

Estimating the design flood under the influence of check dams by removing nonstationarity from the flood peak discharge series

Shi Li, Yi Qin, Yixiu Liu, Xiaoyu Song, Qiang Liu and Ziwen Li

ABSTRACT

The construction of check dams in northwestern China has resulted in nonstationary changes in flood peak discharge series; the stationary assumption of the conventional hydrological frequency analysis is no longer satisfied. According to the characteristics of the construction and operation of check dams, the nonstationarity of flood peak discharge series are largely induced by changes in the effective runoff generation area (i.e., the basin area minus the area controlled by check dams). Knowing the power function relationship between the flood peak discharge and the basin area, we can remove the influence of the effective runoff generation area and convert the original nonstationary series into a stationary series. This de-nonstationarity method can achieve stationarity in the first and second moments simultaneously. Therefore, we can calculate the design value of the reconstructed series using the conventional frequency analysis method. According to the effective runoff generation area under design conditions, we can then obtain the corresponding design flood of the original series. We applied this method to the Mahuyu River basin to obtain the design flood under nonstationarity. Due to the consideration of the deterministic influence of check dams during the de-nonstationarity process, the uncertainty analyzed by the bootstrap method is obviously small.

Key words | check dam, de-nonstationarity, design flood, frequency analysis, nonstationary, reconstruction

Shi Li
Yi Qin (corresponding author)
Xiaoyu Song
 State Key Laboratory of Eco-hydraulics in
 Northwest Arid Region of China,
 Xi'an University of Technology,
 Xi'an,
 China
 E-mail: qinyi@xaut.edu.cn

Yixiu Liu
 Powerchina Northwest Engineering Corporation
 Limited,
 Xi'an,
 China

Qiang Liu
Ziwen Li
 College of Water Conservancy and Civil
 Engineering,
 Shandong Agricultural University,
 Taian,
 China

HIGHLIGHTS

- We proposed a de-nonstationary method to transform nonstationary flood peak discharge series into a stationary series.
- We inferred that nonstationarity in the flood peak discharge series is caused by changes in the effective runoff generation area in the presence of check dams.
- The nonstationarity is removed according to the power function relationship between the flood peak discharge and the basin area.
- This de-nonstationarity method can achieve stationarity in the first and second moments simultaneously.

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

doi: 10.2166/nh.2020.050

INTRODUCTION

Design flood estimation plays an important role in water project design, water resources planning, and flood risk control. It is usually obtained using frequency analysis which is a technique of fitting a probability distribution to a series of observations for defining the probabilities of future occurrences of some events of interest, e.g., an estimate of a flood magnitude corresponding to a chosen risk of failure (Khalil *et al.* 2006). The conventional hydrologic frequency analysis method requires the hydrological series to be independent and identically distributed (Salas & Obeysekera 2014; Qin *et al.* 2015; Read & Vogel 2015). For the flood peak discharge Q with distribution $F_Q(Q|\theta)$, we generally assume that the series is stationary, and that the distribution parameters θ are time invariant. In northwestern China, climate and topographical conditions cause serious water and soil loss. As such, many check dams have been built to try to mitigate this phenomenon. In the context of the construction of thousands of check dams, the stationarity of the flood peak discharge series gradually disappears. The nonstationary flood frequency analysis has become a point of concern for engineers and local managers.

Under nonstationary conditions, in the absence of understanding the physical mechanisms, we usually assume that the distribution parameters of the hydrological series change with time; the time-varying distribution is determined by fitting the time-varying characteristic of the distribution parameters from a statistical perspective. Based on this logic, Strupczewski *et al.* (2001a, 2001b) and Strupczewski & Kaczmarek (2001) proposed the time-varying moments method. The Generalized Additive Models for Location, Scale, and Shape (GAMLSS), which was proposed by Rigby & Stasinopoulos (2005), makes the application of time-varying moments method more flexible. Usually, the model uses time as the covariate so it can describe the trend variation in the hydrological time series; however, this is not sufficient to reflect the driving force of the nonstationary change. Therefore, a number of studies have analyzed the influence of climate change and anthropogenic activity on the nonstationary hydrological time series by using them as the covariates (Villarini *et al.* 2009; López & Francés 2013; Xiong *et al.* 2014; Zhang *et al.*

2015). Using the traditional definition of the return period, where the return period is equal to the reciprocal of the exceedance probability, results in unrealistic design values that are not appropriate for real-world applications because the design value varies from one year to the next (Yan *et al.* 2017).

In order to cope with nonstationarity, some studies have extended the concept of return period to a more general interpretation as the expected waiting time (EWT), which is the expected time interval between two exceedance events (Wigley 1988; Olsen *et al.* 1998). Alternatively, other works have employed the return period as the time period during which the expected number of exceedance (ENE) events is equal to 1 (Parey *et al.* 2007, 2010). Cooley (2013) and Yan *et al.* (2017) pointed out that the EWT definition requires indefinite extrapolation along the trend, while the ENE definition only requires extrapolation for a finite number of years. Based on the EWT or ENE definitions, we can obtain the unique design value for a nonstationary series with a given return period. However, under nonstationary conditions, different initial and end time of calculation period results in different hydrological design values; the EWT and ENE methods do not have a clear definition of the calculation period. In practice, many studies take the project's design lifetime as the calculation period (Rootzén & Katz 2013; Yan *et al.* 2017).

Some studies illustrate that the return period does not reflect the risk of failure of a project, and recommend to replace it with the reliability, which is the probability of no exceedance events occurring over the lifetime of the project (Rootzén & Katz 2013; Salas & Obeysekera 2014; Read & Vogel 2015). Reliability is a more effective index than the return period because it summarizes the joint probability instead of the average probability (Serinaldi & Kilsby 2015). On this basis, some reliability-based methods such as the 'design life level' (Rootzén & Katz 2013), 'equivalent reliability' (Hu *et al.* 2017), and 'average design life level' (Read & Vogel 2015) have been proposed successively.

Whether the return period or the reliability, we usually construct its expression according to the time-varying probability distribution, and then use it to calculate the design

value. However, it is impossible to accurately describe the characteristics of hydrological time series or to predict the influence degree of climate change or human activities in the future simply by pure statistical approaches unless you have a clear understanding of the physical mechanism. Even if meteorological or anthropogenic factors are chosen as the covariates, the nonstationary changes are usually characterized using data-based regression technique. This nonstationary change is usually fitted with linear or some simple nonlinear curves, but the actual conditions must be extremely complicated, and it is difficult to answer how this nonstationary change will evolve in the future. Besides, according to statistical theory, because the nonstationary hydrological series no longer has ergodicity, the population distribution cannot be inferred from it.

For a long time, we have tried to reformulate and extend the conventional hydrological frequency analysis method to adapt the nonstationary conditions. Since there is insufficient knowledge of the physical mechanism, for nonstationary frequency analysis, new challenges and doubts are encountered as one problem after another is solved. However, under the stationary conditions, the conventional frequency analysis method has been widely proved to be scientific and reasonable. The conventional method has been mastered by engineers, and there are a large number of guides and criteria to ensure its reliability. Therefore, the conventional frequency analysis method should not be discarded easily. If the stationarity of a hydrological series can be reconstructed, then the conventional method can still be used to calculate the design value.

For a basin that has a high number of check dams, we analyze the influence of the check dams on the nonstationary changes in the flood peak discharge series according to its construction and operation characteristics. According to this influence law, we try to remove the nonstationary influence of the check dams from the flood peak discharge series and transform it into a new stationary series. The de-nonstationarity idea is not new in hydrology. Previous studies are based on the theory of stochastic hydrology to divide the hydrological series into a deterministic component (trend, periodicity, etc.) and a stochastic component (Kottegoda 1980; Xie *et al.* 2005). After removing the deterministic component, the stationary stochastic component can be obtained. However, this simplification is not

accurate enough for the complex changes of hydrological series. The mechanism-based de-nonstationarity method proposed in this paper does not consider the type of nonstationarity, no matter what type of nonstationary changes of hydrological variable is due to the corresponding nonstationary change of its influence factors. The de-nonstationarity process is, indeed, the process of removing the nonstationary effects of these factors. Moreover, through theoretical derivation and practical application, we prove that the de-nonstationarity method proposed in this study can achieve stationarity in both the first and second moments of the flood peak discharge series. The resulting uncertainty of design flood, determined using the bootstrap method (Efron 1992), is small.

STUDY AREA AND DATA

The check dams constructed in northwestern China have largely alleviated the water and soil loss. There are more than 20,000 check dams in Yulin in the Shaanxi Province alone (Jiang & Shuai 2011). This kind of high-density water project leads to significant nonstationary changes in the flood peak discharge series, which poses a great challenge to the water project construction and flood risk management. We take the Mahuyu River Basin (MRB) in Yulin City, Shaanxi Province, China as an example to calculate the design flood in the presence of check dams. The Mahuyu River is located in the hinterland of the Loess Plateau, sourced from Shiyagou in Hengshan County, and flows from west to east into Wuding River. The MRB lies in the area bounded by 37°44'–37°55'N, 109°37'–110°02'E (Figure 1). The control station for the MRB is the Mahuyu hydrological station, which controls an area of 372 km². The MRB flood peak discharge series recorded at the Mahuyu station from 1962 to 2010 is shown in Figure 2. We employed the Mann–Kendall (M–K) test (Mann 1945; Kendall 1975) and the Breusch–Pagan (B–P) (Breusch & Pagan 1979) test to analyze the trends in the first and second moments, respectively (Table 1). Under the influence of check dams, both the first and second moments of the MRB flood peak discharge series exhibit nonstationary changes, where the nonstationarity in the second moment

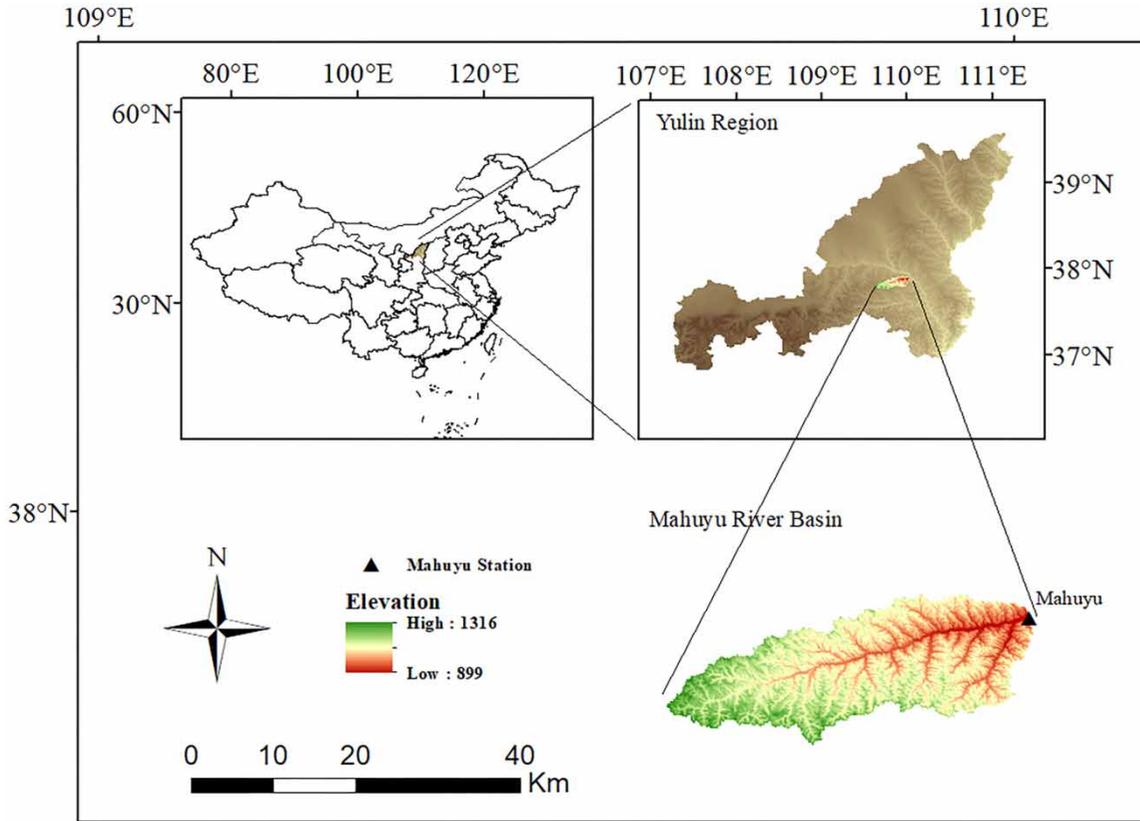


Figure 1 | Topographic map of the Mahuyu River Basin (MRB).

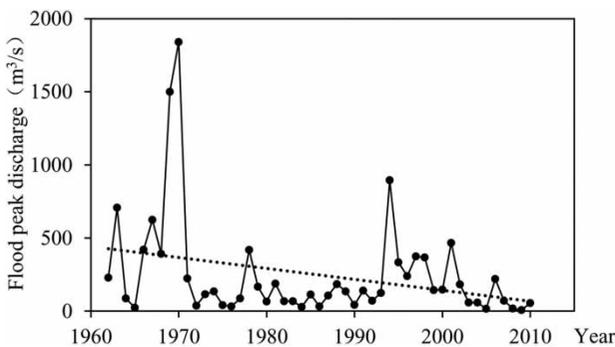


Figure 2 | Flood peak discharge series recorded at the Mahuyu Station.

Table 1 | Nonstationary analysis of the MRB flood peak discharge series

Hydrological series	Record period	M – K test for mean	B – P test for variance	Test result
Q(t)	1962 – 2009	-1.672*	16.881***	Nonstationary

***p < 0.01; *0.05 < p < 0.1.

is more significant. Except for the trend, we do not consider other types of nonstationarity in this study.

A check dam is generally composed of two parts: the dam and the horizontal pipe drainage structure. The slow drainage (about 1.0–1.5 m³/s) of the horizontal pipes in the check dam causes the sediment in the flood flow to be deposited in front of the dam, which helps mitigate the soil loss. The runoff in the control area of the check dams cannot be discharged to the outlet of the basin in time. That is to say, for a flood event, the runoff generated in the control area of the check dam provides little contribution to the flood at the outlet of the basin. Therefore, we infer that the nonstationary changes in the flood peak discharge series are essentially caused by the change in the control area of the check dams. During the flooding process, areas in the basin that are not controlled by the check dams still generate runoff effectively, so this part of the basin area is defined as the effective runoff generation area. Figure 3

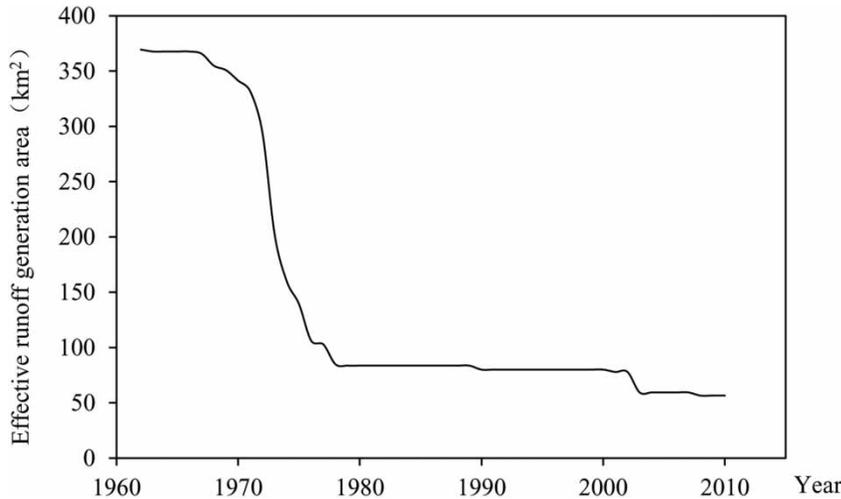


Figure 3 | MRB effective runoff generation area series.

shows the interannual variation of the effective runoff generation area in the MRB. Dozens of check dams in the MRB were built year by year from 1962 to 2010, and the effective runoff generation area also changed gradually.

METHODOLOGY

Removing nonstationarity from the flood peak discharge series

The presence of the check dams resulted in a reduced effective runoff generation area, which leads to nonstationary changes in the flood peak discharge series. After removing the influence of the effective runoff generation area from the original nonstationary series, the remaining part of the flood peak discharge series should be statistically stationary.

Numerous studies have shown that there is a power function relationship between the flood peak discharge and the basin area (Smith 1992; Gupta 2004). This power function relationship in the Yulin region takes the form of (Quan 2016):

$$Q \propto A^{0.73} \quad (1)$$

where A is the basin area. A more complete expression is as follows:

$$Q = A^{0.73} \cdot \xi \quad (2)$$

For the flood peak discharge series in a basin controlled by the check dams, changes in the effective runoff generation area are the main cause of nonstationarity of the series. Of course, other factors, such as slope, vegetation situation, and even some unrecognized factors may also influence the flood peak discharge. Most of these factors are stationary, while some nonstationary factors have only a secondary effect on flood peak discharge relative to the effective runoff generation area. Therefore, we summarized the influence of all these factors as ξ , and assume that it is a stationary random variable with a probability distribution characterized by the mean μ and the variance σ^2 . By rearranging Equation (2), we can remove the nonstationary influence of the effective runoff generation area from the flood peak discharge series. This reconstructed series $RS(t)$ is defined as follows:

$$RS(t) = \frac{Q(t)}{A^{0.73}(t)} \quad (3)$$

The mean and variance of the reconstructed series $RS(t)$ are as follows:

$$E[RS(t)] = E\left[\frac{Q(t)}{A^{0.73}(t)}\right] = E(\xi) = \mu \quad (4)$$

$$D[RS(t)] = D\left[\frac{Q(t)}{A^{0.73}(t)}\right] = D(\xi) = \sigma^2 \quad (5)$$

The mean and variance of the reconstructed series $RS(t)$ are time-invariant constant. That is to say, in the new series $RS(t)$, both the first moment and the second moment are stationary. The simultaneous stationarity of the first and second moments guarantees the reliability of the design flood calculated according to the reconstructed series. It is easy to obtain the design quantile of the reconstructed series $RS(t)$ using the conventional hydrological frequency analysis method. Based on Equation (2), the design value for the original nonstationary flood peak discharge series is as follows:

$$Q_p = rs_p \cdot A_p^{0.73} \quad (6)$$

where rs_p is the design value of the reconstructed series $RS(t)$ and A_p is the effective runoff generation area under design conditions.

Uncertainty analysis of the design flood using the bootstrap method

For hydrological frequency analysis, using sample series to estimate the population distribution will inevitably produce uncertainty. However, because the de-nonstationarity method combines the conventional frequency analysis with the deterministic influence of the check dams, the design flood will have a smaller uncertainty than those calculated using purely statistical approaches. On the other hand, because the nonstationary hydrological series no longer has ergodicity, we cannot actually estimate the population distribution according to the nonstationary sample series. The de-nonstationarity method transforms the nonstationary series into a stationary reconstructed series with ergodicity. However, with potential observation errors and a low sample size, it is still necessary to assess the design value uncertainty.

The bootstrap method proposed by Efron (1992) has been widely used in the quantitative analysis of uncertainty. Serinaldi & Kilsby (2015) pointed out that the bootstrap method strictly depends on the available information without any asymptotic assumptions, and that it does not rely on specific parameter estimation methods. Moreover, this method is relatively easy to implement, regardless of the

complexity of the model. We use the bootstrap method to estimate the confidence intervals of the design flood. The procedure of the bootstrap method is as follows:

1. After removing the nonstationarity from the original flood peak discharge series $Q(t)$, we apply the conventional frequency analysis method to stationary reconstructed series $RS(t)$ and calculate the design value rs_p for a certain return period.
2. The design value Q_p of the original nonstationary flood peak discharge series is calculated with Equation (6).
3. We then resample the reconstructed series $RS(t)$ with replacements to obtain a new sample series $RS^*(t)$.
4. Based on the new sample series $RS^*(t)$, we calculate the new design value rs_p^* for a specific return period.
5. Using Equation (6), we determine the design value Q_p^* of the nonstationary flood peak discharge series for a specific return period.
6. We then repeat the resampling process N times to obtain N design values: $Q_{p,i}^*$, $i = 1, 2, \dots, N$.
7. For a certain significance level α , the confidence interval of the design value Q_p is $(Q_{p,N\alpha/2}^*, Q_{p,N(1-\alpha/2)}^*)$.

RESULTS AND DISCUSSION

Stationary analysis of reconstructed series

After removing the nonstationary influence of the effective runoff generation area, we obtained the stationary reconstructed series $RS(t)$ (Figure 4). The hypothesis test results show that both the first and second moments of the reconstructed series of the MRB flood peak discharge are stationary (Table 2). Because the reconstructed series is stationary, we can use the conventional hydrological frequency analysis method to calculate the design value. The presence of the check dams drives changes in both the magnitude of the flood peak discharge and the rate of change relative to the mean, which reflects the nonlinear characteristics of the hydrological system. Because the statistics of B-P test decrease more significantly after removing the nonstationary effects of the effective runoff generation area from $Q(t)$, we infer that the changes in the effective runoff generation area have a more significant impact on the second

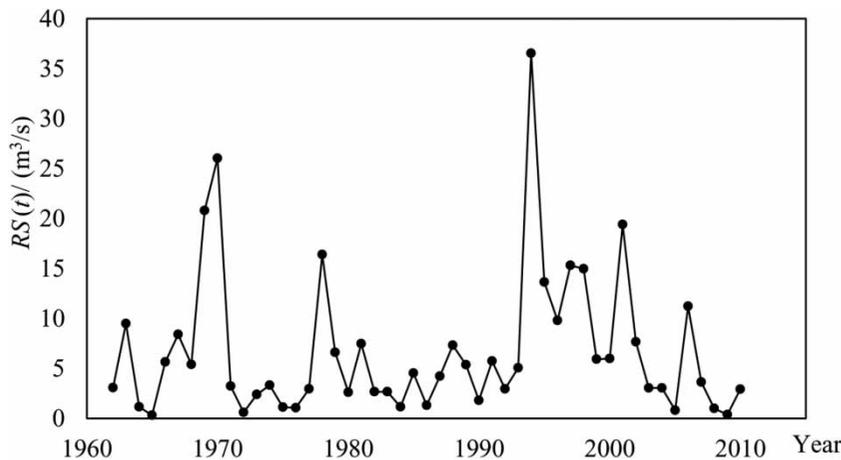


Figure 4 | $RS(t)$ series for the MRB flood peak discharge after removing the nonstationary influence of the effective runoff generation area.

Table 2 | Nonstationary analysis of the MRB $RS(t)$ series

Hydrological series	M-K test for mean	B-P test for variance	Test result
Q	-1.672	16.881	Nonstationary
$RS(t)$	0.483	0.095	Stationary

moment of $Q(t)$ than they do on the first moment. Therefore, in the hydrological frequency analysis, we should account for nonstationarity in both the first moment and the second moment.

The effect of removing the nonstationary influence of the effective runoff generation area by formula (3) is similar to a kind of unitization, that is, the reconstructed series $RS(t)$ is a rough approximation of the flood peak discharge per unit area; when the exponent in that expression is equal to 1, $RS(t)$ represents the flood peak modulus series. The construction of check dams reduces the actual runoff generation area. However, in regions that are beyond the control area of the check dams, the runoff generation process for a flood event does not change, resulting in the flood peak modulus within the effective runoff generation area that is unaffected by the presence or absence of the check dams. Therefore, it is reasonable that the reconstructed series $RS(t)$, which approximates the flood peak modulus, is stationary after the influence of the effective runoff generation area has been removed.

The nonstationary change of the flood peak discharge series and the effective runoff generation area in MRB can

also be characterized as abrupt changes that occurred in the 1970s. The existing nonstationary hydrological frequency analysis methods, which follow the time-varying characteristics of the hydrological series, are always limited to a simple type of nonstationary change in fitting these nonstationary changes, but the actual conditions must be extremely complicated. In fact, nonstationary changes in a hydrological series are caused by the corresponding nonstationary changes in certain influencing factors. No matter what type of nonstationarity, it can be removed from the hydrological series using the de-nonstationarity method. To cope with nonstationarity, it is necessary to identify the nonstationary influence factors and their corresponding physical mechanisms.

Frequency analysis of the nonstationary flood peak discharge series

Determination of the distribution type of the reconstructed series

In this paper, the frequency analysis of the reconstructed series of flood peak discharge (1962–2010) of Mahuyu hydrological station at the outlet of the MRB was carried out. Four distributions, Weibull (WEI), Gumble (GU), General extreme value (GEV), and Pearson type III (P-III), are selected as the alternative distributions in this study (Table 3).

In order to select the best-fitting distribution, we evaluated the overall simulation accuracy of the alternative

Table 3 | Probability density functions of the alternative distributions

Distribution	Probability density function
WEI	$f(x) = \frac{\kappa}{\sigma} \left(\frac{x-\mu}{\sigma}\right)^{\kappa-1} \exp\left[-\left(\frac{x-\mu}{\sigma}\right)^\kappa\right], x > 0, \mu > 0, \sigma > 0, -\infty < \kappa < \infty$
GU	$f(x) = \frac{1}{\sigma} \exp\left\{-\left(\frac{x-\mu}{\sigma}\right) - \exp\left[-\left(\frac{x-\mu}{\sigma}\right)\right]\right\}, -\infty < x < \infty, -\infty < \mu(\infty, \sigma)0$
GEV	$f(x) = \frac{1}{\sigma} \left[1 + \kappa\left(\frac{x-\mu}{\sigma}\right)\right]^{-\frac{1}{\kappa}-1} \exp\left\{-\left[1 + \kappa\left(\frac{x-\mu}{\sigma}\right)\right]^{-\frac{1}{\kappa}}\right\}, -\infty < x < \infty, -\infty < \mu(\infty, \sigma)0, -\infty < \kappa < \infty$
P-III	$f(x) = \frac{1}{\sigma \mu\kappa \Gamma(1/\kappa^2)} \left(\frac{x-\mu}{\mu\sigma\kappa} + \frac{1}{\kappa^2}\right)^{\frac{1}{\kappa^2}-1} \exp\left[-\left(\frac{x-\mu}{\mu\sigma\kappa} + \frac{1}{\kappa^2}\right)\right], -\infty < x(\infty, \sigma)0, \kappa \neq 0, \frac{x-\mu}{\mu\sigma\kappa} + \frac{1}{\kappa^2} \geq 0$

distributions with the Kolmogorov–Smirnov (K–S) test (Kolmogorov-Smirnov et al. 1933), the Nash–Sutcliffe efficiency NSE_{pp} (NSE_{QQ}) (Nash & Sutcliffe 1970) between the theoretical and empirical distributions (quantiles), and the root mean square error (RMSE). The goodness of fit of the four alternative distributions of the reconstructed MRB flood peak discharge series is shown in Table 4.

For stationary series $RS(t)$, the K–S test statistic and NSE_{pp} show that the GEV distribution is optimal, while NSE_{QQ} and RMSE indicate that the WEI distribution is a better fit. In general, the difference between the empirical probability and the theoretical probability for the same sample point at the tail of the distribution is small (both close to 1), but the difference between the empirical and theoretical quantile is relatively larger. Therefore, NSE_{QQ} is more sensitive to the tail than NSE_{pp} . We conclude that the WEI distribution is the optimal distribution for the reconstructed series of the MRB flood peak discharge. The Q–Q plot in Figure 5 also shows the good performance of the optimal distribution.

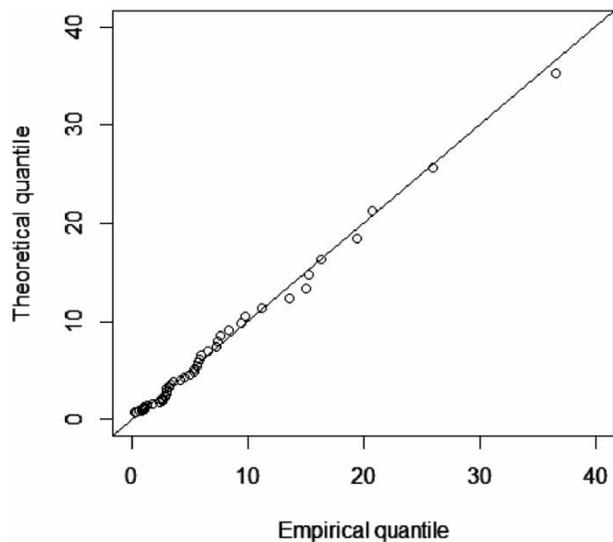


Figure 5 | Q–Q plot of the theoretical and empirical quantiles for the reconstructed series of the MRB flood peak discharge.

Table 4 | Goodness of fit of the four alternative distributions of the reconstructed series of the MRB flood peak discharge

Hydrological series	Alternative distribution	K–S test D_n	NSE_{pp}	NSE_{QQ}	RMSE
$RS(t)$	WEI	0.114	0.984	0.994	0.539
	GU	0.149	0.914	0.869	2.250
	GEV	0.083	0.990	0.982	0.929
	P-III	0.125	0.979	0.993	0.601

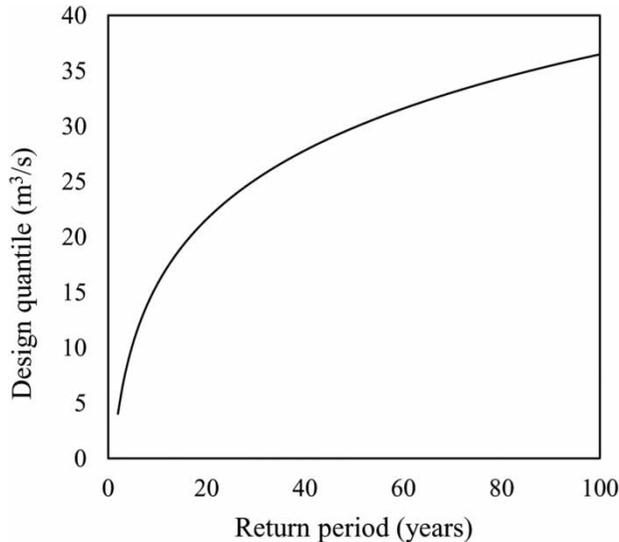
Bold indicates the best fitting.

Calculation of the design quantiles

We estimated the distribution parameters using the L-moments method (Hosking 1990). The L-moment is a linear combination of the probability weighted moments. Compared with the regular moment estimation method, the L-moments method is less affected by individual data points. After removing the influence of the effective runoff generation area from the flood peak discharge series, we estimated the distribution parameters of the reconstructed stationary series (Table 5) and calculated the design quantiles for a certain return period (Figure 6).

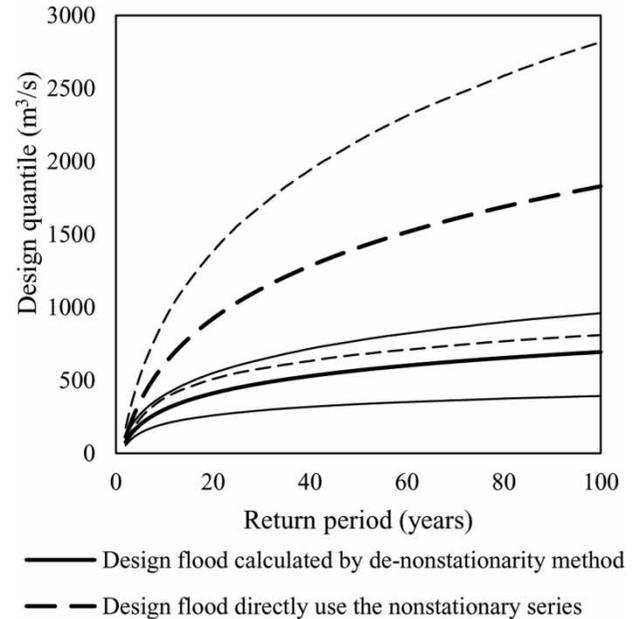
Table 5 | Parameter estimation of the optimal distribution for the reconstructed series of the MRB flood peak discharge

Hydrological series	Optimal distribution	Parameter estimation
$RS(t)$	WEI	$\mu = 0.699, \sigma = 5.290, \kappa = 0.799$

**Figure 6** | Design quantiles of the reconstructed series of the MRB flood peak discharge.

In order to obtain the design value for a certain return period of the original nonstationary flood peak discharge series, the design quantile for the same return period of the reconstructed series $RS(t)$ should be multiplied by the influence of the effective runoff generation area under design conditions according to Equation (6). The changes in the effective runoff generation area mainly occurred in the 1970s. In recent years, the construction of check dams has been saturated, so the effective runoff generation area will not change drastically in the immediate future. By this logic, we can use the effective runoff generation area data from 2010 as the effective runoff generation area under the design conditions. According to Equation (6), the design quantiles for nonstationary flood peak discharge series are shown in Figure 7. The thin lines represent the 95% confidence intervals (CIs) for the design floods. As the return period increases, the CIs get wider, and the uncertainty in the design flood grows.

Because the existing nonstationary frequency analysis techniques are far from mature, most hydrological projects

**Figure 7** | Design quantiles for the MRB flood peak discharge series. Thin lines represent the 95% bootstrapped CIs.

are still designed using the conventional hydrological frequency analysis method, regardless of whether the sample series is stationary. Due to the significant decreasing trend, if we calculate the design flood using the nonstationary MRB flood peak discharge series (the dashed line in Figure 7), the design flood for the same return period is larger than that of the value derived from the series we transformed with the de-nonstationarity method. The design values of the conventional method are nearly three times that of the de-nonstationarity method. The 100-year flood peak discharge is $1,829 \text{ m}^3/\text{s}$, which is generated within the effective runoff generation area of only 56.5 km^2 . This apparently larger, even somewhat impractical, design flood is unreliable. Because of the nonstationarity of the flood peak discharge series, it is obviously impossible to accurately assess the risk of flood by using conventional methods. Besides, the uncertainty of the design flood calculated by the original nonstationary flood peak discharge series is larger than that calculated by the de-nonstationarity method. When the return period is more than 50 years, the width of uncertainty CIs of the design value by the conventional method has exceeded the design value itself. According to the physical mechanism of check dams on flood peak discharge, the de-nonstationarity method

removes the nonstationarity and greatly reduces the uncertainty of the design value.

Comparison with reliability-based methods

Some reliability-based nonstationary hydrological frequency analysis methods were introduced in the Introduction, including the expected number of exceedance (ENE), design life level (DLL), equivalent reliability (ER), and average design life level (ADLL). In this paper, we also use these four methods to calculate the design flood of MRB. With time and effective runoff generation area as the covariates, the GAMLSS model was used to model the nonstationary annual maximum flood peak discharge series. The alternative distributions were also WEI, GU, GEV, and P-III. To determine the best-fitting model (Table 6), the Akaike information criterion (AIC) (Akaike 1974) was employed.

Following the time-varying characteristics of the annual maximum flood peak discharge series, the four reliability-based nonstationary hydrologic frequency analysis methods

were used to calculate the design value. It was assumed that the hydrologic structures of the MRB are planned to be in service from 1962 to 2061, with a lifespan of 100 years. Uncertainties were also calculated with the bootstrap method. The design values of different return periods are listed in Table 7 (other details are not shown).

As shown in Table 7, the design flood calculated by the DLL method is the largest, followed by the ENE method. The design values of these two methods even exceed that of the conventional method, which is obviously unreasonable. The design flood of the ENE method increased rapidly with the increase of the return period, and almost did not change when the return period was more than 20 years. This change characteristic of the design value of the ENE method is also unrealistic. The design value of ER and ADLL methods are relatively small and close to that of the de-nonstationarity method. However, the uncertainty of these two methods increases rapidly as the recurrence period increases, and even exceeds the magnitude of the design value. On the other hand, these methods all take the effective runoff generation area as the covariate to consider its influence. In MRB, although the effective runoff generation area of a flood event is significantly reduced, the runoff generation process outside of the control area of check dams should not change, that is, the flood peak modulus should not change. The flood peak modulus of the design flood obtained by these reliability-based methods and the conventional method is obviously increased. The de-nonstationarity method proposed in this paper has certain advantages in terms of the magnitude of the calculated design value, the uncertainty, and the consideration of the physical mechanism.

Table 6 | The best-fitting nonstationary distribution and estimated parameters for the annual maximum flood peak discharge series of the MRB

Basin	Optimal nonstationary distribution	Estimated parameters	AIC
MRB	WEI (<i>A</i> as covariates): $\log(\mu_t) = \alpha_0 + \alpha_1 A;$ $\log(\sigma_t) = \beta_0 + \beta_1 A$	$\alpha_0 = 4.748930$ $\alpha_1 = 0.004265$ $\beta_0 = 0.0934251$ $\beta_1 = -0.0004531$	630.43

Table 7 | The design floods of the MRB with 95% bootstrapped CIs calculated by different methods

Basin	Method	Design quantile (width of 95% CIs)/m ³ /s			
		20-year	30-year	50-year	100-year
MRB	De-nonstationarity	411 (290)	479 (349)	568 (436)	694 (568)
	Conventional	922 (872)	1,127 (1,115)	1,408 (1,464)	1,829 (2,008)
	ENE	1,338 (2,054)	1,338 (2,022)	1,339 (1,983)	1,340 (1,941)
	DLL	3,196 (6,123)	3,465 (6,829)	3,805 (7,710)	4,269 (8,969)
	ER	630 (553)	773 (820)	997 (1,228)	1,364 (1,886)
	ADLL	595 (452)	734 (698)	960 (1,146)	1,340 (1,833)

The width of 95% CIs is in italics, and the results of the de-nonstationarity method are in bold.

Meantime, the de-nonstationarity method also avoids some problems that are difficult to solve in the current methods. First, the time-varying probability distribution is the basis of the existing methods, but it is extremely difficult to accurately fit the time-varying characteristics of the flood peak discharge series. However, no matter what kind of nonstationarity (trend, abrupt change, etc.) caused by the influence factors, the de-nonstationarity method can remove the influence according to the physical mechanism. Second, in the absence of ergodicity, we cannot use the nonstationary sample series to estimate the population. Third, under nonstationary conditions, the different initial and end time of the calculation period result in different design values. The de-nonstationarity method transforms the nonstationary series into the stationary reconstructed series, thus avoiding this problem.

CONCLUSIONS

Check dams, as a kind of engineering measure for water and soil conservation, lead to the nonstationary changes in the flood peak discharges series. This nonstationarity creates unreliable design flood values when they are calculated with the conventional hydrological frequency analysis method. Previous works have attempted to explore the time-varying characteristics of nonstationary hydrological series using purely statistical approaches. Due to the influence of climate change and human activity, it is difficult to accurately describe the characteristics of hydrological time series or predict how the series will evolve in the future using statistical approaches alone. If we cannot guarantee the reliability of the time-varying probability distribution, then the design value under nonstationary conditions is also unreliable.

The conventional method, which requires hydrological series to be stationary, is a mature and robust scientific strategy. In the presence of check dams, nonstationarity in the flood peak discharge series is caused by changes in the effective runoff generation area. By quantifying the power function relationship between the flood peak discharge and the basin area, we can remove the nonstationarity from the flood peak discharge series. Once the reconstructed series is stationary, we can calculate the design quantile via the conventional method, and then obtain the design value of the

original nonstationary series according to the effective runoff generation area under design conditions.

We applied this de-nonstationarity method to the nonstationary MRB peak flood discharge series. After removing the nonstationary influence of the effective runoff generation area, the nonstationary flood peak discharge series was transformed into the stationary reconstructed series, and the stationarity is simultaneously achieved in the first and second moments. Because the nonstationarity in the effective runoff generation area has a stronger impact on the second moment of the flood peak discharge series than it does on the first moment, we cannot ignore nonstationarity in higher order moments in a changing hydrological environment. The design flood calculated using the de-nonstationarity method is obviously smaller than that calculated according to the original nonstationary flood peak discharge series. By removing the nonstationarity from the flood peak discharge series, this method can greatly reduce the uncertainty of the design flood. Compared to the four reliability-based methods, the de-nonstationarity method has certain advantages in terms of the magnitude of the calculated design value, the uncertainty, and the consideration of the physical mechanism.

The nonstationary problems caused by climate change or human activities cannot be solved through statistical approaches alone. It is necessary to explore the physical mechanism of the nonstationary change and establish a new method of combining physical mechanism with statistical theory. For other cases, it should also be possible to eliminate the nonstationary effects of influence factors on the hydrological series according to the influence law. Therefore, the exploration of the physical mechanism of the interaction between hydrological elements should be the focus of future research, which can also promote our understanding of hydrological processes.

ACKNOWLEDGEMENTS

This research is supported by the National Natural Science Foundation of China (Grant NO.51679184). We are grateful to the editor and the anonymous reviewers for their detailed and insightful comments, which helped us to improve the quality of this paper. We thank LetPub (www.letpub.com)

for its linguistic assistance during the preparation of this manuscript.

DATA AVAILABILITY STATEMENT

All relevant data are included in the paper or its Supplementary Information.

REFERENCES

- Akaike, H. T. 1974 A new look at the statistical model identification. *IEEE Transactions on Automatic Control* **19** (6), 716–723. doi: 10.1109/tac.1974.1100705.
- Breusch, T. S. & Pagan, A. R. 1979 A simple test for heteroscedasticity and random coefficient variation. *Econometrica: Journal of the Econometric Society* **47**, 1287–1294. doi: 10.2307/1911963.
- Cooley, D. 2013 *Return Periods and Return Levels Under Climate Change. Extremes in A Changing Climate*. Springer, Dordrecht, The Netherlands.
- Efron, B. 1992 *Bootstrap Methods: Another Look at the Jackknife*. Springer, New York, USA. doi: 10.2307/2958830.
- Gupta, V. K. 2004 Emergence of statistical scaling in floods on channel networks from complex runoff dynamics. *Chaos, Solitons and Fractals* **19** (2), 357–365. doi: 10.1016/S0960-0779(03)00048-1.
- Hosking, J. R. M. 1990 L-moments: analysis and estimation of distributions using linear combinations of order statistics. *Journal of the Royal Statistical Society. Series B (Methodological)* **52** (1), 105–124. doi: 10.2307/2345653.
- Hu, Y. M., Liang, Z. M., Singh, V. P., Zhang, X., Wang, J., Li, B. & Wang, H. 2017 Concept of equivalent reliability for estimating the design flood under non-stationary conditions. *Water Resources Management* **32** (2), 997–1011. doi: 10.1007/s11269-017-1851-y.
- Jiang, W. & Shuai, Q. 2011 Investigation and analysis on the problem of the dangerous check dams in Yulin. *Shaanxi Water Resources* **5**, 55–57. doi: CNKI:SUN:SXSXN.0.2011-05-026. (in Chinese).
- Kendall, M. G. 1975 *Rank Correlation Methods*. Oxford University Press, New York, USA.
- Khaliq, M. N., Ouarda, T. B. M. J., Ondo, J. C., Gachon, P. & Bobée, B. 2006 Frequency analysis of a sequence of dependent and/or non-stationary hydro-meteorological observations: a review. *Journal of Hydrology* **329** (3–4), 534–552. doi: 10.1016/j.jhydrol.2006.03.004.
- Kolmogorov-Smirnov, A. N., Kolmogorov, A. & Kolmogorov, M. 1933 Sulla determinazione empirica di una legge di distribuzione. (On the empirical determination of a distribution law). *Giornale Istituto Italiano Attuari* **4**, 83–91.
- Kottogoda, N. T. 1980 *Stochastic Water Resources Technology*. Springer, New York, USA.
- López, J. & Francés, F. 2013 Non-stationary flood frequency analysis in continental Spanish rivers, using climate and reservoir indices as external covariates. *Hydrology and Earth System Sciences* **17** (8), 3189–3203. doi: 10.5194/hess-17-3189-2013.
- Mann, H. B. 1945 Nonparametric tests against trend. *Econometrica: Journal of the Econometric Society* **13**, 245–259. doi: 10.2307/1907187.
- Nash, J. E. & Sutcliffe, J. V. 1970 River flow forecasting through conceptual models part I – A discussion of principles. *Journal of Hydrology* **10** (3), 282–290. doi: 10.1016/0022-1694(70)90255-6.
- Olsen, J. R., Lambert, J. H. & Haimes, Y. Y. 1998 Risk of extreme events under nonstationary conditions. *Risk Analysis* **18** (4), 497–510. doi: 10.1111/j.1539-6924.1998.tb00364.x.
- Parey, S., Malek, F., Laurent, C. & Dacunha-Castelle, D. 2007 Trends and climate evolution: statistical approach for very high temperatures in France. *Climatic Change* **81** (3), 331–352. doi: 10.1007/s10584-006-9116-4.
- Parey, S., Hoang, T. T. H. & Dacunha-Castelle, D. 2010 Different ways to compute temperature return levels in the climate change context. *Environmetrics* **21** (7–8), 698–718. doi: 10.1002/env.1060.
- Qin, Y., Quan, Q., Li, K. & Li, X. 2015 Calculation of design flood for regions with no data but influenced by warping dam projects. *Proceedings of the International Association of Hydrological Sciences* **368**, 281–286. doi: 10.5194/piahs-368-281-2015.
- Quan, Q. Z. 2016 *The Study on Design Flood Influenced by Reservoirs and Warping Dams in Loess Area*. Xi'an University of Technology, Xi'an. (in Chinese).
- Read, L. K. & Vogel, R. M. 2015 Reliability, return periods, and risk under nonstationarity. *Water Resources Research* **51** (8), 6381–6398. doi: 10.1002/2015WR017089.
- Rigby, R. & Stasinopoulos, D. 2005 Generalized additive models for location, scale and shape (with discussion). *Journal of the Royal Statistical Society: Series C* **54** (3), 507–554. doi: 10.1111/j.1467-9876.2005.00510.x.
- Rootzén, H. & Katz, R. W. 2013 Design life level: quantifying risk in a changing climate. *Water Resources Research* **49** (9), 5964–5972. doi: 10.1002/wrcr.20425.
- Salas, J. D. & Obeysekera, J. 2014 Revisiting the concepts of return period and risk for nonstationary hydrologic extreme events. *Journal of Hydrologic Engineering* **19** (3), 554–568. doi: 10.1061/(ASCE)HE.1943-5584.0000820.
- Serinaldi, F. & Kilsby, C. G. 2015 Stationarity is undead: uncertainty dominates the distribution of extremes. *Advances in Water Resources* **77**, 17–36. doi: 10.1016/j.advwatres.2014.12.013.
- Smith, J. A. 1992 Representation of basin scale in flood peak distributions. *Water Resources Research* **28** (11), 2993–2999. doi: 10.1029/92wr01718.
- Strupczewski, W. & Kaczmarek, Z. 2001 Non-stationary approach to at-site flood frequency modelling II. Weighted least

- squares estimation. *Journal of Hydrology* **248**, 143–151. doi: 10.1016/S0022-1694(01)00398-5.
- Strupczewski, W., Singh, V. & Feluch, W. 2001a Non-stationary approach to at-site flood frequency modelling I. Maximum likelihood estimation. *Journal of Hydrology* **248**, 123–142. doi: 10.1016/S0022-1694(01)00397-3.
- Strupczewski, W., Singh, V. & Mitosek, H. T. 2001b Non-stationary approach to at-site flood frequency modelling. III. Flood analysis of Polish rivers. *Journal of Hydrology* **248**, 152–167. doi: 10.1016/S0022-1694(01)00399-7.
- Villarini, G., Smith, J. A., Serinaldi, F., Bales, J., Bates, P. D. & Krajewski, W. F. 2009 Flood frequency analysis for nonstationary annual peak records in an urban drainage basin. *Advances in Water Resources* **32** (8), 1255–1266. doi: 10.1016/j.advwatres.2009.05.003.
- Wigley, T. M. L. 1988 The effect of changing climate on the frequency of absolute extreme events. *Climatic Monitoring* **17** (1–2), 44–55. doi: 10.1007/s10584-009-9654-7.
- Xie, P., Chen, G. C. & Xia, J. 2005 Hydrological frequency calculation principle of inconsistent annual runoff series under changing environments. *Engineering Journal of Wuhan University* **2005** (06), 6–9.
- Xiong, L., Jiang, C. & Du, T. 2014 Statistical attribution analysis of the nonstationarity of the annual runoff series of the Weihe River. *Water Science and Technology* **70** (5), 939–946. doi: 10.2166/wst.2014.322.
- Yan, L., Xiong, L., Guo, S., Xu, C. Y., Xia, J. & Du, T. 2017 Comparison of four nonstationary hydrologic design methods for changing environment. *Journal of Hydrology* **551**, 132–150. doi: 10.1016/j.jhydrol.2017.06.001.
- Zhang, Q., Gu, X., Singh, V. P., Xiao, M. & Chen, X. 2015 Evaluation of flood frequency under non-stationarity resulting from climate indices and reservoir indices in the East River basin, China. *Journal of Hydrology* **527**, 565–575. doi: 10.1016/j.jhydrol.2015.05.029.

First received 13 April 2020; accepted in revised form 30 June 2020. Available online 5 August 2020