Environmental Engineering and Management Journal

September 2020, Vol.19, No. 9, 1467-1474
http://www.eemj.icpm.tuiasi.ro/; http://www.eemj.eu

"Gheorghe Asachi" Technical University of Iasi, Romania

DROUGHT LEVEL PREDICTION BASED ON LOG-LINEAR MODEL

Chunyan Huang[1,2]

[1]*School of Mathematics and Statistics, North China University of Water Resources and Electric Power, Zhengzhou, Henan, 450011, China*
[2]*State Key Laboratory Base of Eco-Hydraulic Engineering in Arid Area, Xi'an University of Technology, Xi'an 710048, China*
E-mail address: huangchunyan@ncwu.edu.cn; Phone: +8613938264486

Abstract

Drought prediction is fundamental to the management of drought risks. This paper collects the monthly precipitation data in 1957~2009 from 19 meteorological stations in the Yellow River Basin, and sets up a log-linear model to predict the drought level in the study period using the standardized precipitation index (SPI). Based on the SPI drought classification criterion, the drought level sequences were established, the frequency of drought level conversion was determined, and the dominance and its confidence interval were obtained by the log-linear model. In this way, meteorological drought level in one to two months was predicted. The model verification shows that the expected frequency agrees well with the observed frequency, and the predicted drought level was in line with the measured level. This means our model can accurately forecast the drought level in one to two months, and can be used for early warning of drought in the short term.

Keywords: drought level, log-linear model, standardized precipitation index (SPI)

Received: October, 2019; Revised final: February, 2020; Accepted: March, 2020; Published in final edited form: September, 2020

1. Introduction

Drought, as one of the natural calamities that slowly develop from a complex origin which occurs in the broadest area most frequently, seriously interferes with sustainable development of human society, ecological environment and it presents one of the most severe impacts on natural disasters (Heim, 2002; Liu et al., 2019). The spread of global warming makes the drought exacerbate. We should effectively and precisely forecast the drought to provide a powerful scientific basis for technical measures such as drought early warning and monitoring.

The drought indices underlie quantitative analysis, monitoring, surveying and prediction of drought. With drought indices, we can discern statistical properties of drought, and take appropriate measures against losses caused by it. There are commonly used drought indices, including: precipitation anomaly percent (%) (PAP) (Bhalme and Mooley, 1980; Harman and Elton, 2015), precipitation Z index (Z-index) (Chen et al., 2013; Wu et al., 2001), Palmer drought severity index (PDSI) (Palmer, 1965; Vicente-Serrano et al., 2011; Yang et al., 2018), standardized precipitation index (SPI) (Mckee et al., 1993). SPI features multiple time scales (1, 3, 6, 12 months, etc.). It is widely applied thanks to its simplicity and good stability (McKee et al., 1995; Morid et al., 2010). Forecasts of hydro logic droughts can help in decision making in agriculture and water resources management (Liang et al., 2019; Nicholas and Battisti, 2008; Shah and Mishra, 2016; Yusof et al., 2019). But traditional drought prediction methods include the numerical and statistical prediction types, which mainly adopt the mathematical tools to predict and analyze the probability that drought will occur in the future but at a relatively low rate. The features available during the calculation of the log-linear model (Paulo et al., 2005) and its confidence interval can improve the prediction precision. In addition, the log-linear model enables concise prediction algorithm, objective and simple. With this model, drought level transitions are useful for short term drought warning in the study area.

In this paper, The Yellow River basin was selected as the research object. The Yellow River Basin (YRB) is located between 32ºN–42ºN and 96ºE–119ºE (Fig. 1), controlling a drainage area of 795,000 km 2. From northwest to southeast, the elevation presents a gradually decreased pattern and ranges between 1 and 6199 m above the sea level. The mean annual temperature varies between 4 and 14 ºC.

Precipitation is unevenly distributed over YRB which divides the whole basin into four climate zones from northwest to southeast. In addition, precipitation and temperature have significant seasonality, where summer is generally rainy and hot while winter is cold and dry.

Therefore, this paper attempts to predict the meteorological drought levels by establishing a log-linear model with 3D contingency Table. The SPI of relevant stations in the Yellow River Basin are aimed, providing the clues to fighting against drought and reducing losses caused by the drought calamity.

2. Materials and methods

2.1. SPI

The Standardized Precipitation Index (SPI), proposed by American scholar Mckee in 1993 (McKee et al., 1993), is widely used to describe and compare droughts occurred in different periods and under different climatic conditions. It adopts the Γ (gamma) distribution probability to describe the change in annual precipitation, which can be available from the normal standardization. With multiple time scale, the SPI can supervise, compare and assess different climate conditions at different time scales in an area (Edwards and McKee, 1997). The SPI at 12-month time scale clearly reflects periodic drought changes and the effects that meteorological drought plays on hydrological mechanism and water resources. Table 1 includes the classification standard for drought grades (Moreira et al., 2008; Patel et al., 2007).

2.2. 3D log-linear model

The log-linear model enables to analyze discrete data or data converted in the form of contingency table. There are 3 types of different classification attributes for 3D contingency table: marked as a, b, c. Among them, the attribute a has i *levels*; the attribute b has j *level*s; the attribute c has k *level*s; i, j, k, $1\in\{1, 2, 3, 4\}$. These three attributes represent the drought levels in$t-1$, t and $t+1$ month, respectively. The drought level can be represented by 1-4, that is, 1 means "no drought"; 2 means "light drought"; 3 means "moderate drought"; 4 means "heavy drought/severe drought".

When the drought levels in adjacent months $t-1$, t and $t+1$ are i, j and k, respectively, o_{ijk} in Table 2 represents the frequency appeared. A drought level transition 3D contingency table is shown in Table 2. For example, o_{111} in the table represents the actual frequency when the drought levels in adjacent three months are averaged as 1.

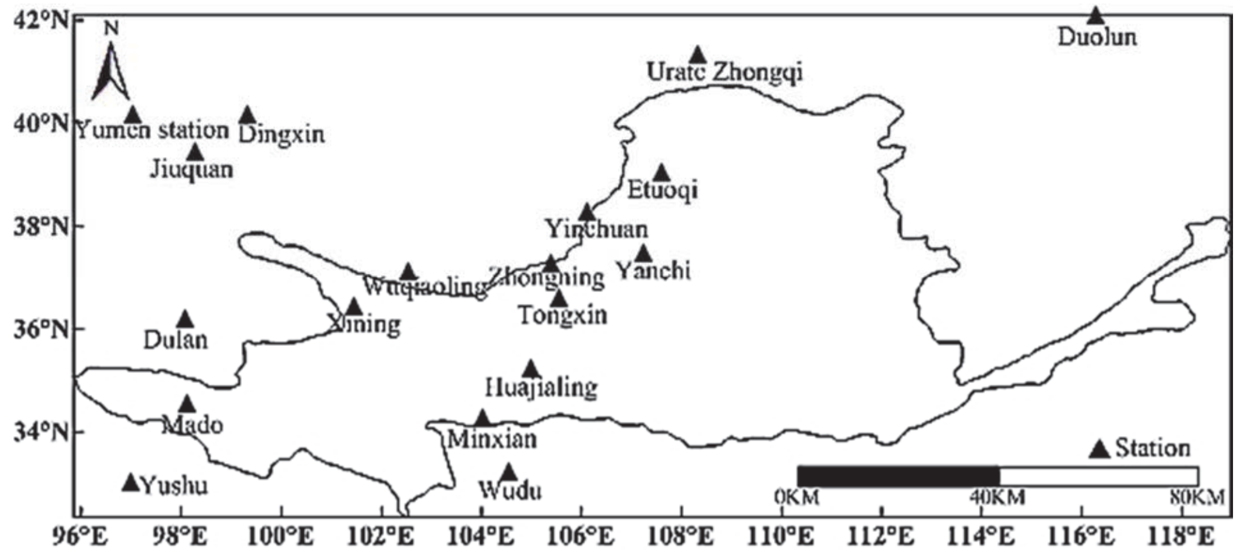

Fig. 1. Location of meteorological station over upper yellow river basin

Table 1. The classification standard for drought of SPI

drought grade	*Type of drought*	*SPI*
1	no drought	SPI≥0
2	light drought	-1≤SPI<0
3	moderate drought	-1.5<SPI≤-1
4	heavy drought /extreme drought	SPI≤-1.5

Table 2. 3D contingency table for transitions between drought classes

drought grade of T-1 month	***drought grade of T month***															
	1				***2***				***3***				***4***			
	drought grade of T+1 month				***drought grade of T+1 month***				***drought grade of T+1 month***				***drought grade of T+1 month***			
	1	***2***	***3***	***4***	***1***	***2***	***3***	***4***	***1***	***2***	***3***	***4***	***1***	***2***	***3***	***4***
1	O_{111}	O_{112}	O_{113}	O_{114}	O_{121}	O_{122}	O_{123}	O_{124}	O_{131}	O_{132}	O_{133}	O_{134}	O_{141}	O_{142}	O_{143}	O_{144}
2	O_{211}	O_{212}	O_{213}	O_{214}	O_{221}	O_{222}	O_{223}	O_{224}	O_{231}	O_{232}	O_{233}	O_{234}	O_{241}	O_{242}	O_{243}	O_{244}
3	O_{311}	O_{312}	O_{313}	O_{314}	O_{321}	O_{322}	O_{323}	O_{324}	O_{331}	O_{332}	O_{333}	O_{334}	O_{341}	O_{342}	O_{343}	O_{344}
4	O_{411}	O_{412}	O_{413}	O_{414}	O_{421}	O_{422}	O_{423}	O_{424}	O_{431}	O_{432}	O_{433}	O_{434}	O_{441}	O_{442}	O_{443}	O_{444}

The saturated form of 3D log-linear model is Eq. (1):

$$lnE_{ijk} = \lambda + \lambda_i^a + \lambda_j^b + \lambda_k^c + \lambda_{ij}^{ab} + \lambda_{jk}^{bc} + \lambda_{ik}^{ac} + \lambda_{ijk}^{abc} \tag{1}$$

where: E_{ijk} represents the frequency as the drought level transition model predicts; λ is a constant; λ^a_i, λ^b_j and λ^c_k represent the main effects of the drought levels in the months *t-1*, *t* and *t+1*, respectively; λ^{ab}_{ij}, λ^{bc}_{jk} and λ^{ac}_{ik} represent the interaction effects (called the first-order interaction effect) between drought levels in the months *t-1*, *t* and *t+1*, respectively; λ^{abc}_{ijk} represents the interaction effect (called secondary interaction effect) among three drought levels in the months *t-1*, *t* and *t+1*, where: $i, j, k, l \in \{1, 2, 3, 4\}$.

2.3. Test on model goodness of fit

The goodness of fit is tested using χ^2 (Agresi, 1990). The degree of freedom of the contingency tables available by subtracting the number of independent parameters from the total number of cells. The test statistics is the logarithm of the ratio of the actual observation frequency (0_{ijk}) to the expected frequency (E_{ijk}) of each cell, which can be used to measure the deviation of actual observation frequency from the expected frequency, called the likelihood ratio statistics G^2, as shown in Eq (2):

$$G^2 = 2\sum_i \sum_j \sum_k O_{ijk}\, log(O_{ijk} / E_{ijk}) \tag{2}$$

The null hypothesis H_0: the model better fits data. By consulting the distribution table χ^2, the boundary value $\chi^2_{\alpha,\ \nu}$ is available when $\alpha=0.05$ and the degree of freedom is ν. When $G^2<\chi^2_{\alpha,\nu}$, and the significance level $P>\alpha$, it is considered that the actual observation frequency has a higher similarity to the frequency model-predicted frequencies, then the null hypothesis is true.

2.4. Dominance and confidence interval

The dominance is the ratio of the model-predicted frequencies available when a drought level is transferred into two other drought levels. It can represent the probabilities when it is converted into certain two drought levels, as shown in Eq. (3):

$$\Omega_{kl|ij} = E_{ijk} / E_{jkl} \tag{3}$$

where: i, j, k, $l \in \{1,2,3,4\}$ and $k \neq 1$.

The logarithmic form of the dominance is shown in Eq. (4):

$$ln\Omega_{kl|ij} = lnE_{ijk} - lnE_{ijl} \tag{4}$$

When the samples of $\Omega_{kl|ij}$ are multiple, it is considered that $\Omega_{kl|ij}$ roughly obeys the normal distribution; for $\ln\Omega_{kl|ij}$, it can approximate the normal distribution at a faster convergence rate. Therefore, by calculating the estimated value of the asymptotic standard error $\sqrt{Var\ (ln\Omega_{kl|ij})}$, the asymptotic confidence interval 1-*a* of $ln\Omega_{kl|ij}$ can be available (Eq. 5).

$$[\ln\Omega_{kl|ij} - Z_{1-\alpha/2}\sqrt{\widehat{\mathrm{Var}}(\ln\Omega_{kl|ij})},\quad \ln\Omega_{kl|ij} + Z_{1-\alpha/2}\sqrt{\widehat{\mathrm{Var}}(\ln\Omega_{kl|ij})}), \tag{5}$$

where: $Z_{1-\alpha/2}$ is the quartile of 1-α/2 for the normal distribution.

The confidence interval of dominance $ln\Omega_{kl|ij}$ is (Eq. 6):

$$[\exp(\ln\Omega_{kl|ij} - Z_{1-\alpha/2}\sqrt{\widehat{\mathrm{Var}}(\ln\Omega_{kl|ij})}),\quad \exp(\ln\Omega_{kl|ij} + Z_{1-\alpha/2}\sqrt{\widehat{\mathrm{Var}}(\ln\Omega_{kl|ij})})] \tag{6}$$

When the dominance confidence interval contains the value 1, it represents the probabilities the drought level transitions i→j→k and i→j→l occur are roughly equal; when the values in the confidence interval of the dominance are all greater than (less than) 1, the probability the drought level transition i→j→k occurs is greater than (less than) that i→j→l does. If the dominance confidence interval is too wide, the interval estimation accuracy of the dominance is lower, and the probability the drought level transition i→j→k occurs is extremely low.

In the 3D contingency table, there are a total of 96 dominances. Among them, there are $\Omega_{12|ij}$, $\Omega_{13|ij}$, $\Omega_{14|ij}$, $\Omega_{23|ij}$, $\Omega_{24|ij}$, $\Omega_{34|ij}$ which contains 16 sub-dominances. In the adjacent months, t-1, t and t+1, if *t* represents the current month, the drought levels in the current month (t) and the previous month (t-1) are known, then the dominances and confidence intervals can be available to realize the drought level capacity after 1 month (t+1) and two months (t+2).

3. Results

3.1. Establishment of log-linear model

(1) Data pre-processing

With the SPI at 19 meteorological stations in the Yellow River Basin, the relevant drought level sequences can be available. The transition frequency of each drought level is calculated and a 3D contingency Table for drought level transition is established. For example, the Dingxin Station, the 3D contingency Table of the observation frequency is shown in Table 3.

(2) Analysis of model parameters

The high-dimensional interaction in the model is statistically tested using the likelihood ratio chi-square and Pearson chi-square. The results of Dingxin Station are shown in Table 4. If the significance levels from the third-order likelihood ratio and the Pearson chi-square tests are both greater than 0.05, both tests show that the 3D interaction has no statistical significance. If the significance from the second-order likelihood ratio and Pearson chi-square tests are 0.000 (<0.05), it is considered that the 2D interaction effect and the main effect are statistically significant.

The parameter estimations for the Dingxin station model are shown in Table 5. The results from the goodness of fit test are shown in Table 6, and the expected frequency of the drought level transition is shown in Table 7. The constant λ is 3.481; the main effects in *t-1* are λ_1^a, λ_2^a and λ_3^a, respectively; the main effects in *t* are λ_1^b, λ_2^b and λ_3^b, respectively; the main effects in *t+1* are λ_1^c, λ_2^c and λ_3^c, respectively; the interaction effects in *t-1* and *t* are λ_{11}^{ab}, λ_{12}^{ab}, λ_{13}^{ab}, λ_{21}^{ab}, λ_{22}^{ab}, λ_{23}^{ab}, λ_{31}^{ab}, λ_{32}^{ab} and λ_{33}^{ab}, respectively; the interaction effects in *t* and *t+1* are λ_{11}^{bc}, λ_{12}^{bc}, λ_{13}^{bc}, λ_{21}^{bc}, λ_{22}^{bc}, λ_{23}^{bc}, λ_{31}^{bc}, λ_{32}^{bc} and λ_{33}^{bc}, respectively; the interaction effects in *t-1* and *t+1* are λ_{11}^{ac}, λ_{12}^{ac}, λ_{13}^{ac}, λ_{21}^{ac}, λ_{22}^{ac}, λ_{23}^{ac}, λ_{31}^{ac}, λ_{32}^{ac} and λ_{33}^{ac},respectively.

As shown in Table 6, the degree of freedom of contingency Table is 27, the likelihood ratio statistic $G^2=24.382<\chi^2_{0.05,27}=40.113$, and the significance level $P=0.609>0.05$. From Table 7, it is found that the predicted frequency of model well fits the actually observed frequency, this means that the model has a good fit.

Table 3. 3D contingency Table for transitions between drought classes of Dingxin station

drought grade of T-1 month	***drought grade of T month***															
	1				***2***				***3***				***4***			
	drought grade of T+1 month															
	1	***2***	***3***	***4***	***1***	***2***	***3***	***4***	***1***	***2***	***3***	***4***	***1***	***2***	***3***	***4***
1	291	17	1	0	3	21	1	0	0	1	0	0	0	0	0	0
2	17	7	0	0	17	98	12	3	1	3	9	4	0	0	0	3
3	2	0	0	0	3	9	3	0	1	11	36	3	0	0	2	8
4	1	0	0	0	1	0	1	0	0	0	6	3	1	4	7	32

Table 4. Test of Model K-Level Interaction effect

K		***df***	***likelihood ratio***		***Pearson***		***Iterative number***
			chi square	***Sig.***	***chi square***	***Sig.***	
K-Way and High-order effects[a]	1	63	2511.276	0	9176.869	0	0
	2	54	1642.329	0	5015.165	0	2
	3	27	24.382	0.609	22.059	0.734	8
K-way effects[b]	1	9	868.947	0	4161.704	0	0
	2	27	1617.947	0	4993.106	0	0
	3	27	24.382	0.609	22.059	0.734	0

Table 5. Parameter estimation of model for Dingxin station

Parameter	***Estimate of Parameters (λ=3.481)***		
Main Effects	$\lambda_1^a=-64.975$	$\lambda_2^a=-9.281$	$\lambda_3^a=-3.823$
	$\lambda_1^b=-64.975$	$\lambda_2^b=-64.975$	$\lambda_3^b=-9.281$
	$\lambda_1^c=-21.672$	$\lambda_2^c=-7.221$	$\lambda_3^c=-4.332$
Interactive Effects	$\lambda_{11}^{ab}=64.975$	$\lambda_{12}^{ab}=64.975$	$\lambda_{13}^{ab}=9.281$
	$\lambda_{21}^{ab}=9.281$	$\lambda_{22}^{ab}=64.975$	$\lambda_{23}^{ab}=11.941$
	$\lambda_{31}^{ab}=3.823$	$\lambda_{32}^{ab}=3.823$	$\lambda_{33}^{ab}=3.823$
	$\lambda_{11}^{bc}=64.975$	$\lambda_{12}^{bc}=7.221$	$\lambda_{13}^{bc}=4.332$
	$\lambda_{21}^{bc}=64.975$	$\lambda_{22}^{bc}=7.221$	$\lambda_{23}^{bc}=13.001$
	$\lambda_{31}^{bc}=3.096$	$\lambda_{32}^{bc}=1.031$	$\lambda_{33}^{bc}=8.045$
	$\lambda_{11}^{ac}=21.672$	$\lambda_{12}^{ac}=7.221$	$\lambda_{13}^{ac}=4.332$
	$\lambda_{21}^{ac}=3.096$	$\lambda_{22}^{ac}=1.031$	$\lambda_{23}^{ac}=-1.616$
	$\lambda_{31}^{ac}=1.275$	$\lambda_{32}^{ac}=-2.354$	$\lambda_{33}^{ac}=1.275$

Table 6. Goodness- of- fit test method for Dingxin station

	chi square	***df***	***Sig.***
likelihood ratio	24.382	27	0.609
Pearson	22.059	27	0.734

Table 7. Expected frequencies of drought class transitions for Dingxin station

drought grade of T-1 month	***drought grade of T month***															
	1				***2***				***3***				***4***			
	drought grade of T+1 month															
	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4
1	288	20	1	0	6	18	1	0	0	1	0	0	0	0	0	0
2	20	3	0	0	14	100	13	3	0	4	8	5	0	0	0	3
3	2	0	0	0	3	9	3	0	1	9	36	4	0	1	2	7
4	1	0	0	0	1	1	0	0	0	1	7	1	1	2	7	34

3.2. Dominance and confidence interval

At Dingxin Station, the confidence intervals of the dominance $\Omega_{12|ij}=E_{ij1}/E_{ij2}$ and $\Omega_{23|ij}=E_{ij2}/E_{ij3}$ are shown in Tables 8 and 9. In the two tables, the values at the upper level of each cell represent the dominances, and those at the lower level are the confidence intervals. The dominance $\Omega_{12|ij}$ represents the ratio of the drought level transitions i→j→1 to i→j→2. If the dominance is greater than 1, the probability that the former transition (i.e. i→j→1) occurs is greater than that the latter transition (i.e. i→j→2) does. It can be used to finish the later predictions. For example, the value of the dominance $\Omega_{12|11}$ is 14.097, and the confidence interval is [1.295, 766.107], the values are all greater than 1, this means the possibilities that drought level transitions 1→1→1 occur is higher than that 1→1→2 does; the value of dominance $\Omega_{23|11}$ is 23.614, the confidence interval is [1.558, 446.519] in which the values are all greater than 1, this means that the probability of drought level transition 1→1→2 occurs is higher than that 1→1→3 does. As described above, it is known that the dominances $\Omega_{12|11}$ and $\Omega_{23|11}$ are both greater than 1, so the probability the drought level transition 1→1→1 occurs is greater than that 1→1→3 does.

3.3. Test on goodness of fit

The logarithmic model of 3D contingency table at 19 meteorological stations in the upstream of the Yellow River is analyzed. The goodness of fit tested is shown in Table 10. each likelihood ratio statistics G^2 is less than the boundary value $\chi^2_{0.05,27}$=40.113 of χ^2, and the significance levels P of the test are all greater than α=0.05: it is known that the real frequency more fits that predicted by the model. It means that the model has a good fit.

3.4. Analysis of expected frequency

For example, Yumen, Wuqiaoling and Huajialing stations, the expected frequency predicted by log-linear model for the drought level transition is shown in Table 11. As show in Table 11, when the drought level transition is 1→1→1, the frequency predicted by model is the maximum, far greater than that of other drought level transitions.

Table 8. $\Omega_{12|ij}$ and confidence interval for Dinxing station

drought grade of T-1 month	***drought grade of T month***							
	1		***2***		***3***		***4***	
1	14.097		0.330		0.147		1.600	
	1.259	766.107	0.006	17.934	0.003	7.989	0.029	217.382
2	6.149		0.144		0.064		0.047	
	0.113	334.17	0.003	7.826	0.001	3.478	0.001	2.554
3	12.908		0.301		0.134		0.102	
	0.238	701.49	0.006	16.358	0.002	7.282	0.002	5.543
4	44.136		1.052		0.469		0.355	
	0.812	1329.858	0.019	57.171	0.009	25.488	0.007	19.263

Table 9. $\Omega_{23|ij}$ and confidence interval for Dinxing station

drought grade of T-2 month	***drought grade of T-1 month***							
	1		***2***		***3***		***4***	
1	23.614		7.832		1.143		2.071	
	1.558	446.519	0.274	190.613	0.074	28.676	0.218	5.326
2	38.054		9.823		2.34		3.405	
	2.301	773.477	0.513	190.267	0.131	46.392	0.167	67.842
3	2.433		0.552		0.125		0.204	
	0.213	43.217	0.031	11.575	0.005	2.753	0.009	3.657
4	4.569		1.416		0.269		0.24	
	0.425	89.198	0.16	31.597	0.051	5.967	0.022	8.027

Table 10. Goodness- of- fit test method for every station

Station Name	***Degree of Freedom***	***Statistic of Likelihood Ratio***	***P Value***
Yumen station	27	24.918	0.579
Wuqiaoling	27	33.344	0.186
Huajialing	27	28.010	0.410
Jiuquang	27	21.059	0.784
Wudu	27	12.977	0.992
Minxian	27	12.926	0.990
Dingxin	27	24.382	0.714
Duolun	27	31.019	0.270
Urate Zhongqi	27	21.664	0.735
Etuoqi	27	25.338	0.513
YinChun	27	13.682	0.972
Zhongning	27	30.285	0.367
Tongxin	27	22.436	0.624
Yanchi	27	24.358	0.603
Mado	27	19.239	0.861
Xinin	27	32.364	0.214
Duran	27	25.643	0.536
Yushu	27	16.529	0.912

In the contingency table, the four drought level transitions 1→1→1, 2→2→2, 3→3→3, 4→4→4 show a distinct strong diagonal trend. It means that the drought is persistent. That is to say, if there is a drought occurred in an area, there is a high probability that there will be persistent drought in the area for several consecutive years.

Another obvious feature of the table is that the drought evolution is a gradual accumulation process, which does not leap forward, that is, the frequency of occurrence of a large interval of drought level transition is much less than that of a small interval of drought level transition. For example, the frequency at which the drought level transition 1→3→1 occurs is far less than 1→1→1 does.

3.5. Model test

After establishing a log-linear model, the 2010 drought level is predicted and compared with 2010 observation data to test the prediction validity. As shown in Tables 12 and 13, the drought level observations and predicted values at Yumen and Huajialing stations are compared, respectively.

In Tables, we can see that the accuracy of prediction in the period of 2 months is slightly less than that in the period of 1 month, and most of the prediction results in the period of 2 months include two possibilities. In general, the actual drought levels observed in most of the months are consistent with those predicted by the model. Although there are some disparities, the model predicted drought levels are all in proximity to the actual drought level as observed. All others are able to accurately or better predict the drought level. All of these show that the prediction accuracy made by the log-linear model is higher.

The reason why there is an inconsistency between the model-predicted and observed actual drought levels is that when the SPI value in t and t-1 months are at the boundary of the drought levels, the actual drought levels observed in t and t+1 month are sensitive to the changes in rainfall.

Table 11. Expected frequencies of drought class transitions for Yumen station, Wuqiaoling station and Huajialing station

drought grade of T -1	***drought grade of T month***															
	1				***2***				***3***				***4***			
	drought grade of T+1 month															
	1	***2***	***3***	***4***	***1***	***2***	***3***	***4***	***1***	***2***	***3***	***4***	***1***	***2***	***3***	***4***
Yumen station																
1	**296.3**	20.3	1.3	0	10.7	16.6	0.7	0	0	1.1	0.9	0	0	0	0	0
2	22.1	7.3	0.6	0	17.8	**132.4**	6.9	0.9	0	3.3	3.4	2.4	0.1	0.1	0.1	0.7
3	0	0	0	0	1.4	7.6	1	0.1	0	4.7	**11.9**	6.4	0.6	0.7	2.1	8.5
4	0.5	0.3	0.1	0	0.1	1.5	0.4	0	0	1	5.7	3.2	0.4	1.2	7.7	**32.8**
Wuqiaoling station																
1	**259.9**	25.5	0.7	1.8	12	21.3	0.9	0.7	0.1	0.3	0.3	0.3	0	0.1	0	2
2	25.5	9.2	0.2	0.2	21.1	**137.3**	5.4	1.1	0.4	3.2	3.3	1	0	0.2	0.1	1.7
3	1.6	0.3	0	0	1.9	6.9	1.2	0	1.5	6.3	**28.6**	1.6	0	0.4	1.2	2.4
4	0	0	0	0	0	0.5	0.5	0.1	0	0.2	5.8	1.1	0	0.3	5.7	**41.9**
Huajialing station																
1	**262**	23.4	0.8	0.8	12.9	18.4	2	0.6	0	0.2	0.7	0.1	0	0	0.4	0.6
2	24.2	10.4	0.2	0.2	19.6	**134.5**	9	2.8	0.2	3	8.2	0.7	0	0.2	1.6	2.3
3	0.8	0.2	0	0	2.5	11.2	0.9	0.5	0.8	8.4	**28.5**	4.2	0	0.2	1.6	4.2
4	0	0	0	0	0	1.9	0.1	0.1	0	3.4	4.5	1.1	0	1.6	5.4	**23.9**

Table 12. Comparison between observed and predicted drought classes for Yumen station

Time	*Month of drought*		*drought grade of T month*		*drought grade of T+1month*	
	t-2	*t-1*	*observed*	*predicted*	*observed*	*predicted*
January	2	2	2	2/3	2	2/3
February	2	2	2	2/3	1	2/3
March	2	2	1	2/3	1	2/3
April	2	1	1	1	1	1
May	1	1	1	1	1	1/2
June	1	1	1	1	1	1/2
July	1	1	1	1	1	1/2
August	1	1	1	1	1	1/2
September	1	1	1	1	1	1/2
October	1	1	1	1	1	1/2
November	1	1	1	1	1	1/2
December	1	1	1	1	1	1/2

Table 13. Comparison between observed and predicted drought classes for Huajialing station

Time	*Month of drought*		*drought grade of T month*		*drought grade of T+1month*	
	t-2	*t-1*	*observed*	*predicted*	*observed*	*predicted*
January	3	3	3	2/3	3	2/3
February	3	3	3	2/3	3	2/3
March	3	3	3	2/3	3	2/3
April	3	3	3	2/3	2	2/3
May	3	3	2	2/3	2	2/3
June	3	2	2	2	2	2/3
July	2	2	2	2/3	3	2/3
August	2	2	3	2/3	2	2/3
September	2	3	2	3	2	2/3
October	3	2	2	2/3	2	2/3
November	2	2	2	2/3	2	2
December	2	2	2	2/3	2	2

At this time, if the predication follows the transition dominance of the drought levels, then the predicted drought level is likely to be different from the actual value. As analyzed at the Yumen station: the result predicted in a period of one month from April to December is exactly equal to the observed values. “1/2” in the predicted result in a period of 2 months indicates that the probabilities that there are no drought (level 1) and minor drought (level 2) exactly equal, also consistent with the observation value “1”; in the results predicted in one month from January to March, the predicted value (“2/3”) in March is inconsistent with the observed value (“1”). In the results predicted in two months, the predicted values (“2/3”) in February and March are inconsistent with the observed values (“1”).

As analyzed at the Huajialing Station: the results predicted in June is exactly equal to the observed value. “2/3” in the observed results of the two-month predication period indicates the probabilities of the minor drought (Level 2) and middle drought (level 3) occur are exactly the same, also consistent with the observed value “2”; the result (“3”) predicted in September is inconsistent with the observed value (“2”). In the results predicted in a period of two months, the result (“2”) from November to December completely coincides with the observed value (“2”). Conceptual data shows that the model has a good prediction accuracy and can well predict the drought level.

4. Discussions

The Yumen station and Huajialing station are selected for verification, which the predicted length is 12. The predicted results which in the period of 1 month and in the period of 2 month are shown in Table 14. According to the table, it can be found that there was only one inconsistency phenomena of the result predicted in a period of 1 month, and two inconsistency phenomena of the result predicted in a period of 2 month of Yumen station for 12 months. Similar results were shown for Huajialing station. All this fully demonstrates the validity of the prediction.

Table 14. Summary of predicted drought level for Huajialing station and Yumen station

Station	*Predicted result in a period of 1 month*		*Predicted result in a period of 2 month*	
	Consistent number	*Inconsistent number*	*Consistent number*	*Inconsistent number*
Yumen	1	11	10	2
Huajialing	1	11	10	2

The analysis shows that the log-linear model can predict meteorological drought level consistent with the actual value in a period of one month. Although there is inconsistency possibly appeared in some data, the predicted level is in proximity to the actually measured value. It is proved that this model has a high prediction accuracy and can be used for short-term drought warning in relevant watersheds.

5. Conclusions

Monthly precipitation data from 19 weather stations in the Yellow River Basin from 1957 to 2009 are used to obtain the standard precipitation index (SPI) sequence on the 12-month time scale. From the statistics of various drought level transitions, a log-linear model is established based on the 3D contingency table to predict the short-term meteorological drought level in the Yellow River Basin.

The ratio of drought level transition frequencies predicted by the model derives appropriate dominance value and its confidence interval. With 96 dominance and confidence intervals, the drought levels in the next 1 ~ 2 month (s) are predicted. In this way, the prediction precision has been greatly improved. The algorithm can improve the availability of drought monitoring information, and provide the clues for water resources managers to take measures against drought calamity.

Acknowledgements

This research was supported by National Natural Science Foundation of China (51709221), National key research and development plan projects (2017YFC0405901-2), Study on the response mechanism of Drought and Flood in Guizhou and countermeasures for disaster reduction (KT201705) and Key Scientific Research Projects Plan of Henan Higher Education Institutions (19A120008). Sincere gratitude is extended to the editor and anonymous reviewers for their professional comments and corrections, which greatly improved the presentation of the paper.

References

Agresi A., (1990), *Categorical Data Analysis*, John Wiley, New York.

Bhalme H.N., Mooley D.A., (1980), Large-scale drought/floods and monsoon circulation, *Monthly Weather Review*, **108**, 1197-1211.

Chen L.L., Liu P.X., Yao Y.L., Zhu Z.J., Zhao M.J., (2013), Variation characteristics of annual and spring standardized precipitation index and Z index in different climate regions of Gansu Province, Northwest China in 1960-2010, *Chinese Journal of Ecology*, **32**, 704-711.

Edwards D.C., McKee T.B., (1997), Characteristics of 20th century drought in the united states at multiple time scales, *Department of Atmospheric Science, Colorado State University*, **634**, 97-102.

Harman J.R., Elton W.M., (2015), The LaPorteprecipitation anomaly, *Annals of the Association of American Geographers*, **61**, 468-480.

Heim R., (2002), A review of twentieth-century drought indices used in the United States, *Bulletin of the American Meteorological Society*, **83**, 1149-1165.

Liang Z.T., Hewitt R.R., Du Y., (2019), Research on design method for the blue-green ecological network system to deal with urban flooding: A case study of Charleston Peninsula, *International Journal of Design and Nature and Ecodynamics*, **14**, 275-286.

Liu Q., Zhang G.L., Ali S., Wang X.P., Wang G.D., Pan Z.K., Zhang J.H., (2019), SPI-based drought simulation and prediction using ARMA-GARCH model, *Applied Mathematics and Computation*, **355**, 96-107.

Mckee T.B., Doesken N.J., Kleist J., (1993), *The Relationship of Drought Frequency and Duration to Time Scales*, 8th Conf. on Applied Climatology, Anaheim, 179-184.

McKee T.B., Doeskin N.J., Kleist J., (1995), *Drought Monitoring with Multiple Time Scales*, Proc. of 9th Conf. on Applied Climatology, Massachusetts: American Meteorological Society, 233-236.

Moreira E.E., Coelho C.A., Paulo A.A., Pereira L.S., Mexia J.T., (2008), SPI-based drought category prediction using loglinear models, *Journal of Hydrology*, **354**, 116-130.

Morid S., Smakhtin V., Bagherzadeh K., (2010), Drought forecasting using artificial neural networks and time series of drought indices, *International Journal of Climatology*, **27**, 2103-2111.

Nicholas R.E., Battisti D.S., (2008), Drought recurrence and seasonal rainfall prediction in the Río Yaqui Basin, Mexico, *Journal of Applied Meteorology and Climatology*, **47,** 991-1005.

Palmer W.C., (1965), *Meteorological drought*, U.S. Department of Commerce, Weather Bureau, Washington, No.45, 1-58.

Patel N.R., Chopra P., Dadhwal V.K., (2007), Analyzing spatial patterns of meteorological drought using standardized precipitation index, *Meteorological Applications*, **14,** 329-336.

Paulo A.A., Ferreira E., Coelho C., Pereira L.S., (2005), Drought class transition analysis through Markov and Loglinear models, an approach to early warning, *Agricultural Water Management*, **77**, 59-81.

Shah R.D., Mishra V., (2016), Utility of Global Ensemble Forecast System (GEFS) reforecast for medium-range drought prediction in India, *Journal of Hydrometeorology*, **17**, 1-44.

Vicente-Serrano S.M., Begueria S., López-Moreno J.I., (2011), Comment on "Characteristics and trends in various forms of the Palmer Drought Severity Index (PDSI) during 1900-2008" by Aiguo Dai, *Journal of Geophysical Research Atmospheres*, **116**, 41-41.

Wu H., Hayes M.J., Weiss A., Hu Q., (2001), An evaluation of the standardized precipitation index, the China-Z-indexand the statistical Z-score, *International Journal of Climatology*, **21**, 745-758.

Yang P., Xia J., Zhang Y., Zhan C.S., Qiao Y.F., (2018), Comprehensive assessment of drought risk in the arid region of Northwest China based on the global palmer drought severity index gridded data, *Science of the Total Environment*, **627**, 951-962.

Yusof K.W., Hussain M., Mustafa M.R.U., (2019), Dam operation under changing climate: Analysing water availability and hydropower production from Murum Dam in Sarawak, *International Journal of Sustainable Development and Planning*, **14**, 237-244.