Hydrol. Earth Syst. Sci., 12, 207–221, 2008
www.hydrol-earth-syst-sci.net/12/207/2008/

Hydrology and Earth System Sciences

Detecting changes in extreme precipitation and extreme streamflow in the Dongjiang River Basin in southern China

W. Wang[1], **X. Chen**[1], **P. Shi**[1], **and P. H. A. J. M. van Gelder**[2]

[1]State Key Laboratory of Hydrology-Water Resources and Hydraulic Engineering, Hohai University, Nanjing, 210098, China
[2]Section of Hydraulic Engineering, Faculty of Civil Engineering & Geosciences, Delft University of Technology, Delft, The Netherlands

Received: 28 June 2007 – Published in Hydrol. Earth Syst. Sci. Discuss.: 23 July 2007
Revised: 22 October 2007 – Accepted: 12 December 2007 – Published: 1 February 2008

Abstract. Extreme hydro-meteorological events have become the focus of more and more studies in the last decade. Due to the complexity of the spatial pattern of changes in precipitation processes, it is still hard to establish a clear view of how precipitation has changed and how it will change in the future. In the present study, changes in extreme precipitation and streamflow processes in the Dongjiang River Basin in southern China are investigated with several nonparametric methods, including one method (Mann-Kendall test) for detecting trend, and three methods (Kolmogorov–Smirnov test, Levene's test and quantile test) for detecting changes in probability distribution. It was shown that little change is observed in annual extreme precipitation in terms of various indices, but some significant changes are found in the precipitation processes on a monthly basis, which indicates that when detecting climate changes, besides annual indices, seasonal variations in extreme events should be considered as well. Despite of little change in annual extreme precipitation series, significant changes are detected in several annual extreme flood flow and low-flow series, mainly at the stations along the main channel of Dongjiang River, which are affected significantly by the operation of several major reservoirs. To assess the reliability of the results, the power of three non-parametric methods are assessed by Monte Carlo simulation. The simulation results show that, while all three methods work well for detecting changes in two groups of data with large sample size (e.g., over 200 points in each group) and large differences in distribution parameters (e.g., over 100% increase of scale parameter in Gamma distribution), none of them are powerful enough for small data sets (e.g., less than 100 points) and small distribution parameter difference (e.g., 50% increase of scale parameter in Gamma distribution). The result of the present study raises the concern of the robustness of statistical change-detection methods, shows the necessity of combined use of different methods including both exploratory and quantitative statistical methods, and emphasizes the need of physically sound explanation when applying statistical test methods for detecting changes.

Correspondence to: W. Wang
(w.wang@126.com)

1 Introduction

Extreme meteorological and hydrological events may have huge impacts on human society. With significant global warming, it seems that the occurrence of extreme events gets more frequent, and therefore more and more efforts have been put on the research of extreme events in various relevant fields in the last decade.

It is widely conceived that with the increase of temperature, the water cycling process will be sped up, which in consequence will possibly result in the increase of precipitation amount and intensity. Many outputs from Global climate models (GCMs) indicate the possibility of substantial increases in the frequency and magnitude of extreme daily precipitation (e.g., Gordon et al., 1992; Fowler and Hennessy, 1995; Hennessy et al., 1997; McGuffie, 1999). The increase also shows itself in observed data. Karl et al. (1995) found that the contribution to total annual precipitation of 1-day precipitation events exceeding 50.8 mm (2.0 in.) increased from about 9% in the 1910s to about 11% in the 1980s and 1990s. Further on, Karl and Knight (1998) found that the 8% increase in precipitation across the contiguous United States since 1910 is reflected primarily in heavy and extreme daily precipitation events. The results of Kunkel et al. (1999) confirmed that the national trend in short duration (1–7 d) extreme precipitation events for the United States is upward at a rate of 3% decade^{-1} for the period 1931–1996. In Australia, much of the country has experienced increases in heavy precipitation events, except in southwestern Australia where there has been a decrease in both the number

Published by Copernicus Publications on behalf of the European Geosciences Union.

Table 1. Meteorological and hydrological Gauging stations.

Station type	Station	Latitude	Longitude	Elevation (m)	Drainage area (km^2)	Period	Length (year)
Meteorology	59096	114°29′	24°22′	214.5	–	1953–2004	52
	59102	115°39	24°57	303.9	–	1956–2004	49
	59293	114°41′	23°44′	40.8	–	1953–2004	52
	59298	114°25′	23°05′	22	–	1953–2004	52
Hydrology	Jiuzhou	114°59[1]	23°07[1]	–	385	1959–2005	47
	Yuecheng	114°16[1]	24°06[1]	–	531	1960–2005	46
	Lantang	114°56[1]	23°26[1]	–	1080	1958–2005	48
	Longchuan	115°15′	24°07′	–	7699	1952–2002	51
	Heyuan	114°42′	23°44′	–	15 750	1951–2002	52
	Boluo	114°18′	23°10′	–	25 325	1953–2002	50

of rainy days and heavy precipitation events (Suppiah and Hennessy 1998; Haylock & Nicholls, 2000). In the United Kingdom increases in heavy wintertime events and decreases in heavy summertime events have been found (Osborn et al., 1999). Moberg et al. (2006) showed that, winter precipitation totals, averaged over 121 European stations north of 40° N, have increased significantly by 12% per 100 yr, and trends in 90th, 95th and 98th percentiles of daily winter precipitation have been similar. New et al. (2001) showed that, on the basis of gridded observed monthly precipitation data, global land precipitation (excluding Antarctica) has increased by about 9 mm over the twentieth century, and data from a number of countries provide evidence of increased intensity of daily precipitation, generally manifested through increased frequency of wet days and an increased proportion of total precipitation occurring during the heaviest events. Roy and Balling (2004) found that, in general, evidence exists for an increase in the frequency of extreme precipitation events in India over the period 1910 to 2000. According to the observed data over half of the land area of the globe, there has been a widespread increase in the frequency of very heavy precipitation in the mid-latitudes during the past 50 to 100 yr (Groisman et al., 2005). The results of Zhai et al. (2005) indicated that while there is little trend in total precipitation for China as a whole, significant increases in extreme precipitation have been found in western China, the mid-lower reaches of the Yangtze River, and parts of the southwestern and southern China coastal areas.

While in many areas increased intensity of heavy rainfall is observed, in quite a number of other areas and other studies no significant increase is observed. For instance, Nicholls et al. (2000) calculated various indices for monitoring variations in Australian climate extremes, and showed that, most of the trends in the various indices of climate extremes investigated were relatively weak and lacked statistical significance, and no clear trend has emerged in the percentage of Australia in extreme rainfall (drought or wet) conditions, since 1910. Zhang et al. (2001) showed that there has been no long-term trend in the frequency or intensity of extreme precipitation events in Canada during the 20 century. Koning & Franses (2005) showed that no statistically significant shift is found in the annual largest values of daily rainfall in the Netherlands over the course of a century. Zhang et al. (2005b) showed that the trends in precipitation indices, including the number of rainy days, the average precipitation intensity, and maximum daily precipitation events in Middle East, are weak in general. Su et al. (2006) analyzed the observed extreme temperature and precipitation trends over Yangtze River Basin in China from 1960 to 2002 on the basis of daily data from 108 meteorological stations, and found no statistically significant change in heavy rain intensity from a basin-wide point of view, although a significant positive trend was found for the number of days with heavy rainfall (daily rainfall ≥50 mm). Klein Tank et al. (2006) found that most regional indices of precipitation extremes show little change between 1961 and 2000 in central and south Asia. New et al. (2006), in their study of trends in daily extremes over mainly southern Africa for the period 1961 to 2000, concluded that there are few consistent and statistically significant trends in the precipitation indices that they calculated.

While evidences of increasing trends are presented for many regions, statistically significant decreasing trends in extreme rainfall events have also been found in some areas, including the Sahel region of Nigeria (Tarhule and Woo, 1998), southwestern and western Australia (Suppiah and Hennessy 1998; Haylock and Nicholls, 2000), Southeast Asia and parts of the central Pacific (Manton et al., 2001; Griffiths et al., 2003), northern and eastern New Zealand (Salinger and Griffiths, 2001), the UK in summer (Osborn et al., 2000), Poland (Bielec, 2001), and some parts in India (Roy and Balling, 2004). Therefore, the spatial pattern of changes in precipitation is complex and varies over the world.

Fig. 1. Location of the study area (left) and locations of gauging stations (right). (Note: the star symbols denote meteorological stations, and the diamond symbols denote hydrological stations).

On the other hand, in the context of significant global changes in many regions, whether or not the streamflow processes has changed is of great concern because streamflow processes are mainly driven by meteorological processes, and possibly more extreme weather may result in higher flood and drought risks. Some results show increases in extreme events. For example, when investigating the relationship of changes in the probability of heavy precipitation and high streamflow over the contiguous United States, Groisman et al. (2001) showed that the variations of high and very high streamflow and heavy and very heavy precipitation are similar. In a recent study, Zhang et al. (2005a) evaluated the relations between the temperature, the precipitation and the streamflow during 1951–2002 of the Yangtze River basin, suggesting that the present global warming will intensify the flood hazards in the basin. At the same time, some others show no significant change in extreme flood events. For instance, Mudelsee et al. (2003) found no upward trends in the occurrence of extreme floods in central Europe; Kundzewicz et al. (2005) showed that the analysis of annual maximum flows does not support the hypothesis of ubiquitous growth of high flows.

In summary, although some notable work has been done on detecting the change in extreme meteorological and hydrological events, it is still not conclusive about how hydro-meteorological events have changed in different regions over the world. Such controversy may arise in two perspectives. Firstly, the change in climate may vary significantly over different regions, and the link between excessive precipitation and hydrologic flooding is affected by several factors, including meteorological factors (such as antecedent precipitation amount and the intensity, duration and spatial pattern of precipitation events), human activities (such as land-use change and dam construction), and basin characteristics (such as the size, topography, control structures, and drainage network of the basin). These factors vary from event to event, from season to season, and from region to region. Hence, we need to have more exploration about how climate changed over different regions so as to get a comprehensive view of changes in water cycle all over the world. Secondly, in detecting the changes in extreme hydro-meteorological events, there are two major issues that are more or less subjective: the definition of extreme events and the methods for assessing the changes. How the extreme events are defined and what method is employed may differ among different researchers, which may lead to different conclusions. Furthermore, although many methods have been applied, it is not common to consider the uncertainty of the results derived with those methods.

The objective of this study is to determine whether the precipitation process, especially the extreme precipitation, in the Dongjiang River Basin in southeastern China has changed in the context of global warming, and whether streamflows, including high flows and low-flows, in the basin have changed as well with the intensified climate change and human intervention. Furthermore, we want to assess the reliability of several techniques, including the Kolmogorov–Smirnov test, Levene's test and quantile test, for detecting changes in the probability distribution of precipitation. In Sect. 2, we will briefly describe our study area and the data used. Description of the change detection methods used in the study will be presented in Sect. 3. Results for detecting changes in extreme precipitation and streamflow are reported in Sect. 4, followed by some discussions and conclusions in Sect. 5.

2 Study area and data used

Dongjiang River originates in Jiangxi Province in southern China and flows through eastern Guangdong Province, converged into the Pearl River. It has a 562 km long mainstream with a drainage area of 35 240 km^2. The streamflow process of Dongjiang River demonstrates strong seasonality due to a sub-tropical monsoon climate. The Dongjiang River is important for not only the local region but also for Hong Kong because about 80% of Hong Kong's water supply comes from Dongjiang River through cross-basin water transfer. Three major reservoirs (see Fig. 1) were built in the basin, inlcuding Xinfengjiang Reservoir (started to operate in 1959), Fengshuba Reservoir (started to operate in 1973) and Baipenzhu Reservoir (started to operate in 1984).

In the present study, daily precipitation data at 4 meteorological stations and daily streamflow data at 6 hydrological stations are used for the analysis (Fig. 1). The descriptions of all the 10 gauging stations are listed in Table 1. Annual precipitation in the center of the basin (at station 59293) is about 1932 mm, with nearly 80% falling in spring and summer from March to August. As shown in Fig. 1, three (Longchuan, Heyuan and Boluo) of the 6 streamflow gauging stations are significantly impacted by reservoir operation, whereas the other three are little impacted by any major hydraulic works. Daily discharges were available for 45 to 50 yr, whereas the daily precipitation for 52 yr. Very few data are missing in these series, and missing data are filled with linear interpolation. For the period 1956 to 2004, the basin daily areal rainfall is estimated from the daily precipitation observed at 4 meteorological stations by using the classical Thiessen polygon method.

3 Methods for detecting climate change

Different statistical tools for assessing changes in extremes exist, and the community has not generally agreed to a "best" approach. In the present study, four non-parametric methods will be applied, including the Mann-Kendall trend test (MK-test), Kolmogorov–Smirnov distribution test (KS-test), Levene's variance homogeneity test (L-test) and quantile test (Q-test). Among these methods, the MK test is widely used in hydrology, the KS-test has been used in some studies, but the L-test and Q-test are not commonly used by the hydrometeorology community. In the present study, the MK test is applied to annual series of extreme indices whereas the KS, L and Q tests are applied to the distribution of daily precipitation amounts. Although the KS-test, L-test and Q-test could be applied also to the distributions of annual extreme indices, the small sample size (around 50 years in total, and only aournd 25 yr if the data are split into two parts) makes the results quite un-reliable. Therefore, the KS, L and Q tests are not applied to annual series. In addition, an exploratory graphical data analysis method, i.e., Quantile-Quantile plot, is used for graphically detecting changes between two samples. Descriptions of these methods are given in this section.

3.1 Mann-Kendall trend test

An important task in hydrological modeling is to determine if any trend exists in the data, not only for the purpose of modeling, because many models have assumptions of stationarity, but also for detecting possible links between hydrological processes and environmental changes. Many methods are available for detecting trends. Non-parametric trend detection methods are less sensitive to outliers (extremes) than are parametric statistics such as Pearson's correlation coefficient. In addition, nonparametric tests can test for a trend in a time series without specifying whether the trend is linear or nonlinear. Therefore, the Mann-Kendall's test (Kendall, 1938; Mann, 1945), referred to as MK test hereafter, which is a rank-based nonparametric method, is applied in this study.

Under the null hypothesis H_0 that a series $x_1, \ldots, x_N$ come from a population where the random variables are independent and identically distributed, the MK test statistic is $S=\sum_{i=1}^{N-1}\sum_{j=i+1}^{N} sgn(\mathrm{x}_j-\mathrm{x}_i)$, where

$$sgn(x_j-x_i)=\begin{cases} +1, & x_j>x_i \\ 0, & x_j=x_i \\ -1, & x_j<x_i \end{cases} \tag{1}$$

And Kendall's τ, which measures the strength of the monotonic trend, is estimated by:

$$\tau=\frac{2S}{N(N-1)}. \tag{2}$$

Table 2. Extreme precipitation indices used in this study.

Index	Description	Unit
CDD	Annual maximum number of consecutive dry days with RR<1 mm	Days
CWD	Annual maximum number of consecutive wet days with RR≥1 mm	Days
R20mm	Annual count of days when RR≥20 mm	Days
R50mm	Annual number of days when RR≥50 mm	Days
RX1day	Annual maximum precipitation in 1 d	mm
RX5day	Annual maximum precipitation in 5 consecutive days	mm
PRCPTOT	Annual total precipitation from wet days (RR≥1 mm)	mm
SDII	Simple pricipitation intensity index, average daily precipitation amount on wet days with RR≥1 mm	mm/day

Note: RR denotes daily precipitation amount.

Table 3. Mann-Kendall trend tests on annual precipitation series.

Annual series	MK test	59096	59102	59293	59298	Areal
RX1day	τ	0.0287	−0.1108	−0.0309	−0.0136	−0.048
	p-value	0.7703	0.2704	0.7523	0.8933	0.635
RX5day	τ	0.0747	−0.0940	−0.0641	−0.0762	−0.146
	p-value	0.4393	0.3507	0.5074	0.4300	0.140
CDD	τ	−0.0706	0.0195	0.0157	−0.0847	−0.045
	p-value	0.4690	0.8516	0.8773	0.3842	0.656
CWD	τ	−0.1704	−0.2961	−0.1342	−0.0694	−0.167
	p-value	0.0739	**0.0026**	0.1580	0.4666	0.091
PRCPTOT	τ	−0.0136	−0.0656	0.0271	−0.0106	−0.021
	p-value	0.8933	0.5165	0.7824	0.9183	0.836
R 20 mm	τ	−0.0890	−0.0745	0.0920	0.1109	−0.052
	p-value	0.3548	0.4593	0.3387	0.2479	0.604
R50mm	τ	0.0837	−0.1507	0.0845	−0.0739	−0.050
	p-value	0.3813	0.1267	0.3785	0.4387	0.614
SDII	τ	0.0543	−0.0895	0.0551	0.0173	−0.020
	p-value	0.5749	0.3741	0.5698	0.8621	0.849

Kendall (1975) showed that the variance of S, $Var(S)$, for the situation where there may be ties (i.e., equal values) in the x values, is given by

$$\sigma_S^2=\frac{1}{18}\left[N(N-1)(2N+5)-\sum_{i=1}^{m}t_i(t_i-1)(2t_i+5)\right], \quad (3)$$

where, m is the number of tied groups in the data set and t_i is the number of data points in the ith tied group.

Under the null hypothesis, the quantity z defined in the following equation is approximately standard normally distributed:

$$z=\begin{cases}(S-1)/\sigma_s & \text{if } S>0\\ 0 & \text{if } S=0\\ (S+1)/\sigma_s & \text{if } S<0\end{cases} \quad (4)$$

At a 0.05 significance level, the null hypothesis of no trend is rejected if $|z|>1.96$.

3.2 Kolmogorov–Smirnov test

The two-sample Kolmogorov–Smirnov (KS) test is one of the most useful and general nonparametric methods for comparing two samples to determine whether they follow the same distribution. The KS-test is a distribution-free test, which is based on looking at the maximum vertical distance between the empirical distribution functions of two samples. Let n_1 and n_2 be the sizes of the two samples, $n_1\geq n_2$. The value of the test statistic for the two-sided two-sample Kolmogorov–Smirnov test is

$$T=\sup_x |F_1(x)-F_2(x)| \quad (5)$$

where F_1 and F_2 are the empirical distribution functions based on the two samples. The asymptotic p value for this statistic as $n_1, n_2\to\infty$ is given by

$$p=Q\left(T\sqrt{\frac{n_1n_2}{n_1+n_2}}\right) \quad (6)$$

Table 4. Mann-Kendall trend tests on annual discharge series.

Series	MK test	Longchuan	Heyuan	Boluo	Jiuzhou	Lantang	Yuecheng
Annual average discharge	τ	0.0573	0.0905	0.0748	−0.0361	−0.0559	−0.0667
	p-value	0.5587	0.3477	0.4533	0.7274	0.5815	0.5196
Annual maximum	τ	−0.3780	−0.4472	−0.2449	−0.1203	−0.1605	−0.2271
	p-value	**0.0001**	**0.0000**	**0.0133**	0.2368	0.1096	**0.0267**
Annual 7-day minimum	τ	0.0918	0.3394	0.4405	0.2081	−0.1118	0.3527
	p-value	0.3461	**0.0004**	**0.0000**	**0.0400**	0.2553	**0.0006**
Timing of annual maximum	τ	−0.0243	−0.0400	0.1216	−0.0130	−0.0399	0.1092
	p-value	0.8074	0.6815	0.2207	0.9051	0.6957	0.2888
Timing of annual 7-day minimum	τ	−0.1671	−0.3167	−0.3010	−0.0740	0.0743	−0.1691
	p-value	0.0850	**0.0009**	**0.0023**	0.4686	0.4515	0.0993

where $Q(z)=2\sum_{k=1}^{\infty}(-1)^{k-1}e^{-2k^2z^2}$. Because the above series converges rapidly, $Q(z)$ can be approximated using $Q(z)\approx 2e^{-2z^2}$, or for even greater accuracy, using $Q(z)\approx 2(e^{-2z^2}-e^{-8z^2})$ (Greenwella and Finchb, 2004). Massey (1951) calculated the exact value of p as an alternative to the use of a symptotic formula given above when the two sample sizes are equal. Kim and Jennrich (1973) developed a more general algorithm for any two sample sizes, and created tables for various values of n_1 and n_2.

3.3 Test for homogeneity of variance

The KS-test is designed to detect a shift in the whole distribution of group 1 relative to the distribution of group 2, and it tends to be more sensitive near the center of the distribution than at the tails (Filliben and Heckert, 2006), whereas when detecting changes in extreme events, we are very interested in the variance and the tails of the data, because the variance difference and tail fatness may indicate the difference of the occurrence of extreme events. Therefore, in addition to the KS-test, we apply Levene's test, a test for the homogeneity of variances between different groups, and the quantile test, a test for the shift of the upper tail.

The F-test is widely used to test if the standard deviations of two populations are equal. But the F-test is extremely sensitive to the normality assumption. This is also the case with another commonly used test method, Bartlett's test (Bartlett, 1937), while precipitation data almost always violate the normality assumption. Thus, in the present study, we use Levene's test (referred to as L-test hereafter), which is less sensitive than the Bartlett test to departures from normality (Conover et al., 1981; Snedecor and Cochran, 1989, p. 252), to detect whether the variances of k groups are identical.

The L-test is based on computing absolute deviations from the group mean within each group. Given a variable Y with sample of size N divided into k subgroups, the L-test statistic is defined as:

$$W=\frac{(N-k)}{(K-1)}\frac{\sum_{i=1}^{k}N_i\left(\bar{Z}_i-\bar{Z}\right)^2}{\sum_{i=1}^{k}\sum_{j=1}^{N_i}\left(z_{ij}-\bar{Z}_i\right)^2} \tag{7}$$

where N_i is the sample size of the ith subgroup; the within-group absolute deviations $z_{ij}=|x_{ij}-\bar{x}_i|$, $i=1,2,\ldots,k$, $j=1,2,\ldots,N_i$, $\bar{x}_i$ is the mean of the ith sub-group; $\bar{Z}_i$ is the group mean of z_{ij}; and $\bar{Z}$ is the overall mean of the z_{ij}.

The L-test rejects the hypothesis of equal variances if

$$W>F(\alpha,k-1,N-k) \tag{8}$$

where $F(\alpha,k-1,N-k)$ is the upper critical value of the F distribution with k- 1 and $N-k$degrees of freedom at a significance level of α.

3.4 Quantile test

When detecting the changes in extreme events, we are also interested in detecting the difference between two distributions where only a portion (especially the lower tail or upper tail) of the distribution of group 1 is shifted relative to the distribution of group 2. The quantile test (referred to as Q-test hereafter) is a two-sample rank test to detect such a shift (Johnson et al., 1987) based on permuting the ranks of the observations in the tail.

Under the null hypothesis, cdfs of group 1 and 2 are the same. If the alternative hypothesis is that the distribution of group 1 is partially shifted to the right of the distribution of group 2, the test combines the observations, ranks them, and computes k, which is the number of observations from group 1 out of the $r$largest observations. The test rejects the null hypothesis if $k$is too large. The p-value is computed as

$$p=\sum_{i=k}^{r}\binom{N-r}{n_1-i}\binom{r}{i}\Big/\binom{N}{n_1} \tag{9}$$

Table 5. P-values for changes in statistical properties of daily rainfall in each month for the periods before and after 1979.

Station	Test	Jan	Feb	March	April	May	June	July	Aug	Sep	Oct	Nov	Dec
59096	KS-test	0.354	0.168	0.680	0.792	0.658	0.128	0.092	0.196	**0.042**	0.076	0.262	0.377
	L-test	0.130	**0.005**	**0.005**	**0.039**	0.446	0.142	0.215	0.365	**0.001**	**0.001**	0.833	0.996
	Q-test	0.121	**0.008**	0.100	**0.010**	0.790	0.770	0.464	0.164	**0.027**	0.989	0.882	0.434
59102	KS-test	0.338	0.126	0.377	0.830	0.059	**0.027**	0.591	0.992	0.426	**0.024**	0.905	0.331
	L-test	0.160	0.251	**0.001**	0.059	0.298	**0.009**	0.135	0.694	0.601	0.155	0.474	0.181
	Q-test	0.131	0.623	0.086	0.230	0.557	0.895	0.446	0.086	0.561	0.916	0.442	0.744
59293	KS-test	0.988	0.963	0.359	0.916	0.608	0.644	0.169	0.167	0.198	0.336	0.723	0.718
	L-test	0.142	0.535	**0.000**	0.538	0.795	0.160	0.011	**0.000**	0.166	**0.037**	0.280	0.950
	Q-test	0.225	0.136	**0.002**	0.286	0.511	0.495	0.171	**0.012**	0.231	0.960	0.271	0.361
59298	KS-test	0.558	0.155	0.332	0.767	0.692	0.999	0.334	0.373	0.169	0.679	0.155	0.373
	L-test	0.422	**0.033**	0.090	0.141	0.330	0.458	0.964	0.098	0.862	**0.037**	0.202	0.968
	Q-test	0.558	0.185	0.148	0.290	0.946	0.722	0.596	0.431	0.616	0.973	0.938	0.211
Areal	KS-test	0.999	0.677	0.274	0.265	0.781	0.074	0.530	**0.010**	0.813	0.130	0.108	0.854
	L-test	0.118	0.098	**0.000**	0.416	0.647	**0.033**	**0.024**	**0.041**	0.497	**0.000**	0.827	0.717
	Q-test	0.521	0.302	**0.022**	0.361	0.802	0.915	0.081	0.019	0.192	1.000	0.716	0.405

Note: Significance level=0.05. The alternative hypothesis of Q-test is that the distribution of data after 1980 is partially shifted to the right of the distribution of data before 1979, and the target quantile is q=0.95.

where n_1 and n_2 are the size of group 1 and group 2, and $N=n_1+n_2$. The value of r is the smallest rank determined by $r/(N+1) > q$, where q is the target quantile.

3.5 Quantile-Quantile (Q-Q) plot

Drawing a Quantile-Quantile (Q-Q) plot is a commonly used technique for checking if the distributions of two data sets are different. The Q-Q plot is a plot of the quantiles of the first data set against the quantiles of the second data set. If the two sets come from a population with the same distribution, the points should fall approximately along a 45-degree reference line. The greater the departure from this reference line, the greater the evidence for the conclusion that the two data sets have come from populations with different distributions.

4 Results for the precipitation processes and streamflow processes in Dongjiang River Basin

4.1 Extreme hydro-meteorological indices

Considerable efforts have been put on defining indices for evaluating changes in extreme climate. For instance, Karl et al. (1996) proposed a Climate Extremes Index (CEI) based on an aggregate set of conventional climate indicators which, after two notable modifications in 2003 (www.ncdc.noaa.gov/oa/climate/research/cei/cei.html), include the following types of data: 1) monthly maximum and minimum temperature; 2) daily precipitation; 3) monthly Palmer Drought Severity Index (PDSI); 4) landfalling tropical storm and hurricane wind velocity. The Expert Team on Climate Change Detection, Monitoring and Indices (ETCCDMI), which was jointly established by the WMO Commission for Climatology and the Research Programme on Climate Variability and Predictability (CLIVAR), developed 27 climate change indices (Peterson et al., 2001), many of which are widely used in evaluating extreme temperature and precipitation in many studies for Middle East, central Asia, etc. (e.g., Zhang et al., 2005b; Klein Tank et al., 2006; Alexander et al., 2006). Similar definitions for extreme climate events are also seen in many other studies (e.g., Nicholls et al., 2000; Frichet al., 2002; STARDEX Project, 2005). In the EMULATE (European and North Atlantic daily to multi-decadal climate variability) project more detailed 64 climate indices are defined (Moberg et al., 2006).

In the present study, 8 indices defined by ETCCDMI are used, as listed in Table 2. RClimDex, which is developed at the Climate Research Branch of Meteorological Service of Canada, and available from the ETCCDMI Web site (http://cccma.seos.uvic.ca/ETCCDMI), was used for calculating these indices except for CDD. Because RClimDex calculates all indices based on calendar year without considering actual seasonality, which is not suitable for calculating CDD for cases where the dry season spans two years, CDD was calculated separately based on a hydrological year starting from 1 October and ending on 30 September.

The analysis of extreme flood flow events can be conducted with the annual maximum flood (AMF) approach, or the peaks-over-threshold (POT) approach, also called partial duration series approach (PDS) (see Lang et al., 1999). An AMF sample is constructed by extracting from a series of flows the maximum value of each year (annual flood), i.e. only one event per year is retained. Due to its simplicity, the AMF approach is adopted in the present study for analyzing extreme flood events.

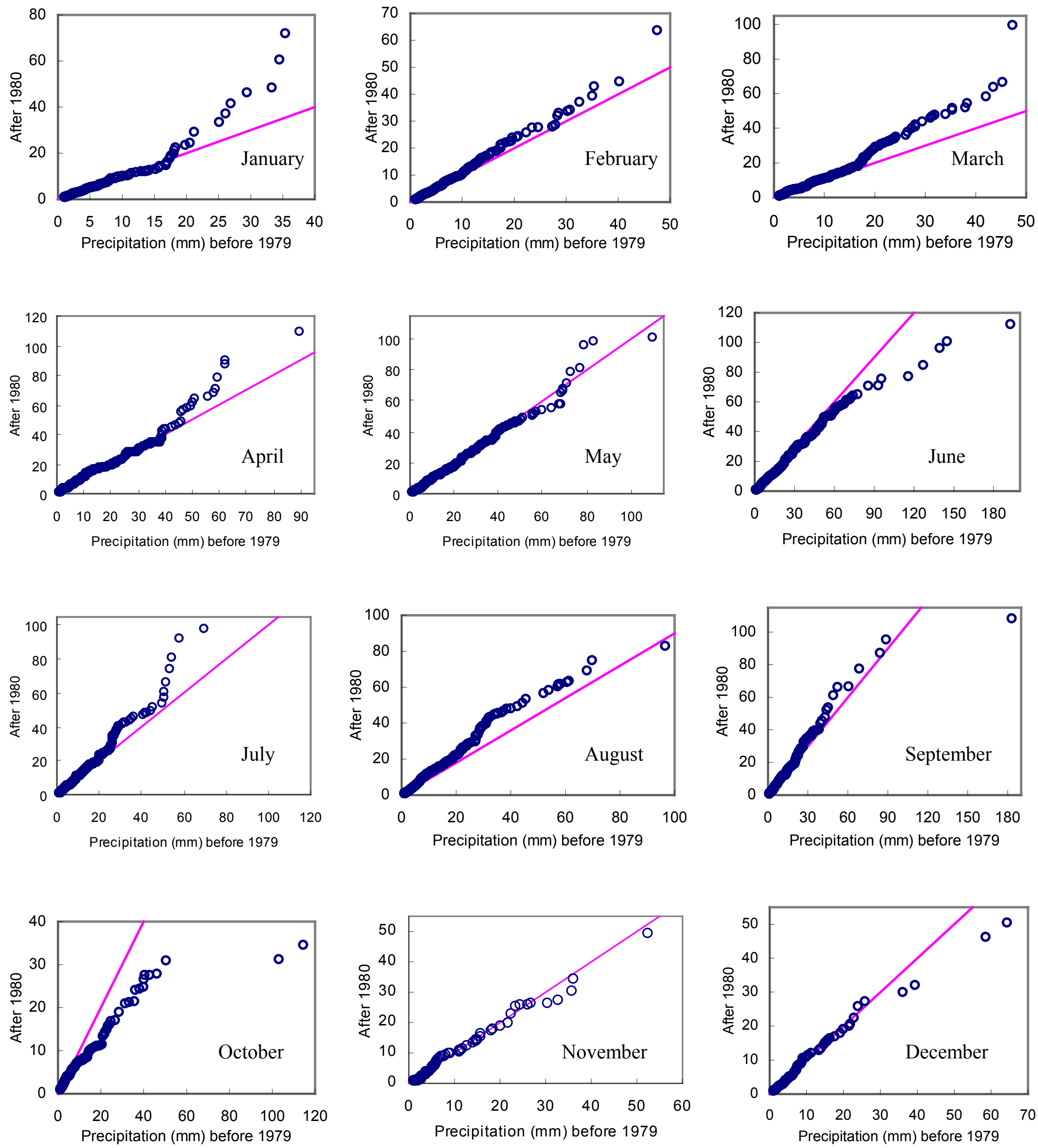

Fig. 2. Q-Q plots for the daily basin average precipitation for each month for periods before 31 December 1979 and after 1 January 1980.

In the minimum low-flow analysis, the minimum 7-day low flow is used. The 7-day low-flow index was chosen for three reasons (Chen et al., 2006): (a) The 7-day low-flow is the most widely used index in the USA, UK and many other countries; (b) Previous studies have shown that, compared with 1-day low flow, an analysis based on a time series of 7-day average flows is less sensitive to measurement errors; (c) Since Dongjiang River Basin is dominated by a humid sub-tropical monsoon climate, the 7-day low flow better represents the drought conditions of concern and can be used more effectively in water management.

In addition, the timing of annual maximum daily discharge and minimum 7-day average discharge is analyzed.

4.2 Trend test for annual hydro-meteorological series

The MK test was applied to all the annual precipitation and streamflow series, including annual total/average series and annual extremal series.

It has been found that the positive serial correlation inflates the variance of the MK statistic S and hence increases the possibility of rejecting the null hypothesis of no trend (von Storch, 1995). In order to reduce the impact of serial correlations, it is common to prewhiten the time series by removing the serial correlation from the series through $y_t = x_t\text{-}\phi x_{t-1}$, where y_t is the prewhitened series value, x_t is the original time series value, and ϕ is the estimated serial correlation coefficient at lag one. However, in our case, none of the data series for detecting trend has significant serial correlation at a 5% level, except the minimum 7-day low-flow series at Boluo with a lag-one autocorrelation coefficient of 0.599. Therefore, prewhitening is not applied in this study. The results of the trend test are listed in Table 3 and Table 4. From Tables 3 and 4 we see that:

1. There is no significant change in either annual total precipitation (PRCPTOT) or annual average discharge.

2. It is shown that no trend is present in annual extreme precipitation series in general at a 0.05 significance level, except consecutive wet days (CWD) at station 59102.

3. Significant trends are detected in several annual streamflow processes, including: annual daily maximum flow series at three stations (Longchuan, Heyuan, and Boluo) along the main channel and one station (Yuecheng) along a tributary which exhibit significant negative trends; annual 7-day minimum flow series at two stations (Heyuan and Boluo) along the main channel and another two stations (Jiuzhou and Yuecheng) along tributaries which exhibit significant positive trends. In addition, the timing of the occurrence of low-flow at the two stations (Heyuan and Boluo) along the main channel gets significantly earlier. For the stations along the main channel, the changes can be explained by the regulation of three major reservoirs. The reason of significant changes in the extreme flows at Yuecheng and Jiuzhou may be a combined effect of land-use/land-cover change and the impacts of small reservoirs.

4.3 Testing changes in precipitation for the periods before and after 1979

As shown by the trend test for various annual indices in Sect. 4.2, no significant trend is present in the annual extreme precipitation series when taking the period from 1950s to early 2000s as a whole. However, it is possible that significant changes occurred in different seasons. On the other hand, it has been found that the climate in China experienced a significant decadal change in the late 1970s (Wang, 1994), which is related to the abrupt change in the large-scale boreal winter circulation pattern over the North Pacific during the late 1970s (Graham, 1994). Dyurgerov and Meier (2000) showed that the time series of change in global glacier volume suggest a significant shift during the late 1970s. Yu and Lin (2002) showed that there is significant difference before and after the late 1970s in terms of the Northern Hemisphere sea level pressure, 500 hPa height and North Pacific sea surface temperature, and such a jump affected the climate of China significantly. Gong and Ho (2002) noticed a significant regime shift in the summer rainfall over the whole eastern China in about 1979. The existence of such a climate shift is also shown in many other research results (e.g., Xu et al., 2005Li et al., 2006). Therefore, we will investigate evidences of changes in daily precipitation on a monthly basis in two periods, i.e., the period before 31 December 1979, and another after 1 January 1980.

Table 6. Quantiles of Gamma distributions with different values of β.

Probability	β=10	β=15	β=20	β=30	β=40
0.99	33.2	49.8	66.3	99.5	132.7
0.999	54.1	81.2	108.3	162.4	216.6

First of all, we draw Q-Q plots for each month in the period 1956–1979 versus 1980–2004 to examine if there is any graphically obvious change. To save space, only the Q-Q plots for the daily basin average precipitation are shown here in Fig. 2.

Q-Q plots give us graphical evidences indicating significant changes in the upper part of the probability distribution in many months, e.g., increase in heavy rainfall in January, February, March, April and July, decrease in heavy rainfall in June and October. Q-Q plots also indicate changes in many months for the precipitation observed at all the 4 meteorological stations, but the results are not in good agreement with each other.

To verify the heuristic results from Q-Q plots, three quantitative statistical tests, i.e., the KS-test, L-test and Q-test (for the upper tail with q=0.95) are applied to the observed data sets of the two periods. The calculation is conducted with the software package *EnvironmentalStats for S-PLUS* (Millard, and Neerchal, 2001), which is an add-on module to the statistical software package S-PLUS. The results are reported in Table 5.

According to the KS-test results, only the overall distribution of the rainfall in September at station 59096 and October at 59102 changed significantly. There are significant changes in variances for several months at each station, but the months are in not in good agreement among the stations. And, the rightward shift of the upper tail is detected in several

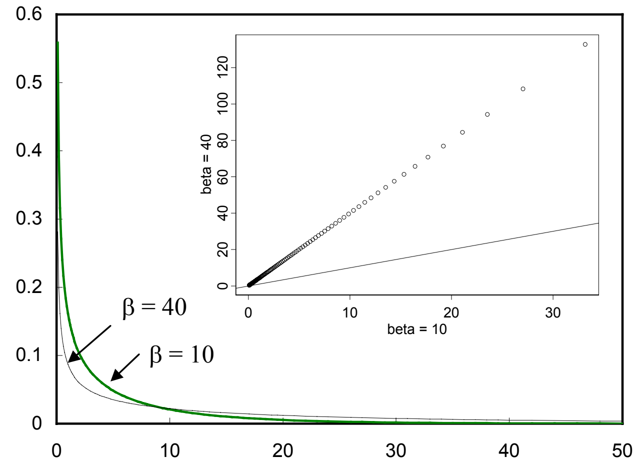

Fig. 3. Two Gamma densities with scale parameter β=10 and 40, and shape parameter α=0.5. (Note: The embedded figure is the Q-Q plot for the two Gamma distributions).

months at two stations (59096 and 59293), but the months are not in good agreement either between the two stations. As for the mean areal precipitation, significant changes in variance are detected for rainfall in March, June, July, August and October, but significant right-ward shift of upper tails is only found in March.

The quantitative statistical test results seem to be more or less different from what we see from the Q-Q plots. For instance, while the Q-Q plot for areal precipitation in October shown in Fig. 2 indicates significant change in the distribution, the test result in Table 5 indicates no change by the KS-test. Are the statistical test results reliable? We therefore make a Monte Carlo simulation to evaluate the effectiveness of these test methods.

4.4 Evaluation of test methods for detecting changes in precipitation

To evaluate the methods for detecting changes in precipitation, we only consider Gamma distributed variables because daily rainfall processes are normally considered to follow a Gamma distribution (e.g., Groisman et al., 1999; Liao et al., 2004) with a probability density function in the form of

$$f(x)=\frac{1}{\Gamma(\alpha)\beta^{\alpha}}x^{\alpha-1}e^{-x/\beta} \quad (10)$$

where α is the shape parameter and β the scale parameter. Groisman et al. (1999) estimated that the scale parameter β changes by an order of magnitude from subarctic regions and deserts (1/0.3) to humid tropics ($\sim$ 1/0.03), and the shape parameter α has little spatial variation, which may vary from 0.5 up to 1.2. Liao et al. (2004) showed that for rainfall data in most areas of China, $\alpha \in$(0.3, 0.5), $\beta \in$(2, 40). Therefore, in our simulation experiment, we concentrate on α=0.5 and β=10$\sim$40. The plots of distribution functions with α=0.5 and β=10, 40 are shown in Fig. 3, and the 0.99 and 0.999 quantiles for each distribution are listed in Table 6. Obviously, the larger the value of β, the more extreme the distribution, and the quantile corresponding to a given probability increases in a rate equal to the rate of increase in the scale parameter.

Now we investigate the robustness and power of the three test methods used in our study in detecting the changes when the variable changes from a distribution with a low β value to a distribution with a higher β value, namely, a more extreme distribution. By robustness, we mean the ability of the test to not falsely detect changes when the underlying data are in fact distributed equally. By power, we mean the ability of the test to detect changes when the distribution indeed changes.

We make 10 000 simulations of Gamma distributed samples with fixed value of α=0.5, varied values of β=10, 15, 20, 30, 40, and varied length of sample size L=50, 100, 200, and 300. In each simulation two groups of data are generated, with one group of length L generated with low value of β=b1, and another group of the same length L but equal or higher value of β=b2 (b2$\geq$b1). That is, the second group of dataset has a equal or larger variance, because it is known that the variance of Gamma distributed variables is given by $\alpha\beta^2$. The Monte Carlo simulations are repeated for b1=10 and 20, b2=10, 15, 20, 30 and 40, and L=50, 100, 200, 300. The results are reported in Table 7. The simulation results in Table 7 show that:

1. While the robustness of the tests has little dependence on data size, the power of all the tests is closely related to the sample size and depends on the magnitude of change (in terms of variance ratio).

2. The KS-test and especially the Q-test are quite robust, with a wrong rejection rate less than 0.05 at a 5% significance level mostly. But the L-test is not robust, with a wrong rejection rate of around 18% at 5% significance level for all cases where variance ratio=1. Therefore, a rejection by the L-test alone does not give a reliable evidence of change, whereas a rejection by the KS-test or Q-test is a good evidence of the presence of change.

3. In case of a sharp 100% increase of the scale parameter β changing from 10 to 20 or 20 to 40, while it is not possible to detect the change between two groups of data set with 50 points each, generally, all the three tests are powerful enough to correctly detect the change for large data sets (200 or 300 points) (over 98% correct rejection of the null hypothesis). But for smaller changes, such as a 50% increase of changing from 20 to 30, the rate of correct rejection of the null hypothesis is low, even for large data sets (e.g., 70.31% for data set with 300 points). Unfortunately, in real world, we often lack enough data, and as shown by Groisman et al. (1999), the change in scale parameter g is normally less than 50%, and seldom over 100% for most areas ug

Table 7. Rejection rate (in percentage) of null hypothesis for testing changes in two groups of data with three test methods.

Size of each group	B1	B2	Variance ratio	KS-test	L-test	Q-test (q=0.95)
50	10	10	1	4.08	18.51	3.16
		15	2.25	13.93	48.88	18.65
		20	4	36.55	80.77	42.52
		30	9	76.64	98.21	77.96
		40	16	93.01	99.86	92.82
	20	20	1	3.96	17.86	2.82
		30	2.25	13.9	49.17	18.71
		40	4	36.95	81.15	42.79
100	10	10	1	3.78	17.58	0.85
		15	2.25	25.05	69.22	17.7
		20	4	65.73	96.48	53.97
		30	9	97.38	99.99	92.77
		40	16	99.9	100	99.19
	20	20	1	3.59	17.98	0.82
		30	2.25	24.65	69.6	18.5
		40	4	65.7	96.55	53.53
200	10	10	1	5.4	17.51	1.95
		15	2.25	54.8	90.32	50.57
		20	4	95.85	99.9	92.58
		30	9	100	100	99.97
		40	16	100	100	100
	20	20	1	5.44	17.43	1.68
		30	2.25	55.66	90.37	50.83
		40	4	95.71	99.9	92.56
300	10	10	1	5.12	17.36	4.39
		15	2.25	73.75	96.82	81.39
		20	4	99.47	100	99.73
		30	9	100	100	100
		40	16	100	100	100
	20	20	1	5.12	17.97	4.52
		30	2.25	72.9	96.9	81.75
		40	4	99.52	100	99.62

Note:The null hypotheses of all three methods are no change. Significance level 0.05. B1 and B2 are respectively the scale parameters of the first and second group of simulated gamma distributed data with the same shape parameter 0.5. The variance ratio is the ratio of the variance of the second dataset over the first dataset.

From the above analysis, we know that, the good news is, if the null hypothesis of no change is rejected by the KS-test and Q-test, it is a good indication of change, whereas the bad news is, if the null hypothesis is accepted, we are still not sure if or not there is any significant change present because non of the three methods are powerful enough for detecting small-scale changes (such as a 50% increase of β) even for a large dataset. But the power increases with the increase in the data size. For two datasets with 300 points each and a 100% increase of β, all three methods can surely detect the changes. By revisiting the analysis in Sect. 4.3, we see that the rainy days for the months from March to September are mostly over 300, while for the months from October to February over 90. Thus the test results for months from March to September should be reliable if sharp changes occur, whereas less reliable for the months from October to February. From Table 5 we know that, changes indeed occur in several months in the Dongjiang Basin, but the changes are not uniform at different locations in the region. On the other hand, because all the test methods are not powerful enough for moderate changes, such as a 50% change in scale parameter , we cannot conclude that no change occurs in other months, which seem to have apparently experienced change according to the Q-Q plots in Fig. 2. Consequently, we suggest combining the use of the Q-Q plot method and statistical test methods to detect changes in extreme events, but the combined use of these methods is still not conclusive.

Table 8. Estimates of Gamma distribution parameters for simulated Gamma distributed samples with various sample sizes (the mean values and standard deviations are calculated based on 5000 simulations.)

L	α		β	
	mean	standard deviation	mean	standard deviation
50	0.5229	0.0913	21.8063	5.8795
100	0.5116	0.0614	20.8684	3.9404
200	0.5052	0.0416	20.4227	2.6423
300	0.5043	0.0339	20.3220	2.1565

5 Discussions and conclusions

(1) For the region under consideration in the present study, little change is observed in various annual extreme precipitation indices, but significant changes are observed in the precipitation processes on a monthly basis, although the seasonal variations are not uniform even in a medium-sized basin such as the Dongjiang River Basin. This is probably because extreme events at a specific location depend not just on the moisture availability and thermodynamic instability, but also on other factors, primarily the frequency and intensity of precipitation-producing meteorological systems (Kunkel, 2003), and the activeness of these systems varies seasonally. To get statistically significant results in detecting changes, we need more robust statistical test methods, and may need some indices that take the changes in seasonality into account as well. In fact, seasonality has been considered in the calculation of the 64 climate indices in the EMULATE project (Moberg et al., 2006).

(2) Despite of little change in extreme precipitation, significant changes are detected at all the three stations along the main river channel, i.e., Longchuan, Heyuan and Boluo. All of the three show significant negative trends in the annual maximum flow, and two of them (Heyuan and Boluo) exhibit significant positive trend in minimum 7-day low-flow. Among three streamflow series observed at tributary stations with medium-size drainage areas and no intervention by major reservoirs, one (Yuecheng) shows significant negative trend in annual maximum flows, and two (Jiuzhou and Yuecheng) show significant positive trend in minimum 7-day low-flows. The changes in annual extremal streamflows at the three stations along the main river channel are obviously due to the operation of several major reservoirs in the basin, whereas the changes at tributary stations are possibly due to land use change and/or operation of small reservoirs. The results indicate that, in the case of little precipitation changes, the operation of major reservoirs is most influential on the extreme streamflow events, whereas land-use/land-cover changes may have secondary impacts. It is common in many studies to examine if extreme high or low flows are associated with climate change or land-use/land-cover change (e.g., Tu et al., 2005; George, 2007; de Wit et al., 2007). But when there are major reservoirs present, in assessing the impacts of environmental changes on streamflow processes, especially flood events, how the reservoirs are operated should be considered first.

(3) It is expected that with the global warming, the water cycling will be sped up, and the enhanced water cycling will result in an increase of evaporation and precipitation. But in fact, in many regions of the world the observed pan evaporation is decreasing (e.g., Peterson and Groisman, 1995; Chattopadhyay and Hulme, 1997; Roderick and Farquhar, 2002; Liu et al., 2004), which is considered as a "paradox". A significant decreasing trend is also observed in the pan evaporation processes in most parts (including the present study area) of China (Liu et al., 2004; Ren and Guo, 2006). It has been demonstrated by some researchers that the actual evaporation is negatively related to pan evaporation (Brutsaert and Parlange, 1998; Lawrimore and Peterson, 2000; Golubev et al., 2001). Whether the actual evaporation has increased with the decrease of pan evaporation for the case of China, specifically for the case of Dongjiang River Basin, is an open question. If it is true, still we have a problem that, with increased evaporation, no significant change is detected in annual total precipitation and annual runoff, and the amplitude of extreme precipitation has not changed much either. Runoff may be affected by the changes of water abstraction for industry and agriculture use (especially irrigation), and the increase/decrease of forest coverage which leads to increased/decreased plant transpiration, because the establishment of forest cover would result in increased transpiration and therefore decreases water yield (e.g., Bosch and Hewlett, 1982). Therefore, how the land cover has changed and how human activities affect the streamflow process in this area will be the subjects of a future study.

(4) In detecting changes in extreme hydro-meteorological events, two approaches are commonly seen in literature, i.e., testing trend in annual extremal series for the entire period under consideration, and comparing probability distribution parameters for data observed during different periods (e.g., Tromel and Schonwiese, 2007). The former approach is typically done with the nonparametric Mann-Kendall test (e.g., Karl and Knight, 1998; Kundzewicz et al., 2005), whereas the later approach is a typical parametric one. In some cases, the two approaches are used in combination (e.g., Osborn and Hulme, 2002). However, by using trend tests, we focus on the change in mean values and cannot find changes in overall statistical properties, while by comparing distribution parameters, the result is subject to parameter estimation uncertainties and statistical tests are not applicable for testing the significance of differences among estimated parameters based on two groups of dataset. To examine the robustness of the parametric method, we can make a simple experiment by generating 5000 simulations for Gamma distributed

samples with α=0.5 and β=20, which are commonly seen in the case of precipitation modeling, with varied length of data size L=50, 100, 200 and 300. Then, we use the method of maximum-likelihood to estimate the shape parameter α and scale parameter β. The results are shown in Table 8, from which we see that the parametric method is not particularly powerful. For instance, with a dataset of 100 points, the 95% confidence intervals for the estimates of β is (13.15, 28.59), which means that estimation uncertainty may cause a over 100% increase of β, whereas in the real world, the estimated change of β is rarely over 100% (e.g., see the analysis of Osborn and Hulme (2002) for the precipitation statistics in the UK).

Therefore, in the present study, besides the Mann-Kendall trend test which has been widely used in the hydrology community, three other non-parametric methods, i.e., Kolmogorov–Smirnov test, Levene's test and quantile test, are applied to test for changes in the distribution, variance and the shift of tails of different groups of data. While all three methods work well for detecting changes in two groups of data with large data size (e.g., over 200 points in each group) and large difference in distribution parameters (e.g., over 100% increase of the scale parameter in the Gamma distribution), none of them are powerful enough for small data sets (e.g., less than 100) and small distribution parameter difference (e.g., 50% increase of the scale parameter in the Gamma distribution). Unfortunately, small dataset sizes and small distribution parameter changes are common in real world applications. Therefore, neither parametric methods nor non-parametric methods is particularly powerful in detecting changes in extreme hydro-meteorological events, and the combined use of graphical exploratory methods, such as Quantile-Qantile plots, and quantitative statistical test methods is recommended.

(5) Caution must be taken when prewhitening a series before conducting Mann-Kendall trend test, because removal of autocorrelation with AR(1) model from time series by prewhitening will remove a portion of trend and hence reduces the possibility of rejecting the null hypothesis while it might be false (Yue and Wang, 2002); on the other hand, when the change in a real-world process has its physical background, the detected trend cannot be ignored even if it is possibly resulted from a significant serial correlation. For instance, in the case of minimum 7-day low-flow series at Jiuzhou, there is a weak autocorrelation coefficient of 0.223 at lag one which is not significant at a 0.05 significance level. When the series is not prewhitened, a positive trend could be detected at a 0.05 significance level, but no trend would be detected after prewhitening. Similar is the case of annual maximum flow at Yuecheng. In another case of minimum 7-day low-flow series at Boluo, the autocorrelation coefficient at lag one is 0.599. If the series is prewhitened, the positive trend is not significant at a 5% level. But the positive trend, we believe, has its physical basis because three major reservoirs, whose major effects are lowering peak flows and increasing low flows, were built in the end of 1950s', the beginning of 1970s' and early 1980s', which regulated streamflow significantly. Therefore, when there is a sound physical basis for the changes in a natural process, we suggest that the original series, rather than the prewhitened series, should be used for detecting the trend.

(6) Before fitting distribution models to a sample precipitation or streamflow data series, it would be wise to investigate the stationarity first, not only the trend in mean value but also the behaviour in variance and even higher moments. The regulation of reservoir outflows and impacts of land-use/land-cover changes have made many streamflow processes exhibit significant changes, which make the flood frequency analysis more tricky than for stationary cases. If no consideration is given to the nonstationary situations in the flood and low-flow frequency analysis (e.g., Chen et al., 2006), the results may be biased. Techniques of flood frequency analysis for nonstationary situations (see Khaliq et al., 2006) should be considered in future research.

Acknowledgements. The financial supports from the Research Fund for the Doctoral Program of Higher Education (No.20060294003), the National Science Foundation of China (Project No. 40771039) and the Key Project of Chinese Ministry of Education (No. 308012) are gratefully acknowledged. Many thanks to Editor H. Madsen for his careful editing which greatly improves the manuscript.

Edited by:H. Madsen

References

Bartlett, M. S.: Properties of sufficiency and statistical tests, P. Roy. Soc. Lond. Ser.-A, 160, 268–282, 1937.

Bielec, Z.: Long-term variability of thunderstorms and thunderstorm precipitation occurrence in Cracow, Poland, in the period 1896–1995, Atmos. Res., 56, 161–170, 2001.

Bosch, J. M., and Hewlett, J. D.: A review of catchment experiments to determine the effect of vegetation changes on water yield and evapotranspiration, J. Hydrol., 55, 3–23, 1982.

Brutsaert, W. and Parlange, M. B.: Hydrological cycle explains the evaporation paradox, Nature, 396, 30–31, 1998.

Chattopadhyay, N. and Hulme, M.: Evaporation and potential evapotranspiration in India under conditions of recent and future climate change, Agr. Forest Meteorol., 87, 55–73, 1997.

Chen Y.D., Huang G., Shao Q., and Xu C.Y.: Regional analysis of low flow using L-moments for Dongjiang basin, South China, Hydrolog. Sci. J., 51(5), 1051–1064, 2006.

Conover, W. J., Johnson, M. E., and Johnson, M. M.: A Comparative Study of Tests for Homogeneity of Variances, with Applications to the Outer Continental Shelf Bidding Data, Technometrics, 23(3), 351–361, 1981.

de Wit, M. J. M., van den Hurk, B., Warmerdam, P. M. M., Torfs, P. J. J. F., Roulin, E., and van Deursen, W. P. A.: Impact of climate change on low-flows in the river Meuse, Climatic Change, 82, 351–372, 2007.

Dyurgerov, M. B. and Meier, M. F.: Twentieth century climate change: Evidence from small glaciers, Proc. Natl. Am. Soc., 97, 1406–1411, doi:10.1073/pnas.97.4.1406, 2000.

Filliben, J. J. and Heckert, A.: Chapter 1: Exploratory Data Analysis, in: NIST/SEMATECH e-Handbook of Statistical Methods, available at: http://www.itl.nist.gov/div898/handbook/, 2006.

Fowler, A. M. and Hennessy, K. J.: Potential impacts of global warming on the frequency and magnitude of heavy precipitation, Natural Hazards, 11, 283–303, 1995.

Frich, P., Alexander, L. V., Della-Marta, P., Gleason, B., Haylock, M., Klein Tank, A. M. G., and Peterson, T.: Observed coherent changes in climatic extremes during the second half of the twentieth century, Climate Res., 19, 193–212, 2002.

George, S. S.: Streamflow in the Winnipeg River basin, Canada: Trends, extremes and climate linkages. J. Hydrol., 332(3–4), 396–411, 2007.

Golubev, V. S., Lawriore, J. H., Groisman, P. Y., et al.: Evaporation changes over the contiguous United States and the former USSR: A reassessment, Geophys. Res. Lett., 28, 2665–2668, 2001.

Gong, D. Y. and Ho, C. H.: Shift in the summer rainfall over the Yangtze River valley in the late 1970s. Geophys. Res. Lett., 29(10), 1436, 10.1029/2001GL014523, 2002.

Gordon, H. B., Whetton, P. H., Pittock, A. B., Fowler, A. M., and Haylock, M. R.: Simulated changes in daily rainfall intensity due to the enhanced greenhouse effect: implications for extreme rainfall events, Clim. Dynam., 8, 83–102, 1992.

Graham, N. E.: Decadal scale variability in the 1970s and 1980s: Observations and model results, Clim. Dynam., 10, 135–162, 1994.

Greenwella, R. N. and Finchb, S. J.: Randomized rejection procedure for the two-sample Kolmogorov–Smirnov statistic, Comput. Stat. Data An., 46, 257–267, 2004.

Griffiths, G. M., Salinger, M. J., and Leleu, I.: Trends in extreme daily rainfall across the South Pacific and relationship to the South Pacific convergence zone, Int. J. Climatol., 23, 847–869, 2003.

Groisman P. Y., Karl T. R., Easterling D. R., Knight R. W., et al.: Changes in the probability of heavy precipitation: important indicators of climatic change, Climatic Change, 42(1), 243–283, 1999.

Groisman, P. Y., Knight, R. W., and Karl, T. R.: Heavy precipitation and high streamflow in the contiguous United States: trends in the twentieth century, B. Am. Meteorol. Soc., 82, 219–246, 2001.

Groisman, P. Y, Knight, R. W., Easterling, D. R., Karl, T. R., et al.: Trends in intense precipitation in the climate record, J. Climate, 18(8), 1326–1349, 2005.

Haylock, M. and Nicholls, N.: Trends in extreme rainfall indices for an updated high quality data set for Australia, 1910–1998, Int. J. Climatol., 20, 1533–1541, 2000.

Hennessy, K. J., Gregor, J. M., and Mitchell, J. F. B.: Changes in daily precipitation under enhanced greenhouse conditions, Clim. Dynam., 13(8), 667–680, 1997.

Johnson, R. A., Verrill, S., and Moore, D. H.: Two-sample rank tests for detecting changes that occur in a small proportion of the treated population, Biometrics, 43, 641–655, 1987.

Karl, T. R., Knight, R. W., Easterling, D. R., and Quayle, R. G.: Indices of climate change for the United States, B. Am. Meteorol. Soc., 77(1), 279–292, 1996.

Karl, T. R. and Knight, R. W.: Secular trends of precipitation amount, frequency, and intensity in the USA, B. Am. Meteorol. Soc., 79, 231–241, 1998.

Karl, T. R., Knight, R. W., and Plummer, N.: Trends in high-frequency climate variability in the twentieth century, Nature, 377, 217–220, 1995.

Khaliq, M. N., Ouarda, T. B. M. J., Ondo, J. C., Gachon, P., and Bobee, B.: Frequency analysis of a sequence of dependent and/or non-stationary hydro-meteorological observations: A review, J. Hydrol., 329, 534–552, 2006.

Kim, P. J.: On the exact and approximate sampling distribution of the two sample Kolmogorov–Smirnov criterion D_{mn}, $m \leq n$, J. Am. Stat. Assoc., 64, 1625–1637, 1969.

Klein Tank, A. M. G., Peterson, T. C., Quadir, D. A., et al.: Changes in daily temperature and precipitation extremes in central and south Asia, J. Geophys. Res., 111, D16105, doi:10.1029/2005JD006316, 2006.

Koning, A., and Franses, P. H.: Are Precipitation Levels Getting Higher? Statistical Evidence for the Netherlands, J. Climate, 18(22), 4701–7414, 2005.

Kunkel, K. E.: North American trends in extreme precipitation, Nat. Hazards, 29, 291–305, 2003.

Kunkel, K. E., Andsager K., and Easterling D. R.: Long-Term Trends in Extreme Precipitation Events over the Conterminous United States and Canada, J. Climate, 12(7), 2515–2527, 1999.

Kundzewicz, Z. W., Graczyk, D., Maurer, T., Pinskwar, I., Radziejewski, M., Svensson, C., Szwed, M.: Trend detection in river flow series: 1. Annual maximum flow, Hydrol. Sci. J., 50(4), 797–810, 2005.

Lang, M., Ouardab, T. B. M. J., and Bobee, B.: Review: Towards operational guidelines for over-threshold modeling, J. Hydrol., 225, 103–117, 1999.

Lawrimore, J. H. and Peterson, T. C.: Pan evaporation trends in dry and humid regions of the United States, J. Hydrometeorology, 1(5), 543–546, 2000.

Levene, H.: Robust tests for the equality of variance. In Contributions to Probability and Statistics: Essays in Honor of Harold Hotelling, edited by Ghurye S. G., Hoeffding W., Madow, W. G., and Mann H. B., Stanford University Press, CA, 278–292, 1960.

Li, X. Z., Ma, Z. G., and Liu, X. D.: Inter-decadal characteristics of aridification over northern China associated with the variation of atmospheric circulation during the past 50 years, Chinese J. Atmos. Sci. (in Chinese), 30(1), 277–284, 2006.

Liao, Y., Zhang, Q., and Chen, D.: Precipitation simulations in China with a weather generator, Acta Geogr. Sinica (in Chinese), 59(4), 689–698, 2004.

Liu, B., Xu, M., Henderson, M., and Gong, A.: A spatial analysis of pan evaporation trends in China, 1955–2000, J. Geophys. Res., 109, D15102, doi:10.1029/2004JD004511, 2004.

Manton, M. J., Della-Marta, P. M., Haylock, M. R., Hennessy, K. J., Nicholls, N., et al.: Trends in extreme daily rainfall and temperature in Southeast Asia and the South Pacific: 1961–1998, Int. J. Climatol., 21, 269–284, 2001.

Massey, F.J.: The distribution of the maximum deviation between two sample cumulative step functions, Ann. Math. Stat. 22, 125–128, 1951.

McGuffie, K., Henderson-Sellers, A., Holbrook, N., Kothavala, Z., Balachova, O., and Hoekstra, J.: Assessing simulation of daily temperature and precipitation variability with global climate models for present and enhanced greenhouse climate, Int. J. Climatol., 19(1), 1–26, 1999.

Meehl, G. A., Karl, T. R., Easterlling, D. R., et al.: An introduction to trends in extreme weather and climate events: Observations, socioeconomic impacts, terrestrial ecological impacts, and model projections, B. Am. Meteorol. Soc., 81, 413–416, 2000.

Millard, S. P. and Neerchal, N. K.: Environmental Statistics with S-PLUS, London: CRC Press, 2001.

Moberg, A., Jones, P. D., Lister, D., Walther, A., Brunet, M., et al.: Indices for daily temperature and precipitation extremes in Europe analyzed for the period 1901–2000, J. Geophys. Res., 111, D22106, doi:10.1029/2006JD007103, 2006.

Mudelsee, M., Borngen, M., Tetzlaff, G., and Grunewald, U.: No upward trends in the occurrence of extreme floods in central Europe, Nature, 425, 166–169, 2003.

New, M., Hewitson, B., Stephenson, D. B., Tsiga, A., et al.: Evidence of trends in daily climate extremes over Southern and West Africa, J. Geophys. Res., 111, D14102, doi: 10.1029/2005JD006289, 2006.

New, M., Todd, M., Hulme, M., and Jones, P.: Precipitation measurements and trends in the twentieth century, Observed Trends in the Daily Intensity of United Kingdom Precipitation, 21, 1899–1922, 2001.

Nicholls, N., Trewin, B., and Haylock, M.: Climate Extremes Indicators for State of the Environment Monitoring, Australia: State of the Environment, Technical Paper Series No.2 (The Atmosphere), Department of the Environment and Heritage, Canberra, 2000.

Osbom, T. J., Hulme, M., Jones, P. D. and Basnett, T. A.: Observed Trends in the Daily Intensity of United Kingdom Precipitation, Int. J. Climatol., 20, 347–264, 2000.

Peterson, T. C., Folland, C., Gruza, G., Hogg, W., Mokssit, A., and Plummer, N.: Report on the activities of the Working Group on Climate Change Detection and Related Rapporteurs 1998–2001, Report WCDMP-47, WMO-TD 1071, World Meteorological Organization, Geneva, Switzerland, 2001.

Peterson, T. C. and Groisman, P. Y.: Evaporation losing its strength, Nature, 377, 687–688, 1995.

Plummer, N., James, M., Neville, S., et al.: Changes in climate extremes over the Australian region and New Zealand during the twentieth century, Climatic Change, 42, 183–202, 1999.

Ren, G. Y. and Guo, J.: Change in pan evaporation and the influential factors over China: 1956–2000, J. Natural Resources (in Chinese) 21(1), 31–44, 2006.

Roderick, M. L. and Farquhar, G.D.: The cause of decreased pan evaporation over the past 50 years, Science, 298, 1410–1411, 2002.

Roy, S. S. and Balling R.C.: Trends in extreme daily rainfall indices in India, Int. J. Climatol., 24, 457–466, 2004.

Salinger, M. J. and Griffiths, G.M.: Trends in New Zealand daily temperature and rainfall extremes, Int. J. Climatol., 21, 1437–1452, 2001.

Snedecor, G. W. and Cochran, W.G.: Statistical Methods, Eighth Edition, Iowa State University Press, Ames, IA, 1989.

STARDEX Project, available at: http://www.cru.uea.ac.uk/cru/projects/stardex/.

Su, B. D., Jiang, T., and Jin, W. B.: Recent trends in temperature and precipitation extremes in the Yangtze River basin, China. Theor. Appl. Climatol., 83, 139–151, 2006.

Suppiah, R. and Hennessy K.: Trends in seasonal rainfall, heavy rain-days, and number of dry days in Australia 1910–1990, Int. J. Climatol., 18, 1141–1155, 1998.

Tarhule, S. and Woo, M.: Changes in rainfall characteristics in northern Nigeria, Int. J. Climatol., 18, 1261–1272, 1998.

Tromel, S. and Schonwiese, C. D.: Probability change of extreme precipitation observed from 1901 to 2000 in Germany, Theor. Appl. Climatol., 87, 29–39, 2007.

Tu, M., Hall, M. J., de Laat, P. J. M., de Wit, M. J. M.: Extreme floods in the Meuse river over the past century: aggravated by land-use changes?, Phys. Chem. Earth, 30, 267–276, 2005.

Wang, S. W.: Diagnostic studies on the climate change and variability for the period of 1880–1990, Acta Meteorol. Sinica (in Chinese), 52 (2), 261–273, 1994.

Wang, Z. L., Chen, X. H., and Li, Y.: Spatial and temporal changes in vegetation coverage over the Pearl River Basin, Ecological Science (in Chinese), 25(3), 303-307, 2006.

Xu, G. Y., Yan, G. X. Q., and Sun, X. G.: Interdecadal and interannual variation characteristics of rainfall in north china and its relation with the northern hemisphere atmospheric circulations, Chinese J. Geophys. (in Chinese), 48 (2), 511–518, 2005.

Yu, S. and Lin, X.: Climatic jump from late 1970s to early 1980s and its effect, Acta Meteorol. Sinica, 16(1), 62–74, 2002

Yue, S., and Wang, C.Y.: Applicability of prewhitening to eliminate the influence of serial correlation on the Mann-Kendall test, Water Resour. Res., 38(5), 1068, doi:10.1029/2001WR000861, 2002.

Zhai, P., Zhang, X., Wan, H., and Pan, X.: Trends in Total Precipitation and Frequency of Daily Precipitation Extremes over China, J. Climate, 18(6), 1096–1107, 2005.

Zhang, Q., Jiang, T., Germmer, M., Becker, S.: Precipitation, temperature and discharge analysis from 1951 to 2002 in the Yangtze River basin, China, Hydrol. Sci. J., 50(1), 65–80, 2005a.

Zhang, X., Aguilar E., Sensoy S., et al.: Trends in Middle East climate extreme indices from 1950 to 2003, J. Geophys. Res., 110, D22104, doi: 10.1029/2005JD006181, 2005b.

Zhang, X., Hogg, W. D., and Mekis, É.: Spatial and temporal characteristics of heavy precipitation events over Canada, J. Climate, 14, 1923–1936, 2001.