Gradation of the significance level of trends in precipitation over China
Ping Xie, Yuxi Zhao, Yan-Fang Sang, Haiting Gu, Ziyi Wu and Vijay P. Singh

ABSTRACT
How to accurately detect and estimate the significance level of trends in hydroclimate time series is a challenge. Building on correlation analysis, we propose an approach for evaluating and grading the significance level of trend in a series, and apply it to evaluate the changes in annual precipitation in China. The approach involved first formulating the relationship between the correlation coefficient and trend’s slope. Four correlation coefficient thresholds are then determined by considering the influence of significance levels and data length, and the significance of trends is graded as five levels: no, weak, moderate, strong and dramatic. A larger correlation coefficient reflects a larger slope of trend and its higher significance level. Results of Monte-Carlo experiments indicated that the correlation coefficient-based approach not only reflects the magnitude of a trend, but also considers the influence of dispersion degree and mean value of the original series. Compared with the Mann-Kendall test used commonly, the proposed approach gave more accurate and specific gradation of the significance level of trends in annual precipitation over China. We find that the precipitation trends over China are not uniform, and the effects of global climate change on precipitation are not strong and limited to some regions.

Key words | detection and attribution, global climate change, hydrological variability, significance, stochastic process, trend

INTRODUCTION
Hydroclimatic variability in many basins and regions worldwide is changing significantly because of global climate change (Allen & Ingram 2002; Elliott et al. 2014; Trenberth et al. 2014). The detection and attribution of hydroclimatic variability is of great socioeconomic importance (Diffenbaugh et al. 2008; IPCC 2013), but considerable methodological challenges remain. The trend is one of the important indicators of hydroclimatic variability (Hamed 2008; Carmona et al. 2014; Rice et al. 2015), and the identification of a trend is the simplest and the most frequent way of detecting hydroclimatic variability (Yue et al. 2002). Various methods have been used for identifying trends in hydrological studies (Adam & Lettenmaier 2008; Ishak et al. 2013; Kirchner & Neal 2013; Sonali & Kumar 2013; Anghileri et al. 2014; Sang et al. 2014; Lopes et al. 2016; etc.). They can be classified into four types: data-fitting, time domain-based, frequency domain-based, and time-frequency domain-based test (Sang et al. 2013). Each type has advantages and disadvantages, which can affect the accuracy with which the trends are identified and our understanding of hydroclimate variability.
It is important to accurately quantify the statistical significance of a trend (Yue et al. 2002), and many methods have been developed for it (Sayemuzzaman & Jha 2014; Gjoneska et al. 2015). The Mann–Kendall (MK) non-parametric test is a widely used method, and has been successfully applied in past studies on the impacts of climate change (Kendall 1975; Burn & Hag Elmur 2002; Kisi & Ay 2014). However, each study uses a single significance level as a threshold for trend identification, which makes the significance evaluation by the MK test dependent on the chosen significance level. Besides, the slope of a trend cannot be directly estimated from the MK test. Furthermore, the detection of a trend not only depends on the magnitude of trend and the pre-assigned confidence level, but also on the probability distribution, sample size, dispersion degree of the time series (Yue et al. 2002; Adamowski et al. 2009; Shao & Li 2011; Hossein et al. 2012). These factors complicate the identification and assessment of the significance level of a trend. These issues can be overcome by a gradation of the significance level of a trend, taking into account the other factors that influence the significance of the trend.

The slope of a time series represents the significance level of its trend, but the slope can theoretically range from the negative infinity to positive infinity, making it unsuitable for the gradation of significance levels. The correlation coefficient (CC) quantifies the linear relationship between two variables, thus it can function as an effective index for the gradation of a trend’s significance level. A higher CC between a hydrological variable and its time order indicates a stronger significance level of its trend. The CC has values in the range of −1 to 1 and is mathematically related to the confidence level of a trend (Troch et al. 2015; McCuen 2016).

Here, we develop a new method for the gradation of the significance of a trend based on the correlation coefficient, and demonstrate its use by investigating the trends in annual precipitation in China. In the following section, we derive the relationship between the correlation coefficient and the slope of a trend, and describe our new method for the gradation of significance level of trend, and test its reliability through Monte-Carlo experiments. The annual precipitation data used in this study is described in the next section. We then apply the method to investigate trends in annual precipitation in China.

## METHODS

### Relationship between the correlation coefficient and the slope of trend

We use linear regression to calculate the monotonic trend in a hydroclimatic time series, following past studies (Sonali & Kumar 2013). The slope of a trend can directly reflect its significance level, but it cannot be used to grade the trend’s significance. We therefore develop a correlation analysis-based approach for grading the significance level of a trend, and begin by deriving the relationship between the correlation coefficient and trend’s slope.

Following stochastic hydrology (Marco et al. 2012), a hydroclimatic time series $x_t$ with periodicities removed can be simply described as:

$$x_t = a + bt + \eta_t$$  \hspace{1cm} (1)

where $a$ is a constant, and $t$ is time order with the total number of $n$; $\eta_t$ is a stationary random variable, with the mean value of zero and a constant variance, and its covariance $\text{cov}(\eta_i, \eta_j) = 0$ for $i \neq j$; $b$ is the slope of a trend and can be estimated as:

$$b = \frac{\sum_{t=1}^{n} (x_t - \bar{x})(t - \bar{t})}{\sum_{t=1}^{n} (t - \bar{t})^2}$$  \hspace{1cm} (2)

where $\bar{x}$ and $\bar{t}$ are the mean values of the series $x_t$ and its time order $t$, respectively. Then, the correlation coefficient $r$ is used to describe the linear relationship between the series $x_t$ and its time order $t$ (McCuen 2016):

$$r = \frac{\sum_{t=1}^{n} (x_t - \bar{x})(t - \bar{t})}{\sqrt{\sum_{t=1}^{n} (x_t - \bar{x})^2} \sqrt{\sum_{t=1}^{n} (t - \bar{t})^2}}$$  \hspace{1cm} (3)

By comparing Equation (2) with Equation (3), the $r \sim b$ relationship can be expressed as:

$$r = \frac{b \sigma_t}{\sigma_x}$$  \hspace{1cm} (4)

where $\sigma_t$ ($\sigma_x$) is the standard deviation of the time order $t$ (series $x_t$).
Rewriting Equation (1) by adding the constant $a$ to the random variable $n_t$, we get:

$$x_t = bt + u_t$$  

(5)

where $u_t$ has the mean value of $a$, and the linear trend component $bt$ is independent from $u_t$. The standard deviation $\sigma_x$ of series $x_t$ can be expressed as:

$$\sigma_x^2 = b^2\sigma_t^2 + \sigma_u^2$$  

(6)

where $\sigma_t$ is determined by the length $n$ of series $x_t$:

$$\sigma_t^2 = \frac{n^2 - 1}{12}$$  

(7)

and $\sigma_u$ is determined by the mean ($\bar{u}$) and coefficient of variation ($C_{vu}$) of series $u_t$:

$$\sigma_u = \bar{u}C_{vu}$$  

(8)

We substitute Equations (6)–(8) into Equation (4), and get a new equation of $r$:

$$r^2 = \frac{1}{1 + (12(\bar{u}C_{vu})^2)/(b^2(n^2 - 1))}$$  

(9)

Equation (9) describes the $r \sim b$ relationship. For a hydrological time series $x_t$, the statistical parameters of $\bar{u}$ and $C_{vu}$ of its random component $u_t$ are constant. From Equations (2)–(4) we know that when the correlation coefficient $r = 0$, the slope $b = 0$, indicating that there is no trend in the series $x_t$. Following Equation (9) we know that when $\bar{u}, C_{vu}$ and $n$ are first determined, the absolute values of $r$ and $b$ have a positive relationship. Therefore, the correlation coefficient can be used to grade the significance level of trends in a hydroclimate time series.

**Correlation coefficient-based approach for the gradation of trend**

After formulating the $r \sim b$ relationship in Equation (9), we need to determine the thresholds of correlation coefficient $r$ for the gradation of trend’s significance at appropriate confidence levels. For the statistical hypothesis test, different confidence levels are used, and each confidence level has a corresponding correlation coefficient $r$ (Lehmann & D’Abrera 2010; Murphy et al. 2014). The higher the confidence level, the stricter the statistical test of the significance level of a trend. In practice, confidence levels of 95 or 99% are often chosen for hydroclimate time series analysis, and the corresponding value of $r$ is denoted as $r_{95\%}$ and $r_{99\%}$ respectively. The values of $r_{95\%}$ and $r_{99\%}$ depend only on the data length. We use $r_{95\%}$ and $r_{99\%}$ as thresholds for the gradation of trend’s significance level.

For hydroclimate time series analysis, the data length should be at least 20 sampling points for robust trend analysis. For a series with a length of 20 or more, the critical (absolute) value of $r$, based on the $F$ test, is smaller than 0.6 for a confidence level of 99% or lower (Table 1). For example, for a length of 20, $r$ equals 0.56 at 99% confidence level (Corder & Foreman 2014). Therefore, we use 0.6 as the third threshold for the gradation of trends. In hydrological correlation analysis the $r$ is usually required to be larger than 0.8 to ensure its statistical significance at different situations. For example, the correlation coefficient value for a time series with the length of 10 must be as high as 0.77 at 99% confidence level. Thus, we use 0.8 as the fourth gradation threshold.

Our proposed approach for assessing and grading the significance level of a trend using the four thresholds identified above is as follows:

1. For a time series $x_t$, calculate the correlation coefficient $r$ between $x_t$ and its time order $t$ using Equation (5).
2. Choose two confidence levels $\alpha$ and $\beta$ ($\alpha < \beta$), and determine the corresponding correlation coefficient thresholds (denoted as $r_\alpha$ and $r_\beta$) for the given data length. Use 0.6 and 0.8 as the two other thresholds.
3. Compare the absolute value of $r$ ($|r|$) in step (1) with the four thresholds selected in step (2).
4. If $|r| < r_\alpha$, the trend is insignificant at the confidence level $\alpha$. We denote this as ‘no trend’.
5. If $r_\alpha \leq |r| < r_\beta$, then the trend is significant at level $\alpha$ but insignificant at level $\beta$. We denote this as ‘weak trend’.
6. If $r_\beta \leq |r| < 0.6$, then the trend is significant at the level $\beta$ but may not be significant at a higher confidence level. We denote this as ‘moderate trend’.
7. If $0.6 \leq |r| < 0.8$, then the trend is significant in most but not all situations, and we denote this as ‘strong trend’.
8. If $|r| \geq 0.8$, then the trend is significant in all situations, and we denote this as ‘dramatic trend’.

Following the above steps, the significance level of trend in a hydroclimate time series can be graded into five ranks (Table 2), and ten ranks if the negative and positive trends are separated.

### Verification of the proposed approach

To verify the reliability of the proposed approach and further investigate the influence of some factors on the gradation of significance level of trend, we have designed the following Monte-Carlo experiments:

1. We generate 30 random time series that follow the Pearson-III probabilistic distribution, which is used commonly for hydrological analysis and design in China. Each time series has a same length ($\text{n} = 100$), mean value ($\mu = 1,000$), variation coefficient ($\text{C}_{v_\mu} = 0.2$), and skewness coefficient ($\text{C}_{\text{s}} = 0.4$). We denote each of this time series as $x_j$, $j = 1, 2, \ldots, 30$.
2. To each series $u_j$ we add a different trend component $b t$ ($b = 0.5, 1, 1.5, \ldots, 15$), and the new time series is denoted as $y_j$, $j = 1, 2, \ldots, 30$.
3. We use Equation (3) to calculate the correlation coefficient $r$ between the series $x_t$ and its time order $t$.
4. We repeat the above steps 1,000 times (i.e. $i = 1, 2, \ldots, 1,000$) to ensure the stability of the result $r_i$.
5. The average value of $r_i$ in each group of experiments is calculated as:

$$r_i = \frac{1}{1,000} \sum_{i=1}^{1,000} r_{ij} (j = 1, 2, \ldots, 30)$$  \hspace{1cm} (10)

Because the true slope $b$ of trend in each synthetic series is known, the true value of the correlation coefficient (denoted as $r_1$) can be calculated by Equation (9). The correlation coefficient calculated by Equation (10) (denoted as $r_2$) is then verified against $r_1$. We find that for all 30 groups, the

### Table 1 | Critical values of correlation coefficient $r$ of series with different length $n$ and under different confidence levels to ensure its statistical significance

<table>
<thead>
<tr>
<th>Degree of freedom $n-m-1$</th>
<th>Confidence level $\alpha$</th>
<th>$90%$</th>
<th>$95%$</th>
<th>$98%$</th>
<th>$99%$</th>
<th>Degree of freedom $n-m-1$</th>
<th>Confidence level $\alpha$</th>
<th>$90%$</th>
<th>$95%$</th>
<th>$98%$</th>
<th>$99%$</th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td>0.549</td>
<td>0.632</td>
<td>0.716</td>
<td>0.765</td>
<td>20</td>
<td>0.360</td>
<td>0.423</td>
<td>0.492</td>
<td>0.537</td>
<td>0.537</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>0.521</td>
<td>0.692</td>
<td>0.765</td>
<td>0.835</td>
<td>25</td>
<td>0.323</td>
<td>0.381</td>
<td>0.445</td>
<td>0.487</td>
<td>0.537</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>0.497</td>
<td>0.576</td>
<td>0.658</td>
<td>0.708</td>
<td>30</td>
<td>0.296</td>
<td>0.349</td>
<td>0.409</td>
<td>0.449</td>
<td>0.537</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>0.476</td>
<td>0.553</td>
<td>0.634</td>
<td>0.684</td>
<td>35</td>
<td>0.275</td>
<td>0.325</td>
<td>0.381</td>
<td>0.418</td>
<td>0.537</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>0.458</td>
<td>0.532</td>
<td>0.612</td>
<td>0.661</td>
<td>40</td>
<td>0.257</td>
<td>0.304</td>
<td>0.358</td>
<td>0.393</td>
<td>0.537</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>0.441</td>
<td>0.514</td>
<td>0.592</td>
<td>0.641</td>
<td>45</td>
<td>0.244</td>
<td>0.288</td>
<td>0.338</td>
<td>0.372</td>
<td>0.537</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>0.426</td>
<td>0.497</td>
<td>0.574</td>
<td>0.623</td>
<td>50</td>
<td>0.231</td>
<td>0.273</td>
<td>0.322</td>
<td>0.354</td>
<td>0.537</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>0.412</td>
<td>0.482</td>
<td>0.558</td>
<td>0.606</td>
<td>60</td>
<td>0.211</td>
<td>0.250</td>
<td>0.295</td>
<td>0.325</td>
<td>0.537</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>0.400</td>
<td>0.468</td>
<td>0.543</td>
<td>0.590</td>
<td>70</td>
<td>0.195</td>
<td>0.232</td>
<td>0.274</td>
<td>0.302</td>
<td>0.537</td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>0.389</td>
<td>0.456</td>
<td>0.529</td>
<td>0.575</td>
<td>80</td>
<td>0.183</td>
<td>0.217</td>
<td>0.257</td>
<td>0.283</td>
<td>0.537</td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>0.378</td>
<td>0.444</td>
<td>0.516</td>
<td>0.561</td>
<td>90</td>
<td>0.173</td>
<td>0.205</td>
<td>0.242</td>
<td>0.267</td>
<td>0.537</td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>0.369</td>
<td>0.433</td>
<td>0.503</td>
<td>0.544</td>
<td>100</td>
<td>0.164</td>
<td>0.195</td>
<td>0.230</td>
<td>0.254</td>
<td>0.537</td>
<td></td>
</tr>
</tbody>
</table>

$n$ represents the data length, and $m$ ($m - 1$ for the study) represents the unknown members of dimension.

### Table 2 | Thresholds of correlation coefficient $r$ used for the gradation of significance level of trends in hydroclimatic time series under confidence level of $\alpha$ and $\beta$

<table>
<thead>
<tr>
<th>Correlation coefficient</th>
<th>Significance level</th>
<th>Correlation coefficient</th>
<th>Significance level</th>
</tr>
</thead>
<tbody>
<tr>
<td>$0 \leq</td>
<td>r</td>
<td>&lt; r_0$</td>
<td>No trend</td>
</tr>
<tr>
<td>$r_0 \leq</td>
<td>r</td>
<td>&lt; r_7$</td>
<td>Weak trend</td>
</tr>
<tr>
<td>$r_7 \leq</td>
<td>r</td>
<td>&lt; 0.6$</td>
<td>Moderate trend</td>
</tr>
</tbody>
</table>
relationship between the slope and the correlation coefficients. In Figure 1, we show the significance levels of trends for five series with slopes, $b$, of 0, 1.5, 4.0, 6.0, 10.0. According to Equation (3), the correlation coefficient $r$ of the five series is 0, 0.212, 0.501, 0.655, and 0.824, respectively. The five series fall into the five different grades of significance (Table 4). Figure 1 shows that the slopes of the trends in the five series increase with $r$, indicating the applicability of Equation (9).

### Influence of other factors on the gradation of trend's significance level

From Equation (9) we know that the gradation of the significance level of a trend (i.e. the magnitude of $r$), depends not only on the slope $b$ of the trend, but also on the mean $\bar{u}$ and variation coefficient $C_{vu}$ of the random component $u_t$ of the time series. We design two sets of Monte-Carlo (MC) experiments to investigate how these two factors influence the $r \sim b$ relationship.

For the first set of MC experiments, we generate a synthetic series with length $n$ of 100 and mean value $\bar{u}$ of 1,000, and vary $C_{vu}$ to investigate its influence on the $r \sim b$ relationship. Figure 2(a) shows that for any $C_{vu}$, $r$ increases with $b$, but the increase is slower when $b$ is larger. Furthermore, the $r \sim b$ curve becomes flatter at higher $C_{vu}$ value. This shows that the dispersion degree of a series has a strong influence on the significance level of its trend. For two series with the same trend but different dispersion degrees, the series with a smaller dispersion degree will have a more obvious trend with a higher significance level and is easily detectable.

In the second set of MC experiments, the length $n$ of synthetic series is 100 and the value of $C_{vu}$ is fixed at 0.2, and $\bar{u}$ is varied to investigate its influence on the $r \sim b$ relationship (Figure 2(b)). $r$ increases with $b$ for all $\bar{u}$ values, but the $r \sim b$ relationship is weaker for a larger $\bar{u}$. Thus, the mean magnitude of a series also has a strong influence on the significance level of its trend. For two series with the same trend but different mean values, for example, $\bar{u}$ of 200 and 1,000, the series with a smaller mean value will have a more significant trend.

The results in Figure 2 show the influence of the mean $\bar{u}$ and variation coefficient $C_{vu}$ of a series on the significance level of its trend.
of its trend. These two factors reflect the different ratio between the trend and the random component of a time series. For a time series with a smaller mean value and a smaller dispersion degree, this ratio is higher and the trend will have a higher significance level. On the other hand, for a time series with a larger mean value and a large dispersion degree, the significance level of its trend is weaker. To clarify this further, we use the signal-to-noise (SNR) index to quantify the influence of the two factors (Herrick 2014) on the magnitude of the correlation coefficient. The SNR index is defined as the ratio between the variance of a trend and its random component:

$$\text{SNR} = \frac{(b\sigma_t)^2}{\sigma_u^2} = \frac{b^2\sigma_t^2}{\bar{u}^2 C_{vu}}$$  \hspace{1cm} (11)
By substituting Equation (11) into Equation (9), the $r \sim \text{SNR}$ relationship can be described as:

$$r^2 = 1 - \frac{1}{\text{SNR} + 1}$$ \hspace{1cm} (12)

Equation (11) shows that if the random component ($\bar{u}$ and $C_{vu}$) of a series is larger, the SNR values are smaller, and the trend would be difficult to identify. Therefore, $r$ and SNR have a positive relationship as shown in Equation (12) and are consistent with Figure 2. The trend in a series with a smaller dispersion degree and mean value would be more easily identified. This, the correlation coefficient $r$ reflects not only the magnitude of a trend, but also considers the influence of the dispersion degree and the mean value of the time series. Therefore, the correlation coefficient-based approach developed in this study is effective for estimating and grading the significance level of trends in hydrological time series.

**STUDY AREA AND DATA**

In this study, 520 meteorological stations (Figure 3) were chosen for investigating the trends in annual precipitation over China. The data were obtained from the China Meteorological Data Sharing Service System (http://cdc.cma.gov.cn/). These stations were chosen by considering the length, consistency and completeness of data records. They are approximately uniformly distributed over China, with somewhat fewer stations in the southwest region. All of the stations have measurements from 1961 to 2013, with no missing values.

**RESULTS AND DISCUSSION**

Precipitation is an important variable for understanding the variability and changes in hydroclimatic systems (Brunetti et al. 2006; Ashouri et al. 2014; Trenberth et al. 2014). There have been many studies on the precipitation variability over China and at regional scales (Zhai et al. 1999; Wang et al. 2004; Ma et al. 2008; Ye 2014; Zhang et al. 2016), but their conclusions and interpretations differ. Some studies indicate that the precipitation in many regions, especially in northwest China, fluctuated considerably and show significant trends over recent decades due to the influence of global climate change (Chen et al. 2013; Sang et al. 2013a, 2013b; Wan et al. 2015; Gu et al. 2017; Yang et al. 2017).

![Figure 3](http://iwaponline.com/hr/article-pdf/49/6/1890/509271/nh0491890.pdf)
while other studies indicate that precipitation in many regions has kept its stochastic characteristics and has not changed significantly (Gao et al. 2012; Sun et al. 2012). Thus, the significance level of trends in precipitation over China remains unclear. An analysis of the spatiotemporal variability of precipitation over China is important for water resources management and many other water activities.

We used our correlation coefficient-based approach to investigate the significance level of trends in the annual precipitation time series measured at 520 stations over China. The confidence levels of 95 and 99% were used for the precipitation series analysis, and the upper and lower limits of correlation coefficient for each rank were calculated. For comparison purposes, trends of all these annual precipitation series were also identified by the MK test. Results in Figure 3 (left) indicate that among the 520 precipitation series, the trends of only 60 series (11.5%) are significant, and the other 460 series do not show any obvious trends, that is, their trends are graded as ‘no trend’. Surprisingly, none of the trends are strong or dramatic. The upward trends of precipitation with a moderate significance level are seen in 21 stations in the northwest corner of China and in the northeast boundary of the Tibet Plateau, and trends with a weak significance level at 12 stations in those regions. The downward trends of precipitation have moderate significance levels at eight stations in the Yunnan-Guizhou Plateau (100°–111°E, 22°–30°N) in southwest China and two stations in the centre of north China. Also, in the Yunnan-Guizhou Plateau and its surrounding regions, precipitation series at nine stations have downward trends at a weak significance level. The upward trends of precipitation with a weak significance level are detected at four stations in southeast coastal areas of China.

In comparison, the results by the MK test in Figure 3 (right) show that precipitation has downward trends in the mid-arid and mid-humid regions from the northeast to southwest China. In northwest China, including the Tibet Plateau and southeast China, precipitation has a mainly upward trend. At the significance of 95% confidence level, the thresholds of ±1.96 are used to distinguish the statistical characters of the MK test, with a whole value range of −4.01 to 4.72. Results show that trends of 64 precipitation series (12.3% of the total series) are identified as significant by the MK test, but the other 456 series indicate no obvious trends.

Figure 3 indicates that the spatial distribution of the significance level of trends in precipitation series obtained from the proposed approach and the MK test are similar. Global climate change has likely led to the strengthening westerlies but the weakening Indian summer monsoon over recent decades (Wu 2005; Thompson et al. 2006), which would influence the precipitation variability over China, especially over the western regions. The increase in precipitation in northwest China can be due to the strengthening westerlies; the precipitation decrease on the Yunnan-Guizhou Plateau and its surrounding regions can be caused by the weakening Indian summer monsoon (Sang et al. 2016).

Moreover, the number of precipitation series with significant trends detected by the two methods (60 in our method and 64 in the MK test) are similar. The indices used to quantify the significance of trends in the two methods also indicate a positive relationship (Figure 4). The similarity of the results of our proposed approach with the MK test, which has been successfully applied for trend identification in the past, demonstrates the reliability of our approach for the significance evaluation of trends. Moreover, the MK test can judge only whether the trend is significant or not at a certain confidence level, but our approach can also be used for the gradation of the significance level of trends (Figure 4). In the MK test, any value greater than 1.96 indicates a significant upward trend, but there is no distinction based on the degree of significance. The relationship between the slope of a trend and the test statistic (Z) in the MK test is also unknown. However, accurate gradation of the significance level of trends is urgently needed in practical analysis, and the approach proposed in this study meets the purpose. We use our approach to grade the significance of the trends and understand the degree of significance of the trends in the annual precipitation data at each station.

In addition, we computed the values of SNR and r of each precipitation series, and show their scatter diagram in Figure 5. As expected, all 520 points fall on the standard curve in Equation (12). The absolute value of |r| increases with SNR, but the increase rate becomes slower at larger SNR values. Those precipitation series with larger SNR values have higher significance level of trends.
Figure 4 | Relationship between the correlation coefficient \( r \) and the statistic character \( Z \) of the MK test obtained from each of the 520 precipitation time series over China, and the different significance levels of their trends.

A No trend   B Weak trend   C Moderate trend   D Strong trend   E Dramatic trend
a Significantly downward trend  b Insignificant trend  c Significantly upward trend

Figure 5 | Scatter diagram between the SNR and the absolute value of \( r \) of each of the 520 precipitation time series over China, and their relationship with the standard curve obtained from Equation (12).
From the above analysis, we conclude that although global climate change has a major influence on hydroclimatic variability worldwide, its influence on the precipitation over China during 1961–2013 is not as strong as one might expect. In most regions in China, precipitation has changed over long timescales, but the change is insignificant. The strengthening westerlies cause a precipitation increase in northwest China, and the weakening Indian summer monsoon causes a precipitation decrease on the Yunnan-Guizhou plateau, but the observed trends are weak or moderate. There is no strong or dramatic trend in precipitation in China during the recent five decades. Thus, in contrast to other studies which provide only an approximation of the significance level of trends in precipitation (Gao et al. 2012; Sun et al. 2012), we conclude that precipitation over China does not show obvious trends, although global climate change causes some precipitation changes in local regions.

CONCLUSIONS

How to accurately detect and estimate the significance level of trends is a challenge for understanding hydroclimatic variability and assessing its potential impacts. In this paper, we use the correlation coefficient and develop an approach for the gradation of the significance level of trends in a hydroclimate time series. We first derive the relationship between the correlation coefficient and the slope of trend. Then, by determining four correlation coefficient thresholds, we propose a gradation of the significance of trends of time series into five levels: no trend, weak trend, moderate trend, strong trend, and dramatic trend. A larger correlation coefficient value implies a larger slope of trend and a higher significance level.

The results of Monte-Carlo experiments indicate that the mean value and dispersion degree of a series have a strong influence on the calculation of the significance level of trends. The correlation coefficient-based approach of ours not only reflects the magnitude of a trend, but also considers the influence of dispersion degree and mean value of the original series. Therefore, it is an effective approach for estimating and grading the significance level of trends in a hydrological time series. Compared with other widely used methods, the main advantage of our method is that it provides a method for the gradation of significance level of trends, and also quantifies the influences of statistical characteristics of the original series. More research is needed to further verify the applicability of this method by considering many other hydroclimatic variables and non-linear trends.

We analyzed the changes in precipitation over China over five recent decades, and find that the significance of trends and their spatial distribution calculated with our approach are similar to the MK test. However, compared to the MK test, our approach provides a gradation of the significance level of trends. We found that although global climate change has a great influence on the hydroclimate variability worldwide, its influence on precipitation over China is not strong. None of the 520 meteorological stations analyzed showed a strong or dramatic trend in precipitation. The precipitation trends over China are not uniform, and the effects of global climate change on precipitation are limited to some regions. Other deterministic characteristics, such as periodicities and step changes, need to be further studied to determine whether precipitation over China mainly shows stochastic characteristics or not.

ACKNOWLEDGEMENTS

The authors gratefully acknowledged the valuable comments and suggestions given by the editors and the anonymous reviewers. The authors thank Viral Shah for language editing. This study was financially supported by the National Natural Science Foundation of China (No. 91547205, 91647110, 51579181), the Program for the ‘Bingwei’ Excellent Talents from the Institute of Geographic Sciences and Natural Resources Research, CAS, the Youth Innovation Promotion Association CAS (No. 2017074), and the National Mountain Flood Disaster Investigation Project (SHZH-IWHR-57).

REFERENCES

Adam, J. C. & Lettenmaier, D. P. 2008 Application of new precipitation and reconstructed streamflow products to


First received 5 December 2017; accepted in revised form 29 May 2018. Available online 12 June 2018