

RESEARCH ARTICLE

Monitoring indexes of concrete dam based on correlation and discreteness of multi-point displacements

Xiangnan Qin^{1,2,3*}, Chongshi Gu^{1,2,3*}, Erfeng Zhao^{1,2,3}, Bo Chen^{1,2,3}, Yanling Yu⁴, Bo Dai^{1,2,3}

1 State Key Laboratory of Hydrology-Water Resources and Hydraulic Engineering, Hohai University, Nanjing, China, **2** National Engineering Research Center of Water Resources Efficient Utilization and Engineering Safety, Hohai University, Nanjing, China, **3** College of Water Conservancy and Hydropower Engineering, Hohai University, Nanjing, China, **4** Henan Electric Power Survey & Design Institute, Zhengzhou, China

* Qin_xn@163.com (XQ); csgu@hhu.edu.cn (CG)



OPEN ACCESS

Citation: Qin X, Gu C, Zhao E, Chen B, Yu Y, Dai B (2018) Monitoring indexes of concrete dam based on correlation and discreteness of multi-point displacements. PLoS ONE 13(7): e0200679. <https://doi.org/10.1371/journal.pone.0200679>

Editor: Alessio Alexiadis, University of Birmingham, UNITED KINGDOM

Received: March 13, 2017

Accepted: July 2, 2018

Published: July 17, 2018

Copyright: © 2018 Qin et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Funding: This work was supported by National Key R&D Program of China (2016YFC0401601, 2017YFC0804607), National Natural Science Foundation of China (Grant Nos. 51739003, 51479054, 51779086, 51579086, 51379068, 51579083, 51579085, 51609074), Project Funded by the Priority Academic Program Development of Jiangsu Higher Education Institutions (YS11001), Jiangsu Natural Science Foundation (Grant No.

Abstract

Monitoring indexes are significant for real-time monitoring of dam performance in ensuring safe and normal operation. Traditional methods for establishing monitoring indexes are mostly focused on single point displacements, and rational monitoring indexes based on multi-point displacements are rare. This study establishes monitoring indexes based on correlation and discreteness of multi-point displacements. The proposed method is applicable when several monitoring points show strong correlation. In this study, principal component analysis (PCA) was introduced for preprocessing the observations of multi-point displacements. Correlation and discreteness of multi-point displacements were extracted and constructed. The correlation and discreteness parts described the integral and local variance of the displacement field. On this basis, the annual maximum values of the correlation and discreteness parts were selected and their probability density functions (PDF) could be generated by employing the principle of maximum entropy. PDF was constructed using maximum entropy method and was least subjective because it barely provided the moment information of the observations. The multi-point monitoring indexes were then determined by the typical low probability method based on the obtained PDFs. Finally, the proposed method was analyzed using a practical engineering and was verified in terms of its feasibility.

Introduction

Hydraulic structure is an indispensable infrastructure in society because of its comprehensive functions and benefits in flood control, irrigation, and power generation [1, 2]. Safety monitoring indexes based on prototype observations are frequently adopted in hydraulic structures [3–7], and displacement is one of the major monitored items for dam safety [8]. As the monitoring items and data are becoming various and multitudinous, much more attention has been paid to monitoring indexes considering multiple monitoring variables [9–11]. A concrete dam is a dynamic system exposed to influences of various non-deterministic settings, such as environmental variances, hydraulic loads and geological factors. These non-deterministic factors

BK20160872), Special Project Funded of National Key Laboratory(20145027612, 20165042112), Key R&D Program of Guangxi (AB17195074), the Fundamental Research Funds for the Central Universities (Grant Nos.2016B05323, 2017B617X14, 2016B41314), Central University Basic Research Project (2017B11114), Postgraduate Research & Practice Innovation Program of Jiangsu Province (Grant Nos. KYZZ16_0283, KYCX17_0424). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

affect the overall displacement of the concrete dam versatily [12, 13]. Therefore, monitoring indexes based on the interrelationship of multi-point displacements must be developed.

Various methods for establishing monitoring indexes based on multi-point displacements have been proposed in several references [14–16]. Huang et al. [16] investigated the reasonable form of presetting factors of two-dimensional (2D) space mathematical model with multiple survey points, and they established a statistical model for the Danjiangkou Dam. Yu et al. [15] adopted the principal component analysis (PCA) in multivariate analysis of dam monitoring data and they established the overall control region for dam safety monitoring. Yang et al. [14] determined multi-stage warning indicators for the overall deformation of concrete dam considering fuzziness and randomness, and they achieved nondeterministic optimal control.

Monitoring indexes based on multi-point displacements must be determined based on the correlation of the observations. In the present study, PCA [17–20] was introduced to reduce the dimension of observations and obtain the correlation of the observed displacements. PCA is a popular variable reduction technique [8, 9] and converts a p-vector of the observed displacements into a p-vector of principal components. The correlation and discreteness of multi-point displacements were also extracted and established with PCA in this study.

Thereafter, maximum entropy method (MEM) was performed to generate the probability density functions (PDFs) of the correlation and discreteness parts. According to the principle of maximum entropy, MEM selects the PDF with least subjectivity, thereby maximizing the entropy subject to the moment constraints. MEM has been used to achieve the probability distribution in many fields [21–25]. The monitoring indexes of the correlation and discreteness parts could then be decided using the typical low probability method.

The rest of the paper is organized as follows. Section 1 introduces PCA for extracting the principal components of multi-point displacements. The correlation and discreteness parts are also obtained in this section. Section 2 outlines the MEM for generating the PDFs of the annual extreme values of the correlation and discreteness parts. Section 3 demonstrates how the monitoring indexes of the correlation and discreteness parts can be decided using the typical low probability method based on PDFs. Finally, Section 4 analyzes a numerical example for verifying the feasibility of the proposed method.

PCA for correlation and discreteness of multi-point displacements

False alters, data redundancy and noise effect [15] are the three key problems in dam safety monitoring based on prototype observations. PCA is effective in reducing the frequency of false alarms in consideration of the correlation between numerous observed items, extracting the principal components, realizing data reduction, and reducing the noise effect on the data analysis. Therefore, PCA was introduced in this study in processing the observed multi-point displacements and extracting the correlation and discreteness parts.

1.1 Rationale of PCA for multi-point displacements

For m monitoring points of dam displacement with certain correlation, m simultaneous observations were selected to form the matrix of observed multi-displacement as follows:

$$X = \begin{bmatrix} x_1(t_1) & x_2(t_1) & \cdots & x_m(t_1) \\ x_1(t_2) & x_2(t_2) & \cdots & x_m(t_2) \\ \vdots & \vdots & \ddots & \vdots \\ x_1(t_n) & x_2(t_n) & \cdots & x_m(t_n) \end{bmatrix} \quad (1)$$

where $x_i(t_j)$ ($i=1,2,\dots,m; j=1,2,\dots,n$) is the displacement at t_j of observation point i .

PCA replaces the displacements at m monitoring points with k principal components, $u_1, u_2, \dots, u_k (k < m)$, which contain the vast majority of the information of the original observations. k principal components are the linear combinations of the observed displacements, which can be expressed as follows:

$$\begin{cases} u_1 = l_{11}x_1 + l_{12}x_2 + \dots + l_{1m}x_m = \sum_{i=1}^m l_{1i}x_i \\ u_2 = l_{21}x_1 + l_{22}x_2 + \dots + l_{2m}x_m = \sum_{i=1}^m l_{2i}x_i \\ \vdots \\ u_k = l_{k1}x_1 + l_{k2}x_2 + \dots + l_{km}x_m = \sum_{i=1}^m l_{ki}x_i \end{cases} \quad (2)$$

where l_{ji} are the coefficients of the linear combinations. The principal components must satisfy the following requirements:

1. u_i and $u_j (i \neq j; i, j = 1, 2, \dots, k)$ are independent, i.e. $u_i \bullet u_j^T = 0$;
2. u_i has the maximum variance among all the linear combinations of x_1, x_2, \dots, x_m which are independent of u_1, u_2, \dots, u_{i-1} .

Based on the aforementioned analysis, the PCA for multi-point displacements determines the linear combinations coefficients l_{ji} , and they are the eigenvectors corresponding to the first k largest eigenvalues of the correlation matrix of the multi-point displacements. A two dimensional example was presented as follows to illustrate the mathematical meaning of PCA.

P_1 and P_2 are two observation points with certain correlation. Fig 1 shows the observations of the two points over a certain time period, and Fig 2 illustrates the two principal directions of these two-point observations. Given the correlation between the two observation points, these observations distribute near a straight line and roughly form an ellipse. The coordinate system

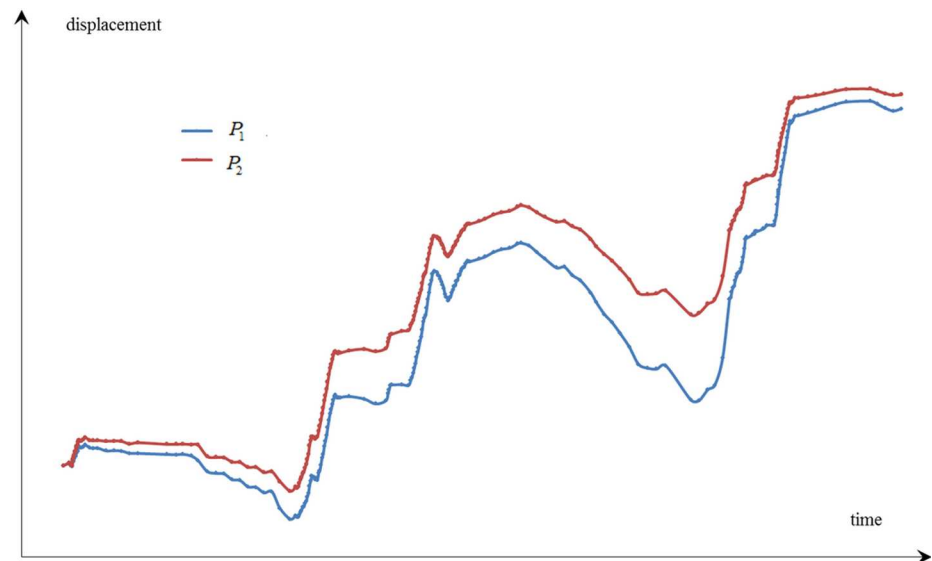


Fig 1. Displacements observations of P_1 and P_2 .

<https://doi.org/10.1371/journal.pone.0200679.g001>

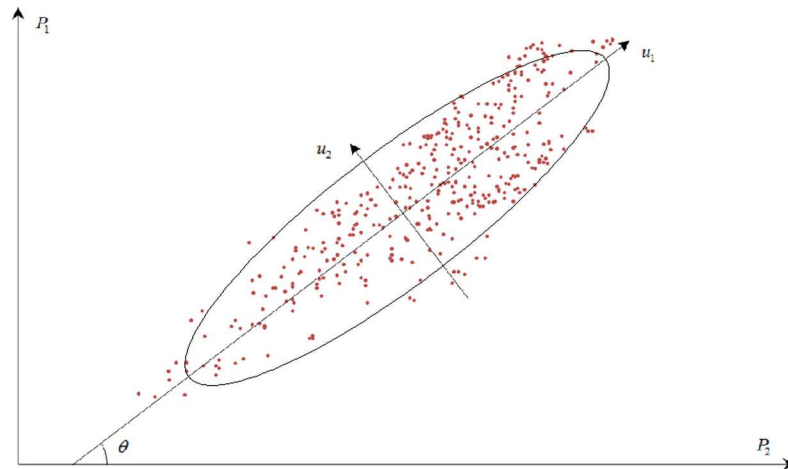


Fig 2. PCA diagram of concrete dam displacements observations in 2D space.

<https://doi.org/10.1371/journal.pone.0200679.g002>

is then rotated with a certain angle θ in 2D space. The major and minor axes of the ellipse are denoted as u_1 and u_2 , respectively. The formula of the rotation is presented as follows:

$$\begin{cases} u_{1j} = x_{1j}\cos\theta + x_{2j}\sin\theta \\ u_{2j} = x_{1j}(-\sin\theta) + x_{2j}\cos\theta \end{cases} \quad (3)$$

where $j = 1, 2, \dots, N$, and N is the times of the observations. The rotation formula can also be expressed in a matrix form as follows.

$$Y = UX \quad (4)$$

where U is the coordination rotation matrix and $U = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}$, such that $U^T = U^{-1}$, $UU^T = I$.

As shown in Fig 1, the major variance of these observations reflects on u_1 axis after the rotation. If the distribution of the multi-point displacements is expressed in one dimension, then u_1 is the best direction that can guarantee the minimum loss of the original observations. The information of the displacement observations on u_1 is regarded as first principal component, which presents the major correlation of the multi-point displacements. Meanwhile, the information on u_2 is regarded as second principal component. When correlation exists between m observation points, the multi-point displacements can be separated in m directions in order of importance with orthogonal transformation in multi dimension. The specific steps of PCA for multi-point displacements are presented in the next section.

1.2 Steps of PCA for multi-point displacements

The steps of PCA for the correlation and discreteness parts of multi-point displacements are as follows:

Step 1. The matrix of multi-point displacement observations was normalized through the following formula:

$$z_{ij} = \frac{(x_{ij} - \bar{x}_j)}{s_j} \quad (5)$$

where $\bar{x}_j = \frac{1}{n} \sum_{i=1}^n x_{ij}$, $s_j^2 = \frac{1}{n-1} \sum_{i=1}^n (x_{ij} - \bar{x}_j)^2$. The normalized matrix was then obtained as follows:

$$Z = \begin{bmatrix} z_{11} & z_{21} & \cdots & z_{m1} \\ z_{12} & z_{22} & \cdots & z_{m2} \\ \vdots & \vdots & \ddots & \vdots \\ z_{1n} & z_{2n} & \cdots & z_{mn} \end{bmatrix} \tag{6}$$

Step 2. The correlation matrix R could be generated through the following formula:

$$R = \frac{1}{n-1} Z^T Z \tag{7}$$

Step 3. The secular equation of the correlation matrix R was presented as follows:

$$|R - \lambda I_p| = 0 \tag{8}$$

The solutions of its secular equation were m eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_m (\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_m \geq 0)$.

Step 4. m components of multi-point displacements were calculated through the following formula:

$$U = ZV^T = \begin{bmatrix} u_{11} & u_{21} & \cdots & u_{m1} \\ u_{12} & u_{22} & \cdots & u_{m2} \\ \vdots & \vdots & \ddots & \vdots \\ u_{1n} & u_{2n} & \cdots & u_{mn} \end{bmatrix} = [u_1 \quad u_2 \quad \cdots \quad u_m] \tag{9}$$

where $V^T = [v_1 \ v_2 \ \cdots \ v_m]$, and v_i is the real eigenvector corresponding to the eigenvalue λ_i .

Step 5. The number of the principal components k could be determined by calculating the cumulative percent variance (CPV) through the following formula:

$$CPV(k) = \frac{\sum_{i=1}^k \lambda_i}{\sum_{i=1}^m \lambda_i} \tag{10}$$

The threshold of CPV, c , should be predetermined. When $CPV(k) \geq c$, the first k components were the principal components of multi-point displacements. c should reach 70%-90% [26].

Based on the aforementioned analysis, m principal components of multi-point displacements could be separated into two parts through the following formula:

$$\begin{aligned} \sum U &= u_1 + u_2 + \cdots + u_k + u_{k+1} + \cdots + u_m \\ &= \sum_{i=1}^k u_i + \sum_{i=k+1}^m u_i = U_p + U_e \end{aligned} \tag{11}$$

where U_p includes the vast majority information of the multi-point displacements; this information shows strong correlation and corresponds to the information on u_1 in Fig 2; meanwhile, U_e is the discrete or even uncorrelated information of the multi-point displacements; this information corresponds to the information on u_2 in Fig 2.

1.3 Analysis of U_p and U_e

The monitoring indexes in this study were determined based on correlation and discreteness of multi-point displacements, corresponding to U_p and U_e .

When U_e changes obviously and U_p remains stable, the discreteness of the multi-point displacements increases and the correlation decreases. The variance of the displacement observations on u_2 also increases. This finding corresponds to the changing process from condition (a) to (c) in 2D space in Fig 3. The multi-point displacements show parallel or consistent variation trend if all the included observation points were analyzed as a whole. The sudden drop in correlation of the displacements observations can illustrate that several individual observations locations were interfered by unknown factors. Therefore, the local abnormality of the multiple observations can be detected if all the observation points are guaranteed to be in proper working order.

When U_p changes obviously and U_e remains stable, the correlation and discreteness parts of multi-point displacements remain stable. Meanwhile, the average value of the displacement observations on u_1 changes significantly. This finding corresponds to the changing process from condition (a) to (b) in 2D space in Fig 3. The integral and consistent change of the multi-point displacements can illustrate that the displacement field in this part changes integrally.

Based on the aforementioned analysis, the monitoring indexes of U_p and U_e can be determined for monitoring the integral and local variation of the multi-point displacements according to their PDFs using the typical low probability method. MEM was introduced to approximate the PDFs of U_p and U_e in the next section.

MEM for the PDFs of correlation and discreteness parts

2.1 Maximum entropy principle (MEP)

The Shannon entropy of discrete random variables was defined by the following formula:

$$H(x) = -\sum_{i=1}^n p_i \ln p_i \tag{12}$$

where p_i is the probability of variable $x_i (i = 1, 2, \dots, n)$; $\sum_{i=1}^n p_i = 1, p_i \geq 0$; $H(x)$ is the information entropy, which expresses the uncertainty of a stochastic system.

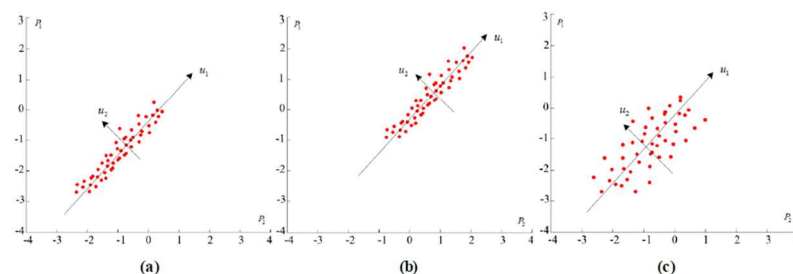


Fig 3. Changing processes related to U_p and U_e in 2D space.

<https://doi.org/10.1371/journal.pone.0200679.g003>

For continuous probability distributions, the Shannon entropy can be defined by the following formula:

$$H(x) = - \int_R f(x) \ln f(x) dx \tag{13}$$

where $f(x)$ is the PDF of the probability distributions, and $\int_R f(x) = 1, f(x) \geq 0$.

Formulas (12) and (13) indicate messages in two conditions. On the one hand, the information entropy can be calculated with the two formulas if the probability p_i or the PDF $f(x)$ is already known; on the other hand, $H(x)$ can be regarded as the functional of the probability p_i or the PDF $f(x)$. $H(x)$ changes with p_i or $f(x)$ and the probability distribution function can be decided with $H(x)$. According to the probability distribution statistical inference rules proposed by Jaynes [27], the probability distribution maximizing the information entropy should be selected when based on partial message. Maximum entropy implies that the contrived hypotheses caused by data deficiencies are the minimum and the obtained solution is the most unaffected and objective. MEM is advantageous because of its simple and rapid calculation.

2.2 Maximum entropy probability density function

According to the principle of maximum entropy, the probability distribution with minimum deviation maximizes the information entropy $H(x)$ subject to certain constraints based on the known samples, which is expressed as follows:

$$\text{MAX } H(x) = - \int_R f(x) \ln f(x) dx \tag{14}$$

subject to

$$\int_R f(x) dx = 1 \tag{15}$$

$$\int_R x^i f(x) dx = \mu_i \quad (i = 1, 2, \dots, N) \tag{16}$$

where R is the domain of integration; $\mu_i (i = 1, 2, \dots, N)$ is the i order origin moment, which can be obtained from the samples.

The first four order origin moments are sufficient in describing the main characteristics of the random variables in many studies [22]. Thus, they were adopted in the present case study. The first origin moment describes the center of the random variables; the second origin moment describes the discreteness of the random variables around the average value; the third origin moment describes the symmetry of the random variables; and the fourth origin moment describes the centralization and decentralization degree of the random variables.

The Lagrange multiplier method was applied to solve this problem and the corresponding Lagrange function was established as follows:

$$L = H(x) + (\lambda_0 + 1) \left[\int_R f(x) dx - 1 \right] + \sum_{i=1}^N \lambda_i \left[\int_R x^i f(x) dx - \mu_i \right] \tag{17}$$

According to stationary value theory ($\partial L/\partial f(x) = 0$), Formula (17) can be transformed into the following expression:

$$\frac{\partial L}{\partial f(x)} = - \int_R [\ln f(x) + 1] dx + (\lambda_0 + 1) \int_R dx + \sum_{i=1}^N \lambda_i \int_R x^i dx = 0 \tag{18}$$

The analytical form of maximum entropy PDF $f(x)$ is presented as follows:

$$f(x) = \exp(\lambda_0 + \sum_{i=1}^N \lambda_i x^i) \tag{19}$$

Formula (19) shows that the solution of maximum entropy PDF can be obtained by determining the Lagrange multipliers $(\lambda_0, \lambda_1, \dots, \lambda_N)$.

Substituting Formula (19) into Formula (15) leads to the following formulas:

$$\int_R \exp(\lambda_0 + \sum_{i=1}^N \lambda_i x^i) = 1 \tag{20}$$

$$\exp(-\lambda_0) = \int_R \exp(\sum_{i=1}^N \lambda_i x^i) \tag{21}$$

$$\lambda_0 = -\ln \left\{ \int_R \exp(\sum_{i=1}^N \lambda_i x^i) \right\} \tag{22}$$

Substituting Formulas (19) and (22) into Formula (16) leads to the following formula:

$$\int_R x^i f(x) dx = \int_R x^i \exp(\lambda_0 + \sum_{j=1}^N \lambda_j x^j) dx = \frac{\int_R x^i \exp(\sum_{j=1}^N \lambda_j x^j) dx}{\int_R \exp(\sum_{j=1}^N \lambda_j x^j) dx} = \mu_i \tag{23}$$

For a more convenient numerical calculation, Formula (23) can be transformed into the following expression:

$$1 - \frac{\int_R x^i \exp(\sum_{j=1}^N \lambda_j x^j) dx}{\mu_i \int_R \exp(\sum_{j=1}^N \lambda_j x^j) dx} = r_i \tag{24}$$

where r_i are the residuals that can approach zero using a numerical technique. A solution of the Lagrange multipliers $(\lambda_0, \lambda_1, \dots, \lambda_N)$ can be generated by non-linear programming to obtain the minimum of the sum of the squared residuals r_i .

$$r = \sum_{i=1}^N r_i^2 \rightarrow \min \tag{25}$$

Convergence is achieved when $r < \epsilon$ or $|r_i| < \epsilon$, where ϵ is the specified acceptable error.

Formula (22) is used to obtain λ_0 with the substitution of the obtained Lagrange multipliers $(\lambda_1, \lambda_2, \dots, \lambda_N)$. Eventually, the maximum entropy PDF $f(x)$ can be determined using Formula (19).

Given that the integration above is difficult to solve by analytical methods, a numerical integration is needed. The domain of integration R should be preprocessed and the upper and lower bounds of the PDFs should be assumed in advance. Considering that the PDF $f(x)$ of random variables x is generally characterized by its thin tails, the integral value of $f(x)$ in the infinite domain $(-\infty, +\infty)$ can be approximated by that in the finite domain (a, b) . The residual error should be reduced to be acceptable if the finite domain (a, b) is sufficiently wide. For example, the residual error is 0.27% if (a, b) was set as $(\mu - 3\sigma, \mu + 3\sigma)$ for random variables x obeying normal distribution. (a, b) was set as $(\mu - 5\sigma, \mu + 5\sigma)$ conservatively in the following numerical example considering that the practical probability distribution may be skewed.

The PDFs of the annual maximum and minimum values of U_p and U_e can be approximated using MEM. The monitoring indexes are determined in the next section.

Monitoring indexes of correlation and discreteness

Based on the PDFs of the annual maximum and minimum values of U_p and U_e , the monitoring indexes $(\delta_{min}^p, \delta_{max}^p)$ and $(\delta_{min}^e, \delta_{max}^e)$ can be determined given certain significance level α [28, 29]. When U_p , the correlation part of observed multi-point displacements, exceeds $(\delta_{min}^p, \delta_{max}^p)$, the integral displacement of the dam is in a warning situation. Therefore, forewarning measures must be adopted against probable danger. When U_e , the discreteness part of observed multi-point displacements, exceeds $(\delta_{min}^e, \delta_{max}^e)$, local abnormality may occur in the observed region. Thus, timely inspection must be adopted to guarantee the safety and forewarning measures must be taken.

The corresponding exceedance probability and monitoring indexes $(\delta_{min}^p, \delta_{max}^p)$ and $(\delta_{min}^e, \delta_{max}^e)$ can be derived with the following formula:

$$\begin{cases} P(\delta > \delta_{max}) = \alpha = \int_{\delta_{max}}^{\infty} f_{max}(\delta) d\delta \\ P(\delta < \delta_{min}) = \alpha = \int_{-\infty}^{\delta_{min}} f_{min}(\delta) d\delta \end{cases} \tag{26}$$

where $f_{min}(\delta)$ and $f_{max}(\delta)$ are the PDFs of the annual maximum and minimum values of U_p and U_e .

The flowchart for the determination of the multi-point displacement monitoring indexes of the correlation and discreteness parts is presented in Fig 4.

Numerical example

The concrete gravity arch dam with variable radii located in the upstream of Qingyi River, a tributary of Yangzi River, is an important part of a comprehensive middle-sized hydropower project. The dam was constructed from August 1958 and was completed in 12 years. The elevation of the dam crest is 126.3 m, the maximum height of dam blocks is 76.3 m, and the dam consists of 28 blocks. The dead water level of the reservoir is 101.0 m and the normal high water level is 119.0 m. The total capacity of the dam is $2.825 \times 10^8 \text{ m}^3$. The vertical observation includes the direct and indirect pendulums, which are mainly buried in dam galleries or dam abutments. The direct pendulums (DP8-UP, DP8-DOWN, DP18-UP, DP18-DOWN, DP26-UP, DP26-DOWN) in 8#, 18#, 26# dam sections (Fig 5), were selected for determining the monitoring indexes. The observations from July 18, 1972 to July 11, 2013 were analyzed. The

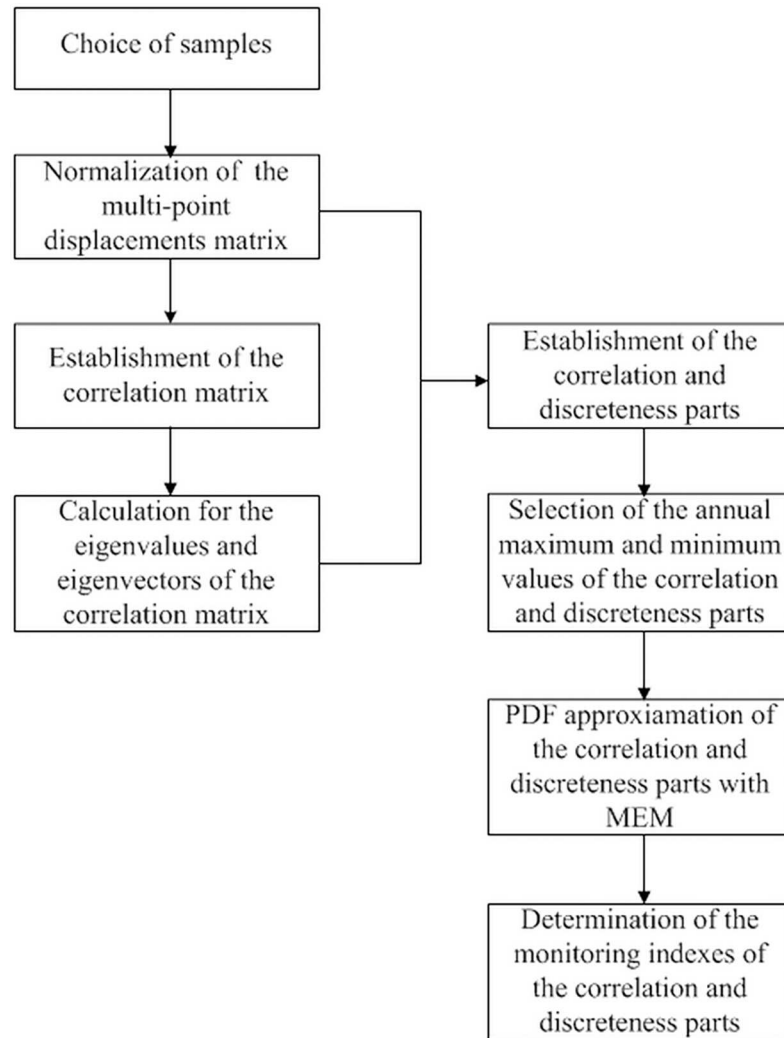


Fig 4. Flowchart for the determination of monitoring indexes of the correlation and discreteness parts.

<https://doi.org/10.1371/journal.pone.0200679.g004>

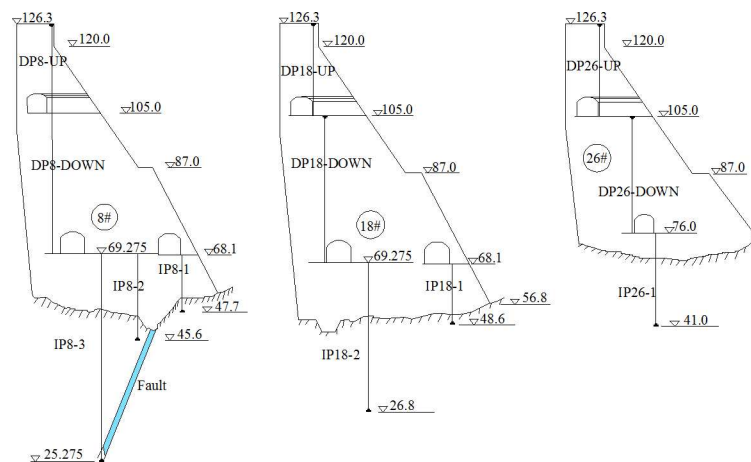


Fig 5. Layout of the direct pendulums in 8#, 18#, 26# dam sections.

<https://doi.org/10.1371/journal.pone.0200679.g005>

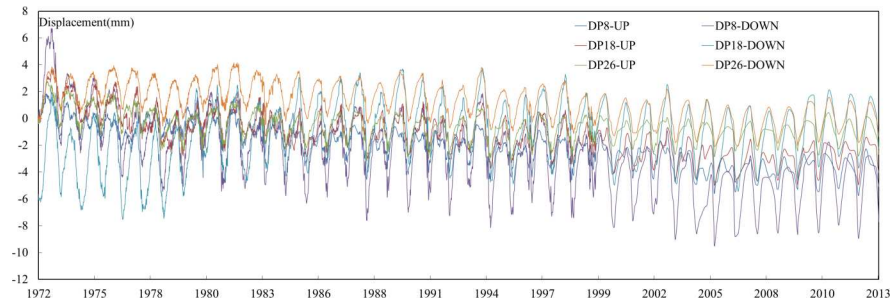


Fig 6. Process lines of the multi-point radial displacements observations.

<https://doi.org/10.1371/journal.pone.0200679.g006>

observations of the multi-point radial displacements and water level are shown in Fig 6. The radial horizontal displacement is positive when it moving towards downstream and negative when it moving towards the upstream.

In this study, the two-stage monitoring indexes of correlation and discreteness were determined according to the practical operation of the project. The significance level α was assumed as 0.05 for the primary monitoring indexes and 0.01 for the secondary monitoring indexes. The reliability coefficients of the monitoring indexes were 95% and 99%, respectively.

4.1 Extraction of correlation and discreteness parts using PCA

PCA was conducted with the observations of the six selected observation points, and the variance contribution percent of the six extracted components are shown in Fig 7. The eigenvalues, variance contribution percent and CPVs are listed in Table 1. The threshold of CPV, c , was set as 90% in this study. The information included in the first two components reached 91.37%, exceeding 90%. Therefore, the correlation part U_p is the sum of the first two components and the discreteness part U_e is the sum of the last four components, which were expressed as follows:

$$U_p = u_1 + u_2 = \sum_{i=1}^2 u_i$$

$$U_e = u_3 + u_4 + u_5 + u_6 = \sum_{i=3}^6 u_i \tag{27}$$

where u_i are the extracted principal components. The process lines of U_p and U_e are shown in Fig 8.

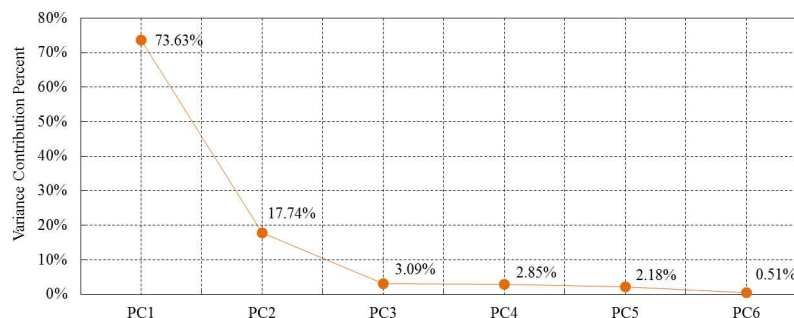


Fig 7. Variance contribution percent of the principal components.

<https://doi.org/10.1371/journal.pone.0200679.g007>

Table 1. Eigenvalues, variance contribution percent and CPVs of the components.

Principal Components	PC1	PC2	PC3	PC4	PC5	PC6
Eigenvalue	1529.94	368.69	64.29	59.19	45.37	10.52
Variance Contribution Percent	73.63%	17.74%	3.09%	2.85%	2.18%	0.51%
CPV	73.63%	91.37%	94.46%	97.31%	99.49%	100.00%

<https://doi.org/10.1371/journal.pone.0200679.t001>

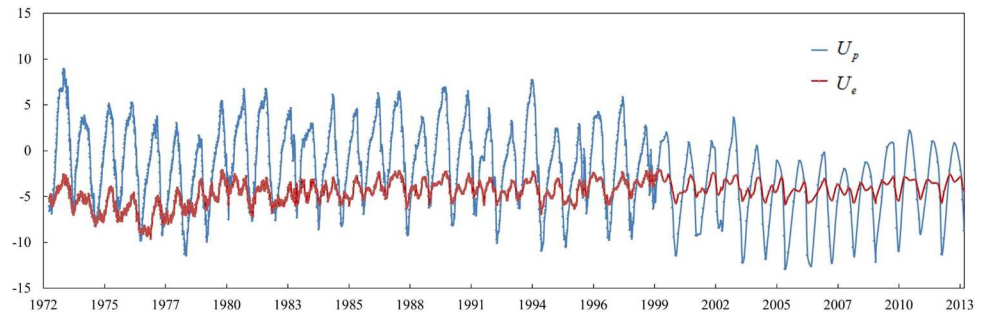


Fig 8. Process lines of correlation and discreteness parts.

<https://doi.org/10.1371/journal.pone.0200679.g008>

4.2 Generation of PDFs of U_p and U_e using MEM

In determining the monitoring indexes ($\delta_{\min}^p, \delta_{\max}^p$) and ($\delta_{\min}^e, \delta_{\max}^e$), the annual extreme values of U_p and U_e from 1972 to 2012 were selected to generate PDFs. The first four moments for the annual extreme values of U_p and U_e are listed in Table 2.

The first four moments for the annual extreme values of U_p and U_e were substituted into Formula (24), and the non-linear least square method was performed to obtain the solution of the five Lagrange multipliers ($\lambda_1, \lambda_2, \lambda_3, \lambda_4$). λ_0 can then be solved by substituting Lagrange multipliers ($\lambda_1, \lambda_2, \lambda_3, \lambda_4$) into Formula (22). Table 3 presents the solutions for these Lagrange multipliers.

Accordingly, the corresponding maximum entropy PDFs for the annual extreme values of U_p and U_e can be expressed as follows:

$$f_{\max}(U_p) = \exp(-2.6149 + 0.2725x - 0.0947x^2 + 0.0245x^3 - 0.0023x^4) \quad (28)$$

$$f_{\min}(U_p) = \exp(-53.6424 - 22.7234x - 3.7109x^2 - 0.2675x^3 - 0.0072x^4) \quad (29)$$

$$f_{\max}(U_e) = \exp(-74.4431 - 37.6518x - 6.9236x^2 - 0.5426x^3 - 0.0157x^4) \quad (30)$$

$$f_{\min}(U_e) = \exp(-29.4397 - 28.0643x - 9.4202x^2 - 1.2896x^3 - 0.0648x^4) \quad (31)$$

4.3 Determination of the monitoring indexes of U_p and U_e

In the case of the correlation part U_p , if $\alpha = 5\%$, then the primary warning monitoring index is (-12.52, 7.52); if $\alpha = 1\%$, then the secondary warning monitoring index is (-13.37, 8.51). In the case of the discreteness part U_e , if $\alpha = 5\%$, then the primary warning monitoring index is (-8.30, -2.07); if $\alpha = 1\%$, then the secondary warning monitoring index is (-9.99, -1.77).

Table 4 presents the two-stage monitoring indexes of U_p and U_e decided using the proposed method and the KS method. The probability density curves of the two methods are shown in Fig 9. The results obtained by the proposed method were close to those by the KS method, and

Table 2. First four moments for the annual extreme values of U_p and U_e .

Moments	U_p		U_e	
	Annual maximum values	Annual minimum values	Annual maximum values	Annual minimum values
0	1.00	1.00	1.00	1.00
1st	3.58	-9.13	-3.00	-6.14
2nd	19.92	87.71	9.43	38.97
3rd	117.04	-881.22	-31.23	-256.61
4th	752.70	9193.06	109.13	1760.01

<https://doi.org/10.1371/journal.pone.0200679.t002>

Table 3. Lagrange multipliers of the PDFs for U_p and U_e .

Lagrange multipliers	U_p		U_e	
	Annual maximum values	Annual minimum values	Annual maximum values	Annual minimum values
λ_0	-2.6149	-53.6424	-74.4431	-29.4397
λ_1	0.2725	-22.7234	-37.6518	-28.0643
λ_2	-0.0947	-3.7109	-6.9236	-9.4202
λ_3	0.0245	-0.2675	-0.5426	-1.2896
λ_4	-0.0023	-0.0072	-0.0157	-0.0648

<https://doi.org/10.1371/journal.pone.0200679.t003>

Table 4. Two-stage monitoring indexes of U_p and U_e .

Part of the multi-point displacements	Method	Confidence level	
		The primary warning monitoring indexes	The secondary warning monitoring indexes
U_p	Proposed method	(-12.52,7.52)	(-13.37,8.51)
	KS method	(-12.61,8.02)	(-14.05,9.90)
U_e	Proposed method	(-8.30,-2.07)	(-9.99,-1.77)
	KS method	(-8.01,-1.91)	(-8.78,-1.46)

<https://doi.org/10.1371/journal.pone.0200679.t004>

the probability density curves of these two methods were relatively consistent. However, the probability distributions approximated by the KS method were derived from the reference distributions and had high subjectivity. MEM generated the probability distributions merely considering the moment information of the multi-point displacement observations and was thus more objective. Therefore, the monitoring indexes of the proposed method were more rational and recommended to be used for the practical monitoring and operation.

Conclusions

This paper presents a method for establishing monitoring indexes of correlation and discreteness of multi-point displacements for concrete dam using PCA and MEM. The correlation and discreteness parts of multi-point displacements were extracted and constructed using PCA, which can describe the integral and local variation trend of the dam displacement. The PDFs of the two parts were approximated by MEM, which is an effective approach to establish a probability density distribution with least subjectivity given a finite number of moments. The monitoring indexes of the two parts could be determined given a certain significance level. The feasibility of the proposed method was demonstrated by a numerical example. The numerical results show that the proposed method could determine rational and accurate multi-point monitoring indexes for concrete dam displacement. A comparison of the results

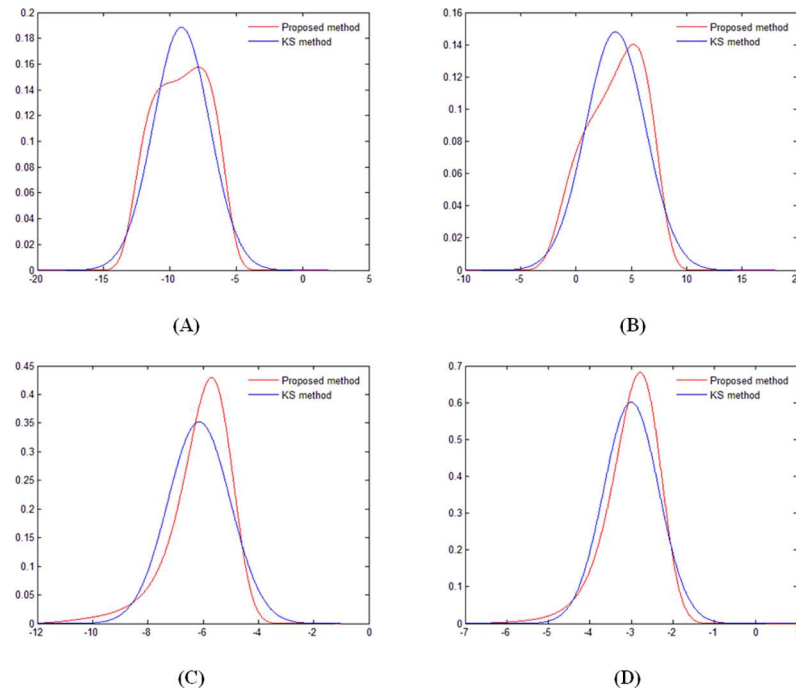


Fig 9. Probability density curves approximated using proposed method and the KS method: (A) Annual minimum values of U_p ; (B) Annual maximum values of U_p ; (C) Annual minimum values of U_e ; (D) Annual maximum values of U_e .

<https://doi.org/10.1371/journal.pone.0200679.g009>

from the proposed method and the KS method confirms the accuracy of the proposed method. The monitoring indexes determined by the proposed method were recommended because of their least subjectivity. However, the accuracy of the monitoring indexes obtained by the proposed method depends on the observation data under most unfavorable load combination, and a structural mechanic analysis is ignored in this study. Further research will be directed to improve the insufficiency and make the monitoring indexes highly rational.

Supporting information

S1 File.
(XLSX)

Acknowledgments

This work was supported by National Key R&D Program of China (2016YFC0401601, 2017YFC0804607), National Natural Science Foundation of China (Grant Nos. 51739003, 51479054, 51779086, 51579086, 51379068, 51579083, 51579085, and 51609074), Project Funded by the Priority Academic Program Development of Jiangsu Higher Education Institutions (YS11001), Jiangsu Natural Science Foundation (Grant No. BK20160872), Special Project Funded of National Key Laboratory(20145027612, 20165042112), Key R&D Program of Guangxi (AB17195074), Central University Basic Research Project (2017B11114), the Fundamental Research Funds for the Central Universities (Grant No. 2017B617X14), Postgraduate Research & Practice Innovation Program of Jiangsu Province (Grant Nos. KYZZ16_0283, KYCX17_0424).

Author Contributions

Conceptualization: Xiangnan Qin, Chongshi Gu.

Data curation: Xiangnan Qin.

Formal analysis: Xiangnan Qin.

Funding acquisition: Chongshi Gu, Erfeng Zhao, Bo Chen.

Investigation: Xiangnan Qin.

Methodology: Xiangnan Qin, Chongshi Gu.

Project administration: Xiangnan Qin.

Resources: Xiangnan Qin, Chongshi Gu.

Software: Xiangnan Qin.

Supervision: Xiangnan Qin, Chongshi Gu, Erfeng Zhao.

Validation: Chongshi Gu, Yanling Yu.

Visualization: Xiangnan Qin, Yanling Yu.

Writing – original draft: Xiangnan Qin.

Writing – review & editing: Chongshi Gu, Erfeng Zhao, Bo Chen, Yanling Yu, Bo Dai.

References

1. Wu ZR, Xu B, Gu CS, Li ZC. Comprehensive evaluation methods for dam service status. *Science China Technological Sciences*, 2012. 55(8): 2300–2312.
2. Wu ZR, Gu CS, Su HZ, Chen B. Review and Prospect of Calculation Analysis Methods in Hydro-Structure Engineering. *Journal of Hohai University (Natural Sciences)*, 2015(05): 395–405.
3. Wieland M. & Kirchen G.F. Long-term dam safety monitoring of Punt dal Gall arch dam in Switzerland. *Frontiers of Structural & Civil Engineering*, 2012. 6(1): 76–83.
4. Isidori D, Concettoni E, Cristalli C, Soria L, & Lenci S. Proof of concept of the structural health monitoring of framed structures by a novel combined experimental and theoretical approach. *Structural Control and Health Monitoring*, 2016. 23(5): 802–824.
5. Cheng L. & Zheng DJ. Two online dam safety monitoring models based on the process of extracting environmental effect. *Advances in Engineering Software*, 2013. 57: 48–56.
6. Su HZ, Wu ZR, Gu YC, Hu J, Wen ZP. Game model of safety monitoring for arch dam deformation. *Science in China Series E: Technological Sciences*, 2008. 51(S2): 76–81.
7. Gu CS, Wang YC, Peng Y, Xu BS. Ill-conditioned problems of dam safety monitoring models and their processing methods. *Science China Technological Sciences*, 2011. 54(12): 3275–3280.
8. Wang SW, Gu CS, and Bao TF, Safety Monitoring Index of High Concrete Gravity Dam Based on Failure Mechanism of Instability. *Mathematical Problems in Engineering*, 2013. 2013: 1–14.
9. Yang YC & Nagarajaiah S. Blind denoising of structural vibration responses with outliers via principal component pursuit. *Structural Control & Health Monitoring*, 2014. 21(6): 962–978.
10. Yang YC & Nagarajaiah S., Data Compression of Structural Seismic Responses via Principled Independent Component Analysis. *Journal of Structural Engineering*, 2014. 140(7): 04014032.
11. Yang YC & Nagarajaiah S. Structural damage identification via a combination of blind feature extraction and sparse representation classification. *Mechanical Systems & Signal Processing*, 2014. 45(1): 1–23.
12. Spiridonov Y.V. Set of geodetic observations of overall displacements of the dam for the Sayano-Shushenskoe hydroelectric power plant. *Hydrotechnical Construction*, 1998. 32(9): 536–539.
13. Zhao Q & Huang SX, A Way for Overall Analysis on Dam's Displacement. *Geomatics and Information Science of Wuhan University*, 2009(12): 1419–1422.
14. Yang G & Yang M. Multistage Warning Indicators of Concrete Dam under Influences of Random Factors. *Mathematical Problems in Engineering*, 2016. 2016: 1–12.

15. Yu H, Wu ZR, Bao TF, Zhang L. Multivariate analysis in dam monitoring data with PCA. *Sci China Tech Sci*, 2010(07): 830–839.
16. Huang M & Li ZZ, Research on the Two-dimensional Space Mathematical Model of Multiple Survey Points for Gravity Dam Displacement. *Engineering Journal of Wuhan University*, 1997(01): 3–7.
17. Datri L.A, Faggi A.M, Gallo L.A, Carmona F. Half a century of changes in the riverine landscape of Limay River: the origin of a riparian neoecosystem in Patagonia (Argentina). *Biological Invasions*, 2016. 18(6): 1713–1722.
18. Okkan U. & Inan G. Statistical downscaling of monthly reservoir inflows for Kemer watershed in Turkey: use of machine learning methods, multiple GCMs and emission scenarios. *International Journal of Climatology*, 2015. 35(11): 3274–3295.
19. Enki D.G, Trendafilov N.T, Jolliffe I.T. A clustering approach to interpretable principal components. *Journal of Applied Statistics*, 2013. 40(40): 583–599.
20. Huang B, Koh B.H, Kim H.S. PCA-based damage classification of delaminated smart composite structures using improved layerwise theory. *Computers & Structures*, 2014. 141: 26–35.
21. Kang H.Y & Kwak B.M. Application of maximum entropy principle for reliability-based design optimization. *Structural and Multidisciplinary Optimization*, 2009. 38(4): 331–346.
22. Deng J, Li XB, Gu DS. Probability Distribution of Rock Mechanics Parameters by using Maximum Entropy Method. *Chinese journal of Rock Mechanics and Engineering*, 2004(13): 2177–2181.
23. Singh VP, Asce F. Hydrologic Synthesis Using Entropy Theory: Review. *Journal of Hydrologic Engineering*, 2011, 16 (5):421–433
24. Li CQ, Wang W, Wang S. Maximum-Entropy Method for Evaluating the Slope Stability of Earth Dams. *Entropy*, 2012. 14(12): 1864–1876.
25. Zhou RX, Cai R, Tong GQ. Applications of Entropy in Finance: A Review. *Entropy*, 2013. 15(11): 4909–4931.
26. Zhao EF. Multi Scale Analysis Methods of Gravity Dam Service Stability under Complex Interaction Factors [dissertation]. Nanjing, China: Hohai University; 2011
27. Jaynes E.T. Information Theory and Statistical Mechanics. *Physical Review*, 1957. 106(4): 620–630.
28. Gu CS. Safety Monitoring of Dams and Dam Foundations-Theories & Methods and Their Application. 2nd ed. Nanjing, China: Hohai University Press; 2006
29. Wu ZR. Safety Monitoring & Its Application of Hydraulic Structures. 2nd ed. Beijing, China: Higher Education Press; 2003