text{-}3d}$, and $P_{max\text{-}7d}$.

The nonstationarity of the hydrologic system in the Weihe basin detected here has also been proven earlier (Zuo et al., 2012, 2014; Du et al., 2015; Xiong et al., 2014, 2015a; Jiang et al., 2015), thereby motivating the interest of extrapolating flood return levels that considers nonstationarity.

4.2. Flood-frequency distribution models for AM and POT

Table 3 lists the stationary models and the nonstationary models calibrated by optimal combination of climatic covariates $\mathbf{E}$ under the assumption of LNO3, LP3, and GEV distributions, respectively. The Q-Q plots for these models visually confirmed the reasonable model performance since the resulting points did not significantly deviate from 1:1 line (Fig. 2). In case of stationarity, LP3 model yielded the smallest AIC/BIC value among the candidates with slightly higher values of $R_{PP}^2 = 92.5\%$ and $R_{QQ}^2 = 91.1\%$ thus being regarded as the optimum. It has been found that in nonstationarity context the optimum should again owe to the model of LP3 distribution whose AIC/BIC values (820.9/830.5) were much less than the remaining ones, with favorable model adequacy suggested by R_{PP}^2 and R_{QQ}^2. In this nonstationary model, the location parameter μ_t negatively correlates to T_{mean} but positively correlates to P_{total}, in accordance with the trend test result of the AM series in section 4.1. The similar results for fitting the AM flood series of the Weihe basin can be found in Xiong et al. (2015a).

Table 4 presents detailed fitting information of the stationary GP models for the three POT magnitude series (POT2, POT3, and POT4) that were found to be better than the models under nonstationarity. It may not be the case respecting their POT arrival rate series (POT_AR2, POT_AR3, and POT_AR4) where nonstationary models produced much improvement over the corresponding stationary models according to the AIC/BIC values. For example, the AIC value of nonstationary Poisson model was 149.9 much lower than that (182.9) of stationary Poisson model in

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-619, 2016
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Published: 9 December 2016

case of POT_AR2. Table 5 shows both the optimal Poisson and NB models with parameters fitted as functions of climatic covariates. The Poisson distribution remained to be preferred over the NB distribution for fitting POT_AR3 and POT_AR4. However, the worm plots in Fig. 3 reveal a better performance of the NB model for fitting POT_AR2 and this model also has a lower AIC/BIC value than the Poisson model. It was found that the best model (NB) fitted for POT_AR2 showed a time-varying scale parameter dependent on P_{total} while the other NB models fitted for POT_AR3 and POT_AR4 had constant scale parameters. Integrating the separate results for magnitudes and arrival rates, the optimal nonstationary POT models are epitomized orderly for POT2, POT3, and POT4:

$$\begin{aligned}
&\mathrm{POT2}\sim\mathrm{GP}[\ln(\sigma_t)=7.03,\xi_t=0.11]+\mathrm{NB}[\ln(\mu_t)=2.791+0.004P_{total}-0.480T_{mean},\ln(\sigma_t)=-0.013P_{total}]\\
&\mathrm{POT3}\sim\mathrm{GP}[\ln(\sigma_t)=6.76,\xi_t=0.13]+\mathrm{PO}[\ln(\mu_t)=-1.088+0.004P_{total}]\\
&\mathrm{POT4}\sim\mathrm{GP}[\ln(\sigma_t)=6.56,\xi_t=0.1]+\mathrm{PO}[\ln(\mu_t)=0.004P_{total}-0.074T_{mean}]
\end{aligned} \tag{20}$$

4.3. Flood projections under the climate scenarios

Given that both P_{total} and T_{mean} have been parameterized in the nonstationary models for AM and POT, the future scenarios of P_{total} and T_{mean} are in need of the investigation of future flood return levels projected by the ENE method. Herein the scenarios generated by the seven GCMs in Du et al. (2015) were applied. Figure 4 shows the average projections (red lines) and their ranges (gray shadow) from the seven GCMs over the future period 2010-2099 for P_{total} and T_{mean}, respectively. The result announced the notably rising T_{mean} (average annual growth of around 0.0596℃) and the negligible increase in P_{total} (average annual growth of approximately 0.13mm) over the future period. Flood return levels are inferred under both stationarity and

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-619, 2016
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Published: 9 December 2016

Hydrology and Earth System Sciences
Discussions
Open Access

nonstationarity denoted as x_T^s and $x_T^{non\_s}$, respectively, for the convenience of explanation. To eliminate the uncertainty brought by single GCM, the averaged projections for P_{total} and T_{mean} (red lines) were finally used to calculate $x_T^{non\_s}$.

Variations of *T*-year flood return levels estimated by using the models in Table 3 are presented in Fig. 5a. The largest flood magnitudes were estimated by the LNO3 model followed by the GEV and LP3 models. With the use of LNO3 model, $x_T^{non\_s}$ presented values above and then below x_T^s as return period *T* prolonged through a transition *T* of around 30 years. For both LP3 and GEV model, $x_T^{non\_s}$ were generally lower than the corresponding x_T^s. However, differences between $x_T^{non\_s}$ and x_T^s appeared to reduce over *T* of 30-50 years for the LP3 model while enlarged evidently in case of the GEV model (The largest magnitude of their difference can reach above 2000 m^3/s) where no overlap of their 95% confidence intervals announced. It is interesting to note that $x_T^{non\_s}$ have similar estimations to x_T^s in both cases of the LNO3 and LP3 models with *T* around 30, in which there is much overlap of confidence intervals between $x_T^{non\_s}$ and x_T^s.

Figure 5b displays the results of flood return levels for the POT series (POT2, POT3, and POT4) where x_T^s estimated from stationary POT model, i.e., stationary GP with constant arrival rate m, and $x_T^{non\_s}$ from climatic covariates-dependent POT model, i.e., a combination of stationary GP and nonstationary (Poisson/NB) models. Three important findings were delivered: (i) the overall differences between x_T^s and $x_T^{non\_s}$ became larger as *T* increased. $x_T^{non\_s}$ were always lower than x_T^s whether the Poisson or NB distribution was employed; (ii) the difference of $x_T^{non\_s}$

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-619, 2016
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Published: 9 December 2016

arising from the use of Poisson and NB models expands orderly for POT4, POT3, and POT2 floods; (iii) no matter for x_T^s or $x_T^{non\_s}$, flood estimations dropped in order of POT2, POT3, and POT4 series for any given T. For example, it was observed that 50-year x_T^s had the different estimations of 7858 m^3/s, 6897 m^3/s, and 5429 m^3/s for POT2, POT3, and POT4, respectively.

Comparing the results for AM and POT series, flood return levels x_T^s estimated with POT was larger than those with AM if the threshold was set relatively high such as u^{POT2} (1060 m^3/s) and vice versa. However, no similar features were found in $x_T^{non\_s}$ on nonstationarity conditions.

4.4. Sensitivity of flood estimations to changing climate

How the flood return levels would co-vary with the parameters of climatic covariates P_{total} and T_{mean} is checked by the Sobol' sensitivity analysis. The parameter samples were generated randomly with ranges defined by the seven climatic scenarios in Fig. 4. Sensitive parameters are designated as those that have a contribution of at least 10 percent. Parameters controlling 50 percent of the overall model variance are thought to be highly sensitive.

The first-order and total-order Sobol' indices in nonstationary models fitted for AM over the return period of 90 years are shown in Fig. 6. In Fig. 6a, Sobol' indices computed with the LNO3 model discerned both P_{total} and T_{mean} as sensitive parameters. The total contribution of P_{total} (averagely 65%) to overall output variance was larger than that of T_{mean} (averagely 47%). High parameter sensitivity was captured in Fig. 6b for the LP3 model in which the total-order indices of above-50

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-619, 2016
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Published: 9 December 2016

present occurred in both P_{total} and T_{mean} with a steadily rising trend in indices. The results for GEV model in Fig. 6c showed the low indices (at about 0.3) for P_{total} but anyway demonstrated the sensitivity to it whereas T_{mean} was classified as highly sensitive parameter that would exhibit more effect on output variance. An imperceptibly increasing tendency was found in Sobol' indices for both parameters P_{total} and T_{mean} in the GEV model. In all of the results presented for the three models, the highest sensitivity for P_{total} existed in the LP3 model reaching up to 0.75, followed by the LNO3 model explaining about 67% of the total variance at most, and the GEV model presented the lowest sensitivity to P_{total} (no more than 0.35). These models were all greatly sensitive to T_{mean} with small difference in the values of total-order indices and the highest value again occurred in the LP3 model controlling 44 to 63 percent of flood response. The temporal dynamics of respective parameter interactions (shown in Fig. 6 with the shading area) indicated that all the model sensitivities to T_{mean}, as would be expected, were more highly interactive, with approximately 11-27 percent of its influence on model output coming from interactions with other parameters, than that to P_{total}

Figure 7 shows the results for POT floods which are in general not as sensitive to climate change as AM floods on nonstationarity conditions. Overall, the total-order or first-order sensitivity indices became larger in sequence of POT4, POT3, and POT2. T_{mean} was seen as sensitive though its Sobol' indices were not very high (averagely above 0.1), while the sensitivity to P_{total} presented much lower values than that to T_{mean}, especially in case of POT4 using the Poisson model where P_{total} was assigned to be non-sensitive parameter. Sensitivities of POT flood response to P_{total} and T_{mean} were also compared when using the Poisson and NB models, respectively. Adopting the NB instead of Poisson model for fitting POT_AR2 is likely to attach less uncertainty as discovered in Fig. 7a where the sensitivity indices with the NB were mostly below 0.22 in contrast to that with the Poisson model. However, the opposite results were found with POT3 and POT4 for which climate effect is stronger with NB model. These findings might increase confidence in promotion of the NB distribution for significantly heterogeneous POT arrival rates (e.g., POT_AR2 with variant scale parameter) while for nonstationary POT arrival rates without significantly changing variance, a time-varying Poisson process can be competent.

Changes in $x_T^{non\_s}$ computed with different values of the average scenarios increased by increments of 0-20% (each scenario series was altered alone with other parameters fixed) are shown in Fig. 8 with the specific examples corresponding to the return level of 5, 10 and 80 years. It is seen that for the LNO3 model, an increase only imposed in T_{mean} caused a declining flood response, and such response would be stronger as return period prolonged or increment enlarged. A shift in flood response to single variation of P_{total} from the escalating to moderating trend was also noted. Analyses conducted on the LP3 and GEV models shows that their derived flood return levels both corresponded to rising values in response to increasing P_{total} and were quickened to descend by a large growth in T_{mean}. However, with the use of GEV model, the variation of return level as a response to increasing P_{total} is rarely reflected (the biggest rise is roughly 360m^3/s). Similar results were observed for POT floods that a rising T_{mean} leading to the decrease of flood response has a larger influence than an increasing P_{total} (corresponding to higher flood estimation). Based on the analyses, flood estimations under nonstationarity would presumably be lower with a single effect of increasing air temperature or declining precipitation.

These analyses make sense of our results in section 4.3 explaining somewhat why there is not much difference between $x_T^{non\_s}$ and x_T^s while a downward trend in floods has been verified. In the LNO3 model, separate change in T_{mean} causes a continuing decrease in $x_T^{non\_s}$ as *T* increases while single variation of P_{total} presents different effect that makes $x_T^{non\_s}$ first increase with a short *T* and then decrease for a longer *T*. Taking into account that P_{total} has a higher overall importance than T_{mean} (Fig. 6) and the latter shows a significantly upward trend (Fig. 4), their short-term inverse effects are likely to generate $x_T^{non\_s}$ similar to x_T^s and with the growth of return period, the agreement of effects between them might result in a significantly lower $x_T^{non\_s}$ than x_T^s. Analogously, the accumulation of inverse effects between P_{total} and T_{mean} in the LP3 model has rendered a very small difference between $x_T^{non\_s}$ and x_T^s. In the GEV model, a gently increasing T_{mean} is mainly responsible for the markedly declining $x_T^{non\_s}$ as the strong effect of T_{mean} on $x_T^{non\_s}$ has been notified. Likewise, changes in T_{mean} also control POT flood response that lower values of $x_T^{non\_s}$ would be caused given the long-term significant growth of temperature scenarios used here.

5. Discussion

The ENE method provides a new path to expand flood design to nonstationarity conditions with both AM and POT samplings conveniently with an input of future climate scenarios into the pre-constructed flood-frequency distribution model. A preliminary challenge is providing the faithful evidence with real nonstationarity (Villarini et al., 2009a) if we allowing for the nonstationary modeling with historical flood. Encouraged by the data analyses (section 4.1) and the preceding studies of the Weihe basin (Zuo et al., 2012; Xiong et al., 2014, 2015a; Du et al., 2015), the present study is designed to release nonstationarity for future flood extrapolation under changing climate. In addition to climate change, further research could examine other physical covariates like human impact, an important factor to influence flood process (Zuo et al., 2014; Jiang et al., 2015) which, however, is not included here given current difficulty in prediction of future anthropogenic factors that may requires specific studies of other disciplines (e.g., sociology, economics).

The confirmed nonstationarity is parameterized by modeling different probability distributions with time-varying parameters as functions of climatic covariates so that the effect of climate on complex flood response can be explained (Villarini et al., 2009a, b; Prosdocimi et al., 2015). In this sense, these physically-based nonstationary models adopted here lay a more reliable basis to ensure the quality of flood projection than those purely using explanatory covariates like time without clear causality, which, however, has long been used before (Du et al., 2015). The optimal POT models that combined stationary magnitudes with time-variant arrival rates were discovered

(Tables 4 and 5), similar to that in Parey et al. (2010), but different from Silva et al. (2015) who found that both POT magnitudes and arrival rates changed dependent on the physical covariates. This result conforms to the preliminary test for nonstationarity in section 4.1, which also motivates the proposal of the NB distribution instead of traditional usage of the Poisson for POT arrival rate modeling. A comparison between the Poisson and NB distributions highlights the superiority of the latter for fitting POT_AR2 (Fig. 3) where the time-varying scale parameter was found (Table 5). This might implicitly assume the inapplicability of the homogeneous Poisson while POT arrival rate shows high variability in variance considering other comparable studies conducted elsewhere (e.g., Ben-Zvi, 1991; Villarini et al., 2012). However, there are some divergent voices, such as Cunnane (1979) and Önöz and Bayazit (2001), who suggested the use of Poisson distribution even when the Poisson distribution assumption was rejected by statistical tests, and Bezak et al. (2014), who found that the NB distribution did not offer improvements over the Poisson distribution for fitting POT arrival rates. It is necessary to point out that the climate-dominated nonstationary model as well as the climate scenarios we implemented here are not mandatory but rather identified for a specific basins of interest (López and Francés, 2013). The proposal for future climate scenarios here is to use GCMs, an advanced tool used worldwide for replicating current climate condition and predicting unknown future climate. While beyond the scope of this paper, it must be noted that there have been still difficulties in adequate climate projection based on GCMs due to their inherent defects such as the oversimplified conceptualization of nonlinear processes, coarse resolution, and moderate performance in modeling rainfall characteristics like the frequency, intensity, and extremes (e.g., Raff et al., 2009; Koutsoyiannis, 2011; Chen et al., 2012; Du et al., 2015). To relieve the negative impact and attain more credible climate scenarios in study area of interest, endeavors on complete assessment related to the choice of the scenarios, climate models and downscaling methods are of realistic significance.

Appling the future changing climate scenarios to the nonstationary climatic covariate-dependent POT models, return levels $x_T^{non\_s}$ derived with the nonstationary Poisson distribution are invariably higher than those from the NB distribution, and their differences become more evident with the increasing POT threshold. A typical example can be found in Fig. 5 comparing the results between POT2 and POT4, specifically for a shorter return period when the gaps are negligibly small in case of POT4 but easily recognizable for POT2. Similar outcome has been reported early for stationarity strategy in Önöz and Bayazit (2001) that flood estimates were nearly identical based on both Poisson and NB distributions. It is natural to suppose that the difference levels of flood estimations between the Poisson and NB distribution are associated with the given POT threshold. Likewise, this surmise is tenable when comparing estimated floods between the AM and POT samplings. POT sampling does not always give higher flood designs than AM under either stationarity or nonstationarity assumption (e.g., POT4). Various results have been found in other research, e.g., Önöz and Bayazit (2001) applied POT series with fewer than average three events per year to stationary flood estimation and found that POT always gave lower estimates than AM, whether using Poisson or NB distribution, Bezak et al. (2014) recently showed that POT series with an average of five events per year produced higher flood estimations than AM when the Poisson distribution was assumed. One plausible explanation for these phenomena is that the POT series extracted above a low threshold lose the significance of 'real flood', while the POT series with a very high threshold are more liable to expose the nonstationarity in response to changing climate as have been stated by the increasing MK statistics orderly for POT_AR4, POT_AR3, and POT_AR2 (section 4.1).

Results for either AM or POT floods declare that the ENE method could yield design floods much and/or little lower than those derived from the stationary model with identical distribution assumption. For example, $x_T^{non\_s}$ from the optimal nonstationary LP3 model deviated slightly from x_T^s estimated with stationary LP3 model for the return periods of 30-50 years. POT4 shows a minor difference between x_T^s and $x_T^{non\_s}$ using the Poisson model. However, such result may not be dependable for other discussed distributions like GEV where the striking difference (>2000m^3/s) between x_T^s and $x_T^{non\_s}$ has been informed (Fig. 6c). Accordingly, it is revealed that under changing climate scenarios, nonstationary flood frequency model embedded in the ENE method does not necessarily lead to the results that are significantly different from those obtained by traditional stationarity strategy.

A sensitivity analysis of flood estimations to changing climate in section 4.4 effectively illustrates our reports with, e.g., the LP3 model for extrapolating AM floods that nonstationarity cannot readily be conjectured to be the transformed parlance of "changes" (Koutsoyiannis and Montanari, 2015) due to the underlying parameter interactions emerged in the ENE inference. This model captures the most nonstationary variation of flood thus being preferred at first glance for application to the ENE method. However, the high sensitivity to changing climate detected in it has to be taken into account cautiously because of the implication that uncertainty problem involved might be added. Contrasting to the AM floods, the application to POT series seems not to be that sensitive (Fig. 7), and from this perspective, it should be henceforth devoted sufficient attention for nonstationary flood return level analysis in view of its relatively lower uncertainty originating from climate scenarios. Comparing the Poisson and NB models respecting POT flood estimations, climate variability has a lower influence on flood projection (or cause less uncertainty) when using the NB distribution for the POT series with a high threshold, e.g., POT2 and that the best model fitted for nonstationary POT2 series is confirmed with the NB instead of Poisson distribution. The promotion of the NB distribution in nonstationarity context is in this regards of necessity. It is presumably that a weaker reliability of flood projection on nonstationarity condition would be made if future climate is poorly predicted and/or nonstationarity in POT arrival rate inappropriately represented by a homogeneous Poisson process.

In light of the analyses above, we believe that the ENE method has the potential for flood projections as it is easily understandable and computationally efficient. Nevertheless, the ENE method still faces barriers to reliable flood predictions as the method is by its very nature subjugated to so many mathematical hypotheses.

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-619, 2016
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Published: 9 December 2016

Therefore, carefulness should be taken in practice for nonstationary flood projections that have been carried out herein given the limitations incumbent upon the method, such as the cause-effect mechanism of nonstationarity, the extent to which climate scenarios asymptotically converge to the reality in the future and the uncertainty associated with nonstationarity inference (e.g., Lins and Cohn, 2011; Koutsoyiannis, 2011; Montanari and Koutsoyiannis, 2014; Salas and Obeysekera, 2014; Serinaldi and Kilsby, 2015; Silva et al., 2015).

6. Conclusions

Flood return levels have been projected under future changing climate scenarios by applying the expected number of exceedances (ENE) method to both Annual Maximum (AM) and Peaks over Threshold (POT) series of the Weihe basin, China. To evaluate the climatic effect on flood projection, the sensitivity of flood response to future changing climate is explored via the Sobol' method. The initial detection of nonstationarity confirms a significantly decreasing trend in the observed AM floods while the POT records are characterized by stationary flood magnitudes with the heterogeneous occurrences. The findings can therefore motivate the proposal of the Negative Binomial (NB) distribution for fitting POT arrival rates given the previous report that the common assumption of homogeneous Poisson process might be invalid under nonstationarity. Time-varying flood-frequency models parameterized for describing nonstationarity are constructed via functional relation between distribution parameters and climatic covariate, which have been proven to yield superiority over stationary ones for the series of AM and POT arrival rates. The variations of AM floods are best captured by the physical covariates-dependent LP3 model. For fitting the over-dispersed POT series (with significantly variant variance higher than the mean in its arrival rates, hereinafter), the NB distribution model is demonstrated to be preferable to the Poisson model.

The comparison of differences between the return levels calculated with AM and POT floods reveals that the AM flood projections are mostly lower than the POT estimation except when POT series are sampled with small threshold (attributed to the damage of real flood information). The AM-based flood extrapolation is more vulnerable to climate change than flood estimation with POT. From this perspective, we suggest that POT series should be warranted more attention in nonstationary flood frequency analysis, as the relatively complicated sampling criteria has long limited its application.

Comparison with respect to the POT-based flood projection shows that the presence of over-dispersed flood occurrences could lead to overestimation of return levels if treating the assumption of homogeneous Poisson process without discretion. The gaps based on the choices between the Poisson and NB models would enlarge with the increasing POT threshold value. Referring to the advantage in model fitting and low detrimental impact on future flood extrapolation (i.e., incurring less uncertainty originating from changing climate than the Poisson model), the NB distribution would be a better choice when POT arrival rates exhibit significant heterogeneity (e.g., POT_AR2 tested in this study).

Under future changing climate, flood return levels derived with the ENE method are usually but not always more different from those analyzed by traditional stationarity strategy although the significant representation of nonstationarity in flood samples has been notified. Such results could be indirectly approved due to the generally opposite impacts of increasing air temperature and precipitation (used as the climatic covariates in this study) as well as the different extent to which the flood estimation responds to them. This information has important implication to the influence of multifactorial interactions included in the ENE inferences which could perhaps maintain the dynamic balance between stationarity and nonstationarity. It is therefore as stressed that nonstationarity cannot be taken as equivalent to change. However, assuming a separate variation of increasing air temperature or declining precipitation, lower flood estimation under nonstationarity could be induced in general.

This study can be useful in guiding decisions with flood design under changing climate and as an attempt for future inference to contribute to the further development of relieving the concomitant problems attached with nonstationary flood extrapolation. Given that this region-specific research only considers the impact of climate change, it is suggested a more sufficient consideration of physical covariates relevant to flood response for application to other studies.

Acknowledgements

This research is financially supported by the National Natural Science Foundation of

China (Grants NSFC 51525902, 51190094, and 51479139), which is greatly appreciated.

References

Akaike, H.: A new look at the statistical model identification, IEEE Trans. Automat. Control, 19, 716-723, doi:10.1109/TAC.1974.1100705, 1974.

Anscombe, F.J.: Sampling theory of the negative binomial and logarithmic series distributions, Biometrika, 37, 358-382, doi:10.2307/2332388, 1950.

Atkinson, P.M., and Tate, N.J.: Spatial scale problems and geostatistical solutions: a review, Professional Geographer, 52, 607-623, doi: 10.1111/0033-0124.00250, 2000.

Ben-Zvi, A.: Observed advantage for negative binomial over Poisson distribution in partial duration series, Stoch. Hydrol. Hydraul., 5, 135-146, doi:10.1007/BF01543055, 1991.

Bezak, N., Brilly, M., and Šraj, M.: Comparison between the peaks-over-threshold method and the annual maximum method for flood frequency analysis, Hydrol. Sci. J., 59, 959-977, doi:10.1080/02626667.2013.831174, 2014.

Böhning, D.: A note on a test for Poisson overdispersion, Biometrika, 81, 418-419, doi: 10.1093/biomet/81.2.418, 1994.

Buuren, S.V., and Fredriks, M.: Worm plot: a simple diagnostic device for modelling growth reference curves, Stat. Med., 20, 1259-1277, doi:10.1002/sim.746, 2001.

Chen, H., Xu, C.-Y., and Guo, S.: Comparison and evaluation of multiple GCMs, statistical downscaling and hydrological models in the study of climate change impacts on runoff, J. Hydrol., 434-435, 36-45, doi:10.1016/j.jhydrol.2012.02.040, 2012.

Coles, S.: An introduction to statistical modeling of extreme values, Springer, London, 2001.

Cooley, D.: Return periods and return levels under climate change, in Extremes in a changing climate: detection, analysis and uncertainty, edited by AghaKouchak, A. et al., Springer, Dordrecht, Netherlands, 97-114, 2013.

Cunnane, C.: A note on the Poisson assumption in partial duration series models, Water Resour. Res., 15, 489-494, doi:10.1029/WR015i002p00489, 1979.

Du, T., Xiong, L., Xu, C.-Y., Gippel, C.J., Guo, S., and Liu, P.: Return period and risk analysis of

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-619, 2016
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Published: 9 December 2016

nonstationary low-flow series under climate change, J. Hydrol., 527, 234-250, doi:10.1016/j.jhydrol.2015.04.041, 2015.

Dunn, P.K., and Smyth, G.K.: Randomized quantile residuals, J. Comput. Graph. Stat., 5, 236-244, doi:10.2307/1390802, 1996.

Gringorten, I.I.: A plotting rule for extreme probability paper, J. Geophys. Res., 68, 813-814, doi:10.1029/JZ068i003p00813, 1963.

IPCC: 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, United Kingdom and New York, USA, 2013.

Ishak, E.H., Rahman, A., Westra, S., Sharma, A., and Kuczera, G.: Evaluating the non-stationarity of Australian annual maximum flood, J. Hydrol., 494, 134-145, doi:10.1016/j.jhydrol.2013.04.021, 2013.

Jenkinson, A.F.: The frequency distribution of the annual maximum (or minimum) values of meteorological elements, Q. J. Roy. Meteor. Soc., 81, 158-171, doi:10.1002/qj.49708134804, 1955.

Jiang, C., Xiong, L., Wang, D., Liu, P., Guo, S., and Xu, C.-Y.: Separating the impacts of climate change and human activities on runoff using the Budyko-type equations with time-varying parameters, J. Hydrol., 522, 326-338, doi:10.1016/j.jhydrol.2014.12.060, 2015.

Kendall, M.G.: Rank correlation methods, Charles Griffin, London, 1975.

Khaliq, M.N., Ouarda, T.B.M.J., Ondo, J.-C., Gachon, P., and Bobée, B.: Frequency analysis of a sequence of dependent and/or non-stationary hydro-meteorological observations: A review, J. Hydrol., 329, 534-552, doi:10.1016/j.jhydrol.2006.03.004, 2006.

Koutsoyiannis, D.: Hurst-Kolmogorov dynamics and uncertainty, J. Am. Water Resour. Assoc., 47, 481-495, doi:10.1111/j.1752-1688.2011.00543.x, 2011.

Koutsoyiannis, D., and Montanari, A.: Negligent killing of scientific concepts: the stationarity case, Hydrol. Sci. J., 60, 1174-1183, doi:10.1080/02626667.2014.959959, 2015.

Lang, M., Ouarda, T.B.M.J., and Bobée, B.: Towards operational guidelines for over-threshold modeling, J. Hydrol., 225, 103-117, doi:10.1016/S0022-1694(99)00167-5, 1999.

Libiseller, C., and Grimvall, A.: Performance of partial Mann-Kendall tests for trend detection in the presence of physical covariates, Environmetrics, 13, 71-84, doi:10.1002/env.507, 2002.

Lins, H.F., and Cohn, T.A.: Stationarity: wanted dead or alive?, J. Am. Water Resour. Assoc., 47, 475-480, doi:10.1111/j.1752-1688.2011.00542.x, 2011.

López, J., and Francés, F.: Non-stationary flood frequency analysis in continental Spanish rivers, using climate and reservoir indices as external covariates, Hydrol. Earth Syst. Sci., 17, 3189-3203, doi:10.5194/hess-17-3189-2013, 2013.

Machado, M.J., Botero, B.A., López, J., Francés, F., Díez-Herrero, A., and Benito, G.: Flood frequency analysis of historical flood data under stationary and non-stationary modeling, Hydrol. Earth Syst. Sci., 19, 2561-2576, doi:10.5194/hess-19-2561-2015, 2015.

Madsen, H., Rasmussen, P.F., and Rosbjerg, D.: Comparison of annual maximum series and partial duration series methods for modeling extreme hydrologic events: 1. At-site modeling, Water Resour. Res., 33, 747-757, doi:10.1029/96WR03848, 1997.

Mann, H.B.: Nonparametric tests against trend, Econometrica, 13, 245-259, available at http://www.jstor.org/stable/1907187, 1945.

Milly, P.C.D., Betancourt, J., Falkenmark, M., Hirsch, R.M., Kundzewicz, Z.W., Lettenmaier, D.P., and Stouffer, R.J.: Stationarity is dead: whither water management?, Science, 319, 573-574, doi:10.1126/science.1151915, 2008.

Milly, P.C.D., Betancourt, J., Falkenmark, M., Hirsch, R.M., Kundzewicz, Z.W., Lettenmaier, D.P., Stouffer, R.J., Dettinger, M.D., and Krysanova, V.: On Critiques of "Stationarity is Dead: Whither Water Management?", Water Resour. Res., 51, 7785-7789, doi:10.1002/2015WR017408, 2015.

Montanari, A., and Koutsoyiannis, D.: Modeling and mitigating natural hazards: Stationarity is immortal!, Water Resour. Res., 50, 9748-9756, doi:10.1002/2014WR016092, 2014.

Olsen, J.R., Lambert, J.H., and Haimes, Y.Y.: Risk of extreme events under nonstationary conditions, Risk Anal., 18, 497-510, doi:10.1111/j.1539-6924.1998.tb00364.x, 1998.

Önöz, B., and Bayazit, M.: Effect of the occurrence process of the peaks over threshold on the flood estimates, J. Hydrol., 244, 86-96, doi:10.1016/S0022-1694(01)00330-4, 2001.

Parey, S., Malek, F., Laurent, C., and Dacunha-Castelle, D.: Trends and climate evolution: statistical approach for very high temperatures in France, Climatic Change, 81, 331-352, doi:10.1007/s10584-006-9116-4, 2007.

Parey, S., Hoang, T.T.H., and Dacunha-Castelle, D.: Different ways to compute temperature

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-619, 2016
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Published: 9 December 2016

return levels in the climate change context, Environmetrics, 21, 698-718, doi:10.1002/env.1060, 2010.

Peters, G.P., Andrew, R.M., Boden, T., Canadell, J.G., Ciais, P., Quéré, C.L., Marland, G., Raupach, M.R., and Wilson, C.: The challenge to keep global warming below 2 °C, Nature Climate Change, 3, 4-6, doi:10.1038/nclimate1783, 2013.

Pickands, J.: Statistical inference using extreme order statistics, Ann. Stat., 3, 119-131, . available at http://www.jstor.org/stable/2958083, 1975.

Prosdocimi, I., Kjeldsen, T.R., and Miller, J.D.: Detection and attribution of urbanization effect on flood extremes using nonstationary flood frequency models, Water Resour. Res., 51, 4244-4262, doi:10.1002/2015WR017065, 2015.

Raff, D.A., Pruitt, T., and Brekke, L.D.: A framework for assessing flood frequency based on climate projection information, Hydrol. Earth Syst. Sci., 13, 2119-2136, doi:10.5194/hessd-6-2005-2009, 2009.

Rigby, R.A., and Stasinopoulos, D.M.: Generalized Additive Models for Location, Scale and Shape, J. Roy. Stat. Soc., 54, 507-554, doi:10.1111/j.1467-9876.2005.00510.x, 2005.

Rosbjerg, D.: Estimation in partial duration series with independent and dependent peak values, J. Hydrol., 76, 183-195, doi:10.1016/0022-1694(85)90098-8, 1985.

Saltelli, A.: Making best use of model evaluations to compute sensitivity indices. Comput. Phys. Commun. 145, 280-297, doi:10.1016/S0010-4655(02)00280-1, 2002

Sobol', I.M.: Sensitivity analysis for nonlinear mathematical models. Math. Model. Comput. Exp. 1, 407-414, 1993.

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

Schwarz, G.: Estimating the dimension of a model. Ann. Statist. 6, 461-464, doi: 10.1214/aos/1176344136, 1978.

Sen, P.K.: Estimates of the regression coefficient based on Kendall's tau, J. Am. Stat. Assoc., 63, 1379-1389, doi:10.1080/01621459.1968.10480934, 1968.

Serinaldi, F., and Kilsby, C.G.: Stationarity is undead: uncertainty dominates the distribution of extremes, Adv. Water Resour., 77, 17-36, doi:10.1016/j.advwatres.2014.12.013, 2015.

Shane, R.M., and Lynn, W.R.: Mathematical model for flood risk evaluation, J. Hydraul. Div. Am. Soc. Civ. Eng., 90, 4119-4122, 1964.

Silva, A.T., Naghettini, M. and Portela, M.M.: On some aspects of peaks-over-threshold modeling of floods under nonstationarity using climate covariates, Stoch. Environ. Res. Risk Assess., 1-18, doi:10.1007/s00477-015-1072-y, 2015.

Sivapalan, M., and Samuel, J.M.: Transcending limitations of stationarity and the return period: process-based approach to flood estimation and risk assessment, Hydrol. Process., 23, 1671-1675, doi:10.1002/hyp.7292, 2009.

Stedinger, J.R., and Griffis, V.W.: Getting from here to where? Flood frequency analysis and climate, J. Am. Water Resour. Assoc., 47, 506-513, doi:10.1111/j.1752-1688.2011.00545.x, 2011.

Tang, Y., Reed, P., Wagener, T., and van Werkhoven, K.: Comparing sensitivity analysis methods to advance lumped watershed model identification and evaluation, Hydrol. Earth Syst. Sci., 11, 793-817, doi:10.5194/hess-11-793-2007, 2007.

USWRC: Guidelines for Determining Flood Flow Frequency, United States Water Resources Committee, Washington, DC, USA, 1982.

Villarini, G., Serinaldi, F., Smith, J.A., and Krajewski, W.F.: On the stationarity of annual flood peaks in the continental United States during the 20th century, Water Resour. Res., 45, W08417, doi:10.1029/2008WR007645, 2009a.

Villarini, G., Smith, J.A., Serinaldi, F., Bales, J., Bates, P.D., and Krajewski, W.F.: Flood frequency analysis for nonstationary annual peak records in an urban drainage basin, Adv. Water Resour, 32, 1255-1266, doi:10.1016/j.advwatres.2009.05.003, 2009b.

Villarini, G., Smith, J.A., Serinaldi, F., Ntelekos, A.A., and Schwarz, U.: Analyses of extreme flooding in Austria over the period 1951-2006, Int. J. Climatol., 32, 1178-1192, doi:10.1002/joc.2331, 2012.

von Storch, H.: Misuses of statistical analysis in climate research, in Analysis of climate variability: applications of statistical techniques, edited by von Storch, H., and Navarra, A., Springer-Verlag, Berlin, Germany, 11-26, 1995.

Wigley, T.M.L.: The effect of changing climate on the frequency of absolute extreme events, Climatic Change, 97, 67-76, doi:10.1007/s10584-009-9654-7, 2009.

Wilby, R.L., Dawson, C.W., and Barrow, E.M.: SDSM - a decision support tool for the assessment of regional climate change impacts, Environ. Modell. Softw., 17, 145-157, doi:10.1016/S1364-8152(01)00060-3, 2002.

Xiong, L., Du, T., Xu, C.-Y., Guo, S., Jiang, C., and Gippel, C.J.: Non-stationary annual maximum flood frequency analysis using the norming constants method to consider non-stationarity in the annual daily flow series, Water Resour. Manag., 29, 3615-3633, doi:10.1007/s11269-015-1019-6, 2015a.

Xiong, L., Jiang, C., Xu, C.-Y., Yu, K.-X., and Guo, S.: A framework of change-point detection for multivariate hydrological series, Water Resour. Res., 51, 8198-8217, doi:10.1002/2015WR017677, 2015b.

Xiong, L., Jiang, C., and Du, T.: Statistical attribution analysis of the nonstationarity of the annual runoff series of the Weihe River, Water Sci. Technol., 70, 939-946, doi:10.2166/wst.2014.322, 2014.

Yue, S., Pilon, P., Phinney, B., and Cavadias, G.: The influence of autocorrelation on the ability to detect trend in hydrological series, Hydrol. Process., 16, 1807-1829, doi:10.1002/hyp.1095, 2002.

Zuo, D., Xu, Z., Yang, H., and Liu, X.: Spatiotemporal variations and abrupt changes of potential evapotranspiration and its sensitivity to key meteorological variables in the Wei River basin, China, Hydrol. Process., 26, 1149-1160, doi:10.1002/hyp.8206, 2012.

Zuo, D., Xu, Z., Wu, W. Zhao, J., Zhao, F.: Identification of Streamflow Response to Climate Change and Human Activities in the Wei River Basin, China, Water Resour. Manage., 28, 833, doi:10.1007/s11269-014-0519-0, 2014.

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-619, 2016
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Published: 9 December 2016

Hydrology and Earth System Sciences
Open Access
Discussions

Tables

Table 1. Summary of the probability distribution functions and parametric link functions $g(\cdot \mid \mu_t, \sigma_t, \xi_t)$ for AM and POT floods.

Series	Distribution	Function	Link functions
AM	LNO3	$F(x\|\mu_t,\sigma_t,\xi_t) = \int_{\xi_t}^{x} \frac{1}{(x-\xi_t)\sigma_t\sqrt{2\pi}} \exp\left[-\left(\log(x-\xi_t)-\mu_t\right)^2 / 2\sigma_t^2\right] dx$ $\mu_t > 0, \sigma_t > 0, \xi_t > 0$	$g(\mu_t) = \mu_t$ $g(\sigma_t) = \ln(\sigma_t)$ $g(\xi_t) = \xi_t$
	LP3	$F(x\|\mu_t,\sigma_t,\xi_t) = \int_{\exp(\mu_t - \frac{\mu_t\sigma_t}{\xi_t})}^{x} f(x\|\mu_t,\sigma_t,\xi_t) dx$ $f(x\|\mu_t,\sigma_t,\xi_t) = \frac{\exp\left[-\left(\frac{\ln(x)-\mu_t}{\mu_t\sigma_t\xi_t} + \frac{1}{\xi_t^2}\right)\right]}{x\sigma_t \left\|\mu_t\xi_t\right\| \Gamma(1/\xi_t^2)} \left(\frac{\ln(x)-\mu_t}{\mu_t\sigma_t\xi_t} + \frac{1}{\xi_t^2}\right)^{\frac{1}{\xi_t^2}-1}$ $\sigma_t > 0, \xi_t \neq 0, \frac{\ln(x)-\mu_t}{\mu_t\sigma_t\xi_t} + \frac{1}{\xi_t^2} \geq 0$	$g(\mu_t) = \mu_t$ $g(\sigma_t) = \ln(\sigma_t)$ $g(\xi_t) = \xi_t$
	GEV	$F(x\|\mu_t,\sigma_t,\xi_t) = \begin{cases} \exp\left\{-[1+\xi_t(x-\mu_t)/\sigma_t]^{-1/\xi_t}\right\} & \xi_t \neq 0 \\ \exp\left\{-\exp[-(x-\mu_t)/\sigma_t]\right\} & \xi_t = 0 \end{cases}$ $1+\xi_t(x-\mu_t)/\sigma_t > 0, -\infty < \mu_t < \infty, \sigma_t > 0, -\infty < \xi_t < \infty$	$g(\mu_t) = \mu_t$ $g(\sigma_t) = \ln(\sigma_t)$ $g(\xi_t) = \xi_t$
POT	GP	$H(x \mid \sigma_t,\xi_t,u) = \begin{cases} 1-[1+\xi_t(x-u)/\sigma_t]^{-1/\xi_t} & \xi_t \neq 0 \\ 1-\exp[-(x-u)/\sigma_t] & \xi_t = 0 \end{cases}$ $1+\xi_t(x-\mu_t)/\sigma_t > 0, -\infty < \mu_t < \infty, \sigma_t > 0, -\infty < \xi_t < \infty$	$g(\sigma_t) = \ln(\sigma_t)$ $g(\xi_t) = \xi_t$
POT_AR	Poisson	$\Pr(M_t = k\|\mu_t) = \frac{\mu_t^k}{k!}\exp(-\mu_t); \quad k = 0,1,2,\ldots$ $E(M_t) = Var(M_t) = \mu_t$	$g(\mu_t) = \ln(\mu_t)$
	NB	$\Pr(M_t = k\|\mu_t,\sigma_t) = \frac{(\mu_t\sigma_t)^k \Gamma(k+1/\sigma_t)}{\Gamma(1/\sigma_t)\Gamma(k+1)} \left(\frac{1}{1+\mu_t\sigma_t}\right)^{k+1/\sigma_t}$ $k = 0,1,2,\ldots$ $E(M_t) = \mu_t;\ Var(M_t) = \mu_t + \mu_t^2\sigma_t$	$g(\mu_t) = \ln(\mu_t)$ $g(\sigma_t) = \ln(\sigma_t)$

Table 2. Temporal trends tested (↑ for increase and ↓ for decrease) by the statistic of the MK (Z_{MK}), PW (Z_{PW}), and TFPW (Z_{TFPW}). The PMK test discerns the potential drivers (*p*-values highlighted in bold) for the series with significant trends.

Series	Z_{MK}	Z_{PW}	Z_{TFPW}	Potential drivers				
				P_{total}	P_{max_1d}	P_{max_3d}	P_{max_7d}	T_{mean}
AM	-3.72(↓)	-3.61((↓)	-3.58(↓)	**0.07**	0.00	0.00	0.01	**0.15**
POT2	-0.96	-0.76	-0.84			(-)		
POT3	-1.92	-1.55	-1.86			(-)		
POT4	-1.49	-1.38	-1.47			(-)		
POT_AR2	-4.12(↓)	-3.19(↓)	-4.26(↓)	**0.09**	0.00	0.00	0.03	**0.11**
POT_AR3	-4.00(↓)	-2. 72(↓)	-3.80(↓)	**0.08**	0.00	0.01	0.01	0.04
POT_AR4	-3.35(↓)	-2.52(↓)	-3.73(↓)	0.03	0.00	0.01	0.02	**0.06**

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-619, 2016
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Published: 9 December 2016

Hydrology and Earth System Sciences
Discussions
Open Access

Table 3. Optimal models for AM under stationarity and nonstationarity.

Model	Estimated parametric functions (standard error)	AIC BIC	R^2_{PP} (%)	R^2_{QQ} (%)
Stationarity				
LNO3	$\mu_t = 7.564$ (0.086) $\ln(\sigma_t) = -0.495$ (0.100) $\xi_t = 110.032$ (4.199)	854.8 860.5	92.1	90.4
LP3	$\mu_t = 7.629$ (0.088) $\ln(\sigma_t) = -2.489$ (0.176) $\xi_t = -0.588$ (0.241)	852.9 858.6	92.5	91.1
GEV	$\mu_t = 1805.973$ (189.035) $\ln(\sigma_t) = 6.906$ (0.162) $\xi_t = 0.053$ (0.016)	857.4 863.1	96.7	92.4
Nonstationarity				
LNO3	$\ln(\mu_t) = 9.424 + 0.003P_{total} - 0.364T_{mean}$ (1.321, 0.001, 0.101) $\ln(\sigma_t) = 1.151 - 0.003P_{total}$ (0.136, 0.001) $\xi_t = 110.002$ (4.209)	833.8 845.3	86.3	92.6
LP3	$\mu_t = 8.741 + 0.003P_{total} - 0.307T_{mean}$ (1.149, 0.001, 0.102) $\ln(\sigma_t) = -2.740$ (0.102) $\xi_t = 0.451$ (0.181)	820.9 830.5	87.7	90.8
GEV	$\mu_t = 1789.594 + 3.818P_{total} - 215.657T_{mean}$ (2161.661, 1.717, 230.536) $\ln(\sigma_t) = 9.736 - 0.336T_{mean}$ (4.102, 0.151) $\xi_t = 0.108$ (0.048)	832.3 843.8	88.4	97.9

Table 4. Optimal GP models for the POT magnitudes.

POT magnitudes	POT2	POT3	POT4
Threshold (u)	1060	780	530
$\ln(\sigma_t)$ (standard error)	7.03(0.163)	6.76(0.149)	6.56(0.114)
ξ_t (standard error)	0.11(0.049)	0.13(0.062)	0.10(0.043)
AIC	1611.3	2371.9	3067.5
BIC	1616.5	2377.9	3074.1
R^2_{PP} (%)	99.5	98.7	99.7
R^2_{QQ} (%)	97.9	98.2	99.4

Table 5. Optimal nonstationary Poisson and NB models for the POT arrival rates.

Distribution	Estimated parametric functions (standard error)	AIC	BIC	R^2_{PP} (%)	R^2_{QQ} (%)
Poisson					
POT_AR2	$\ln(\mu_t) = 0.005P_{total} - 0.241T_{mean}$ (0.001,0.048)	149.9	153.6	90.2	89.7
POT_AR3	$\ln(\mu_t) = -1.088 + 0.004P_{total}$ (0.315,0.001)	171.1	174.9	90.4	90.1
POT_AR4	$\ln(\mu_t) = 0.004P_{total} - 0.074T_{mean}$ (0.001,0.033)	181.8	185.6	92.7	90.2
NB					
POT_AR2	$\ln(\mu_t) = 2.791 + 0.004P_{total} - 0.480T_{mean}$ (0.567,0.001) $\ln(\sigma_t) = -0.013P_{total}$ (0.007)	143.8	151.3	92.4	91.7
POT_AR3	$\ln(\mu_t) = 0.004P_{total} - 0.113T_{mean}$ (0.001,0.042) $\ln(\sigma_t) = -5.06$ (0.302)	171.2	177.0	90.0	86.9
POT_AR4	$\ln(\mu_t) = 2.327 + 0.003P_{total} - 0.276T_{mean}$ (1.023,0.001,0.132) $\ln(\sigma_t) = -6.07$ (0.498)	183.3	190.9	89.5	87.8

Figures

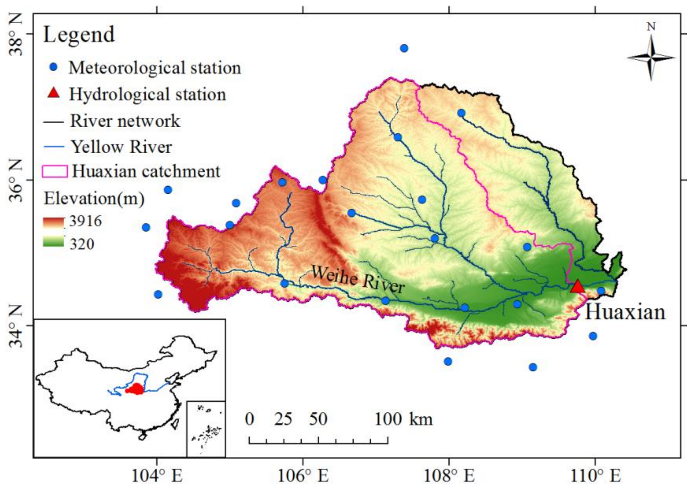

Figure 1. Geographic positions of the hydrological and meteorological stations in the Weihe basin.

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-619, 2016
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Published: 9 December 2016

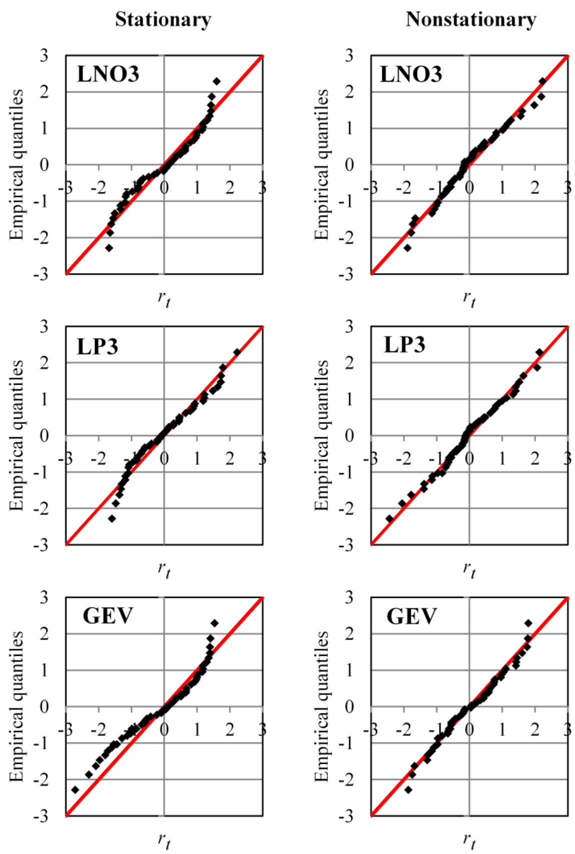

Figure 2. Q-Q plots of standard normal quantiles r_t against empirical quantiles for the models in Table 3.

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-619, 2016
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Published: 9 December 2016

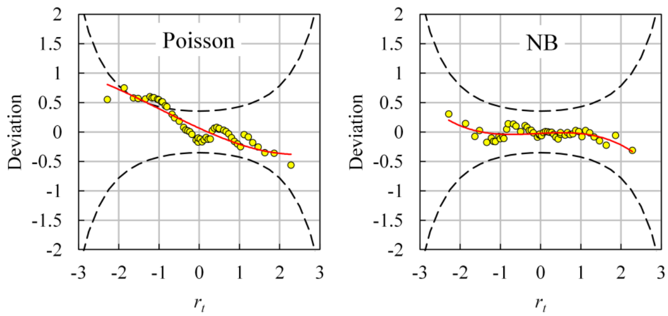

Figure 3. Worm plots of nonstationary Poisson and NB models for POT_AR2 in Table 5.

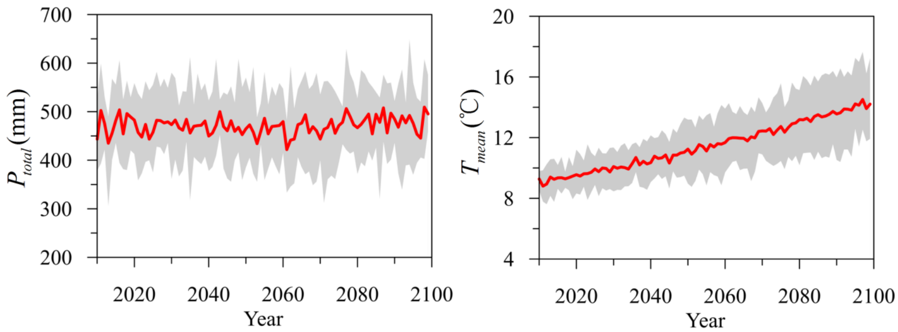

Figure 4. Projected series of P_{total} and T_{mean} for the future period of 2010-2099 averaged from the seven GCMs (red lines) with their ranges shown by gray shadows.

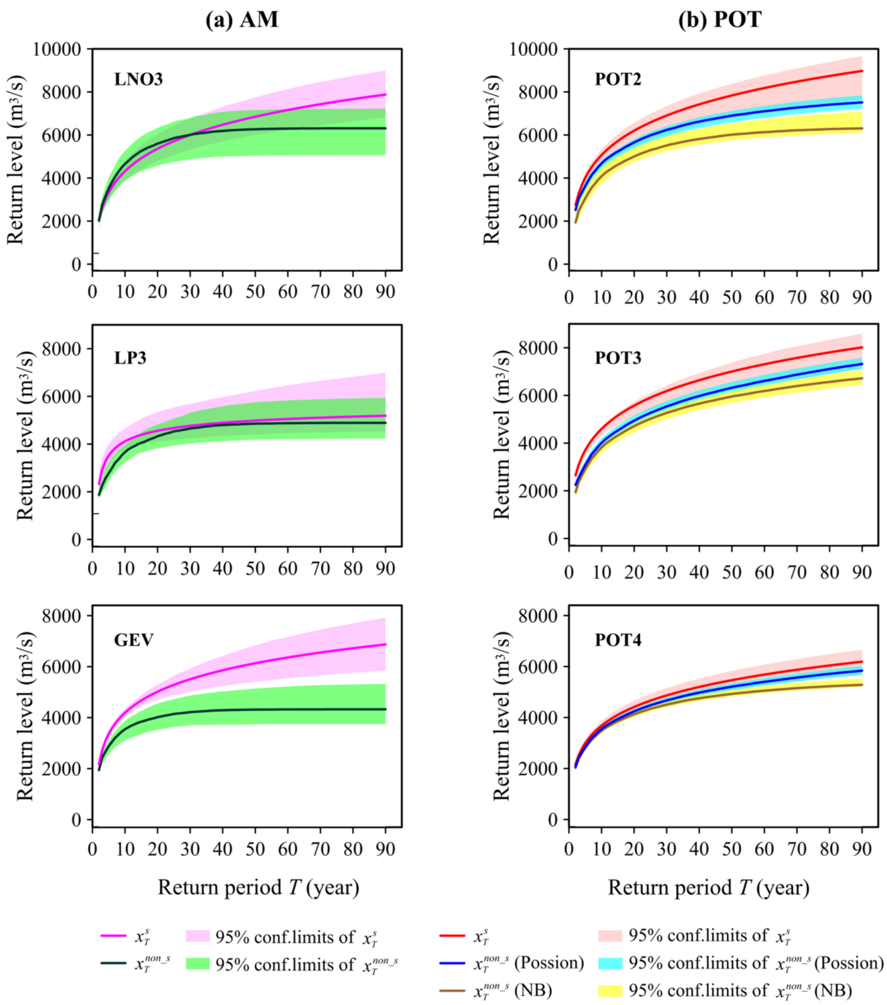

Figure 5. Flood return levels x_T estimated from the models for AM in Table 3(left) and for POT with the results of Tables 4 and 5 (right).

Hydrology and Earth System Sciences
Open Access
Discussions

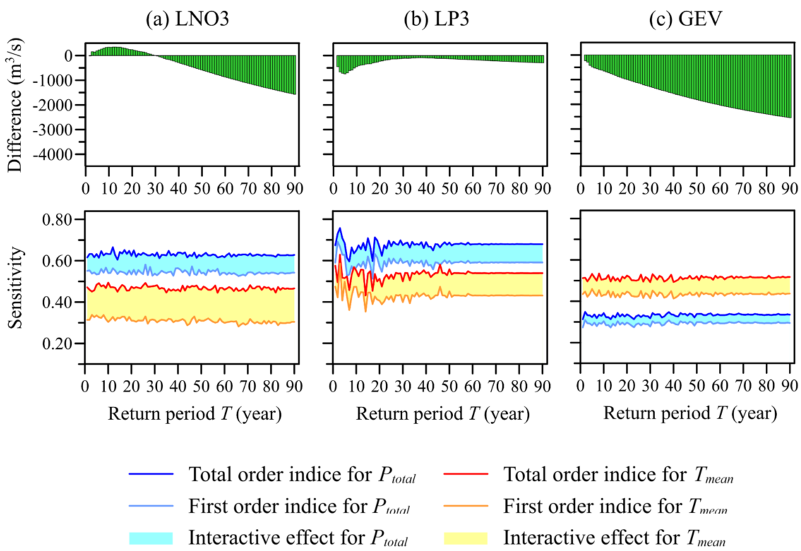

Figure 6. Difference in T-year AM-based return levels between stationarity and nonstationarity ($x_T^{non\_s}$ - x_T^s) in case of LNO3, LP3, and GEV models, respectively, and the Sobol' sensitivity indices of parameters P_{total} and T_{mean} for each model.

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-619, 2016
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Published: 9 December 2016

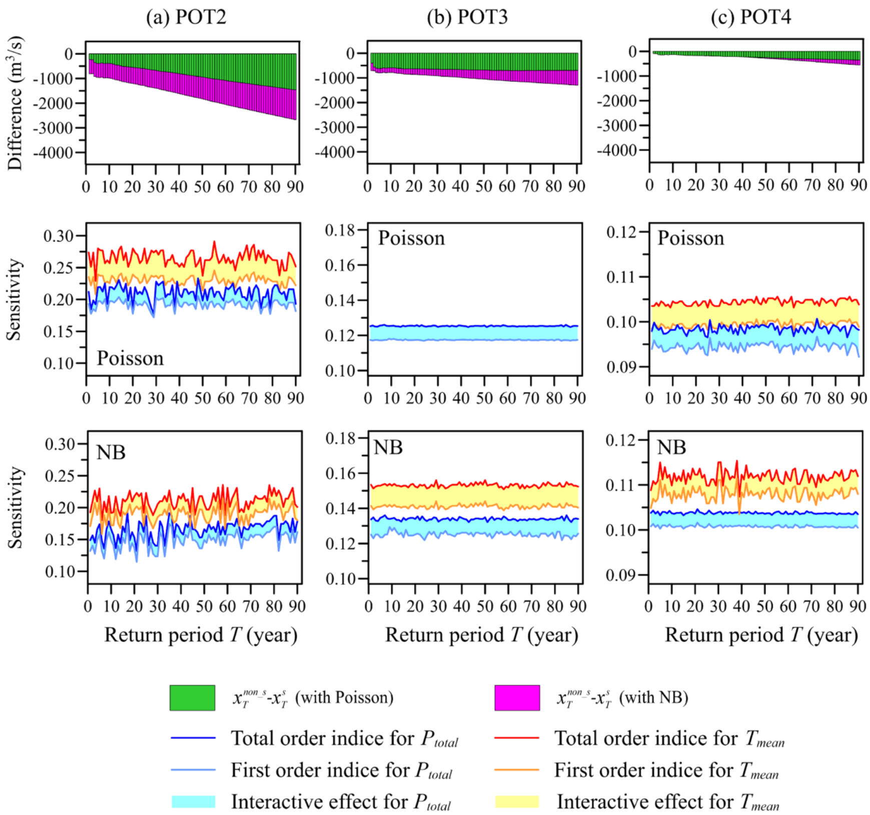

Figure 7. Difference in T-year POT-based return levels between stationarity and nonstationarity ($x_T^{non\_s}$ - x_T^s) when using the Poisson and NB models, respectively, and the Sobol' sensitivity indices of parameters P_{total} and T_{mean} for each model.

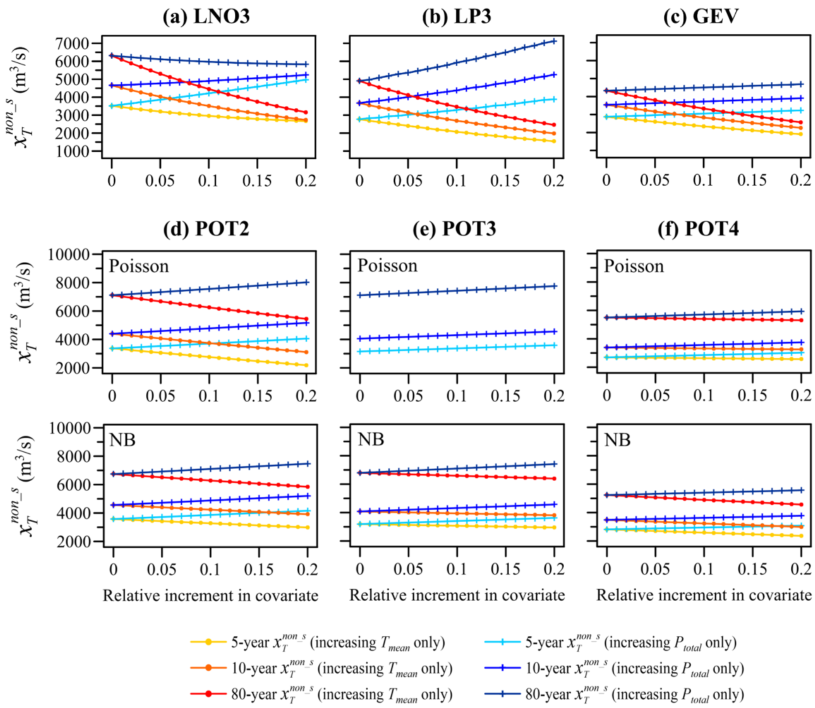

Figure 8. Variation of return level $x_T^{non\_s}$ (T = 5, 10, 80) computed with a separate increase of 0-20% in series P_{total} and T_{mean}, respectively, for AM (with nonstationary LNO3, LP3, and GEV distributions) and POT models (using the Poisson and NB distribution).