

## Research Article

# Time-Varying Identification Model for Crack Monitoring Data from Concrete Dams Based on Support Vector Regression and the Bayesian Framework

Bo Chen,<sup>1,2,3,4</sup> Zhongru Wu,<sup>1,3,4</sup> Jiachen Liang,<sup>1,3,4</sup> and Yanhong Dou<sup>4</sup>

<sup>1</sup>State Key Laboratory of Hydrology-Water Resources and Hydraulic Engineering, Hohai University, Nanjing 210098, China

<sup>2</sup>Key Laboratory of Earth-Rock Dam Failure Mechanism and Safety Control Techniques, Ministry of Water Resources, Nanjing 210029, China

<sup>3</sup>National Engineering Research Center of Water Resources Efficient Utilization and Engineering Safety, Hohai University, Nanjing 210098, China

<sup>4</sup>College of Water Conservancy and Hydropower Engineering, Hohai University, Nanjing 210098, China

Correspondence should be addressed to Bo Chen; chenbohhu@126.com

Received 26 October 2016; Revised 20 December 2016; Accepted 19 January 2017; Published 19 February 2017

Academic Editor: Salvatore Caddemi

Copyright © 2017 Bo Chen et al. 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 modeling of cracks and identification of dam behavior changes are difficult issues in dam health monitoring research. In this paper, a time-varying identification model for crack monitoring data is built using support vector regression (SVR) and the Bayesian evidence framework (BEF). First, the SVR method is adopted for better modeling of the nonlinear relationship between the crack opening displacement (COD) and its influencing factors. Second, the BEF approach is applied to determine the optimal SVR modeling parameters, including the penalty coefficient, the loss coefficient, and the width coefficient of the radial kernel function, under the principle that the prediction errors between the monitored and the model forecasted values are as small as possible. Then, considering the predicted COD, the historical maximum COD, and the time-dependent component, forewarning criteria are proposed for identifying the time-varying behavior of cracks and the degree of abnormality of dam health. Finally, an example of modeling and forewarning analysis is presented using two monitoring subsequences from a real structural crack in the Chencun concrete arch-gravity dam. The findings indicate that the proposed time-varying model can provide predicted results that are more accurately nonlinearly fitted and is suitable for use in evaluating the behavior of cracks in dams.

## 1. Introduction

Cracks are the most common structural threats to concrete dams. According to an investigative report issued by the International Committee on Large Dams, a total of 243 dams have collapsed as a result of cracking problems [1]. As an example, in the case of the Koelnbrein arch dam in Austria, diagonal cracks emerged at the heel of the dam, causing the grouting curtain to break. Under the enormous water pressure at the crack surfaces, these cracks expanded throughout the foundation and heel of the dam, which led to the full head of uplift pressure being applied over the entire surface of the base [2]. Therefore, the dam could no longer operate safely. In another example, soon after the pouring

of the Phase II concrete for the Chencun dam in China, a crack of 300 m in length and more than 5 m in depth, which stretched across dam sections #5 to #28, emerged at the interface between the concrete layers poured in Phases I and II. Serious crack problems have also occurred during the operation of the Dworshak solid gravity dam in America, the Sayano-Shushenskaya arch-gravity dam in Russia, and the Xiaowan arch dam in China [3–5]. Cracks are destructive. Severe cracks can weaken the strength and rigidity of a dam, destroy its integrity and antipermeability, accelerate the corrosion and carbonation of the concrete, and endanger the safe operation of the dam.

The research on structural health monitoring of cracks is mainly classified into two categories [6, 7]. The first one

focuses on the nondestructive detection and characterization of cracks, including location, size, depth, and other information. Numerous sensing techniques are proposed and applied, such as strain gauges [8], optical fiber sensor [9], piezoceramic sensor [10], carbon nanotube thread [11], and other vibration-based methods [12, 13]. The other one focuses on the assessment and prediction of crack propagation, as well as structural behavior, based on quantification monitoring data analysis. However, these data are still difficult to be obtained accounting for the limitations of sensing techniques, heterogeneity of concrete material, and bulkiness of dam concrete [7]. Luckily, there are still some real-time cracking series being obtained by later buried crack gauges in dam engineering, such as Chencun and Xiaowan dam [3].

To date, methods of propagation analysis and abnormality diagnosis for concrete crack behavior have mainly relied on fracture mechanics methods and applied mathematical methods. Currently, criteria such as the stress intensity factor criterion [14–16], the strain energy density factor criterion [17], the maximum energy release rate criterion [18], the mixed energy-stress criterion [19], and the dual-K fracture criterion [20] are frequently adopted to determine whether unstable propagation is occurring. But the variables on these criteria are difficult to monitor accurately and inconvenient to apply in practical engineering. The other approach [19, 21] is based on the consideration that the crack will lose its stability rapidly when the crack tip opening displacement (CTOD) reaches its critical value ( $CTOD_C$ ). Several mathematical models based on the CTOD have been established based on the relationship between the CTOD and the crack aperture. However, CTOD theory is also difficult to use in the diagnosis of crack propagation behavior, mainly because the CTOD is extremely difficult to monitor. Moreover, the indicator,  $CTOD_C$ , must be measured in fracture tests using large-sized specimens, which is difficult to achieve in typical laboratories [22]. Therefore, these methods are subject to limitations in the diagnosis of structural behavior in dam engineering.

For research on dam health monitoring, especially with regard to major cracks, Wu et al. [23–25] considered the contributions of water pressure, temperature, and structural aging to establish a statistical model for crack monitoring data. In this model, the water pressure component is constructed based on the relationship between water level and COD derived from dam engineering theory; the structural aging component is derived based on the creep theory of concrete; and the temperature component is obtained by using the results of temperature field simulation or fitting a semiexperienced periodic function. Li and Gu [3–5, 26] suggested that the degree of abnormality of a crack in a dam can be diagnosed based on a time-dependent analysis. A semiparametric statistical model and a nonparametric statistical model [4, 26] were established by considering the relationship between the degree of abnormality of dam crack behavior and its statistical point of change to improve the fitting and explanation of crack monitoring data from concrete dams. Li et al. [3] analyzed the dynamic properties of crack monitoring series and proposed a dynamic structural diagnosis method for crack abnormality based on a fuzzy cross-correlation factor exponent. Li et al. [5] proposed a

fluctuation-based method of regression coefficients for crack monitoring series, in which each abnormal event is detected from a cumulative sum of regression model residuals. These models play an important role in COD monitoring. However, the fitting and prediction capabilities of the models need to be improved accounting for the nonlinearity of crack monitoring series, and feasible crack monitoring criteria need to be proposed for more effective application in dam engineering. In pursuit of nonlinear modeling and prediction methods, Panizzo and Petaccia [27] developed predictive models based on random forests (RFs), boosted regression trees (BRTs), neural networks (NNs), support vector machines (SVMs), and multivariate adaptive regression splines (MARSs), and a comparison of the prediction accuracy of these models revealed that they showed poorer performance on average than did a conventional statistical model. Therefore, further research is needed to construct nonlinear modeling approaches for crack monitoring that are able to better fit and predict crack opening behavior and evaluate anomalous characteristics of crack propagation and structural health.

A technique based on the use of SVMs for regression, namely, support vector regression (SVR), is an advanced statistical method that has recently been used in the modeling of nonlinear systems [28–32]. In SVR, a nonlinear transformation is employed to map monitoring data into a higher-dimensional feature space, and the nonlinear relationships between the monitored values of dam behavior quantities and their influencing factors are linearized by searching for an optimized fitting function. With the introduction of the insensitive error loss function as proposed by Vapnik, SVR has emerged as one of the best methods of regression because of its improved robustness and generalization performance [33]. Several researchers have employed SVR in modeling methods for structural identification to verify its potential capabilities in this respect. Ranković et al. [34] established a nonlinear autoregressive SVR model that yields accurate results for the prediction of the tangential displacement of a concrete dam. Lee et al. [35] successfully predicted the strength of concrete based on its mix proportions using the SVR technique and NNs, as indicated by a comparison against experimental results, and concluded that the SVR method can be used to predict the compressive strength of concrete with higher estimation accuracy and within a shorter computation time compared with the NN method. Su et al. [36] developed a time-varying identification model for dam behavior before and after structural reinforcement based on the SVR method, and this model was found to yield more accurate fitted and forecasted results compared with classical statistical models. In these studies, one key factor in establishing a useful model is the selection of suitable SVR parameters, such as the penalty coefficient, the loss coefficient, and the width coefficient of the kernel function. There is an urgent need for the development of an appropriate algorithm for parameter optimization.

Based on the brief review presented above, it is clear that the modeling of crack monitoring data is still an important problem in research on concrete dams. In the present study, SVR is applied to establish a parametric statistical model for the safety monitoring of cracks in concrete dams, with the

intent of improving the extraction of the nonlinear effects of the various influencing factors on COD. The BEF method is then introduced to optimize the model parameters. Finally, forewarning criteria based on the predicted components of the time-varying features of the crack monitoring model are proposed for the real-time diagnosis of abnormal dam behavior.

## 2. Time-Varying Forewarning Criteria for the Crack Monitoring of Concrete Dams

Monitoring equipment such as crack measurement meters and optical fibers provides a large amount of COD data for concrete dams. The statistical model for crack monitoring data that is considered in this paper has been confirmed to be effective for analyzing the crack evolution process. However, there is still a need for suitable forewarning criteria to be proposed to characterize the relationship between the crack propagation process and abnormal structural behavior in dam engineering applications. In this section, the prediction of the components of the statistical model and their time-varying features are introduced to establish these criteria.

### 2.1. Prediction Using the Statistical Crack Monitoring Model.

The factors that affect the monitored crack quantities in a concrete dam can be attributed to external loads (such as water pressure and temperature) and time-varying effects. Therefore, the monitoring data for dam cracks can be decomposed into contributions from different influencing factors as follows:

$$y(t) = y_H(t) + y_T(t) + y_\theta(t). \quad (1)$$

In the formula above,  $y(t)$  is a measured value in the crack opening monitoring sequence;  $y_H(t)$  and  $y_T(t)$  are the crack opening components associated with the action of water pressure and temperature, respectively; and  $y_\theta(t)$  is the time-varying component, which characterizes the other effects on crack behavior that are induced by creep or structural changes in the concrete.

Once the factors influencing the COD have been decomposed, a regression model and a prediction model for the monitored crack displacement values can be established based on the measured sequence, and then the estimated values for each component, namely,  $\hat{y}_H(t)$ ,  $\hat{y}_T(t)$ , and  $\hat{y}_\theta(t)$ , can be calculated. The statistical prediction model for crack monitoring data is as follows:

$$\begin{aligned} \hat{y}(t) &= \hat{y}_H(t) + \hat{y}_T(t) + \hat{y}_\theta(t) \\ S &= \sqrt{D(y(t) - \hat{y}(t))}. \end{aligned} \quad (2)$$

In the formula above,  $\hat{y}(t)$  is the predicted value of the COD at time  $t$ ;  $\hat{y}_H(t)$ ,  $\hat{y}_T(t)$ , and  $\hat{y}_\theta(t)$  are the estimated water pressure, temperature, and time-varying components, respectively;  $S$  is the residual standard deviation of the model; and  $D$  is the variance.

**2.2. Forewarning Criteria for Dam Behavior.** According to the prediction model presented above, the predicted value

of the COD at a future time  $t$  is calculated to be  $\hat{y}(t)$ , and the measured value of the COD at that time is denoted by  $y(t)$ . If it is assumed that the residual standard deviation of the model satisfies a normal distribution, then, according to the mathematical theory of probability and statistics, the relationship between the distance from the predicted value to the measured value and the residual standard deviation of the model is as follows: the probability that  $|y(t) - \hat{y}(t)|$  falls into the interval  $[0, 2S]$  is 95.5% and the probability that  $|y(t) - \hat{y}(t)|$  falls into  $[0, 3S]$  is 99.7%. That is to say, both  $2S < |y(t) - \hat{y}(t)| \leq 3S$  and  $|y(t) - \hat{y}(t)| > 3S$  are events that will occur with low probability. However, the time-dependent component of the COD reflects the irreversible changes to the concrete caused by creep and damage and is consequently an important index for characterizing the dam safety trend. Accordingly, by comparing between the measured values and the values predicted by the model and their eigenvalues and analyzing the trend of the development of the time-dependent component, time-varying forewarning criteria for the COD of a concrete dam can be established. Thus, a measured time sequence of data can be used to determine whether the structural state of the dam is changing. The forewarning criteria are as follows:

If

$$\begin{aligned} |y(t) - \hat{y}(t)| &\leq 2S \\ \text{or } \hat{y}(t) &\leq y_m, \end{aligned} \quad (3)$$

then the safety state of the concrete dam structure is normal.

If

$$\begin{aligned} 2S < |y(t) - \hat{y}(t)| &\leq 3S, \quad \hat{y}(t) > y_m, \\ \frac{d\hat{y}_\theta(t)}{dt} &> 0, \\ \frac{d^2\hat{y}_\theta(t)}{dt^2} &= 0, \end{aligned} \quad (4)$$

then the safety state of the concrete dam structure is abnormal.

If

$$\begin{aligned} |y(t) - \hat{y}(t)| &> 3S, \quad \hat{y}(t) > y_m, \\ \frac{d\hat{y}_\theta(t)}{dt} &> 0, \\ \frac{d^2\hat{y}_\theta(t)}{dt^2} &> 0, \end{aligned} \quad (5)$$

then the safety state of the concrete dam structure is significantly abnormal.

In the formulas above,  $\hat{y}(t)$  and  $y(t)$  are the predicted and measured values, respectively, of the monitored COD;  $y_m$  is the historical maximum value of the COD in the measured monitoring sequence; and  $\hat{y}_\theta(t)$  is the time-dependent component extracted from the prediction model.

## 3. Crack Prediction Model Based on B-SVR

The stepwise regression method is used in the traditional statistical modeling process. However, problems with modeling

accuracy often arise because of the nonlinear characteristics of crack monitoring data. To address this problem, a nonlinear modeling approach, namely, the SVR method, is adopted here to establish a bridge between the monitoring data and the components of the statistical model. The Bayesian evidence framework is then applied to determine the optimal SVR modeling parameters for achieving the best possible accuracy of the crack monitoring model.

**3.1. Modeling of Crack Monitoring Data Based on SVR.** Considering the nonlinear relationships between the effects on the monitored crack quantities and their influencing factors, let  $X_i$  ( $i = 1, 2, \dots, n$ ) be the input sequence of the set of crack-influencing factors and let  $y_i$  be the input sequence of monitored crack quantities. Then, the independent identically distributed sample set  $((X_i, y_i)$  ( $i = 1, 2, \dots, n$ )) of the crack monitoring sequence is subject to a certain distribution  $F(x, y)$ ,  $x \in R^l$ ,  $y \in R$ , in the function space  $R^l \times R$ . Based on the method of support vector regression (SVR) [28], the basic idea used to solve the above nonlinear prediction problem in a multidimensional function space is as follows: the set of monitoring data is mapped from the input space to a high-dimensional feature space via a nonlinear transformation  $\phi(\cdot)$ , and the hyperplane that yields the optimal classification of the linearized regression problem is found to fit the relationships between the influencing factors  $X$  and the effects on the input crack quantities  $y$ . Finally, the optimization problem for reproducing the monitored quantities based on the SVR model is established as follows:

$$\begin{aligned} \min \quad & R(\mathbf{w}, \xi) = \frac{1}{2} \mathbf{w}^T \mathbf{w} + C \sum_{i=1}^n \xi_i \\ \text{S.t.} \quad & y_i [\mathbf{w}^T \phi(x_i) + b] \geq 1 \\ & \xi_i \geq 0 \quad i = 1, 2, \dots, n. \end{aligned} \quad (6)$$

In the formulas above,  $R$  represents the risk function of the optimization problem;  $\mathbf{w} = (w_1, w_2, \dots, w_l)^T$  is the vector of the weights of all influencing factors;  $C > 0$  is a penalty parameter, which describes the tradeoff between the empirical risk and the model complexity;  $\xi_i$  is a slack variable such that  $\sum_{i=1}^n \xi_i$  describes the deviation of the monitored crack quantity data with respect to the ideal conditions of the classification;  $b$  is a constant bias term; and  $\phi(x)$  is the function used to map the input space to the high-dimensional feature space. The basic principle of the modeling of monitored crack quantities based on SVR to solve the corresponding nonlinear problem is illustrated in Figure 1.

In accordance with the optimization theory [37], the risk expectation minimization problem for the monitoring data model presented in formula (6) is transformed into a dual problem. The Lagrange multiplier method is used to seek the optimal solution at the saddle point of the above function, and

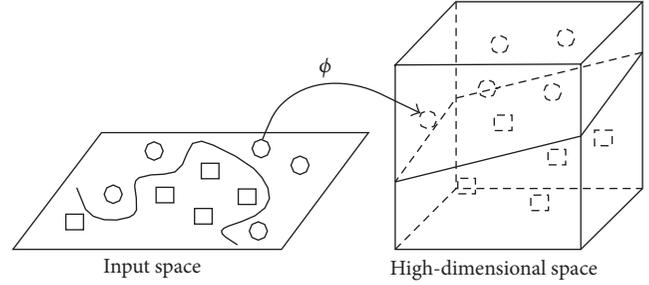


FIGURE 1: Principle of using a support vector machine to solve a nonlinear problem.

the Wolfe duality technique [38] is applied to solve the dual planning problem with the following inequality constraints:

$$\begin{aligned} \max: \quad & L(\boldsymbol{\alpha}, \mathbf{w}, b) = \sum_{i=1}^n \alpha_i + \frac{1}{2} \sum_{i,j=1}^n \alpha_i \alpha_j y_i y_j \phi(x_i) \phi(x_j) \\ \text{S.t.} \quad & \sum_{i=1}^n \alpha_i y_i = 0 \\ & 0 \leq \alpha_i \leq C \\ & i = 1, 2, \dots, n, \end{aligned} \quad (7)$$

where  $\boldsymbol{\alpha} = \{\alpha_1, \alpha_2, \dots, \alpha_n\}^T$  is the vector of the Lagrange multipliers and  $\phi(x_i)$  and  $\phi(x_j)$  are nonlinear transformation functions.

The solution to the nonlinear problem above is obtained by mapping the low-dimensional input samples to the high-dimensional feature space by introducing a kernel function to transform the nonlinear support vector regression problem into a linear problem. Finally, the hyperplane that represents the optimal solution to the problem is constructed. According to functional theory, the inner product function  $(K(x_i, x_j))$  that satisfies Mercer's condition  $(K(x_i, x_j) = \phi(x_i) \phi(x_j))$  is the kernel function that is sought. Commonly used kernel functions include a polynomial kernel function, a sigmoid kernel function, and the radial basis function. Research shows that the radial basis function kernel can be used to extract fewer support vectors with the same classification accuracy. Therefore, the radial basis function kernel  $(K(x_i, x_j) = \exp(-|x_i - x_j|^2 / 2\sigma^2))$  can be selected when modeling the measured crack monitoring data for a concrete dam. However, the value of  $\sigma$  must be determined according to the specifics of the modeling problem.

To enhance the robustness and sparsity of the algorithm, the nonsensitive loss function  $(L(y, f) = L(|y - f|_\epsilon))$  is defined as follows:

$$|y - f|_\epsilon = \begin{cases} 0, & |y - f| \leq \epsilon \\ |y - f| - \epsilon, & |y - f| > \epsilon \end{cases} \quad \epsilon \geq 0, \quad (8)$$

where  $\epsilon$  is the loss coefficient.

For the optimal regression problem shown in formula (6), by adopting the nonsensitive loss function as a constraint

condition and using the Lagrange multiplier method to construct the function to be optimized, the dual optimization problem can be obtained as follows:

$$\begin{aligned} \max: \quad & -\varepsilon \sum_{i=1}^n (\alpha_i + \alpha_i^*) + \sum_{i=1}^n y_i (\alpha_i^* - \alpha_i) \\ & - \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n (\alpha_i^* - \alpha_i) (\alpha_j^* - \alpha_j) K(x_i - x_j) \quad (9) \\ \text{s.t.} \quad & \sum_{i=1}^n (\alpha_i^* - \alpha_i) = 0 \quad \alpha_i, \alpha_i^* \in [0, C], \end{aligned}$$

where  $\alpha_i^*$  ( $i = 1, 2, \dots, n$ ) is the Lagrange multiplier corresponding to  $\alpha_i$ , such that  $\alpha_i^* \alpha_i = 0$ . The corresponding sampled data point is defined as a support vector only when  $\alpha_i^* \neq \alpha_i$ .

The SVR model and the estimated bias parameter of the crack monitoring sequence are expressed by formula (10) and formula (11), respectively, after the above quadratic optimization problem is solved. In the modeling of COD values, the vector of the inputs to the model,  $\mathbf{x}$ , should be defined to consist of the crack-influencing factors, namely, the water pressure, the temperature, and the time-varying factor. The output of the model can be defined as the monitored quantity ( $\delta$ ) of the COD. A structural representation of the modeling process using this support vector regression method is shown in Figure 2.

$$f(x) = \sum_{i=1}^n (\alpha_i^* - \alpha_i) K(x_i, \mathbf{x}) + b \quad (10)$$

$$\begin{aligned} b = \frac{1}{N_{\text{NSV}}} & \left\{ \sum_{0 < \alpha_j < C} \left[ y_i - \sum_{x_j \in \text{SV}}^l (\alpha_i - \alpha_i^*) K(x_j, x_i) \right. \right. \\ & \left. \left. - \varepsilon \right] + \sum_{0 < \alpha_j < C} \left[ y_i - \sum_{x_j \in \text{SV}}^l (\alpha_i - \alpha_i^*) K(x_j, x_i) \right. \right. \\ & \left. \left. + \varepsilon \right] \right\}, \quad (11) \end{aligned}$$

where  $n$  is the number of support vectors and  $N_{\text{NSV}}$  is the number of standard support vectors. SV denotes the set of standard support vectors.

**3.2. Deduction of Model Parameters Based on the Bayesian Evidence Framework.** The penalty coefficient ( $C$ ), the loss coefficient ( $\varepsilon$ ), and the width coefficient of the radial kernel function ( $\sigma$ ) need to be determined before the SVR modeling process for an observed series of concrete dam crack opening data. There are two traditional optimization approaches that are used to choose the best model parameters. The first method is to select the model parameters based on the user's experience, which is likely to result in prediction distortion due to subjective factors. The other method is cross validation, in which the samples are divided into a

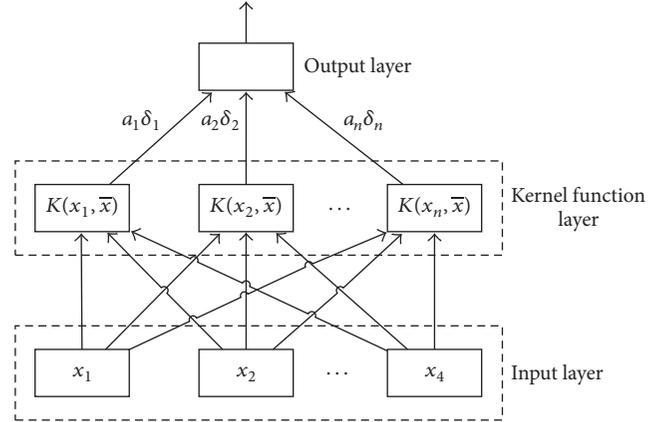


FIGURE 2: Schematic illustration of the support vector regression modeling process.

training set, a validation set, and a test set; the SVR model is trained using the training set for each  $(\varepsilon, C, \sigma)$ , and the reliability of the model is verified using the validation set. Once the constructed model has been confirmed to be reliable, the model parameters  $(\varepsilon, C, \sigma)$  that yield, on average, the minimum mean squared errors of prediction for the above SVR model can be adopted. To reduce the effort required in the cross-validation process, the optimal solution for the model parameters is deduced via the following steps by maximizing the posterior probability of the parameter distribution based on the Bayesian evidence framework. The steps of obtaining the optimal SVR parameters are as follows.

*Step 1.* First,  $\mathbf{w}$  and  $b$  are deduced. Suppose that the training set for model  $H$  is  $(X_i, y_i) \in D, i = 1, 2, \dots, n$ ; then, the posterior probability of the parameters can be expressed as in (12), according to the Bayes criterion [39, 40]:

$$\begin{aligned} P(\mathbf{w}, b | D, \lg C, \lg \varepsilon, H) \\ = \frac{P(D | \mathbf{w}, b, \lg C, \lg \varepsilon, H) P(\mathbf{w}, b | \lg C, \lg \varepsilon, H)}{P(D | \lg C, \lg \varepsilon, H)}, \quad (12) \end{aligned}$$

where  $P(D | \lg C, \lg \varepsilon, H)$  is the generalized constant that is related the deduced parameters  $\mathbf{w}$  and  $b$ ,  $P(\mathbf{w}, b | \lg C, \lg \varepsilon, H)$  is the joint prior probability of  $\mathbf{w}$  and  $b$ , and  $P(D | \mathbf{w}, b, \lg C, \lg \varepsilon, H)$  is the likelihood function.

If  $P(D | \mathbf{w}, b, \lg C, \lg \varepsilon, H)$  is independent of parameter  $C$ , then the likelihood function can be transformed into the following:

$$\begin{aligned} P(D | \mathbf{w}, b, \lg C, \lg \varepsilon, H) &= \prod_{i=1}^n P(y_i | X_i, \mathbf{w}, b, \lg \varepsilon, H) \\ &\cdot P(X_i | \mathbf{w}, b, \lg \varepsilon, H) = \prod_{i=1}^n P(X_i | \mathbf{w}, b, \lg \varepsilon, H) \quad (13) \\ &\cdot \left( \frac{\varepsilon}{2\pi} \right)^{0.5} \exp \left[ -\frac{\varepsilon}{2} (y_i - \mathbf{w}^T \phi(x_i) - b)^2 \right]. \end{aligned}$$

In (13), let the parameter  $\mathbf{w}$  be independent of  $b$ ; thus,  $P(\mathbf{w}, b | \lg C, \lg \varepsilon, H) = P(\mathbf{w} | \lg \varepsilon, H) P(b | H)$ . Let

$P(\mathbf{w} | \lg \varepsilon, H) = P(\mathbf{w} | \lg \varepsilon, H)P(b | H)$  and  $P(b | H)$  obey a Gaussian distribution and a uniform distribution, respectively; then, (12) satisfies the following relation:

$$P(\mathbf{w}, b | D, \lg C, \lg \varepsilon, H) \propto \exp\left(-\frac{C}{2}\mathbf{w}^T\mathbf{w}\right) \exp\left(-\frac{\varepsilon}{2}\sum_{i=1}^n \xi_i\right). \quad (14)$$

Therefore, the optimal parameters corresponding to the maximum a posteriori estimation, denoted by  $\mathbf{w}_{MP}$  and  $b_{MP}$ , can be obtained by minimizing the negative logarithm of the risk function given in formula (6).

*Step 2.* Second,  $C$  and  $\varepsilon$  are deduced. The Bayesian evidence framework is introduced to deduce the parameters  $C$  and  $\varepsilon$  from the training set  $D$ .  $C$  and  $\varepsilon$  are assumed to be independent; that is,  $P(\lg C, \lg \varepsilon | H) = P(\lg \varepsilon | H)P(\lg C | H)$ . Thus, it can be seen that

$$P(\lg C, \lg \varepsilon | D, H) = \frac{P(D | \lg C, \lg \varepsilon, H) P(\lg \varepsilon | H) P(\lg C | H)}{P(D | H)}. \quad (15)$$

The optimal solution for  $(C, \varepsilon)$  can be obtained by solving the following minimization problem, adopting the same assumption as in formula (14):

$$\begin{aligned} \min J(C, \varepsilon) = & \frac{C}{2}\mathbf{w}_{MP}^T\mathbf{w}_{MP} + \varepsilon \sum_{i=1}^n \xi_i \\ & + \frac{1}{2} \sum_{i=1}^{N_{\text{eff}}} \lg(C + \varepsilon \lambda_i) - \frac{N_{\text{eff}}}{2} \lg C \\ & - \frac{n-1}{2} \lg \varepsilon, \end{aligned} \quad (16)$$

where  $\lambda_i$  denote the nonzero eigenvalues and  $N_{\text{eff}}$  is the number of nonzero eigenvalues.

*Step 3.* Finally,  $\sigma$  is deduced. For a certain submodel  $H_k$ ,  $P(H_k)$  is assumed to obey a uniform distribution, whereas  $\lg C$  and  $\lg \varepsilon$  obey the Gaussian distributions  $N(0, \sigma_{\lg C}^2)$  and  $N(0, \sigma_{\lg \varepsilon}^2)$ , respectively. From the Bayes rules, it can be deduced that

$$P(H_k | D) \propto P(D | H_k) P(H_k) \propto P(D | \lg C_{MP}, \lg \varepsilon_{MP}, H_k) \frac{\sigma_{\lg C|D} \sigma_{\lg \varepsilon|D}}{\sigma_{\lg C} \sigma_{\lg \varepsilon}}. \quad (17)$$

According to [41], the standard deviations of the Gaussian distributions can be approximated as follows:  $\sigma_{\lg C|D} \approx 2/(\gamma_{\text{eff}} - 1)$  and  $\sigma_{\lg \varepsilon|D} = 2/(n - \gamma_{\text{eff}})$ . The optimal kernel function coefficients of the SVR model can be obtained from the following formula:

$$P(H_k | D) \propto \sqrt{\frac{C_{MP}^{N_{\text{eff}}} \varepsilon_{MP}^{n-1}}{(\gamma_{\text{eff}} - 1)(n - \gamma_{\text{eff}}) \prod_{i=1}^{N_{\text{eff}}} (C_{MP} + \varepsilon_{MP} \lambda_i)}}, \quad (18)$$

where  $\gamma_{\text{eff}}$  is obtained in the parameter estimation of the Gaussian distributions;  $n$  is the number of support vectors; and  $C_{MP}$  and  $\varepsilon_{MP}$  are the optimal penalty parameter and loss parameter, respectively, obtained in the last step.

In conclusion, the steps of analyzing monitoring data for early warning of crack opening in a concrete dam using the B-SVR model can be summarized as follows. First, the prediction model expressed by formulas (10) and (11) is established using support-vector-based modeling based on measured monitoring data consisting of a series of COD. Then, the optimal parameters of the model are deduced using formulas (14), (16), and (18). Finally, the trend of crack variation can be predicted using the established model with the optimal parameters, and an early warning analysis of the variation in the structural behavior of the dam can be conducted by comparing the prediction results with the forewarning criteria expressed in formulas (3)–(5).

## 4. Example Engineering Application

*4.1. Engineering Background.* The application case considered here involves the Chencun concrete arch-gravity dam, with crest elevation of 126.3 m and a maximum dam height of 76.3 m, which is located in China. The dam consists of a total of 28 dam sections from the left bank to the right bank. The dam was built in two phases: construction was begun in 1958, halted in 1962, restarted in 1968, and essentially completed in 1972. When construction was shut down in 1962, most of the upstream dam sections had been poured to a height of 105.50 m; the corresponding zone is referred to as Phase I concrete. When construction was restarted after an interval of 6 years, the remaining Phase I concrete in the river bed sections was also poured to a height of 105.50 m before October 1970. Subsequently, the Phase II concrete in the downstream dam sections was poured very rapidly in the year 1971. The pouring temperature was up to 29°C, and the temperature of the cooling water in the preburied pipeline also reached 28–30°C because of the high summer temperatures. Some dam sections were also erected without water cooling. Therefore, the dam concrete was subjected to a relatively minor cooling effect. Moreover, the top part of the Phase II concrete was poured on top of the lower Phase I and II concrete within a very short interval of time after the lower Phase II concrete had risen up to the level of the Phase I sections. For these reasons, under the cold wave that arrived in the winter of 1971, a horizontal crack emerged as a result of the stress concentration at approximately the height of the joint between the Phase I and II concrete layers. The length of the crack is greater than 300 meters; it lies near the top of the Phase I concrete and extends from section #5 to section #28 of the dam, as seen in Figures 3 and 4.

*4.2. Crack Monitoring Data and Modeling Factors.* In this paper, a total of 1146 monitoring samples of COD, which were measured using a joint meter buried in dam section #18 from Jan-1-1973 to Dec-18-2006, are used to demonstrate the effectiveness of the proposed time-varying model for dam



FIGURE 3: Exterior photograph of the Chencun arch-gravity dam project and the crack location.

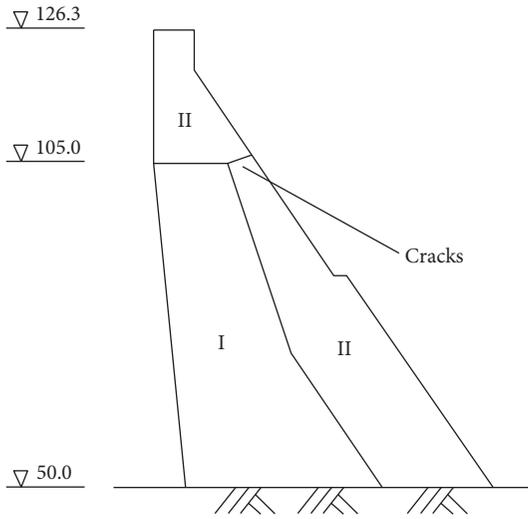


FIGURE 4: Transverse section of the Chencun arch-gravity dam.

safety prediction. The process curve corresponding to this time sequence is presented in Figure 5.

To verify the effectiveness of the proposed method, two subsequences of the COD data were selected for modeling and prediction. The modeling periods and forecast analysis periods are summarized in Table 1.

For the modeling process, the kernel function used for support vector regression was chosen to be the Gaussian radial basis function:  $K(x_i, x_j) = \exp(-|x_i - x_j|^2/2\sigma^2)$ . The factors influencing COD were decomposed in accordance with previous dam engineering and mechanical experience, allowing the monitoring model of COD to be summarized as follows:

$$\begin{aligned}
 y(\delta) &= y_H(\delta) + y_T(\delta) + y_\theta(\delta) \\
 &= \sum_{i=1}^{m_1} a_i H^i \\
 &\quad + \sum_{i=1}^{l_1} \sum_{j=1}^m \left( a_{ij} \sin \frac{2\pi i t}{365} + b_{ij} \cos \frac{2\pi i t}{365} \right) T^j \\
 &\quad + \sum_{i=1}^{l_2} \sum_{j=1}^m c_{ij} t^i T^j + \sum_{i=1}^{l_3} \sum_{j=1}^m d_{ij} H^i T^j + c_1 \theta + c_2 \ln \theta,
 \end{aligned} \tag{19}$$

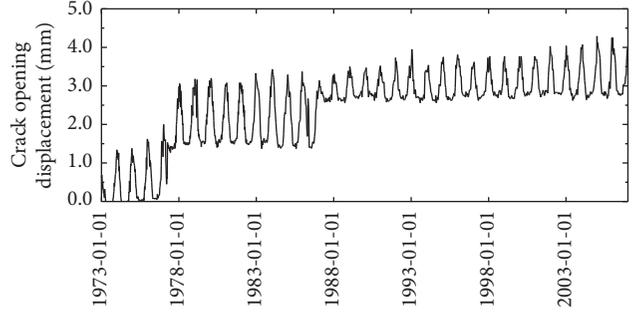


FIGURE 5: Measured process curve of COD values.

where  $y(\delta)$  is the COD as measured by the joint meter;  $H$  is the water level in the reservoir;  $T$  is the temperature change between the first recorded date and the monitoring date;  $t$  is the cumulative number of days from the first recorded date to the monitoring date;  $\theta$  is the aging factor, which is generally taken to be  $\theta = t/100$ ;  $m_1$  is a coefficient that is generally assigned a value of 3 for a gravity dam and a value of 4 or 5 for an arch dam;  $a_i, a_{ij}, b_{ij}, c_{ij}, d_{ij}, c_1,$  and  $c_2$  are the regression model parameters; and  $m, l_1, l_2,$  and  $l_3$  are the numbers of influencing factors.

These influencing factors were taken as the input vectors  $\mathbf{x}$  for the modeling of the COD. Afterward, time-varying identification and dam safety analysis were conducted based on formulas (3)–(5).

**4.3. Analysis of Optimal Modeling Parameters.** SVR dam safety monitoring modeling was performed within the Bayesian framework using the data of subsequence 1. The first 245 monitoring data points (for the modeling period of 1997–2005) were used to infer the best model parameters, resulting in values of  $\epsilon = 0.09, C = 176,$  and  $\sigma = 0.42$ . These model parameters were held fixed for the prediction of 28 subsequent displacement events in the year 2006. The average deviation between the model-predicted displacement and the monitoring results was 0.044 mm.

To verify the validity of the abovementioned optimal parameters, a series of SVR models were established using the same data samples (subsequence 1). In the modeling process, the set of values for the kernel parameter,  $\sigma$ , was preliminarily established to be  $\{0.2, 0.5, 1.0\}$  based on the previous modeling experience. Then, 75 SVR regression models were established using the same 245 data samples (subsequence 1) with a loss factor set of  $\epsilon = \{0.02, 0.05, 0.1, 0.2, 1\}$  and a penalty factor set of  $C = \{50, 100, 150, 500, 1000\}$ . The prediction deviations of each model are shown in Figure 6. The following conclusions can be drawn from Figure 6. (1) The average prediction deviations range from 0.049 mm to 0.331 mm for all 75 tests; the smallest of these values is still slightly larger than that for the B-SVR model (0.044 mm). (2) The optimal value of the SVR parameter obtained using the Bayesian framework, that is,  $\sigma = 0.42$ , is consistent with the observation that, for fixed values of  $\epsilon$  and  $C$ , the deviations decrease as  $\sigma$  increases from 0.2 to 0.5 and subsequently increase as  $\sigma$  further increases to 1.0. (3) A

TABLE 1: Regression and prediction sequences.

Monitoring sequence	Modeling period	Prediction period
Subsequence 1	1987-01-01–2005-12-31	2006-01-01–2006-12-31
Subsequence 2	1978-01-01–1986-12-31	1987-01-01–1987-12-31

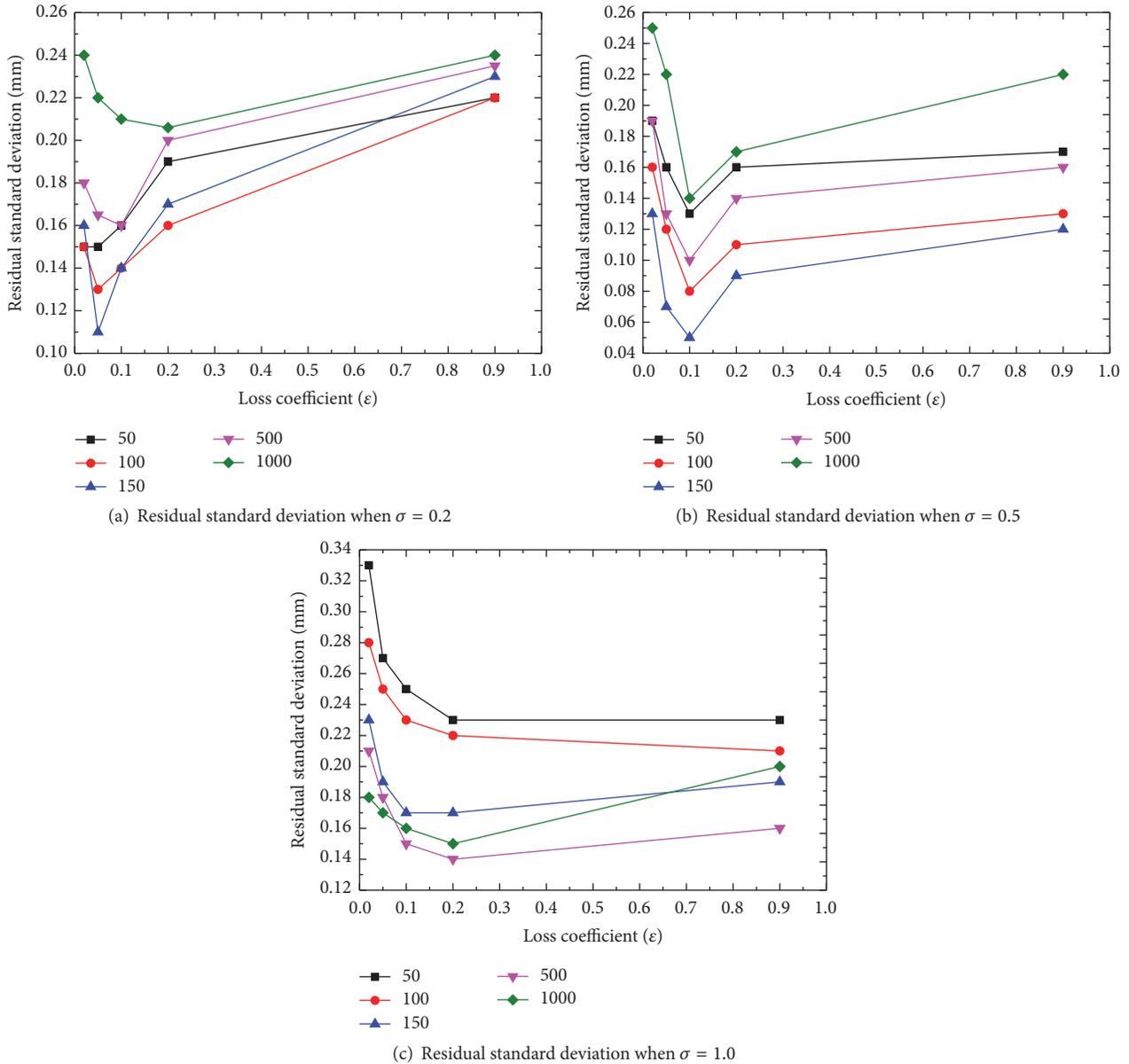


FIGURE 6: Correspondence relationships between the prediction deviations and the model parameters.

similar phenomenon can be observed if the value of  $\sigma$  is fixed. In general, the best model parameters obtained by the Bayesian framework are acceptable in comparison with the results of the 75 modeling tests.

4.4. Time-Varying Prediction Analysis Based on the B-SVR Model. To validate the effectiveness of the proposed forewarning criteria, COD modeling and prediction were

conducted based on the optimal parameters  $(\epsilon, C, \sigma) = (0.09, 176, 0.42)$  using the two subsequences represented in Table 1. The monitored values, fitted values, and residual process curves for the two subsequences are shown in Figures 7 and 8. The residual standard deviations for the monitored and model-fitted values are 0.08 mm and 0.09 mm, respectively. Then, predictions were generated for the two subsequences. The corresponding monitored values, predicted values, and residual process curves are shown in Figures 9 and 10.

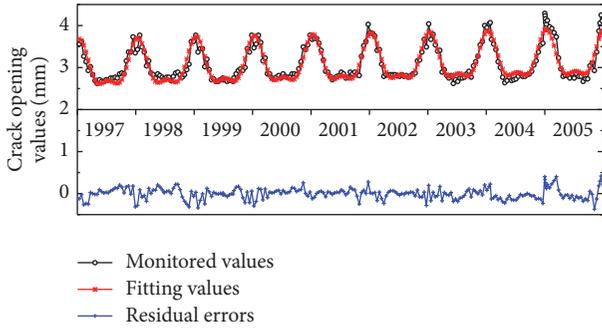


FIGURE 7: Monitored values, fitted values, and residual process curve for subsequence 1.

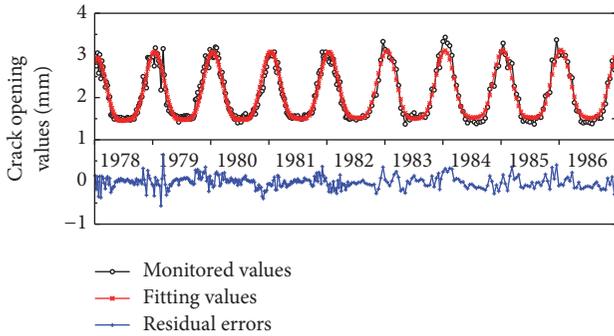


FIGURE 8: Monitored values, fitted values, and residual process curve for subsequence 2.

It can be seen from Figures 7 and 8 that a high fitting degree was obtained for both subsequences using the B-SVR model. These findings demonstrate that the established model accurately reflects the variation in the COD and can be used for forewarning analyses of dam behavior.

It can be seen from Figure 9 and the forewarning criteria given in (3)–(5) that (1) the deviation between the monitored values of the COD and the predicted values output by the model for subsequence 1 satisfies the equation  $|K(t) - \widehat{K}(t)| \leq 2S = 0.22 \text{ mm}$ ; (2) all monitored values for subsequence 1 are smaller than the maximum COD throughout the entire history; that is,  $K(t) < K_m = 4.29 \text{ mm}$ ; (3) the rate of change of the aging component remained essentially equal to zero,  $dK_\theta(t)/dt \approx 0$ . In general, these three conditions imply that the monitored process of COD (subsequence 1) is consistent with the normal pattern of dam behavior, and the time-varying crack process remained safe during the analyzed period.

However, a different result is revealed by the forewarning analysis of subsequence 2 (Figure 10): (1) a serious deviation appears between most of the monitored values of the COD and the predicted values during the prediction period for subsequence 2 (i.e.,  $|K(t) - \widehat{K}(t)| > 3S = 0.42 \text{ mm}$ ) and (2) at some time points the time-dependent component of the COD behavior satisfies the equations  $d\widehat{K}_\theta(t)/dt > 0$  and  $d^2\widehat{K}_\theta(t)/dt^2 > 0$ . According to the forewarning criteria expressed in (3)–(5), obvious abnormality is evident in the variation of the crack opening displacement. At this time, an

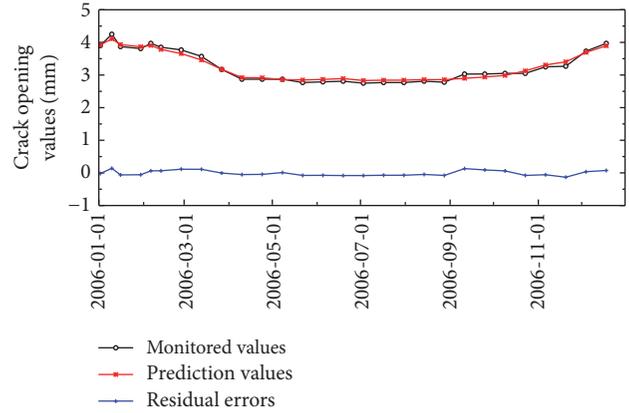


FIGURE 9: Monitored values, predicted values, and residual process curve for subsequence 1.

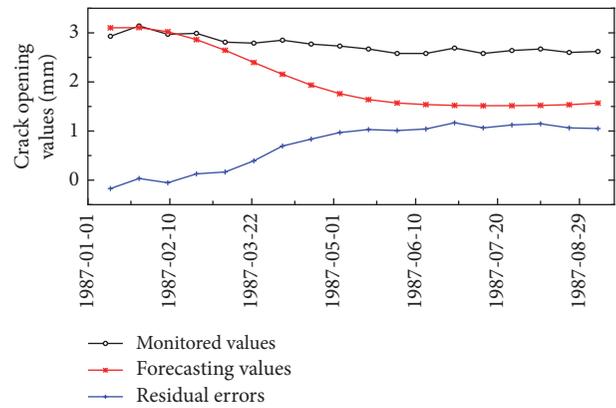


FIGURE 10: Monitored values, predicted values, and residual process curve for subsequence 2.

urgent dam safety alarm should be sent out, and a further comprehensive inspection of the overall safety conditions of the dam should be conducted immediately.

## 5. Conclusions

An accurate determination of the nonlinear relationships between the COD and its influencing factors is useful for predicting and diagnosing COD variations for a forewarning analysis of dam health. Based on a monitored COD time series, the SVR method is adopted to describe its nonlinearly varying behavior, and the BEF method is applied to determine the best SVR modeling parameters. With the establishment of a suitable mathematical model, the identification of crack behavior and the degree of abnormality of dam health can be achieved in accordance with the proposed forewarning criteria considering the predicted and monitored time-varying COD behavior.

In the case considered as an example here, the optimal modeling parameters were obtained and shown to be acceptable according to the results of 75 SVR modeling tests with a range of specified parameters. These optimal parameters were

then used for the COD modeling of two monitoring subsequences from an actual crack in the Chencun concrete dam. According to the proposed forewarning criteria, both normal and abnormal behaviors were identified. Therefore, the time-varying identification model presented here is suitable for use in evaluating the behavior of cracks in dams.

## Competing Interests

The authors declare that there are no competing interests regarding the publication of this paper.

## Acknowledgments

The authors would like to acknowledge the research funds provided by the National Natural Science Foundation of China (Grants nos. 41323001, 51479054, 51579086, 51379068, 51579083, 51279052, 51579085, and 51609074), the Jiangsu Natural Science Foundation (Grants nos. BK20140039, BK20160872), a project funded by the Priority Academic Program Development of Jiangsu Higher Education Institutions (Grant no. YS11001), the National Key Research Program (Grant no. 2016YFC0401601), and the Key Laboratory of Earth-Rock Dam Failure Mechanism and Safety Control Techniques, Ministry of Water Resources (Grant no. YK914002).

## References

- [1] ICOLD, *World Register of Dams*, International Commission on Large Dams, Paris, France, 1988.
- [2] L. M. Feng, O. A. Pekau, and C. H. Zhang, "Cracking analysis of arch dams by 3D boundary element method," *Journal of Structural Engineering*, vol. 122, no. 6, pp. 691–699, 1996.
- [3] Z. Li, C. Gu, and Z. Wu, "Nonparametric change point diagnosis method of concrete dam crack behavior abnormality," *Mathematical Problems in Engineering*, vol. 2013, Article ID 969021, 13 pages, 2013.
- [4] C. Gu, Z. Li, and B. Xu, "Abnormality diagnosis of cracks in the concrete dam based on dynamical structure mutation," *Science China Technological Sciences*, vol. 54, no. 7, pp. 1930–1939, 2011.
- [5] Z. Li, C. Gu, Z. Wang, and Z. Wu, "On-line diagnosis method of crack behavior abnormality in concrete dams based on fluctuation of sequential parameter estimates," *Science China Technological Sciences*, vol. 58, no. 3, pp. 415–424, 2015.
- [6] Y. Yao, S.-T. E. Tung, and B. Glisic, "Crack detection and characterization techniques—an overview," *Structural Control and Health Monitoring*, vol. 21, no. 12, pp. 1387–1413, 2014.
- [7] O. V. Shiryayev and J. C. Slater, "Detection of fatigue cracks using random decrement signatures," *Structural Health Monitoring*, vol. 9, no. 4, pp. 347–360, 2010.
- [8] M. I. Friswell and J. E. T. Penny, "Crack modeling for structural health monitoring," *Structural Health Monitoring*, vol. 1, no. 2, pp. 139–148, 2002.
- [9] Y. Bao, F. Tang, Y. Chen, W. Meng, Y. Huang, and G. Chen, "Concrete pavement monitoring with PPP-BOTDA distributed strain and crack sensors," *Smart Structures and Systems*, vol. 18, no. 3, pp. 405–423, 2016.
- [10] S. Ritdumrongkul and Y. Fujino, "Identification of the location and size of cracks in beams by a piezoceramic actuator-sensor," *Structural Control and Health Monitoring*, vol. 14, no. 6, pp. 931–943, 2007.
- [11] A. Hehr, M. Schulz, V. Shanov, and Y. Song, "Micro-crack detection and assessment with embedded carbon nanotube thread in composite materials," *Structural Health Monitoring*, vol. 13, no. 5, pp. 512–524, 2014.
- [12] R. Rangaraj, B. Pokale, A. Banerjee, and S. Gupta, "Investigations on a particle filter algorithm for crack identification in beams from vibration measurements," *Structural Control and Health Monitoring*, vol. 22, no. 8, pp. 1049–1067, 2015.
- [13] B. Omondi, D. G. Aggelis, H. Sol, and C. Sitters, "Improved crack monitoring in structural concrete by combined acoustic emission and digital image correlation techniques," *Structural Health Monitoring*, vol. 15, no. 3, pp. 359–378, 2016.
- [14] S. Kumar and S. V. Barai, "Equivalence between stress intensity factor and energy approach based fracture parameters of concrete," *Engineering Fracture Mechanics*, vol. 76, no. 9, pp. 1357–1372, 2009.
- [15] S. Kumar, S. R. Pandey, and A. K. L. Srivastava, "Determination of double-K fracture parameters of concrete using peak load method," *Engineering Fracture Mechanics*, vol. 131, pp. 471–484, 2014.
- [16] R. Ince, "Determination of the fracture parameters of the Double-K model using weight functions of split-tension specimens," *Engineering Fracture Mechanics*, vol. 96, pp. 416–432, 2012.
- [17] H. Bang, S.-K. Lee, C. Cho, and J. U. Cho, "Study on crack propagation of adhesively bonded DCB for aluminum foam using energy release rate," *Journal of Mechanical Science and Technology*, vol. 29, no. 1, pp. 45–50, 2015.
- [18] C. Yang, J. S. Tomblin, and L. Salah, "Stress model and strain energy release rate of a prescribed crack in scarf joint/repair of composite panels," *Journal of Composite Materials*, vol. 49, no. 29, pp. 3635–3663, 2015.
- [19] T. Bao, D. Qin, X. Zhou, and G. Wu, "Abnormality monitoring model of cracks in concrete dams," *Science China Technological Sciences*, vol. 54, no. 7, pp. 1914–1922, 2011.
- [20] S. Xu and H. W. Reinhardt, "Determination of double-K criterion for crack propagation in quasi-brittle fracture, Part III: compact tension specimens and wedge splitting specimens," *International Journal of Fracture*, vol. 98, no. 2, pp. 179–193, 1999.
- [21] Z. Li, C. Gu, and Z. Wu, "Abnormality diagnosis of cracks in the concrete based on double crack tip opening displacement criterion," *Science China Technological Sciences*, vol. 56, no. 8, pp. 1915–1928, 2013.
- [22] B. Xu, W. G. Lu, Z. C. Li et al., "Order typical small probability diagnosis method of concrete dam crack behavior abnormality," *Science China Technological Sciences*, vol. 44, pp. 1035–1042, 2014 (Chinese).
- [23] Z. R. Wu, *Safety Monitoring Theory and Its Application of Hydraulic Structures*, Higher Education Press, Beijing, China, 2003 (Chinese).
- [24] Z. Dongjian, H. Zhongyan, and L. Bo, "Arch-dam crack deformation monitoring hybrid model based on XFEM," *Science China Technological Sciences*, vol. 54, no. 10, pp. 2611–2617, 2011.
- [25] Z. R. Wu, C. S. Shen, and H. X. Ruan, "Factor selection of displacement statistical models for concrete dams," *Journal of Hohai University*, vol. 16, pp. 1–9, 1988.
- [26] 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.

- [27] A. Panizzo and A. Petaccia, *Analysis of Monitoring Data for the Safety Control of Dams Using Neural Networks*, Tsinghua University Press, 2007.
- [28] W. C. Chan, C. W. Chan, K. C. Cheung, and C. J. Harris, "On the modelling of nonlinear dynamic systems using support vector neural networks," *Engineering Applications of Artificial Intelligence*, vol. 14, no. 2, pp. 105–113, 2001.
- [29] X. Zhang, P. Wang, D. Liang, C. Fan, and C. Li, "A soft self-repairing for FBG sensor network in SHM system based on PSO-SVR model reconstruction," *Optics Communications*, vol. 343, pp. 38–46, 2015.
- [30] A. Gretton, A. Doucet, R. Herbrich, P. J. W. Rayner, and B. Schölkopf, "Support vector regression for black-box system identification," in *Proceedings of the IEEE Workshop on Statistical Signal Processing*, pp. 341–344, IEEE, Singapore, August 2001.
- [31] M. Martínez-Ramón, J. L. Rojo-Álvarez, G. Camps-Valls et al., "Support vector machines for nonlinear Kernel ARMA system identification," *IEEE Transactions on Neural Networks*, vol. 17, no. 6, pp. 1617–1622, 2006.
- [32] Z. Lu and J. Sun, "Non-Mercer hybrid kernel for linear programming support vector regression in nonlinear systems identification," *Applied Soft Computing Journal*, vol. 9, no. 1, pp. 94–99, 2009.
- [33] S. Balasundaram, D. Gupta, and Kapil, "Lagrangian support vector regression via unconstrained convex minimization," *Neural Networks*, vol. 51, pp. 67–79, 2014.
- [34] 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.
- [35] J. J. Lee, D. K. Kim, S. K. Chang, and J.-H. Lee, "Application of support vector regression for the prediction of concrete strength," *Computers and Concrete*, vol. 4, no. 4, pp. 299–316, 2007.
- [36] H. Su, Z. Wen, X. Sun, and M. Yang, "Time-varying identification model for dam behavior considering structural reinforcement," *Structural Safety*, vol. 57, pp. 1–7, 2015.
- [37] N. Cristianini and J. Shawe-Taylor, *An Introduction to Support Vector Machines and Other Kernel-based Learning Methods*, Cambridge University Press, Cambridge, UK, 2000.
- [38] H. Su, Z. Wen, and Z. Wu, "Early-warning model of dam safety based on SVM theory," *Journal of Basic Science and Engineering*, vol. 17, no. 1, pp. 40–48, 2009.
- [39] C. M. Bishop, *Neural networks for pattern recognition*, The Clarendon Press, Oxford University Press, New York, NY, USA, 1995.
- [40] R. Wang, F. Luo, H. Yang et al., "Observer design for nonlinear systems based on regression LS-SVM with Bayesian methods," *Automation & Instrumentation*, vol. 7, pp. 5–9, 2011.
- [41] J. T.-Y. Kwok, "The evidence framework applied to support vector machines," *IEEE Transactions on Neural Networks*, vol. 11, no. 5, pp. 1162–1173, 2000.



# Hindawi

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

