Nature Environment and Pollution Technology
An International Quarterly Scientific Journal

ISSN: 0972-6268	Vol. 12	No. 2	pp. 249-254	2013

Original Research Paper

Application of Sample Entropy on Measuring Precipitation Series Complexity in Jiansanjiang Branch Bureau of China

Liu Dong and Liu Meng
School of Water Conservancy & Civil Engineering, Northeast Agricultural University, 59#, Mucai Street, Gongbin Road, Xiangfang Strict, Harbin, China

Nat. Env. & Poll. Tech.
Website: www.neptjournal.com
Received: 20-2-2013
Accepted: 22-3-2013

Key Words:
Sample entropy
Precipitation series
ArcGIS technology
Jiansanjiang Branch Bureau

ABSTRACT

In order to describe the complexity characteristics of precipitation series in Jiansanjiang Branch Bureau of China, the complexities of monthly average precipitation series were measured by the sample entropy method and the regional differences of precipitation resources system complexity in the farms belonging to Jiansanjiang Branch Bureau were analyzed by utilizing the visualization function of ArcGIS technology. The measure and analysis results show that the complexities of each monthly average precipitation have obvious regional differences in Jiansanjiang Branch Bureau. The monthly average precipitation series complexities in Farm Hongwei and Farm Bawujiu are the strongest; in Farm Daxing, Farm Chuangye, Farm Qianjin, Farm Qinglongshan, Farm Yalvhe, Farm Qianfeng and Farm Shengli are in the middle; and in Farm Qindeli, Farm Qixing, Farm Honghe, Farm Qianshao and Farm Erdaohe are the weakest. Sample entropy can reflect the dynamic change degree of precipitation series and has advantages like less data needing and is stable, which provide a new way to research hydrological series complexity. The research achievements reveal the complexity and its areal variation of local precipitation resources system, and screen boundary parameters for structuring water resources optimization allocation models, and provide scientific basis for rationally utilizing and developing precipitation resources in Jiansanjiang Branch Bureau and even in the whole Sanjiang Plain of China.

INTRODUCTION

Jiansanjiang Branch Bureau locates in Sanjiang Plain hinterland in the northeast of Heilongjiang province in China, which is in 46°49'47" N~48°12'58" N and 132°31'38" E~134°32'19"E. The total area is 12350km^2, with cultivated acreage of 682000hm^2 and it has fifteen large-sized and medium-sized state-run farms. It is, both important grain reserve base and marketable grain production base, in our country, which is the key for realizing the strategic target of 50 billion kilogram food production engineering of Heilongjiang Province, and for ensuring the food security in China. Jiansanjiang Branch Bureau locates in the climate region of mild and humid in Sanjiang Plain which is continental monsoon climate, with mean annual precipitation of 500~600mm (Guo 2009), but the interannual variation of precipitation is large and the annual distribution is very uneven in this area. Together with the long-term influence of various kinds of natural factors (solar cycle, landform, geographical position) and human activities (afforestation and tree planting, hydraulic engineering constructing) have led to the complexity characteristics of precipitation series, which are increasing obviously in Jiansanjiang Branch Bureau. However, the complexities of water resources system were ignored when scholars are studying the optimized utilization of water resources issues at past, which made it difficult to discover the internal information of water resources system adequately and difficult to truly realize the optimized utilization of regional water resources after the water resources optimal allocation. Therefore, it is necessary to diagnose the complexity of precipitation series in Jiansanjiang Branch Bureau in order to provide the basis for the precipitation development trend analysis and the effective use of precipitation resources.

At present, complexity measure methods have complexity statistics measure (Feng 1998), fractal theory (Backes 2012), symbols dynamics (Itokawa 2012), chaos theory (Dhanya 2010), fluctuation complexity (Bates 1993) and so on. But more or less there are several weaknesses like large quantity of data, weak capacity in anti-noise and so on. The length of some hydrological time series may be short, so it is necessary to use the algorithm that can measure the complexity characteristic of series effectively with short data. Sample entropy suggested by Richman et al. (2000) is a good complexity measure method for time series, which needs less data and has some advantages like high consistency and not sensitive for losing data (Bai 2007). And has been applied to some fields widely (Widodo 2011, Alcaraz 2010, Vexler 2010, Ramdani 2009) but less in the complexity research of hydrological series. So we try to measure the complexity of precipitation series based on sample entropy and to analyze the regional differences of precipitation resources system complexity by utilizing the visualization function of ArcGIS technology in Jiansanjiang Branch Bureau. And we hope these can provide scientific guiding for the optimal utilization and management of local water resources and the local agricultural production.

COMPLEXITY MEASURE OF PRECIPITATION SERIES

Data sources: The monthly precipitation data of 14 Farms (there are no precipitation monitoring data in Farm Nongjiang) from 1997 to 2009 were collected from the Meteorological Observatory of Jiansanjiang Branch Bureau of China. And the monthly precipitation series variation curves were drawn (Fig. 1). Fig. 1 shows that the periodic (1 year) variation characteristic of precipitation series of each farm in Jiansanjiang Branch Bureau is obvious.

Preferences: Sample entropy is an improved regularity measure based on approximate entropy but different from approximate entropy because it is a statistic without self matches. Sample entropy means the probability in producing new information of nonlinear dynamic system, and it mainly describes the regularity and complexity of the system quantificationally. The bigger sample entropy value is, the lower the self-similarity of series is, the higher the probability in producing new information is, more complex the series are. On the contrary, the smaller sample entropy value is, the higher the self-similarity of series is, the lower the probability in producing new information is, the simpler the series are. SampEn (m, r, N) can stand for sample entropy, being N the length of the time series. Two input parameters, a run length m and a tolerance window r. SampEn can be defined in Widodo (2011). As we know, the values of parameters m and r are keys for SampEn and according to Pincus S M, the most widely established values for m and r are m = 2 and r between 0.10 and 0.25 times the standard deviation of the original time series. Here, m = 2, r = kSD, k = 0.10~0.25 step by 0.01. SampEn values of each monthly precipitation series under different k values were calculated by the software Matlab R2009a (Fig. 2).

From Fig. 2, one can see that the SampEn values and the rangeability of SampEn values after k values of each precipitation series has the trend of overall decreasing gradually. The rangeability of SampEn values after k values of farm should be confirmed to get the k value and the k value which has the smallest rangeability is the final k value (Yan & Gao 2007). When k is 0.21, SampEn values of each precipitation series compared with SampEn values of k is 0.20, the rangeabilities are 0.67%, 2.24%, 2.39%, 3.05%, 1.13%, 2.77%, 1.67%, 3.43%, 1.02%, 4.01%, 0.06%, 2.42%, 0.97% and 6.38%, which are smaller than 7% and the average rangeability is 2.30% which is the smallest and the most reliable. So k is 0.21 finally, in other words, r = 0.21SD.

Stability analysis: Stable algorithm (Yang 2005) is a kind of algorithm that the micro changes of initial data can only cause the micro changes of the results. For algorithm M, $y = M(x)$, if the micro changes of x can only cause the micro changes of y, then the algorithm M is stable. In another word, for the algorithm M: $y = M(x)$, if there is $\Delta x \geq \Delta y$ under any Δx then we can say that the algorithm M is stable. So we can use this theorem to check the stability of sample entropy in measuring the precipitation series complexity in Jiansanjiang Branch Bureau. Then the data were calculated by Matlab R2009a and the results show that the changes of precipitation series data are always bigger than the change of sample entropy of the precipitation series data in Jiansanjiang Branch Bureau. Therefore, we can say that the sample entropy algorithm is stable in measuring the precipitation series complexity in Jiansanjiang Branch Bureau.

REGIONAL DIFFERENCES OF PRECIPITATION RESOURCES SYSTEM COMPLEXITY

Taking SampEn (2,0.21SD) as the final measure value of each monthly average precipitation series complexity, the complexity order of monthly average precipitation series of each farm in Jiansanjiang Branch Bureau from strong to weak is Farm Hongwei, Farm Bawujiu, Farm Daxing, Farm Yalvhe, Farm Qinglongshan, Farm Qianjin, Farm Shengli, Farm Chuangye, Farm Qianfeng, Farm Qindeli, Farm Qixing, Farm Erdaohe, Farm Honghe, Farm Qianshao. Complexity degree was divided into three levels by the final measure values (Level I: 0.7425-0.8299, Level II: 0.8300-1.0590, Level III: 1.0591-1.2137). The higher the level, the stronger the complexity. In order to see the regional differences of precipitation resources system complexity in Jiansanjiang Branch Bureau directly, the complexity space distribution (Fig. 3) was drawn by utilizing the visualization function of ArcGIS.

From Fig. 3, one can see that the complexities of each monthly average precipitation have obvious regional differences in Jiansanjiang Branch Bureau. The monthly average precipitation series complexities of Farm Hongwei and Farm Bawujiu are in Level III, which explain that the fluctuation range of the two farms are the biggest, the forecasting feasibility are the lowest, the influencing factors are the most, the uncertainty components are the most and the complexities of related precipitation system structures are the strongest.

The monthly average precipitation series complexities of Farm Daxing, Farm Chuangye, Farm Qianjin, Farm

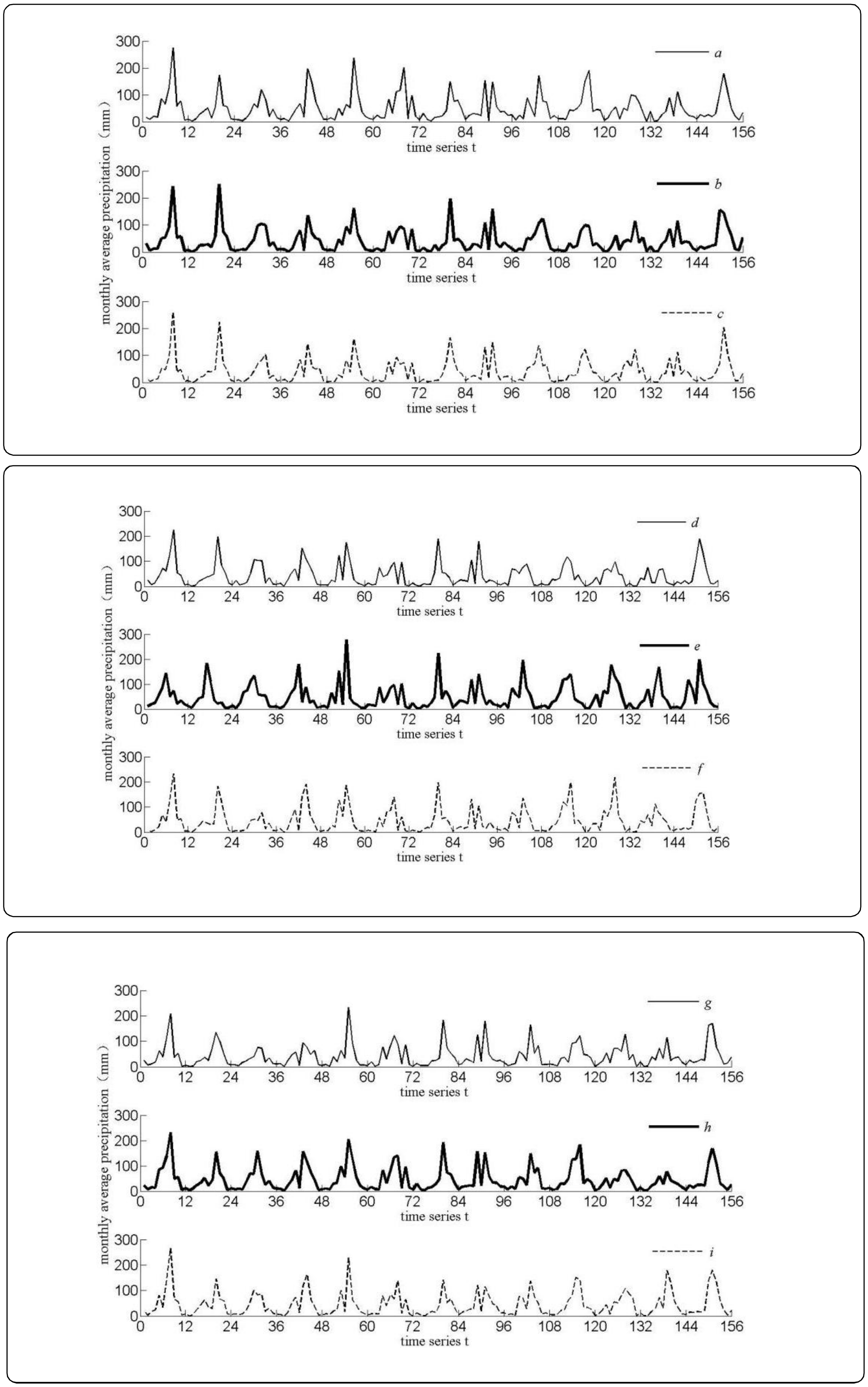
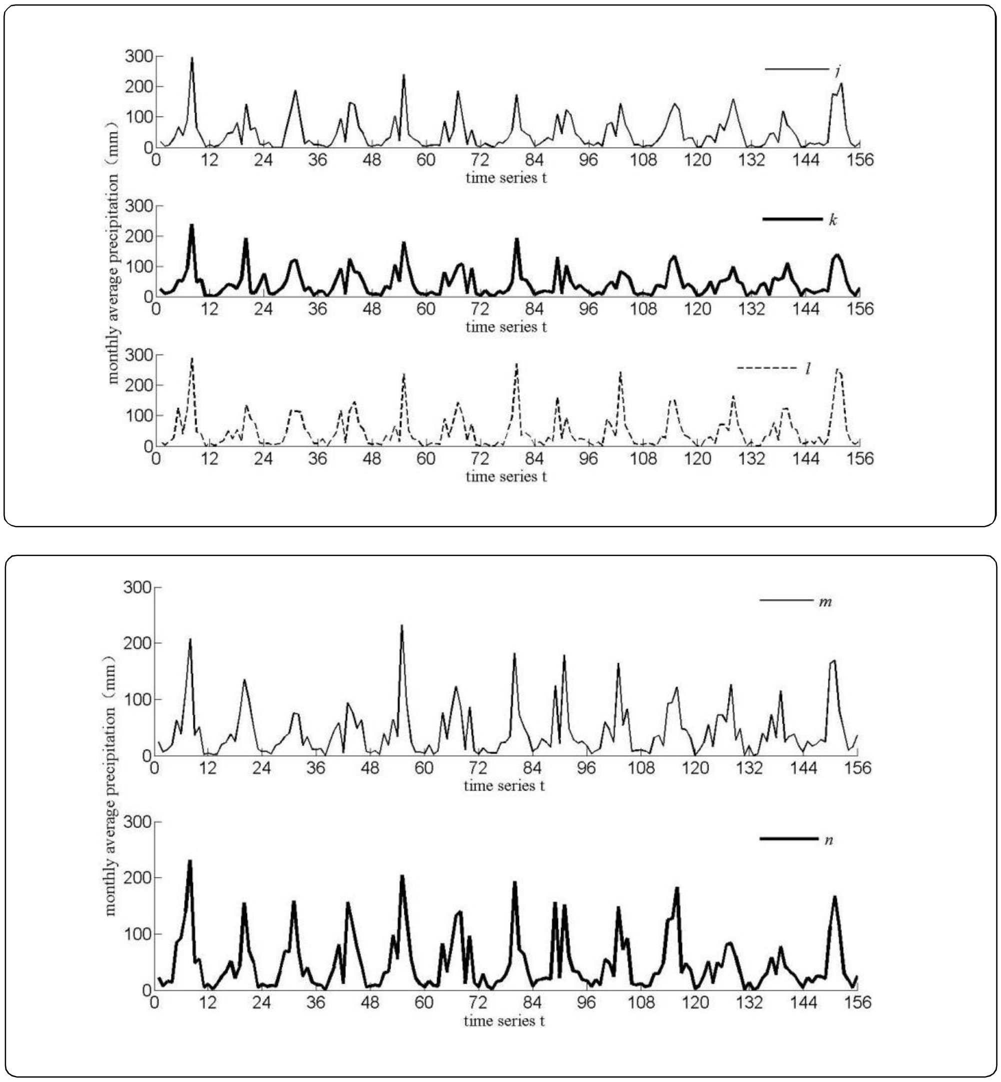

monthly average precipitation（mm）
time series t
a
b
c
d
e
f
g
h
i
0
100
200
300
0 12 24 36 48 60 72 84 96 108 120 132 144 156

Fig. 1 Cont....



a. Farm Bawujiu; b. Farm Chuangye; c. Farm Qixing; d. Farm Qianjin; e. Farm Yalvhe; f. Farm Honghe; g. Farm Daxing; h. Farm Shengli; i. Farm Erdaohe; j. Farm Qianshao; k. Farm Hongwei; l. Farm Qindeli; m. Farm Qinglongshan; n.Farm Qianfeng

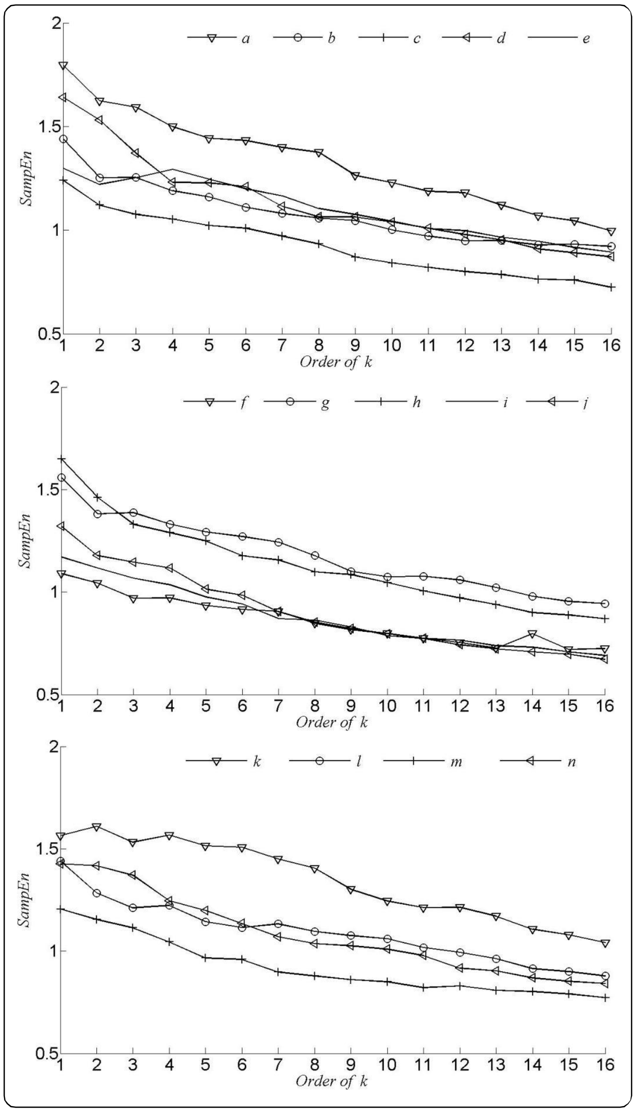
Fig. 1: Variation curves of monthly average precipitation series of each farm in Jiansanjiang Branch Bureau (1997~2009).x

Qinglongshan, Farm Yalvhe, Farm Qianfeng and Farm Shengli are in Level II, which explain that the fluctuation range of the seven farms are in the middle, the forecasting feasibility are in the middle, the influencing factors are in the middle, the uncertainty components are in the middle and the complexities of related precipitation system structures are in the middle.

The monthly average precipitation series complexities of Farm Qindeli, Farm Qixing, Farm Honghe, Farm Qianshao and Farm Erdaohe are in Level I, which explain that the fluctuation range of the two farms are the smallest, the forecasting feasibility are the highest, the influencing factors are the least, the uncertainty components are the least and the complexities of related precipitation system structures are the weakest.

CONCLUSIONS

Precipitation is the key climate factor for crops growing. Analyzing and researching the complexity variation characteristics of precipitation are important to improve the regional rainfall use efficiency and scientific water resources planning and allocation. Sample entropy method can reflect the nonlinearity characteristic of time series only with less data and it is not sensitive for losing data. The algorithm is easy to work and is stable. Sample entropy was introduced in measuring the complexities of monthly average



***a~k, n*: same as Figure 1; *l*. Farm Qinglongshan; *m*. Farm Qindeli**

Fig. 2: Variation curves of *SampEn* after *k* of monthly average precipitation series of each farm in Jiansanjiang Branch Bureau.

precipitation series of each farm in Jiansanjiang Branch Bureau, the complexity sort of monthly average precipitation series of each farm was confirmed and its space distribution was given by utilizing the visualization function of ArcGIS technology. The measure and analysis results show that the complexities of each monthly average precipitation have obvious regional differences in Jiansanjiang Branch Bureau. The monthly average precipitation series complexities in Farm Hongwei and Farm Bawujiu are the strongest; in Farm Daxing, Farm Chuangye, Farm Qianjin, Farm Qinglongshan, Farm Yalvhe, Farm Qianfeng and Farm Shengli are in the middle; and in Farm Qindeli, Farm Qixing, Farm Honghe, Farm Qianshao and Farm Erdaohe are the weakest. Therefore,

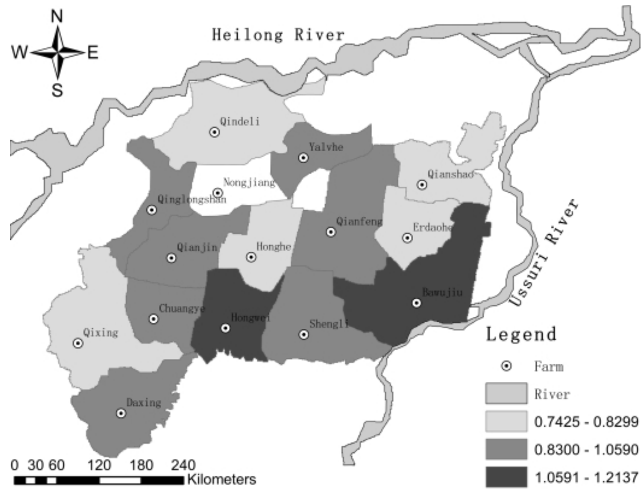


Fig. 3: Complexity space distribution of each monthly average precipitation series in Jiansanjiang Branch Bureau

according to the complexity diagnosis results of precipitation series and the regional actual situation, it should make reasonable water resources utilization policies to provide the reference basis for the sustainable utilization of water resources and the sustainable development of the economic society. In addition, it should structure the regional water resources optimal allocation under the complexity combining with the complexity diagnosis results in the future water resources optimal allocation to truly realize the optimal utilization of regional water resources.

ACKNOWLEDGMENTS

This study is supported by the National Natural Science Foundation of China (No.41071053), Special Fund of China Postdoctoral Science Foundation (No.201003410), Specialized Research Fund for the Doctoral Program of Higher Education of China (No.20102325120009), Natural Science Foundation of Heilongjiang Province of China (No. C201026), Postdoctoral Scientific Research Start-up Fund of Heilongjiang Province of China (No. LBH-Q11154) and Doctoral Start-up Fund of Northeast Agricultural University of China (No.2009RC37).

REFERENCES

Alcaraz R. and Rieta, J.J. 2010. A review on sample entropy applications for the non-invasive analysis of atrial fibrillation electrocardiograms. Biomedical Signal Processing and Control, 5(1): 1-14.

Backes, A.R., Casanova, D. and Bruno, O.M. 2012. Color texture analysis based on fractal descriptors. Pattern Recognition, 45(5): 1984-1992.

Bai, D.M., Qiu, T.S. and Li, X.B. 2007. The sample entropy and its application in EEG based epilepsy detection. Journal of Biomedical Engineering, 24(1): 200-205.

Bates, J.E. and Shepard, H.K. 1993. Measuring complexity using information fluctuation. Physics Letters A, 172(6): 416-125.

Dhanya, C.T. and Kumar, D.N. 2010. Nonlinear ensemble prediction of chaotic daily rainfall. Advances in Water Resources, 33(3): 327-347.

Feng, G.Z., Song, S.B. and Li, P.C. 1998. A statistical measure of complexity in hydrological systems. Journal of Hydraulic Engineering, 11, Article No. 14.

Guo, L., Ma, K.M., and Zhang, Y. 2009. Denitrification potential of different land-use types in Jiansanjiang District. Journal of Agro-Environment Science, 28(5) :941-945

Itokawa, Y., Katoh, K., Uchida, T. and Shoudai, T. 2010. Algorithm using expanded LZ compression scheme for compressing tree structured data. Lecture Notes in Electrical Engineering, 52, DOI 10.1007/978-90-481-3517-2_26.

Ramdani, S., Seigle, B., Lagarde, J., Bouchara, F. and Bernard, P.L. 2009. On the use of sample entropy to analyze human postural sway data. Medical Engineering & Physics, 31(8): 1023-1031.

Richman, J. S. and Moorman, J.R. 2000. Physiological time-series analysis using approximate entropy and sample entropy. Am. J. Physiol. Heart Circ. Physiol., 278(6): H2039-H2049.

Vexler, A. and Gurevich, G. 2010. Empirical likelihood ratios applied to goodness-of-fit tests based on sample entropy. Computational Statistics & Data Analysis, 54(2): 531-545.

Widodo, A., Shim, M.C., Caesarendra, W. and Yang, B.S. 2011. Intelligent prognostics for battery health monitoring based on sample entropy. Expert Systems with Applications, 38(9): 11763-11769.

Yan, R. and Gao, R.X. 2007. Approximate entropy as a diagnostic tool for machine health monitoring. Mechanical Systems and Signal Processing, 21(2): 824-839.

Yang, X.L., Xu, J., Hu, J.C., Zhang, H.K. and Zhang, S.D. 2005. Stability analysis of fast max-min fair bandwidth allocation algorithm. Aerospace Control.