Multifractal detrended fluctuation analysis of streamflow series of the Yangtze River basin, China

Qiang Zhang,1,2* Chong-Yu Xu,3 Yongqin David Chen4 and Zuguo Yu5,6
1 State Key Laboratory of Lake Science and Environment, Nanjing Institute of Geography and Limnology, Chinese Academy of Sciences, Nanjing 210008, China
2 Institute of Space and Earth Information Science, The Chinese University of Hong Kong, Hong Kong, China
3 Department of Geosciences, University of Oslo, Norway
4 Department of Geography and Resource Management, The Chinese University of Hong Kong, Hong Kong, China
5 School of Mathematical Sciences, Queensland University of Technology, GPO Box 2434, Brisbane, Q4001, Australia
6 School of Mathematics and Computational Science, Xiangtan University, Hunan 411105, China

Abstract:
Scaling and multifractal properties of the hydrological processes of the Yangtze River basin were explored by using a multifractal detrended fluctuation analysis (MF-DFA) technique. Long daily mean streamflow series from Cuntan, Yichang, Hankou and Datong stations were analyzed. Using shuffled streamflow series, the types of multifractality of streamflow series was also studied. The results indicate that the discharge series of the Yangtze River basin are non-stationary. Different correlation properties were identified within streamflow series of the upper, the middle and the lower Yangtze River basin. The discharge series of the upper Yangtze River basin are characterized by short memory or anti-persistence; while the streamflow series of the lower Yangtze River basin is characterized by long memory or persistence. $h(q)$ vs $q$ curves indicate multifractality of the hydrological processes of the Yangtze River basin. $h(q)$ curves of shuffled streamflow series suggest that the multifractality of the streamflow series is mainly due to the correlation properties within the hydrological series. This study may be of practical and scientific importance in regional flood frequency analysis and water resource management in different parts of the Yangtze River basin. Copyright © 2008 John Wiley & Sons, Ltd.

KEY WORDS multifractal property; multifractal detrended fluctuation analysis; hydrological processes; the Yangtze River basin

Received 10 February 2008; Accepted 26 June 2008

INTRODUCTION

As one of the major components of the hydrological cycle, streamflow analysis and modelling have received considerable attention in hydrological science in recent decades. The last 20 years or so has witnessed great progress in studying the scaling behaviour of some geophysical fields, including streamflow, rainfall, temperature, etc. (Lovejoy and Mandelbrot, 1985; Pandey et al., 1998). It is accepted that there exists persistence in climatological and hydrological series over a wide range of time scales. Generally, the persistence on long time scales is larger than that on short time scales (Pelletier and Turcotte, 1997). Hurst (1951) and Hurst et al. (1965) demonstrated persistence using the rescaled-range technique, identifying a power-law rescale-range plot in the climatological and hydrological series with an average exponent of 0.73. Actually, many signals present complex behaviour that can exhibit long-range power law correlation and/or nonstationary trends, e.g. DNA sequences (Yu et al., 2001; Yu et al., 2004) and meteorological measurements (Olsson, 1996; Lin and Fu, 2008). This complex behaviour can be characterized by the famous Hurst exponent, or scaling exponent, which quantifies the correlation properties of a signal. It is feasible to characterize these diverse phenomena by using critical exponents and thus to identify similarities between different systems (Chianca et al., 2005).

The currently well-evidenced global warming is believed to accelerate the hydrological cycle (Menzel and Bürger, 2002; Xu and Singh, 2005; Zhang et al., 2008). Therefore, increasing attention has been paid to sustainability and hydro-environmental protection, which require modelling of dynamic processes such as runoff-induced wash-off from impermeable surfaces and flood prediction from ungauged basins. Hydrologists and meteorologists come to realize the importance of investigating scaling properties of hydro-meteorological series in that good understanding of scaling properties of hydrological system is of great importance for hydrological modelling, regionalization of flood frequency and assessment of hydrological conditions in ungauged area based on gauged regions. Peng et al. (1994) introduced the detrended fluctuation analysis (DFA), which has since then been widely used to detect the long-range correlations in stationary and nonstationary time series (Maraun et al., 2004; Bunde et al., 2006). The DFA method has been applied successfully in diverse fields such as DNA and protein sequences, heart rate dynamics, weather records (Kantelhardt et al., 2002; Yu et al.,...
The multifractal detrended fluctuation analysis (MF-DFA) proposed by Kantelhardt et al. (2002) is a modified version of DFA to detect multifractal properties of time series. It allows a reliable multifractal characterization of nonstationary time series such as geophysical phenomena (Kantelhardt et al., 2002). It is now known that multifractal is the appropriate framework for scaling fields and time series and thus can provide the natural framework for analysing and modelling various geophysical processes (Pandey et al., 1998). Gupta et al. (1994) reported that streamflows are multiscaling with basin area. Tessier et al. (1996) analysed multifractal properties of river flow series from 30 small river basins in France. These researches provided a broader framework in modelling the rainfall-runoff processes, such as topography and river network, that generate and modify the streamflow processes through the basin (Gupta et al., 1996; Movahed and Hermanis, 2008).

The Yangtze River (Changjiang) (91°E–122°E, 25°N–35°N) is the longest river in China and the third longest river in the world. It has a drainage area of about 1.8085 × 106 km2 with a mean annual discharge of 3.400 m3 s−1 measured at Hankou Station. The Yangtze River originates in the Qinghai-Tibet Plateau and flows about 6300 km eastwards to the East China Sea (Zhang et al., 2006). The upper Yangtze River basin is from the origin to Yichang, with a length of 4504 km and drainage area of about 1.0 × 106 km2. The river reach between Yichang and Hankou (the outlet of Poyang Lake) is the middle Yangtze River basin, with a length of 955 km and drainage area of 6.8 × 106 km2. The lower Yangtze River basin is from Hankou to the river mouth, with a length of 938 km and drainage area 0.12 × 106 km2 (CWRC, 2000). The river is located in the monsoon region of East Asia subtropical zone, and has a mean annual precipitation of about 1090 mm (Zhang et al., 2000). The river reach between Yichang and Hankou has a mean annual discharge of 4504 km and drainage area of about 1.0 × 106 km2. The river reach between Yichang and Hankou (the outlet of Poyang Lake) is the middle Yangtze River basin, with a length of 955 km and drainage area of 6.8 × 106 km2. The lower Yangtze River basin is from Hankou to the river mouth, with a length of 938 km and drainage area of 0.12 × 106 km2.

**DATA AND METHODOLOGY**

**Data**

Long daily mean streamflow series extracted from four hydrological stations, i.e. Cuntan, Yichang, Hankou and Datong, along the mainstream of the Yangtze River basin have been analysed (Figure 2). The location of these four stations is shown in Figure 1. Detailed information about the hydrological dataset is given in Table I. The streamflow data represent the hydrological conditions of the upper, middle and lower Yangtze River basins, respectively. The quality of the streamflow series is controlled by the Changjiang Water Resources Commission, Ministry of Water Resources, China. The length of the data series is more than 50 years, with some more than 100 years.

**Methodology**

Multifractal detrended fluctuation analysis (MF-DFA) is a generalization of standard DFA by identifying the scaling of the qth-order moments of the time series, which may be non-stationary (Kantelhardt et al., 2002). Movahed et al. (2006) described the procedure of MF-DFA. Actually, the first three steps are the same as those in the conventional DFA procedure. Assuming x_k is a time series, k = 1, ..., N.

**Step 1.** Determine the 'profile' \( Y(i) = \sum_{k=i}^{N} [x_k - \langle x \rangle], \)

\( i = 1, ..., N, \) where \( \langle x \rangle \) is the mean of \( x_k \).

**Step 2.** Divide the profile \( Y(i) \) into \( N_s = \text{int}(N/s) \) non-overlapping segments of equal length \( s \); here \( \text{int}(N/s) \) is

<table>
<thead>
<tr>
<th>Station name</th>
<th>Drainage area (×10^4 km^2)</th>
<th>Time interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cuntan station</td>
<td>86.66</td>
<td>1 Jan. 1893–31 Dec. 2004</td>
</tr>
<tr>
<td>Yichang station</td>
<td>100.55</td>
<td>1 Jan. 1946–31 Dec. 2004</td>
</tr>
</tbody>
</table>

Figure 1. Location of the Yangtze River basin and hydrological stations

Figure 2. Streamflow series of Cuntan, Yichang, Hankou and Datong stations of the Yangtze River basin
the integer part of $N/s$. Since the length $N$ of the series is often not a multiple of the timescale $s$ considered, a short part at the end of the profile may remain. To retain this part of the series, the same procedure is repeated starting from the opposite end. Thereby, 2Ns segments are obtained.

**Step 3.** Calculate the local trend for each of the 2Ns segments by a least squares fit of the series. Then determine the variance

$$F^2(s, v) = \frac{1}{s} \sum_{i=1}^{s} \{Y[(v-1)s + i] - y_s(i)\}^2$$

for each segment $v$,

$$v = 1, \ldots, N_s,$$

and $F^2(s, v) = \frac{1}{s} \sum_{i=1}^{s} \{Y[N - (v - N_s)s + i] - y_s(i)\}^2$,

for $v = N_s + 1, \ldots, 2N_s$.

Here, $y_s(i)$ is the fitting polynomial in segment $v$. Linear, quadratic, cubic or higher order polynomials can be used in the fitting procedure (DFA1, DFA2, DFA3... DFAm).

**Step 4.** Average over all segments to obtain the $q$th-order fluctuation function, defined as $F_q(s) = \left\{ \frac{1}{2N_s} \sum_{i=1}^{2N_s} [F^2(s, v)]^{q/2} \right\}^{1/q}$, where $q \neq 0$, $s \geq m + 2$.

**Step 5.** Determine the scaling behaviour $F_q(s) \propto s^{h(q)}$ of the fluctuation functions by the log–log plot of $F_q(s)$ versus $s$ for each value of $q$.

For stationary time series, the exponent $h(2)$ for small time scales is identical to the well-known Hurst exponent $H$. For non-stationary signal, the relation between the exponent $h(2)$ for small scales and the Hurst exponent $H$ is $H = h(2) - 1$ (Hu et al., 2001). For small scales where the effect of the sinusoidal trend is not pronounced, $h(2) > 1$ indicates that the time series is non-stationary (Movahed et al., 2006). It is well known that for uncorrelated series, the scaling exponent $H$ equals to $0.5$; $0.5 < H < 1$ indicates long memory or persistence; $0 < H < 0.5$ indicates short memory or anti-persistence. Hence we can use the value of $h(2)$ to determine whether a time series is stationary or nonstationary and detects its correlations.

In general, two different types of multifractality in time series can be distinguished (Movahed et al., 2006): (1) multifractality due to a broadening of the probability density function (PDF) of the time series. In this case the multifractality cannot be removed by shuffling the series; (2) multifractality due to different correlations in small and large scale fluctuations. In this case the time series have a PDF with finite moments. Therefore the shuffled time series will show mono-fractal scaling because all long-term correlations are destroyed by the shuffling procedure. The easiest way to clarify the type of multifractality is by analysing the corresponding shuffled and surrogate time series. The random shuffling of time series destroys the long range correlation. Therefore, if the multifractality is only due to the long range correlation, $h_{	ext{shuf}}(q) = 0.5$. The multifractality nature due to the broadening of the PDF signals is not affected by the shuffling procedure. However, if both kinds of multifractality are present, the shuffled series will show weaker multifractality when compared with the original time series.

**RESULTS AND DISCUSSION**

Figure 3 displays the DFA1 exponent of the streamflow series of Cuntan, Yichang, Hankou and Datong stations of the Yangtze River basin. Based on log-log plots of $F_q(s)$ versus $s$ of the streamflow series, one crossover point can be detected for the three curves of $F_q(s)$ versus $s$. The timing for these crossover points is between 346 and 398 days, and is due to annual periodicity. To determine the statistical properties of the fluctuations of streamflow series, we compute the three scaling exponents for smaller time scales. The $h(2)$ values of the streamflow series of these four stations are $1.3027 \pm 0.0141$ (Cuntan station), $1.3637 \pm 0.0095$ (Yichang station), $1.4911 \pm 0.0063$ (Hankou station) and $1.6659 \pm 0.0132$ (Datong station) respectively. The $h(2)$ values of the streamflow series of Cuntan, Yichang, Hankou and Datong are all larger than 1. These numerical results indicate that these four streamflow series are nonstationary signals. Therefore, based on the relationship between the Hurst exponent and the exponent $h(2)$ for small scales, i.e. $H = h(2) - 1$, we obtained associated Hurst exponents for the streamflow series of Cuntan, Yichang, Hankou and Datong stations as $0.3027 \pm 0.0141, 0.3637 \pm 0.0095, 0.4911$ and $0.6659 \pm 0.0132$. Therefore, for Cuntan, Yichang and Hankou stations, the streamflow fluctuations are characterized by short memory or anti-persistance. Moreover, the Hurst exponent of the hydrological series from the Hankou station is close to 0.5, implying that the streamflow series of the Hankou station is close to being an uncorrelated series. For the streamflow fluctuations at Datong station in the lower Yangtze River basin, the numerical results suggest long memory or persistence.

The above results indicate that the streamflow series of the four stations of the Yangtze River basin are non-stationary processes. Furthermore, the streamflow series of the Cuntan, Yichang and Hankou stations are characterized by short memory, and the streamflow series of the Datong stations by long memory. The result of the MF-DFA procedure is the family of the generalized Hurst exponents $h(q)$ (Figure 4). For an actual multifractal signal, $h(q)$ is a decreasing function of $q$; while for a monofractal signal, $h(q)$ is of a constant value. It can be seen from Figure 4 that $h(q)$ vs $q$ curves of original streamflow series indicate strong dependence of $h(q)$ on $q$, suggesting that the streamflow series of the four hydrological stations in the Yangtze River basin are characterized by multifractality. Here we are also interested in the possible source of multifractality. To clarify the
The slope is 1.3027 ± 0.0141

Cuntan station

The slope is 0.2598 ± 0.0134

Crossover point P1

The slope is 1.3637 ± 0.0095

Yichang station

The slope is 0.2364 ± 0.0099

Hankou station

The slope is 1.4911 ± 0.0063

Crossover point P3

The slope is 0.3340 ± 0.0132

Datong station

Crossover point P4

Cuntan station

Yichang station

Hankou station

Datong station

Crossover points

Figure 3. Log–log plots of $F_q(s)$ versus $s$ of streamflow series of Cuntan, Yichang, Hankou and Datong stations in the Yangtze River basin

Figure 4. Generalized Hurst exponent $h(q)$ as a function of $q$ for original and shuffled streamflow series of Cuntan, Yichang, Hankou and Datong stations

type of multifractality, we used a shuffled streamflow series. Figure 4 shows $h(q)$ vs $q$ curves of the shuffled streamflow series: it can be seen that $h(q)$ vs $q$ curves of the shuffled streamflow series indicate almost independence of $h(q)$ on $q$. The $h(q)$ values are largely equal to 0.5. If only the breadth of the PDF is responsible for the multifractality, $h(q) = h_{shuf}(q) = 0$ can be expected. However, if only correlation multifractality is present, one can expect $h_{shuf}(q) = 0.5$ (Movahed et al., 2006). It can be seen from Figure 4 that the $h(q)$ vs $q$ curves of shuffled streamflow series are almost independent of $q$ values and $h(q)$ values are mostly equal to 0.5. Therefore, we can conclude that the multifractality of the streamflow series of the Yangtze River basin is mainly due to the correlation properties (short-term correlation for Cuntan, Yichang and Hankou stations; long-term correlation for Datong station).
Streamflow series are the combined results of the precipitation phenomena and the impact of other basin factors such as soil moisture, plant coverage, and channel geometry. Therefore the gauged streamflow series are the results of the overall complex interactions between precipitation input and the basin factors that modify it (Pandey et al., 1998). It can be seen from Figure 3 that similarities of $h(2)$ values can be identified between Cuntan and Yichang stations, but different $h(2)$ values between Cuntan, Yichang, Hankou and Datong station. Gupta et al. (1996) exploited the hypothesis of statistical self-similarity, or scaling invariance, in the spatial variability of rainfall, channel network structures and floods and also the distributed rainfall-landform-runoff relationships. Xu et al. (2006) divided the whole Yangtze River basin into three parts along the longitude from west to east, which correspond well with the decrease in altitude. The upper region ($<104^\circ$E) has a mean altitude of 2551 m above sea level (m.a.s.l), and the middle ($>104^\circ$E and $<112^\circ$E) and lower regions ($>112^\circ$E) have mean altitudes of 627 and 113 m.a.s.l, respectively. Cuntan and Yichang stations are located in the upper Yangtze River basin, being characterized by similar topography and channel network structures. Hankou station is located in the middle Yangtze River basin. Datong station is located in the lower part of the Yangtze River basin, and the topographical properties are different from those of the middle and upper Yangtze River. The different scaling properties of the streamflow series from Cuntan, Yichang, Hankou and Datong stations are probably due to these factors mentioned above.

CLOSING REMARKS

The scaling and multifractal properties of long streamflow series of the Yangtze River basin have been established by using a multifractal detrended fluctuation analysis technique. Furthermore, using a shuffling procedure, the types of multifractality of the streamflow series of the Yangtze River basin have been established. Several interesting findings have been obtained as follows:

(1) Log-log plots of $F_q(s)$ versus $s$ of streamflow series show that the $h(2)$ values of the streamflow series of these four stations are 1-3027 ± 0.0141 (Cuntan station), 1-3637 ± 0.0095 (Yichang station), 1-4911 ± 0.0063 (Hankou station) and 1-6659 ± 0.0132 (Datong station) respectively. All these $h(2)$ values suggest that the hydrological processes of the Yangtze River basin are non-stationary. Moreover, the $h(2)$ values together with the relationship between Hurst exponent and the exponent $h(2)$ for small scales, i.e., $H = h(2) - 1$, indicate that the streamflow series of the upper and middle Yangtze River basin (Xu et al., 2006) are characterized by short memory or anti-persistence; while the streamflow series of the lower Yangtze River basin is characterized by long memory or persistence.

(2) To decide the types of multifractality of the streamflow series, the original streamflow series was shuffled. Comparison of $h(q)$ curves of original and shuffled streamflow series indicates that the multifractality properties of the streamflow series are mainly due to correlation characteristics within the hydrological series of the Yangtze River basin.

(3) Different scaling properties for Cuntan, Yichang, Hankou and Datong are probably due to the different topographical characteristics in the upper, the middle and the lower parts of the Yangtze River basin. Different rainfall—landform—runoff relationships in the upper, middle and lower Yangtze River basin may be the major factors inducing different scaling properties of the hydrological processes of the Yangtze River basin. This research is helpful for a better understanding of the scaling properties of hydrological processes in the Yangtze River basin and for regional flood frequency analysis.

ACKNOWLEDGEMENTS

The work described in this paper was supported financially by National Natural Science Foundation of China (Grant No.: 40701015 and 30570426), the Innovative Project of Nanjing Institute of Geography and Limnology, CAS (CXTLS00814), the Outstanding Overseas Chinese Scholars Fund from CAS (The Chinese Academy of Sciences) and by the Fok Ying Tung Education Foundation of China (Grant No.: 101004). Cordial thanks are extended to two anonymous reviewers and the editor-in-chief, Professor Malcolm G Anderson, for their valuable comments and suggestions, which greatly improved the quality of this paper.

REFERENCES


