

# Assessment of dam safety using the approximate entropy method

Zhongkai Xie<sup>1,2</sup>, Wenliang Chen<sup>1,2</sup>, Liang Wang<sup>1,2</sup>, Chengfa Deng<sup>1,2</sup> and Chi Guan<sup>3,\*</sup>

<sup>1</sup>Zhejiang Institute of Hydraulics and Estuary, Hangzhou, China

<sup>2</sup>Zhejiang Institute of Marine Planning and Design, Hangzhou, China

<sup>3</sup>School of Management Science and Engineering, Zhejiang Gongshang University, Hangzhou, China

\*Corresponding author e-mail: [guanchi0509@zjgsu.edu.cn](mailto:guanchi0509@zjgsu.edu.cn)

**Abstract.** One of the most important research focuses in today's dam engineering is related to health monitoring of reservoir dams. A novel method named Approximate Entropy (ApEn) is introduced to evaluate the operation condition of gravity dam by monitoring data in this paper, which has been proved to be very successful while applied in many other disciplines, such as the fields of economics, medicals and mechanical, etc. The theoretical framework of ApEn together with a modified programming algorithm are proposed first, and then a quantitative relationship between the horizontal displacement of 14 different measure points located on a practical gravity dam is established. The results of the practical applications have shown that ApEn can effectively and efficiently diagnose the health state of the dam, with high robustness and strong ability to resist noise.

## 1. Introduction

Dam is an important part of flood control system in hydraulic engineering, which also plays a great role in promoting social economic development and improving people's livelihood. However, with the development of dam construction, the safety of reservoir dams (e.g. cracks and leaks etc.) has been constantly exposed. Dam safety is directly related to people's livelihood, once an accident happens, the consequences will be very serious. Hence, it appears particularly necessary to avoid serious safety accidents by proper handling of the potential safety hazard existed in dam engineering [1]. In recent years, the safety evaluation of reservoir dams has been attached great importance. Benefit from the rapid development and popularization of computer technology, dam safety monitoring and analysis technology has been developed rapidly and much improved. Nowadays dam project is facing the challenge of broad in scale as well as more and more complicated construction conditions, new monitoring instrument has been put into application continuously, which puts forward higher requirements for the analysis of monitoring data. On the premise that the calculation model becomes quick and easy to obtain, how to use the appropriate calculation method to objectively and accurately evaluate the operation state of the dam has become an important problem in present dam safety evaluation.



In recent years, a lot of useful research results in dam safety monitoring model has been achieved by many scholars, it can be summed up as [2]: qualitative analysis methods, such as eigenvalue statistic method, drawing method and comparison method, etc.; quantitative analysis methods, such as grey system, neural network, genetic algorithm, wavelet theory, chaos theory, vector machine and other intelligent algorithms. As the capability of numerical computation continually improved, the deterministic relationship between the external environment quantity and the effect quantity acquired by constant optimization fitting of the finite element method (deterministic model or hybrid model) and the actual monitoring value is getting attention by many scholars [3,4]. It is worth noting that finite element method is not omnipotent, the selection of calculation parameters and the determination of boundary conditions are very empirical [5,6]. Wu [7] suggests that the prototype observation method is the most intuitive and reliable method for dam safety monitoring. Although the intelligent algorithms mentioned before have made a great success in dam safety evaluation, the shortcoming should be recognized. For example, gray number algorithm in grey system is difficult to determine, the prediction accuracy is often poor when the sample size is great; the oversized training samples of neural networks often lead to difficulty in convergence problems; the selection of key parameters in genetic algorithm and wavelet theory mainly depends on experience, etc.

Recently, a novel method named Approximate Entropy (ApEn) algorithm [8] draws a large number of researchers' attention due to its advantages of relatively less data required, high anti-noise capability and suitable for both deterministic and random signals, etc. In fact, this method has been applied in many disciplines and has achieved great success [9-12]. ApEn can be used to quantitatively describe the complexity and regularity of time series, so it is possible to get good results by using this algorithm to analyze the dam safety monitoring data and then to evaluate the safety state of the dam. It might be challenging and meaningful to use ApEn as an evaluation tool in dam safety monitoring system, which has never been done before to the best of the author's knowledge.

## 2. Methodology

### 2.1. Approximate entropy

$S$  is the evenly spaced time series of the  $n$  data points for a given data sequence  $\{S\} = \{u(1), u(2), \dots, u(n)\}$ . Firstly, specify the dimension  $m$  and construct a set of  $m$ -dimensional vectors from the original data. And  $1 \leq i \leq n-m+1$ :

$$\mathbf{x}_i = \{u_i, u_{i+1}, \dots, u_{i+m-1}\}, 1 \leq i \leq n - m + 1. \quad (1)$$

Then, define the distance between two different vectors  $\mathbf{x}(i)$  and  $\mathbf{x}(j)$ , and  $1 \leq i, j \leq n-m+1$ :

$$d(\mathbf{x}_i, \mathbf{x}_j) = \max \{ |u_{i+k-1} - u_{j+k-1}| : k = 1, 2, \dots, m \}. \quad (2)$$

Calculate the probability of all the vectors except  $\mathbf{x}(i)$  under a given threshold of  $\varepsilon$  for the specified vector  $\mathbf{x}(i)$ :

$$C_i(m, \varepsilon) = \frac{1}{n - m + 1} \sum_{j \neq i} \Theta \{ \varepsilon - d(\mathbf{x}_i, \mathbf{x}_j) \}. \quad (3)$$

In which,  $1 \leq j \leq n-m+1$ , and  $\Theta$  is the unit step function:

$$\Theta \{ \varepsilon - d(\mathbf{x}_i, \mathbf{x}_j) \} = \begin{cases} 1 : \varepsilon - d(\mathbf{x}_i, \mathbf{x}_j) \geq 0 \\ 0 : \varepsilon - d(\mathbf{x}_i, \mathbf{x}_j) < 0 \end{cases} \quad (4)$$

After getting the natural logarithm of  $C_i(m, \varepsilon)$ , the average is:

$$\Phi^m(\varepsilon) = \frac{1}{n-m+1} \sum_{i=1}^{n-m+1} \ln[C_i(m, \varepsilon)]. \quad (5)$$

Finally, the approximate entropy of the system can be approximated as follows considering the limited data length in actual operation:

$$ApEn(m, \varepsilon, n) = \Phi^m(\varepsilon) - \Phi^{m+1}(\varepsilon). \quad (6)$$

The parameters  $(m, \varepsilon, n)$  of approximate entropy theory was analyzed by Pincus[8]. And it was suggested that the value of approximate entropy was stable when the dimension  $m=2$ ,  $\varepsilon=0.1\sim 0.2SD$  ( $SD$  is the standard deviation of the sequence), and  $n=1000$ . Approximate entropy can be used to measure the regularity of the time series. The value of approximate entropy is smaller when the time series is more regular, and vice versa, such as the approximate entropy value of sine signal is close to zero, and the approximate entropy value of white noise sequence is close to 1.5.

## 2.2. programming

The program can also be improved from two aspects as follows:

The dimension  $m$  becomes  $m+1$  in formula (6), and then the upper limit of the loop variable is turned from  $n-m+1$  to  $n-m$ . It seems that the program has to go through a two-layer loop. However, formula (2) and (3) can be broken up as:  $\max(\text{abs}(\mathbf{x}_i - \mathbf{x}_j))_m \leq \varepsilon$ ,  $\text{abs}(u_{i+m} - u_{j+m}) \leq \varepsilon$  when the dimension is  $m+1$  and  $\max(\text{abs}(\mathbf{x}_i - \mathbf{x}_j))_{m+1} \leq \varepsilon$  is equal to when the dimension is  $m$ . In which,  $\text{abs}(\mathbf{x}_i - \mathbf{x}_j)$  is the distance of each element between recombinant vectors of  $\mathbf{x}_i$  and  $\mathbf{x}_j$  obtained from formula (1), and  $\max(\text{abs}(\mathbf{x}_i - \mathbf{x}_j))$  is the maximum of these distances.

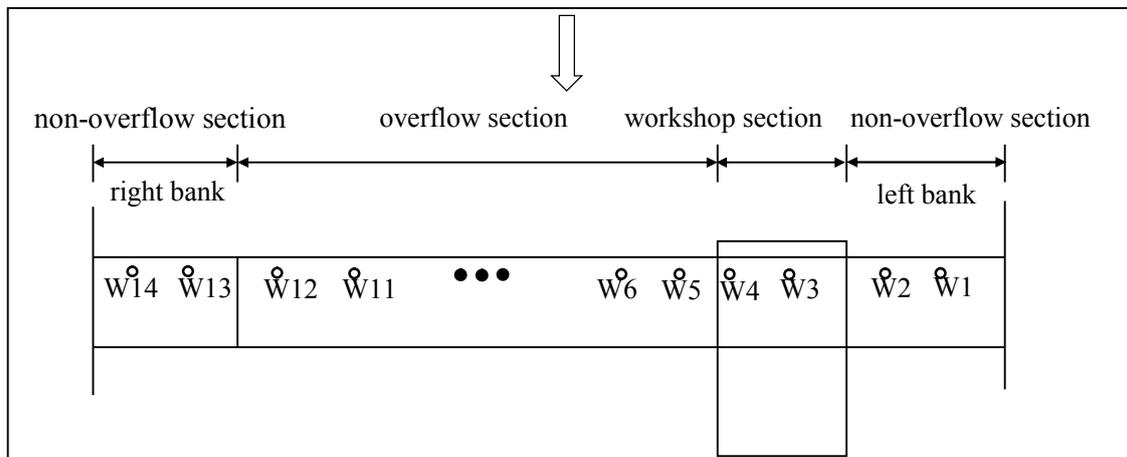
Thus, the two layers of circulation can be turned into a single layer cycle.

(2) Generally, formula (2) indicates that  $d(\mathbf{x}_i - \mathbf{x}_j) \leq \varepsilon$  and  $d(\mathbf{x}_j - \mathbf{x}_i) \leq \varepsilon$  would be judged separately by the procedure. Actually these two conditions are equivalent. Thus, the size of the single-layer loop can be reduced to approximately half of the size of the original cycle.

The algorithm procedure of approximate entropy is reprogrammed based on the above two improvement measures. The entire program can be completed in a single loop instead of two big loops. It has been approved by the author that the computation time would be reduced to 20% of ordinary programming method with 1 000 data points.

## 3. Monitoring items and layout

The concrete gravity dam of Fenshui River hydraulic engineering, with a height of 52.2m and a length of 268.5m, is located in Zhejiang Province in China. Automatic monitoring system of this dam started from June 18th, 2005. Two main items could be obtained from the automatic monitoring system: horizontal displacement and seepage pressure, which are measured with a tension line equipment (14 measuring points, as illustrated in Fig.1, named as W1, W2, W3, ..., W14) and 18 piezometers (most are out of work, not shown in Fig.1) respectively.

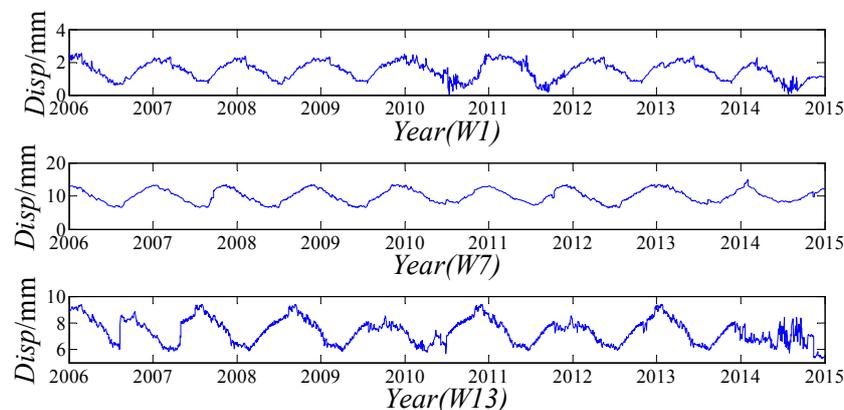


**Fig. 1** Plan view of the dam and monitoring layout

## 4. Results and discussion

### 4.1. Typical horizontal displacement

As mentioned above, 14 measure points of horizontal displacement are relatively evenly arranged on the dam, the data are recorded automatically by the monitoring system day by day. Due to limited space, only three typical groups of horizontal displacement of the dam from the beginning of 2006 to the end of 2014 are chosen, as shown in Fig.2, two (W1 and W13) of which represent the non-overflow sections, and the rest (W7) stands for the overflow section.



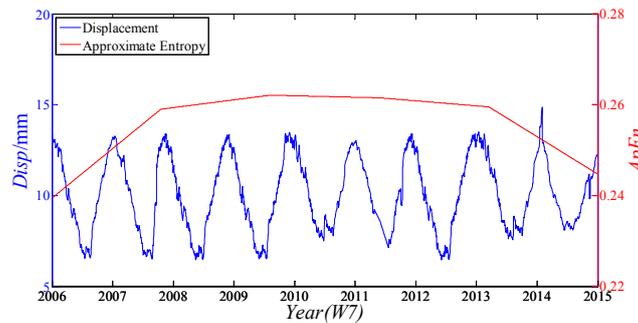
**Fig. 2** Typical horizontal displacement time series of the dam

Fig. 2 shows: (1) Horizontal displacement ( $Disp$  in vertical axis) changes periodically with the seasons, the axis of the dam extended in high temperature seasons, which leads to upstream displacement and vice versa. (2) The largest variation of horizontal displacement is occurred in the overflow section of the dam, which has the maximum height and more likely to be affected by temperature than the non-overflow sections consequently. (3) Overflow section located in the middle of the river bed obtains maximum horizontal displacement, which gradually reduces to both sides.

A qualitative conclusion based on Fig. 2 could be summarized: The horizontal displacement of this engineering is coinciding with the general deformation law of gravity dams. However, the evolution trend of horizontal displacement is still unknown, the current situation of the dam cannot be diagnosed quantitatively.

## 4.2. ApEn analysis

**4.2.1. ApEn of horizontal displacement.** In fact, all the time series of horizontal displacement on the dam have certain regularities, which could be easily and quantitatively measured by ApEn index. Herein, taking horizontal displacement of W7 as an example again, ApEn values, as well as the time series itself are shown in Fig.3 simultaneously.



**Fig. 3** ApEn results of a typical time series

To satisfy the stability and accuracy of ApEn algorithm in this project, the data length should be 1 000 at a minimum via a series of parameter analysis studied by the author. Hence, each ApEn value is calculated by four straight years' data with a length of approximate 1 200(it takes about 3s to complete a calculation on a laptop with a 2.4 GHz CPU and 8 GB RAM). As shown in Fig.3, the ApEn value of time series W7 remains remarkably stable at the level between 0.24 and 0.26, which means the time series of each year possesses high regularity and the evolution trend of measure point W7 is being under good condition.

To get overall information of the safety state of the dam, the ApEn indexes of all measure points are calculated and shown in Table 1 as below. In this table, statistical indicators of ApEn values are also provided, among which  $\mu$  and  $\sigma$  stand for the mean value and standard deviation of ApEn at each point through 2006 to 2015 respectively. From this table, it is pretty obvious that the ApEn indexes fluctuate around mean values of each measure points with relatively small amplitude, which indicates that the dam runs well through these years and the ApEn algorithm is quite stable and robust.

**Table 1.** ApEn value at each measure points from 2006 to 2015

Measure point	Year						Sample statistics	
	2006-2010	2007-2011	2008-2012	2009-2013	2010-2014	2011-2015	$\mu$	$\sigma$
W1	0.463	0.586	0.634	0.631	0.625	0.547	0.581	0.067
W2	0.488	0.480	0.507	0.461	0.380	0.359	0.446	0.062
W3	0.684	0.681	0.596	0.573	0.410	0.401	0.558	0.126
W4	0.309	0.312	0.302	0.274	0.225	0.235	0.276	0.038
W5	0.281	0.349	0.357	0.362	0.399	0.382	0.355	0.041
W6	0.257	0.273	0.285	0.287	0.273	0.271	0.274	0.011
W7	0.239	0.259	0.262	0.262	0.260	0.245	0.254	0.010
W8	0.236	0.255	0.259	0.252	0.256	0.258	0.253	0.009
W9	0.222	0.230	0.239	0.241	0.245	0.235	0.235	0.008
W10	0.249	0.252	0.273	0.272	0.269	0.261	0.263	0.010
W11	0.226	0.254	0.260	0.277	0.298	0.276	0.265	0.025
W12	0.224	0.262	0.260	0.281	0.322	0.300	0.275	0.034
W13	0.571	0.673	0.683	0.714	0.717	0.789	0.691	0.072
W14	0.813	0.796	0.826	0.782	0.780	0.837	0.806	0.024

4.2.2. *Anti-noise ability study.* In practice, unavoidable noise in nature could pollute the monitoring data and therefore the noise resistance ability of ApEn is necessary and worthy of research. Adding different intensity of noise to monitoring data, signal-to-noise ratio(SNR) is defined as.

$$SNR = 10 \log_{10} \frac{\sum X(t)^2}{\sum e(t)^2}. \quad (7)$$

where  $X(t)$  represents original monitoring system, and  $e(t)$  represents noise signal.

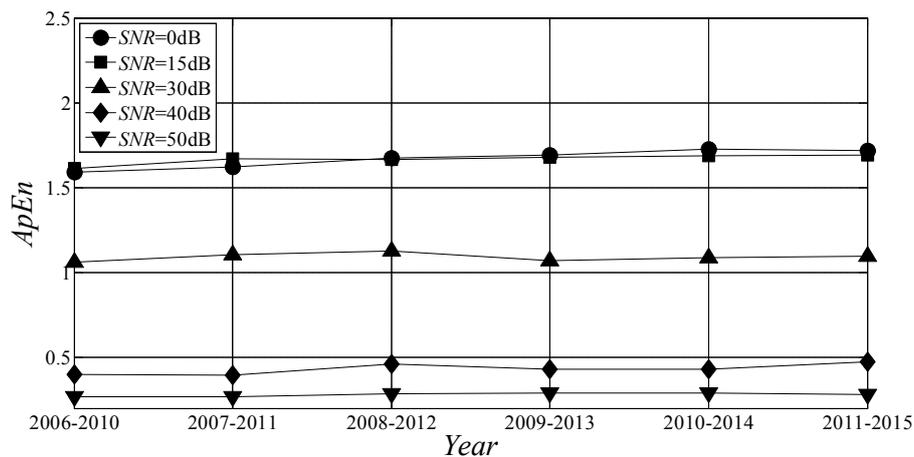


Fig. 4 ApEn of W10 at different SNR

Any one of the time series mentioned above could be applied to validate the anti-noise ability of ApEn algorithm, W10 is applied here as shown in Fig.4. Fig.4 presents the influence on ApEn index when various intensity of noise added to the original data: (1) On vertical direction, bottom to top, bigger intensity of noise leads to higher ApEn values, which means the monitoring data become irregular when noise added to the original data, known as data pollution. (2) On horizontal direction, left to right, the ApEn values remain stable at different level of SNR from beginning to end, which means the ApEn algorithm has powerful anti-noise ability and has great potential in practical application in engineering.

## 5. Conclusion

ApEn index is proposed to quantitatively evaluate the safety state of practical gravity dam via relatively short data length. A modified programming method of ApEn algorithm is briefly presented to accelerate the computation time. Practical engineering research on a gravity dam has shown that the evolution trend of the operation condition of dam can be quantitatively described by ApEn theory, and this method has been proved to be very robust and stable by calculating the ApEn indexes of 14 different monitoring points located on the dam. Under different SNR conditions, the ApEn theory still can accurately capture the real state of the dam and maintain pretty good stability throughout. The efficiency and effectiveness indicate that ApEn is quite suitable for online applications of practical dam engineering.

## Acknowledgments

This work was financially supported by Zhejiang Provincial Water Science and Technology Project (RC1601) and National Natural Science Foundation of China (51508508).

**References**

- [1] J.C. Cao, Application of power engineering technology in the construction of smart grid, *Eng. Technol. Appl.* 3 (2017) 35+56.
- [2] Y.J. Wang, J.B. Zhou, Y. Dong, Current situation and prospect for safety monitoring data analysis of hydropower dams, *Dam Saf.* 5 (2015) 50-57.
- [3] D. Gaffoi, A. Abdulamit and D. Stematiu, Assessment of Gura Raului Dam Safety Using Measurements of Structural Response to Ambient Vibrations, *Procedia Eng.* 161 (2016) 722-728.
- [4] B. Sevim, A. C. Altunisik and A. Bayraktar. Structural identification of concrete arch dams by ambient vibration tests, *Adv. Concr. Constr.* 1 (2013)227-237.
- [5] B. Hu, G.B. Liu, S.J. Wang, X.L. Liu, Evaluation and research on replacement reinforcement quality at abutment of ultra-high arch dam based on prototype monitoring, *J. Hydraul. Eng.* 39 (2011) 876-882.
- [6] H. Z. Su, Z. P. Wen and Z. R. Wu, Study on an intelligent inference engine in early-warning system of dam health, *Water Resour. Manag.* 25 (2011) 1545-1563.
- [7] Z. R. Wu, Safety monitoring theory and its application to hydraulic structures, Higher Education, Beijing, 2003.
- [8] S.M. Pincus, Approximate entropy as a measure of system complexity, *Proc. Natl. Acad. Sci. USA.* 88 (1991)2297-2301.
- [9] D. Sampaio, R. Nicoletti, Detection of cracks in shafts with the Approximated Entropy algorithm, *Mech. Syst. Signal Pr.* 72-73 (2016) 286-302.
- [10] K. Manickam, C. Moore, T. Willard and N. Slevin, Quantifying aberrant phonation using approximate entropy in electrolaryngography, *Speech Commun.* 47 (2005) 312-321.
- [11] J. Cullen, A. Saleem, R. Swindell, P. Burt and C. Moore, Measurement of cardiac synchrony using Approximate Entropy applied to nuclear medicine scans, *Biomed. Signal Proces. Control.* 5 (2010) 32-36.
- [12] G.J. Wang, C. Xie and F. Han, Multi-Scale Approximate Entropy analysis of foreign exchange markets efficiency, *Syst. Eng. Procedia.* 3 (2012)201-208.