Hydrological drought forecasting under a changing environment in the Luanhe River basin

Min Li[1,2], Mingfeng Zhang[1], Runxiang Cao[3], Yidi Sun[1], Xiyuan Deng[4,5],
([1]College of Hydraulic Science and Engineering, Yangzhou University, Yangzhou, China, 225000; [2]State Key Laboratory of Hydraulic Engineering Simulation and Safety, Tianjin University, Tianjin, China, 300072; [3]College of Water Resources, North China University of Water Resources and Electric Power, Zhengzhou, China, 450046; [4]Nanjing Hydraulic Research Institute, Nanjing, China, 210029; [5]State Key Laboratory of Hydrology-Water Resources and Hydraulic Engineering, Nanjing, China, 210029)

Abstract: Forecasting the occurrence of hydrological drought according to a forecasting system is an important disaster reduction strategy. In this paper, a new drought prediction model adapted to changing environments was constructed. Taking the Luan River basin in China as an example, first, nonstationarity analysis of hydrological sequences in the basin was carried out. Then, conditional distribution models with the human activity factor as an exogenous variable were constructed to forecast hydrological drought based on meteorological drought, and the results were compared with the traditional normal distribution model and conditional distribution model. Finally, a scoring mechanism was applied to evaluate the performance of the three drought forecasting models. The results showed that the runoff series of the Luanhe River basin from 1961 to 2010 were nonstationary; moreover, when human activities were not considered, the hydrological drought class tended to be the same as the meteorological drought class. The calculation results of the models involving *HI* as an exogenous variable were significantly different from the models that did not consider human activities. When the current drought class tended towards less severe or normal, the meteorological drought tended to turn into more severe hydrological drought with the increase in human index values. According to the scores of the three drought forecasting models, the conditional distribution models involving the human index can further improve the forecasting accuracy of drought in the Luanhe River basin.

Keywords: Changing environment; Drought forecasting; Human activity factor; Luanhe River basin

Acknowledgements: This work was supported by the State Key Laboratory of Hydraulic Engineering Simulation and Safety (Tianjin University) program (No. HESS-2206 and No. HESS-2222).

[1] Corresponding author: Min Li Mail: limintju@126.com
Xiyuan Deng Mail: xydeng@nhri.cn

1 Introduction

Typically, meteorological drought is regarded as the beginning of a drought event; after the occurrence of meteorological drought, other drought phenomena occur, such as hydrological drought (Miriam et al., 2018; Fuentes et al., 2022; Wang et al., 2021). However, there is a delay period from meteorological drought to hydrological drought (Ding et al., 2021; Xu et al., 2019; Charles, 2017; Carmelo and Jürgen, 2018). Therefore, the occurrence of hydrological drought can be forecasted according to meteorological drought monitoring. Accurate hydrological forecast information is beneficial to reduce the losses caused by hydrological drought (Behzad and Hamid, 2019; Melanie et al., 2018 Dixit et al.,2022; Muhammad et al.,2020).

To identify the drought characteristics of the region, scholars have developed drought indices. For example, the standardized precipitation index (SPI) is typically used to identify and capture the characteristics of meteorological drought (McKee, 1993). Considering the influence of precipitation and temperature, Vicente-Serrano et al. (2010) proposed the standardized precipitation evapotranspiration index (SPEI) to characterize meteorological drought. The standardized runoff index (SRI), which focuses on the surface runoff of catchments, is typically used to indicate hydrological drought (Shukla, 2008). Aghelpour and Varshavian (2021) proposed the multivariate standardized precipitation index (MSPI) to forecast hydrological drought in Iran.

Statistical technology is an effective prediction method that has been widely used in drought forecasting in recent years (Alquraish et al., 2021; Abbasi et al., 2021; Bagher et al., 2013). For instance, neural network models have been proposed to combine multiple data for drought prediction (Mehdi et al., 2016; Maryam et al., 2017; Ahnadi et al., 2011), and time series models can be used to analyse the variation in time series such as rainfall and runoff to achieve drought prediction (Mohammad et al., 2020; Natsagdorj et al., 2021; Stojković et al., 2020). The conditional probability model was proposed by Cancelliere et al. (2007) and developed for drought forecasting by Bonaccorso et al. (2015). Bonaccorso et al. (2015) showed that the conditional probability model can calculate the transition probabilities from the current drought index values to the future drought classes, and this is a more robust method that can be used to forecast drought than the traditional probability prediction models (such as the multivariate normal distribution model and Markov method).

A change in the environment may lead to the nonstationarity of the relationship between hydrological series (for example, precipitation and runoff series), which also occurs in the Luanhe River basin (Wang et al., 2018; Li et al., 2015; Wang et al., 2016). Traditional drought prediction methods need to be further improved to adapt to nonstationary conditions (Wang et al., 2022; Zhao et al., 2018; Chen et al. (2021)). Ren et al. (2017) found that the conditional distribution model using large-scale climatic indices as covariates can improve the accuracy of meteorological drought forecasting in the Luanhe River basin. Although some progress has been made in the study of drought forecasting, there are relatively few studies considering the impact of the changing environment.

In this paper, to analyse the impact of human activities on hydrological drought, we constructed the human activity index (*HI*) based on the restoration method. Subsequently, conditional distribution models with the *HI* as the exogenous variable were developed to forecast hydrological drought based on meteorological drought, and then the results were compared with the traditional normal distribution model and conditional distribution model; as a result, the impact of human activities on transition probabilities was illustrated. A scoring mechanism was applied to the evaluation of the three probability models.

In addition to the introduction, this paper contains the following sections. Section 2 introduces the study area and data. Section 3 briefly describes the methods used in the research. Section 4 introduces the model construction and calculation results and analyses the results. Section 5 presents the prospects.

2 Study area and data

The Luanhe River basin, located in the subtropical monsoon region, covers an area of approximately 33700 square kilometres. Its geographical location is shown in Figure 1. Due to the influence of geographical location and topography, the annual average north–south temperature difference in the basin is 11.5 °C, and the annual rainfall distribution is uneven. Less rain in spring and winter makes the area prone to meteorological drought and hydrological drought, while there is relatively more rainfall in summer. The average rainfall in summer is approximately 200-560 mm, resulting in highly variable annual runoff in the basin. The concentrated rainfall in summer has also become one of the remarkable features of the climate in this area. In recent years, the precipitation and inflow of the Luanhe River basin have gradually decreased, the water level of the Panjiakou Reservoir in the lower reaches of the basin has decreased, the runoff has decreased, and the frequency of meteorological drought and hydrological drought has significantly increased. Especially after entering the 21st century, the river basin has exhibited continuous drought and even extreme drought. With the change in the global climate and the impact of human activities on the basin environment, drought disasters in the Luanhe River basin occur frequently, causing significant social and economic losses.

Influenced by topography, meteorology, hydrology and hydrogeological conditions, the spatial distribution of groundwater resources in the Luanhe River basin is quite different. The recharge and storage conditions of shallow groundwater in plain areas and intermountain basins are relatively superior, and the content of groundwater in mountainous areas is relatively small (the area of mountainous areas in the Luanhe River basin accounts for 98.2%). Therefore, the total amount of water resources in the Luanhe River basin is mainly considered to be affected by the amount of surface water resources.

In this paper, the monthly rainfall data from 26 stations in the Luanhe River basin from 1961 to 2010 were provided by the Hebei Provincial Hydrology and Water Resources Investigation Bureau. The average monthly rainfall data of the area were obtained by the inverse distance weighting interpolation method. The runoff data from 1961 to 2010 came from the inflow runoff series of the Panjiakou Reservoir. The SPI and SRI can be calculated for 1-month, 3-month, 6-month, and 12-month time scales to characterize meteorological drought and hydrological drought based on these data.

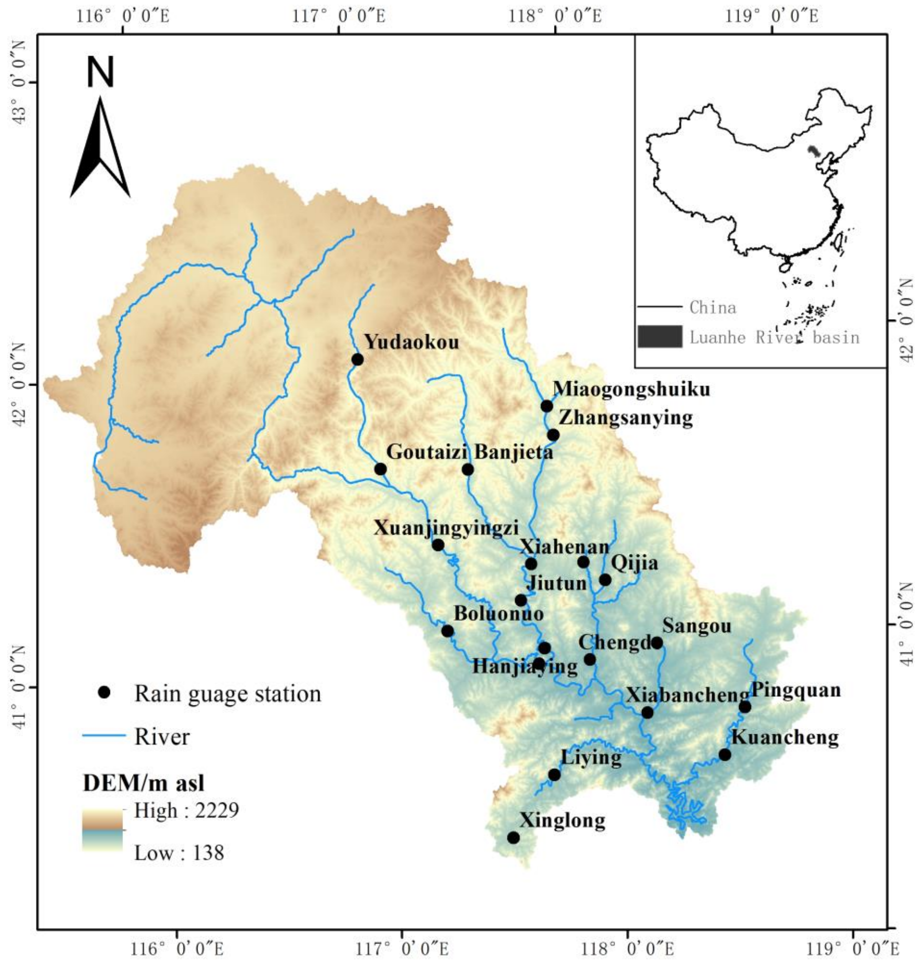


Figure 1 The geographical location of the Luanhe River basin

3 Methods

3.1 Nonstationarity test method

In the case of environmental changes, nonstationarity may occur in hydrological series. The Pettitt test, as one of the important methods to test whether there is nonstationarity in time series, can identify whether there are change points in the sample series (Malede et al., 2022). Assuming that the sample sequence is $x=(x_1,x_2,\cdots x_n)$, the formula is as follows:

$$U_{t,n}=U_{t-1,n}+\sum_{i=1}^{n}\operatorname{sgn}(x_t-x_i)\quad (t=2,3,\cdots n)t_0 \tag{1}$$

$$\operatorname{sgn}(x_t-x_i)=\begin{cases}1 & x_t-x_i>0\\ 0 & x_t-x_i=0\\ -1 & x_t-x_i=0\end{cases} \tag{2}$$

where $U_{t,n}$ is the test statistic, which indicates the cumulative number of the values at time t greater than or less than the values at time i. In addition, if $K_{t0,n}$ satisfies the following:

$$K_{t0,n} = \max\left|U_{t,n}\right| \quad (\text{t=1,2,}\cdots\text{,n}) \tag{3}$$

Then, t_0 is considered to be the change point, and the cumulative probability of possible change is determined by $K_{t0,n}$:

$$P_{t0,n} = 2\exp\left(-\frac{6K_{t0,n}^2}{n^3+n^2}\right) \tag{4}$$

Given the significance level $\alpha = 0.05$, if $P_{t0,n}$>0.95, it means that the point is a significant change point (Li et al., 2022; Koudahe et al., 2018). Furthermore, combined with the Mann-Kendall test, the trend characteristics of the sample series can be obtained (Linchao et al., 2018).

The sliding T test is a basic method commonly used in statistics. According to the mean and variance of the two sample sequences before and after the change points in the runoff time series, the two sample sequences are tested (Li et al., 2020):

$$t = \frac{\overline{x}_1 - \overline{x}_2}{S\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}} \tag{5}$$

$$S = \sqrt{\frac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}} \tag{6}$$

$$S_1^2 = \frac{1}{n_1-1}\sum_{t=1}^{n_1}(x_t-\overline{x}_1)^2 \tag{7}$$

$$S_2^2 = \frac{1}{n_2-1}\sum_{t=1}^{n_1+n_2}(x_t-\overline{x}_2)^2 \tag{8}$$

where the change point is x_t, n_1 and n_2 represent the sample size before and after the change point, and S_1^2 and S_2^2 represent the variance of the samples before and after the change point, respectively. If the t statistic satisfies $t > t_\alpha$ at the significance level of $\alpha = 0.05$, the point can be considered a change point.

The Spearman correlation test can be applied to test the trend of time series, and the specific description refers to the article of Bishara and Hittner (2012).

3.2 Human activity index

The rainfall and runoff series of the watershed are usually strongly correlated. However, under the interference of human activities, the relationship between rainfall and runoff changes.

The double cumulative curve method can test the nonstationarity of the bivariate correlation between rainfall series and runoff series, and the point where the underlying surface is significantly altered by human activities can be determined according to the position of the slope change of the curve. Due to the short data series before and after the change point (20 years before the change point and 30 years after the change point), a linear equation was used to fit the relationship between precipitation and runoff.

The linear regression relationship of the cumulative rainfall and runoff series can be calculated according to the following formula:

$$\sum x = k \sum y + b \tag{9}$$

where x is the runoff series; y is the rainfall series; k is the correlation coefficient of the regression equation; and b is the intercept of the regression equation.

Human activities are the main reason for the nonstationarity of the runoff series in the watershed, so the *HI* can be constructed to quantify the impact of human activities on runoff. Based on the linear regression relationship established between the accumulated precipitation and the accumulated runoff before the change point, the theoretical runoff sequence during the human activity period can be calculated from the measured precipitation sequence. SRI' represents the standardized runoff index value without human activity interference, and SRI represents the normalized runoff index value calculated based on the measured runoff sequence under the disturbance of human activities. The *HI* is obtained by subtracting the theoretical SRI' and the actual SRI, and the calculation formula is as follows:

$$HI = SRI' - SRI \tag{10}$$

When *HI*>0, it can be assumed that human activities exacerbate hydrological drought, *HI*<0 means that the actual SRI is greater than the theoretical SRI without human activities, and when *HI*=0, the watershed is considered undisturbed by human activities.

3.3 Multivariate normal distribution model

The SPI is one of the important indicators for evaluating meteorological drought in the basin, and the SRI is an important indicator for evaluating hydrological drought in the basin. According to the rainfall data and runoff data in the basin, the SPI and SRI can be calculated at different time scales. Table 1 provides the drought class classification and corresponding SPI values and SRI values (Kolachian and Saghafian, 2021).

Table 1 Drought class classification and corresponding SPI values and SRI values

SPI/SRI values	Class
> -0.99	Normal
-1.00 to -1.49	Moderate
-1.50 to -1.99	Severe
≤ -2.00	Extreme

As a traditional drought class forecasting model, the multivariate normal distribution model (Model 1) can forecast the future SRI class according to the current SPI class. Assuming that the current SPI and SRI series both satisfy a multivariable normal distribution, the joint probability density can be expressed as follows (Chang et al., 2022):

$$f_{Z^{(k)}_{v,\lambda} W^{(k)}_{v,\lambda+M}}(t,s) = \frac{1}{2\pi|\Sigma|} \cdot \exp\left(-\frac{1}{2} X^{T} \Sigma^{-1} X\right) \qquad (11)$$

where $\sum$ is the covariance matrix, and $X = [t, s]^{T}$. The form of the covariance matrix is as follows:

$$\Sigma = \begin{bmatrix} 1 & \operatorname{cov}\left[Z^{(k)}_{v,\lambda}, W^{(k)}_{v,\lambda+M}\right] \\ \operatorname{cov}\left[Z^{(k)}_{v,\lambda}, W^{(k)}_{v,\lambda+M}\right] & 1 \end{bmatrix} \qquad (12)$$

Furthermore, according to the joint probability density function of the SPI value $Z^{(k)}_{v,\lambda}$ at year v and month λ and the future M month's SRI value $W^{(k)}_{v,\lambda+M}$, the analytical formula of the transition probability of the future SRI drought class can be obtained (Zhang et al., 2017):

$$P\left[W_{v,\lambda+M}^{(k)} \in C_M\right] = \frac{\iint_{C_N C_M} f_{Z_{v,\lambda}^{(k)} W_{v,\lambda+M}^{(k)}}(t,s) \cdot dt \cdot ds}{\int_{C_N} f_{Z_{v,\lambda}^{(k)}}(t) \cdot dt} \tag{13}$$

where C_M represents the drought class, and $f_{Z_{v,\lambda}^{(k)}}(t)$ represents the marginal density function of $Z_{v,\lambda}^{(k)}$ in the current λ month.

3.4 Conditional distribution model

The conditional distribution model (Model 2) proposed by Bonaccorso et al. (2015) is described as follows: when one group of sample data X obeys a normal distribution and satisfies $X \sim N(\mu_1, \Sigma_1)$, and another group of sample data Y also obeys a normal distribution, namely, $Y \sim N(\mu_2, \Sigma_2)$, then the total sequence can be written as follows:

$$B = \begin{bmatrix} X \\ Y \end{bmatrix} \begin{matrix} r \\ p-r \end{matrix} \sim N_p\left(\begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}, \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix}\right) \tag{14}$$

When sequence Y obeys a normal distribution, the distribution of sequence X under the Y condition still satisfies a normal distribution, namely, the distribution of $(X \mid Y)$ is as follows (Gong et al. 2021):

$$(X \mid Y) \sim N(\mu_3, \Sigma_3) \tag{15}$$

where μ_3 represents the expected value under the conditional distribution, and Σ_3 is the conditional covariance matrix:

$$\mu_3 = \mu_1 + \Sigma_{12}\Sigma_{22}^{-1}(y - \mu_2) \tag{16}$$

$$\Sigma_3 = \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21} \tag{17}$$

Then, the probability of the current SPI value transitioning to the future SRI drought class can be deduced as follows (Ren et al., (2017)):

$$P\left[W_{v,\lambda+M} \in C_M \mid Z_{v,\lambda} = z_0\right] = \int_{C_{Mi}}^{C_{Ms}} \frac{1}{\sqrt{2\pi}\sigma_Z} e^{-\frac{1}{2}\left(\frac{x-\rho z_0}{1-\rho^2}\right)^2} dx \tag{18}$$

where $Z_{v,\lambda}$ represents the SPI value of the current month λ , $W_{v,\lambda+M}$ represents the SRI value of the $\lambda+M$ month, C_{Ms} and C_{Mi} are the upper and lower limits of drought class C_M , and the correlation coefficient between the current SPI value and the future SRI value is ρ . Furthermore, the current SPI and future SRI can be expressed as the standard normal cumulative distribution function Φ :

$$P\left[W_{v,\lambda+M} \in C_M \mid Z_{v,\lambda} = z_0\right] = \Phi\left[\frac{C_{Ms} - \rho \bullet z_0}{1-\rho^2}\right] - \Phi\left[\frac{C_{Mi} - \rho \bullet z_0}{1-\rho^2}\right] \qquad (19)$$

The calculation of the correlation coefficient ρ is as follows:

$$\rho = \frac{\operatorname{cov}[Z_{v,\lambda}^{(k)}, W_{v,\lambda+M}^{(k)}]}{\sqrt{\operatorname{var}(Z_{v,\lambda}^{(k)})\operatorname{var}(W_{v,\lambda+M}^{(k)})}} \qquad (20)$$

where K represents the time scale of the drought index. Assuming that the cumulative rainfall Y and runoff X satisfy a normal distribution, after the standardization process, the SPI value $Z_{v,\lambda}^{(k)}$ corresponding to cumulative rainfall Y and SRI value $W_{v,\lambda+M}$ corresponding to runoff X obey the standard normal distribution, namely:

$$\operatorname{var}(Z_{v,\lambda}^{(k)}) = \operatorname{var}(W_{v,\lambda+M}^{(k)}) = 1 \qquad (21)$$

$\operatorname{cov}[Z_{v,\lambda}^{(k)}, W_{v,\lambda+M}^{(k)}]$ represents the covariance between the current SPI and the Sri value with a forecast period of M months. The calculation is as follows:

$$\operatorname{cov}[Z_{v,\lambda}^{(k)}, W_{v,\lambda+M}^{(k)}] = \frac{1}{\sqrt{\sum_{i=0}^{k-1}\sigma_{\lambda+M-i}^2 \sum_{j=0}^{k-1}\sigma_{\lambda-j}^2}} \cdot \sum_{i=0}^{k-1}\sum_{j=0}^{k-1}\operatorname{cov}\left[X_{v,\lambda+M-j}, Y_{v,\lambda-i}\right] \qquad (22)$$

3.5 Conditional distribution model involving the *HI* as an exogenous variable

According to the above conditional probability model, when considering the *HI* as an exogenous variable, the model (Model 3) can be extended as follows:

$$P\left[W_{v,\lambda+M} \in C_M \mid Z_{v,\lambda} = z_0, H_{v,\lambda} = h_0\right] = \int_{C_{Mi}}^{C_{Ms}} \frac{1}{\sqrt{2\pi\sigma_z}} e^{-\frac{1}{2}\left(\frac{x-\mu_z}{\sigma_z}\right)^2} dx \qquad (23)$$

$$\mu_z = E\left[W_{v,\lambda+M} \mid \mathrm{Z}_{\mathrm{v},\lambda}, H_{\mathrm{v},\lambda}\right] = \Sigma_{12}'(\Sigma_{22}')^{-1}\begin{bmatrix} z_0 \\ h_0 \end{bmatrix} \qquad (24)$$

$$\sigma_z^2 = \text{var}\left[W_{v,\lambda+M} \mid Z_{v,\lambda}, H_{v,\lambda}\right] = 1 - \Sigma'_{12}(\Sigma'_{22})^{-1}\Sigma'_{21} \quad (25)$$

where:

$$\Sigma'_{12} = \left[\text{cov}(W_{v,\lambda+M}, Z_{v,\lambda})\ \ \text{cov}(W_{v,\lambda+M}, H_{v,\lambda})\right] \quad (26)$$

$$\Sigma'_{22} = \begin{bmatrix} 1 & \text{cov}(Z_{v,\lambda}, H_{v,\lambda}) \\ \text{cov}(H_{v,\lambda}, Z_{v,\lambda}) & 1 \end{bmatrix} \quad (27)$$

$$\Sigma'_{21} = (\Sigma'_{12})^T \quad (28)$$

3.6 Scoring mechanism

A scoring mechanism was applied to evaluate the performance of the drought forecasting models. In this method, the monthly drought transition probability is summed to evaluate the model (Chen et al., 2013), where $p_{s,t}$ characterizes the transition probability in month t of year s, and n is the length of the validation period.

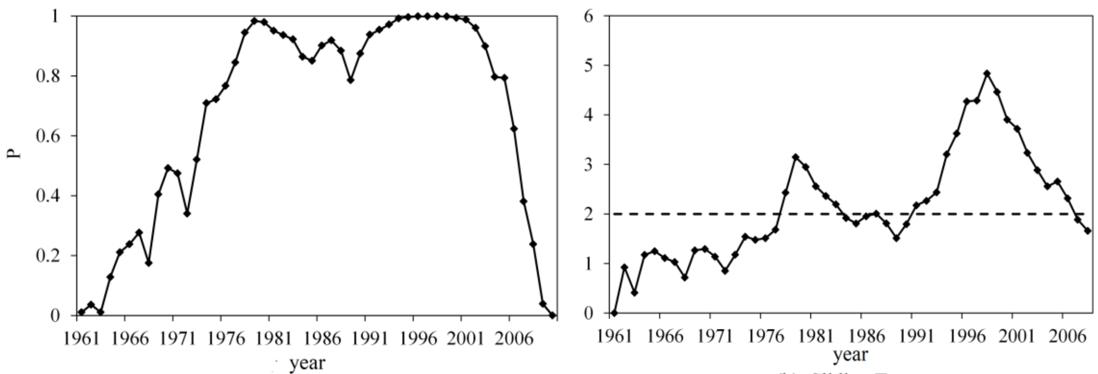
$$Score = \frac{1}{12n}\sum_{t-1}^{12}\sum_{s=1}^{n} p_{s,t} \quad (29)$$

4 Results and discussion

4.1 Nonstationarity analysis

In this paper, the area average monthly rainfall data of the Luanhe River basin from 1961 to 2010 were obtained by spatial interpolation. The runoff data came from the inflow runoff series of the Panjiakou Reservoir. Given the significance level $\alpha = 0.05$, the nonstationarity test results are shown in Figure 2.

Figure 2 (a) shows that the years of possible runoff change were 1979, 1996, 1997, 1998, and 1999. The P values in 1979 and 1998 were infinitely close to 1, which were considered to be extremely significant runoff change points. Among all the possible points satisfying $t > t_\alpha$, there were two maximum points (Figure 2 (b)), namely, 1979 and 1998, which were considered to be possible runoff change points. The final change point needs to be judged based on the actual situation of the watershed.



(a) Pettitt test (b) Sliding T test

Figure 2 The change points of the runoff series

The results of the Spearman correlation test (Table 2) indicate that the runoff series showed an upwards trend before 1979, but the trend was not significant. However, there was a significant downwards trend in the series after 1979. In general, the runoff series showed a significant downwards trend.

Table 2 Spearman correlation test results of runoff series trend

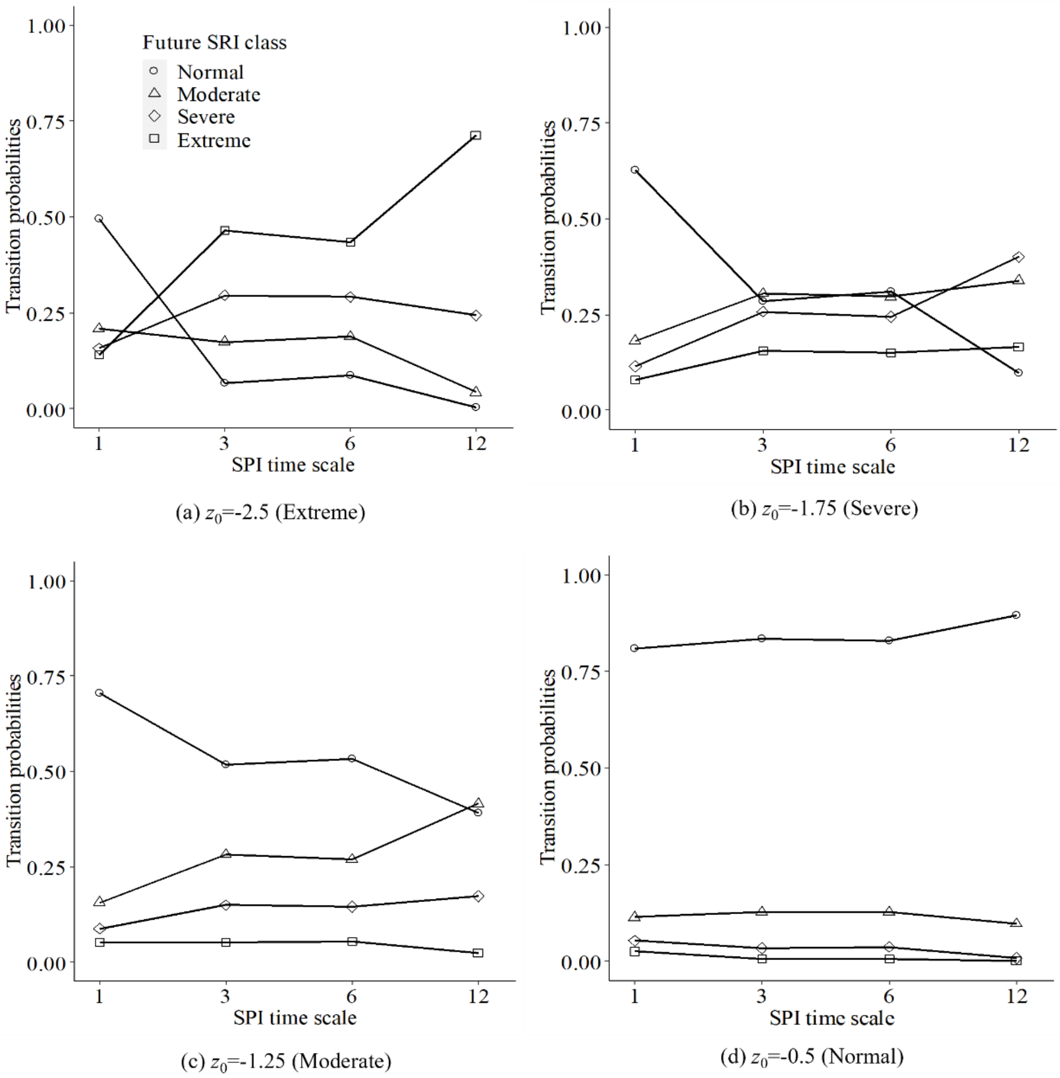
Runoff series	statistic t	Critical value t_α
The whole series	-3.471	±2.009
Serie before 1979	0.691	±2.009
Serie after 1979	-2.292	±2.009

In addition, according to historical records, local human activities (such as land use change and reservoir construction) are regarded as the main factors influencing runoff (Yan et al.,2018; Chen et al., 2021). Synthesizing the above analysis, 1979 was determined as the change point for the runoff sequence in the basin, and this conclusion was consistent with Li et al. (2015) and Wang et al. (2015).

4.2 Transition probabilities from current SPI values to future SRI classes

According to the normality test results of rainfall and runoff series, it was reasonable to apply the conditional distribution model. To analyse the influence of different time scales of the SPI on the transition probabilities, using the forecast period as one month and the SPI time scales at 1 month, 3 months, 6 months and 12 months as examples, the probabilities of converting SPI values to SRI classes were calculated (Figure 3).

As shown in Figure 3, when meteorological drought is categorized as extreme drought, the probabilities of maintaining the SRI class in extreme drought increased with the increasing SPI time scale. While the SPI had a short time scale, the response of the future SRI class to rainfall was fast, so the hydrological drought was more likely to tend to a normal status. This situation also occurred when the current meteorological drought was in another status.



(a) z_0=-2.5 (Extreme)

(b) z_0=-1.75 (Severe)

(c) z_0=-1.25 (Moderate)

(d) z_0=-0.5 (Normal)

Figure 3 Influence of the SPI time scale on transition probabilities (z_0 : initial value of SPI)

In addition, the transition probabilities of drought were distinct for different forecast periods. As seen in Figure 4, when the forecast periods were short (*M*=1 or 2), the hydrological drought classes obtained from the transition of meteorological drought tended to be the same as those of meteorological drought. With the extension of the forecast period (M=2 or 3), the hydrological drought classes obtained from the transition tended to be lower than the meteorological drought or the normal status.

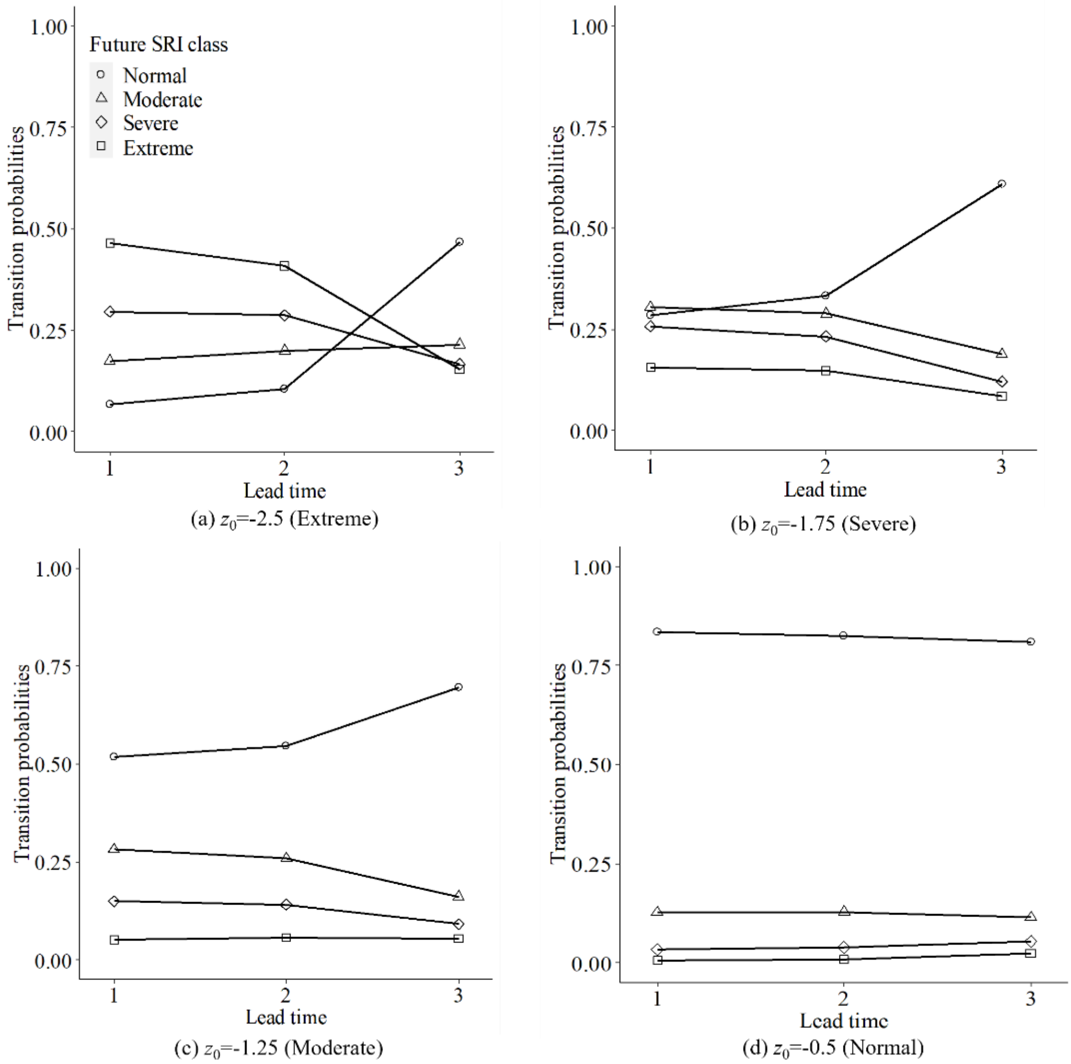


Figure 4 Influence of forecast period on transition probabilities (z_0 : initial value of SPI)

4.3 Transition probabilities involving the *HI* as the covariate

The effects of human activities are complex. To quantify the impact of human activities, the change point was identified, and then it was believed that the difference in the relationship between precipitation and runoff before and after the change point was caused by human activities. Moreover, the *HI* is easy to calculate and can approximately replace the influence of human activities. According to historical records, local human activities (such as land use change and reservoir construction) were regarded as the main factors influencing runoff (Yan et al., 2018; Chen et al., 2021). According to the above nonstationarity test results, 1979 was the change point, and the linear regression relationship of the cumulative rainfall and runoff series before and after the change point was established. The calculation results are shown in Table 3:

Table 3 Linear regression relationship between cumulative precipitation (x / mm) and cumulative runoff ($y/10^6\mathrm{m}^3$)

Period	Linear regression equation	Correlation coefficient
1961~1979	$x=0.0276y+2.7566$	0.99
1980~2010	$x=0.0307y-30.652$	0.98

The HI results for different time scales are shown in Figure 5.

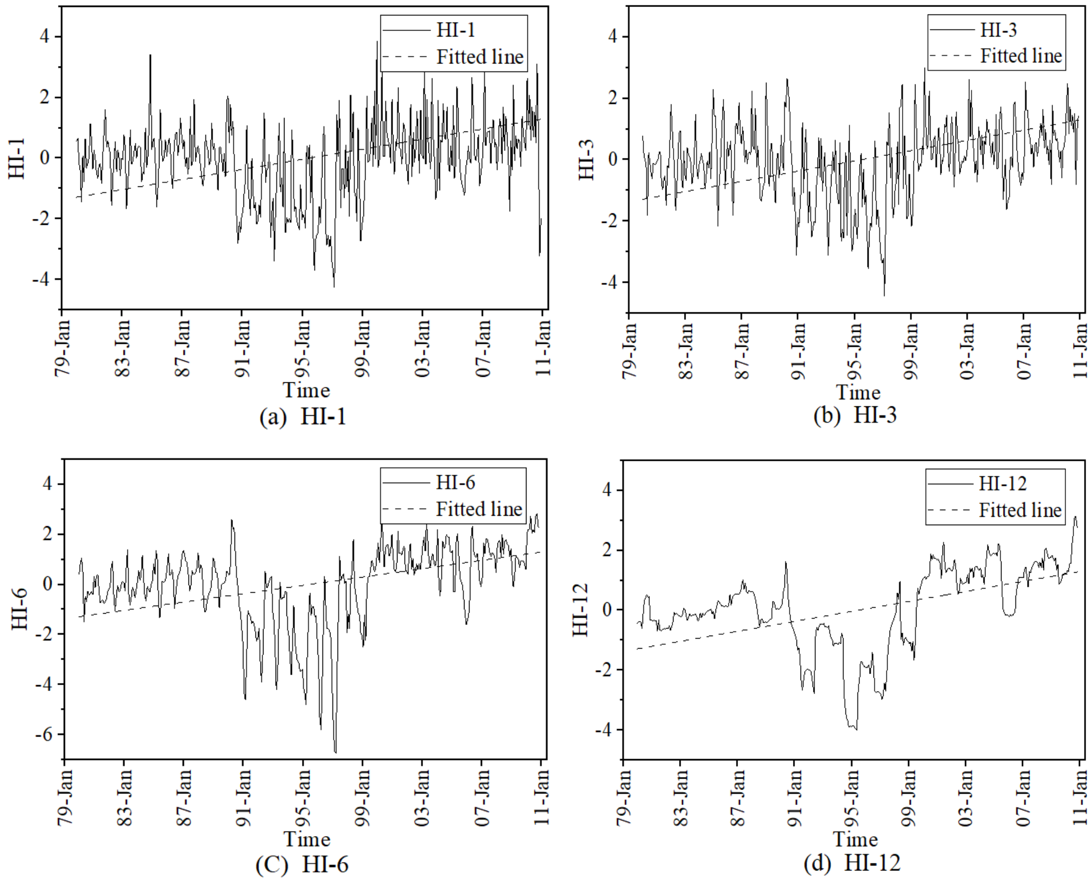


Figure 5 Different average periods of the *HI* (*HI*-1: *HI* with a 1-month time scale; *HI*-3: *HI* with a 3-month time scale; *HI*-6: *HI* with a 6-month time scale; *HI*-12: *HI* with a 12-month time scale)

As shown in Figure 5, the *HI* at all monthly scales generally ranged upwards, which means that human activities have intensified the occurrence of hydrological drought. According to historical statistics, many water conservancy projects were built in the basin from 1980 to 2000, and the construction and operation of large reservoirs in the mid-1990s may be the main reason for the serious negative values of the *HI*.

The *HIs* of different monthly scales were standardized, taking the 12-month time scale as an example, and the results were calculated as shown in Table 3.

Table 3 *HI*-12 monthly mean and standard deviation

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sept	Oct	Nov	Dec
Mean	-0.04	-0.03	-0.03	-0.03	-0.03	0.00	0.06	0.06	0.10	0.10	0.09	0.06
Sd	1.36	1.37	1.38	1.41	1.41	1.51	1.40	1.40	1.45	1.44	1.44	1.43

Furthermore, the drought transition probabilities involving the *HI* can be calculated from Eq. (23). Using the forecast period of one month from December and the SPI time scale of 12 months as an example, the drought transition probabilities from the current SPI values to the future SRI classes were calculated (Figure 6). To analyse the effect of human activities on the drought transition probability more clearly, the calculation results of the three models are compared here separately. The horizontal coordinate indicates the drought classes corresponding to the SRI for the coming month, and the vertical coordinate is the drought transition probability.

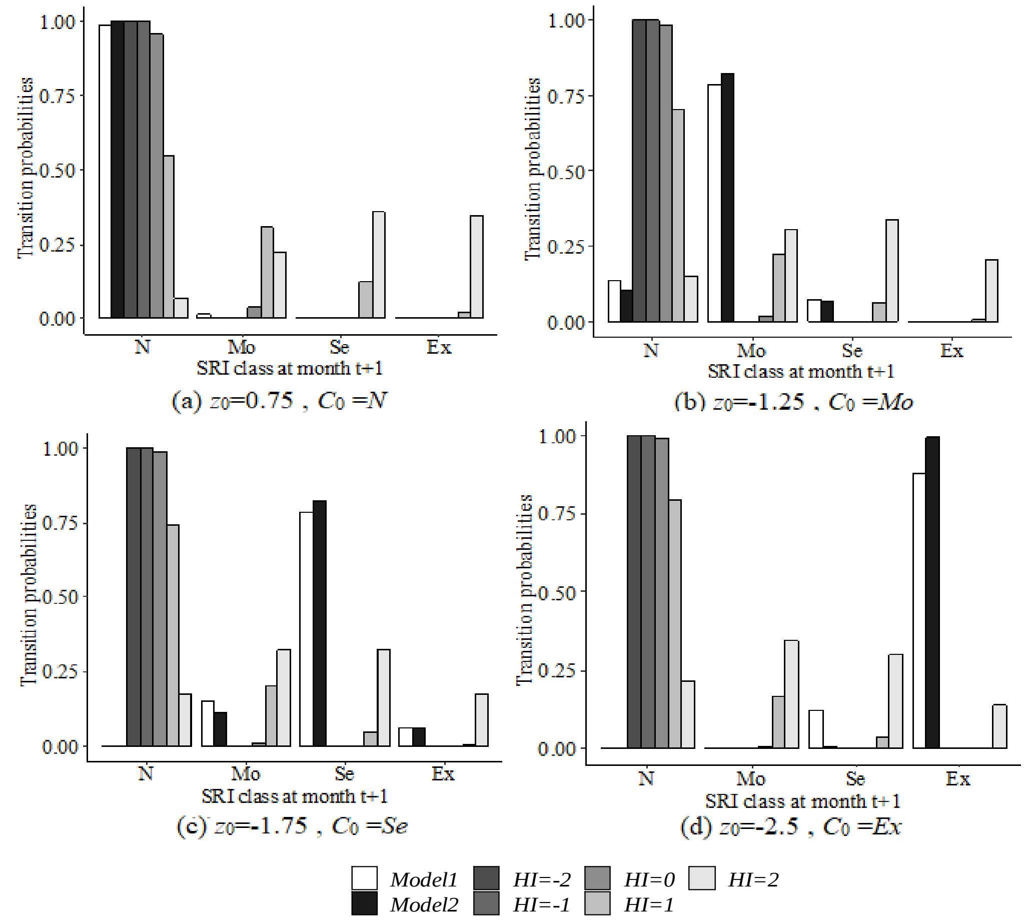


Figure 6 Drought transition probability under the influence of human activities (c_0 denotes the initial drought class of the SPI in the multivariate normal model; z_0 represents the initial value of the SPI in the conditional distribution model; Model 1: The normal distribution model; Model 2: The conditional distribution model; Model 3: The conditional distribution model involving the *HI*)

In Figure 6 (a), when the initial z_0=0.75 and c_0=*N*, the results shown in Model 1 and Model 2 were similar, and the probabilities of the SPI values transitioning to the SRI classes in the future month in the normal class were close to 1. However, the results of Model 3 indicated that the probabilities of maintaining the SRI in the normal class in the future decreased as the *HI* increased. When *HI*=2, the future hydrological drought classes were more likely to transition to severe drought or extreme drought.

From the initial z_0=-1.25 and c_0=*Mo* (Figure 6 (b)), the results of Model 3 showed that the transition probabilities of the SPI values to a normal SRI class in the coming month were higher when the *HI* was less than 1. As the *HI* increased, the transition probabilities of the SPI values to a moderate drought or even a more severe drought in the future increased. In addition, the probabilities of maintaining moderate drought were the highest when human activities were not considered, and Model 2 showed a higher probability than Model 1.

While the initial meteorological drought class was severe drought (Figure 6 (c)), the probabilities of the future SRI drought class being in the normal class became larger as the *HI* decreased. When the effect of human activities was not considered, the probability that the current SPI value transitioned to the SRI class under severe drought in the future month was the highest, and the probability of being in the normal class was the lowest. For Model 2, the probability of the SRI classes transitioning to severe drought was higher than the result of Model 1.

It was noteworthy that when the initial z_0=-2.5 and c_0=*Es* (Figure 6 (d)), the probabilities of transition of the SPI values to the future SRI classes at the normal class were close to 1 as *HI*<0. However, hydrological drought was more likely to be moderate drought or severe drought, as *the HIs* were greater than 0, and the transition probabilities exceeded 0.25. For Model 1 and Model 2, the probabilities of transition of the current SPI values or classes to the future month SRI classes in extreme drought were both higher than 0.75, and Model 2 showed a higher probability than Model 1.

In general, for the evaluation of drought transition probabilities in the future month, hydrological drought classes tended to be the same as meteorological drought when human activities were not considered, and this situation was more significant in Model 2 than in Model 1. The calculation results of the model involving the *HI* as an exogenous variable were significantly different from those of the models that did not consider human activities. The calculation results of Model 1 and Model 2 showed that the future hydrological drought classes were more likely to be the same as the meteorological drought classes in the current period, and they were more significant in Model 2. In addition, it was obvious that the drought transition probabilities of Model 3 were significantly different from those of Model 1 and Model 2. Taking Figure 6 (b) as an example, when z_0 =−1.25 and C_0=Mo, the result of Model 1 showed that the probability of the SPI values transitioning to the SRI classes in the future month in the normal class was close to 0.15, the result of Model 2 was close to 0, and the result of Model 3 (*HI*=0) was close to 0.95. The results of Model 3 (*HI*=0) indicated that hydrological drought was likely to remain at the normal class in the future month. Moreover, the value of the *HI* had a great impact on the results of Model 3; for example, when HI=−2 or −1, the probabilities of the SPI values transitioning to the SRI classes in the future month in the normal class were both close to 1, but the probability was close to 0.65 and 0.17 when *HI*=1 and 2, respectively.The results further indicated that meteorological drought tended to turn into more severe hydrological drought with increasing *HI* values.

4.4 Model evaluation and analysis

To quantitatively evaluate the prediction accuracy of Model 1, Model 2 and Model 3, the study period was divided into a correction period (1961-2003) and a verification period (2004-2010), and then the drought transition probability from the SPI value or class to the SRI class in the future M-month was calculated. The calculation results are shown in Table 4.

With the same time scale of the SPI, the model scores of Model 1 and Model 2 decreased as the forecast period *M* lengthened, while the model scores of Model 3 were not significantly affected by the forecast period *M*. Model 1 had the highest rating of 0.36 at an SPI of a 1-month time scale and a forecast period of one month; Model 2 reached the highest model rating of 0.74 at a 12-month time scale and a forecast period of one month; and Model 3 performed well at an SPI of 1-month time scale and a 12-month time scale. Overall, Model 3 had the highest rating, and Model 1 had the lowest rating for the same SPI time scale and the same forecast period, which also indicated that the forecast accuracy of the conditional distribution model considering the *HI* was higher for short-term forecasts with a forecast period of 3 months or less, and including the *HI* could further improve the forecast accuracy of the model.

Table 4 Model evaluation (Model 1: Multivariate normal distribution model; Model 2: Conditional distribution model; Model 3: Conditional distribution model with the *HI*)

Model type	Lead time *M*	SPI time scale			
		1	3	6	12
Model 1	1	0.36	0.36	0.28	0.22
	2	0.11	0.35	0.27	0.22
	3	0.02	0.34	0.26	0.22
Model 2	1	0.69	0.52	/	0.74
	2	0.69	0.47	/	0.67
	3	0.69	0.44	0.39	0.60
Model 3	1	0.72	0.64	0.59	0.71
	2	0.71	0.64	0.59	0.71
	3	0.72	0.64	0.60	0.71

5 Conclusions

Many studies have noted that human activities have a significant impact on watershed runoff in the Luanhe River basin. In this paper, three probability models were constructed to calculate the transition probabilities from the current SPI classes or values to the future SRI classes; then, a scoring mechanism was applied to evaluate the performance of the models.

The calculation results of Model 1 and Model 2 showed that the future hydrological drought classes were more likely to be the same as the meteorological drought classes in the current period, and they were more significant in Model 2. In addition, it was obvious that the drought transition probabilities of Model 3 were significantly different from those of Model 1 and Model 2. Under the condition of considering the *HI*, the results of the drought transition probability showed that when $HI<0$, the future hydrological drought classes tended to normal status, and this situation was more obvious with the decrease in the *HI values*, which indicates that human activities mitigate the degree of hydrological drought when $HI<0$. However, when $HI>0$, the future hydrological drought classes generally transitioned to more severe drought with increasing *HI* values. Thus, it was indicated that human activities exacerbate the degree of hydrological drought as $HI>0$.

Finally, a scoring mechanism was applied to the evaluation of the models, and the forecast results of the three models were evaluated. The results demonstrate that when the SPI time scale was the same, the scores of Model 1 and Model 2 decreased as the forecast period lengthened. In most cases, Model 2 performed better than Model 1, and the performance of Model 3 was the most stable of the three models and had the highest score. The conditional probability model considering the *HI* was more suitable for the Luanhe River basin, where human activities have a high influence.

Although this study has made some progress in the forecasting of hydrological drought in a changing environment, only the *HI* was considered as the exogenous variable in this paper, and human activities were generalized. In future studies, the *HI* can be analysed specifically, for example, the impact of land use and socioeconomics, on drought prediction can be specifically analysed. In addition, climate factors can be further considered in future research.

Ethical Approval: This work meets the ethical and moral requirements.

Consent to Participate: M L., MF Z., RX C., YD S., and XY D. all agreed to participate in the research for the article.

Consent to Publish: M L., MF Z., RX C., YD S., and XY D. all agreed to publish this article.

Author Contributions:

M L (First Author and Corresponding Author): Conceptualization, Methodology, Software, Investigation, Formal Analysis, Writing-Original Draft;

MF Z: Data Curation, Writing-Original Draft;

RX C: Visualization, Investigation;

YD S: Resources, Supervision;

XY D: Visualization, Writing-Review & Editing.

Competing interest: M L., MF Z., RX C., YD S., and XY D. all declare that there are no conflicts of interest.

Data availability statement: We are grateful to the Hydrology and Water Resource Survey Bureau of Hebei Province for providing runoff data. The data and materials of the research are available.

References

Abbasi, A., Khalili K., Behmanesh, J., and Shirzad, A.: Estimation of ARIMA model parameters for drought prediction using the genetic algorithm, Arabian Journal of Geosciences., 14(10), 841-841, https://doi.org/10.1007/s12517-021-07140-0, 2021.

Aghelpour, P., and Varshavian, Vahid.: Forecasting Different Types of Droughts Simultaneously Using Multivariate Standardized Precipitation Index (MSPI), MLP Neural Network, and Imperialistic Competitive Algorithm (ICA), COMPLEXITY., 2021, https://doi.org/10.1155/2021/6610228, 2021.

Ahnadi, M.: Climatic drought forecasting using artificial neural network in Hamedan region, New York Science Journal., 4(8),15-19, https://doi.org/10.7537/marsnys040811.03, 2011.

Alquraish, M. A., Abuhasel, K. S., and Alqahtani, A. K.: SPI-Based Hybrid Hidden Markov–GA, ARIMA–GA, and ARIMA–GA–ANN Models for Meteorological Drought Forecasting, Sustainability., 13(22), 12576-12576, https://doi.org/10.3390/su132212576, 2021.

Behzad, A., and Hamid, M.: Revisiting hydrological drought propagation and recovery considering water quantity and quality, Hydrological Processes., 33(10), 1492-1505, https://doi.org/10.1002/hyp.13417, 2019.

Bishara, A. J., and Hittner, J. B.: Testing the significance of a correlation with nonnormal data: comparison of Pearson, Spearman, transformation, and resampling approaches, Psychological methods., 17(3), 399-417, https://doi.org/10.1037/a0028087, 2012.

Bonaccorso, B., Cancelliere, A., and Rossi G.: Probabilistic forecasting of drought class transitions in Sicily (Italy) using Standardized Precipitation Index and North Atlantic Oscillation Index, Journal of Hydrology., 526, 136-150, https://doi.org/10.1016/j.jhydrol.2015.01.070, 2015.

Carmelo, C., and Jürgen, V. V.: Non-stationarity in MODIS fAPAR time-series and its impact on operational drought detection, International Journal of Remote Sensing., 40(4), 1428-1444, https://doi.org/10.1080/01431161.2018.1524603, 2018.

Chang, G. B., Zhang, S. B., and Liu, Z. P.: Understanding the adjusment from an information theoretic perspective, Geomatics and Information Science of Wuhan University., 2022,1-17, https://kns.cnki.net/kcms/detail/42.1676.TN.20211125.1422.006.html, 2022. (in Chinese)

Charles, O.: On Rigorous Drought Assessment Using Daily Time Scale: Non-Stationary Frequency Analyses, Revisited Concepts, and a New Method to Yield Non-Parametric Indices, Hydrology., 4(4), 48-48, https://doi.org/10.3390/hydrology4040048, 2017.

Chen, J., Brissette, F. P., Chaumont, D., and Braun, M.: Performance and uncertainty evaluation of empirical downscaling methods in quantifying the climate change impacts on hydrology over two North American river basins, Journal of Hydrology., 479, 200-214, https://doi.org/10.1016/j.jhydrol.2012.11.062, 2013.

Chen, X., Han, R. G., Feng, P., and Wang, Y. J.: Combined effects of predicted climate and land use changes on future hydrological droughts in the Luanhe River basin, China, Natural Hazards., 2021, 1-33, https://doi.org/10.1007/s11069-021-04992-3, 2021.

Ding, Y. B., Xu, J. T., Wang X. W., Cai, H.J., Zhou Z.Q., Sun Y.N., and Shi H.Y.: Propagation of meteorological to hydrological drought for different climate regions in China, Journal of Environmental Management., 283, 111980-111980, https://doi.org/10.1016/j.jenvman.2021.111980Get rights and content, 2021.

Dixit, S., and Jayakumar, K. V.: A Non-stationary and Probabilistic Approach for Drought Characterization Using Trivariate and Pairwise Copula Construction (PCC) Model, Water

Resources Management., 2022, 1-20, https://doi.org/10.1007/s11269-022-03069-5, 2022.

Fuentes, I., Padarian J., Vervoort, R. W.: Spatial and Temporal Global Patterns of Drought Propagation, Frontiers in Environmental Science., 2022,788248-788248, https://doi.org/10.3389/fenvs.2022.788248, 2022.

Zhao, G., Gao, H. L., Kao, S. C., Nathalie V., Bibi S. N.: A modeling framework for evaluating the drought resilience of a surface water supply system under non-stationarity, Journal of Hydrology., 563, 22-32, https://doi.org/10.1016/j.jhydrol.2018.05.037, 2018.

Gong, H. N., Xie, B. T., Wang J. R.: Long-term prediction of extreme response of deepwater floating platform based on environmental contour method, The Ocean Engineering., 39(05), 28-38, https://doi.org/10.16483/j.issn.1005-9865.2021.05.003, 2021. (in Chinese)

Kolachian, R., and Saghafian B.: Hydrological drought class early warning using support vector machines and rough sets, Environmental Earth Sciences., 80(11), 390-390, https://doi.org/10.1007/s12665-021-09536-3, 2021.

Koudahe, K., Koffi, D., Kayode, J. A., Awokola, S. O., Adebola, A. A.: Impact of Climate Variability on Crop Yields in Southern Togo, Environment Pollution and Climate Change., 2(1): 1-9, https://doi.org/10.4172/2573-458X.1000148 , 2018.

Li, J. Z., Wang, Y. X., Li, S. F., and Hu, R.: A Nonstationary Standardized Precipitation Index incorporating climate indices as covariates, Journal of Geophysical Research: Atmospheres., 120(23), 082-095, https://doi.org/10.1002/2015JD023920, 2015.

Li, L. C., Yao, N., Li, Y., Liu D. L., Wang, B., and Ayantobo, O. O.: Future projections of extreme temperature events in different sub-regions of China, Atmospheric Research., 217, 150-164, https://doi.org/10.1016/j.atmosres.2018.10.019, 2018.

Li, X., Fang, G. H., Wen, X., Xu, M., and Zhang, Y.: Characteristics analysis of drought at multiple spatiotemporal scale and assessment of CMIP6 performance over the Huaihe River Basin, Journal of Hydrology: Regional Studies., 41, 101-103, https://doi.org/10.1016/j.ejrh.2022.101103, 2022.

Li, X. X., Ma, X. X., Li, X. D., and Zhang, W. J.: Method Consideration of Variation Diagnosis and Design Value Calculation of Flood Sequence in Yiluo River Basin, China, Water., 12(10), 2722-2722, https://doi.org/10.3390/w12102722, 2020.

Majid, D., Bahram, S., and Mansoor, Z.: Probabilistic hydrological drought index forecasting based on meteorological drought index using Archimedean copulas, Hydrology Research., 50(5), 1230-1250, https://doi.org/10.2166/nh.2019.051, 2019.

Malede, D. A., Agumassie, T.A., Kosgei, J.R., Linh, N. T. T., and Andualem, T. G.: Analysis of rainfall and streamflow trend and variability over Birr River watershed, Abbay basin, Ethiopia Environmental Challenges., 7, 100528-100528, https://doi.org/10.1016/J.ENVC.2022.100528, 2022.

Ganeshchandra, M., Shivam, T., Sergey, K., and Rao S. G.: Probabilistic Assessment of Drought Characteristics Using Hidden Markov Model, Journal of Hydrologic Engineering., 18(7), 834-845. https://doi.org/10.1061/(ASCE)HE.1943-5584.0000699, 2013.

Maryam, M., Farzad, E., Nima, J. V., and Alireza, A.: Drought forecasting by ANN, ANFIS, and SVM and comparison of the models, Environmental Earth Sciences., 76(21), 729-729, https://doi.org/10.1007/s12665-017-7064-0, 2017.

Mckee, T. B., Doesken, N. J., Kleist, J.: The relationship of drought frequency and duration to time scale, in: Proceedings of the Eighth Conference on Applied Climatology. American Meteorological Society, pp. 179–184. 1993.

Mehdi R., Alfred S., and Jonathan P. C.: Drought Forecasting using Markov Chain Model and Artificial Neural Networks, Water Resources Management., 2016, 30(7), 2245-2259, https://doi.org/10.1007/s11269-016-1283-0, 2016.

Melanie, O., Francisco, J. M., Jorge, G., Christopher, A. S., Facundo, R., and Nicolás, P.: Drought Propagation in Semi-Arid River Basins in Latin America: Lessons from Mexico to the Southern Cone, Water., 10(11), 1564-1564. https://doi.org/10.3390/w10111564, 2018.

Miriam, F., Tobias, G., Lívia, L., Dana, V., Zuzana, D., Marián, F., and Pavla, P.: Analysing 21st century meteorological and hydrological drought events in Slovakia, Journal of Hydrology and Hydromechanics., 66(4), 393-403, https://doi.org/10.2478/johh-2018-0026, 2018.

Mohammad, M. M., Abdol, R. Z., and Mohammad, R. M.: Seasonal drought forecasting in arid regions, using different time series models and RDI index, Journal of Water and Climate Change.,11(3), 633-654, https://doi.org/10.2166/wcc.2019.009, 2020.

Muhammad, J., Sabab, A.S., Jiyoung, Y., and Tae-Woong, K.: Investigating the impacts of climate change and human activities on hydrological drought using non-stationary approaches, Journal of Hydrology., 588, 125052-125052, https://doi.org/10.1016/j.jhydrol.2020.125052, 2020.

Natsagdorj, E., Renchin, T., Maeyer, P. D., and Darkhijav, B.: Spatial Distribution of Soil Moisture in Mongolia Using SMAP and MODIS Satellite Data: A Time Series Model (2010–2025), Remote Sensing., 13(3), 347-347, https://doi.org/10.3390/RS13030347, 2021.

Ren, W., Wang, Y., Li, J., Feng, P., and Smith R. J.: Drought forecasting in Luanhe River basin involving climatic indices, Theoretical and Applied Climatology., 130(3-4), 1133-1148, https://doi.org/10.1007/s00704-016-1952-1, 2017.

Shraddhanand, S., and Andrew, W. W.: Use of a standardized runoff index for characterizing hydrologic drought, Geophysical Research Letters., 35(2), 2405-1-2405-7, https://doi.org/10.1029/2007GL032487, 2008.

Stojković, M., Plavšić, J., Prohaska, S., Pavlović, D, and Despotović, J.: A two-stage time series model for monthly hydrological projections under climate change in the Lim River basin (southeast Europe), Hydrological Sciences Journal., 65(3), 387-400, https://doi.org/10.1080/02626667.2019.1699241, 2020.

Vicente-Serrano, S., Beguería, S., and López-Moreno, J.: A Multiscalar Drought Index Sensitive to Global Warming: The Standardized Precipitation Evapotranspiration Index, Journal of Climate., 23(7), https://doi.org/10.1175/2009JCLI2909.1, 2010.

Wang, M., Jiang, S., Ren, L., Xu, C-Y., Menzel, L., Yuan, F., Xu, Q., Liu, Y., Yang, X.: Separating the effects of climate change and human activities on drought propagation via a natural and human-impacted catchment comparison method, Journal of Hydrology., 603(PA), 126913-123613, https://doi.org/10.1016/J.JHYDROL.2021.126913, 2021.

Wang, Y., Duan, L., Liu, T., Li, J., and Feng, P.: A Non-stationary Standardized Streamflow Index for hydrological drought using climate and human-induced indices as covariates, Science of the Total Environment., 699(C), 134278-134278, https://doi.org/10.1016/j.scitotenv.2019.134278, 2020.

Wang, Y., Li, J., Feng, P., and Hu, R.: Analysis of drought characteristics over Luanhe River basin using the joint deficit index, Journal of Water and Climate Change., 7(2), 340-352, https://doi.org/10.2166/wcc.2015.108, 2016.

Wang, Y., Li, J., Feng, P, and Chen, F.: Effects of large-scale climate patterns and human activities on hydrological drought: a case study in the Luanhe River basin, China, Natural Hazards.,76(3), 1687-1710, https://doi.org/10.1007/s11069-014-1564-y, 2015.

Wang, Y., Zhang, T., Chen, X., Li, J., and Feng, Ping.: Spatial and temporal characteristics of droughts in Luanhe River basin, China, Theoretical and Applied Climatology., 131(3), 1369-1385, https://doi.org/10.1007/s00704-017-2059-z, 2018.

Wang, Y., Peng, T., Lin, Q., Singh, V., Dong, X., Chen, C., Liu, J., Chang, W., and Wang, G.: A New Non-stationary Hydrological Drought Index Encompassing Climate Indices and Modified Reservoir Index as Covariates, Water Resources Management., 36(7), 2433-2454, https://doi.org/10.1007/S11269-022-03151-Y, 2022.

Xu, Y., Zhang, X., Wang, Xiao., Hao, Z., Singh, V., and Hao, F.: Propagation from meteorological drought to hydrological drought under the impact of human activities: A case study in northern China, Journal of Hydrology., 579(C), 124147-124147, https://doi.org/10.1016/j.jhydrol.2019.124147, 2019.

Yan, X., Bao, Z., Zhang, J., Wang, G., He, R., and Liu, C.: Quantifying contributions of climate change and local human activities to runoff decline in the upper reaches of the Luanhe River basin, Journal of Hydro-environment Research., 28(C), 67-74, https://doi.org/10.1016/j.jher.2018.11.002, 2018.

Zhang, T., et al.: Drought class transition analysis through different models: a case study in North China, Water Science & Technology: Water Supply., 17(1), 138-150, https://doi.org/10.2166/ws.2016.123, 2017.