Research Article

Method of Fusion Diagnosis for Dam Service Status Based on Joint Distribution Function of Multiple Points

Zhenxiang Jiang¹ and Jinping He^{1,2}

¹*School of Water Resources and Hydropower Engineering, Wuhan University, Wuhan 430072, China*

²*State Key Laboratory of Water Resources and Hydropower Engineering Science, Wuhan University, Wuhan 430072, China*

Correspondence should be addressed to Jinping He; whuhjp@163.com

Received 13 April 2016; Accepted 31 July 2016

Academic Editor: Eric Florentin

Copyright © 2016 Z. Jiang and J. He. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The traditional methods of diagnosing dam service status are always suitable for single measuring point. These methods also reflect the local status of dams without merging multisource data effectively, which is not suitable for diagnosing overall service. This study proposes a new method involving multiple points to diagnose dam service status based on joint distribution function. The function, including monitoring data of multiple points, can be established with t-copula function. Therefore, the possibility, which is an important fusing value in different measuring combinations, can be calculated, and the corresponding diagnosing criterion is established with typical small probability theory. Engineering case study indicates that the fusion diagnosis method can be conducted in real time and the abnormal point can be detected, thereby providing a new early warning method for engineering safety.

1. Introduction

A dam is a kind of important infrastructures that form a reservoir, and its safety, which is a major public safety issue, is related not only to reservoir benefits but also to human lives, national economic development, social stability, and many other aspects. However, some dams confront some security risks and crash risks because of the complexity of their surroundings and because of other deficiencies in management and engineering; in addition, dam burst events have also occurred occasionally in the world [1, 2].

Dam safety monitoring is an important means to ensure the normal operation of dams. The service status of dams can be reflected and the abnormal situation can be detected by analyzing the effect variable of dams, such as deformation, seepage, and stress, thereby providing a useful way to monitor these structures and to provide an early warning system [3].

At present, the work of analyzing monitoring data and diagnosing dam status mainly relies on the monitoring model of single point. Since the Italian scholar Tonini [4] described the main factors that affected dam deformation as hydraulic press, temperature, and time, in 1956, further studies aimed at

the classical monitoring model of single point have been conducted and have been widely used in practical engineering, forming the statistical model [5, 6], the deterministic model [7], and the hybrid model [8]. Meanwhile, with the development of machine learning algorithms [9, 10], the modeling approaches of single point have developed into the nonlinear direction. Stojanovic et al. [11] proposed the artificial neural networks to establish the nonlinear monitoring model between environmental variables and dam displacement, improving the adaptability of model. Ranković et al. [12] used support vector machine to train the original monitoring data, improving the forecast accuracy of model in the nonlinear situation. Some scholars also introduced intelligence algorithms, such as particle swarm optimization [13], ant colony optimization algorithm [14], and genetic algorithm [15], to construct the monitoring model of single point, thereby enriching the modeling approach and achieving accurate result.

However, the monitoring model of single point is only a reflection of the local structure of a dam, and the overall status of dam cannot be described easily by this method. In general, the monitoring information of each measuring point has

a strong correlation. Thus, the overall service state of a dam should also be diagnosed through the fusion of multipoint monitoring information. The current research in this area is relatively scarce. He et al. [16] introduced D-S evidence theory to realize the fusion diagnosis for dam on the basis of expert evaluations. De Sortis and Paoliani [17] used the modulus of elasticity as diagnostic indicator and diagnosed dam behavior on the basis of finite element model and monitoring data. Su et al. [18] used fractal dimension as diagnostic indicator and applied rescaled range analysis to fuse multipoint monitoring information. In addition, some fusion diagnosis methods in other engineering fields are also worth learning. Rafiq et al. [19] proposed Bayesian networks to diagnose the bridge behavior according to expert evaluations. Masoumi et al. [20] constructed finite element model of steel structure and introduced the optimization algorithm to improve diagnostic precision about the modulus of elasticity. Georgoulas et al. [21] diagnosed the rotor bar with Markov model on condition that the monitoring data obey the normal distribution.

Although a useful attempt to diagnose dam services was derived from the results of these studies, some shortcomings remain; for instance, expert evaluations are always subjective, and the diversity of different experts may affect the diagnosis. Then the diagnostic model based on these methods regards the characteristic parameter (such as fractal dimension and modulus of elasticity) as diagnosis basis. Therefore, making a real time diagnosis for the whole dam based on the measuring values of each point in each monitoring day is impossible. In addition, the monitoring values should presumably obey the normal distribution when the aforementioned methods are used to conduct diagnosis of whole dam. However, these values do not strictly obey the normal distribution. In particular, the distribution is not normal when the abnormal data exist in the monitoring series. Therefore, the fusion diagnosis methods for dam behavior under the multiple points need to be further studied.

This study aims at the aforementioned shortcomings and proposes a new fusion diagnostic method related to joint distribution function based on the in situ monitoring data of dams. The distribution of a single point can be calculated with kernel density estimation (KDE), thereby obtaining the relatively real distribution of each point. Then, the distribution of different single points can be connected to the multidistribution function with t-copula function, and the possibility of measuring values in different combinations can be calculated, thereby providing benefits in analyzing the synergetic changing feature of multiple points. Then, the diagnostic criterion is established with the typical small probability method, and the diagnostic process is described, thereby providing a real time and efficient way to diagnose the overall service status of dams. In addition, the proposed method can also be applied to the fusion diagnosis of other engineering fields.

2. Research Methods

2.1. Single Point Distribution. In general, the monitoring data of single point are considered normal distribution to facilitate analysis and calculation [22]. However, deviations always

exist between the true distribution and normal distribution. The use of approximation may ignore the real characteristic of the monitoring data, thereby leading to calculation error. KDE [23] uses the sample data to estimate the probability density function of the population distribution. The advantage of this method is that the estimation depends on the sample data without a priori assumptions. Therefore, the real trait of monitoring data, which reflect the real feature of single point, can be retained. The expression of KDE can be written as follows:

$$f(x) = \frac{1}{nh} \sum_{i=1}^n K\left(\frac{x - X_i}{h}\right), \quad (1)$$

where $f(x)$ is the probability density function of a single point (S-PDF), X_i is the measuring value of the sample $\{X_1, X_2, \dots, X_n\}$, n is the number of samples, $K(\cdot)$ is the kernel function, and h is the bandwidth. Equation (1) indicates that the accuracy of KDE is affected by both kernel function and bandwidth. In general, a different kernel function is regarded to have a slight effect on the S-PDF. Gauss kernel function, which is always applied to engineering, is expressed as

$$K\left(\frac{x - X_i}{h}\right) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(x - X_i)^2}{2h^2}\right). \quad (2)$$

Therefore, the precision of S-PDF mainly relies on the bandwidth, which is a key factor determining the shape and smoothness of the S-PDF curve. The cumulative distribution function of single point (S-CDF) can be obtained with the integration of S-PDF. The expression of S-CDF is presumably $F(x)$, and the measuring value is presumably X in one observation; thus, the S-CDF value, indicating that the possibility (P) of X is less than x , can be expressed as

$$P(X \leq x) = F(x) = \int_{-\infty}^x f(x) dx, \quad (3)$$

where $0 \leq F(x) \leq 1$. $X_{(1)} \leq X_{(2)} \leq \dots \leq X_{(n-1)}$ is assumed to be the order statistics of the sample to compare the accuracy in a different bandwidth, and the empirical cumulative distribution function of a single point (S-ECDF) can be written as

$$\hat{F}(x) = \begin{cases} 0 & x \leq X_{(1)} \\ \frac{k}{n} & X_{(k)} \leq x \leq X_{(k+1)} \\ 1 & x > X_{(n)} \end{cases} \quad k = 1, 2, \dots, n-1. \quad (4)$$

The fitting goodness can be estimated with root mean square error (RMSE) by comparing the S-CDF value with the S-ECDF value in the sample values X_i , and low RMSE value means high fitting accuracy. The expression of RMSE is written as

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^n (F(X_i) - \hat{F}(X_i))^2}{n}}. \quad (5)$$

2.2. Multipoint Distribution Based on t-Copula Function. The S-PDF and S-CDF only reflect the operational state of one point. However, the probability density function of multiple points (M-PDF) and the cumulative distribution function of multiple points (M-CDF) should be constructed based on S-PDF and S-CDF to diagnose the overall service status of dams. Sklar theorem [24, 25] provides a theoretical basis to generate the M-CDF based on nonnormal S-CDF. If the number of points is m and the expressions of their S-CDF are $F_1(x_1), F_2(x_2), \dots, F_m(x_m)$, then a unique copula cumulative distribution function (C-CDF) $C(\cdot; \theta)$ connecting the S-CDFs to M-CDF $F(x_1, x_2, \dots, x_m; \theta)$ exists. For the measuring values $(X_1, X_2, \dots, X_m)$, M-CDF can be expressed as

$$\begin{aligned} F(x_1, x_2, \dots, x_m; \theta) \\ = P(X_1 \leq x_1, X_2 \leq x_2, \dots, X_m \leq x_m) \\ = C[F_1(x_1), F_2(x_2), \dots, F_m(x_m); \theta], \end{aligned} \quad (6)$$

where θ is the parameter of copula function and the domain of $C(\cdot; \theta)$ is $[0, 1]^n$. For any value $(F_1(x_1), F_2(x_2), \dots, F_m(x_m))$ in the domain, the expression $0 \leq C(\cdot; \theta) \leq 1$ is established.

Sklar theorem has shown M-CDF can be decomposed into several S-CDFs and one C-CDF that describes the relevant information among these variables. C-CDF has many forms. In this study, t-copula function is chosen as the C-CDF because of its symmetrical tail and sensitive feature in capturing tail correlation of variables. For t-copula function, the parameters are ρ (one matrix) and k (one constant), which describe the relationship of various points and degrees of freedom in the function. The parameters can be calculated with maximum likelihood estimation (EML). If the S-PDFs of m points are $f_1(x_1), f_2(x_2), \dots, f_m(x_m)$, then M-PDF can be expressed as (7) when the derivation is conducted on (6):

$$\begin{aligned} f(x_1, x_2, \dots, x_m; \rho, k) &= \frac{\partial^2 F}{\partial x_1 \partial x_2 \dots \partial x_m} \\ &= c(F_1(x_1), F_2(x_2), \dots, F_m(x_m); \rho, k) \prod_{i=1}^m f_i(x_i), \end{aligned} \quad (7)$$

where $c(\cdot; \theta)$ is the probability density function of t-copula. Let n be the number of times for observation; thus, the likelihood function can be written as

$$\begin{aligned} L(\rho, k) &= \prod_{i=1}^n f(x_{1i}, x_{2i}, \dots, x_{mi}; \rho, k) \\ &= \prod_{i=1}^n c[F_1(x_{1i}), F_2(x_{2i}), \dots, F_m(x_{mi}); \rho, k] \\ &\quad \cdot f_1(x_{1i}) f_2(x_{2i}) \dots f_m(x_{mi}). \end{aligned} \quad (8)$$

The corresponding logarithmic likelihood function can be written as

$$\begin{aligned} \ln L(\rho, k) \\ = \sum_{i=1}^n \ln c[F_1(x_{1i}), F_2(x_{2i}), \dots, F_m(x_{mi}); \rho, k] \\ + \sum_{i=1}^n \ln f_1(x_{1i}) + \dots + \sum_{i=1}^n \ln f_m(x_{mi}). \end{aligned} \quad (9)$$

The EML for ρ and k can be obtained by solving the maximum value of (9):

$$\hat{\rho}, \hat{k} = \arg \max \ln L(\rho, k). \quad (10)$$

The fitting accuracy can also be evaluated by comparing the M-CDF values and the empirical cumulative distribution function of multiple points (M-ECDF).

2.3. Diagnostic Criteria Based on Typical Small Probability Method. As shown in (6), M-CDF, which is the function of measured values collected by different points, has a positive correlation with these measured values. Therefore, M-CDF can be used as an important parameter to diagnose the service status of an entire dam. If the value changes in a fixed interval, then the service status of an entire dam can be considered normal. If the value exceeds the interval, its status can be judged as abnormal. In this study, the monitoring data are divided into two periods: modeling period and diagnosis period. The data of modeling period are applied to establish M-CDF, and the interval can be calculated based on M-CDF values in this period and typical small probability method.

Let the length of the monitoring data in the modeling period be r years; thus M-CDF can be established, and the corresponding M-CDF value in each monitoring day can be calculated. Then, the unfavorable M-CDF value I in each year indicates that the annual maximum I and the annual minimum I form the samples of $I_{\max}$ and $I_{\min}$:

$$\begin{aligned} I_{\max} &= \{I_{\max 1}, I_{\max 2}, \dots, I_{\max r}\}; \\ I_{\min} &= \{I_{\min 1}, I_{\min 2}, \dots, I_{\min r}\}. \end{aligned} \quad (11)$$

The mean values $\overline{I_{\max}}$ and $\overline{I_{\min}}$ and variances $\sigma_{\max}$ and $\sigma_{\min}$ are written as follows:

$$\begin{aligned} \overline{I_{\max}} &= \frac{1}{n} \sum_{i=1}^r I_{\max i}; \\ \overline{I_{\min}} &= \frac{1}{n} \sum_{i=1}^r I_{\min i}; \\ \sigma_{\max} &= \sqrt{\frac{1}{r-1} \left(\sum_{i=1}^r I_{\max i}^2 - r \overline{I_{\max}}^2 \right)}; \\ \sigma_{\min} &= \sqrt{\frac{1}{r-1} \left(\sum_{i=1}^r I_{\min i}^2 - r \overline{I_{\min}}^2 \right)}. \end{aligned} \quad (12)$$

According to the characteristic values, K-S testing approach is applied to check the distribution pattern of $I_{\max}$ and $I_{\min}$. Then, the probability density functions $f_{\max}(I_{\max})$ and $f_{\min}(I_{\min})$ and the cumulative distribution functions $F_{\max}(I_{\max})$ and $F_{\min}(I_{\min})$ are confirmed. Let the threshold

values of M-CDF values be $I_{\max \alpha}$ and $I_{\min \alpha}$; thus, the possibility of $I > I_{\max \alpha}$ or $I < I_{\min \alpha}$, which indicates the abnormal situation of dam, can be written as

$$P(I > I_{\max \alpha}) = \alpha = \int_{I_{\max \alpha}}^{+\infty} f(I_{\max}) dI_{\max}; \quad (13)$$

$$P(I < I_{\min \alpha}) = \alpha = \int_{-\infty}^{I_{\min \alpha}} f(I_{\min}) dI_{\min}.$$

Based on the small probability principle [26], α is always set as 0.05. Thus, $I_{\max \alpha}$ and $I_{\min \alpha}$ can be calculated as (14). If M-CDF exceeds threshold, then the measured values collected by different points are in the small probability event area:

$$I_{\max \alpha} = F_{\max}^{-1}(\overline{I_{\max}}, \sigma_{\max}, 0.95), \quad (14)$$

$$I_{\min \alpha} = F_{\min}^{-1}(\overline{I_{\min}}, \sigma_{\min}, 0.05).$$

2.4. Diagnosis Process for Service Status of the Entire Dam. M-CDF reflects the joint distribution of various points based on historical monitoring data. Therefore, if the value of the diagnostic period exceeds the threshold level, then two cases should be considered. (1) An abnormal M-CDF value is caused by multiple points. This phenomenon indicates that the whole dam is in a state of high risk. Attention should be given to the changing trend of the value in the following days, and related measures should be taken to make the value be in the normal range. (2) An abnormal value is caused by the measuring data of several points. When this phenomenon occurs, the abnormal points should be removed, and a new M-CDF should be reestablished. If the new M-CDF is in a normal state, then the analysis should be focused on these abnormal points. If the new M-CDF value still exceeds the threshold level, then its trend should be observed in the following days. If necessary, the emergency action should also be conducted to recover the value to the reasonable region. The flowchart based on the preceding discussion is illustrated in Figure 1 to describe clearly the diagnosing process for service behavior of dams.

3. Case Study

3.1. Description of the Project. Wanan Water Conservancy Project is located in the middle reaches of Gan River in Jiangxi province, China. This project is constituted by a concrete gravity dam, earth- and rock-filled dams, and a ship lock. Figure 2 shows the photo of the dam. The normal water level, design flood level, and maximum flood level of the project are 96 m, 100 m, and 100.70 m, respectively. The crest elevation of the concrete gravity dam is 104 m, and the height of gravity dam is 46.04 m. Thus, the project has a complete monitoring system for displacement and seepage. A total of 25 dam foundations are located in the gravity dam. One wire alignment named EX4 with 25 measuring points is arranged at the observation gallery set in its crest to monitor the static horizontal displacement of each dam foundation. The sketch of point arrangement is shown in Figure 3. In this study, four points (i.e., EX422, EX423, EX424, and EX425) in

FIGURE 1: Flowchart of diagnosing process.

FIGURE 2: Wanan Water Conservancy Project.

the powerhouse dam are selected as example to analyze the diagnosis method (the EX421 is excluded because it has insufficient monitoring data). After processing and synchronizing these measuring values of four points, 4091 sets of data from January 1, 1999, to December 31, 2013, are set as modeling period. These data are applied to establish the M-CDF and diagnosis criteria. Therefore, the data in 2014 are set as diagnosis period (including 295 sets), and the service status of the powerhouse dam in this period is diagnosed.

3.2. Distribution of a Single Point and Multiple Points. The process lines in the modeling period and diagnosing period of the four points are shown in Figure 4. The definition of the orientation is as follows: toward downstream is positive, whereas toward upstream is negative. Although the changing trend of EX422 from 2005 to 2013 is toward downstream, all points have consistency and regularity, thereby facilitating the construction of M-CDF.

According to (1), the S-PDF values of different points in the modeling period are determined with KDE. The fitting results of these points for frequency histogram are shown in Figure 5 with Gauss kernel function, and the bandwidths

FIGURE 3: Diagram of wire alignment in dam crest.

FIGURE 4: Recorded static horizontal deformation of EX422, EX423, EX424, and EX425.

are set as 0.1, 0.5, and 1. Then, S-CDF can be calculated based on S-PDF, and the S-ECDF values are calculated with (4). Table 1 indicates the RMSE in different bandwidths of different points. As shown in Figure 5 and Table 1, if the bandwidth is small, then the S-PDF curve is close to the frequency histogram. This scenario indicates a high fitting accuracy. When the bandwidth is set as 0.1, the features of frequency histograms are simulated preferably by the S-PDF. The small RMSE indicates that the corresponding S-CDF values are close to S-ECDF values. If the samples of EX422, EX423, EX424, and EX425 are $\{X_{1i}\}$, $\{X_{2i}\}$, $\{X_{3i}\}$, and $\{X_{4i}\}$, respectively, then the S-PDF of these points in this bandwidth can be expressed as

$$f_1(x_1) = \frac{1}{409.1} \sum_{i=1}^{4091} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(x_1 - X_{1i})^2}{0.02}\right);$$

$$f_2(x_2) = \frac{1}{409.1} \sum_{i=1}^{4091} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(x_2 - X_{2i})^2}{0.02}\right),$$

TABLE 1: RMSE in different bandwidths of different points.

Bandwidth	EX422	EX423	EX424	EX425
1.0	7.35×10^{-3}	1.07×10^{-2}	1.19×10^{-2}	1.68×10^{-2}
0.5	3.06×10^{-3}	4.18×10^{-3}	4.20×10^{-3}	6.42×10^{-3}
0.1	8.77×10^{-4}	9.80×10^{-4}	9.05×10^{-4}	1.07×10^{-3}

$$f_3(x_3) = \frac{1}{409.1} \sum_{i=1}^{4091} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(x_3 - X_{3i})^2}{0.02}\right);$$

$$f_4(x_4) = \frac{1}{409.1} \sum_{i=1}^{4091} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(x_4 - X_{4i})^2}{0.02}\right). \quad (15)$$

The corresponding S-CDF can be written as

$$F_1(x_1) = \int_{-\infty}^{x_1} f_1(x_1) dx;$$

$$F_2(x_2) = \int_{-\infty}^{x_2} f_2(x_2) dx;$$

$$F_3(x_3) = \int_{-\infty}^{x_3} f_3(x_3) dx;$$

$$F_4(x_4) = \int_{-\infty}^{x_4} f_4(x_4) dx. \quad (16)$$

Then, the M-CDF values of the four points are established with (6) to (10), and the M-ECDF in each sample point is calculated. The following is the expression of M-CDF:

$$F(x_1, x_2, x_3, x_4; \rho, k) = C(F_1(x_1), F_2(x_2), F_3(x_3), F_4(x_4));$$

$$F_4(x_4); \rho, k) = C\left(F_1(x_1), F_2(x_2), F_3(x_3), F_4(x_4); \begin{bmatrix} 1.00 & 0.53 & 0.75 & 0.56 \\ 0.53 & 1.00 & 0.89 & 0.91 \\ 0.75 & 0.89 & 1.00 & 0.93 \\ 0.56 & 0.91 & 0.93 & 1.00 \end{bmatrix}, 13.56\right). \quad (17)$$

FIGURE 5: S-PDF of EX422, EX423, EX424, and EX425 in different bandwidths.

Figure 6 describes the relationship of the M-CDF and the M-ECDF of each sample point. All values of the sample points are obviously located near the diagonal line. This observation indicated that the M-CDF values are near the M-ECDF values. In addition, the RMSE is 2.04×10^{-2} , thereby indicating the high accuracy of M-CDF in simulating the joint distribution of the four points and the benefits for fusing diagnosis.

3.3. Diagnosis of Service Status. According to (11) and (12), the annual maximum $I_{\max}$ and minimum $I_{\min}$ of the four-point M-CDF in the modeling period are calculated and shown in Table 2. With K -S test method, $I_{\max}$ obeys normal distribution $N(7.54 \times 10^{-1}, 7.95 \times 10^{-2})$, whereas $I_{\min}$ obeys

logarithmic normal distribution $N(-5.69, 2.03)$. When α is set as 0.05, $I_{\max\alpha}$ and $I_{\min\alpha}$ are 8.84×10^{-1} and 1.2×10^{-4} , respectively. The process line of the four-point M-CDF values in the diagnosis period is shown in Figure 7.

As shown in Figure 7, all M-CDF values in the diagnostic period are not lower than $I_{\min\alpha}$. This observation indicates that the displacement of various dam sections toward upstream is harmonious, and an excessive large displacement toward the upstream of all dam sections or individual section does not occur. However, from March 11 to 15, the value exceeds $I_{\max\alpha}$, and the maximum value in March 13 is reached. This observation indicates that the displacement of all sections or partial sections toward downstream in this period is excessively large, and further analysis should be

TABLE 2: Maximum and minimum values of the four-point M-CDF.

Year	$I_{\max}$	$I_{\min}$
1999	6.01×10^{-1}	7.14×10^{-3}
2000	8.33×10^{-1}	1.75×10^{-5}
2001	6.88×10^{-1}	2.62×10^{-3}
2002	7.63×10^{-1}	2.57×10^{-2}
2003	8.31×10^{-1}	5.52×10^{-4}
2004	8.02×10^{-1}	1.44×10^{-2}
2005	6.76×10^{-1}	2.56×10^{-3}
2006	8.14×10^{-1}	2.68×10^{-2}
2007	7.39×10^{-1}	1.81×10^{-3}
2008	8.63×10^{-1}	2.07×10^{-3}
2009	8.02×10^{-1}	4.94×10^{-3}
2010	7.04×10^{-1}	1.04×10^{-3}
2011	7.85×10^{-1}	7.43×10^{-2}
2012	7.83×10^{-1}	1.37×10^{-2}
2013	6.21×10^{-1}	9.28×10^{-4}

FIGURE 6: M-CDF value and M-ECDF value.

conducted. According to the diagnosis process of Figure 1, the three-point M-CDFs after eliminating one point are established, and the expressions can be expressed as

$$F_{\text{EX422,EX423,EX424}}(x_1, x_2, x_3; \rho, k) = C \left(F_1(x_1), F_2(x_2), F_3(x_3); \begin{bmatrix} 1.00 & 0.53 & 0.75 \\ 0.53 & 1.00 & 0.89 \\ 0.75 & 0.89 & 1.00 \end{bmatrix}, 14.76 \right),$$

$$F_{\text{EX422,EX423,EX425}}(x_1, x_2, x_4; \rho, k) = C \left(F_1(x_1), \right.$$

FIGURE 7: Process line of the four-point M-CDF in diagnosis period.

$$F_2(x_2), F_4(x_4); \begin{bmatrix} 1.00 & 0.52 & 0.55 \\ 0.52 & 1.00 & 0.91 \\ 0.55 & 0.91 & 1.00 \end{bmatrix}, 15.35 \right),$$

$$F_{\text{EX422,EX424,EX425}}(x_1, x_3, x_4; \rho, k) = C \left(F_1(x_1), \right.$$

$$F_3(x_3), F_4(x_4); \begin{bmatrix} 1.00 & 0.75 & 0.57 \\ 0.75 & 1.00 & 0.93 \\ 0.57 & 0.93 & 1.00 \end{bmatrix}, 59.00 \right),$$

$$F_{\text{EX423,EX424,EX425}}(x_2, x_3, x_4; \rho, k) = C \left(F_2(x_2), \right.$$

$$F_3(x_3), F_4(x_4); \begin{bmatrix} 1.00 & 0.89 & 0.91 \\ 0.89 & 1.00 & 0.93 \\ 0.91 & 0.93 & 1.00 \end{bmatrix}, 8.05 \right).$$

(18)

The results of K -S test show that $I_{\max}$ obeys normal distribution, whereas $I_{\min}$ obeys the logarithmic normal distribution for all three-point M-CDFs. The characteristic parameters of these distributions, such as expectation, μ , and variance, σ , are shown in Table 3. For $I_{\max}$ and $I_{\min}$ of M-CDF involving EX423, EX424, and EX425, μ is larger whereas σ is smaller than those of other M-CDFs. This finding indicates that the annual extreme value change is relatively stable, and the dispersion degree is small. The process lines of the three-point M-CDF values in diagnosis period are shown in Figure 8.

As shown in Figures 8(a)–8(c), the M-CDF values involving EX422 exceed $I_{\max \alpha}$ from March 11 to 15, and the maximum value is reached in March 13. This scenario is similar with that in Figure 7. However, the M-CDF values involving EX423, EX424, and EX425 are in the normal range from March 11 to 15, 2014. Therefore, M-CDF value anomaly is regarded to be caused by a single point (i.e., EX422) from the qualitative perspective. Furthermore, the quantitative analysis can be applied on EX422. The measured information

FIGURE 8: Process line of the three-point M-CDF values in the diagnosis period.

of various single points and the M-CDF values on March 13 are shown in Tables 4 and 5, respectively.

Based on (6), the M-CDF of the four points on March 13 can be expressed as

$$\begin{aligned}
 F(x_1, x_2, x_3, x_4) &= P(\text{EX422} \leq 13.35, \text{EX423} \\
 &\leq 3.84, \text{EX424} \leq 4.92, \text{EX425} \leq 4.42) = C[9.93 \\
 &\times 10^{-1}, 9.02 \times 10^{-1}, 9.83 \times 10^{-1}, 9.38 \times 10^{-1}] \\
 &= 8.90 \times 10^{-1}.
 \end{aligned} \quad (19)$$

The M-CDF involving EX423, EX424, and EX425 is in normal region on March 13. Therefore, under this condition, when the M-CDF values of the four points reach $I_{\max\alpha} = 8.84 \times 10^{-1}$, the S-CDF of EX422 can be calculated as

$$\begin{aligned}
 \text{S-CDF}_{\text{EX422}} &= F^{-1}(I_{\max\alpha}, x_2, x_3, x_4) = F^{-1}(8.84 \\
 &\times 10^{-1}, 9.02 \times 10^{-1}, 9.83 \times 10^{-1}, 9.38 \times 10^{-1}) \\
 &= 9.80 \times 10^{-1}.
 \end{aligned} \quad (20)$$

Then, the conditional probability for EX422, which means the S-PDF value, is in the region of $[9.80 \times 10^{-1}, 9.93 \times 10^{-1}]$ under the condition of the normal M-CDF value involving EX423, EX424, and EX425, and the corresponding measured

value of EX422, which is in the region of $[11.21, 13.35]$, can be calculated as

$$\begin{aligned}
 &P(11.21 \leq \text{EX422} \leq 13.35 \mid \text{EX423} \leq 3.84, \text{EX424} \\
 &\leq 4.92, \text{EX425} \leq 4.42) = P(\text{EX422} \leq 13.35 \mid \text{EX423} \\
 &\leq 3.84, \text{EX424} \leq 4.92, \text{EX425} \leq 4.42) - P(\text{EX422} \\
 &\leq 11.21 \mid \text{EX423} \leq 3.84, \text{EX424} \leq 4.92, \text{EX425} \leq 4.42) \\
 &= \frac{C(9.93 \times 10^{-1}, 9.02 \times 10^{-1}, 9.83 \times 10^{-1}, 9.38 \times 10^{-1})}{C(9.02 \times 10^{-1}, 9.83 \times 10^{-1}, 9.38 \times 10^{-1})} \\
 &\quad - \frac{C(9.80 \times 10^{-1}, 9.02 \times 10^{-1}, 9.83 \times 10^{-1}, 9.38 \times 10^{-1})}{C(9.02 \times 10^{-1}, 9.83 \times 10^{-1}, 9.38 \times 10^{-1})} \\
 &= 6.72 \times 10^{-3}.
 \end{aligned} \quad (21)$$

According to the analysis, the measured value of EX422 is in the region of $[11.21, 13.35]$ with the condition that the M-CDF involving EX423, EX424, and EX425 is in normal region while the M-CDF of the four points is in the abnormal region. The probability of the matter is 6.72×10^{-3} , which is much less than the significance level of small probability event (0.05). However, the measured values of EX422 are in this region from March 11 to 15, 2014. This observation indicates that the small possibility event has already occurred. Therefore, the change of EX422 is not coordinated with other points. Thus, the displacement toward the downstream of 4# powerhouse

TABLE 3: Characteristic parameter of three-point M-CDF.

M-CDF	$I_{\max}$		$I_{\max \alpha}$	Logarithmic $I_{\min}$		$I_{\min \alpha}$
	μ	σ		μ	σ	
EX422, EX423, and EX424	7.74×10^{-1}	7.46×10^{-2}	8.96×10^{-1}	-5.36×10^0	2.01×10^0	1.69×10^{-4}
EX422, EX423, and EX425	7.55×10^{-1}	7.98×10^{-2}	8.87×10^{-1}	-5.67×10^0	2.06×10^0	1.16×10^{-4}
EX422, EX424, and EX425	7.90×10^{-1}	7.76×10^{-2}	9.18×10^{-1}	-5.36×10^0	2.13×10^0	1.41×10^{-4}
EX423, EX424, and EX425	8.27×10^{-1}	7.43×10^{-2}	9.50×10^{-1}	-5.03×10^0	1.92×10^0	2.78×10^{-4}

TABLE 4: Measured information of various single points on March 13.

	Point			
	EX422	EX423	EX424	EX425
Measured value (mm)	13.35	3.84	4.92	4.42
S-CDF value	9.93×10^{-1}	9.02×10^{-1}	9.83×10^{-1}	9.38×10^{-1}

TABLE 5: M-CDF values on March 13.

	EX422, EX423, EX424, and EX425	EX422, EX423, and EX424	EX422, EX423, and EX425	EX422, EX424, and EX425	EX423, EX424, and EX425
M-CDF	8.90×10^{-1}	9.03×10^{-1}	8.89×10^{-1}	9.35×10^{-1}	8.93×10^{-1}

dam section is excessively large, whereas those of the other sections are normal. As shown in Figure 4, the displacement toward the downstream of EX422 in the modeling period shows an increasing trend. Thus, this trend is fully considered in the process of building M-CDF. In this case, the M-CDF values related to EX422 from March 11 to March 15, 2014, still exceed the high threshold level. This finding indicates that the changing trend toward the downstream of this dam section continues to increase. Therefore, focusing on the monitoring values of EX422 is essential, and the maintaining or strengthening measures can be applied to 4# powerhouse dam section if necessary.

4. Conclusion

In this study, a diagnosis method for the service status of the whole dam based on joint distribution of multiple points is proposed. The distributions of single point and multiple points are researched. The method of parameter estimation and the accuracy of S-CDF and M-CDF are discussed. Then, the diagnostic criteria and processes are established with typical small probability method. The following conclusions are obtained.

- (1) KDE can calculate the S-PDF on the basis of protecting in situ monitoring information effectively. When the bandwidth is set as 0.1, S-PDF has a high fitting precision, which can overcome the shortage of a priori assumption and can provide a new method to estimate the distribution of single point.
- (2) M-CDF based on t-copula can reflect the relationship of multiple points. It also has high fitting precision for joint distribution. The M-CDF effectively fuses the measuring values of multiple points. Therefore, it can be considered an important index to diagnose the service status of an entire dam.

- (3) The diagnosis criteria based on the typical small probability method reflect the distribution of the extremum of M-CDF. This extremum is beneficial to diagnose the service status of an entire dam and to identify abnormal points.

Competing Interests

The authors declare that they have no competing interests.

Acknowledgments

This research is supported by the National Natural Science Foundation of China (51379162 and 51079114); Water Conservancy Science and Technology Innovation Project of Guangdong Province (2016-06); and the Fundamental Research Funds for the Central Universities (2042016kf1081).

References

- [1] ICOLD, *Automated Dam Monitoring Systems—Guidelines and Case Histories*, International Commission on Large Dams, Paris, France, 2000.
- [2] ICOLD, *Guidelines for Automated Dam Monitoring Systems*, International Commission on Large Dams, Paris, France, 1999.
- [3] J. P. He, *Safety Monitoring Theory and Its Application of Dam*, China Water Power Press, Beijing, China, 2010.
- [4] D. Tonini, "Observed behaviour of several Italian arch dams," *Journal Power Division*, vol. 3, pp. 82–86, 1956.
- [5] C. Gu, D. Qin, Z. Li, and X. Zheng, "Study on semi-parametric statistical model of safety monitoring of cracks in concrete dams," *Mathematical Problems in Engineering*, vol. 2013, Article ID 874629, 9 pages, 2013.
- [6] M. Tatin, M. Briffaut, F. Dufour, A. Simon, and J.-P. Fabre, "Thermal displacements of concrete dams: accounting for water temperature in statistical models," *Engineering Structures*, vol. 91, pp. 26–39, 2015.

- [7] F. Salazar, M. A. Toledo, E. Oñate, and R. Morán, "An empirical comparison of machine learning techniques for dam behaviour modelling," *Structural Safety*, vol. 56, pp. 9–17, 2015.
- [8] F. Ganelli, A. Beko, and A. Mordini, "Structural assessment of the Mignano solid-gravity dam based on the monitoring data," *Structural Engineering International*, vol. 22, no. 1, pp. 130–138, 2012.
- [9] C. Bishop, *Pattern Recognition and Machine Learning*, Springer, Berlin, Germany, 2007.
- [10] K. P. Murphy, *Machine Learning*, MIT Press, Cambridge, Mass, USA, 2012.
- [11] B. Stojanovic, M. Milivojevic, M. Ivanovic, N. Milivojevic, and D. Divac, "Adaptive system for dam behavior modeling based on linear regression and genetic algorithms," *Advances in Engineering Software*, vol. 65, pp. 182–190, 2013.
- [12] V. Ranković, N. Grujović, D. Divac, and N. Milivojević, "Development of support vector regression identification model for prediction of dam structural behaviour," *Structural Safety*, vol. 48, pp. 33–39, 2014.
- [13] G. E. Tsekouras and J. Tsimikas, "On training RBF neural networks using input-output fuzzy clustering and particle swarm optimization," *Fuzzy Sets and Systems*, vol. 221, pp. 65–89, 2013.
- [14] P. Sitarz and B. Powalka, "Modal parameters estimation using ant colony optimisation algorithm," *Mechanical Systems and Signal Processing*, vol. 76–77, pp. 531–554, 2016.
- [15] M. Bagheri, M. Miri, and N. Shabakhty, "Modeling of epistemic uncertainty in reliability analysis of structures using a robust genetic algorithm," *Iranian Journal of Fuzzy Systems*, vol. 12, no. 2, pp. 23–40, 2015.
- [16] J. P. He, C. B. Ma, and Y. Q. Shi, "Multi-effect-quantity fusion model of high arch dam based on improved D-S evidence theory," *Geomatics and Information Science of Wuhan University*, vol. 37, no. 12, pp. 1397–1400, 2012.
- [17] A. De Sortis and P. Paoliani, "Statistical analysis and structural identification in concrete dam monitoring," *Engineering Structures*, vol. 29, no. 1, pp. 110–120, 2007.
- [18] H. Z. Su, J. Hu, and Z. R. Wu, "A study of safety evaluation and early-warning method for dam global behavior," *Structural Health Monitoring*, vol. 11, no. 3, pp. 269–279, 2012.
- [19] M. I. Rafiq, M. K. Chryssanthopoulos, and S. Sathananthan, "Bridge condition modelling and prediction using dynamic Bayesian belief networks," *Structure and Infrastructure Engineering*, vol. 11, no. 1, pp. 38–50, 2015.
- [20] M. Masoumi, M. Masoumi, and E. Jamshidi, "Damage diagnosis in steel structures with different noise levels via optimization algorithms," *International Journal of Steel Structures*, vol. 15, no. 3, pp. 557–565, 2015.
- [21] G. Georgoulas, M. O. Mustafa, I. P. Tsoumas et al., "Principal component analysis of the start-up transient and Hidden Markov Modeling for broken rotor bar fault diagnosis in asynchronous machines," *Expert Systems with Applications*, vol. 40, no. 17, pp. 7024–7033, 2013.
- [22] X. Liu, Z.-R. Wu, Y. Yang, J. Hu, and B. Xu, "Information fusion diagnosis and early-warning method for monitoring the long-term service safety of high dams," *Journal of Zhejiang University: Science A*, vol. 13, no. 9, pp. 687–699, 2012.
- [23] F. Nasrfard-Jahromi and M. Mohammadi, "Probabilistic harmonic load flow using an improved kernel density estimator," *International Journal of Electrical Power & Energy Systems*, vol. 78, pp. 292–298, 2016.
- [24] S. Sadri and D. H. Burn, "Copula-based pooled frequency analysis of droughts in the Canadian prairies," *Journal of Hydrologic Engineering*, vol. 19, no. 2, pp. 277–289, 2014.
- [25] M. J. Reddy and P. Ganguli, "Bivariate flood frequency analysis of upper godavari river flows using archimedean copulas," *Water Resources Management*, vol. 26, no. 14, pp. 3995–4018, 2012.
- [26] Z. Sheng, S. Q. Xie, and C. Y. Pan, *Probability and Statistics*, Higher Education Press, Beijing, China, 2008.

Hindawi

Submit your manuscripts at
<http://www.hindawi.com>