AMERICAN METEOROLOGICAL SOCIETY

Journal of Applied Meteorology and Climatology

EARLY ONLINE RELEASE

The DOI for this manuscript is doi: 10.1175/JAMC-D-15-0295.1

A hybrid index for characterizing drought based on a nonparametric kernel estimator

Shengzhi Huang[1], Qiang Huang[1], Guoyong Leng[2,*], and Jianxia Chang[1]

1 State Key Laboratory Base of Eco-Hydraulic Engineering in Arid Area, Xi'an University of Technology, Xi'an 710048, China

2 Joint Global Change Research Institute, Pacific Northwest National Laboratory, College Park, MD 20740, USA

* Corresponding author address: Guoyong Leng, Joint Global Change Research Institute, Pacific Northwest National Laboratory. E-mail: guoyong.leng@pnnl.gov

Abstract This study develops a nonparametric multivariate drought index, namely, the Nonparametric Multivariate Standardized Drought Index (NMSDI), by considering the variations of both precipitation and streamflow. Building upon previous efforts in constructing Nonparametric Multivariate Drought Index, we use the nonparametric kernel estimator to derive the joint distribution of precipitation and streamflow, thus providing additional insights in drought index development. The proposed NMSDI are applied in the Wei River Basin (WRB), based on which the drought evolution characteristics are investigated. Results indicate: (1) generally, NMSDI captures the drought onset similar to Standardized Precipitation Index (SPI) and drought termination and persistence similar to Standardized Streamflow Index (SSFI). The drought events identified by NMSDI match well with historical drought records in the WRB. The performances are also consistent with that by an existing Multivariate Standardized Drought Index (MSDI) at various timescales, confirming the validity of the newly constructed NMSDI in drought detections (2) An increasing risk of drought has been detected for the past decades, and will be persistent to a certain extent in future in most areas of the WRB; (3) the identified change points of annual NMSDI are mainly concentrated in the early 1970s and middle 1990s, coincident with extensive water use and soil reservation practices. This study highlights the nonparametric multivariable drought index, which can be used for drought detections and predictions efficiently and comprehensively.

Keywords Hybrid drought index; Nonparametric kernel estimator; Wavelet analysis; Wei River Basin

1. Introduction

Drought, caused by water deficit over extended periods, has huge adverse impacts on agriculture, society, economy and ecosystems (Wilhite et al., 2000; Mishra and Singh, 2010; Dai, 2012; Sheffield et al. 2012; Van Loon and Van Lanen, 2012). According to the assessment of Ministry of Water Resources of China (MWRC, 2011), annual grain losses induced by drought in China is nearly 39.2 billion kg, and the average economic loss is about 1.47% of the country's gross domestic product. With global warming, China is projected to experience more frequent, pronounced and severe droughts with long-term droughts increasing more than short term droughts (Leng et al. 2015a, b). Drought monitoring and prediction are therefore of great importance for taking prompt and effective measures to reduce the influences of droughts (WMO, 2006; AghaKouchak and Nakhjiri 2012; Sheffield et al., 2014; Hao et al., 2014; Huang et al. 2014a, b).

Unfortunately, albeit there is a strong desire to develop an algorithm which can be used for accurately describing and assessing droughts, it is not easy to be achieved either by means of statistical or physical analyses. The primary obstacle is lack of a universally accepted definition of droughts (Wells et al. 2004; Vicente-Serrano et al. 2010; Mishra and Singh 2010, 2011; Winslow et al. 2011; Sherwood and Fu 2014; Spinoni et al. 2014; Van et al. 2013; Hao and Singh 2015). Drought indicator is extremely critical and fundamental for drought monitoring and prediction because of its ability to simplify complex interplay amongst many hydro-climatic parameters. Drought Indices allow the quantitative evaluation of droughts in terms of frequency, duration, severity and spatial extent (Wilhite et al. 2000).

Various drought indices have been developed and adopted for drought characterization and monitoring. The Palmer drought severity index (PDSI) is used and applied in drought assessment and

monitoring studies (Palmer 1968; Dai 2011). The standardized precipitation index (SPI) proposed by McKee et al. (1993) is used for meteorological drought characterization and has been demonstrated as an effective tool to identify the early onset of drought (Shukla et al. 2011; Hao and Aghakouchak 2014). The standardization concept of the SPI has been extended to other variables to develop drought indexes such as the standardized runoff index (SRI; Shukla and Wood, 2008), Standardized Streamflow Index (SSFI, Telesca et al. 2012) and the standardized soil moisture index (SSI; Hao and AghaKouchak 2013) for hydrological and agricultural drought characterization respectively.

However, the performance of drought indicators based on different variables differs in identifying the drought onset and termination. A meteorological drought (shortage in precipitation) may develop and end quickly, whilst the onset and persistence of a hydrologic drought (deficit in streamflow) respond to a meteorological drought with some delays (Pandey and Ramasastri 2001; Heim 2002). A drought index based on precipitation shortage is appropriate for capturing the drought onset, while a drought index based on streamflow shortage is efficient for detecting drought persistence. The discrepancies in the physical bases of drought-related variables make it difficult to construct an efficient and comprehensive drought index based on one single variable (e.g., precipitation or streamflow). Wilhite (2005) pointed out that the utilization of a single index is one of the major obstacles for accurate drought characterization and prediction. It is well recognized that no single index can characterize all aspects of meteorological, hydrological, and agricultural droughts and that a multi-variable approach should be adopted to cope with this issue (Quiring et al. 2007; Hao and AghaKouchak 2013).

Recently, several efforts have been made to develop a variety of integrated drought indicators which combine multiple drought related variables (e.g. precipitation, soil moisture and streamflow),

88 such as the Standardized Precipitation Evapotranspiration Index (SPEI) (Vicente-Serrano et al.,
89 2010), the Aggregate Drought Index (ADI) (Keyantash and Dracup, 2004), the Combined Drought
90 Indicator (CDI) (Sepulcre-Canto et al., 2012), the Joint Deficit Index (JDI) (Kao and Govindaraju,
91 2010), and the Multivariate Standardized Drought Index (MSDI) (Hao and AghaKouchak, 2014).
92 The MSDI based on copula functions has been demonstrated as an effective tool in drought
93 prediction and monitoring practice (AghaKouchak 2015; Hao and Singh 2015). It is featured with
94 denoting the droughts at a specific time scale, which is similar to SPI but is extended to bivariate
95 version combing the information of precipitation and soil moisture. However, copulas, which are in
96 essence a parametric approach, strongly rely on assumptions that data samples follow a particular
97 probability density function (PDF) (Huang et al. 2014b). Since the complex interrelationships
98 between atmosphere, vegetation, surface water, soil and groundwater have large impacts on drought
99 processes, any of the given distributions may fail to accurately reflect drought conditions (Sadri
100 2012). In fact, no worldwide parametric distribution exists for hydrologic and meteorological
101 variables (Silverman 1986; Moon and Lall 1994; Smakhtin 2001; Farahmand and AghaKouchak
102 2015). Application of parametric distribution tends to cause a big deviation of tails such as the low or
103 high quantiles (Sharma 2000).

104 To our knowledge, only Hao and AghaKouchak (2014) introduced a nonparametric multivariate
105 drought index which is based on an empirical joint probability in the bivariate case using the
106 Gringorten plotting position formula. Based on the weighted moving means of sample data from a
107 small neighborhood of the estimated point, nonparametric function estimation has the advantage of
108 reproducing the attributes denoted by the sample (Lall 1995; Sharma 2000). In this study, we build
109 upon the efforts of Hao and AghaKouchak (2014) by proposing an additional nonparametric

multivariate drought index using Kernel Estimator. The majority of nonparametric density estimation methods can be expressed through a kernel density estimator entailing a weighted moving mean of an empirical frequency distribution of the sample data (Scott 1992; Sharma 2000). Our study is complement to existing research and can provide additional insights for drought index developments. The newly proposed drought index was applied in the Wei River Basin (WRB), based on which the drought evolution characteristics including its trends, persistence and period are then investigated.

2. Study area and data

The WRB lies between 103.5°E~110.5°E and 33.5°N~37.5°N (Fig. 1), with a total land area up to 1.35×10^5 km^2. The altitude increases from the lowest Guanzhong Plain in south of the basin to the highest northwest mountainous areas. The regional mean annual precipitation of the basin is approximately 559 mm with a decreasing trend for the past decades (Zhang et al. 2008).

In this study, observed precipitation records from 21 meteorological stations in the WRB are obtained from the National Climate Center (NCC) of the China Meteorological Administration (CMA) (Fig. 1). The records covering the period from 1960 to 2010 at the daily time step have been well used in characterizing the spatial-temporal hydro-climatic changes in the WRB (e.g. Chang et al. 2014). Daily observed streamflow series for Beidao (BD), Linjiacun (LC), Weijiapu (WP), Zhangjiashan (ZJS), Xianyang (XY), Lintong (LT), Huaxian (HX), and Zhuangtou (ZT) hydrological stations in the main stream and major tributaries of the WRB are obtained from the Yellow River Conservancy Commission. Specifically, BD, LC, WP, XY, LT, and HX stations are located in the main stream, while ZJS and ZT stations lie in the first grade tributaries i.e. the Jinghe River and the Beiluohe River, respectively (Fig. 1). Figure 1 also shows the boundary of eight sub-areas for our detailed investigation in this study.

3. Methodology

3.1. Nonparametric Multivariate Standardized Drought Index (NMSDI)

3.1.1 Univariate Kernel Estimator

Nonparametric function estimates, which are based on the weighted moving mean values, have the advantage that they can effectively capture the attributes of the sample (Lall 1995; Kim et al. 2003) and have been well used in hydrology for frequency analysis (Lall 1995). A kernel density estimator is the key in most nonparametric density estimate approaches in calculating the weighted moving mean of the empirical frequency distribution of the given sample (Scott 1992; Kim et al. 2003). By inter-comparing with the parametric Gumbel, Lognormal, and Pearson III distributions, Kim et al. (2003) demonstrated that the nonparametric kernel estimator has a better performance in fitting the probability density function (PDF) of input time series.

Given n observations $x_1,...,x_n$, the univariate kernel probability density estimator can be derived as follows:

$$\hat{f}_X(x) = \frac{1}{nh}\sum_{i=1}^{n} K\left(\frac{x - x_i}{h}\right) \quad (1)$$

where K is a kernel function; h is the bandwidth controlling the variance of the kernel function. Previous studies applying the kernel functions in water resources and hydrology research fields are summarized in Table 1 (Silverman 1986; Lall et al. 1996). Here, we use the Gaussian kernel function following Kim et al. (2003) in this study. With the decrease of h, the assessed function tends to exhibit the variance resembling the individual observations and vice versa. In a given data space, a larger h would lead to a smoother function. With K as the Gaussian density, the estimation of h is expressed as follows:

$$h_d = [\frac{4}{n(p+2)}]^{1/(p+4)} \sigma_d \quad (2)$$

where h_d is the optimal bandwidth; p is the number of dimensions with p = 1 and 2 indicating the univariate and bivariate kernel estimator, respectively); σ_d is the standard deviation of the distribution in dimension d.

3.1.2 Bivariate Kernel Estimator

Bivariate kernel density estimators are constructed in the similar approaches as the univariate estimators but based on a set of bandwidths, (h_x, h_y). The joint PDF for two variables of interest (precipitation and streamflow in this study) is computed as follows:

$$\hat{f}_{XY}(x,y) = \frac{1}{nh_x h_y} \sum_{i=1}^{n} \left\{ K(\frac{x-x_i}{h_x}) K(\frac{y-y_i}{h_y}) \right\} \quad (3)$$

where n is the total number of observations (x_i, y_i); . K is the Gaussian kernel function; h_x and h_y are the bandwidths for variable x and y, respectively. The determination of h was referred to Kim et al. (2003). h depends on the sample data and only one optimal value is derived.

3.1.3 Calculation of NMSDI

Through constructing the joint PDF of precipitation and streamflow using the bivariate kernel density estimators, their joint Cumulative Distribution Function (CDF) is expressed as:

$$P(X \le x, Y \le \mathrm{y}) = p \quad (4)$$

where p is the joint CDF of precipitation (x) and streamflow (y). The NMSDI is then calculated as:

$$NMSDI = \phi^{-1}(\mathrm{p}) \quad (5)$$

where ϕ^{-1} is the inverse of the standard normal distribution function. The derivation of NMSDI is similar to the Standardized Precipitation Index (SPI) (McKee et al. 1993; Huang et al. 2014b) but based on multivariate distribution function. Classification of the developed NMSDI index was

175 exhibited in Table 2.

176 The trends of NMSDI are detected using the modified Mann-Kendall (MMK) trend test method
177 (Mann, 1945, Kendall, 1955 and Hamed and Rao, 1998). Compared to original MK trend test
178 method, the MMK trend test method can overcome issues arising from the persistence of the
179 hydro-meteorological series (Hamed and Rao, 1998).The wavelet analysis (Torrence and Compo
180 1998) is used to reveal the periodic characteristics of droughts based on the annual NMSDI. The
181 technique can give a quantitative measure of changes in variance on interdecadal timescales. The
182 MMK and wavelet analysis techniques have been well in our previous studies (Huang et al. 2014a).

183 *3.2. The Rescaled Range (R/S) analysis*

184 The R/S analysis (Hurst 1956) has been widely used in the fields of morphology, cell reproduction,
185 hydrology and earthquake activity in forecasting their long memory processes (Hosking 1984; Wang
186 et al. 2013). The basic thought of the R/S analysis is to change the timescale of a given series based
187 on which its statistical characteristics is examined (Oliver and Ballester 1996).

188 For a time series with *n* observations $X = x_1, x_2, ..., x_n$, the average series is calculated as follows:

189

190 $$y(\tau) = \frac{1}{\tau}\sum_{t=1}^{\tau} x(\mathrm{t}) \quad \tau = 1, 2, ... \tag{6}$$

191 The deviation is calculated as follows:

192 $$F(\mathrm{t}, \tau) = \sum_{t=1}^{\tau} x(\mathrm{t}) \quad 1 \le t \le \tau \tag{7}$$

193 The range is calculated as following:

194 $$R(\tau) = \max_{1 \le t \le \tau} F(\mathrm{t}, \tau) - \min_{1 \le t \le \tau} F(\mathrm{t}, \tau) \tag{8}$$

195

196 And its standard deviation is calculated as following:

197 $$S(\tau) = \left| \frac{1}{\tau} \sum_{t=1}^{\tau} (\mathrm{x(t)} - \mathrm{y}(\tau)) \right|^{2}^{1/2} \quad (9)$$

198 The rescaled range based on its standard deviation and range is calculated as:

199 $$R(\tau) / \mathrm{S}(\tau) = (\mathrm{C}\tau)^{H} \quad (10)$$

200 where *C* is a constant, and *H* is the calculated Hurst index.

201 A Hurst index value greater than 0.5 indicates the characteristics of future change are expected to
202 be similar with those in the past, and a larger *H* value in the range of 0.5~1.0 indicates higher
203 probabilities of this tendency. A Hurst index value equal to 0.5 indicates that the series is totally
204 independent. Hurst index value less than 0.5 denotes that the change trend in the future would be
205 opposite to the past, and a smaller H in the range of 0~0.5 indicates higher probabilities of this
206 tendency (Oliver and Ballester 1996).

207 *3.3. The heuristic segmentation method*

208 Existing statistical test approaches such as the sliding F test and sliding T test are frequently
209 adopted for detecting the change points. However, the assumption of stationarity has to be made,
210 which is not true in hydrologic time series with nonlinear features due to high variability of
211 hydrologic process (Feng et al. 2005). The heuristic segmentation method proposed by Pedro et al.
212 (2001) is based on the sliding T test but modified to partition a non-stationary series into several
213 stationary series, thus overcoming the drawbacks of the above methods.

214 In order to divide a non-stationary time series into several stationary segments, a sliding pointer is
215 designed which moves step-by-step along the time series from its left side to the right side (Pedro et
216 al. 2001). The average values of the subsets to the left of the pointer (μ_1) and to the right (μ_2) are then
217 calculated. The difference between the averages of the two series μ_1 and μ_2 is estimated using the

Student's *t*-test statistic as following:

$$t = \left| \frac{\mu_1 - \mu_2}{S_D} \right| \tag{11}$$

where

$$S_D = \left(\frac{(N_1\text{-}1)s_1^2 + (N_2\text{-}1)s_2^2}{N_1 + N_2 - 2} \right)^{1/2} \left(\frac{1}{N_1} + \frac{1}{N_2} \right)^{1/2} \tag{12}$$

is the pooled variance, N_1 and N_2 are the number of points of the two series, s_1 and s_2 are the standard deviations of the two series. A larger t indicates more statistically different mean values between the two time series. Hence, the largest t value is taken as the cut point, based on which the statistical significance $P(t_{\max})$ is calculated. Note that $P(t_{\max})$ is not the standard Student's t-test as the series are not independent and cannot be derived in a closed analytical form; therefore, $P(t_{\max})$ is approximately calculated as follows:

$$P(t_{\max}) \approx \left\{ 1 - I_{\left[v/(v+t_{\max}^2) \right]}(\delta v, \delta) \right\}^{\eta} \tag{13}$$

where $\eta = 4.19 \ln N - 11.54$ and $\delta = 0.4$ are obtained from the Monte Carlo simulations, N is the length of the time series, v=N-2, and $I_x(a, b)$ is the incomplete beta function (Pedro et al. 2001). No split is made to the time series if the difference of the mean values is not statistically significant (e.g., $P(t_{\max})$ is smaller than a threshold of 0.95). Conversely, the time series is cut into two segments and the iteration of the above procedures continues on each new segment until the significant value is smaller than the threshold or the length of the new segments is smaller than the predefined minimum segment length (ℓ_0).

4. Results and discussion

4.1. The performance of the NMSDI

Fig. 2A shows the time series of 1-month NMSDI, the multivariate standardized drought index

(MSDI) developed by Hao and AghaKouchak (2013), SPI and SSFI time series covering 1960~2011 for the whole Wei River basin. It is found that the change pattern of 1-month NMSDI is generally consistent with that of 1-month SPI or SSFI as well as MSDI. The correlation coefficients of MSDI with SPI and SSFI are 0.80 and 0.79, respectively. The performance of NMSDI is comparable to MSDI and the correlation coefficients with SPI and SSFI are up to 0.83 and 0.81, respectively, indicating the reliability of the newly constructed NMSDI. For better visual comparisons, the 1-month NMSDI, MSDI, SPI, and SSFI values covering 2002~2011 are shown in Fig. 2B. Some discrepancies are found among the four drought indices. Generally, as the persistent deficit of precipitation is commonly the beginning of drought, SPI is sensitive to capturing the drought onset, while SSFI is effective to identify drought persistence mainly due to the complex runoff process (Hao and AghaKouchak 2013). As shown by the black rectangles in Figure 2B, the timing of NMSDI and SPI values smaller than -1 (which is identified as the drought onset) is earlier than that of SSFI, indicating that NMSDI capture the onset of drought as early as SPI. In addition, the timing of NMSDI and SSFI with values larger than -1 is later than that of SPI, indicating that NMSDI capture the termination of drought similar to SSFI. The performances of NMSDI are also consistent with the MSDI at various time scales. That is, NMSDI well captures the drought onset similar to SPI and drought termination and persistence similar to SSFI. Notably, NMSDI is more sensitive in capturing drought onset and termination than MSDI at the 1-month timescale, demonstrating the added value of the non-parametric approach with Kernel estimator in drought characteristics. By examining the temporal variability of NMSDI, MSDI, SPI, and SSFI at the sub-basin level, similar findings are obtained (figures not shown for brevity). Since NMSDI is coupled with the information of both precipitation and streamflow, it is sensitive and effective in capturing the onset, persistence and

261 termination of droughts, and similar features are also exhibited at the 3-month and 6-month scales
262 (Fig. 3). Statistics show that the drought frequency identified by SPI is the highest, whilst that
263 identified by SSFI is the lowest (Table 3). This is because meteorological droughts characterized by
264 SPI develop quickly and end abruptly, whereas hydrological droughts characterized by SSFI respond
265 to meteorological droughts with some lag time and show a strong persistence (Heim 2002). Since
266 MSDI and NMSDI are based on the information of precipitation and streamflow deficits, the
267 detected drought frequencies are between those of SPI and SSFI. The drought duration and intensity
268 identified by MSDI and NMSDI are found to be larger than those by SPI and SSFI due to the fact
269 that MSDI and NMSDI capture the drought onset similar to SPI and drought termination and
270 persistence similar to SSFI.

271 In addition, the historical severe droughts records in the WRB are obtained to examine the
272 performance of NMSDI. Our results show that the timing of the low NMSDI values match well with
273 the drought occurrence indicated by the available records. Especially, two typical drought events
274 which occurred in 2003 and 2008 (YRCC, 2004; 2009) are well identified by the NMSDI (Fig. 2B).
275 Given the comprehensiveness and effectiveness of NMSDI for drought characterization, it was
276 adopted for our further investigations on the historical drought evolution characteristics of the WRB.

277 *4.2. The spatial pattern of NMSDI in the WRB*

278 Fig. 4 shows the averages of annual NMSDI in each sub-area. Negative values indicate dry
279 conditions and vice versa for the positive values. The detailed classifications are shown in Table 2.
280 Overall, all of the sub-areas have experienced a certain degree of drought, which exhibited a distinct
281 spatial pattern at the sub-basin level. Specifically, the historical drought risks experienced in the
282 middle basin including ZJSP, LHP, XLP, and BLP are lower than other areas, as indicated by the

283 smaller magnitude of negative NMSDI values. The highest risks of droughts occurred in WRB are
284 located in the sub-basin of LWP and ZTP as shown by the largest magnitude of negative NMSDI
285 values.

286 *4.3. The temporal pattern of NMSDI*

287 The modified Mann-Kendall (MMK) trend test method was utilized to calculate the trend of
288 monthly and annual NMSDI of each sub-area in the WRB (Table 4). An increasing risk of droughts
289 is found by the negative trends of monthly NMSDI in the WRB, especially in March, April, May,
290 July, and September. For BDP, the monthly NMSDI has decreased significantly for all months at the
291 95% confident level and has experienced the most significant decreasing trend compared with other
292 sub-areas. Non-significant changes trends are found in January and February in the eastern part of
293 WRB (including XLP, LHP, and ZTP). Annually, a statistically significant downward tendency is
294 detected for all sub-areas at the 99% confidence level. In general, our results show that the
295 drought-prone region has experienced an increasing risk of droughts, which posed great challenges to
296 the sustainable development of local agriculture, economy and society.

297 *4.4. The persistence of drought in the WRB*

298 How will the historical droughts persist in the future? Fig. 5 shows the Hurst index values of the
299 annual NMSDI of each sub-area in the WRB. It can be found that the Hurst index values are larger
300 than 0.5 in all sub-basins except for ZJSP partition with a Hurst index value slightly smaller than 0.5.
301 This indicates that the decreasing trend of annual NMSDI may be persistent to a certain extent in
302 most parts of the WRB. In other words, the historical droughts in the WRB will probably continue in
303 the future. In the contrary, the decreasing trend of annual NMSDI in ZJSP will be opposite in the
304 future, implying the drought condition in this region may be mitigated to a certain degree. Our results

have great implications for making effective adaptation strategies to reduce the possible damages caused by droughts at the sub-basin level. However, it should be noted that the drought persistence as analyzed here is in nature a projection of possible conditions in a probabilistic way. Physically-based approaches such as driving a hydrological model by global climate model (GCM) or regional climate model (RCM) outputs should be adopted in the future.

4.5. The period of annual NMSDI in the WRB

In addition to the change trends, we examine the period of annual NMSDI, which can help better understand the historical droughts and facilitate drought predictions. Here, the wavelet analysis on the basis of the Morlet wavelet (Huang et al. 2014a) is utilized and the time-frequency distribution of annual NMSDI series is exhibited in Fig. 6. The red color in Fig. 6 represents positive phase while the blue color denotes negative phase, green and yellow colors are between positive and negative phases. The alternation between positive and negative phases is taken as the period of historical droughts. Results show that a primary period up to 20 years and a secondary period of about 7 years of annual NMSDI series exists, indicating the inter-decadal oscillations of droughts in the WRB. Dou and Yan (2013) studied the linkage of drought/flood in the Guanzhong Plain (located in WRB) with sunspot activities, and found a negative linkage in 1960s and a positive relationship after 1970s. Since the major objective is to develop a reliable integrated drought index and characterize the drought evolution characteristics, a further physical explanation for these periods should be explored in the future.

4.6. The identification of change points of annual NMSDI

The heuristic segmentation method was applied to identify the change points of annual NMSDI series in each sub-area in the WRB. The threshold P_0 was selected as 0.95 and ℓ_0 was chosen as 25

following Pedro et al. 2001. The results of detected change points in BDP and LHP are shown in Fig. 7. One change point was detected in BDP during its first iteration (Fig. 7A). However, for the second iteration, no change points were identified because the probability of the largest T was smaller than the threshold. Fig. 7B indicates two change points in LHP during the two iterations as the probabilities of their largest T was larger than the threshold. The identified change points in other partitions were also carried out and are summarized in Table 5.

Generally, the annual NMSDI series in the WRB is not stationary, with change points detected mainly in early 1970s and middle 1990s (Table 5). It is worth mentioning that a large number of reservoirs and irrigation canals were constructed in the early 1970s, e.g. the Yangmaowan and Fengcun reservoirs in 1970, the Baojixia canal in 1971, and the Dayu and Shimen reservoirs in 1972. These water projects were used for withdrawal of large amounts of water from the Wei River and have greatly changed the intra-annual characteristic of streamflow. Additionally, the soil conservation measures in the WRB were performed in the 1950s, and have expanded on a large scale since early 1970s. The soil conservation practice further alters local micro-topography, rainfall interception, and infiltration rate, thereby reducing local streamflow and resulting in hydrological drought (Chang et al. 2014). Since the precipitation has a non-significant decreasing trend in WRB (Huang et al. 2014c), the construction of water projects and the soil conservation practices may be the major driving forces resulting in the identified change points. The developed NMSDI in this study is based on the coupled information of precipitation and streamflow. Hence, human activities in the WRB have lowered the drought index to a certain extent mainly through affecting streamflow, although the quantitative effects require more investigations in the future.

5. Summary and Conclusions

Drought index is the key for water resource management and drought mitigation. Drought index based on a single variable is not sufficient for characterizing droughts comprehensively and effectively. By assuming that samples follow a particular probability density function, most parametric multivariate approaches would lead to large deviations of tails (e.g. the low or high quantiles). In this study, we build upon our efforts on Hao and AghaKouchak (2014) by proposing an additional nonparametric multivariate drought index using Kernel Estimator. The newly proposed drought index was applied in the WRB in China, based on which the change trends, persistence, periods and change points are investigated. The primary conclusions are as follows:

(1) Generally, NMSDI captures the drought onset similar to SPI and drought termination and persistence similar to SSFI. Consistence is also found when comparing NMSDI with an existing Multivariate Standardized Drought Index (MSDI) at 1, 3 and 6-month timescales. Through comprehensive validations against existing drought index and collected historical drought records, our results demonstrated that NMSDI is an effective tool for drought characterizations.

(2) The NMSDI in the WRB shows a noticeable spatial gradient, with larger drought risk experienced in the LWP than other areas. The drought-prone region has experienced an increasing trend of droughts in most of sub-basin at the monthly and annual scale, which may persist to a certain extent in the future.

(3) A primary period of about 20 years and a secondary period of about 7 years of annual NMSDI series in the WRB are found.

(4) The identified change points of annual NMSDI are mainly concentrated in early 1970s and middle 1990s, which may be caused by the construction of water projects and the soil conservation practices in the WRB.

Similar to existing multi-variable drought indices, advantages and disadvantages do exist. The advantages of the NMSDI developed in this study are that the index is distribution-free and has exhibited good performance when comparing with the corresponding SPI and SSFI (the correlation coefficient is 0.83 and 0.81 respectively) and the available historical drought records. Generally, the variations of NMSDI can be used to capture the drought onset similar to SPI and the drought termination and persistence similar to SSFI. That is, it is more sensitive to capturing the onset, persistence, and termination of droughts, thereby integrating the useful information from SPI and SSFI. Compared to the existing MSDI, high consistence is found for the temporal variations of the two drought indexes at various timescales, confirming the validity of the newly constructed NMSDI in drought detections. However, NMSDI is more sensitive in capturing drought onset and termination than MSDI at the 1-month timescale, demonstrating the added value of the newly developed non-parametric drought index in capturing drought characteristics. However, it should be noted that this paper does not claim that NMSDI alone is always sufficient for capturing drought characteristics. In fact, previous studies have highlighted the limitations of various drought indicators. Having an additional source of information can provide additional insights in drought index development and improve our understanding of the drought development. Whether and where the new drought index would be particularly advantageous do require more in-depth research over larger research domain with varying hydro-climatic regimes in the future efforts. Moreover, the NMSDI developed in this study is only coupled with the deficits of precipitation and streamflow due to the limitation of the nonparametric kernel estimator. More efforts are still needed in the future toward developing a drought index which can couple meteorological, hydrological, and agricultural factors in an efficient way.

Acknowledgements

This research was supported by the National Department Public Benefit Research Foundation of Ministry of Water Resources (201501058), the National Major Fundamental Research Program (2011CB403306-2), the National Natural Fund Major Research Plan (51190093), the Natural Science Foundation of China (51179148, 51179149, 51309188, and 51379014), the Key Innovation Group of Science and Technology of Shaanxi (2012KCT-10) and the China Scholarship Council. Guoyong Leng was supported by U.S. Department of Energy (DOE) Office of Science Biological and Environmental Research (BER) through the Integrated Assessment Research program. The Pacific Northwest National Laboratory (PNNL) is operated for the U.S. DOE by Battelle Memorial Institute under contract DE-AC05-76RL01830

References

Adamowski, K., 1996: Nonparametric estimation of low-flow frequencies. *J. Hydrol. Eng.*, **122**(1), 46-49.

AghaKouchak, A.,. 2015: A multivariate approach for persistence-based drought prediction: Application to the 2010–2011 East Africa drought. *J. Hydrol.*, 526, 127-135.

AghaKouchak, A., N. Nakhjiri. 2012: A near real-time satellite-based global drought climate data record. *Environ. Res. Lett.*, 7 (4), 044037

Chang, J.X., Y. Wang, and E. Istanbulluoglu, 2014: Impact of climate change and human activities on runoff in the Weihe River Basin, China. *Quarter. Int.*, **380**, 169–179

Daufresne, M., K. Lengfellner, and U. Sommer, 2009: Global warming benefits the small in aquatic ecosystems. *P. Natl. Acad. Sci. USA.*, **106** (31), 12788-12793.

Dai, A., 2011: Characteristics and trends in various forms of the Palmer Drought Severity Index during 1900-2008. *J. Geophys. Res.*, **116**, D12115.

Dou, L., and A. Yan, 2013: Relationships between drought and flood disasters in Guanzhong plain and the activities of sunspot. *J. Arid Land Resour. Environ.*, **27** (8), 76-82.

Federal Emergency Management Agency, 1995: National Mitigation Strategy: Partnerships for Building Safer Communities. FEMA, Washington DC.

Feng, G.L., Gong, Z.Q., Dong, W.J., et al., 2005. Abrupt climate change detection based on heuristic segmentation algorithm. ACTA PHYSICA SINICA, 54(11), 5494-5499.

Farahmand, A., and A. AghaKouchak, 2015: A generalized framework for deriving nonparametric standardized drought indicators. *Adv. Water Res.*, **76**, 140-145.

Hurst, H. E., 1956: Method of using long-term storage in reservoirs. ICE Proceedings **5**(5), 519-543.

Hosking, J. R. M., 1984: Modeling persistence in hydrological time series using fractional differencing. *Water Resour. Res.*, **20**(12), 1898-1908.

Hamed, K. H., and A. R. Rao, 1998: A modified Mann–Kendall trend test for autocorrelated data. *J. Hydrol.*, **204**, 182-196.

Heim, R. R., 2002: A review of twentieth-century drought indices used in the United States. *Bull. Amer. Meteor. Soc.*, **83**, 1149-1166.

Hayes, M., and Coauthors, 2011: The Lincoln declaration on drought indices: Universal meteorological drought index recommended. *Bull. Amer. Meteor. Soc.*, **92**, 485-488.

Hao, Z., V. P. Singh, 2015: Drought characterization from a multivariate perspective: A review. *J. Hydrol.*, 527, 668–678

Hao, Z., and A. AghaKouchak, 2013: Multivariate Standardized Drought Index: A multi-index parametric approach for drought analysis. *Adv. Water Resour.*, **57**, 12-18.

Hao, Z., and A. Aghakouchak, 2014: A Nonparametric Multivariate Multi-Index Drought Monitoring Framework. *J. Hydrometeor.*, **15**, 89-101.Hamed, K. H., and A. R. Rao, 1998: A modified Mann–Kendall trend test for autocorrelated data. *J. Hydrol.*, **204**, 182–196

Heim, R.R., 2002: A review of twentieth-century drought indices used in the United States. Bull. Amer. Meteor. Soc., 83, 1149-1166.

Huang, Q., and J. Fan, 2013: Detecting runoff variation of the mainstream in Weihe River. *J. Appl. Math.*, doi:10.1155/2013/356474.

Huang, S., and Coauthors, 2014a: Copulas-based probabilistic characterization of the combination of dry and wet conditions in the Guanzhong Plain, China. *J. Hydrol.*, **519**, 3204-3213.

Huang, S., and Coauthors, 2014b: Spatio-temporal Changes and Frequency Analysis of Drought in the Wei River Basin, China. *Water Res. Manage.*, **28**(10), 3095-3110.

Huang, S., and Coauthors, 2014c: Identification of abrupt changes of the relationship between rainfall and runoff in the Wei River Basin, China. *Theor. Appl. Climatol.*, **120**(1-2), 299-310.Kao, S.C., R.S. Govindaraju. 2010: A copula-based joint deficit index for droughts. *J. Hydrol*., 380, 121–134

Kendall, M. G., 1955: Rank Correlation Methods. Griffin, London.

Keyantash, J.A., J.A. Dracup, 2004: An aggregate drought index: assessing drought severity based on fluctuations in the hydrologic cycle and surface water storage, *Water Resour. Res*., 40, W09304Kendall, M.G., 1955: Rank Correlation Methods. Griffin, London.

Kim, T. W., J. B. Valdés, and C. Yoo, 2003: Nonparametric Approach for Estimating Return Periods of Droughts in Arid Regions. *J. Hydrol. Eng*., **8**, 237-246.

Kao, S. C., and R. S. Govindaraju, 2010: A copula-based joint deficit index for droughts. *J. Hydrol.*, **380**, 121-134.

Lall, U., 1995: Recent advance in nonparametric function estimation: Hydrologic application. *Rev. Geophys*., **33**(S), 1093-1102.

Lall, U., B. Rajagopalan, D. G. Tarboton, 1996: A nonparametric wet/dry spell model for resampling daily precipitation. *Water Resour. Res*., **32**(9), 2803-2823.

Leng, G., Q. Tang, and S. Rayburg, 2015a: Climate change impacts on meteorological, agricultural and hydrological droughts in China. *Global Planet. Change*, **126**, 23-34.

Leng, G., Q. Tang, S. Huang, X. Zhang, and J. Cao, 2015b: Assessments of joint hydrological extreme risks in a warming climate in China. *Int. J. Climatol*., doi:10.1002/joc.4447.

Mann, H. B., 1945: Nonparametric tests against trend. *Econometrica*, **13**, 245-259

McKee, T. B., N. J. Doesken, J. Kleist, 1993: The relationship of drought frequency and duration to time scales. Preprints, Eighth Conf. on Applied Climatology, Anaheim, CA, Amer. Meteor. Soc., 179-184.

Moon, Y., and U. Lall, 1994: Kernel quantile function estimator for flood frequency analysis. *Water Resour. Res*., **30**(11), 3095-3103.

Mishra, A. K., and V. P. Singh, 2010: A review of drought concepts. *J. Hydrol.*, **391**(1), 202-216.

Mishra, A. K., and V. P. Singh, 2011: Drought modeling–A review. *J. Hydrol.*, **403**(1), 157-175.

Oliver, R., and J. L. Ballester, 1996: Rescaled range analysis of the asymmetry of solar activity. *Solar Physics* **169**(1), 215-224.

Palmer, W. C., 1968: Keeping track of crop moisture conditions, nationwide: The new crop moisture index. *Weather wise* **21**(4), 156-161.

Pandey, R. P., and K. S. Ramasastri, 2001: Relationship between the common climatic parameters and average drought frequency. *Hydrol. Process.,* **15**, 1019-1032.

Pedro, B. G., and Coauthors, 2001: Scale invariance in the nonstationarty of human heart rate. *Phys. Rev. Lett.,* **87**(16), 160815.

Quiring, S., and Coauthors, 2007: Drought monitoring index for Texas. Tech. Rep. to Texas Water Development Board, Austin, TX, 262 pp.

Sepulcre-Canto, G., S. Horion, A. Singleton, et al. 2012: Development of a combined drought indicator to detect agricultural drought in Europe. *Nat. Hazards Earth Syst. Sci*., 12, 3519–3531

Sheffield, J., E.F. Wood, N. Chaney, K. Guan, S. Sadri, X. Yuan, L. Olang, A. Amani, A. Ali, S. Demuth, L. Ogallo. 2014: A drought monitoring and forecasting system for sub-Sahara African water resources and food security. *Bull. Am. Meteor. Soc*., 95, 861–882

Silverman, B.W., 1986: Density estimation for statistics and data analysis, Chapman & Hall, London.

Scott, D.W., 1992: Multivariate density estimation: Theory, practice and visualization, Wiley, New York.

Sharma, A., 2000: Seasonal to interseasonal rainfall probabilistic forecasts for improved water supply management: Part 3-A nonparametric probabilistic forecast model. *J. Hydrol*., **239**, 249-258.

Sheffield, J., E. F. Wood, and M. L. Roderick, 2012: Little change in global drought over the past 60 years. *Nature* **491**(7424), 435-438.

Sherwood, S., and Q. Fu, 2014: A Drier Future? *Science,* **343**(6172), 737-739

Shukla, S., and W. Andrew, 2008: Use of a standardized runoff index for characterizing hydrologic drought. *Geophys. Res. Lett.,* **35,** L02405, 1-7.

Smakhtin, V. U., 2001: Low flow hydrology: A review. *J. Hydrol*., **240**, 147-186.

Shukla, S., A. C. Steinemann, D. P. Lettenmaier, 2011: Drought monitoring for Washington State: Indicators and applications. *J. Hydrometeor*., **12**, 66-83.

Sadri, S., and D. H. Burn, 2012: Nonparametric methods for drought severity estimation at ungauged sites. *Water Resour. Res*., **48**, W12505.

Spinoni, J., G. Naumann, H. Carrao, P. Barbosa, and J. V. Vogt, 2014: World drought frequency, duration, and severity for 1951–2010. *Int. J. Climatol*., **34**(8), 2792-2804.

Telesca L, Lovallo M, Lopez-Moreno L, et al., 2012. Investigation of scaling properties in monthly streamflow and Standardized Streamflow Index (SSI) time series in the Ebro basin (Spain). Physica A: Statistical Mechanics and its Applications 391(4), 1662-1678.

Trenberth, K.E., and Coauthors, 2014: Global warming and changes in drought. *Nature Climate Change* **4**(1), 17-22

Torrence, C., and G. P. Compo, 1998: A practical guide to wavelet analysis. *Bull. Am. Meteorol. Soc*.,**79**(1), 61-78.

Wells, N., S. Goddard, and M. J. Hayes, 2004: A self-calibrating Palmer drought severity index. *J. Clim.,* **17**(12),

2335-2351.

Wilhite, D. A., 2000 Drought as a natural hazard: Concepts and definitions, in Drought: A Global Assessment, Hazards Disasters Ser., vol. I, edited by D. A. Wilhite, pp. 3-18, Routledge, New York.

Wilhite, D. A., 2005: Drought and Water Crises: Science, Technology, and Management Issues. Taylor and Francis, 406 pp.

Winslow, M.D., and Coauthors, 2011: Science for improving the monitoring and assessment of dryland degradation. *Land Degrad. Dev.*, **22**(2), 145-149.

Wang, C.S.H., L. Bauwens, C. Hsiao, 2013: Forecasting a long memory process subject to structural breaks. *J. Econo.*, **177**(2), 171-184.

Wang, D., M. Hejazi, X. Cai, and A. J. Valocchi, 2011: Climate change impact on meteorological, hydrological, and agricultural drought in central Illinois. *Water Resour. Res.*, **47**, W09527, doi:10.1029/2010WR009845.

WMO, 2006. Drought Monitoring and Early Warning: Concepts, Progress and Future Challenges, Report No. 1006.

Van der Schrier, G., J. Barichivich, K. R. Briffa, and P. D. Jones, 2013: A scPDSI-based global dataset of dry and wet spells for 1901–2009. *J. Geophys. Res.*, 118, 4025-4048.

Vicente-Serrano, S. M., S. Beguería, and J. I. López-Moreno, 2010: A multiscalar drought index sensitive to global warming: the standardized precipitation evapotranspiration index. *J. Clim.*, **23**(7), 1696-1718.

Xu, K., and Coauthors, 2015: Copula based drought frequency analysis considering the spatio-temporal variability in Southwest China. *J. Hydrol.*, **527**, 630-640.

YRCC (Yellow River Conservancy Commission), 2004. Yellow River Water Resources Bulletin. Beijing: China WaterPower Press.

YRCC (Yellow River Conservancy Commission), 2009. Yellow River Water Resources Bulletin. Beijing: China WaterPower Press.

Zhang, H., Y. Chen, and G. Ren, 2008: The characteristics of precipitation variation of Weihe River Basin in Shaanxi Province during recent 50 years. *Agr. Res. Arid Areas* **26**(4): 236-242. (In Chinese)

List of tables and figures

Table 1

Examples of univariate kernel functions

Kernel	$K(t)$
Epanechnikov	$K(t)=0.75(1-t^2)$, $\|t\|\leq 1$ $K(t)=0$, otherwise
Triangular	$K(t)=1-\|t\|$, $\|t\|\leq 1$ $K(t)=0$, otherwise
Gaussian (normal)	$K(t)=(2\pi)^{-1/2}e^{-(t*t)/2}$
Rectangular	$K(t)=0.5$, $\|t\|\leq 1$ $K(t)=0$, otherwise

Table 2

Classification of the NMSDI index

Category	Index value
Extreme drought	NMSDI≤-2.0
Severe drought	-2.0< NMSDI≤-1.5
Moderate drought	-1.5< NMSDI≤-1.0
Near normal	-1.0< NMSDI≤1.0
Moderately wet	1.0< NMSDI≤1.5
Very wet	1.5< NMSDI≤2.0
Extremely wet	NMSDI≥2.0

Table 3

Statistics summarizing the frequency (events), mean duration (months) and intensity of droughts identified by SPI, SSI, MSDI and NMSDI for the study period in the whole WRB.

Statistics	SPI	SSI	MSDI	NMSDI
Frequency	95	47	73	69
Duration	1.35	3.11	3.26	3.43
Intensity	0.87	0.84	1.16	1.23

Note: MSDI denotes the multivariate standardized drought index (MSDI) developed by Hao and AghaKouchak (2013) and the calculation code is obtained from the link at http://amir.eng.uci.edu/software.php.

606

607 **Table 4**

608 The trends of monthly and annual NMSDI of each sub-area in the WRB.

Sub-areas	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	AT
BDP	-2.1*	-2.6**	-3.8**	-4.4**	-3.3**	-2.5*	-3.0**	-2.6**	-3.8**	-2.2*	-4.2**	-3.9**	-4.8**
BLP	0.4	-1.1	-2.9**	-5.0**	-3.3**	-2.2*	-2.9**	-2.0*	-3.1**	-2.2*	-3.9	-1.1	-4.2**
LWP	-2.4*	-1.7	-3.7**	-5.0**	-3.5**	-1.5	-2.8**	-1.0	-2.3*	-2.1*	-4.0	-1.6	-4.7**
WXP	-2.2*	-1.6	-2.7**	-4.2**	-3.6**	-1.0	-2.5*	-0.8	-2.3*	-2.2*	-3.3	-1.2	-3.8**
ZJSP	-0.5	0.1	-2.2*	-4.4**	-3.1**	-1.2	-2.0*	-1.3	-2.9**	-1.9*	-3.6	-1.2	-4.5**
XLP	0.3	0.1	-1.2	-3.5**	-2.3*	-0.2	-1.3	0.9	-2.0*	-1.6	-2.8	-0.1	-2.7**
LHP	0.6	0.1	-1.9	-4.0**	-2.7**	-0.4	-2.3*	-0.2	-2.2*	-1.6	-2.4	-0.2	-3.2**
ZTP	0.9	1.1	-1.9	-2.8**	-1.9	-0.9	-2.0*	-1.4	-2.4*	-1.3	-2.2	-0.4	-3.4**

609 Note: AT denotes annual trend; '*' and '**' represent significance at the 95% and 99% confidence level, respectively;
610 BDP, BLP, LWP, WXP, ZJSP, XLP, LHP, and ZTP stand for Beidao, Beidao-Linjiacun, Linjiacun-Weijiabu,
611 Weijiabu-Xianyang, Zhangjiashan, Xianyang-Lintong, Lintong-Huaxian, and Zhuangtou partitions, respectively.

612

613

614

615

616 **Table 5**

617 The detected change points in each sub-area in the WRB.

Subareas	Change points
BDP	1994
BLP	1994
LWP	1969, 1994
WXP	1971, 1994
ZJSP	1994
XLP	1971, 1994
LHP	1969, 1993
ZTP	1994

618

619

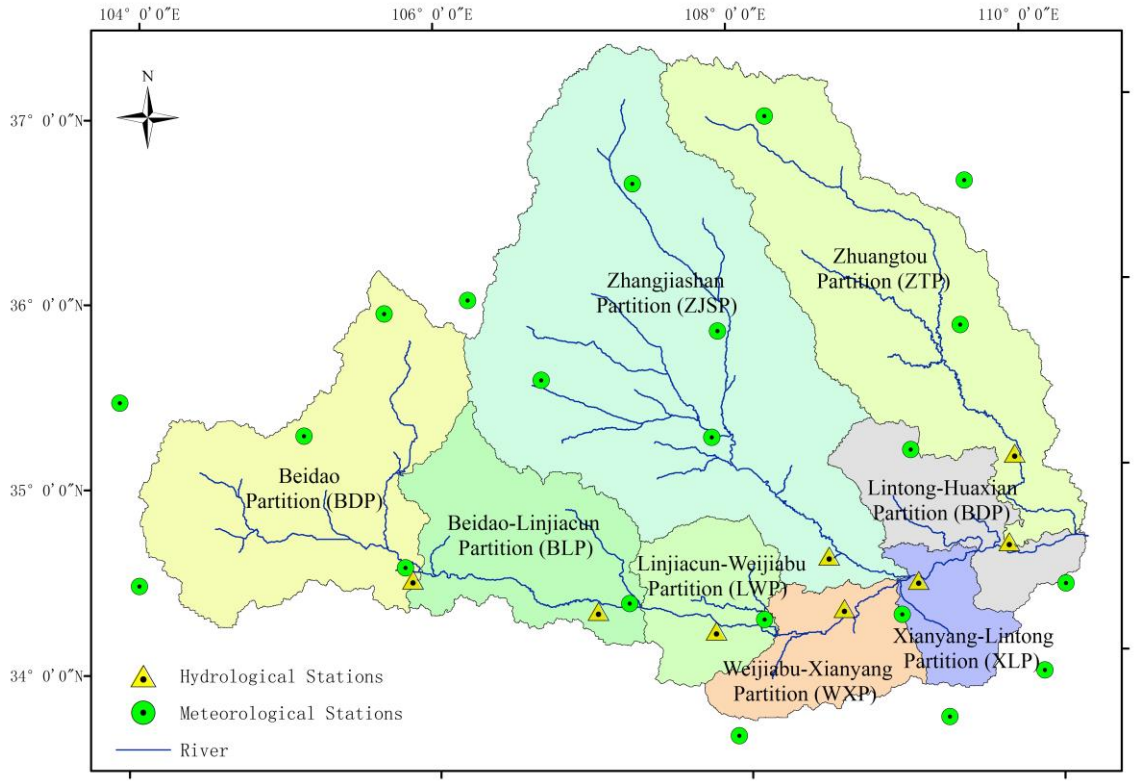

Fig.1. The study region and locations of hydro-meteorological stations.

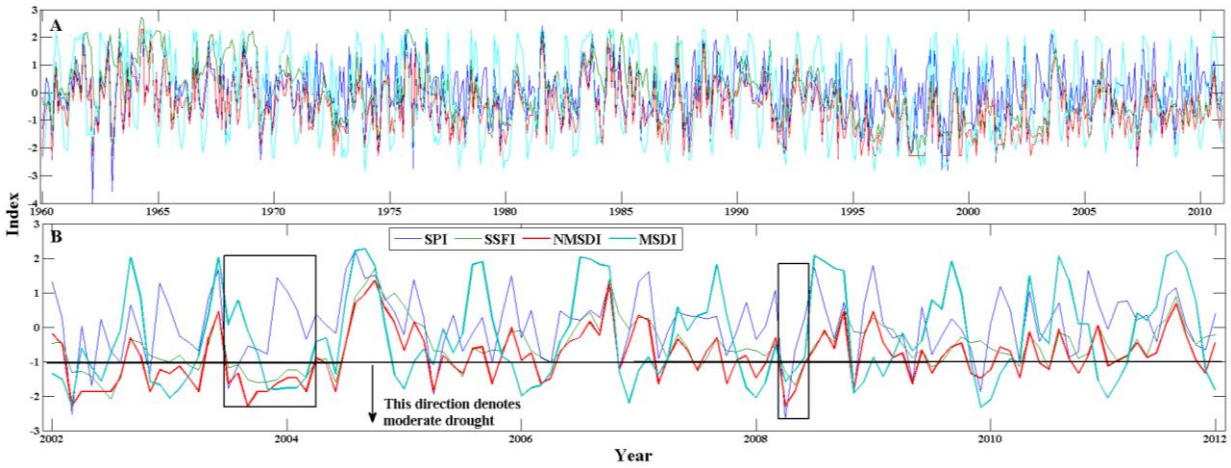

Fig. 2. The comparison between 1-month NMSDI, MSDI, SPI and SSFI series covering 1960~2011 (A) and 2002~2011 (B) in the whole WRB, the solid black line is for the threshold of moderate drought in the drought indices.

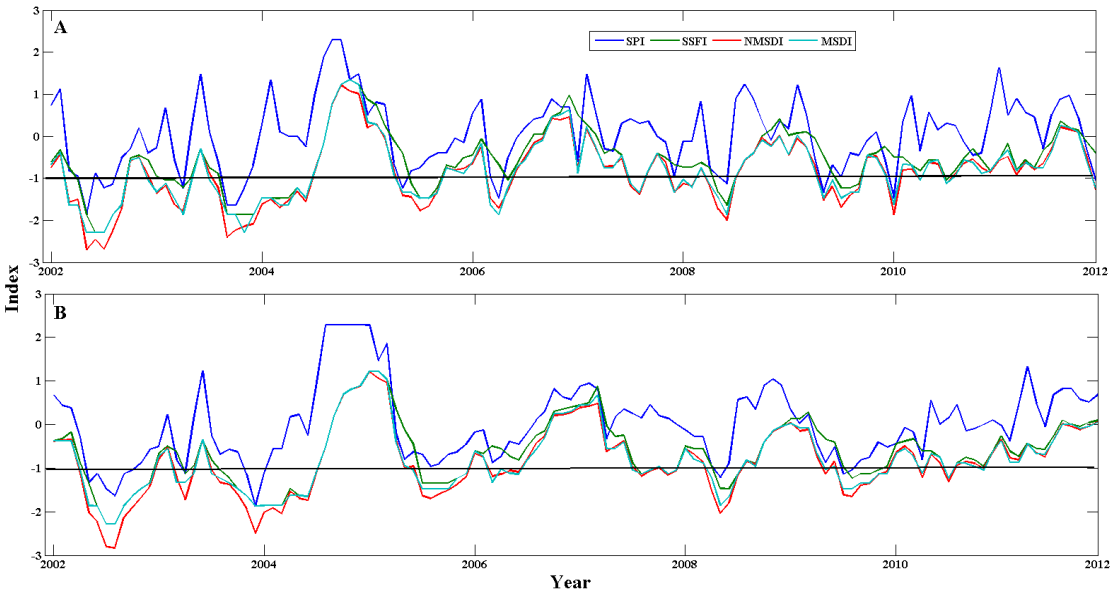

Fig. 3. The comparison between 3-month (A) and 6-month (B) NMSDI, MSDI, SPI and SSFI series covering 2002~2011.

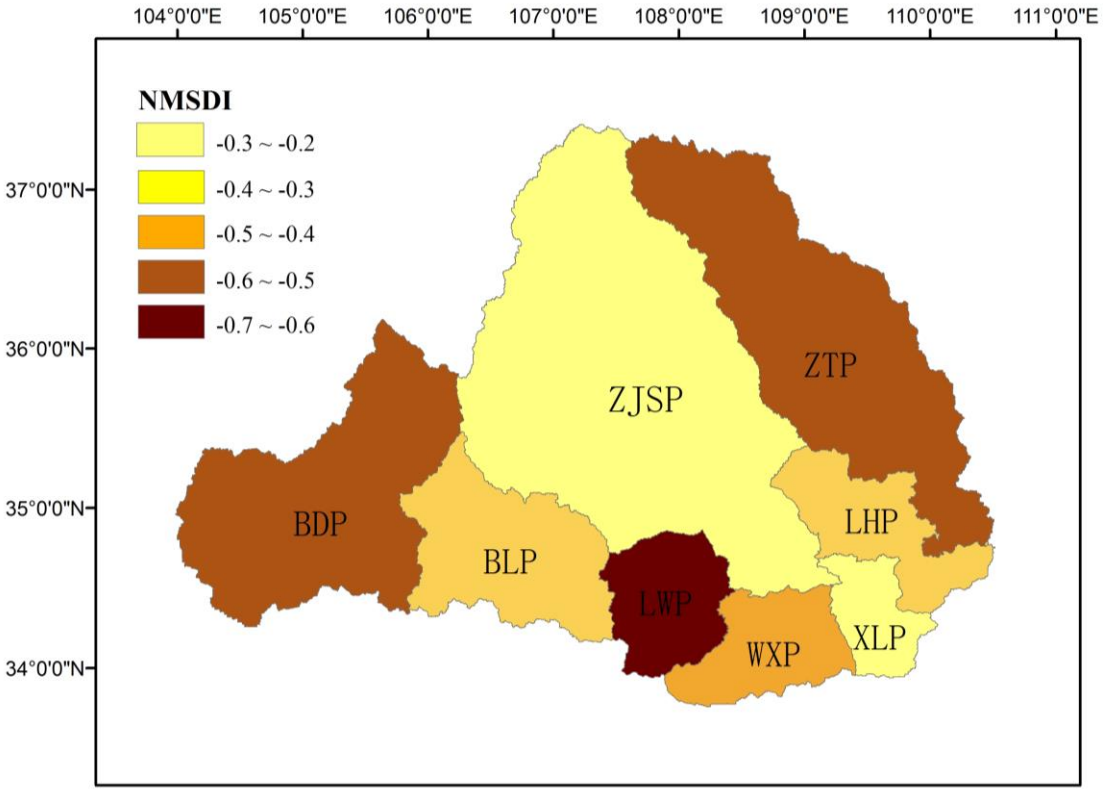

Fig. 4. The spatial distribution of mean annual 1-month NMSDI in each sub-area in the WRB.

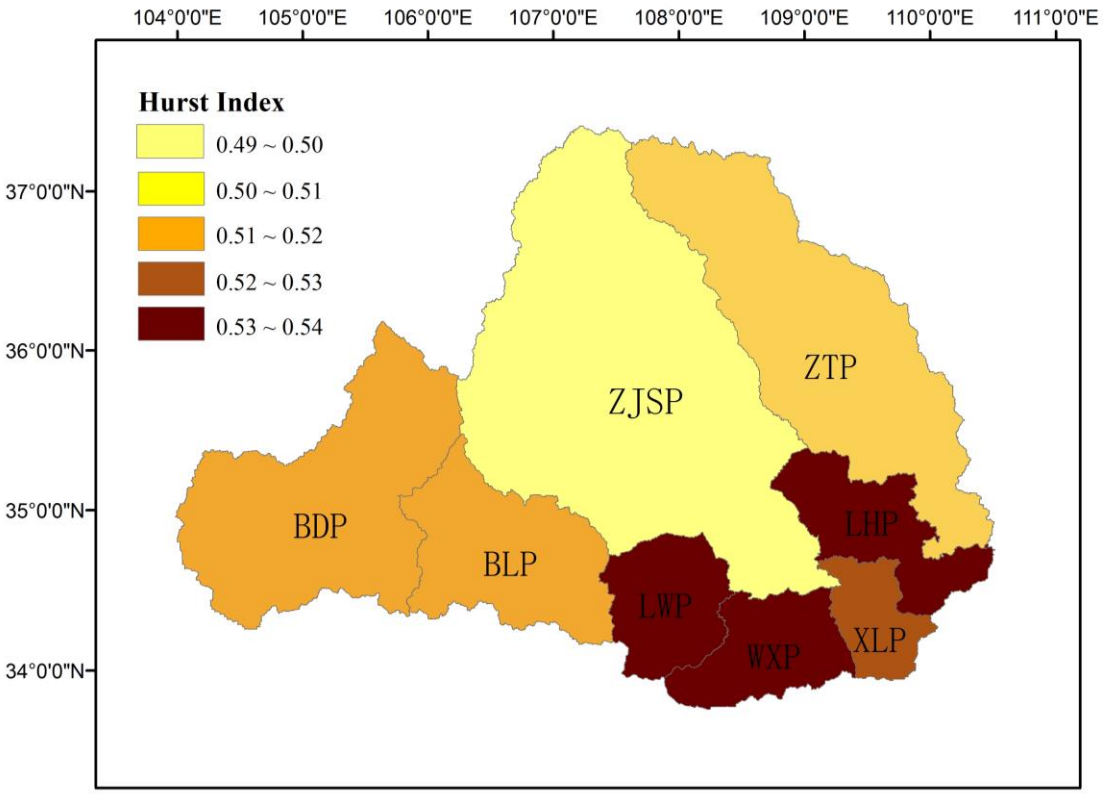

Fig. 5. The spatial distribution of Hurst index in each sub-area in the WRB.

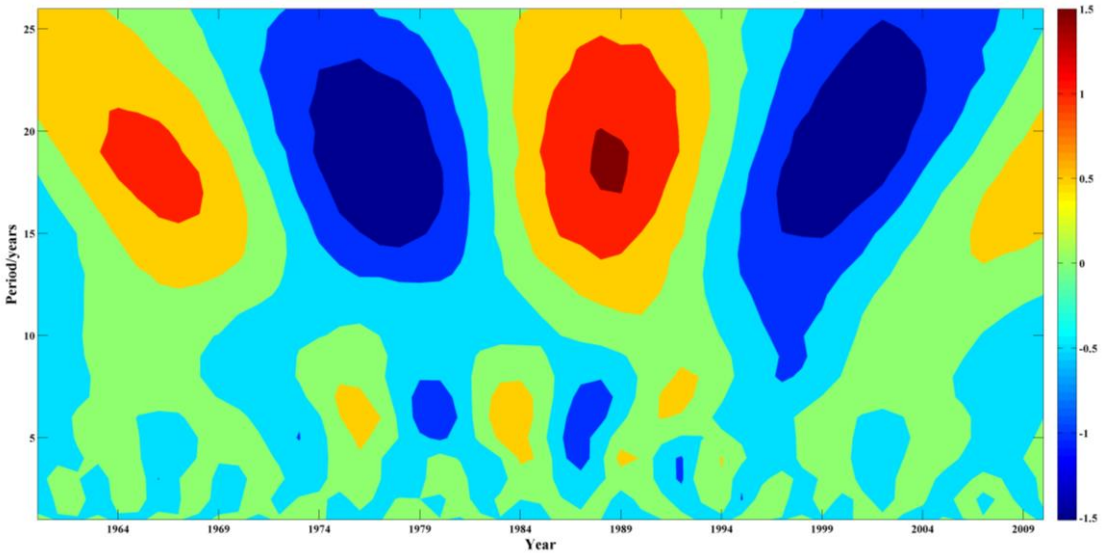

Fig. 6. The time-frequency distribution of annual 1-month NMSDI in the whole basin.

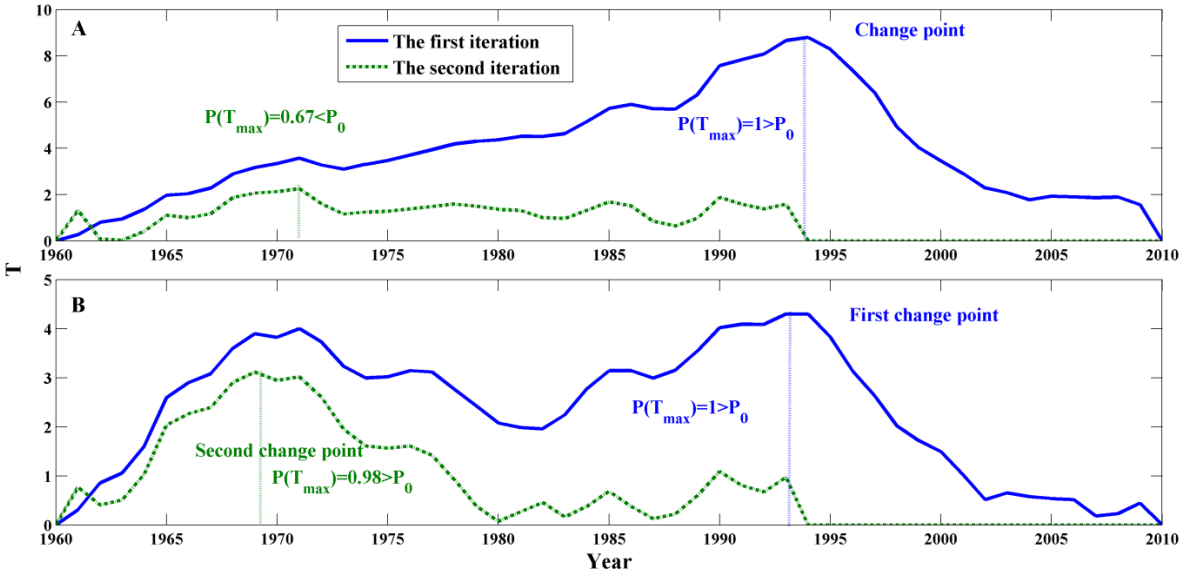

Fig. 7. The identification results of change points in Beidao (BDP) (A) and Lintong-Huaxian (LHP) (B) partitions.