water

MDPI

*Article*

# Settlement Prediction for Concrete Face Rockfill Dams Considering Major Factor Mining Based on the HHO-VMD-LSTM-SVR Model

**Xueqin Zheng [1], Taozhe Ren [1], Fengying Lv [1], Yu Wang [1] and Sen Zheng [2,3,4,5,*]**

[1] Pump-Storage Technological & Economic Research Institute, State Grid Xinyuan Group Co., Ltd., Beijing 100761, China
[2] The National Key Laboratory of Water Disaster Prevention, Hohai University, Nanjing 210098, China
[3] College of Water Conservancy and Hydropower Engineering, Hohai University, Nanjing 210098, China
[4] National Engineering Research Center of Water Resources Efficient Utilization and Engineering Safety, Hohai University, Nanjing 210098, China
[5] ENAC/IIC/LHE, Ecole Polytechnique Fédérale de Lausanne, 1015 Lausanne, Switzerland
* Correspondence: zheng_sen@hhu.edu.cn; Tel.: +86-15150688990

**Abstract:** Some important discoveries have been revealed in some studies, including that the settlement of concrete face rockfill dams (CFRDs) may cause cracks in the concrete face slabs, which may lead to dam collapse. Therefore, deformation behavior prediction of CFRDs is a longstanding and emerging aspect of dam safety monitoring. This paper aims to propose a settlement prediction model for CFRDs combining the variational mode decomposition (VMD) algorithm, long short-term memory (LSTM) network, and support vector regression algorithm (SVR). Firstly, VMD is applied in the decomposition of dam settlement monitoring data to reduce its complexity. Furthermore, feature information on settlement time series is extracted. Secondly, the LSTM and SVR are optimized by the Harris hawks optimization (HHO) algorithm and modified least square (PLS) method to mine the major influencing factors and establish the prediction model with higher precision. Finally, the proposed model and other models are applied to predict the deformation behavior of the Yixing CFRD. Prediction results indicate that the proposed method possesses particular advantages over other models. The proposed VMD-LSTM-SVR model might help to evaluate the settlement trends and safety states of CFRDs.

**Keywords:** settlement prediction model; concrete face rockfill dam; variational mode decomposition; Harris hawks optimization; support vector regression; long short-term memory; factor mining



**Citation:** Zheng, X.; Ren, T.; Lv, F.; Wang, Y.; Zheng, S. Settlement Prediction for Concrete Face Rockfill Dams Considering Major Factor Mining Based on the HHO-VMD-LSTM-SVR Model. *Water* **2024**, *16*, 1643. https://doi.org/10.3390/w16121643

Academic Editor: Giuseppe Oliveto

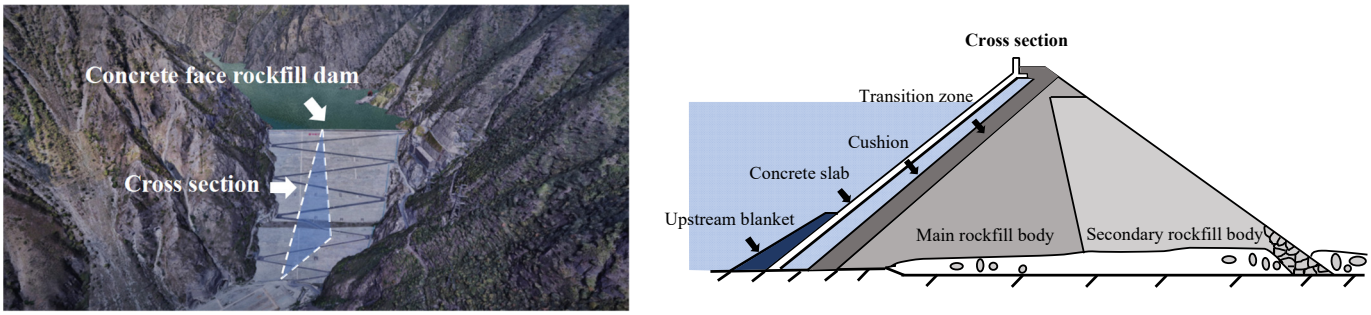
Received: 24 April 2024
Revised: 22 May 2024
Accepted: 6 June 2024
Published: 8 June 2024



## 1. Introduction

For decades, concrete face rockfill dams (CFRDs) have been widely adopted in the field of dam construction. Recent engineering applications have revealed that CFRDs become the most competitive reservoir dam type due to their safety, economy, and construction convenience [1–3]. As shown in Figure 1, CFRDs adopt a rockfill body as the support structure. The concrete slab is an important part of a concrete face rockfill dam, its major role is to prevent seepage. The findings seem to demonstrate the concrete panel is much stiffer than the rockfill body, which leads to myriad problems in compatibility and deformation of the two. The foremost problems are the facts that a large or uneven settlement under load appears in the dam [4,5]. A challenging problem that arises in this domain is that large settlements can cause cracks in the concrete panel and damage the water stop, which then affects the dam safety [6,7]. On 27 August 1993, a catastrophic dam failure occurred at the Gouhou CFRD in Qinghai Province, China [8]. This incident is the first failure of a high CFRD worldwide, resulting in widespread destruction of farmland and homes and claiming over 300 lives. Numerous dams have experienced structural cracks and extrusion damage in their face slabs. Notable examples include the Aguamilpa dam in Mexico [9], the

Tianshengqiao-I dam in China [10], and the Xingo dam in Brazil [11], all of which exhibited various degrees of structural cracking. Similar extrusion damages have been observed in several international CFRD projects. For instance, during the initial impoundment stage of the Mohale dam, the face slab along the riverbed developed extrusion damage extending to the bottom of the wave wall. Similarly, the Barra Grande and Campos Novos dams exhibited significant extrusion damage at the central vertical joints of the face slabs under initial water pressure [12]. These structural cracks and extrusion damage are primarily caused by excessive and uneven deformation of the dam body. Dams in the early stage, such as the early Salt Spring dam, the Barra Grande dam, and the Campos Novos dam during their development stages, all experienced leakage rates exceeding 1000 L/s. The leakage rates of Alto Anchacaya and Shiroro dams in the modern stage both exceeded 1800 L/s [13,14]. The occurrence of these issues is closely linked to excessive deformation of the dam body. The findings indicate that excessive deformation does not significantly impact the dam body itself, it directly damages the face slabs and joint waterstop structures, leading to substantial leakage and potentially catastrophic seepage failure. Despite better control of deformations in modern CFRDs compared to earlier counterparts, the increasing height of dams and the interplay of various complex geological conditions continue to face challenges, including excessive deformation and face slab damage.




(**a**) (**b**)

**Figure 1.** (**a**) Digital three-dimensional model of a CFRD; (**b**) the structure of the CFRD.

Consequently, a large number of alternative research have been developed over the past few years to analyze the deformation behaviors of CFRDs [15–17]. Among them, the dam prediction models enjoy great advantages in reflecting the internal relationship between settlement and influencing factors compared to other means [18,19]. As a result, various settlement prediction models for CFRDs have been developed.

Generally, one of the most popular prediction models in the dam safety monitoring area is the monitoring data-driven model. This type of prediction model can be divided into two primary categories: statistical models and artificial intelligence models [20,21]. As a common statistical model, regression models such as linear regression and polynomial regression models are utilized universally due to their quick running speed. However, it should be noted that they have low interpretability for nonlinear series, which affects their prediction accuracy. In order to rectify the problem, artificial intelligence technology is introduced to combine with the statistical models. As a result, many artificial intelligence prediction models demonstrate particular advantages in nonlinear time series prediction in the CFRD safety monitoring field over other methods [22,23].

For instance, Su et al. [24] used the wavelet support vector machine to reveal the dam settlement characteristics. The results indicated that the prediction model owed high prediction accuracy. Owing to the definite ability to reflect nonlinear relationships between variables, the artificial neural network (ANN) algorithms are utilized to establish the prediction models. The prediction results point to the likelihood that the ANN model possesses better learning ability in extracting deformation trends from dam monitoring data [25–27]. Recently, characterized by the specialized architecture and gating mechanisms, the long short-term memory (LSTM) algorithm offers distinct advantages over conventional

neural network algorithms [28]. Consequently, many researchers introduced LSTM to establish the prediction model for dam deformation. Hu et al. [29] adopted an optimized LSTM model to construct the settlement prediction model for CFRD by integrating multiple monitoring data. The findings would seem to imply that the model has a better prediction performance. Xu et al. [30] considered the impact of random error on dam deformation and utilized the LSTM model to propose a prediction model for CFRDs. The results showed that the model showed perfect long-term prediction performance.

Furthermore, according to [14], the ANN model is easy to fall into the local minima, and its convergence speed of the learning process is slow due to the CFRD monitoring data with complex features. However, the support vector regression (SVR) model can be introduced to resolve the regression problem for this kind of monitoring data under the influence of various influencing factors. Therefore, some scholars have used the SVR algorithm to predict the deformation performance of dams. Li et al. [31] presented a settlement prediction model for CFRDs by introducing an improved hybrid SVR algorithm. To improve the efficiency of the prediction model, Su et al. [32] presented an optimization method for the SVM model. Despite its advantages, these models suffer from several major drawbacks: slow convergence and weak generalizability. Due to the highly nonlinear and complexity of the monitoring deformation for CFRDs, machine learning models exhibit limitations in terms of prediction accuracy, generalization capability, and robustness [33–35]. Therefore, in our work, the deep learning models and SVR model are combined to predict the deformation of CFRDs. In particular, few studies, to our knowledge, have focused on this research topic.

Moreover, scholars have delved into mining the primary influencing factors of dam deformation by analyzing the correlation between influencing factors and deformation monitoring sequences [36]. This aims to avoid mitigating the adverse effects on prediction performance caused by excessive redundant information. For instance, Ma et al. [37] adopted the fuzzy neural network to mine the main factor affecting the dam performance. Gu et al. [38] presented a factor mining model based on a random forest algorithm and evidence theory. In recent years, some artificial algorithms such as clustering algorithms, neural networks, and genetic algorithms have been utilized in factor mining. Among these, one of the most popular ideas in factor mining is the partial least squares (PLS) method, as a multivariate statistical analysis technique, it retains the correlations among variables, reduces the dimensionality of influencing factors, and eliminates the effects of multicollinearity among factors.

Furthermore, studies have focused on the application of signal decomposition algorithms in the establishment of settlement prediction models. It has played an important role in reducing the nonlinearity of time series and refining features of monitoring data, which contributes to the interpretability of models [39,40]. Considering that analyzing the correlation between influencing factors and decomposed sub-sequences of settlement monitoring data can enable easier factor mining, in this study, the variational mode decomposition (VMD) algorithm is applied in reducing the nonlinearity and complexity of settlement monitoring time series at first. Subsequently, the improved PLS method is adopted to assess the correlation between influencing factors and decomposed sub-sequences, which could contribute a lot to enhancing the prediction accuracy of the model.

In this paper, a prediction model based on the HHO, VMD, LSTM, and SVR algorithms is proposed. Primarily, the VMD method optimized by HHO is applied to decompose the settlement sequence, thereby reducing data complexity. Secondly, PLS is employed to assess the main influencing factors as input variables for the prediction model. Subsequently, an integrated model combining the optimized LSTM and SVR models is established for dam settlement prediction. The proposed model sufficiently takes the correlation between influencing factors and dam settlement into account. In this way, by leveraging the advantages of two artificial intelligence algorithms, the settlement safety of CFRDs can be analyzed and evaluated accurately. Finally, the developed model is able to be applied to a practical CFRD health monitoring system.

The overall structure of the study takes the form of six sections, including this introductory section. In Section 2, the Hydrostatic-season-time (HST) model for the settlement of CFRDs is introduced. In Section 3, the framework of the presented model in this paper is illustrated. In Section 4, the monitoring data of a practical engineering project is utilized to test and verify the prediction performance of the proposed model. And three common machine learning models are applied in the case study to be taken as the comparisons. In Section 5, comparison results and prediction results are discussed. Ultimately, conclusions are discussed in Section 6.

## 2. The HST Model for the Settlement of CFRDs

The Hydrostatic-season-time (HST) model stands out as a frequently utilized statistical framework for the interpretation of dam settlement based on structural analysis and mechanical theory. The specific formula of the HST model is expressed as follows [41].

$$\delta(h,t,s) = \delta_H(h,t) + \delta_S(s) + \delta_T(t) + \varepsilon \tag{1}$$

For rockfill materials, settlement primarily occurs due to soil compression, and dam settlement typically depends on the load and time. The settlement of CFRDs during operation is similarly influenced by water pressure ($\delta_H$), aging effect ($\delta_S$), and temperature effect ($\delta_T$). Therefore, the settlement at a certain point of the dam can be expressed using Equation (1), with the expressions for each component as follows:

(1) Settlement caused by water pressure ($\delta_H$)

Water pressure is the predominant factor contributing to dam settlement. As water is impounded, the dam deformation due to water pressure can be categorized into three primary types: upward buoyancy deformation, water pressure deformation, and deformation due to water saturation. The upstream water pressure can be decomposed into horizontal and vertical forces, and the vertical force induces the dam settlement. Meanwhile, the dam body below the phreatic line experiences the action of buoyancy, which reduces its downward settlement. Additionally, the pore water within the rockfill decreases friction, allowing the particles to rearrange and compress under their own weight, resulting in further settlement. In summary, the settlement caused by water pressure can be represented by a Taylor series expansion, considering the first three terms and incorporating the impact of creep under water pressure over time. The influence factor is defined as the average value of upstream water levels in the early stage, the expression $\delta_H$ can be demonstrated as follows:

$$\delta_H = f\left(H, \overline{H}\right) = \sum_{i=0}^{3} a_{1i} H^i + \sum_{i=0}^{m_1} a_{2i} \overline{H}^i \tag{2}$$

where $H$ represents the upstream water level, $\overline{H}$ denotes the average value of upstream water levels in the early stage, and $m_1$ indicates the dates considered as the previous stage.

(2) Settlement caused by temperature ($\delta_T$)

The settlement induced by the linear expansion of the dam is related to temperature variations, especially in cold regions where soil tends to undergo frost heaving, leading to significant settlement. Since temperature changes generally follow an annual cycle, the relationship between settlement caused by frost heaving and temperature can be represented by a periodic function of time. Therefore, sinusoidal and cosine functions are employed to represent the temperature components. The expression for the temperature component can be demonstrated as follows:

$$\delta_T = \sum_{i=1}^{m_2} \left( b_{1i} \sin \frac{2\pi i t}{365} + b_{2i} \cos \frac{2\pi i t}{365} \right) \tag{3}$$

where $m_2$ can be taken as 1 for CFRD, $b_{1i}$, $b_{2i}$ are pending coefficients, $t$ means the time.

(3) Settlement caused by time effect ($\delta_S$)

Analysis of monitoring data from constructed panel rockfill dams indicates that after impoundment, the dam settlement appears due to the water load, the settlement usually persists for 2–3 years. In the later stages of dam settlement, the rate of settlement increase gradually decreases over time, eventually stabilizing. The functional form should conform to the actual settlement characteristics of CFRDs, thereby the expression for the aging component can be exhibited as follows:

$$\delta_S = c_1\theta + c_2 \ln\theta \tag{4}$$

where $c_1$ and $c_2$ are pending coefficients; $t$ is the time, $\theta = t/100$.

Above all, the formula of the HST model for the settlement of CFRDs can be described as follows:

$$\delta = \delta_H + \delta_S + \delta_T + cons \tag{5}$$

where *cons* denotes the constant term.

## 3. VMD-LSTM-SVR Model for the Settlement Prediction of CFRDs

In this section, the HHO algorithm, the VMD algorithm, and the LSTM and SVR models are introduced to establish the settlement prediction of CFRDs.

### *3.1. Improved VMD Algorithm Based on HHO Algorithm*

#### 3.1.1. HHO Algorithm

Research conducted by Seyedali Mirjalili in 2017 found that the hunting behavior of Harris' hawks exhibits a unique social hunting strategy involving cooperation and competition. Inspired by the hunting behavior, the Harris' hawk optimization (HHO) algorithm, which is a kind of metaheuristic algorithm [42,43], has been utilized in different fields. The HHO algorithm consists of three stages: exploration, transition, and exploitation.

(1) Exploration stage

During the exploration stage, Harris' hawks, randomly perched in a certain location, search for prey through two strategies. When iterating, they update their positions with a probability $q$.

$$X(t+1) = \begin{cases} X_{rand}(t) - r_1|X_{rand}(t) - 2r_2X(t)| & q \geq 0.5 \\ \left[X_{R(t)} - X_m(t)\right] - r_3[lb + r_4(ub - lb)] & q < 0.5 \end{cases} \tag{6}$$

where $X(t+1)$ and $X(t)$ represent the positions of individuals at the next iteration and the current iteration, respectively, $t$ denotes the iteration count, $X_{rand}(t)$ represents the position of a randomly selected individual of Harris' hawks, $X_R(t)$ represents the position of the prey, $r_1$, $r_2$, $r_3$, $r_4$, and $q$ are random numbers, ranging between 0 and 1, $X_m(t)$ denotes the average position of individuals, which can be expressed as follows:

$$X_m(t) = \sum_{k=1}^{M} \frac{X_k(t)}{M} \tag{7}$$

where $X_k(t)$ represents the position of the $k$-th individual in the population, and $M$ denotes the population size.

(2) Transition stage

Since the energy expended by the prey during escape can be converted between exploration and different exploitation behaviors, the HHO algorithm defines its escape energy $E$ as follows:

$$E = 2E_0\left(1 - \frac{t}{T}\right) \tag{8}$$

where $E_0$ represents the initial escape energy of the prey, the value of $E_0$ can be selected randomly within the range $[-1, 1]$. When $-1 < |E_0| < 0$, the prey is in a state of weak energy; when $0 \leq |E_0| < 1$, the prey is in a state of energy recovery. The escape energy $E$

exhibits a decreasing trend during the iteration process. When $|E| < 1$, Harris' hawks search for the position of prey in different areas, the HHO algorithm executes the exploration stage. When $|E| < 1$, the HHO algorithm performs local search on adjacent solutions, and executes the exploitation stage. At this point, the update of individual positions varies according to the value of $E$.

(3) Exploitation stage

Define a random number $z$ within the range of [0, 1]. Different exploitation strategies can be selected based on the random number $z$. When $0.5 < |E| < 1$ and $z \geq 0.5$, the algorithm adopts soft besieging technique for location updating:

$$X(t+1) = \Delta X(t) - E|JX_R(t) - X(t)| \tag{9}$$

where $\Delta X(t) = X_R(t) - X(t)$ represents the difference between the prey's position and the current position of the individual, $J$ is a random number between 0 and 2, indicating the jump distance of the prey during its escape process.

When $|E| < 0.5$ and $z \geq 0.5$, the algorithm adopts the hard besieging method for location updating:

$$X(t+1) = X_R(t) - E|\Delta X(t)| \tag{10}$$

When $0.5 \leq |E| < 1$ and $z < 0.5$, the algorithm adopts the soft besieging strategy with asymptotic rapid descent for updates. The first update strategy is as follows:

$$Y = X_R(t) - E|JX_R(t) - X(t)| \tag{11}$$

If the first strategy is ineffective, the second strategy is executed:

$$Z = Y + S \times LF(D) \tag{12}$$

Therefore, the final update strategy for this stage is as follows:

$$X(t+1) = \begin{cases} Y & f(Y) < f[X(t)] \\ Z & f(Z) < f[X(t)] \end{cases} \tag{13}$$

where $f(\ )$ represents the fitness function, $Y$ and $Z$ denote the current fitness positions of individuals, $D$ denotes the dimensionality of the problem, $S$ represents a $1 \times D$ random vector, and $LF(\ )$ is the mathematical expression for Levy flight.

When $|E| < 0.5$ and $z < 0.5$, the algorithm adopts the hard besieging strategy with asymptotic rapid descent for location updating:

$$X(t+1) = \begin{cases} Y & f(Y) < f[X(t)] \\ Z & f(Z) < f[X(t)] \end{cases} \tag{14}$$

$$Y = X_R(t) - E|JX_R(t) - X_m(t)| \tag{15}$$

$$Z = Y + S \times LF(D) \tag{16}$$

3.1.2. Improved VMD Algorithm

The VMD algorithm involves numerous parameters, and the selection of the number of modal components $K$ and the penalty factor $\alpha$ play an important role in the decomposition effectiveness. If the number of modal components $K$ is too small, it may fail to capture all the details and features of the sequence, leading to information loss. Conversely, if $K$ is too large, it may result in excessive information redundancy, increasing computational and analytical complexity. For the penalty factor $\alpha$, an excessively large value can overly smooth the decomposed modal components, leading to the loss of crucial information. On the contrary, if $\alpha$ is configured to be too small, it may introduce significant instability and roughness to the sub-sequences, exacerbating the impact of data noise on the decomposition

results. Hence, determining appropriate values for $K$ and $\alpha$ is essential before performing data decomposition. Typically, these parameters are determined empirically, which is subjective and can affect the generality of the model. In this paper, the HHO algorithm is applied to the parameter selection of the VMD algorithm by selecting minimum average envelope entropy.

In the signal decomposition, the smaller the envelope entropy of each sub-sequence, the lower the complexity of the decomposed sequence. Consequently, the average envelope entropy is utilized as the fitness function. Specifically, for a signal $X(t)$, the average envelope entropy of its decomposed sequences can be calculated through the following steps:

(1) Calculate the envelope $E(t)$ of a signal using the Hilbert transform.

(2) Segment the signal into $M$ data blocks according to a fixed length.

(3) For each data block, compute its probability density function $p_i(n)$, where $n$ represents the envelope falling within the $n$-th interval.

(4) Calculate the envelope entropy using the following formula:

$$H_E = -\sum_{i=1}^{M}\sum_{n=1}^{X} p_i(n)\ln_2(p_i(n)) \tag{17}$$

where $X$ denotes the number of internals.

In this study, the variance percentage (VP) indicator is applied to calculate the proportion of each sub-sequence in the original sequence. Firstly, the variance of each component and the total variance of the original sequence are computed to obtain the proportion of each sub-sequence in the original sequence. Secondly, the proportion of each sub-sequence is multiplied by the envelope entropy of the sub-sequence and summed to obtain the average envelope entropy of the decomposed sub-sequences. The main steps of the optimization process of the VMD method based on HHO are as follows:

(1) Initialize the parameters of the HHO algorithm, including the number of population quantities, maximum iteration times, and the population location.

(2) Set the fitness function as the average envelope entropy and calculate the fitness value of each hawk to determine the optimal location of an individual.

(3) Estimate $|E|$ and $z$. Update the location according to Equations (9)–(16).

(4) Calculate the updated fitness values of the Harris hawks' individual. If the fitness of the new individual surpasses that of its predecessor, the original location is replaced by the new location. Otherwise, the original location is retained.

(5) Adjust the search range and direction of the particle swarm adaptively through the algorithm's fitness. Continuously search for the minimum fitness value according to the fitness function.

(6) Evaluate if the maximum iteration times has been reached. If affirmative, output the current optimal individual's location. If negative, repeat operations from step (2) to step (5).

(7) Obtain the parameters corresponding to the minimum fitness as the optimal parameters for the VMD algorithm.

Subsequently, the optimized VMD algorithm can be applied to decompose the settlement monitoring data series of CFRD.

### *3.2. HHO-LSTM-SVR Prediction Model for the Settlement of CFRD*

#### 3.2.1. LSTM Network

To resolve the problem of vanishing or exploding gradients during model training, scholars have introduced "gate" control structures on the hidden layer nodes of Recurrent Neural Networks (RNNs). This innovation aims to regulate the flow of information, proposing the LSTM network. The LSTM network exhibits strong temporal dependencies that are capable of storing relevant information over arbitrary time intervals, thereby effectively extracting patterns from historical data. Moreover, it has demonstrated promising results in practical applications [44,45].

As depicted in Figure 2, the architecture of the LSTM network operates with a unidirectional hidden layer. In Figure 2, $x_t$ represents the input value at time step $t$, $y_t$ denotes the output value at time step $t$, $S_t$ signifies the cell state at time step $t$, $h_t$ signifies the output value of the hidden layer at time step $t$.

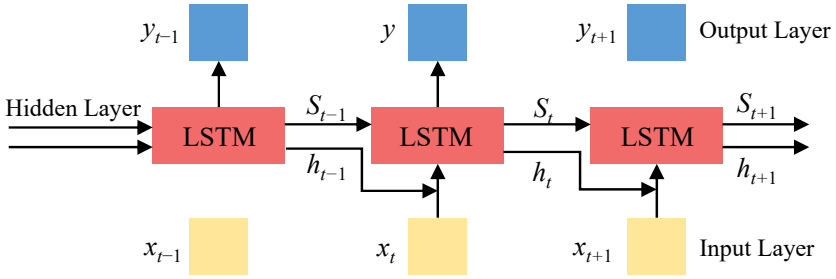



**Figure 2.** The structure of the LSTM network.

### 3.2.2. SVR Algorithm

The support vector regression (SVR) algorithm employs a nonlinear mapping technique to project training samples into a high-dimensional space, facilitating linear regression in this space without introducing additional computational complexity [46]. The SVR algorithm does well in addressing nonlinear regression prediction tasks, particularly demonstrating strong performance with moderate- to small-scale datasets. The SVR algorithm exhibits robustness and is effective in mitigating the issue of local extremum. Moreover, the SVR algorithm shows promising results in practical applications [47].

The computational expression of the SVR model can be represented as follows:

$$y(x) = \sum_{j=1}^{N} \left[ \left( \alpha_j^* - \alpha_j \right) \times K\left( x_j, x \right) \right] + d \tag{18}$$

where $x$ is the input vector, $y$ is the corresponding output value, $\alpha_j$ and $\alpha_j^*$ are Lagrange multipliers, $K\left( x_j, x \right)$ is the kernel function, $d$ is the bias term, $N$ is the total number of samples in the dataset.

### 3.2.3. HHO-LSTM-SVR Model

With the aim of determining the parameters of the LSTM and SVR models effectively, the HHO algorithm is utilized. In this work, the HHO-LSTM and HHO-SVR models are presented, respectively. By integrating multiple homogeneous or heterogeneous learners, the prediction performance of individual learners can be enhanced effectively [48]. Consequently, the merits of HHO-LSTM and HHO-SVR models can be further integrated by using partial least squares (PLS) regression. The HHO-LSTM-SVR model is proposed to predict the dam settlement with high accuracy.

The computational workflow of the HHO-LSTM-SVR model begins with the parallel computation of the HHO-LSTM and HHO-SVR models. Subsequently, the sum of squared prediction errors of the two single models and a penalty term of a weighting coefficient is selected as the objective function. Then, the optimal weights for each single model are computed to minimize the objective function, as shown in Equation (19). Finally, on the basis of the weights, the HHO-LSTM-SVR model is tested. The PLS regression can effectively address the issue of excessive bias weighting due to differences in the predictive accuracy of single models in ensemble combination models. The computational process of the HHO-LSTM-SVR model is illustrated in Figure 3.

$$\min Q = \sum_{k=1}^{m} [\xi y_{1k}^* + (1 - \xi) y_{2k}^* - y_k]^2 + \lambda \|\xi\|^2 \tag{19}$$

where $y_k$ is the actual value, $y_{1k}^*$ and $y_{2k}^*$ are the prediction values of the HHO-LSTM model and the HHO-SVR model, respectively; $\xi$ presents the weight of the model, $\lambda$ is the coefficients of the penalty term, and $m$ is the number of the sample.

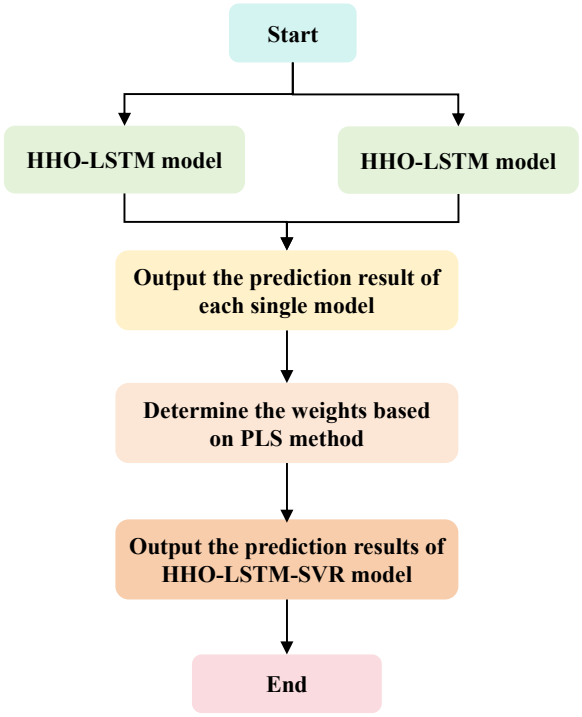



**Figure 3.** The calculation flow of the HHO-LSTM-SVR model.

### *3.3. HHO-VMD-LSTM-SVR Model Considering Factor Mining for Settlement Prediction of CFRD*

Based on the methods and theories discussed above, the proposed workflow framework for the settlement prediction model of CFRDs is illustrated in Figure 4. The specific steps of the process are outlined as follows:

(1) Select the target monitoring points and preprocess the monitoring data of water level, temperature, dam settlement, and time series data. The settlement sequences are decomposed using the VMD method optimized by the HHO algorithm, yielding several IMF sub-sequences (refer to Modules 1 and 2 in Figure 4).

(2) Determine the optimal number of input influencing factors for each IMF sub-sequence through cross-validation. Subsequently, the PLS algorithm is utilized to extract the optimal input influencing factor set for each IMF sub-sequence, achieving the mining of influencing factors (refer to Module 3 in Figure 4).

(3) Analyze the local characteristics of each IMF sub-sequence and establish the HHO-LSTM-SVR prediction model. The iterations of these steps are executed to construct and optimize the HHO-LSTM-SVR prediction model for each IMF component (refer to Module 4 in Figure 4).

(4) Predict each IMF sub-sequence component using the HHO-LSTM-SVR model sequentially to obtain output sequences. These sequences are then superimposed to form the final prediction results (refer to Module 4 in Figure 4).

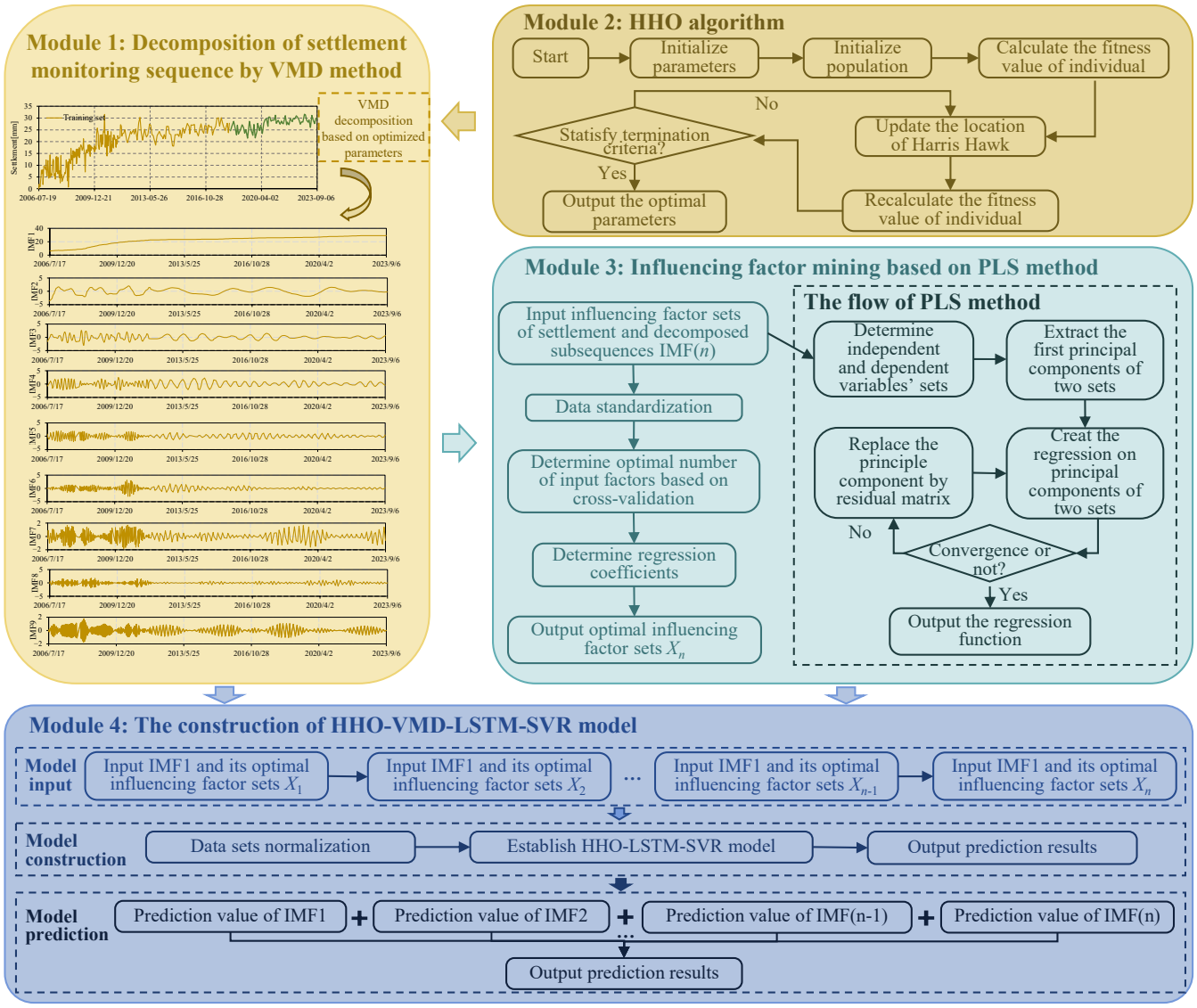



**Figure 4.** The flowchart of the HHO-VMD-LSTM-SVR model construction of settlement prediction for CFRDs (in Module 1, yellow line represents the fitting data sets, and green line represents the prediction data sets).

## 4. Project Overview

The Yixing pumped-storage power station (PSPS) is situated on the southwest side of Tongguan Mountain in Yixing city, Jiangsu Province. The station comprises upper and lower reservoirs, water transmission systems, underground powerhouses, and switchyards. With a total installed capacity of 1000 MW and an annual electricity generation of 1.49 billion kWh, the station primarily serves peak shaving, valley filling, frequency regulation, phase adjustment, and emergency backup functions. The upper reservoir is located in the valley on the north side of Tongguan Mountain, surrounded by the main dam, auxiliary dam, and surrounding mountains. The main dam of the upper reservoir is a concrete face rockfill dam, with a crest elevation of 474.20 m and a crest length of 494.90 m. The schematic of the dam structure and distribution of monitoring points of deformation for Yixing PSPS are exhibited in Figure 5b,c, separately.

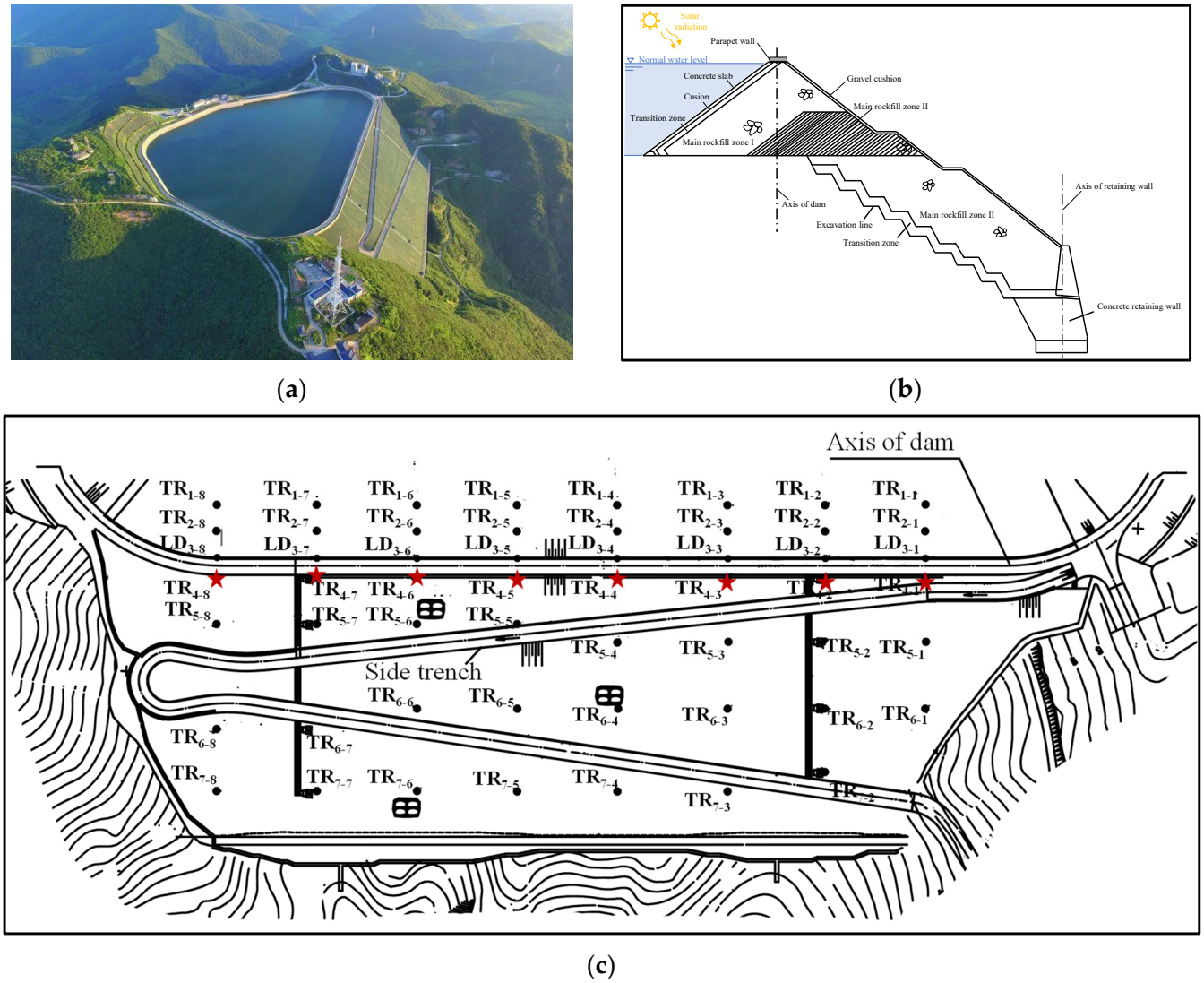



**Figure 5.** (**a**) Photo of the upstream reservoir and CFRD, (**b**) schematic of structure of CFRD, (**c**) schematic of monitoring points of deformation of CFRD for Yixing PSPS (the locations of the selected monitoring points ($TR_{4\text{-}1}$ to $TR_{4\text{-}8}$) are marked with red star-shaped labels).

The settlement data of $TR_{4\text{-}1}$ to $TR_{4\text{-}8}$ from 4 December 2006 to 6 September 2023 are selected for analysis in this case study. The locations of the monitoring points ($TR_{4\text{-}1}$ to $TR_{4\text{-}8}$) are marked in Figure 5c with red star-shaped labels. After data preprocessing, the valid settlement data of the selected monitoring points are shown in Figure 6. The upstream water level and the environmental temperature in Yixing PSPS are illustrated in Figure 7. It can be seen that all the settlement monitoring data increase slowly, and tend to converge with time.

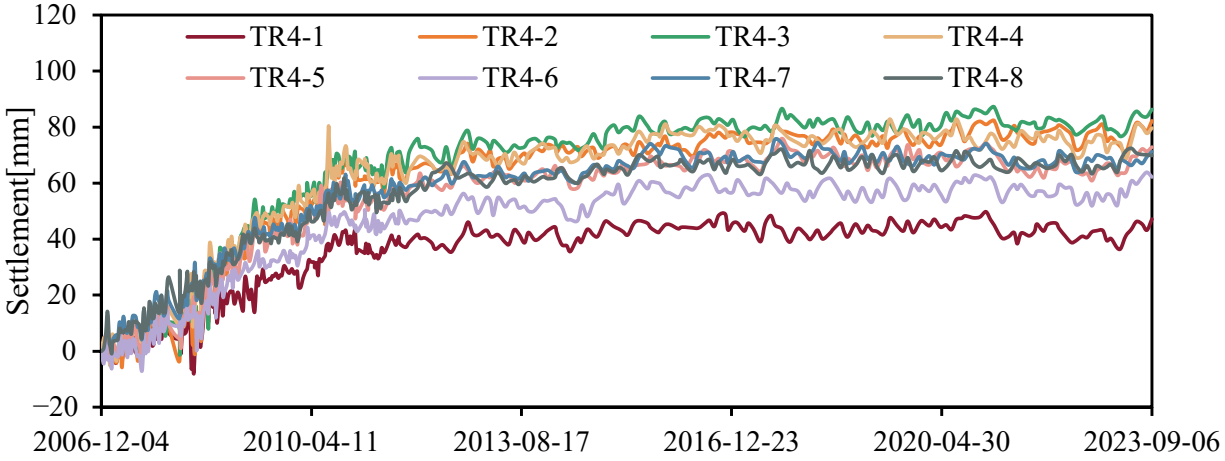



**Figure 6.** Settlement monitoring data sequences of the selected monitoring points.

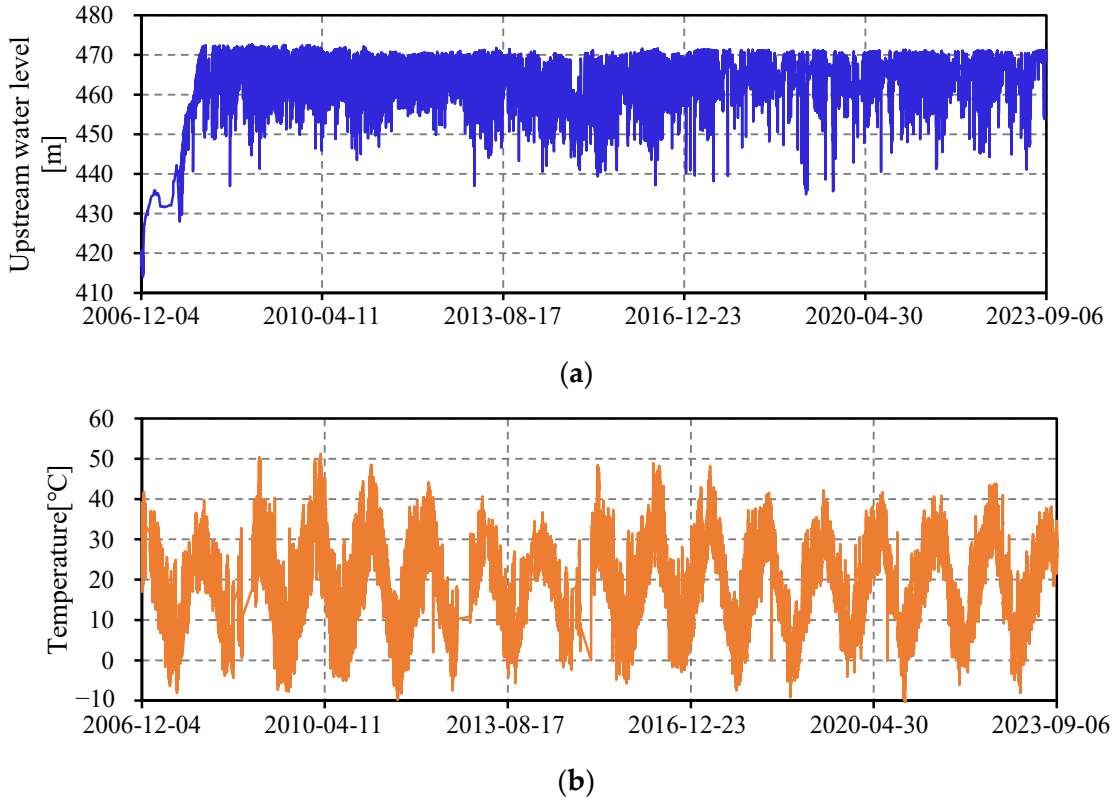



**Figure 7.** Monitoring data series: (**a**) upstream water level, (**b**) environmental temperature.

## 5. Results and Discussion

### *5.1. Decomposition of Settlement Sequences Based on VMD Method*

The parameters of the VMD method are determined by using the HHO algorithm. The optimization process aims to synthesize the most suitable parameters based on the complexity of the data sequences, employing the average envelope entropy as the fitness function. Ultimately, the optimal number of modal components $k$ is set at 9, and the penalty factor $\alpha$ is set at 664. The optimization process of the HHO algorithm is exhibited in Figure 8.

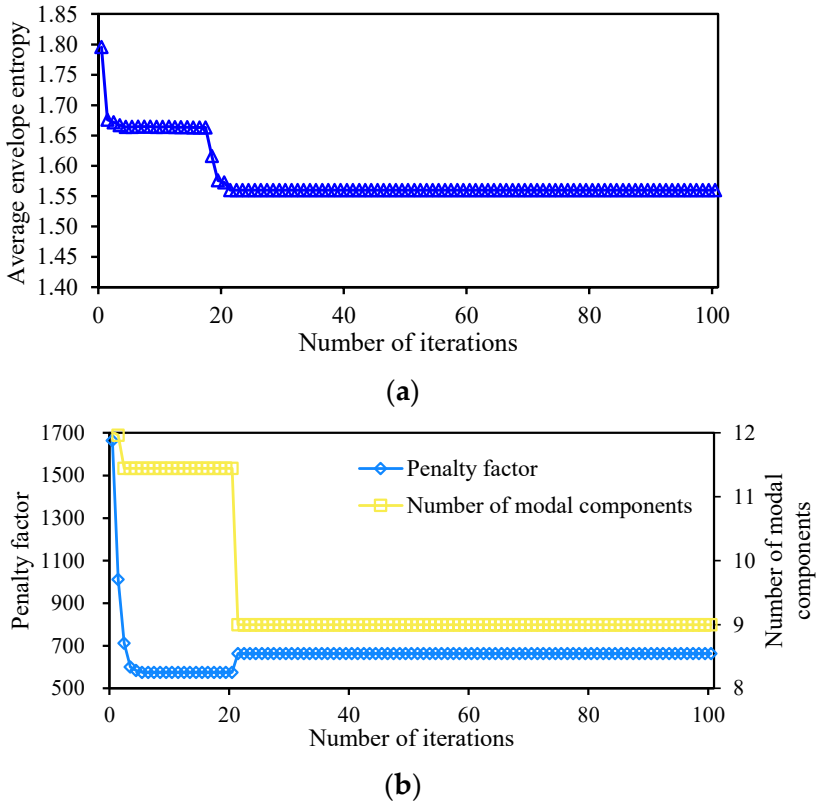



**Figure 8.** The optimization process of the HHO algorithm. (**a**) The variation process of average envelope entropy, (**b**) the optimization process of penalty factor and number of modal components.

For example, Figure 9 depicts the decomposition results by the VMD method of the settlement monitoring sequence at monitoring point $TR_{4\text{-}7}$. Based on the characteristic features of the decomposition sub-sequences, IMF1 is defined as the trend component of the original sequence, portraying the sustained deformation of the rockfill body of the Yixing CFRD underwater load, particularly under prolonged duration and high water levels. The deformation characteristic exhibits a rheological characteristic with a trending tendency. IMF2 and IMF4 represent periodic sequences gradually, influenced significantly by cyclic factors such as upstream water level and temperature changes. IMF5 to IMF9 demonstrate characteristics of short-period and low-amplitude oscillations, possibly attributed to rapid and transient deformation components induced by factors like seismic events, self-weight, or monitoring errors in dam structures. Due to the differing primary influencing factors of each component, it is essential to construct and optimize the factor sets for the decomposed sub-sequences, subsequently establishing a corresponding prediction model.

### *5.2. Major Factor Mining Based on PLS Method*

To simultaneously consider the correlation between multiple independent variables and dependent variables, partial least squares (PLS) regression is introduced for morphological sequence influence factor mining. To prevent overfitting, the first 80% of the data is selected for training to obtain the input primary influence factors for each decomposed sub-sequence. Furthermore, to accurately determine the number of input impact factors for different decomposed sub-sequences and to improve the fitting effects of sub-sequences, cross-validation is employed to obtain the mean absolute error, serving as the criterion for determining the optimal number of input factors for each sub-sequence.

Finally, the optimal number of input impact factors and the corresponding regression coefficients are determined. By computing the mean absolute error under different numbers of input impact factors, the primary influencing factors are identified according to the

minimum mean absolute error. When the mean absolute errors are close, fewer input impact factors should be selected to enhance the modeling efficiency. Taking the determination process of the primary influence factor mining for IMF1 and IMF2 as examples, Figure 10 illustrates the mean absolute error of IMF1 and IMF2 under different numbers of impact factors. It is evident that IMF1 achieves the minimum mean absolute error of 0.748 when the number of input factors is 3. As for IMF2, the minimum mean absolute error is attained when the number of input influence factors is 13 or 15. However, it is noted that the mean absolute error is close to the minimum value when the number of influence factors is 3. Considering training and prediction efficiency, the number of influence factors should be taken as 3. Similarly, the optimal number of input influence factors and mean absolute errors for each sequence are obtained, as shown in Table 1.

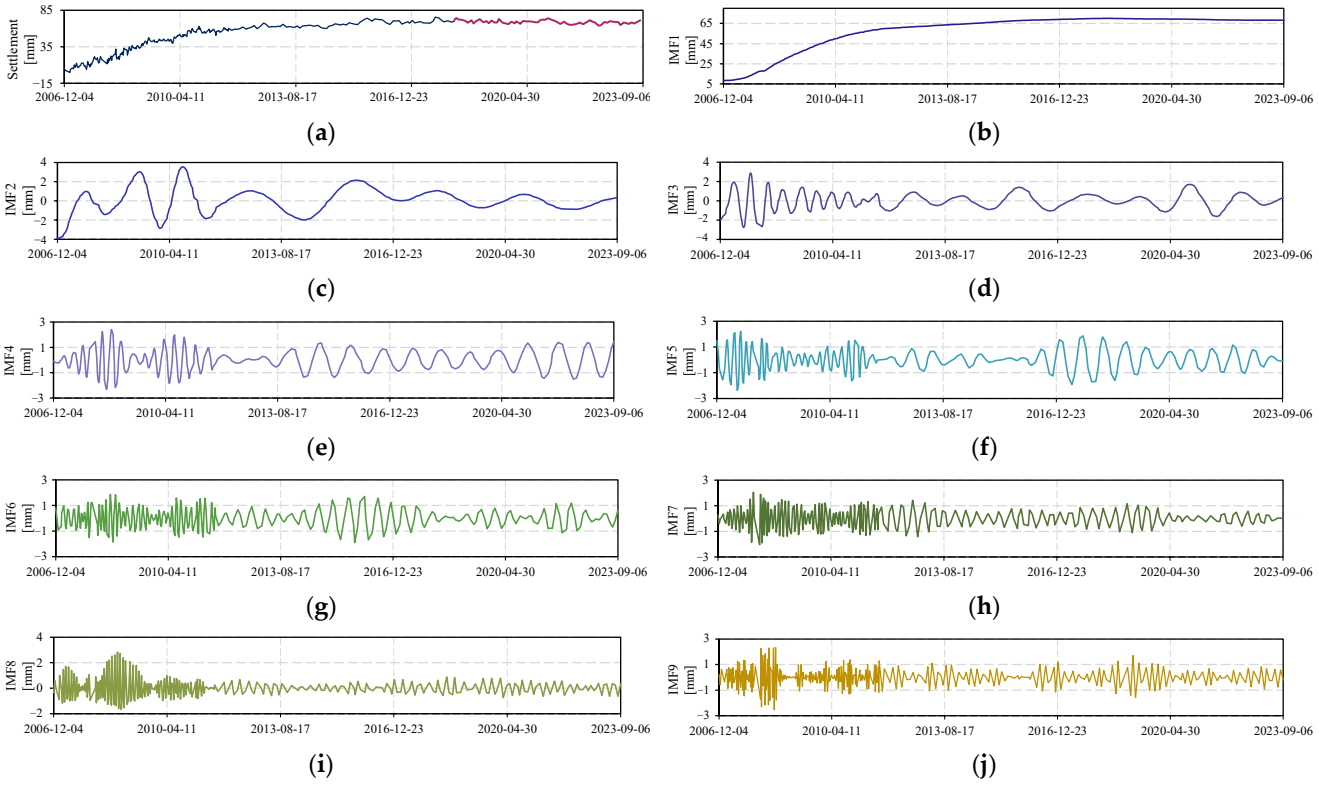



**Figure 9.** The results of VMD decomposition. (**a**) Monitoring data sequence (blue line represents fitting data sets, and red line represents the prediction data sets), (**b**) decomposition subsequence IMF1, (**c**) decomposition subsequence IMF2, (**d**) decomposition subsequence IMF3, (**e**) decomposition subsequence IMF4, (**f**) decomposition subsequence IMF5, (**g**) decomposition subsequence IMF6, (**h**) decomposition subsequence IMF7, (**i**) decomposition subsequence IMF8, (**j**) decomposition subsequence IMF9.

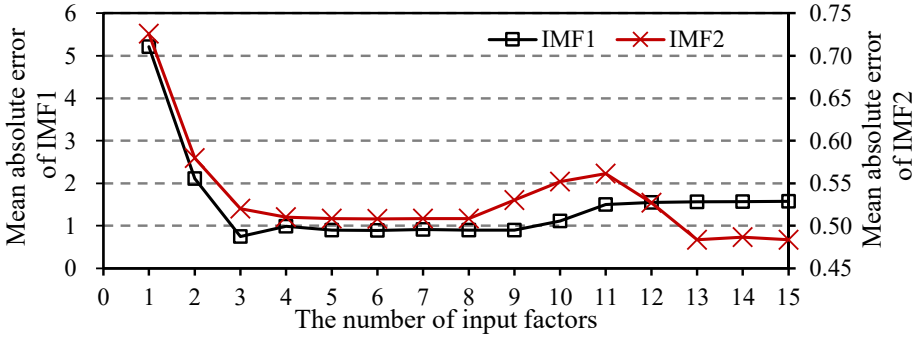



**Figure 10.** Mean absolute errors of IMF1 and IMF2 corresponding to various numbers of input factors.

**Table 1.** Number of the optimal input influence factors of the decomposition sequence and the corresponding average absolute errors.

| Decomposition Sequence | Mean Absolute Error | Number of the Optimal Input Influence Factors | Decomposition Sequence | Average Absolute Error | Number of the Optimal Input Influence Factors |
|---|---|---|---|---|---|
| IMF1 | 0.748 | 3 | IMF6 | 0.051 | 1 |
| IMF2 | 0.520 | 4 | IMF7 | 0.059 | 4 |
| IMF3 | 0.419 | 3 | IMF8 | 0.063 | 5 |
| IMF4 | 0.157 | 5 | IMF9 | 0.061 | 6 |
| IMF5 | 0.062 | 1 | - | - | - |

### *5.3. Fitting and Prediction Results*

On the basis of parameter optimization by the HHO algorithm, the selected hyperparameters of the LSTM model are set as follows: hidden layers can be taken as 2 (each hidden layer has 8 neurons), iteration period is 1000, initial learning rate is 0.05, and descending factor is 0.15. In the same way, the hyperparameters of the SVR model are set as follows: penalty parameter $C$ can be selected as 0.2, kernel function parameter $\gamma$ can be taken as 0.8. 80% of the monitoring data are applied to train the prediction model.

Take the monitoring settlement data of monitoring point $TR_{4\text{-}7}$ again, all the prediction results of the HHO-LSTM-SVR model for each sub-sequence are illustrated in Figure 11. Comparisons between the true monitoring values and prediction values of settlement reveal that the model exhibits great fitting and prediction accuracy.

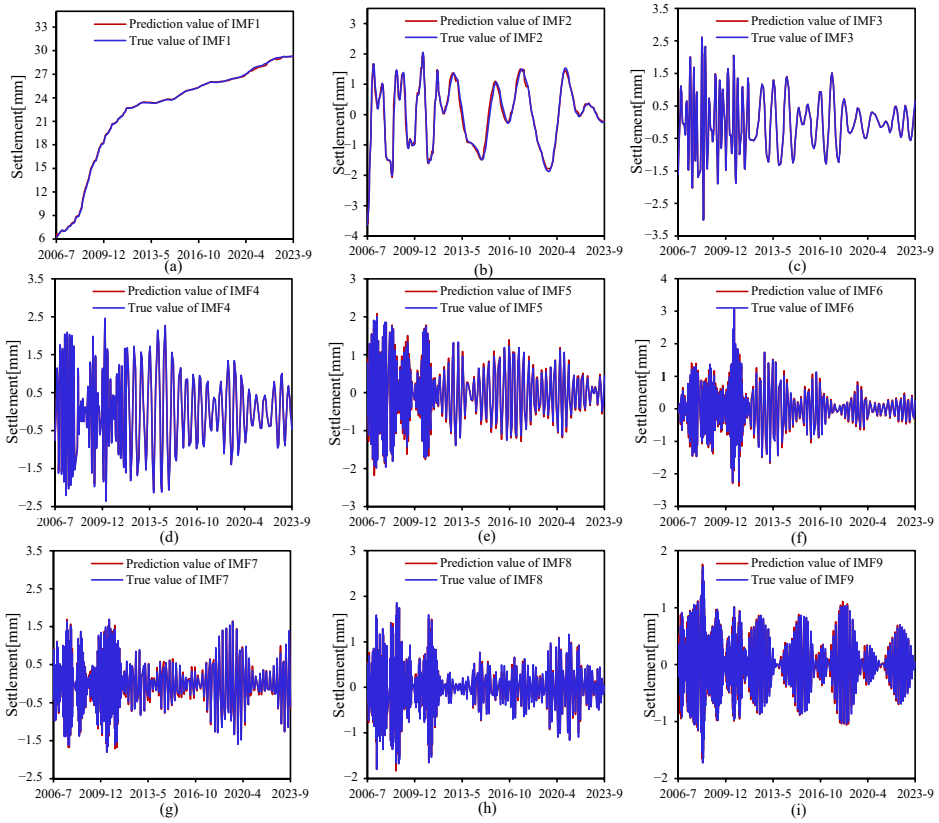



**Figure 11.** Comparisons between true monitoring values and prediction values of (**a**) decomposed subsequence IMF1, (**b**) decomposed subsequence IMF2, (**c**) decomposed subsequence IMF3, (**d**) decomposed subsequence IMF4, (**e**) decomposed subsequence IMF5, (**f**) decomposed subsequence IMF6, (**g**) decomposed subsequence IMF7, (**h**) decomposed subsequence IMF8, (**i**) decomposed subsequence IMF9.

In order to evaluate the fitting and prediction ability of the VMD-LSTM-SVR model, two evaluation indicators, the coefficient of determination $R^2$ and the root mean square error *RMSE*, are utilized. The calculation formula for the above indicators is as follows:

$$R^2 = \frac{\sum_{i=1}^{N}(\hat{\delta}_i - \overline{\delta})}{\sum_{i=1}^{N}(\delta_i - \overline{\delta})} \tag{20}$$

$$RMSE = \sqrt{\frac{1}{M}\sum_{j=1}^{M}(\delta_j - \hat{\delta}_j)^2} \tag{21}$$

where $\delta_i$ represents the training sets of the settlement monitoring data, $\overline{\delta}$ is the average value of the training datasets, $\hat{\delta}_i$ denotes the fitting results of the training datasets, $\delta_j$ indicates the test sets of the settlement monitoring data, $\hat{\delta}_j$ is the prediction results of the test datasets, $M$ and $N$ represent the numbers of data of training and test datasets, respectively.

At the same time, we introduced three other models: the stepwise regression (SR) model, LSTM optimized by the HHO (HHO-LSTM) model, and SVR optimized by the HHO (HHO-SVR) model to establish a prediction model for dam monitoring settlement. The comparison results of $R^2$ and *RMSE* of all the selected monitoring points of different prediction models are exhibited in Table 2 and Figure 12, respectively. It should be noted that the results of $R^2$ of the SR model, the HHO-LSTM model, the HHO-SVR model, and the proposed model vary from 0.801 to 0.862, 0.886 to 944, 0.873 to 0.943, and 0.980 to 0.995, respectively. Therefore, the proposed model possesses high fitting accuracy. As exhibited in Figure 12, the values of *RMSE* of the SR model, the HHO-LSTM model, the HHO-SVR model, and the proposed model vary from 3.5355 to 5.3826, 2.1386 to 3.6316, 2.4835 to 3.8471 and 1.5543 to 2.1537, respectively. It is obvious that the values of *RMSE* of the proposed model for all the selected monitoring points are the smallest.

**Table 2.** $R^2$ of different prediction models for all the selected monitoring points.

| Monitoring Point | SR Model | HHO-LSTM Model | HHO-SVR Model | The Proposed Model |
|---|---|---|---|---|
| $TR_{4\text{-}1}$ | 0.854 | 0.931 | 0.936 | 0.995 |
| $TR_{4\text{-}2}$ | 0.862 | 0.927 | 0.899 | 0.982 |
| $TR_{4\text{-}3}$ | 0.801 | 0.894 | 0.887 | 0.981 |
| $TR_{4\text{-}4}$ | 0.847 | 0.886 | 0.873 | 0.980 |
| $TR_{4\text{-}5}$ | 0.835 | 0.921 | 0.915 | 0.981 |
| $TR_{4\text{-}6}$ | 0.829 | 0.907 | 0.924 | 0.982 |
| $TR_{4\text{-}7}$ | 0.813 | 0.939 | 0.943 | 0.985 |
| $TR_{4\text{-}8}$ | 0.817 | 0.944 | 0.938 | 0.986 |

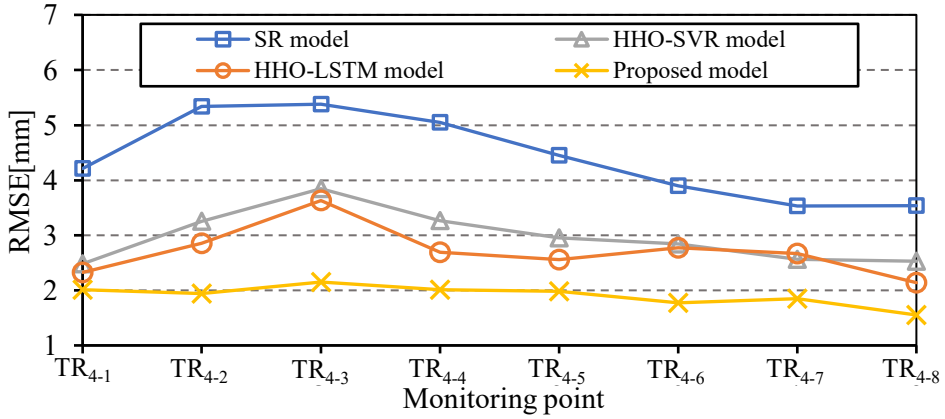



**Figure 12.** *RMSE* of different prediction models for all the selected monitoring points.

The settlement monitoring data from 4 December 2006 to 30 April 2020 was utilized to establish the prediction model. Then, the settlement data from 1 May 2020 to 6 September 2023 is applied in model testing. As shown in Figure 13, for all the settlement monitoring points, the fitting and prediction performance of the SR model is the poorest, while the proposed model demonstrates significantly better fitting and prediction performance, with results closer to the actual monitoring values. Despite significant fluctuations in the settlement monitoring data of all monitoring points, the fitting and prediction results of the proposed model closely match the actual monitoring data, demonstrating a high level of accuracy in both fitting and prediction outcomes for all the monitoring points.

Furthermore, the prediction residual was introduced as an evaluation index to assess the effectiveness of the models. Figure 14 implies the boxplots of prediction residuals for all the monitoring points of different prediction models. For all the monitoring points, the residuals of the SR model, the HHO-LSTM model, the HHO-SVR model, and the proposed model vary from −22.43 to 22.05, −7.10 to 6.07, −18.46 to 20.55, and −3.4 to 4.10, respectively.

### *5.4. Discussion*

In statistics and data analysis, the coefficient of determination $R^2$ is a measure that indicates the strength and direction of the linear relationship between a set of predictor variables and a response variable. The value of $R^2$ ranges from 0 to 1, and a larger value suggests the higher prediction accuracy of the models. Therefore, as given in Table 2, the prediction model possesses the largest values of $R^2$ for all the monitoring points, which indicates a stronger ability in prediction performance. Moreover, the root mean square error *RMSE* as another statistical indicator was introduced, and the smaller value of *RMSE* means higher prediction precision of the models. As depicted in Figure 12, *RMSE* is adopted to quantitatively assess the fitting and prediction ability of the models. The smallest value range of *RMSE* implies the proposed model is reliable.

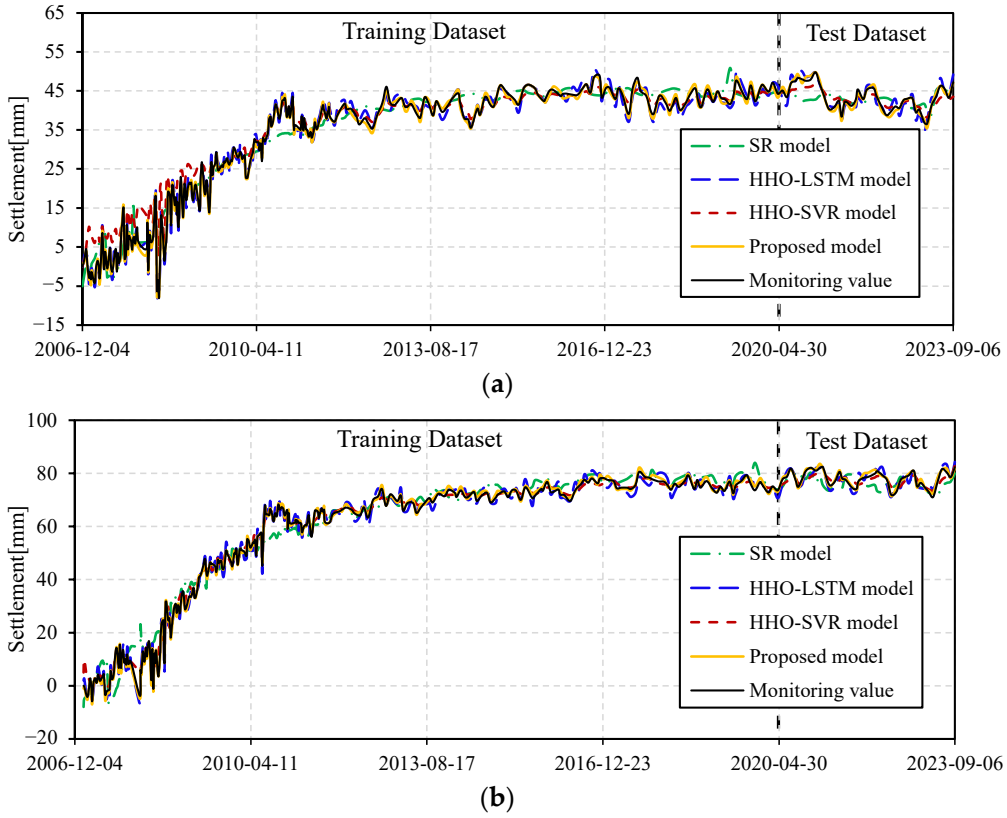



**Figure 13.** *Cont.*

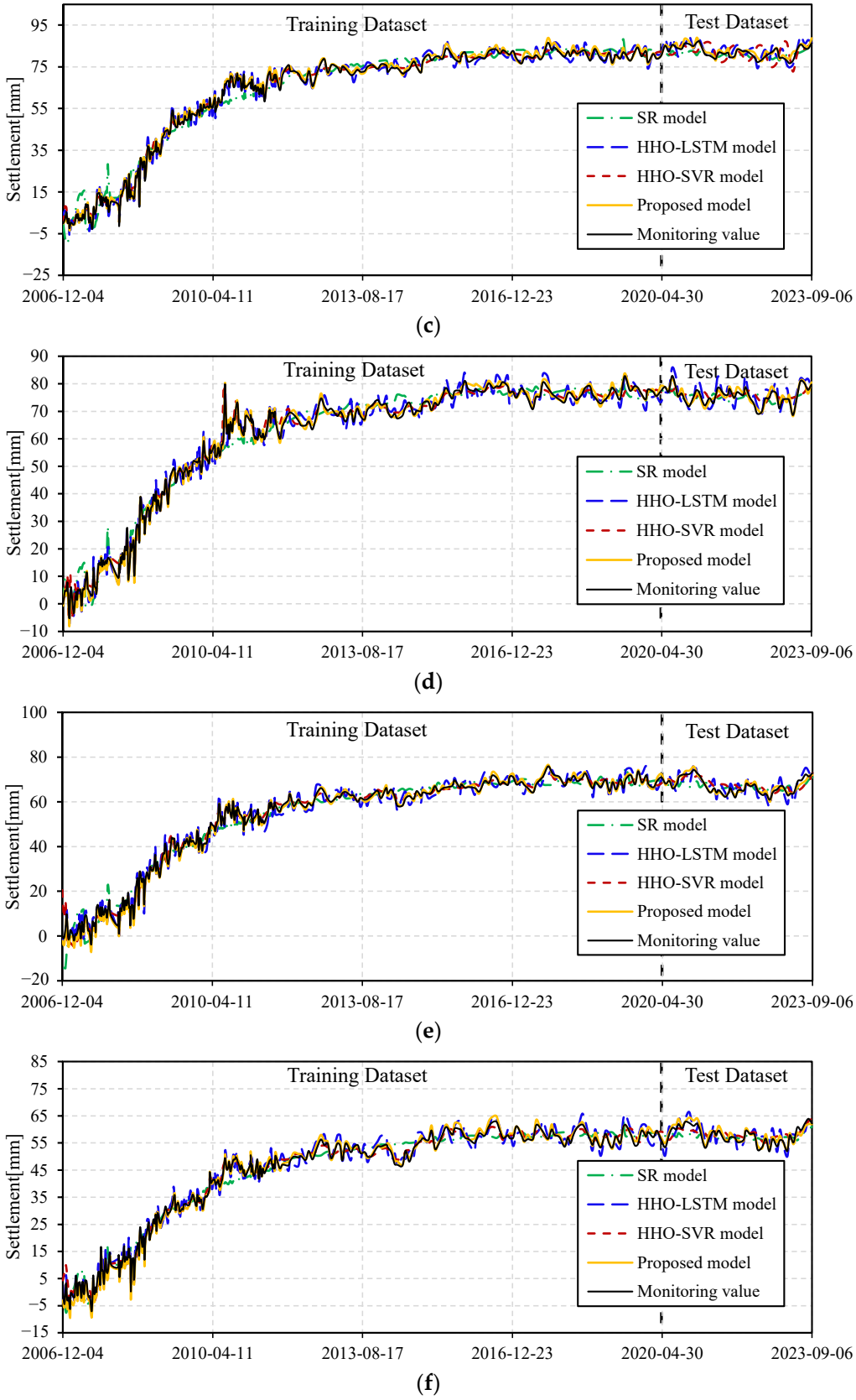



**Figure 13.** *Cont.*

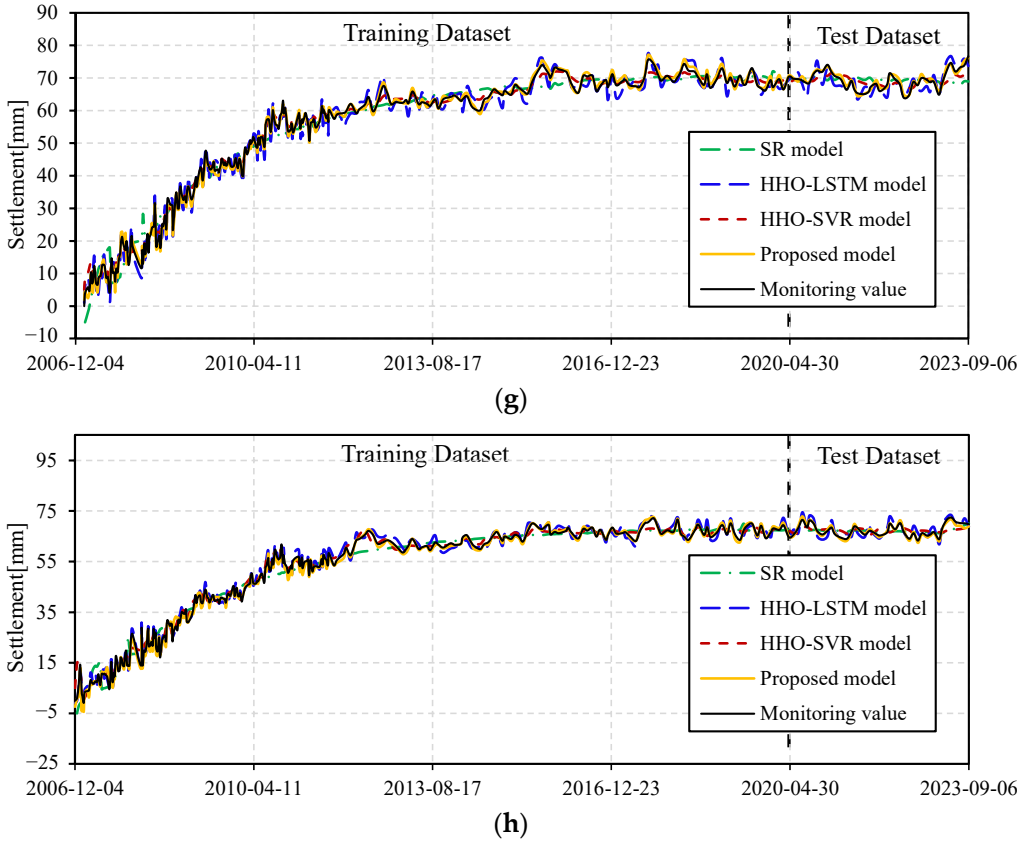



**Figure 13.** Fitting and prediction results of different prediction models for all the selected monitoring points: (**a**) $TR_{4\text{-}1}$, (**b**) $TR_{4\text{-}2}$, (**c**) $TR_{4\text{-}3}$, (**d**) $TR_{4\text{-}4}$, (**e**) $TR_{4\text{-}5}$, (**f**) $TR_{4\text{-}6}$, (**g**) $TR_{4\text{-}7}$, (**h**) $TR_{4\text{-}8}$.

The same conclusion is further confirmed in Figure 13, the prediction model exhibits excellent fitting and prediction ability for all the monitoring points. Figure 14 shows that the absolute values of the prediction residuals for the SR model are relatively large, the SR model has the poorest prediction performance. In Figure 14, it is evident that the prediction residual distributions of the proposed models for all the selected monitoring points are tighter, with overall smaller residuals, and notably fewer extreme outliers compared to other models. Consequently, for different monitoring sequences at various points, the proposed model exhibits excellent prediction performance, indicating its strong generalizability.

With the combination of utilizing the HHO algorithm, the LSTM model, and the SVR model, for the total selected 8 monitoring points, the proposed shows the strong nonlinear data mining ability. In this case study, the fitting and prediction results indicate the higher prediction accuracy and generalization ability of the proposed model, which is expected to be applied to the health diagnosis of CFRDs.

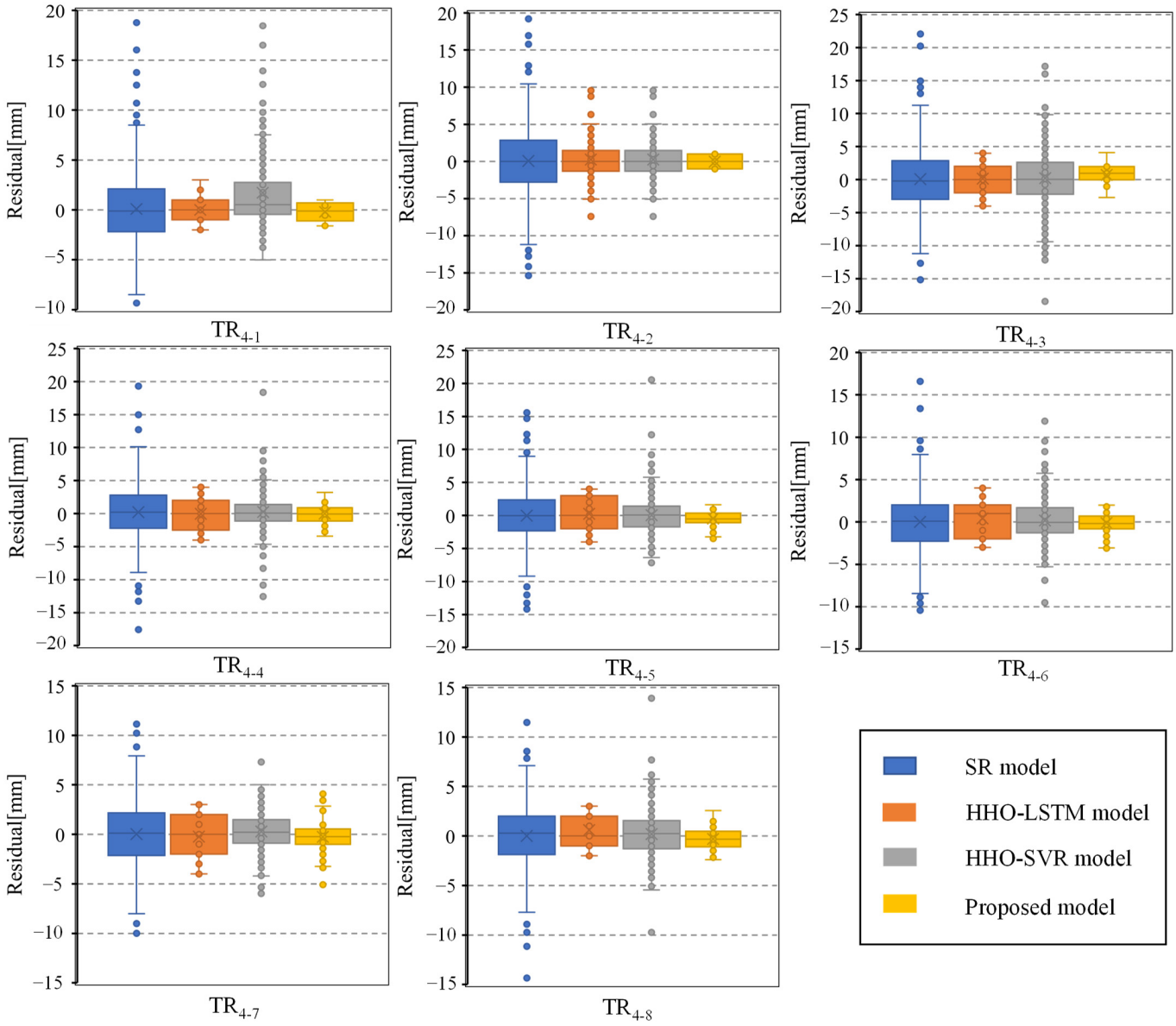



**Figure 14.** Boxplots of fitting and prediction residuals of different prediction models for all the selected monitoring points.

## 6. Conclusions

In summary, this paper has presented a useful study of a settlement prediction model for CFRDs. In this study, the VMD signal processing combined with LSTM and SVR models was introduced to establish the settlement prediction model of CFRDs. Through optimizing parameters of the above machine learning algorithms, the dam settlement monitoring data were processed via multiscale decomposition, enabling the accurate major influencing factors mining and subsequent time series prediction. In constructing the SR model, the HHO-LSTM model, and the HHO-SVR model, the fitting and prediction performance of the proposed model was evaluated. Some conclusions can be drawn as follows.

(1) On the basis of introducing the HHO algorithm, the training efficiency of the LSTM model and SVR model has been significantly enhanced. The integration of the VMD algorithm effectively decomposes the dam deformation time series, transforming the complex nonlinear sequence into multiple relatively stable sub-sequences with different frequency scales, thereby substantially improving the prediction ability of the model. Utilizing the PLS method to analyze the decomposed deformation components allows for

the major influencing factors mining, the correlation between the influencing factors and the deformation can be improved. It not only improves prediction accuracy but also reduces the input dimensionality of the prediction model. The better prediction performance is a consequence of accurate major influencing factor mining. In addition, the model achieves improved prediction accuracy by significantly reducing the input dimensions of the prediction model.

(2) The combined application of the LSTM and SVR models results in the proposed model demonstrating advantages over comparative models in terms of prediction accuracy, residual distribution, and operational efficiency. The proposed model can play an important role in decomposing the deformation monitoring data, extracting data features, mining major influencing factors, and high prediction efficiency. Consequently, the proposed model exhibits certain advantages in contrast to other comparative models in practical engineering applications for the safety monitoring of CFRDs.

(3) It should be noted that the proposed model only focuses on data from individual monitoring points, without considering the interrelationships between different monitoring points. Therefore, it seems that considering the spatiotemporal relationships among various monitoring points when establishing a prediction model may be capable of improving the prediction accuracy in future work.

**Author Contributions:** Conceptualization, X.Z. and S.Z.; methodology, X.Z.; software, T.R.; validation, F.L., Y.W. and S.Z.; formal analysis, X.Z.; investigation, T.R.; data curation, F.L.; writing—original draft, X.Z. and S.Z.; writing—review and editing, T.R., Y.W. and F.L.; supervision, X.Z.; funding acquisition, X.Z. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by the National Natural Science Foundation of China (Grant Nos. U2243223, 52209159), the Fundamental Research Funds for the Central Universities (Grant No. B230201011), the Water Conservancy Science and Technology Project of Jiangsu (Grant No. 2022024), the Anhui Provincial Natural Science Foundation "Water Sciences" Joint Fund (Grant No. 2208085US17), and the State Grid Xinyuan (Holding) Company Ltd. Science and Technology Project (Grant No. SGXYKJ-2023-059).

**Data Availability Statement:** The data presented in this study are available on request from the corresponding author.

**Conflicts of Interest:** X.Z., T.R., F.L., and Y.W. were employed by the Pumped-storage Technological & Economic Research Institute State Grid Xinyuan Group Co., Ltd. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as potential conflicts of interest.

## References

1. Shao, C.; Gu, C.; Meng, Z.; Hu, Y. Integrating the finite element method with a data-driven approach for dam displacement prediction. *Adv. Civ. Eng.* **2020**, *2020*, 4961963. [CrossRef]
2. Shao, C.; Gu, C.; Yang, M.; Xu, Y.; Su, H. A novel model of dam displacement based on panel data. *Struct. Control Health Monit.* **2018**, *25*, e2037. [CrossRef]
3. Qin, X.; Gu, C.; Shao, C.; Chen, Y.; Luis, V.; Zhao, E. Safety evaluation with observational data and numerical analysis of Langyashan reinforced concrete face rockfill dam. *Bull. Eng. Geol. Environ.* **2020**, *79*, 3497–3515. [CrossRef]
4. Zhang, G.; Zhang, J. Large-scale Apparatus for Monotonic and Cyclic Soil-Structure Interface Test. *Geotech. Test. J.* **2006**, *29*, 401–408. [CrossRef]
5. Meng, Z.; Wang, Y.; Zheng, S.; Wang, X.; Liu, D.; Zhang, J.; Shao, Y. Abnormal Monitoring Data Detection Based on Matrix Manipulation and the Cuckoo Search Algorithm. *Mathematics* **2024**, *12*, 1345. [CrossRef]
6. Liu, Y.; Zheng, D.; Cao, E.; Wu, X.; Chen, Z. Cracking risk analysis of face slabs in concrete face rockfill dams during the operation period. *Structures.* **2022**, *40*, 621–632. [CrossRef]
7. Jia, J.; Xu, Y.; Hao, J.; Zhang, L. Localizing and quantifying leakage through CFRDs. *J. Geotech. Geoenviron. Eng.* **2016**, *142*, 06016007. [CrossRef]
8. Chen, Q.; Zhang, L. Three-dimensional analysis of water infiltration into the Gouhou rockfill dam using saturated unsaturated seepage theory. *Can. Geotech. J.* **2006**, *43*, 449–461. [CrossRef]
9. Ma, H.; Chi, F. Technical progress on researches for the safety of high concrete-faced rockfill dams. *Engineering* **2016**, *2*, 332–339. [CrossRef]

10. Zhang, B.; Wang, J.; Shi, R. Time-dependent deformation in high concrete-faced rockfill dam and separation between concrete face slab and cushion layer. *Comput. Geotech.* **2004**, *31*, 559–573. [CrossRef]
11. Seo, M.; Ha, I.; Kim, Y.; Olson, S. Behavior of concrete-faced rockfill dams during initial impoundment. *J. Geotech. Geoenviron. Eng.* **2009**, *135*, 1070–1081. [CrossRef]
12. Won, M.; Kim, Y. A case study on the post-construction deformation of concrete face rockfill dams. *Can. Geotech. J.* **2008**, *45*, 845–852. [CrossRef]
13. Bodtman, W.; Wyatt, J. Closure to "Design and Performance of Shiroro Rockfill Dam" by William L. Bodtman and John D. Wyatt. *J. Geotech. Eng.* **1987**, *113*, 1164–1171. [CrossRef]
14. Wen, L.; Li, Y.; Zhao, W.; Cao, W.; Zhang, H. Predicting the deformation behaviour of concrete face rockfill dams by combining support vector machine and AdaBoost ensemble algorithm. *Comput. Geotech.* **2023**, *161*, 105611. [CrossRef]
15. Wen, L.; Li, Y.; Chai, J. Multiple nonlinear regression models for predicting deformation behavior of concrete-face rockfill dams. *Int. J. Geomech.* **2021**, *21*, 04020253. [CrossRef]
16. Wang, P.; Song, W. Analysis on the measured settlement and prediction modeling of statistic settlement of Shuibuya CFRD. *J. Hydroelectr. Eng.* **2009**, *28*, 81–86.
17. Chen, Y.; Gu, C.; Shao, C.; Qin, X. Parameter sensitivity and inversion analysis for a concrete face rockfill dam based on CS-BPNN. *Adv. Civ. Eng.* **2019**, *2019*, 1–17. [CrossRef]
18. Shao, C.; Zheng, S.; Gu, C.; Hu, Y.; Qin, X. A novel outlier detection method for monitoring data in dam engineering. *Expert Syst. Appl.* **2022**, *193*, 116476. [CrossRef]
19. Hu, Y.; Gu, C.; Meng, Z.; Shao, C. Improve the model stability of Dam's displacement prediction using a numerical-statistical combined model. *IEEE Access* **2020**, *8*, 147482–147493. [CrossRef]
20. Shao, C.; Gu, C.; Meng, Z.; Hu, Y. A Data-Driven Approach Based on Multivariate Copulas for Quantitative Risk Assessment of Concrete Dam. *J. Mar. Sci. Eng.* **2019**, *7*, 353. [CrossRef]
21. Hu, Y.; Shao, C.; Gu, C.; Meng, Z. Concrete Dam Displacement Prediction Based on an ISODATA-GMM Clustering and Random Coefficient Model. *Water* **2019**, *11*, 714. [CrossRef]
22. Behnia, D.; Ahangari, K.; Noorzad, A.; Moeinossadat, S.R. Predicting crest settlement in concrete face rockfill dams using adaptive neuro-fuzzy inference system and gene expression programming intelligent methods. *J. Zhejiang Univ.–Sci. A Appl. Phys. Eng.* **2013**, *14*, 58–71. [CrossRef]
23. Zheng, S.; Gu, C.; Shao, C.; Hu, Y.; Xu, Y.; Huang, X. A Novel Prediction Model for Seawall Deformation Based on CPSO-WNN-LSTM. *Mathematics* **2023**, *11*, 3752. [CrossRef]
24. Su, H.; Chen, J.; Wen, Z.; Wang, F. Wavelet-fractal diagnosis model and its criterion for concrete dam crack status. *Trans. Inst. Meas. Control* **2018**, *40*, 1846–1853. [CrossRef]
25. Kao, C.; Loh, C. Monitoring of long-term static deformation data of Fei-Tsui arch dam using artificial neural networkbased approaches. *Struct. Control Health Monit.* **2011**, *20*, 282–303. [CrossRef]
26. Kim, Y.; Kim, B. Prediction of relative crest settlement of concrete-faced rockfill dams analyzed using an artificial neural network model. *Comput. Geotech.* **2008**, *35*, 313–322. [CrossRef]
27. Hamzic, A.; Avdagic, Z. Multilevel prediction of missing time series dam displacements data based on artificial neural networks voting evaluation. In Proceedings of the 2016 IEEE International Conference on Systems, Man, and Cybernetics (SMC), Budapest, Hungary, 9–12 October 2016; Volume 2016, pp. 3–6.
28. Smagulova, K.; James, A. A survey on LSTM memristive neural network architectures and applications. *Eur. Phys. J. Spec. Top.* **2019**, *228*, 2313–2324. [CrossRef]
29. Hu, Y.; Gu, C.; Meng, Z.; Shao, C.; Min, Z. Prediction for the settlement of concrete face rockfill dams using optimized LSTM model via correlated monitoring data. *Water* **2022**, *14*, 2157. [CrossRef]
30. Xu, Y.; Pan, P.; Xing, C. Dam settlement prediction based on random error extraction and multi-input LSTM network. *J. Surv. Eng.* **2022**, *148*, 04022006. [CrossRef]
31. Li, Y.; Yin, Q.; Zhang, Y.; Qiu, W. Prediction of long-term maximum settlement deformation of concrete face rockfill dams using hybrid support vector regression optimized with HHO algorithm. *J. Civ. Struct. Health Monit.* **2023**, *13*, 371–386. [CrossRef]
32. Su, H.; Chen, Z.; Wen, Z. Performance improvement method of support vector machine based model monitoring dam safety. *Struct. Control Health Monit.* **2016**, *23*, 252–266. [CrossRef]
33. Kang, F.; Liu, J.; Li, J.; Li, S. Concrete dam deformation prediction model for health monitoring based on extreme learning machine. *Struct. Control Health Monit.* **2017**, *24*, e1997. [CrossRef]
34. Bui, K.; Bui, D.; Zou, J.; Doan, C.; Revhaug, I. A novel hybrid artificial intelligent approach based on neural fuzzy inference model and particle swarm optimization for horizontal displacement modeling of hydropower dam. *Neural Comput. Appl.* **2018**, *29*, 1495–1506. [CrossRef]
35. Dai, B.; Gu, H.; Zhu, Y.; Chen, S.; Rodriguez, E.F. On the use of an improved artificial fish swarm algorithm-backpropagation neural network for predicting dam deformation behavior. *Complexity* **2020**, *2020*, 5463893. [CrossRef]
36. Dai, B.; Gu, C.; Zhao, E.; Qin, X. Statistical model optimized random forest regression model for concrete dam deformation monitoring. *Struct. Control Health Monit.* **2018**, *25*, e2170. [CrossRef]
37. Ma, Z.; Zhu, W.; Han, S.; Li, Z. Concrete crack genesis mining based on fuzzy neural network. *J. Hydraul. Eng.* **2010**, *37*, 33–35. (In Chinese)

38. Gu, H.; Yang, M.; Gu, C.; Huang, X. A factor mining model with optimized random forest for concrete dam deformation monitoring. *Water Sci. Eng.* **2021**, *14*, 330–336. [CrossRef]
39. Cao, E.; Bao, T.; Gu, C.; Li, H.; Liu, Y.; Hu, S. A novel hybrid decomposition-ensemble prediction model for dam deformation. *Appl. Sci.* **2020**, *10*, 5700. [CrossRef]
40. Yuan, R.; Su, C.; Cao, E.; Hu, S.; Zhang, H. Exploration of multi-scale reconstruction framework in dam deformation prediction. *Appl. Sci.* **2021**, *11*, 7334. [CrossRef]
41. Huang, L.; Chen, S.; Li, C.; Si, Z.; Zhang, F. Application of Gaussian process regression model based on wavelet denoising in settlement prediction of CFRD. *J. Water Resour. Water Eng.* **2023**, *34*, 144–150. (In Chinese)
42. Heidari, A.; Mirjalili, S.; Faris, H.; Aljarah, I.; Mafarja, M.; Chen, H. Harris hawks optimization: Algorithm and applications. *Future Gener. Comput. Syst.* **2019**, *97*, 849–872. [CrossRef]
43. Yu, J.; Kim, C.; Rhee, S. The comparison of lately proposed Harris Hawks optimization and Jaya optimization in solving directional overcurrent relays coordination problem. *Complexity* **2020**, *2020*, 3807653. [CrossRef]
44. Huang, Z.; Gu, C.; Peng, J.; Wu, Y.; Gu, H.; Shao, C.; Zheng, S.; Zhu, M. A statistical prediction model for sluice seepage based on MHHO-BiLSTM. *Water* **2024**, *16*, 191. [CrossRef]
45. Wang, S.; Yang, B.; Chen, H.; Fang, W.; Yu, T. LSTM-Based Deformation Prediction Model of the Embankment Dam of the Danjiangkou Hydropower Station. *Water* **2022**, *14*, 2464. [CrossRef]
46. Ranković, V.; Grujović, N.; Divac, D.; Milivojević, N. Development of support vector regression identification model for prediction of dam structural behavior. *Struct Saf.* **2014**, *48*, 33–39. [CrossRef]
47. He, P.; Wu, W. Levy flight-improved grey wolf optimizer algorithm-based support vector regression model for dam deformation prediction. *Front. Earth Sci.* **2023**, *11*, 1122937. [CrossRef]
48. Man, K.; Cao, Z.; Liu, X.; Song, Z.; Liu, R. Predicting TBM tunnelling parameters with HHO-LSTM-SVR model. *Railw. Stand. Des.* **2024**, *69*, 1–12. (In Chinese)