Research Article

A Deformation Prediction Model of High Arch Dams in the Initial Operation Period Based on PSR-SVM-IGWO

Mingjun Li,1 Jiangyang Pan,1 Yaolai Liu,1 Hao Liu,1 Junxing Wang,2 and Zhou Zhao3

1Power China Zhongnan Engineering Corporation Limited, Changsha 410019, China
2State Key Laboratory of Water Resources and Hydropower Engineering Science, Wuhan University, Wuhan 430072, China
3Changjiang Institute of Survey, Planning, Design and Research, Wuhan 430010, China

Correspondence should be addressed to Mingjun Li; 1059066537@qq.com

Received 12 August 2021; Revised 20 September 2021; Accepted 6 October 2021; Published 5 November 2021

The deformation prediction of the dam in the initial stage of operation is very important for the safety of high dams. A hybrid model integrating chaos theory, support vector machine (SVM), and an improved Grey Wolf Optimization (IGWO) algorithm is developed for deformation prediction of dam in the initial operation period. Firstly, the chaotic characteristics of the dam deformation time series will be identified, mainly using the Lyapunov exponent method, the correlation dimension method, and the Kolmogorov entropy method. Secondly, the SVM-IGWO model based on phase space reconstruction (PSR) is established for deformation forecasting of the dam in the initial operation period. Taking SVM as the core, the deformation time series is reconstructed in phase space to determine the input variables of SVM and the GWO algorithm is improved to realize the optimization of SVM parameters. Finally, take the actual monitoring displacement of Xiluodu super-high arch dam as an example. The engineering application example shows that, compared with the existing models, the prediction accuracy of the PSR-SVM-IGWO model established in this paper is improved.

1. Introduction

The establishment of an accurate deformation prediction model based on the prototype observation data is of great significance to the long-term safe operation of the dam [1]. The uncertainty in the dam construction process reduces the prediction accuracy of the numerical model (also known as the deterministic model) based on the finite element method (FEM) [2]. Moreover, conventional statistical models are difficult to adapt to the complex nonlinear relationship between multiple factors and effect sizes, and the accuracy of model predictions is hard to guarantee [3]. In addition, the existing prediction models only consider the main influencing factors, such as water pressure, temperature, and aging, but do not consider the chaotic components that may be included in the deformed time series, which further limits the improvement of fitting accuracy [4].

It is difficult to establish a stable and reliable deformation prediction model in the initial stage of high dam operation, because the initial stage of operation is a temporary stage and the most uncertain period. Studies [5, 6] have shown that nearly 50% of dam failures occur in the first 5 years of operation, which emphasizes the necessity and importance of early statistical modelling methods for dam safety monitoring. The purpose of this paper is to propose a hybrid model based on chaos theory, SVM, and IGWO algorithm and apply it to the deformation prediction of the dam in the initial operation stage.

Chaos is an irregular random behavior with initial sensitivity and ergodicity [7–9]. Some studies have shown that there is chaos in the measured displacement data of dams. The efficient extraction of the chaotic component contained in the observation data is of great practical significance for improving the accuracy of prediction models. Wei [7] added a residual prediction item to the conventional statistical model of dam deformation prediction and applied chaos theory to discuss the calculation method of the residual prediction item and thus proposed a new hybrid...
prediction model. Zhang [10] proposed a new method based on empirical mode decomposition and phase space reconstruction theory to analyze the time-varying characteristics of dams. Su [11] combined SVM with phase space reconstruction, wavelet analysis, particle swarm optimization, and other methods to establish a dam deformation prediction model. Obviously, when the phase space reconstruction method is used to design the input variable, the machine learning algorithm shows high applicability in improving the prediction accuracy. However, there is little research on the chaotic characteristics in the deformation observation data during the initial operation period of the dam.

SVMs are widely used in dam behavior prediction because of their obvious advantages in solving low-sample high-dimensional nonlinear problems [12]. The SVM can transform the original nonlinear problem into a high-latitude linear problem through the kernel parameters, so it is well adapted to the nonlinear deformation prediction [13–15]. However, the selection of the kernel parameters and penalty factors of SVM will directly affect its performance in dealing with nonlinear problems.

Grey Wolf Optimizer (GWO) is a new intelligent optimization algorithm proposed by Mirjalili with reference to the social hierarchy and hunting behavior of grey wolves [16, 17]. The algorithm realizes intelligent optimization through the process of tracking, encircling, hunting, and attacking the grey wolf population [17, 18]. Studies have shown that the GWO algorithm is better than the other evolutionary algorithms [17, 19–21] in terms of quality, speed, and stability of the final solution [17, 22, 23]. However, the possibility of premature convergence reduces the probability of the algorithm finding the global optimum. The initial population is an important factor influencing the optimization performance of an intelligent algorithm, but it is difficult to guarantee the diversity of the initial population with conventional GWO algorithms.

Many improved GWO (IGWO) algorithms have been proposed to overcome the shortcomings of conventional GWO and have been widely used to solve engineering problems. Li [24] proposed an IGWO algorithm combining tent chaos, Gaussian perturbation, and cosine control strategy and applied it to the solution of four typical complex engineering optimization problems: pressure vessel design, tension spring design, welded beam design, and truss design. Zhao [25] proposed an IGWO algorithm based on the hybrid update equation of SM algorithm and GWO algorithm and verified the better performance of this IGWO algorithm through simulation experiments. Wang [26] introduced biological evolution and natural SOF principles into the standard GWO algorithm, while Liu [27] proposed an IGWO algorithm based on the differential evolution (DE) algorithm and the OTSU algorithm. In the IGWO algorithm proposed by Mohammad [28], the movement strategy is developed by introducing the hunting (DLH) search strategy based on dimensional learning. Whether it is the performance test result of the simulation test function or the actual engineering problem solving, the IGWO algorithm proposed by the above scholars shows better convergence speed and optimization performance than other intelligent optimization algorithms, such as the DE algorithm, the PSO algorithm, and the ABC algorithm. However, there are still shortcomings in some aspects. For example, it cannot be effectively applied to all problems, and the convergence accuracy can be further improved.

To optimize the optimal kernel parameters and penalty factors of SVMs, an improved GWO (IGWO) algorithm is proposed in this paper. The IGWO algorithm enriches the diversity of the initial population through differential mutation and uses the global optimal search ability of the standard GWO algorithm to ensure convergence characteristics. On the basis of identifying the chaotic characteristics of the deformation observation data, a hybrid deformation prediction model of the high arch dam at the initial stage of operation is established.

The structure of this article is as follows: Section 2 introduces a variety of methods to identify chaotic characteristics of deformation time series. In Section 3, a new dam deformation prediction model is proposed on the basis of phase space reconstruction. An improved GWO algorithm based on the fusion of the DE algorithm and GWO algorithm is introduced to optimize the parameter settings of SVM. In Section 4, taking the measured displacement data of Xiluodu super-high arch dam in the initial operation stage as an example, the method proposed in this paper and other common methods are used to establish the deformation prediction models. The prediction accuracies of models are compared and analyzed. Section 5 summarizes the main conclusions reached here.

2. Identification of Chaotic Characteristics

Chaotic systems are deterministic and sensitive to initial conditions [8]. It is worth noting that it is neither random nor disordered. At present, the identification of chaotic characteristics of time series is mainly based on phase space reconstruction, which can obtain more hidden information by recovering the chaotic attractor in the so-called high-dimensional phase space. The Lyapunov exponent, correlation dimension, and Kolmogorov entropy of the singular attractor are calculated to correctly distinguish the chaotic system from the random system [29].

2.1. Phase Space Reconstruction. The reconstruction of the phase space is the basis for the quantitative analysis of chaotic time series, in which the embedding dimension and the delay coordinate are the two most critical parameters [30]. The time delay method is currently the most commonly used method. For univariate chaotic time series \( \{x_i, i = 1, 2, \ldots, n\} \),

\[
y_j = (x_i, x_{i+\tau}, \ldots, x_{i+(m-1)\tau}), \quad i = 1, 2, \ldots, n - (m - 1)\tau.
\]

According to Takens’ Theorem, the appropriate choice of the embedding dimension \( m \) and the delay time \( \tau \) can restore the dynamics properties of the original state space in the sense of topological equivalence.
2.1.1. Delay Time. The mutual information method [31] is introduced to determine the delay time \( \tau \) of the measured displacement sequence of the dam, shown as follows:

\[
I(\tau) = \sum_{x_i} P(x_i, x_{i+\tau}) \log \frac{P(x_i, x_{i+\tau})}{P(x_i)P(x_{i+\tau})},
\]

where \( P(x_i) \) is the normalized distribution of \( x_i \), \( P(x_i + \tau) \) is the normalized distribution of \( x_i + \tau \), and \( P(x_i, x_{i+\tau}) \) is the joint distribution of \( x_i \) and \( x_i + \tau \). The time at which the first minimum point appears in the \( \tau \sim I(\tau) \) curve is often selected as the optimal delay time.

2.1.2. Embedding Dimension. The embedding dimension \( m \) is determined by the Caao method [32]. The distance \( a(t, m) \) between the phase point and the nearest neighbor point is shown as follows:

\[
a(t, m) = \frac{\|y_i^{m+1} - y_f^{m+1}\|}{\|y_i^m - y_f^m\|},
\]

where \( y_i^{m+1} \) is the \( t \) th vector in the reconstructed phase space with the embedding dimension \( m + 1 \) and \( y_f^{m+1} \) is the nearest neighbor to \( y_i^{m+1} \), \( y_f^m \) is the nearest neighbor to \( y_i^m \) in the reconstructed phase space with the embedding dimension \( m \).

The average value \( E(m) \) of \( a(t, m) \) is calculated as follows:

\[
E(m) = \frac{1}{n - mt} \sum_{i=1}^{n-mt} a(t, m).
\]

Then, the change of \( E(m) \) is as follows:

\[
E_1(m) = \frac{E(m + 1) - E(m)}{E(m)}.
\]

When \( m > m_0 \), if \( E_1(m) \) no longer changes, \( m_0 \) represents the minimum embedding dimension. In order to avoid the situation that the change of \( E_1(m) \) is difficult to judge, the Cao method adds another definition:

\[
E'(m) = \frac{1}{n - mt} \sum_{i=1}^{n-mt} |y_i^{m+1} - y_f^{m+1}|,
\]

\[
E_2(m) = \frac{E'(m + 1)}{E'(m)}.
\]

For chaotic time series, there will always be some value of \( m \), so that \( E_2(m) \neq 1 \). Whether \( E_2(m) \) tends to be stable and whether the value of \( m \) can achieve \( E_2(m) \neq 1 \) are observed to determine the minimum \( m \) of the reconstructed phase space.

2.2. Lyapunov Exponent Method. The maximum Lyapunov exponent \( (\lambda_{\text{max}}) \) is usually regarded as an indicator of chaotic motion [33]. \( \lambda_{\text{max}} > 0 \) indicates that the system is in a chaotic state. The specific process of calculating \( \lambda_{\text{max}} \) by the small data amount method is as follows:

Firstly, determine the nearest neighbor point \( y_j \) of the phase point \( y_i \), and its short-term distance is \( d_j(0) = \|y_j - y_i\| \).

Secondly, calculate the distance \( d_j(i) = \|y_{j+i} - y_{j+i}^r\| \) between \( y_j \) and \( y_j \) after the \( i \) th discrete time steps.

Thirdly, get the average of all \( \ln d_j(i) \), which is \( y(i) \):

\[
y(i) = \frac{1}{q \Delta t} \sum_{i=1}^q \ln(d_j(i)),
\]

where \( q \) represents the number of \( i \) satisfying the condition \( d_j(i) \neq 0 \) and \( \Delta t \) represents the time interval.

Finally, calculate \( \lambda_{\text{max}}, \lambda_{\text{max}} \) is the slope of the regression line obtained by the least-squares method.

2.3. Correlation Dimension Method. The correlation dimension \( D_2 \) is mainly determined by the Grassberger-Procaccia algorithm [34]. Suppose that \( r \) is the radius of the sphere centered on \( y_i \) and \( y_j \), then the correlation integral \( C(r) \) is given by

\[
C(r) = \lim_{n \to \infty} \frac{2}{n(n-1)} \sum_{i,j=1}^n \theta \left[ r - \|y_i - y_j\| \right],
\]

where \( \theta(\cdot) \) is the Heaviside function:

\[
\theta(u) = \begin{cases} 
0, & u \leq 0, \\
1, & u > 0,
\end{cases}
\]

\[
\lim_{r \to 0} C(r) \sim r^D (r \to 0),
\]

where \( D_2 \) is the correlation dimension:

\[
D_2 = \frac{\log C_n(r)}{\log r}.
\]

Thus, draw \( \log C_n(r)/\log r \) curve, and then the value of \( D_2 \) can be determined according to the slope of the curve. As the embedding dimension \( m \) gradually increases, the slope of the curve converges, and the limit of convergence is the correlation dimension \( D_2 \). The slope of the curve of a stochastic system will continue to increase with the increase of the embedding dimension, and there will be no convergence phenomenon.

2.4. Kolmogorov Entropy Method. Kolmogorov entropy [35] describes the generation rate of chaotic orbital information over time. Kolmogorov entropy reflects the chaos level of nonlinear dynamic systems, and the Kolmogorov entropy proposed by Grassberger and Procaccia is most commonly used as its estimate. The Kolmogorov entropy could be estimated by the correlation integral method:

\[
K_2 = \lim_{r \to 0} \lim_{d \to \infty} \frac{1}{\Delta m} \log_2 C_n(r),
\]

When the embedding dimension \( m \) is continuously increasing at intervals of \( \Delta m \), the stable estimation of Kolmogorov entropy can be obtained through the equal slope regression of equation (11). It should be noted that the
minimum value of \( m \) must be an integer greater than the correlation dimension \( D_2 \).

In a phase space with embedding dimension \( i \), there is

\[
\begin{align*}
x_{ij} &= [\log_2 (r)]_{ij}, \\
y_{ij} &= [\log_2 (C(r))]_{ij}
\end{align*}
\]

where

\[
y_{ij} = a x_{ij} + b.
\]

Let \( a = D_2 \), and, for the embedding dimensions \( i \) and \( i + \Delta m \), there is

\[
K_2 = \lim_{i \to \infty} \frac{\Delta b_i}{\Delta m}.
\]

where \( \Delta b = b_i - b_{i+m} \).

The Kolmogorov entropy estimate \( K_2 \) can be used to judge the motion properties of the nonlinear system: \( K_2 = 0 \) means the nonlinear system performs regular motion, \( K_2 > 0 \) means the nonlinear system performs chaotic motion, and \( K_2 < 0 \) means the nonlinear system performs random motion.

### 3. Chaotic Time Series Prediction

For the dam deformation time series \( x_i (i = 1, 2, \ldots, n) \), when the delay time \( \tau \) and the embedding dimension \( m \) have been determined, the phase space reconstruction results of series \( x_i \) are as follows:

\[
\begin{align*}
X &= \begin{bmatrix} x_1 & x_{1+\tau} & x_{1+2\tau} & \cdots & x_{1+(m-1)\tau} \\
x_2 & x_{2+\tau} & x_{2+2\tau} & \cdots & x_{2+(m-1)\tau} \\
x_3 & x_{3+\tau} & x_{3+2\tau} & \cdots & x_{3+(m-1)\tau} \\
\vdots & \vdots & \vdots & \cdots & \vdots \\
x_{n-(m-1)\tau} & x_{n-(m-1)+\tau} & x_{n-(m-1)+2\tau} & \cdots & x_n
\end{bmatrix} \\
Y &= \begin{bmatrix} x_{2+(m-1)\tau} \\
x_{3+(m-1)\tau} \\
\vdots \\
x_{n+1}
\end{bmatrix}
\end{align*}
\]

After determining the input variables and output variables of the model, a novel model based on phase space reconstruction for dam deformation predicting is proposed.

To clarify the influence of the reconstructed phase space as an input variable on the prediction performance of the model, a conventional model with dam deformation in input variable on the prediction performance of the model, a conventional model with dam deformation in-
and the MAPE are and the better the prediction effect of the model is.

3.1. Support Vector Machine. SVM [13] usually needs to establish a suitable function \( f(x) \) to describe the nonlinear relationship between the characteristic value \( x_i \) and the target value \( y_i \), shown as follows:

\[
f(x_i) = w \cdot \varphi(x_i) + b,
\]

where \( w \) is the coefficient vector, \( \varphi(x_i) \) is the transformation function, and \( w \) and \( b \) represent the weight and bias, respectively. \( w \) and \( b \) are estimated by minimizing the regularized hazard function, shown as follows:

\[
R(w) = \frac{1}{2} \|w\|^2 + C \sum_{i=1}^{n} L_{\varepsilon}(y_i, f(x_i)),
\]

where \((1/2)\|w\|^2\) is the regularization term, \( C \) is the penalty coefficient, and \( L_{\varepsilon}(y_i, f(x_i)) \) is the \( \varepsilon \)-insensitive loss function, calculated by the following equation:

\[
L_{\varepsilon}(y_i, f(x_i)) = \max(0, |y_i - f(x_i)| - \varepsilon),
\]

where \( \varepsilon \) represents the allowable error threshold. The optimization object can be deducted as follows:

\[
\min f(w, \xi^-, \xi^+) = \frac{1}{2} \|w\|^2 + C \sum_{i=1}^{n} (\xi^+, \xi^-).
\]

It is subject to

\[
\begin{cases}
    y_i - [w \cdot \varphi(x_i)] - b \leq \varepsilon + \xi^- , \quad \xi^- \geq 0, \\
    [w \cdot \varphi(x_i)] + b - y_i \leq \varepsilon + \xi^+, \quad \xi^+ \geq 0,
\end{cases}
\]

where \( \xi^+ \) and \( \xi^- \) represent slack variables.

The key is to establish the Lagrangian function.

\[
\max H(\delta^-, \delta^+) = \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} (\delta^-_i - \delta^-_j)(\delta^-_j - \delta^-_i)K(x_i, x_j)
\]

\[
+ \sum_{i=1}^{n} y_i(\delta^-_i - \delta^+_i) - \varepsilon \sum_{i=1}^{n} y_i(\delta^-_i + \delta^+_i).
\]

It is subject to

\[
\sum_{i=1}^{n} (\delta^-_i - \delta^+_i) = 0, \quad \delta^-_i, \delta^+_i \in [0, C].
\]

Therefore,

\[
f(x) = \sum_{i=1}^{n} (\delta^-_i - \delta^+_i)K(x_i, x_j) + b,
\]

where \( K(x_i, x_j) \) represents kernel function, including polynomial, radial basis function, and sigmoid.

3.2. Improved Grey Wolf Optimizer (IGWO)

3.2.1. Grey Wolf Optimizer. Grey wolves [17, 38] have a very strict social dominance hierarchy, which is mainly divided into four parts: \( a, b, \delta, \) and \( \omega \). \( a \) is the best solution, followed by \( b \) and \( \delta \), and the remaining solutions belong to \( \omega \). The top three best wolves that are closest to their prey are \( a, \beta, \) and \( \delta \), and they guide \( \omega \) to search for prey in promising search areas [17]. During the hunting process, the wolf will update its position around \( a, \beta, \) and \( \delta \), shown as follows [17]:

\[
\vec{t}^\prime = \vec{C} \cdot \vec{p}(t) - \vec{t},
\]

where \( t \) is the current iteration number, \( \vec{p}(t) \) and \( \vec{t} \) represent the current positions of the prey and the wolf, respectively; and \( \vec{D} \) is the distance between the wolf and the prey.

The coefficient vectors \( \vec{A} \) and \( \vec{C} \) can be expressed as follows:

\[
\vec{A} = 2\vec{r}_1 - \vec{a},
\]

\[
\vec{C} = 2\vec{r}_2 - \vec{r}_1,
\]

where \( \vec{r}_1 \) and \( \vec{r}_2 \) are two vectors randomly generated between \([0, 1]\) and the convergence factor \( \vec{a} \) decreases from 2 to 0 in the iterative process.

Update the position of grey wolves.

\[
\begin{cases}
    \vec{D}_a = \vec{C}_1 \cdot \vec{X}_a - \vec{X}, \\
    \vec{D}_b = \vec{C}_2 \cdot \vec{X}_b - \vec{X}, \\
    \vec{D}_\beta = \vec{C}_3 \cdot \vec{X}_\beta - \vec{X},
\end{cases}
\]

\[
\begin{cases}
    \vec{X}_1 = \vec{X}_a - A_1 \cdot (\vec{D}_a), \\
    \vec{X}_2 = \vec{X}_b - A_2 \cdot (\vec{D}_b), \\
    \vec{X}_3 = \vec{X}_\beta - A_3 \cdot (\vec{D}_\beta),
\end{cases}
\]

\[
\vec{X}(t + 1) = \frac{\vec{X}_1 + \vec{X}_2 + \vec{X}_3}{3},
\]

where \( \vec{X}_a, \vec{X}_b, \) and \( \vec{X}_\beta \) represent the positions of \( a, \beta, \) and \( \delta \), respectively, \( \vec{X} \) represents the location of the current solution, \( \vec{C}_1, \vec{C}_2, \) and \( \vec{C}_3 \) represent randomly generated vectors, and \( \vec{D}_a, \vec{D}_b, \) and \( \vec{D}_\beta \) represent the distances of \( a, \beta, \) and \( \delta \) from other grey wolves, respectively. \( A_1, A_2, \) and \( A_3 \) are random vectors, and \( \vec{X}(t + 1) \) is the final position of \( \omega \).
three operations of mutation, crossover, and selection. The population size is \( N \), \( D \) represents the dimensionality in the solution space, \( x_i = (x_{i1}, x_{i2}, \ldots, x_{iD}) \) is the individual vector, and \( G = 0, 1, \ldots, G_{\text{max}} \) is the iteration time. \( x_{G}^i = (x_{G1}^i, x_{G2}^i, \ldots, x_{GD}^i) \) is the \( i \)-th individual in the \( G \)-th generation. \( x_L = (x_{1L}, x_{2L}, \ldots, x_{DL}) \) and \( x_U = (x_{1U}, x_{2U}, \ldots, x_{DU}) \) record the lower limit and upper limit, respectively, of search space.

The initial population \( P_0 \) is mainly randomly generated within the upper limit and lower limit \((x_1, x_2, \ldots, x_i, \ldots, x_N)\), and the \( j \)-th index of the \( i \)-th individual is obtained by the following equation:

\[
x_{j,i}^0 = x_{jL} + \text{rand}(0, 1) \times (x_{jU} - x_{jL}).
\] (34)

The mutation operation generates a new mutation vector \( v^G_j \), shown as follows:

\[
v^G_j = x^G_{j1} + F (x^G_{j2} - x^G_{j3}),
\] (35)

where \( r_1, r_2, r_3 \in \{1, 2, \ldots, N\} \) are randomly generated integers and \( F \) is the magnification ratio of the control difference vector, which is a real number with a varying range between \([0, 2]\).

The crossover operation is as follows:

\[
u^G_j = \begin{cases} v^G_j, & \text{if } \text{rand}_{ij} \leq \text{CR or } j = \text{rand}, \\ x_{j,i}^G, & \text{otherwise}, \end{cases}
\] (36)

where \( \text{CR} \) is the cross probability, which takes a value between \([0, 1]\).

Finally, the selection operation is performed. For specific problems, all mutation crossover individuals are evaluated. For specific problems, all mutation crossover individuals are evaluated. If the fitness of the current individual exceeds the previous generation, it means that the mutation crossover operation is successful, and the current individual is retained; if the current individual’s fitness is not as good as the previous generation, the better individual is retained. The individual with the optimal fitness will become the optimal value of this generation of individuals. When the termination condition is met, the evolution will stop; otherwise, the next round will continue.

### 3.2.3. Improved Grey Wolf Optimization (IGWO) Algorithm

The conventional GWO algorithm randomly generates the initial population, which may fall into the dilemma of local optimization [17, 40]. The DE algorithm generates group intelligence through mutual cooperation and competition between individuals. Based on the respective advantages and disadvantages of the GWO and DE algorithms, a more efficient improved algorithm (DE-GWO, IGWO) is proposed. Specific steps are as follows:

(1) Set the relevant parameters of the IGWO algorithm, such as the population size \( N \), the maximum number of iterations \( t_{\text{max}} \), and the upper and lower limits of the search range, \( ub \) and \( lb \).

(2) Initialize parameters \( a, A \), and \( C \). Generate intermediates (variant populations) through evolutionary mutation operations, and then generate initial population individuals through competitive selection operations, and set the iteration time \( t = 1 \).

(3) Calculate the objective function value of a single grey wolf individual and determine the best three individuals as \( X_a, X_b, \) and \( X_d \), respectively.

(4) Calculate the distance between other grey wolf individuals and optimal \( X_a, X_b, \) and \( X_d \) according to equation (31), and update the position of each grey wolf according to equation (32).

(5) Update the values of \( a, A \), and \( C \). Crossover and competition operations are applied to retain better individual positions and generate new individuals, respectively.

(6) Update the position of the first three grey wolves \( X_a, X_b, \) and \( X_d \).

(7) Determine whether the maximum number of iterations \( t_{\text{max}} \) has been reached. If yes, exit and output the current objective function value of \( X_d \); otherwise, \( t = t + 1 \), and move to the third step to continue.

### 3.3. Construction Process of PSR-SVM-IGWO Model Predicting Dam Deformation

A hybrid model combining chaos theory and SVM is proposed, and the improved GWO algorithm is used to select optimal parameters for concrete dam deformation analysis and prediction. Use the method mentioned in Section 2 to identify the chaotic characteristics of the deformed time series, and then determine the input variables of the support vector machine. According to the IGWO algorithm proposed in Section 3.2, the parameter optimization of SVM is realized. By performing training operations, a PSR-SVM-IGWO prediction model can be established. The specific process is shown in Figure 1.

(1) Obtain dam deformation data.

(2) Reconstruct the phase space of the observed displacement time series. The mutual information method and the Cao method are adopted to determine the optimal delay time \( \tau \) and the minimum embedding dimension \( m \), and the phase space is reconstructed accordingly.

(3) Identification of chaotic characteristics of the observed displacement time series. Estimate the maximum Lyapunov exponent \( \lambda_{\text{max}} \), the correlation dimension \( D_2 \), and the Kolmogorov entropy \( K \). \( \lambda_{\text{max}} > 0 \). The Kolmogorov entropy estimate \( K_2 \) is a finite positive value and the saturation of the correlation dimension indicates that the observed displacement time series has the chaotic characteristics.

(4) Determine the input and output variables according to equation (15).

(5) Divide the data set into training data set and test data set, and normalize the data set.
Start

Input the dam deformation observation data series

Reconstruct the phase space of dam deformation observation data series

Identify chaotic characteristic?

Y

Determine the input and output variables

Training data

The initial position of the wolf pack is generated by the DE algorithm

Encode the parameters

Calculate the fitness value of the wolf pack

Save the three wolves with the largest fitness value $\alpha, \beta, \delta$

Number of iterations +1

Update wolf pack location

Encode the parameters

Calculate the fitness of the wolf pack

Update $\alpha, \beta, \delta$

Does the maximum iteration number reach?

N

Y

Output the optimal parameter combination ($C, \gamma$)

End

Test data

Train the SVM model and make predictions

Performance evaluation

Figure 1: Construction process of dam deformation prediction model based on PSR-SVM-IGWO.
(6) Initialize the parameters of the grey wolf population. Set the parameters such as the population size and the maximum number of iterations to determine the scope of the search space, and at the same time encode the penalty coefficient C and the kernel parameter γ.

(7) Use DE algorithm to initialize grey wolf population.

(8) Initialize a, A, and C. Use equations (29) and (30) to initialize coefficients a and A, and the convergence factor a = 2.

(9) Calculate the fitness value of the grey wolf individual. The combination parameters (C and γ) are determined by the grey wolf’s position, and, based on this combination of parameters, the SVM model is trained with training data, and the calculated MSE is used to determine the fitness value of the grey wolf individual.

(10) Keep the three grey wolves with the best fitness value, namely, α wolf, β wolf, and δ wolf.

(11) Refer to the positions of α wolf, β wolf, and δ wolf, and use equations (31)–(34) to update the positions of individual grey wolves.

(12) Calculate the fitness values of all grey wolves. Determine the combination parameters (C and γ) based on the grey wolf’s position, and use the training data to train the SVM model based on C and γ, and determine the individual fitness value of the grey wolf through the calculated MSE.

(13) Update the positions and fitness values of α wolf, β wolf, and δ wolf.

(14) Update a, A, and C. Use equations (29) and (30) to initialize coefficients A and C, and the convergence factor a here starts from a = 2 and decreases linearly until a = 0.

(15) Determine whether the maximum number of iterations has been reached. If it is reached, output the best parameter combination (C, γ); if it is not reached, return to step 11, and increase the number of iterations by 1.

(16) Output the optimal parameter combination (C, γ), and retrain the SVM based on the training set.

(17) Use the IGWO-SVM model to predict the sample to be tested, and evaluate the performance of the model’s prediction results.

4. Case Study

4.1. Engineering Overview. The Xiluodu dam is a concrete arch dam on the Jinsha River in China, and its geographic location is shown in Figure 2. The dam crest elevation is 610.0 m, and the maximum dam height is 285.5 m. The normal water storage level is 600.0 m, and the dead water level is 540.0 m. The layout of the dam vertical line monitoring system is shown in Figure 3.

The monitoring point PL15-1 is taken as an example. From June 2014 to June 2018, the radial displacement and water level changes of the measuring point are shown in Figure 4. Take the observation data from June 2014 to December 2015 as the training set, and take the rest as the prediction set. As shown in Figure 4, the symbol “-” indicates downstream displacement, and the symbol “+” indicates upstream displacement. During the study period, the water level change range is [540 m, 610 m], and the displacement showed obvious periodic changes with the water level change, and the change range is [-60 mm, 20 mm]. The displacement of the measuring point presents an obvious periodic change with time, but the peak displacement gradually decreases with the change of time. The main reason is that the Xiluodu arch dam was in the primary impoundment stage during the study period, and the displacement was affected by creep.

4.2. Prediction Model Construction. The delay time τ of the observed displacement time series is estimated by the mutual information method, as shown in Figure 5. According to the principle of mutual information, the optimal delay time τ = 2.

According to the determined τ, the Cao method is adopted to calculate m. The change curves of $E_1(m)$ and $E_2(m)$ as m increases are shown in Figure 6. When $m > 4$, the change of $E_1(m)$ starts to become smaller, and thus the optimal embedding dimension is determined; $m = 4$. It can also be found that there are some values of $m$ such that $E_2(m) \neq 1$, which can prove that the observed displacement data series comes from a deterministic process.

According to the determined τ, the G-P method is adopted to calculate the correlation dimension of the observed displacement data series. When the value of the embedding dimension d ranges from 1 to 10, the corresponding double logarithmic curve of ln $C(r) \sim \ln(r)$ is shown in Figure 7. It is shown in Figure 7 that the correlation index is saturated to a constant. According to the definition, it is concluded that the observed displacement data series is a chaotic series and the correlation dimension $D_2 = 0.9043$.

The small data set method is introduced to calculate the maximum Lyapunov exponent $\lambda_{\text{max}}$ of the observed displacement data series, as shown in Figure 8. $\lambda_{\text{max}}$ can be calculated from the fitting straight line of the corresponding curve: $\lambda_{\text{max}} = 0.001511$. $\lambda_{\text{max}} > 0$ indicates that the observed displacement data series has chaotic characteristics.

The correlation integral method is used to calculate the Kolmogorov entropy estimate $K_2$ of the observed
displacement data series. According to the double logarithmic curve of $\ln C(r) \sim \ln r$, as shown in Figure 7, the double logarithmic curve of $\log_2 (r) \sim \log_2 (C(r))$ can be derived, so that the Kolmogorov entropy estimate $K_2$ of the observed displacement data series can be obtained when the delay time is determined, and $K_2 = 0.0306$. $K_2$ takes a finite positive value, which indicates the chaotic characteristics of the observed displacement data series at the measuring point PL15-1.

Based on the abovementioned results, it can be concluded that the observed displacement data series of the measuring point PL15-1 has chaotic characteristics, and a chaotic prediction model of dam displacement can be established.

The proposed IGWO algorithm is adopted to realize the parameter optimization of SVM. The phase space of
reconstructed observation displacement and the influencing factors of dam deformation are used as input variables, respectively, to evaluate the predicting performance of the IGWO and GWO algorithms. The population size $N = 20$, the maximum number of iterations $t_{\text{max}} = 200$, the penalty factor $C \in [0.01, 100]$, and the kernel parameter $\gamma \in [0.01, 1000]$. The calculation is terminated when the number of iterations reaches 200.

The fitness curves of the GWO and IGWO algorithms are shown in Figure 9. For the conventional model with dam deformation influencing factors as input variables, the optimal parameters $[C, \gamma]$ of SVM obtained by the IGWO algorithm are $[0.1088, 24.2515]$. For the chaotic model with the reconstructed observed deformation phase space as the input variable, the optimal parameters $[C, \gamma]$ of the SVM obtained by the IGWO algorithm are $[3.5801, 29.4460]$.

It can be seen from Figure 9 that, whether it is a conventional model with dam deformation influencing factors as input variables or a chaotic model with reconstructed observation displacement data series phase space as the input variable, the IGWO algorithm reduces the number of iterations and can find the solution closest to the best goal faster. The main explanation is that the IGWO algorithm enriches the diversity of the initial population and improves the global search capability, thereby accelerating the convergence speed and convergence accuracy.

4.3. Results. For the PSR-SVM-IGWO-based dam observation displacement prediction model, the relevant information is introduced as follows. The SVM is at the heart of this innovative combination model. The input variable is the reconstructed phase space of the measured displacement data sequence, and the IGWO algorithm is used to realize the parameter optimization of SVM. In addition, the kernel function of SVM is a radial basis function.

In order to better analyze the predictive performance of the PSR-SVM-IGWO model, the PSR-SVM and PSR-SVM-GWO models that take the reconstructed phase space of the observed displacement data sequence as input variables are established, respectively. In addition, the SVM, SVM-GWO, and SVM-IGWO models with the factors affecting dam deformation as input variables are established to explore the influence of different input variables on the accuracy of model prediction.

The prediction performances of the SVM, SVM-GWO, SVM-IGWO, PSR-SVM, PSR-SVM-GWO, and PSR-SVM-IGWO models are shown in Table 1 and Figure 10. From Table 1, we can see that when the influencing factors of dam deformation are used as the input variables of the model, the square correlation coefficient ($R^2$) is ranked from large to small as: SVM-IGWO model > SVM-GWO model > SVM model, mean absolute percentage error (MAPE) is ranked from small to large as: SVM-IGWO model < SVM-GWO model < SVM-GWO model, and mean square error (MSE) is ranked from small to large as: SVM-IGWO model < SVM-GWO model < SVM model. The SVM-IGWO model has the largest $R^2$ of 0.9747, and the smallest MAPE and MSE are 0.0569 and 6.1547, respectively. When the reconstructed phase space of the dam deformation time series is used as the input variable of the model, $R^2$ is sorted from large to small as PSR-SVM-IGWO model > PSR-SVM-GWO model > PSR-SVM model. The ranking of MAPE and MSE from small to large is PSR-SVM-IGWO model < PSR-SVM-GWO model < PSR-SVM model. The PSR-SVM-IGWO model has the largest $R^2$ of 0.9902, and the smallest MAPE and MSE are 0.0423 and 2.6086, respectively.

As shown in Figure 10, the measured displacement shows an obvious periodic change, but, due to the influence of creep, the peak displacement presents a gradual decrease trend. SVM, SVM-GWO, SVM-IGWO, PSR-SVM, PSR-SVM-GWO, and PSR-SVM-IGWO models can effectively predict the trend of dam displacement; and the PSR-IGWO-SVR model has the highest prediction accuracy and the smallest fluctuation range of the prediction error. For the conventional model, as time goes by, the deviation between the predicted value of the model and the observed value becomes larger and larger, while for the PSR model, the deviation is greatly improved. The calculation results show the following: ① The GWO optimized SVM algorithm is applicable in dam deformation prediction. ② The IGWO algorithm proposed in this paper has more outstanding optimization ability in optimizing SVM parameters than the conventional GWO. ③ For ultrahigh arch dams, the PSR model with the reconstructed phase space as the input variable has higher prediction accuracy and smaller prediction error than its corresponding conventional model with dam deformation influence factors as input variables.

In general, SVM, SVM-GWO, and SVM-IGWO models can predict dam deformation trends well. The prediction accuracy of the PSR model with the phase space of the reconstructed observation displacement data sequence as the input variable is higher, which provides a strong support for the discussion of the chaotic phenomenon in the dam deformation time series. The calculation results also prove the applicability of the GWO algorithm in the field of dam deformation prediction. IGWO has more outstanding optimization capabilities than GWO. However, the algorithm proposed in this paper has only obtained good application results on dam deformation prediction. The application prospects of dam seepage and stress prediction are temporarily unknown, which will be tried in future research work.
Conclusions

This research proposes an innovative model combining chaos theory, support vector machine, difference algorithm, and grey wolf algorithm, namely, the PSR-SVM-IGWO model, to predict dam deformation. Taking the measured displacement data of the Xiluodu super-high arch dam during the initial operation as an example, the prediction effect of the PSR-SVM-IGWO model is compared and verified. The main conclusions are as follows:

1. As the correlation dimension of the deformation time series tends to be saturated ($D_2 = 0.900$), the largest Lyapunov exponent ($\lambda_{\text{max}} = 0.001511$) is greater than 0 and the Kolmogorov entropy estimate ($K_2 = 0.0306$) is a finite positive value; it can be seen that there is chaos in the deformation observation data during the initial operation period of the dam.

2. The optimization performance of the IGWO algorithm is superior to that of the GWO algorithm. Using the DE algorithm to ensure the initial population diversity can effectively improve the grey wolf optimization algorithm's ability to find high-quality solutions.

3. The predictive performances of the PSR-SVM, PSR-SVM-GWO, and PSR-SVM-IGWO models with the reconstructed observation data sequence phase space as the input variable are superior to those of the corresponding SVM, SVM-GWO, and SVM-IGWO conventional models with deformation influence factors as input variables. When the conventional model predicts the deformation of an ultrahigh arch dam, although the accuracy meets the requirements, the predicted value will gradually deviate from the measured value as time goes by. On the contrary, it is difficult to observe such large deviations in models that adopt the reconstructed phase space of the measured displacement sequence as the input variable. Among all the models calculated in this paper, the PSR-SVM-IGWO model has the best prediction performance. In the next step, we will try to apply the method proposed in this paper to the prediction and analysis of seepage, stress, and strain of concrete dams.

Figure 9: Relation curves between fitness value and iteration number: conventional model (a) and chaotic model (b).

Table 1: Predictive performances of SVM, SVM-GWO, and SVM-IGWO models.

| Prediction model   | MSE      | MAPE   | $R^2$   |
|--------------------|----------|--------|---------|
| SVM                | 33.1295  | 0.1148 | 0.9517  |
| SVM-GWO            | 21.0568  | 0.1395 | 0.9616  |
| SVM-IGWO           | 6.1547   | 0.0569 | 0.9747  |
| PSR-SVM            | 6.6619   | 0.0618 | 0.9802  |
| PSR-SVM-GWO        | 2.9586   | 0.0444 | 0.9842  |
| PSR-SVM-IGWO       | 2.6086   | 0.0423 | 0.9902  |

Figure 10: The predicted results of all models.
Data Availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Conflicts of Interest

The authors declare no conflicts of interest.

References

[1] B. Li, J. Yang, and D. Hu, "Dam monitoring data analysis methods: a literature review," Structural Control and Health Monitoring, vol. 27, p. 3, 2019.
[2] G. Liang, Y. Hu, and Q. Li, "Safety monitoring of high arch dams in initial operation period using vector error correction model," Rock Mechanics and Rock Engineering, vol. 51, no. 8, pp. 2469–2481, 2017.
[3] F. Salazar, R. Morán, M. Á. Toledo, and E. Oñate, "Data-based models for the prediction of dam behaviour: a review and some methodological considerations," Archives of Computational Methods in Engineering, vol. 24, no. 1, pp. 1–21, 2015.
[4] C. Zhou, S.-H. Jiang, J. Huang, and F. Huang, "Prediction of groundwater levels using evidence of chaos and support vector machine," Journal of Hydroinformatics, vol. 19, no. 4, pp. 586–606, 2017.
[5] S. Y. Wu, W. Cao, and J. J. Zheng, "Analysis of working behavior of Jinping-I Arch Dam during initial impoundment," Water Science and Engineering, vol. 9, no. 3, pp. 244–248, 2016.
[6] S. M. Sayah and F. Amberg, "Behavior of the cerro del guila rcc dam during initial operation," in Proceedings of the Paper Presented at the 8th International Symposium on Roller Compacted Concrete (RCC) Dams, Kunming, China, November 2019.
[7] B. Wei, D. Yuan, Z. Xu, L. S. C. Li, and H. Monitoring, "Modified hybrid forecast model considering chaotic residual errors for dam deformation," vol. 25, no. 8, 2017.
[8] H. Su, Z. Wen, Z. Chen, and S. Tian, "Dam safety prediction model considering chaotic characteristics in prototype monitoring data series," Structural Health Monitoring, vol. 15, no. 6, pp. 639–649, 2016.
[9] T. Jin, H. Ding, H. Xia, and J. Bao, "Reliability index and Asian barrier option pricing formulas of the uncertain fractional first-hitting time model with caputo type," Chaos, Solitons & Fractals, vol. 142, 2021.
[10] Z. J. Zhang, C. S. Gu, T. F. Bao, and L. Zhang, "Application analysis of empirical mode decomposition and phase space reconstruction in dam time-varying characteristic," Science China Technological Sciences, vol. 53, no. 6, pp. 1711–1716, 2010.
[11] H. Su, X. Li, B. Yang, and Z. Wen, "Wavelet support vector machine-based prediction model of dam deformation," Mechanical Systems and Signal Processing, vol. 110, pp. 412–427, 2018.
[12] S. Djemai, B. Brahmi, and M. O. Bibi, "A primal-dual method for SVM training," Neurocomputing, vol. 211, pp. 34–40, 2016.
[13] S. A. Besseddik, R. Djekidel, and A. Ameur, "Performance of different kernel functions for Ls-Svm-Gwo to estimate flashover voltage of polluted insulators," IET Science, Measurement & Technology, vol. 12, no. 6, pp. 739–745, 2018.
[14] B. Chen, X. Fu, X. Guo, C. Gu, C. Shao, and X. Qin, "Zoning elastic modulus inversion for high arch dams based on the Psogsa-Svm method," Advances in Civil Engineering, vol. 2019, pp. 1–13, 2019.
[15] X. Liao, G. Zhou, Z. Zhang, J. Lu, and J. Ma, "Tool wear state recognition based on Gwo–Svm with feature selection of genetic algorithm," International Journal of Advanced Manufacturing Technology, vol. 104, no. 1–4, pp. 1051–1063, 2019.
[16] S.-X. Li and J.-S. Wang, "Dynamic modeling of steam condenser and design of pi controller based on grey wolf optimizer," Mathematical Problems in Engineering, vol. 2015, Article ID 120975, 9 pages, 2015.
[17] M. Li, J. Wang, J. Pan, C. Peng, and S. Li, "Dam deformation prediction based on modified grey wolf optimization algorithm and support vector machine," Journal of Physics: Conference Series 2005, vol. 1, Article ID 012084, 2021.
[18] M. Wang, H. Chen, H. Li et al., "Grey wolf optimization evolving kernel extreme learning machine: application to bankruptcy prediction," Engineering Applications of Artificial Intelligence, vol. 63, pp. 54–68, 2017.
[19] W. Deng, J. Xu, H. Zhao, and Y. Song, "A novel gate resource allocation method using improved pso-based qea," IEEE Transactions on Intelligent Transportation Systems, pp. 1–9, 2020.
[20] W. Deng, J. Xu, X.-Z. Gao, and H. Zhao, "An enhanced msope algorithm with novel multiple strategies for global optimization problems," IEEE Transactions on Systems, Man, and Cybernetics: Systems, pp. 1–10, 2020.
[21] W. Deng, S. Shang, X. Cai et al., "Quantum differential evolution with cooperative coevolution framework and hybrid mutation strategy for large scale optimization," Knowledge-Based Systems, vol. 224, 2021.
[22] K. Liao, Y. Wu, F. Miao, L. Li, and Y. Xue, "Using a kernel extreme learning machine with grey wolf optimization to predict the displacement of step-like landslide," Bulletin of Engineering Geology and the Environment, vol. 79, no. 2, pp. 673–685, 2019.
[23] Y. Wei, N. Ni, D. Liu et al., "An improved grey wolf optimization strategy enhanced Svm and its application in predicting the second major," Mathematical Problems in Engineering, vol. 2017, pp. 1–12, 2017.
[24] Y. Li, X. Lin, and J. Liu, "An improved gray wolf optimization algorithm to solve engineering problems," Sustainability, vol. 13, p. 6, 2021.
[25] J. Zhao and Z.-M. Gao, "An improved grey wolf optimization algorithm with multiple tunnels for updating," Journal of Physics: Conference Series, vol. 1678, 2020.
[26] J. S. Wang and S. X. Li, "An improved grey wolf optimizer based on differential evolution and elimination mechanism," Scientific Reports, vol. 9, no. 1, p. 7181, 2019.
[27] Y. Liu, J. Sun, H. Yu, Y. Wang, and X. Zhou, "An improved grey wolf optimizer based on differential evolution and OtSU algorithm," Applied Sciences, vol. 10, no. 18, 2020.
[28] M. H. Nadimi-Shahraki, S. Taghian, and S. Mirjalili, "An improved grey wolf optimizer for solving engineering problems," Expert Systems with Applications, vol. 166, 2021.
[29] Y. Jiang, X. Bao, S. Hao, H. Zhao, and X. Wu, "Monthly streamflow forecasting using Elm-ipsobased on phase space reconstruction," Water Resources Management, vol. 6, 2020.
[30] M. B. Kennel, R. Brown, and H. D. I. Abarbanel, "Determining embedding dimension for phase-space reconstruction using a geometrical construction," Physical Review A, vol. 45, no. 6, pp. 3403–3411, 1992.
[31] W. Wang, L. Ying, and J. Zhang, "On the relation between identifiability, differential privacy, and mutual-information
privacy,” *IEEE Transactions on Information Theory*, vol. 62, no. 9, pp. 5018–5029, 2016.

[32] L. Cao, “Practical method for determining the minimum embedding dimension of a scalar time series,” *Physica D: Nonlinear Phenomena*, vol. 110, no. 1-2, pp. 43–50, 1997.

[33] M. T. Rosenstein, J. J. Collins, and C. Luca, “A practical method for calculating largest Lyapunov exponents from small data sets,” *Physica D: Nonlinear Phenomena*, vol. 65, no. 1-2, pp. 117–134, 1993.

[34] P. Grassberger and I. Procaccia, “Characterization of strange attractors,” *Physical Review Letters*, vol. 50, no. 5, pp. 346–349, 1983.

[35] G. Benettin, L. Galgani, and J. M. Strelcyn, “Kolmogorov entropy and numerical experiments,” *Physical Review A*, vol. 14, p. 6, 1976.

[36] M. Li and J. Wang, “An empirical comparison of multiple linear regression and artificial neural network for concrete dam deformation modelling,” *Mathematical Problems in Engineering*, vol. 2019, Article ID 7620948, 13 pages, 2019.

[37] W. Yin, E. Zhao, C. Gu, H. Huang, and Y. Yang, “A nonlinear method for component separation of dam effect quantities using kernel partial least squares and pseudosamples,” *Advances in Civil Engineering*, vol. 2019, Article ID 1958173, 12 pages, 2019.

[38] Z. Zhou, C. Wang, Z. Zhu, Y. Wang, and D. Yang, “Sliding mode control based on a hybrid grey-wolf-optimized extreme learning machine for robot manipulators,” *Optik*, vol. 185, pp. 364–380, 2019.

[39] Z. Zhou, R. Zhang, Y. Wang, Z. Zhu, and J. Zhang, “Color difference classification based on optimization support vector machine of improved grey wolf algorithm,” *Optik*, vol. 170, 2018.

[40] H. Lu, “Cognitive internet of things: frameworks, tools and applications,” *Studies in Computational Intelligence*, Springer, Berlin, Germany, 2020.