



EPiC Series in Engineering

Volume 3, 2018, Pages 2310–2317

HIC 2018. 13th International  
Conference on Hydroinformatics



# Investigating the complexity of runoff series in the Yangtze River using sample entropy

Yuankun Wang<sup>1\*</sup>, Dong Wang<sup>1</sup>, Xiaorui Shi<sup>1</sup>

<sup>1</sup>Key Laboratory of Surficial Geochemistry, Ministry of Education, Department of Hydrosociences, School of Earth Sciences and Engineering, State Key Laboratory of Pollution Control and Resource Reuse, Nanjing University, Nanjing, P.R. China  
yuankunw@nju.edu.cn

## Abstract

Based on sample entropy, we investigated spatial distribution and dynamic change of runoff series complexity with the long-term daily runoff series of main stem in the Yangtze River. The results showed that, complexity of the runoff series show an obvious spatial difference, and an increasing trend from upstream to downstream in the Yangtze River. There are negative relationship between average of runoff sliding window and the corresponding sample entropy, and their peak-to-valley value shows the well corresponding relationships. Complexity of the runoff series at Yichang and Datong stations show a continuous increasing trend, while that of Hankou station an increasing trend after 2000. It could provide scientific reference for understanding of runoff series dynamic evolution in the Yangtze River.

## 1 Introduction

Natural flow regimes play a central role in maintaining the health of river, wetland, estuarine and floodplain ecosystems. However, human activities and climate change have significantly altered the hydrological regime of rivers worldwide. Hydrologic system is an open, complex giant system and a dynamic, nonlinear compound system [1]. The stochastic nature of a hydrologic system can be attributed to randomness. Investigation of long-term flow series plays an important role in water resources planning and management in watershed.

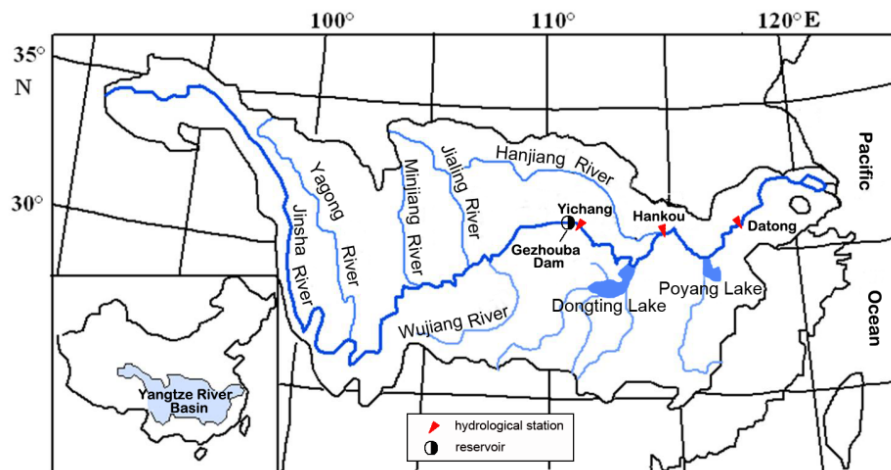
As a measure of the degree of uncertainty (i.e., disorderliness, randomness, or irregularity) or the rate of information production, a greater value of Sample Entropy (SampEn) indicates more complexity in the time series, whereas a lower value of SampEn indicates more orderliness, regularity, or self-similarity. It is well known that the flow formation process in hydrologic basins is complex, and that the flow regime in natural channels is composed of flood and low water with different frequencies, which indicates a high rate of information production [2].

The Yangtze River is the largest and longest river in China and the third largest in annual runoff and the fourth largest in sediment load in the world. Past work mainly focused on flow and sediment load changes, including trend analysis and the impacts of human activities. Jiang et al. (2008) analysed monthly trends in precipitation and discharge for the period 1961–2000 (prior to the operation of the Three Gorges Dam) and found significant upward trends in discharge from the upper catchment, as measured at Yichang hydrological station [3]. Zhang et al. (2009) analysed general trends within long sediment load and streamflow series at the Datong station and investigated streamflow and sediment load variations on different time scales of Yangtze River [4]. Li et al. (2011) analysed the alterations in seasonal and annual sediment load, found that the impacts varied with reservoir storage capacity, operation modes and the distance between the target reservoir and the case study site [5]. Dai et al. (2016) analysed the characteristics of daily suspended sediment concentration (SSC) in the Changjiang estuary from 1956 to 2013 and detected the relationship between SSC and water discharge [6]. Chen et al. (2016) analysed the hydrology regime of the Yangtze River at the monthly time scale using data that cover the period 1955–2014, including 11 years during which the Three Gorges Dam has been completed and brought into full operation [7]. Similar studies also have been conducted in different scales variability of sediment load and streamflow [8–12]. However, the complexity of flow series change and possible reasons in the Yangtze River has yet to be determined.

This study uses sample entropy to characterize the complexity change of runoff series in the main stem of the Yangtze River. Changes in entropy values provide insight into how the complexity has changed, and therefore uncertainty, of hydrological processes

## 2 Study site

The Changjiang River is about 6380 km long with a drainage area of  $1.8 \times 10^6$  km<sup>2</sup> [5, 13] (Figure 1). Originating from the Tibetan Plateau, the terrain of the basin is shaped like a ladder with three stairs. The Qinghai-Tibet Plateau in the west, the highest stair, has an average elevation of over 3000 m above sea level; the second stair, where the Sichuan basin is located, has an average elevation of 1000 m; the third stair in east China Plain has an average elevation of about 100 m. In this study, the basin is divided into three sub-regions that correspond to the three stairs, the upper the middle and the lower regions, respectively. The classification of the upper, middle and lower regions of the basin in this study is different from what is determined by the “Changjiang River Water Resources Commission (CWRC)” in China, where flood control is the main concern for the classification.

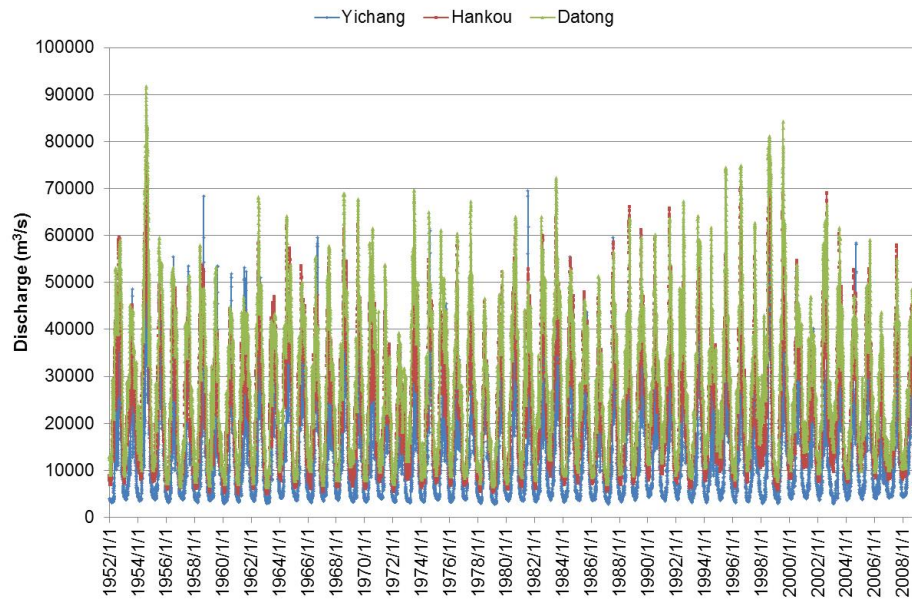


**Figure 1** Location of the study region and hydrological stations.

According to CWRC, the section above Yichang station (where the three gorges dam is located) is called the Upper Reach, 4500 km long, with a controlled catchment area of 1 million km<sup>2</sup> accounting for 70.4% of Yangtze's total area. From Yichang to Hukou is the Middle Reach, 955 km long with a catchment area of 680,000 km<sup>2</sup>. The remaining part from Hukou to the estuary is called the Lower Reach, 938 km long with a catchment area of 120,000 km<sup>2</sup> [13]. Human activities have changed the hydrological processes in the Yangtze River, especially dam construction. According to incomplete statistics, 47, 842 dams and reservoirs have been built in the Yangtze River basin for flood control, navigation, and hydropower generation [14].

### 3 Material and methods

To investigate the complexity of runoff series in the Yangtze River daily flow data for three hydrological stations (Yichang, Hankou and Datong) were analyzed (Figure 1 and Figure 2). Daily data were obtained from the Yangtze (Changjiang) Water Resources Commission for the three stations. Yichang station is located in the upper reaches of the Yangtze River, about 38 km downstream from the TGD, Hankou station is located on the middle reach of the Yangtze River between Dongting and Poyang lakes, and Datong station is located on the lower reach of the Yangtze River between Poyang Lake and the East China Sea (Figure 1). Entropy is used to characterize the complexity of a signal or time series with large entropy values corresponding to high degrees of complexity.



**Figure 2** Flows at Yichang, Hankou, and Datong hydrological stations

A brief summary of SampEn calculation is provided here [15].

For a time series of  $N$  points  $\{x(i) : 1 \leq i \leq N\}$ , SampEn can be calculated as follows:

1 Create an  $m$ -dimensional vector as

$$X(i) = \{x(i), x(i+1), \dots, x(i+m-1)\} \quad (i = 1, 2, \dots, N - m + 1) \quad (1)$$

2 Define the Euclidean distance between  $X(i)$  and  $X(j)$ ,  $d[X(i), X(j)]$  as

$$d[X(i), X(j)] = \text{Max}\{|X(i+k) - X(j+k)|\} \quad 0 \leq k \leq m-1 \quad (2)$$

3 Take  $n_i^m$  as the number of sequences in the time series that match (without self-matching) the template with the length  $m$  within the tolerance criterion  $r$ . Based on this value, the functions  $C_i^m(r)$  and  $C^m(r)$  are defined as:

$$C_i^m(r) = \frac{n_i^m}{N-m+1} \quad (3)$$

$$C^m(r) = \frac{\sum_{i=1}^{N-m+1} C_i^m(r)}{N-m+1} \quad (4)$$

where  $i = 1, 2, \dots, N-m+1$

4 Change the dimension of the vector  $X(i)$  to  $m+1$  and repeat steps 1-3 to yield  $C^{m+1}(r)$ . The SampEn is defined as

$$\text{SampEn}(m, r, N) = -\ln \frac{C^{m+1}(r)}{C^m(r)} \quad (5)$$

SampEn ( $m, r, N$ ) estimates the entropy of a time-series dataset consisting of  $N$  data points by computing the negative natural logarithm of the conditional probability that a given short sequence of data, or template, is repeated during the time series [16]. It is a conditional probability measure that quantifies the likelihood that sequences of  $m$  consecutive data points of the same length that match one another within a tolerance of  $r$  remain similar when the length of the sequences is increased by one sample (sequences of length  $m+1$ ) [17].

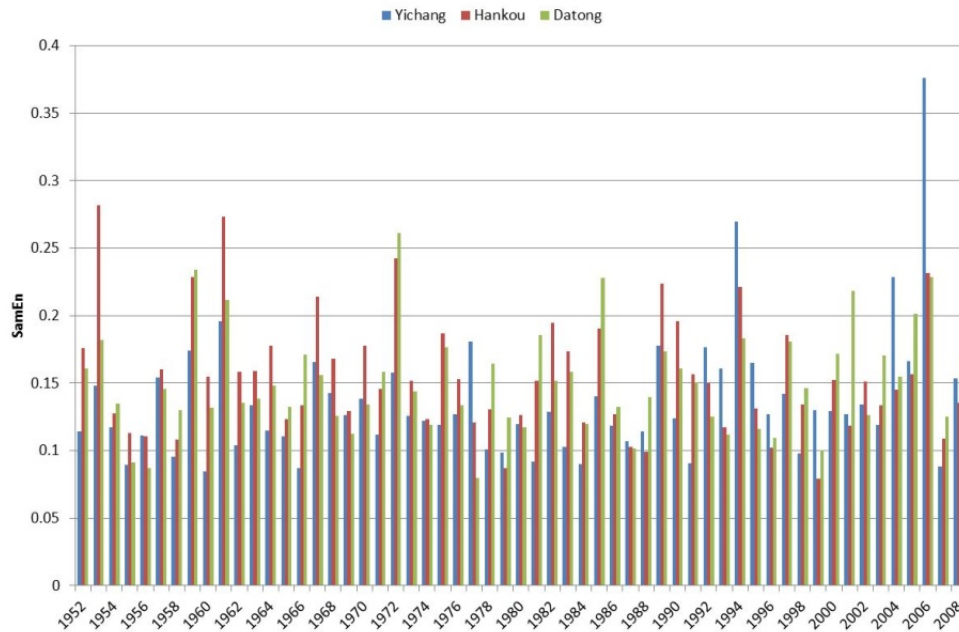
## 4 Results and discussion

### 4.1 Spatial distribution of runoff series complexity

Taking the measured runoff time series data of Yichang station, Hankou station and Datong station as the sample, the sample entropy value of three stations was calculated according to the sample entropy theory. The entropy values of the three stations of Yichang station, Hankou station and Datong station are 0.13, 0.15 and 0.16 respectively. It can be seen from Figure 3 that the sample entropy of the hydrological station in the Yangtze River station is increasing gradually, and the appearance of the sample entropy series and the emergence of the trough are obviously consistent. Before the 1970s, the entropy of the Hankou station in the middle reaches was significantly higher than that of the Yichang station and the Cuntan station. Since the Hankou station was the largest inflow station in the Hanjiang River, the largest tributary of the Yangtze River Basin, the reason for this phenomenon was related to the storage of Danjiangkou Reservoir. The entropy of the Yichang and Hankou stations increased in recent years, which was related to the change of climate and environment in the Yangtze River Basin, the increase of water intake outside the river and the change of the underlying surface caused by human activities.

The Overall, the upper reaches of the Yichang station sample entropy is less than the downstream station value, but after the 2003 part of the change occurred, especially in 2004 and 2006 Yichang station appeared great value, and the Three Gorges reservoir water and weather Drought related. Because the reservoir water storage time of the Three Gorges Reservoir is short, the runoff series after

the impoundment is less, and its influence on the complexity of the runoff sequence is difficult to be fully grasped.



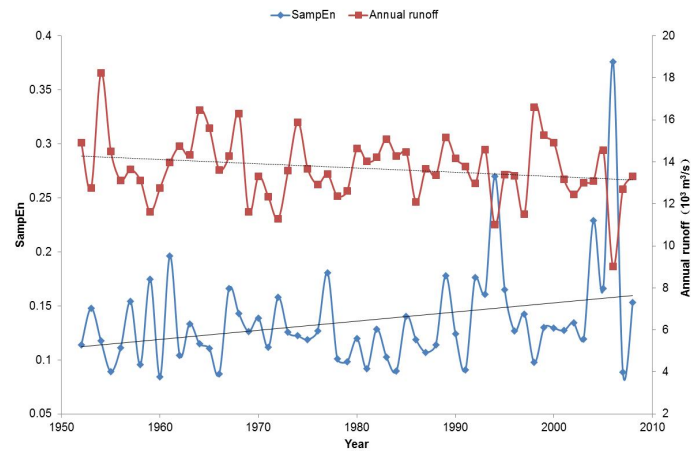
**Figure 3** Comparison of sample entropy values between Yichang, Hankou, and Datong hydrological stations

## 4.2 Temporal distribution of runoff series complexity

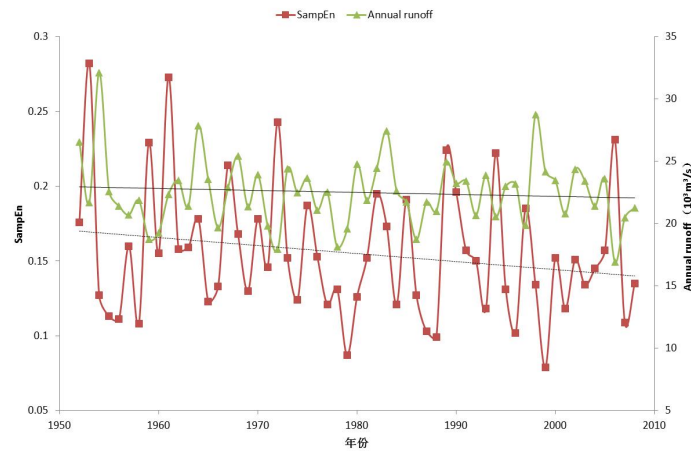
Yichang station sample entropy calculation results shown in Figure 4(a). It can be seen that the entropy of the sample in Yichang station before the 1990s showed an upward trend, and there was a large fluctuation after the 1990s. Analysis of the reasons and found that this and 1994, 2004 and 2006, the major drought occurred. Due to the occurrence of drought, the trend of sample entropy in 1990s was seriously changed. In 2006, the drought in the Yangtze River Basin and the integrated water storage function of the Three Gorges Reservoir made the annual runoff of Yichang station the lowest value, leading to the emergence of the large sample entropy (0.38). It can be seen that under the combined action of human activities and meteorological and environmental conditions, the uncertain components of runoff series are obviously increased, and the structure of runoff series is changed to a certain extent, which makes the complexity of runoff increase.

Hankou station sample entropy calculation results shown in Figure 4(b). It can be seen from Figure 3(b) that the entropy of the Hankou station shows a decreasing trend, and the reservoirs of the Danjiangkou reservoir at the end of 1960s disturb the original sequence structure, so that the entropy of the sample is very large. But after the 20th century showed an increasing trend, which may be related to climate change and climate change in recent years, outside the river to increase the amount of water, human activities caused by changes in the underlying surface.

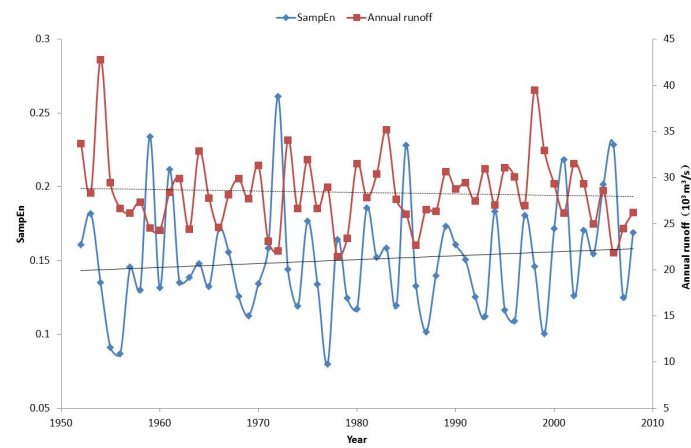
The calculation of the entropy value of the sample is shown in Figure 4(c). It can be known that the entropy change of the sample on the whole is increasing. Datong station is located in the lower reaches of the Yangtze River, is the Yangtze River estuary control station, due to control of catchment area and the upper and lower reaches of different, such as the increase in Dongting Lake, under the surface conditions are different, by the Yangtze River water conservancy project distance and degree of different Increased complexity and uncertainty of runoff sequences.



(a) Yichang



(b) Hankou



(c) Datong

**Figure 4** Variation trend in average runoff and sample entropy values

There is a negative correlation between the sample entropy and the annual runoff of Yichang station, Hankou station and Datong station as a whole. There is a reverse correspondence between the annual runoff and the sample entropy. From the time point of view, the complexity of runoff sequence of Yichang station and Datong station is increasing, and Hankou Station has shown an increasing trend since the beginning of this century. This is the background of global climate change since the 21st century, the intensification of human activities, especially water conservancy construction and river water intake are closely linked, which should be the focus of research work.

## 5 Conclusions

In this paper, the runoff series of Yangtze River is analyzed by using sample entropy theory. The following conclusions can be obtained. (a) The complexity of the runoff sequence of the Yangtze River has shown obvious spatial variability. The peaks - troughs of the sample entropy values are consistent and show an increasing trend from upstream to downstream. (b) In terms of time distribution, the complexity of runoff series of Yichang station and Datong station is increasing. Hankou Station has shown an increasing trend, which is closely related to the background of global climate change since the 21st century, and the intensification of human activities, especially water conservancy construction and river water abstraction. (c) The sample entropy of Yichang station, Hankou station and Datong station showed a negative correlation with annual runoff, and there was a reverse correspondence between annual runoff and peak value of sample entropy.

A number of large dams and reservoirs have been schemed and under construction in the upper stream of the Yangtze River. The combined influences of these dams and reservoirs will affect the hydrological processes greatly. The complexity of runoff series in the middle and lower reaches of the Yangtze River will be influenced, which should be the focus of research work. We hope the outputs presented in this paper could provide a baseline for the assessment of discharge series changes on the large rivers and the health and stability of its ecosystem.

## References

- [1] T. Sun, M. L. Feng. (2013). *Multistage analysis of hydrologic alterations in the yellow river, China*. River. Res. Appl. 29, 991–1003.
- [2] F. Huang, Z. Xia, N. Zhang, Y. Zhang, J. Li. (2011). *Flow-complexity analysis of the upper reaches of the Yangtze River, China*. J. Hydrol. Eng, 16, 914-919.
- [3] T. Jiang, Z.W. Kundzewicz, B. Su. (2008). *Changes in monthly precipitation and flood hazard in the Yangtze River Basin, China*. Int. J. Climatol., 28, 1471–1481
- [4] Q. Zhang, C.Y. Xu, V.P. Singh, T. Yang. (2009). *Multiscale variability of sediment load and streamflow of the lower Yangtze River basin: possible causes and implications*. J. Hydrol. 368, 96–104.
- [5] Q. Li, M. Yu, G. Lu, T. Cai. (2011). *Impacts of the Gezhouba and Three Gorges reservoirs on the sediment regime in the Yangtze River, China*. J. Hydrol. 403, 224–233.
- [6] Z. Dai, S. X. Fagherazzi, J. Mei. (2016). *Decline in suspended sediment concentration delivered by the Changjiang (Yangtze) river into the East China Sea between 1956 and 2013*. Geomorphology, 268, 123-132.
- [7] J. Chen, B. Finlayson, T. Wei, Q. Sun, M. Webber, M. Li, Z. Chen. (2016). *Changes in monthly flows in the Yangtze River, China—With special reference to the Three Gorges Dam*. J. Hydrol, 536, 293-301.

- [8] Q. Hu, S. Feng, H. Guo, G. Chen. (2007). *Interactions of the Yangtze River flow and hydrologic processes of the Poyang Lake, China*. J. Hydrol, 347, 90-100.
- [9] T. Jiang, B. Su, H. Hartmann. (2007). *Temporal and spatial trends of precipitation and river flow in the Yangtze River Basin, 1961–2000*. Geomorphology, 85, 143-154.
- [10] Y. Wang, B. L. Rhoads, D. Wang. (2016). *Assessment of the flow regime alterations in the middle reach of the Yangtze River associated with dam construction: potential ecological implications*. Hydrol. Process, 30, 3949-3966.
- [11] S. L. Yang, Xu, K. H., Milliman, J. D, H. F. Yang, & C. S. Wu. (2015). *Decline of Yangtze River water and sediment discharge: Impact from natural and anthropogenic changes*. Sci. Rep–UK, 5, 12581.
- [12] Y. Wang, N. Zhang, D. Wang, J. Wu, X. Zhang. (2018). *Investigating the impacts of cascade hydropower development on the natural flow regime in the Yangtze River, China*. Sci. Total. Environ, 624, 1187-1194.
- [13] L. Gong, C. Y. Xu, D. Chen, S. Halldin, Y. D. Chen. (2006). *Sensitivity of the Penman–Monteith reference evapotranspiration to key climatic variables in the Changjiang (Yangtze River) Basin*. J. Hydrol, 329, 620-629.
- [14] Y. Wang, B. Rhoads, D. Wang, J. Wu, X. Zhang. (2018). *Impacts of large dams on the complexity of suspended sediment dynamics in the Yangtze River*. J Hydrol, 558, 184–195.
- [15] J.S. Richman, J. R. Moorman. (2000). *Physiological time-series analysis using approximate entropy and sample entropy*. Am. J. Physiol. 278, H2039–H2049.
- [16] Y. Tochigi, N.A. Segal, T. Vaseenon, T.D. Brown. (2012). *Entropy analysis of tri-axial leg acceleration signal waveforms for measurement of decrease of physiological variability in human gait*. J. Orthop. Res. 30, 897-904.
- [17] A. Hameau-Heurtier. (2015). *The multi-scale entropy algorithm and its variants: a review*. Entropy, 17, 3110-3123.