Theor Appl Climatol (2014) 117:169–174
DOI 10.1007/s00704-013-0986-x

ORIGINAL PAPER

Extreme precipitation events identified using detrended fluctuation analysis (DFA) in Anhui, China

Qi Zhang · Jiquan Zhang · Denghua Yan · Yongfang Wang

Received: 12 March 2013 / Accepted: 25 July 2013 / Published online: 30 August 2013

Abstract Extreme weather events include unusual, severe or unseasonal weather, and weather at the extremes of the historical distribution. They have become more frequent and intense under global warming, especially in mid-latitude areas. They bring about great agricultural and economic losses. It is important to define the threshold of extreme weather event because it is the starting point of extreme weather event research, though it has been of seldom concern. Taking extreme precipitation events in Anhui, China as an example, the detrended fluctuation analysis (DFA) method is introduced to define the threshold of extreme weather events. Based on it, the spatial and temporal distributions of extreme precipitation events are analyzed. Compared to the traditional percentile method, DFA is based on the long-term correlation of time series. Thresholds calculated by DFA are much higher than the 99th percentile and the values are higher in the south and lower in the north. This spatial pattern is similar to the annual precipitation spatial pattern. There is an obvious increasing trend in the number of days with extreme precipitation, especially after the 1980s. This observation supports the point that more extreme events happen under global warming.

1 Introduction

It is very likely that global average temperature rises, where extreme high or low temperature days and extreme high precipitation days increase on over 70 % of land area. This trend becomes more obvious during several decades under climate change, especially in mid-latitude areas; this trend will likely continue (Alexander et al. 2006; IPCC 2007). Extreme weather events lead to great losses. It seems that extreme weather events cause more damage to society and ecosystems than simple shifts in the mean values; developing countries are much vulnerable to it (Mirza 2003). In 2011, USA experienced the most extreme high precipitation year on record during the last 100 years (Samenow 2011). Moscow suffered a heat wave with an average unprecedented July temperature at least since the nineteenth century, when measurements began; there is approximate 80 % probability that the 2010 July heat record would not have happened without climate warming (Rahmstorf and Coumou 2011). The number of extreme weather events costing more than a billion dollars was 11 in the USA in 2012, including Superstorm Sandy, the Waldo Canyon fire in Colorado, and the Great Drought, which was the largest since 1939. These extreme events bring great economic losses. Society's rising interest in extreme weather events promotes this research.

The starting point for extreme weather events research is to define a threshold. ETCCDI (The joint CCl/CLIVAR/JCOMM Expert Team on Climate Change Detection and Indices; http://cccma.seos.uvic.ca/ETCCDI) and STARDEX project (STAtistical and Regional dynamical Downscaling of EXtremes for European regions; http://www.cru.uea.ac.uk/cru/projects/stardex/) recommend more than 20 indexes about precipitation and temperature, which are the most commonly used indexes to identify extreme events. The percentile method is the most important one. Extreme precipitation events are defined as those that exceed a threshold percentile of daily precipitation; for example, Bell et al. (2004), Wang and Zhou (2005), Wan Zin et al. (2010) used the 95th percentile for extreme daily precipitation events analysis. Besides the percentile methods, greatest 5-day total rainfall, longest dry period, maximum number of consecutive days with precipitation, etc. are used to indicate extreme climate events (Li et al. 2010; Hu et al. 2012; Jiang et al. 2012). Friederichs (2010) used extreme value theory to provide appropriate distributions of precipitation and to decide the threshold of extreme

Q. Zhang · J. Zhang (✉) · Y. Wang
College of Environment, Northeast Normal University,
Changchun 130024, China
e-mail: zhangjq022@nenu.edu.cn

D. Yan
China Institute of Water Resources and Hydropower Research,
Beijing 100038, China

precipitation events. It can be found that these indexes are arbitrarily delineated. They could be considerably influenced by subjective factors. This method cannot describe the long-term correlation in climate system. So, although the percentile method is widely used, it still has large uncertainties.

Studies show that meteorological and climatological records have positive long-term correlations (Fraedrich 2002; Yang et al. 2008). In recent years, detrended fluctuation analysis (DFA) has become a widely used technique for the determination of fractal scaling properties and for the detection of long-range correlations in time series (Kantelhardt et al. 2002). It was first proposed by Peng et al. (1994) when studying DNA sequences. It can be used to characterize the fractal properties more reliably without being affected by nonstationarity and trends (Yuan et al. 2012). It has been successfully applied to diverse fields such as noise analysis, long-time weather records, geology, and DNA sequences (Weron 2002). Chen et al. (2007) used DFA to study the long-range correlation in daily relative humidity. Kurnaz (2004) used DFA to separate the trends and the correlations in data series and then to analyze the persistence of weather and climate regimes. Long-term correlation in hydroclimate records results in a clustering of extreme events; the return interval depends on the previous one (Bunde et al. 2005). The DFA method could quantify the presence of long-range power law correlations in both gait and heartbeat fluctuations; long-range correlation of gait timing reduces in Parkinson's disease patients (Ivanov et al. 2009; Bartsch et al. 2007).

Extreme data loss could affect the behavior of anticorrelated signals even for a very small fraction; but for positively correlated signals, it practically remained unchanged even for extreme data loss of up to 90 % (Ma et al. 2010). Recently, researchers used the DFA method to study extreme weather events based on the theory that extreme events are the abnormal condition in the climate system and cannot influence the long-term correlations of the climate time series (Yang et al. 2008).

In this study, the thresholds of extreme precipitation events were calculated by both DFA and percentile method for the purpose of estimating the spatial–temporal characteristic of extreme precipitation event in Anhui. The main objective of this study is to demonstrate the effectiveness and advantage of DFA on the detection of local threshold of extreme weather conditions over other traditional approaches such as percentile method.

2 Data and methods

2.1 Study area

Anhui province is located in southeast China between 114°54′–119°37′E and 29°41′–34°38′N (Fig. 1). It has a warm semihumid monsoon climate in the north and a subtropical humid monsoon climate in the south divided by Huai River. The Yangtze River flows through the southern part of Anhui province. Annual mean temperature is 14–17 °C. Annual mean precipitation is around 1,200 mm, which varies greatly from year to year. Drought and flood disasters happen frequently. For example, south Anhui province experienced severe flood disaster after drought in June 2011, which caused 0.22 billion dollar of losses.

2.2 Data

The time series of the daily precipitation record from 1 January 1960 to 31 December 2009 at 14 meteorological stations (Table 1) in Anhui province were used in this study. The data were collected from China Meteorological Data Sharing Service System (http://cdc.cma.gov.cn/home.do).

2.3 Methods

The DFA has proved useful in revealing the extent of long-range correlations in time series which would not be affected by system noise. Extreme events can be regarded as climate system noises. The DFA method is based on random walk theory. The detailed methodology is as follows (Yang et al. 2008; Bashan et al. 2008):

First, we consider a time series (x_i), $i=1,\ldots,N$ (with mean$\langle x \rangle$), integrating this time series into X_t, which can be considered as the position of a random walk on a linear chain after t steps.

$$X_t = \sum_{i=1}^{t} [x_i - < x >] \tag{1}$$

Next, X_t is divided into $N_s \equiv [N/s]$ non-overlapping segments of equal size s. Then, we estimate a piecewise polynomial trend $y_s(n)$ within each segment. The detrended profile function $\widetilde{x}_s$ on scale s is determined by:

$$\widetilde{x}_s = X(n) - y_s(n) \tag{2}$$

The degree of the polynomial can be varied in order to eliminate linear, quadratic, or higher-order trends of the profile function. The variance of $\widetilde{x}_s$ yields the fluctuation function on scale s

$$F(s) = \left\{ \frac{1}{N} \sum_{n=1}^{N} \widetilde{X}_s(n) \right\}^{1/2} \tag{3}$$

The above computation for different range of scale s is repeated to see how $F(s)$ depends on the segment length. If

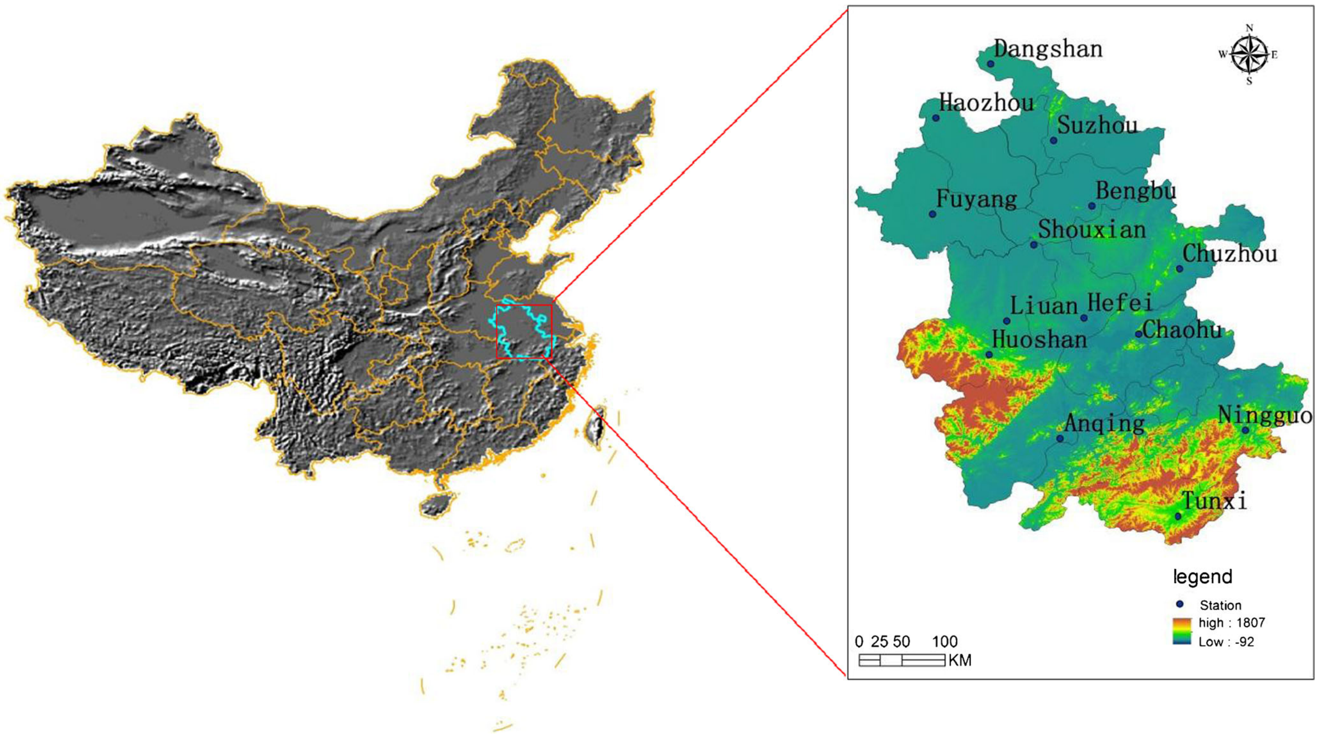

Fig. 1 Study area and spatial distribution of selected meteorological stations

$F(s)$ increases for increasing s, asymptotically it is written as:

$$F(s) \sim s^{a} \tag{4}$$

one finds that the scaling exponent α is related to the correlation exponent. If $0.5<\alpha<1$, the data are positively long-term correlated. The higher the α is, the stronger the correlations in the series are. A value of $\alpha=0.5$ thus indicates that there are no correlations. The case $\alpha<0.5$ corresponds to long-term anti-correlations. In this research scale, s ranges from 10 to $10^{1.5}$.

Extreme events are small probability events in time series. They are the anomalous condition of the climate system. Climate system has long-term correlation. We assume that the extreme event will not influence the long-range correlation of the system. DFA index can reflect the long-term correlation. Based on this hypothesis, researchers use DFA to determine the threshold of extreme weather events. We have a time series $\{x_i\}$:

1. We first need to find out the maximum ($x_{\max}$) and minimum value ($x_{\min}$) of series $\{x_i\}$
2. Then, we need to find out the median value R of the series $\{x_i\}$
3. After leaving out the values x_i ($x_i \geq (x_{\max} - d \times k)$) in series $\{x_i\}$, we get the new series Y_k. Holding on to this process, $k=1,2,\ldots,(x_{\max}-R)/d$, finally we get $(x_{\max}-R)/d$ new series. d is the interval; it influences the accuracy of the model. Here, d is 10; the amount of rainfall is 1 mm.
4. By calculating the long-range correlation indexes α of each new series Y_k, we can get the map of $\alpha \sim x_i$ The index α will not change much when leaving out the extreme large values in $\{x_i\}$. This means that those data do not have influence on the long-range correlation of the time series. When the index α has an obvious trend to diverge

Table 1 Characteristics of selected meteorological stations

Station	Longitude	Latitude	Altitude (m)	Station	Longitude	Latitude	Altitude (m)
Ningguo	118.98	30.61	89.4	Dangshan	116.33	34.43	44.2
Hefei	117.3	31.78	27.0	Haozhou	115.76	33.87	37.7
Liuan	116.5	31.75	60.5	Bengbu	117.38	32.95	18.7
Chaohu	117.86	31.61	22.4	Suzhou	116.98	33.63	25.9
Huoshan	116.31	31.40	68.1	Fuyang	115.73	32.87	32.7
Chuzhou	118.30	32.30	27.5	Tunxi	118.28	29.72	142.7
Shouxian	116.78	32.55	22.7	Anqing	117.05	30.533	19.8

from the original value, we define the corresponding $\{x_i\}$ as the threshold of extreme event.

3 Results and discussion

3.1 The threshold calculated by DFA and the 99th percentile

We used the DFA to calculate the threshold of extreme precipitation events in Anhui province. Figure 2 shows the change of DFA index of the daily precipitation time series as taking the maximum datum out of the time series gradually. We just presented three typical stations, which are Dangshan, Ningguo, and Chaohu located in the north, middle, and south Anhui province, respectively. It can be seen that all the DFA values are between 0.5 and 1, so positive long-term correlation exist in the daily precipitation time series and is much stronger when the time series is complete. In Fig. 2b, the point from where the DFA value began to diverge from the original value is easy to identify; but on Fig. 2a,c, it is hard to identify. There are more than one point at where the value seems to begin to diverge. On this condition, we need to refer to the percentile method to find a point which is most similar to the result of the percentile method. Figure 2d shows how $F(s)$ change with scale s before and after leaving out extreme values in Ningguo; the DFA index α changed a lot from 0.6382 to 0.6205. Then, we compared the thresholds calculated by the DFA and percentile method. Thresholds calculated by the 99th percentile are much smaller than by DFA method. The thresholds are around 80 mm by the 99th percentile and 90 mm by DFA method (Table 2). The extreme events measured by the structure of precipitation time series itself are much more serious than that measured by the 99th percentile method.

3.2 Spatial distribution of the threshold of extreme precipitation events

Average annual precipitation in Anhui is about 1,200 mm. It is much higher in the south around 1,500 mm and lower in the north around 800 mm. The spatial pattern varies considerably with latitude. From Fig. 3a, it can be found that the thresholds calculated by the 99th percentile are higher in the north and south, and lower in the middle. The pattern map by DFA method is different, which is higher in the southwest and lower in the north along with latitude. The DFA threshold pattern looks similar to the precipitation pattern.

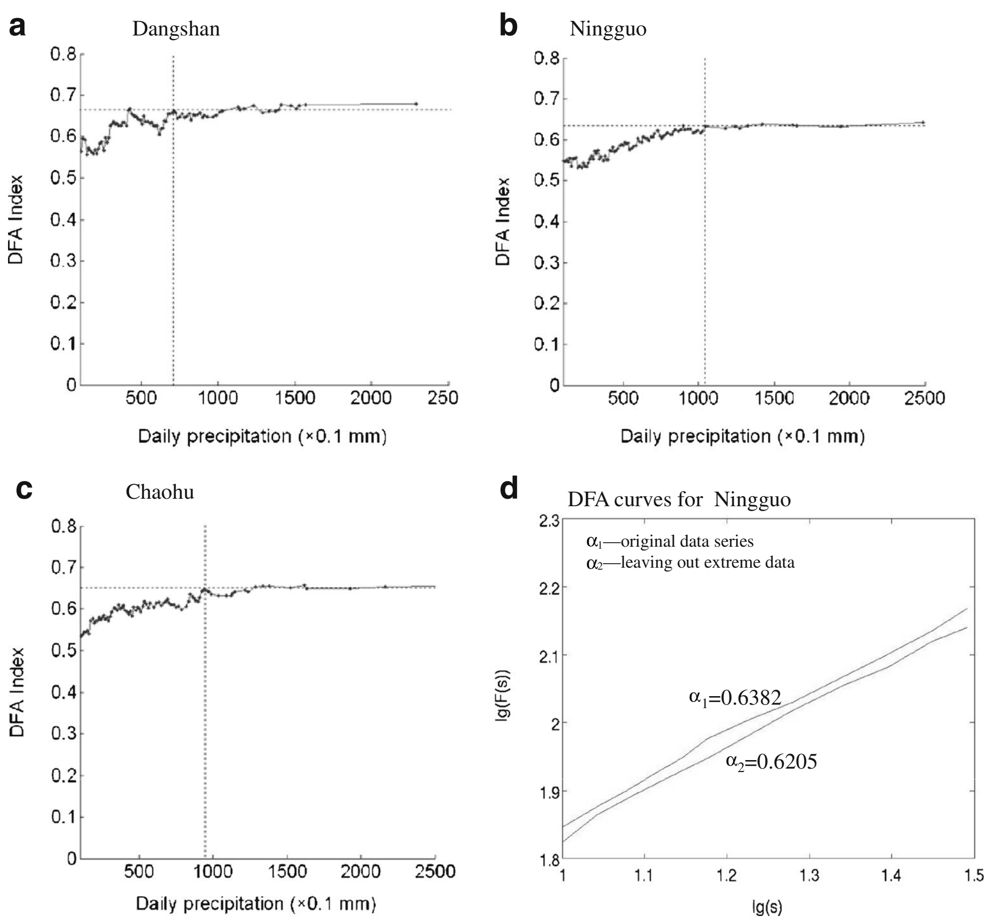

Fig. 2 Thresholds of three typical stations determined by DFA; **a** Dangshan, **b** Ningguo, and **c** Chaohu. The x-axis is the precipitation cutoff for DFA calculation; y-axis is the DFA index α after leaving out the data above the cutoff. **d** DFA curves before and after leaving out extreme values in Ninguo, scale s range from 10 to $10^{1.5}$, α_1 and α_2 are the DFA index

Table 2 Thresholds calculated by the 99th percentile and DFA method (mm)

Stations	99th percentile	DFA	Stations	99th percentile	DFA
Ningguo	66.3	104.5	Dangshan	81.3	71.0
Hefei	68.9	102.0	Haozhou	82.2	80.5
Liuan	72.4	88.0	Bengbu	82.5	86.0
Chaohu	73.5	92.0	Suzhou	82.9	94.0
Huoshan	75.0	104.0	Fuyang	86.2	85.0
Chuzhou	78.9	102.5	Tunxi	87.6	84.0
Shouxian	80.6	100.0	Anqing	92.0	114.0

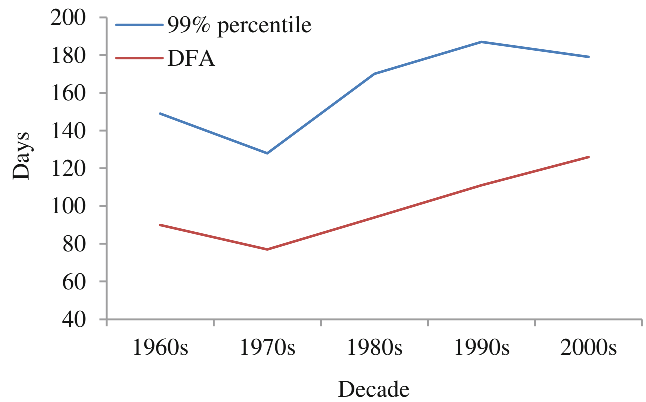

Fig. 4 Extreme daily high precipitation days that happened in each decade

3.3 The number of extreme precipitation days in each decade

We count the extreme precipitation days that happened in each decade in Anhui. Figure 4 shows that there is a considerable increasing trend for the extreme precipitation days. It supports the conclusion that there are more extreme events under global warming. The numbers of days measured by the 99th percentile method are larger than by the DFA method. That is because the threshold by the 99th percentile method is much lower.

4 Conclusion

Extreme weather events bring great agriculture losses and have become more frequent under global climate change. DFA can detect the long-term correlation in daily precipitation time series. It can be used to define the threshold of extreme weather events based on the inner structure of data series. After comparing the thresholds calculated by DFA and the 99th percentile method, it was found that the thresholds are much larger by DFA. In addition, the spatial distribution from the threshold by DFA is in more accordance with annual precipitation distribution. The thresholds are considerably higher in areas with higher annual precipitation. The numbers of extreme daily precipitation events in each decade have been markedly increasing from the 1960s to the 2000s. This study supports the effectiveness and advantage of introducing DFA into the field of extreme weather events threshold determination. But uncertainties still exist in this method. In some cases, it is not clear when the DFA values begin to diverge from the

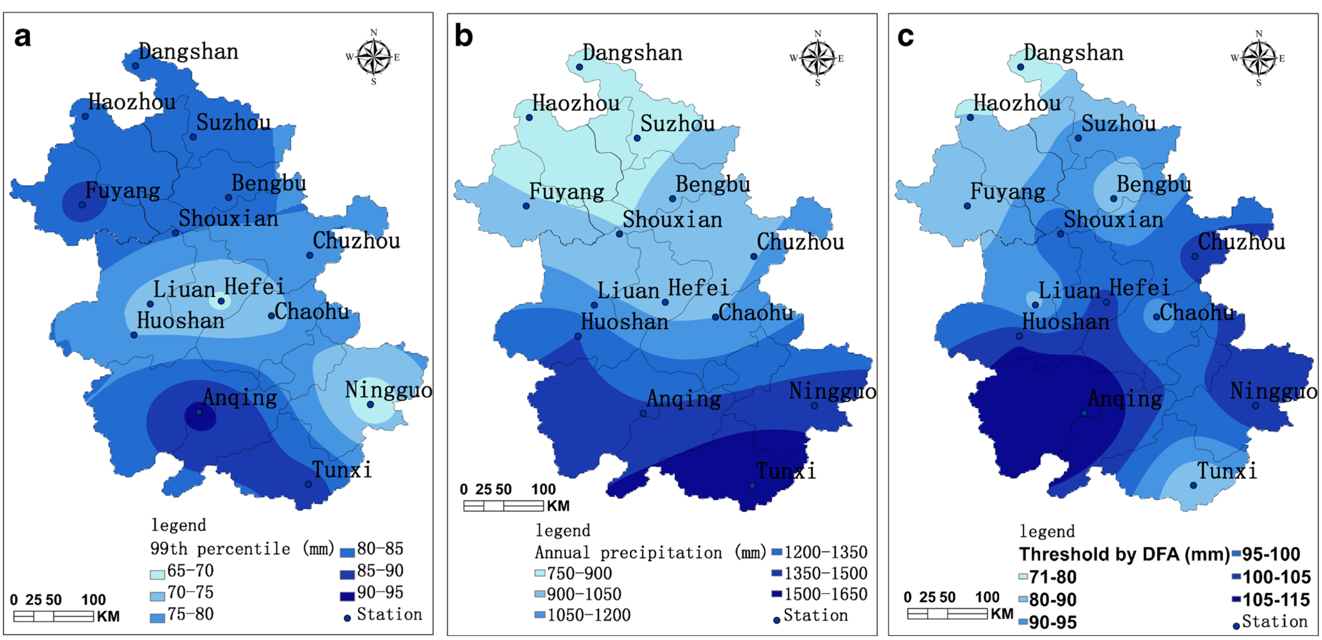

Fig. 3 Annual precipitation map and the thresholds of extreme daily precipitation events. **a** The pattern map of the threshold of extreme precipitation event calculated by the 99th percentile, **b** the spatial pattern of annual precipitation, and **c** the spatial pattern of the threshold of extreme precipitation event calculated by DFA

original number, so it is with some kind of subjectivity that the threshold is detected.

Acknowledgments This study is supported by the National Key Technology R&D Program of China under grant no. 2011BAD32B00-04, the National Grand Fundamental Research 973 Program of China under grant no. 2010CB951102, the National Natural Science Foundation of China under grant nos. 41071326 and 40871236, and the National Scientific Research Special Project of Public sectors (Agriculture) of China under grant no. 200903041.

References

Alexander LV, Zhang X, Peterson TC et al (2006) Global observed changes in daily climate extremes of temperature and precipitation. J Geophys Res Atmos 111, D05109. doi:10.1029/2005JD006290

Bashan A, Bartsch R, Kantelhardt JW, Havlin S (2008) Comparison of detrending methods for fluctuation analysis. Physica A 387:5080–5090

Bartsch R, Plotnik M, Kantelhardt JW et al (2007) Fluctuation and synchronization of gait intervals and gait force profiles distinguish stages of Parkinson's disease. Physica A 383(2):455–465

Bell JL, Sloan LC, Snyder MA (2004) Regional changes in extreme climate events: a future climate scenario. J Clim 17:81–87

Bunde A, Eichner JF, Kantelhardt JW, Havlin S (2005) Long-term memory: a natural mechanism for the clustering of extreme events and anomalous residual times in climate records. Phys Rev Lett 94:048701

Chen X, Lin G, Fu Z (2007) Long-range correlations in daily relative humidity fluctuations: a new index to characterize the climate regions over China. Geophys Res Lett 34, L07804

Fraedrich K (2002) Fickian diffusion and Newtonian cooling: a concept for noise induced climate variability with long-term memory? Stochast Dyn 02:403

Friederichs P (2010) Statistical downscaling of extreme precipitation using extreme value theory. Extremes 13:109–132

Hu Y, Maskey S, Uhlenbrook S (2012) Trend in temperature and rainfall extremes in the Yellow River source region, China. Clim Chang 110:403–429

IPCC (2007) Climate change 2007: the physical science basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, UK

Ivanov PC, Ma QD, Bartsch RP et al (2009) Levels of complexity in scale-invariant neural signals. Phys Rev E 79:041920

Jiang D, Li Z, Wang Q (2012) Trends in temperature and precipitation extremes over Circum-Bohai-Sea Region, China. Chin Geogr Sci 22(1):75–87

Kantelhardt JW, Zschiegner SA, Koscielny-Bunde E, Bunde A, Havlin S, Stanley HE (2002) Multifractal detrended fluctuation analysis of nonstationary time series. Physica A 361:87–114

Kurnaz ML (2004) Application of detrended fluctuation analysis to monthly average of the maximum daily temperatures to resolve different climates. Fractals 12:365–373

Li Z, Zheng F, Liu W et al (2010) Spatial distribution and temporal trends of extreme temperature and precipitation events on the Loess Plateau of China during 1961–2007. Quat Int 266:92–100

Ma QDY, Bartsch RP, Bernaola-Galvan P et al (2010) Effect of extreme data loss on long-range correlated and anti-correlated signals quantified by detrended fluctuation analysis. Phys Rev E 81:031101

Mirza MMQ (2003) Climate change and extreme weather events: can developing countries adapt? Clim Police 3:233–248

Peng CK, Buldyrev SV, Havlin S (1994) Mosaic organization of DNA nucleotides. Phys Rev E 49(2):1685–1689

Rahmstorf S, Coumou D (2011) Increase of extreme events in a warming world. Proc Natl Acad Sci USA 108:17905–17909

Samenow J (2011) U.S. had most extreme precipitation on record in 2011. The Washington post local. 12 January 2012

Wan Zin WZ, Jamaludin S, Deni SM, Jemain AA (2010) Recent changes in extreme rainfall events in Peninsular Malaysia: 1971–2005. Theor Appl Climatol 99:303–314

Wang Y, Zhou L (2005) Observed trends in extreme precipitation events in China during 1961–2001 and the associated changes in large-scale circulation. Geophys Res Lett 32, L09707. doi:10.1029/2005GL022574

Weron R (2002) Estimating long-range dependence: finite sample properties and confidence intervals. Physica A 312:285–299

Yang P, Hou W, Feng G (2008) Determining the threshold of extreme events with detrended fluctuation analysis. Acta Phys Sin 58:5333–5343

Yuan N, Fu Z, Li H, Mao J (2012) Effect of extreme value loss on long-term correlated time series. Theor Appl Climatol 109:133–140