Hindawi
Structural Control and Health Monitoring
Volume 2023, Article ID 2979822, 17 pages
https://doi.org/10.1155/2023/2979822

Research Article

An Explainable Probabilistic Model for Health Monitoring of Concrete Dam via Optimized Sparse Bayesian Learning and Sensitivity Analysis

Chaoning Lin,[1,2] **Siyu Chen**,[3] **Mohammad Amin Hariri-Ardebili**,[4,5] **and Tongchun Li**[1]

[1]*College of Water Conservancy and Hydropower Engineering, Hohai University, Nanjing, Jiangsu, China*
[2]*College of Civil and Transportation Engineering, Hohai University, Nanjing, Jiangsu, China*
[3]*Dam Safety Management Department, Nanjing Hydraulic Research Institute, Nanjing, Jiangsu, China*
[4]*Department of Civil Environmental and Architectural Engineering, University of Colorado, Boulder, CO, USA*
[5]*College of Computer, Mathematical and Natural Sciences, University of Maryland, College Park, MD, USA*

Correspondence should be addressed to Siyu Chen; siyuchen@nhri.cn and Tongchun Li; ltchhu@163.com

Received 29 March 2023; Revised 9 May 2023; Accepted 7 July 2023; Published 26 July 2023

Academic Editor: Andrea Del Grosso

Machine learning has become increasingly popular for modeling dam behavior due to its ability to capture complex relationships between input parameters and dam behavior responses. However, the use of sophisticated machine learning methods for monitoring dam behaviors and making decisions is often hindered by model uncertainty and a lack of interpretability. This paper introduces a novel model for dam health monitoring, focused on monitoring radial displacement and seepage, using optimized sparse Bayesian learning and sensitivity analysis. The model hyperparameters are optimized using an intelligent optimization method integrating the multi-population Rao algorithm and blocked cross-validation, while sensitivity analysis is employed to calculate the relative importance of input variables for a better understanding of the dam's state. The effectiveness of the proposed model is verified by using long-term monitoring data of a prototype concrete arch dam. The results confirm that the proposed model provides satisfactory performance on both the point predictions and the interval predictions for dam structural behaviors while obtaining effective explainability.

1. Introduction

Concrete dams play a crucial role in flood control, power generation, water supply, and irrigation and are thus significant assets in the social and economic domains. Over their service lifetime, these structures are exposed to diverse operational and environmental loads, including rare or extreme events such as earthquakes, prolonged droughts, and severe flooding. Moreover, the overall performance of concrete dams may gradually decline over time due to a range of factors, including hydraulic erosion, age-related deterioration, and other factors. The displacement and seepage are important indicators that can intuitively reflect the structural performance of a concrete dam. Through analysis of historical monitoring data, it is possible to develop a data-driven model that can effectively monitor the structural response of a dam and thereby provide early warning of any anomalies.

Hydrostatic-season-time (HST) and hydrostatic-temperature-time (HTT) are common statistical models for monitoring dam structural responses [1]. These models utilize measured reservoir water height, seasonal parameters (or temperature data), and time-related parameters as input, with dam displacement as the output. To monitor dam uplift or seepage behaviors, rainfall parameters are incorporated as additional model inputs [2].

The advent of artificial intelligence has spurred increasing interest in the use of machine learning techniques

for dam health monitoring [2–5]. Various ML methods have been utilized for data-driven modeling of dam structural behavior, such as feed-forward neural networks [6–9], extreme learning machines [10–12], recurrent neural network (RNN) [13–15], support vector regression (SVR) [16–19], Gaussian process regression [20–22], and decision trees-based ensemble models [23–25]. Besides, some novel data-driven methods or models have been proposed for dam health monitoring, including switching Kalman filter [26], dynamic time warping [27], panel data model [28], cloud model [29], correlated multi-target stacking [30], and spatiotemporal association mining [31]. Recently, the concept of automated machine learning (AutoML) has been also applied in dam response prediction. AutoML allows the user to search among a range of machine learning methods with optimized hyperparameters and identifies the best one [32].

Based on the literature reviewed above, it is evident that most existing methods in the field of dam health monitoring provided deterministic values (i.e., point prediction) for dam structural behaviors, without adequately accounting for the uncertainties associated with these results. Uncertainty quantification (UQ) plays a key role in monitoring and decision making during the dam operation period [33–35]. The uncertainties in dam engineering are mainly divided into two categories, i.e., aleatoric and epistemic uncertainty, where the former is referred to as the data noise, and the latter is associated with the model uncertainty (i.e., the uncertainties of model input, structure, and parameters) [36]. Probabilistic predictions are generally considered more valuable and reliable than point predictions, as they provide information about the predicted values in the form of prediction intervals. These intervals capture the uncertainty associated with the predictions and can be particularly useful for decision making and risk management under a UQ framework.

Another significant concern in the field of probabilistic prediction is model explainability. While existing probabilistic prediction methods [20, 37] often provide accurate results, they can be considered black-box machine learning models, lacking the ability to provide practitioners with information about the inner workings or features of the model. Although some data mining methods have been used in the monitoring of dam structural behaviors to determine the relative importance of each input [23, 38, 39], their applicability to probabilistic machine learning methods can be limited.

The objective of this study is to establish an explainable probabilistic prediction model that is capable of providing reliable and efficient prediction intervals of concrete dam behaviors, as well as visualization of input influencing variables, thereby offering a tool for risk-based decision making in dam health monitoring. To fulfill these objectives, an optimized sparse Bayesian learning (SBL) approach is proposed for generating high-quality prediction intervals, with the model hyperparameters determined via the integration of the Rao algorithm with blocked cross-validation. To address the issue of model explainability, a new sensitivity analysis technique is employed to display the influencing features of model inputs. In summary, the main novelties and contributions of this paper are as follows:

(i) Developing a Bayesian-based probabilistic model to make interval predictions of concrete dam structural behaviors.

(ii) Presenting an efficient optimization framework for selecting model hyperparameters adaptively.

(iii) Utilizing novel sensitivity analysis techniques to render the dam health monitoring model explainable.

(iv) Verifying the performance of the health monitoring model by comparing it with other popular ML-based models.

The remainder of the paper is summarized as follows. Section 2 provides the relevant theory and the implementation procedure of the proposed monitoring model. The case study and detailed performance evaluation are demonstrated in Section 3. Finally, the concluding remarks and future work are presented in Section 4.

2. Methodology

The sparse Bayesian learning, strategy of hyperparameter optimization, and sensitivity analysis-based data mining approach are the important components of the proposed dam health monitoring model. The basic theories of algorithms and the implementation procedure are introduced in the following subsections, and the diagram of the model framework is shown in Figure 1.

2.1. Sparse Bayesian Learning. Sparse Bayesian learning (SBL) is a nonparametric and probabilistic machine learning method. SBL shares a similar form to SVM, but its operation speed is faster and provides more flexibility for the choice of the basis function [40]. Moreover, SBL provides the probabilistic output by estimating the probabilistic density of the weights of kernel functions, which can quantify the uncertainty of the prediction results.

Given the input dataset $D_{\mathrm{SBL}} = \{(\mathbf{X}_i, y_i)|, i = 1, 2, \cdots, N\}$, the quantitative relationship between $\mathbf{X}_i$ and y_i computed by SBL is as follows:

$$y_i = \mathrm{SBL}(\mathbf{X}_i; \mathbf{w}) + \varepsilon_i, \tag{1}$$

where $\mathbf{X}$ is the matrix of the influencing variables of the dam structural behaviors, y is the measured displacement or seepage, $\mathbf{w}$ denotes the weight vector, and $\varepsilon_i \sim N(0, \sigma^2)$ is the error term. Considering the predictive uncertainty, the model output y_i follows the normal distribution:

$$p\left(y_i|\mathbf{X}_i, \mathbf{w}, \sigma^2\right) = N\left(\mathrm{SBL}(\mathbf{X}_i; \mathbf{w}), \sigma^2\right). \tag{2}$$

The output of SBL is represented by the linearly weighted sum of N kernel functions, which is shown below:

$$\widehat{y} = \sum_{i=1}^{N} w_i K(\mathbf{X}_i, \mathbf{X}) + w_0, \tag{3}$$

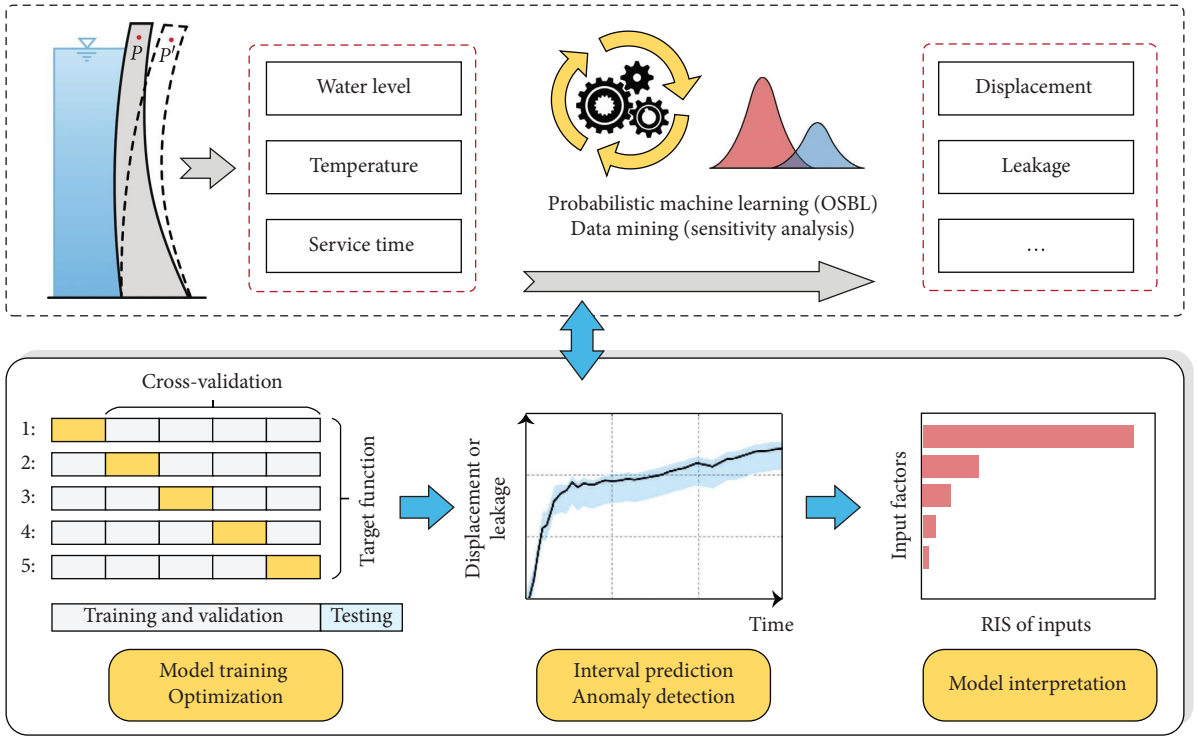


Figure 1: Framework of the proposed dam health monitoring model.

where w_i denotes the weight value and $K(\mathbf{X}_i, \mathbf{X})$ denotes the kernel function.

For the N input vectors, the likelihood of output $\mathbf{y}$ is expressed as follows:

$$p\left(\mathbf{y}|\mathbf{w}, \sigma^2\right) = \left(2\pi\sigma^2\right)^{-N/2} \exp\left\{-\frac{1}{2\sigma^2}\|\mathbf{y} - \mathbf{\Phi}\mathbf{w}\|^2\right\}, \tag{4}$$

where $\mathbf{y} = (y_1, \cdots, y_N)^{\mathrm{T}}$ represents target vector, $\mathbf{w} = (w_0, \ldots, w_N)^{\mathrm{T}}$ represents weight vector, $\mathbf{\Phi} = [\phi(\mathbf{X}_1), \ldots, \phi(\mathbf{X}_N)]^{\mathrm{T}}$ is the $N \times (N+1)$ design matrix, and $\phi(\mathbf{X}_i) = [1, K(\mathbf{X}_i, \mathbf{X}_1), \ldots, K(\mathbf{X}_i, \mathbf{X}_N)]^{\mathrm{T}}$.

It is noted that the maximum likelihood estimation of the unknown parameters $\mathbf{w}$ and σ^2 above could lead to overfitting. To overcome this problem, effective practice is imposing a set of explicit priors over these unknown parameters to constrain the variation, as shown below:

$$p(\mathbf{w}|\boldsymbol{\alpha}) = \prod_{i=0}^{N} \mathcal{N}\left(w_i|0, \alpha_i^{-1}\right), \tag{5}$$

where $\boldsymbol{\alpha}$ is a vector containing $N+1$ hyperparameters that implies how far from zero each weight is allowed to deviate.

Based on Bayesian posterior inference, the posterior distribution over $\mathbf{w}$ is obtained as follows:

$$\begin{aligned} p\left(\mathbf{w}|\mathbf{y}, \alpha, \sigma^2\right) &= \frac{p\left(\mathbf{y}|\mathbf{w}, \sigma^2\right) p(\mathbf{w}|\alpha)}{p\left(\mathbf{y}|\alpha, \sigma^2\right)} \\ &= (2\pi)^{-(1+N)/2} |\textstyle\sum|^{-1/2} \exp\left[-\frac{1}{2}(\mathbf{w} - \boldsymbol{\mu})^T \sum^{-1} (\mathbf{w} - \boldsymbol{\mu})\right], \end{aligned} \tag{6}$$

where the posterior covariance is $\sum = (\sigma^{-2}\mathbf{\Phi}^T\mathbf{\Phi} + \mathbf{A})^{-1}$ with $\mathbf{A} = \mathrm{diag}(\alpha_0, \alpha_1, \ldots, \alpha_N)$ and the mean vector is $\boldsymbol{\mu} = \sigma^{-2} \sum \mathbf{\Phi}^T \mathbf{y}$.

For the uniform hyperpriors over σ^2 and α, the learning of SBL is transformed to maximizing $p(\mathbf{y}|\alpha, \sigma^2)$ as follows:

$$p(\mathbf{y}|\boldsymbol{\alpha},\sigma^2)=\int p(\mathbf{y}|\mathbf{w},\sigma^2)p(\mathbf{w}|\boldsymbol{\alpha})d\mathbf{w}$$
$$=(2\pi)^{-N/2}\left|\sigma^2\mathbf{I}+\boldsymbol{\Phi}\mathbf{A}^{-1}\boldsymbol{\Phi}^T\right|^{-1/2}\exp\left[-\frac{1}{2}\mathbf{y}^T\left(\sigma^2\mathbf{I}+\boldsymbol{\Phi}\mathbf{A}^{-1}\boldsymbol{\Phi}^T\right)^{-1}\mathbf{y}\right]. \tag{7}$$

Values of σ^2 and α are obtained iteratively by using the following formulas:

$$(\alpha_i)^{\text{New}}=\frac{\gamma_i}{\mu_i^2}\quad(\sigma_i^2)^{\text{New}}=\frac{\|\mathbf{y}-\boldsymbol{\Phi}\boldsymbol{\mu}\|^2}{N-\sum_i\gamma_i}, \tag{8}$$

where γ_i denotes the ith posterior mean weight that can be defined by $\gamma_i\equiv 1-\sigma_i^2\Sigma_{ii}$ and Σ_{ii} represents the ith diagonal element of the posterior weight covariance.

With the convergence of the iterative process, the most probable (MP) values of $\boldsymbol{\alpha}$ and σ^2 at termination are denoted as $\boldsymbol{\alpha}_{\text{MP}}$ and σ^2_{MP}. For the trained SBL model, the predictive distribution given the new inputs $\mathbf{X}_*$ is as follows:

$$p(y_*|\mathbf{y},\boldsymbol{\alpha}_{\text{MP}},\sigma^2_{\text{MP}})=\int p(y_*|\mathbf{w},\sigma^2_{\text{MP}})p(\mathbf{w}|\mathbf{y},\boldsymbol{\alpha}_{\text{MP}},\sigma^2_{\text{MP}})d\mathbf{w}. \tag{9}$$

In equation 9, considering both terms in the integrand is Gaussian distribution, the predictive mean y_* and variance σ_*^2 are computed by

$$\widehat{y}_*=\mu^T\phi(\mathbf{X}_*),\sigma_*^2=\sigma_{MP}^2+\phi(\mathbf{X}_*)^T\sum\phi(\mathbf{X}_*), \tag{10}$$

where $\widehat{y}_*$ denotes the mean value of predictive distribution while the uncertainty of predictive distribution is quantified by variance σ_*^2. Once the significance level α is determined, the prediction interval under the confidence level $(1-\alpha)\%$ is obtained:

$$\left[\widehat{L}^{(\alpha)},\widehat{U}^{(\alpha)}\right]=[\widehat{y}_*-z_{1-\alpha/2}\sigma_*,\widehat{y}_*+z_{1-\alpha/2}\sigma_*], \tag{11}$$

where $\widehat{L}^{(\alpha)}$ and $\widehat{U}^{(\alpha)}$ denote the lower bound and upper bound, respectively. $z_{1-\alpha/2}$ is determined by the significance level α. For example, the 95% prediction interval of dam structural behavior is $[\widehat{y}_*-2\sigma_*,\widehat{y}_*+2\sigma_*]$, while the 99% prediction interval is $[\widehat{y}_*-3\sigma_*,\widehat{y}_*+3\sigma_*]$.

Various kernel functions can be adopted in SBL modeling [40], such as Gaussian, linear, polynomial, and sigmoid kernel functions. In this paper, the widely used Gaussian kernel K_G is adopted in SBL, as follows:

$$K_G(\mathbf{X}_i,\mathbf{X})=\exp\left(-\left\|\mathbf{X}_i-\mathbf{X}\right\|^2/2l^2\right), \tag{12}$$

where l denotes the kernel width, which is the hyperparameter of SBL. It is noted that a small value of the kernel width may lead to a low convergence speed and overfitting problem. Conversely, a large value of kernel width may cause underfitting [41]. Therefore, it is necessary to determine the suitable values of l to improve the model performance. The technical details of hyperparameters optimization are introduced in the next subsection.

2.2. Tuning Hyperparameters Using MP-Rao Algorithm and Cross-Validation. It can be inferred from equation (3) that the hyperparameters of SBL have influential effects on the prediction performance. To prevent overfitting and obtain sensible hyperparameters with less human intervention, the modified Rao algorithm called the multi-population Rao algorithm (MP-Rao) and blocked cross-validation [42] are integrated. MP-Rao is a parallel and metaphor-less optimization algorithm, which contains the common controlling parameters (i.e., population size and the number of iterations) and is improved based on the Rao-3 algorithm [43]. The main principle of the Rao algorithm and its variant lies in that the obtained solutions get away from the worst solution and move closer to the best one. Once the target function of optimizing model hyperparameters is given, the target values are dynamically updated based on the differences between the best candidate, the existing one, and the worst solution as follows:

$$x'_{j,k,i}=x_{j,k,i}+r_{1,j,i}\left(x_{j,\text{best},i}-\left|x_{j,\text{worst},i}\right|\right)+r_{2,j,i}\left(\left|x_{j,k,i}\text{ or }x_{j,l,i}\right|-\left(x_{j,l,i}\text{ or}x_{j,k,i}\right)\right), \tag{13}$$

where $x_{j,k,i}$ is the value of the jth variable for the kth candidate during the ith iteration, $r_{1,j,i}$ and $r_{2,j,i}$ are the random numbers within the range of $[0,1]$, $x_{j,\text{best},i}$ represents the value of the jth variable for the best candidate, and $x_{j,\text{worst},i}$ represents the value of the jth variable for the worst candidate. The term $|x_{j,k,i}$ or $|-(x_{j,l,i}$ or $x_{j,k,i})$ indicates the stochastic interaction among the population, where the current solution $x_{j,k,i}$ is compared with a randomly selected solution $x_{j,l,i}$. If the target function value of kth solution is smaller than that of lth solution, then "$x_{j,k,i}$ or $x_{j,l,i}$" becomes $x_{j,k,i}$. Otherwise, "$x_{j,k,i}$ or $x_{j,l,i}$" becomes $x_{j,l,i}$.

Compared to the original Rao algorithm, the MP-Rao algorithm splits the entire population into p subpopulations, and the solution $x_{j,k,i}$ in each subpopulation is

independently updated by equation (13) in parallel. Therefore, the algorithm is expected to reach an optimum solution.

Moreover, to balance the model fitting and predictive performance, the weighted mean squared error of prediction in cross-validation ($\mathrm{WMSE}_{\mathrm{cv,k}}$) is built as the target function:

$$\mathrm{WMSE}_{\mathrm{cv,k}} = \frac{1}{K}\sum_{k=1}^{K}\left[\alpha_1\frac{\sum_{i=1,k}^{N}\left(y_{i,v}-\widehat{y}_{i,v}\right)^2}{N}+\alpha_2\frac{\sum_{i=1,k}^{N}\left(y_{i,t}-\widehat{y}_{i,t}\right)^2}{N}\right], \tag{14}$$

where N is the number of samples, K is the number of folds in blocked cross-validation (K is determined as 5 in this research), $\widehat{y}_{i,v}$ denotes the predicted value of the inner validation samples, $y_{i,v}$ denotes the measured value of the inner validation samples, $\widehat{y}_{i,t}$ denotes the simulated value of the inner training samples, and $y_{i,t}$ denotes the measured value of the inner training samples. α_1 and α_2 are controlling weight coefficients ($\alpha_1+\alpha_2=1,\ 0<\alpha_1<1,\ 0<\alpha_2<1$). The values of α_1 and α_2 are set as 0.3 and 0.7, respectively [11].

After the optimization of hyperparameters, SBL utilizes the obtained optimal parameters to perform the prediction task on the new testing set. The procedure of the adaptive hyperparameter tuning is summarized in Algorithm 1.

2.3. Evaluation Metrics for Model Performance. Both point prediction metrics and interval prediction error metrics are adopted for evaluating the performance of the proposed probabilistic model. The formulas for the error metrics are introduced in the following subsections.

2.3.1. Point Prediction. Correlation coefficient (r), normalized root mean squared error (NRMSE), and mean absolute error (MAE) are selected as point prediction metrics, which have been defined in the literature [3, 11]. While for r, the larger value is preferable (i.e., closer to 1.0), for NRMSE and MAE, the smaller values indicate the better performance of the model. It is worth noting that NRMSE (shown in equation (15)) is a normalized indicator based on the RMSE, which can be used to compare the model prediction performance on different response variables.

$$\mathrm{NRMSE} = \frac{1}{\max(\mathbf{y})-\min(\mathbf{y})}\cdot\sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(\widehat{y}_i-y_i\right)^2}, \tag{15}$$

where y_i denotes the ith measured value and N is the sample number.

2.3.2. Interval Prediction. To assess the effectiveness of the prediction intervals obtained from probabilistic prediction, the prediction interval coverage probability (PICP), normalized mean prediction interval width (NMPIW), and coverage width-based criterion (CWC) are chosen [44].

Prediction interval coverage probability denotes the percentage of data points lying within the yielded intervals:

$$PICP = \frac{1}{N}\sum_{N=1}^{N}p_i,\ p_i=\begin{cases}1, & y_i\in\left[\widehat{L}_i^{(\alpha)},\widehat{U}_i^{(\alpha)}\right],\\ 0, & y_i\notin\left[\widehat{L}_i^{(\alpha)},\widehat{U}_i^{(\alpha)}\right],\end{cases} \tag{16}$$

where y_i denotes the measured value, N denotes the total amount of the target variables, and $\widehat{U}_i^{(\alpha)}$ and $\widehat{L}_i^{(\alpha)}$ are the upper and lower values of the ith prediction interval, respectively. Theoretically, PICP should be close to the confidence level $(1-\alpha)\%$ [44]. However, this hardly happens due to the influence of outliers in testing samples, model uncertainty, and imperfect model training (i.e., overfitting or underfitting).

NMPIW is another important metric that measures the average width of the interval, as follows:

$$\mathrm{NMPIW} = \frac{1}{N\xi}\sum_{N=1}^{N}\left[\widehat{U}_i^{(\alpha)}-\widehat{L}_i^{(\alpha)}\right], \tag{17}$$

where ξ denotes the range of the target variable. If the width of the achieved interval is too small, the corresponding PICP would be lower, and more target variables would lie out of the yielded prediction interval. Conversely, the larger interval width would make the prediction interval invalid as the abnormal variations of the dam behavior are not easily detected.

It is recognized that a reliable monitoring model should provide high-quality prediction intervals, achieving high coverage of the measured value under the scenario of a narrow-width interval. CWC is a comprehensive metric for evaluating the prediction intervals in terms of the coverage probability and interval width:

$$\mathrm{CWC} = \mathrm{NMPIW}\cdot\left[1+\eta e^{-\lambda(PICP-\mu)}\right],\ \eta=\begin{cases}0, & \mathrm{PICP}\geq\mu,\\ 1, & \mathrm{PICP}<\mu,\end{cases} \tag{18}$$

where μ denotes the nominal confidence level associated with prediction intervals. λ denotes the penalty parameter, which should be large enough, and λ is assumed as 5 in this study. η is a boolean variable. If PICP is less than the nominal confidence level, CWC would be large regardless of the width of prediction intervals. On the contrary, if PICP is equal to or greater than the nominal confidence level, then $\eta=0$, and NMPIW would be the controlling parameter. As a result, the reliability of the obtained prediction intervals increases as the CWC decreases.

Inputs: The training dataset D_{SBL} and search range of hyper-parameters l.
Outputs: The optimal hyperparameter l_{opt}
Initialize: Population size n_p, sub-population size n_{sp}.
for $j = 1, \cdots, n_p$ **do**
  for $i = 1, \cdots, n_{sp}$ **do**
    (∗) Compute the target function value f using equation (14).
    Identify the best and worst solutions, respectively.
    Update the solutions by equation (13), and compute the new target function value $f^{'}$.
    if $f^{'} < f$ **then**
      Accept the new solution, $f = f^{'}$.
    else
      Reject the new solution.
    end if
    if The termination condition is achieved **then**
      Output the search results in ith sub-population.
    else
      Return to the key step (∗).
    end if
  end for
  Find the best solutions f_i among n_{sp} sub-populations.
end for
Find the best solutions f_j from n_p populations, as well as l_{opt}.

Algorithm 1: Pseudocode of hyperparameter tuning for SBL model.

2.4. Interpretation of Dam Behaviors Using Sensitivity Analysis. The primary shared trait between SBL and other black-box ML models is their lack of transparency, meaning that the inner workings of the model are difficult to comprehend. Inspired by the pioneering work of Cortez and Embrechts [45], the sensitivity analysis (SA) method is introduced to mine the hidden information from the dam monitoring data. The SA is commonly considered a post hoc data mining algorithm [46]. Its primary principle involves computing relative importance scores for the managed inputs to clarify the inner workings of a model [47]. These scores quantify the sensitivity of each feature on the model outputs. A comparison of the scores among different groups of input variables unveils the importance granted by the model to each of such input variables when producing its outputs.

The detailed procedure of the SA is summarized in Algorithm 2, where M and N represent the number of influencing variables and the sample number, respectively, m denotes the number of subgroups (each subgroup contains at least one influencing variable, $m \le M$), $\mathbf{X}_{(j)}^{(N\times M)}$ denotes the generated meta-inputs by holding all the variables at their mean values except the jth influencing variable ($j \le m$), and $\widehat{y}_{i(j)}$ denotes the corresponding fitted results.

In sensitivity analysis, a larger RIS indicates a greater influence of the load effects on dam structural behavior. In Algorithm 2, the variance and average absolute error (AAE) are adopted to build sensitivity factors ($S_{V(j)}$ and $S_{\mathrm{AAE}(j)}$), as follows:

$$S_{V(j)} = \frac{1}{N-1}\sum_{i=1}^{N}\left(\widehat{y}_{i(j)} - \overline{y}\right)^2,$$
$$S_{\mathrm{AAE}(j)} = \frac{1}{N}\sum_{i=1}^{N}\left|\widehat{y}_{i(j)} - \widetilde{y}\right|, \tag{19}$$

where y_i denotes the measured value, $\widehat{y}_{i(j)}$ denotes the ith fitted value (posterior mean value), $\widetilde{y}$ denotes the median value of the target variable y, $\overline{y}$ denotes the mean value of the target variable y, and N is the sample number.

2.5. The Procedure of the SA-OSBL Model for Dam Health Monitoring. In this section, an explainable probabilistic model for monitoring concrete dam structural behaviors based on SA-OSBL is proposed. The detailed steps are described below, and the flowchart of the model implementation procedure is shown in Figure 2.

2.5.1. Step 1: Generating the Model Input and Output. Generate the model inputs of the dam structural behaviors using hydrostatic, temperature, aging, and rainfall-related variables based on in situ monitoring data of the dam (see equations (20) and (21)).

2.5.2. Step 2: Splitting the Data. The dataset is divided into the training set and testing set, where the former is utilized for the construction of the OSBL-based probabilistic model, and the latter is utilized for the validation of the model's effectiveness. Before model training, the samples are standardized in the range of [0, 1] using the min-max normalization approach.

2.5.3. Step 3: Optimizing and Training the Model. The kernel parameters of the model are estimated via the framework combining the MP-Rao algorithm and blocked cross-validation (5-fold). Then, the SBL model is trained using the optimized hyperparameters based on the training dataset.

Inputs: The training dataset $D_{\mathrm{SBL}} = \left\{(\mathbf{X}_i^{(N\times M)}, y_i)|i = 1, 2, \cdots, N\right\}$.
Outputs: Relative importance score RIS_j $(j = 1, 2, \cdots, m)$.
Initialize: Divide $\mathbf{X}^{(N\times M)}$ into m subgroups, $m \leq M$.
for $j = 1, \ldots, m$ **do**
  Generate the jth meta-inputs$\mathbf{X}_{(j)}^{(N\times M)}$.
  Compute the jth outputs: $\widehat{y}_{(j)} = \mathrm{SBL}[\mathbf{X}_{(j)}^{(N\times M)}]$.
  Compute the jth sensitivity factors $S_{(j)}$.
end for
for $k = 1, \ldots, m$ **do**
  Compute the relative importance score: $\mathrm{RIS}_k = S_{(k)}/\sum_{k=1}^{m} S_{(k)} \times 100\%$
end for

Algorithm 2: Pseudocode of the sensitivity analysis for model interpretation.

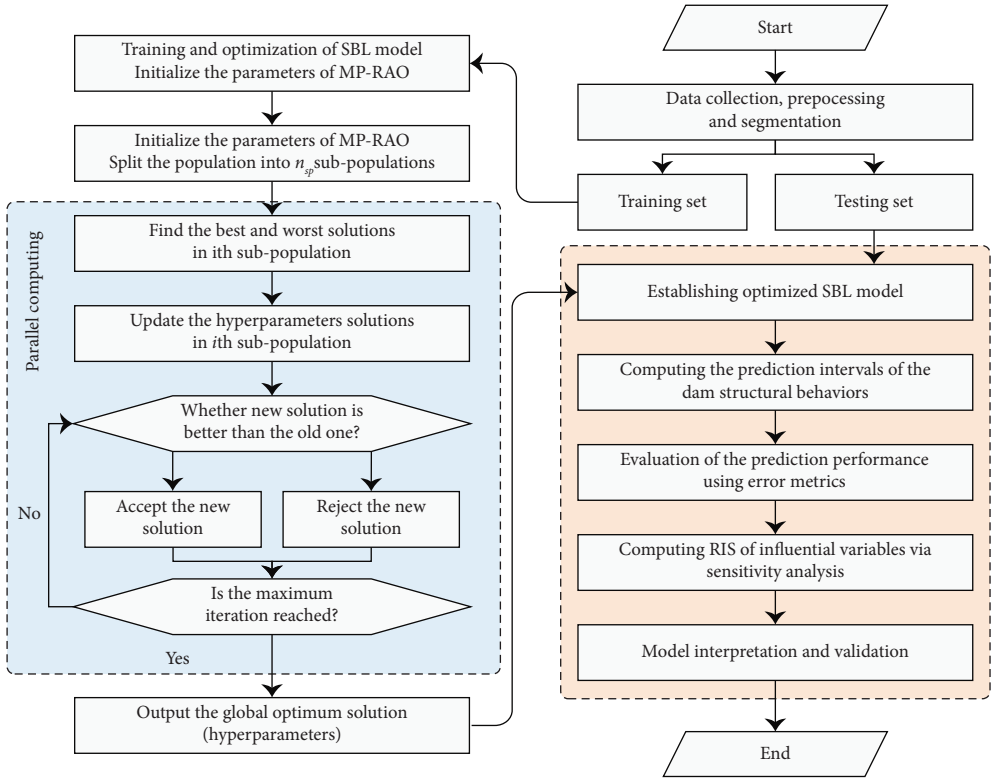


Figure 2: Flowchart of the model implementation procedure.

2.5.4. Step 4: Validating the Performance of the Simulation. The training and testing samples set are fed into the OSBL model to obtain the fitted and predicted results, respectively. Then, evaluate the simulation performance of the monitoring model comprehensively based on different error metrics and interval prediction metrics.

2.5.5. Step 5: Interpreting the Model. The relative importance scores for different groups of influencing variables are computed via the sensitivity analysis based on the historical monitoring data, to offer the interpretation of the dam's structural behaviors.

3. Case Study

In this paper, a double curvature arch dam located in the south of France is chosen as a case study. The dam is owned by EDF and was finished in 1960. The arch dam consists of 13 blocks, and the maximum dam height above the

foundation is about 45 m with the crest length being 166 m. To monitor the dam service status, the dam is equipped with a comprehensive monitoring system and instruments. Figure 3 presents the layout of the dam pendulum monitoring instruments.

3.1. Data Collection and Preparation. The radial dam displacement and seepage of the arch dam are taken as modeling targets. For the radial dam displacement, the measurements of pendulums on the central block are selected, where CB2 denotes the radial displacement between the altitudes 236 m (dam crest) and 196 m (dam toe), while CB3 denotes the radial displacement in the foundation between the altitudes 195 m and 161 m. It is noted that the positive value of the radial displacement indicates the movement of the monitoring point to the downstream direction. For the seepage, the flow rate is measured using a weir located in the gallery at the downstream dam toe. The measured seepage is the total amount of water originated from different locations including the seepage from surrounding rock, potential leakages in concrete cracks, and moisture transport in concrete. The dam behavior data are measured and collected regularly, and the time series of the long-term monitoring data are illustrated in Figure 4.

The corresponding ambient data include the water level, temperature, and daily rainfall (see Figure 5). Since the dam is located on the top of a glacial threshold, the reservoir water height is zero once the water level is lower than 196 m. The air temperature is not measured at the location of the dam. Therefore, the ambient temperature is provided by interpolation based on the measured temperature near the dam region, where the interpolation takes into account the altitude of the dam and is computed on a mesh of 1 km^2. Daily rainfall precipitation is collected from a rain gauge located about 5 km from the dam. The time series of the in situ monitoring data covers a period of thirteen years (i.e., from January 2000 to December 2012). The data from the first ten years (i.e., from January 2000 to December 2009) are used for model training and calibration, while the remaining data (from January 2010 to December 2012) are used for testing. The detailed statistical information of the measured data is summarized in Table 1.

It is widely accepted that the structural behaviors of the concrete dam are mainly influenced by hydrostatic load, temperature load, rainfall effect, and aging effect [3]. For the hydrostatic load, the reservoir water height-related parameters H^i ($i = 1 \sim 4$) or the segmented moving averages of the reservoir water height [24] are selected as factors. The calculation of temperature effects depends on the layouts of the thermometers; considering that the ambient temperature is continuous and available, the segmented moving average of ambient temperature [48] is selected to characterize the temperature effect for displacement and seepage. Likewise, the segmented moving average of daily rainfall is selected to characterize the rainfall effect of seepage [39]. The aging effect reflects the irreversible changing characteristics of dam behaviors, which can be quantitatively represented by the linear, logarithmic, exponential, or hyperbolic functions regarding the time variable t [20].

Consequently, a total of fourteen factors are selected to build the input variables for displacement modeling, which are comprised of four hydrostatic load factors, six temperature factors, and four time factors (aging factors), as follows:

$$\begin{aligned} X_\delta &= \{x_{\delta 1}, x_{\delta 2}, \cdots, x_{\delta 14}\} \\ &= \left\{H, H^2, H^3, H^4, T_0, T_{1-2}, T_{3-7}, T_{8-15}, T_{16-30}, T_{31-60}, \theta, \ln\theta, \left(1 - e^{-\theta}\right), \left(\frac{\theta}{\theta+1}\right)\right\}, \end{aligned} \tag{20}$$

where H is the reservoir water height, T_0 denotes the measured air temperature of the monitoring day, T_{p-q} denotes the moving average air temperature from p to q days before the monitoring day, t represents the cumulative day numbers from the initial day to the monitoring day, and $\theta = t/100$.

As for seepage, the lag effect of external load is considered. Therefore, six hydrostatic load factors, six temperature factors, and six rainfall factors are selected to build the input variables, as shown in equation (21). Based on the engineering judgment, the time factors are discarded as there is no significant trend change in seepage time series.

$$\begin{aligned} X_s &= \{x_{s1}, x_{s2}, \cdots, x_{s18}\} \\ &= \{H_0, H_{1-2}, H_{3-7}, H_{8-15}, H_{16-30}, H_{31-60}, T_0, T_{1-2}, T_{3-7}, T_{8-15}, T_{16-30}, T_{31-60} \\ &\quad R_0, R_{1-2}, R_{3-7}, R_{8-15}, R_{16-30}, R_{31-60}\}, \end{aligned} \tag{21}$$

where H_0 denotes the reservoir water height of the monitoring day, R_0 denotes the measured daily rainfall of the monitoring day, and H_{p-q} and R_{p-q} denote the moving average values of reservoir water height and daily rainfall from p to q days before the monitoring day, respectively. The meanings of the rest symbols are the same as those in equation (20).

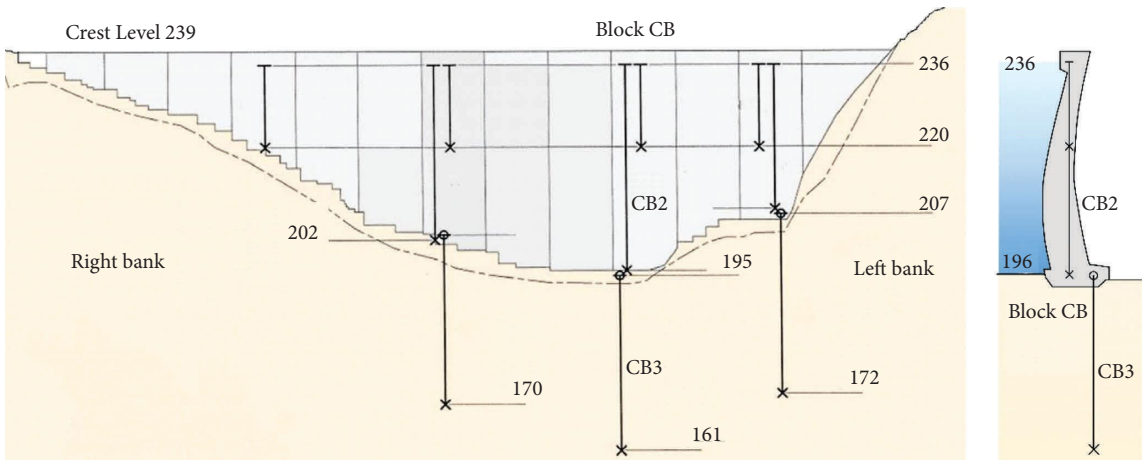


Figure 3: Downstream view and section view of the dam pendulum monitoring instruments (unit: m).

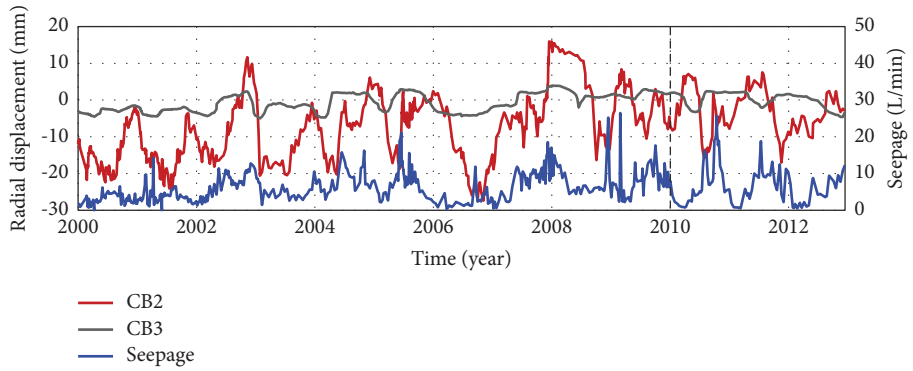


Figure 4: Time series of the long-term monitoring data: CB2, CB3, and seepage.

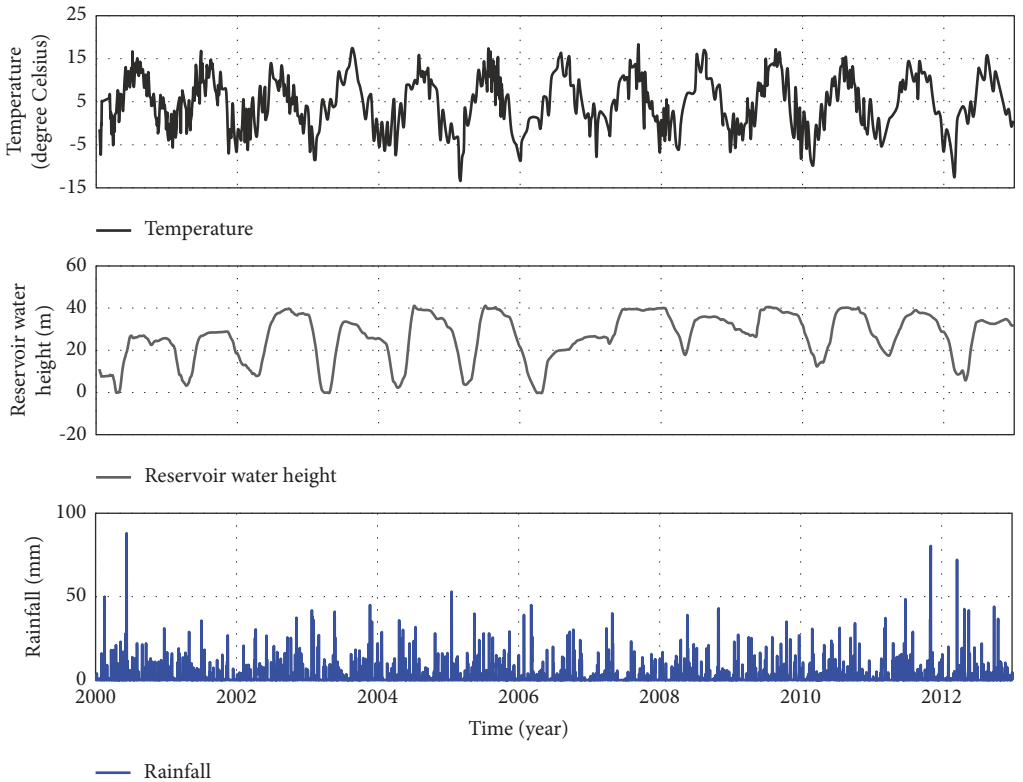


Figure 5: Time series of the temperature, reservoir water height, and daily rainfall.

Table 1: Statistical information about the dam behaviors and ambient data.

Data category (unit)	Maximum	Minimum	Mean	Monitoring frequency	Data amount
CB2 displacement (mm)	15.95	−27.48	−6.79	1.5 weeks	688
CB3 displacement (mm)	3.91	−5.00	−0.33	1.5 weeks	681
Seepage flowrate (L/min)	26.46	0.01	6.54	1.5 weeks	661
Reservoir water height (m)	41.15	0	28.02	1 day	4749
Temperature (°C)	18.35	−13.35	5.01	1 day	4749
Daily rainfall (mm)	88.00	0	2.15	1 day	4749

The provided data are automatically checked, and there is no need for data cleaning.

3.2. Hyperparameter Optimization and Model Training. The kernel width l is automatically determined using the proposed optimization framework, in which the search range of kernel width l is (0.1,5.0]. As for initial parameters in the MP-Rao optimization algorithm, the population size n_p and subpopulation size n_{sp} are set to 40 and 20, respectively. The classical particle swarm optimization (PSO) and the Rao-1 algorithm (Rao) are adopted for performance comparison, where the population size of PSO and Rao is set to 20. As for initial parameters in PSO, c_1 and c_2 are set to 1.4 and 1.6, respectively. For all the algorithms, the maximum number of iterations is 40.

Figure 6 illustrates the convergence curves of $\mathrm{WMSE}_{\mathrm{cv,k}}$ error function using different optimization algorithms. It is obvious that MP-Rao performs best in terms of convergence rate and searching capability, indicating that the proposed MP-Rao algorithm-based optimization framework for hyperparameter tuning is reliable and effective. Based on the optimization results, the hyperparameters of OSBL used for modeling CB2, CB3, and seepage are 2.46, 2.50, and 2.59, respectively.

3.3. Performance Comparison on Point Prediction. The fitting and prediction results using OSBL for CB2, CB3, and seepage are shown in Figure 7, which provides an intuitive reflection on the goodness of fit via time series plots and scatter plots. The corresponding error metrics(r, NRMSE, and MAE) are also summarized in Table 2. Based on the results, it is clear that the fitting and prediction results of radial displacement are better than those of seepage, and the prediction deviations are larger than fitting deviations.

In this research, three kernel-based regression algorithms, namely, KELM, RBFNN, and GPR, are chosen as benchmark models to compare the point prediction performance with that of OSBL. In KELM, the Gaussian kernel is selected as the kernel function, and the regularization parameter C along with kernel width are two hyperparameters to be optimized [11]. Gaussian kernel is also applied in the RBFNN model, the number of hidden neurons is the same as the number, and the spread value is the hyperparameter that needs to be optimized [48]. In GPR, different covariance functions are tested and the squared exponential covariance function is finally adopted in the model. The kernel width and noise term are the main parameters of GPR, which can be estimated by solving the maximum likelihood estimation problem [20, 21].

The long-term prediction results of the radial displacement and seepage using different models are shown in Figure 8. It can be observed that all models perform well on the displacement prediction task, and the predicted values are close to the measured values.

Table 3 summarizes the error metrics, with the best values highlighted in boldface. As seen, the OSBL obtains the smallest NRMSE and MAE values in the three tasks, outperforming the other three models in terms of prediction accuracy.

3.4. Performance Comparison on Interval Prediction. In this research, OSBL is compared with the GPR and multiple linear regression (MLR) in terms of the quality of the achieved prediction intervals. GPR is regarded as the probabilistic ML model that can provide interval prediction directly by the prediction mean and prediction variance. Specific methods for prediction interval generation using GPR can be found in [20, 49]. For the MLR model, the internal coefficients are achieved through the least square method, and the prediction interval is achieved through the confidence interval method [50], where the fitting residuals should be normally distributed.

In the application of the established interval prediction-based health monitoring model, the confidence level is set to 99% (i.e., $\alpha = 0.01$). The PICP, NMPIW, and CWC metrics of the prediction intervals are summarized in Table 4 with the best results highlighted in boldface. It is observed that the PICP values of OSBL in the case of CB2 and CB3 are almost close to the preset confidence level. Besides, the NMPIW values of the OSBL model are smaller, suggesting that the predicted intervals are narrower in each scenario. Based on the CWC results, it is concluded that OSBL is the best model. It is also found that the GPR model performs worst in the case of CB3 while the MLR model performs the worst in the cases of CB2 and seepage. Notably, the prediction interval of the MLR model in the case of seepage is invalid because the prediction intervals are too wide with the NMPIW value even larger than 1.0.

Figures 9 and 10 illustrate the obtained prediction intervals for displacement for the OSBL and GPR models, respectively. As seen, most of the displacement points are within the prediction interval except for a few abnormal values. The obtained prediction intervals for seepage are illustrated in Figure 11. Combined with results listed in Table 4, it can be seen that nearly 30% of the seepage points

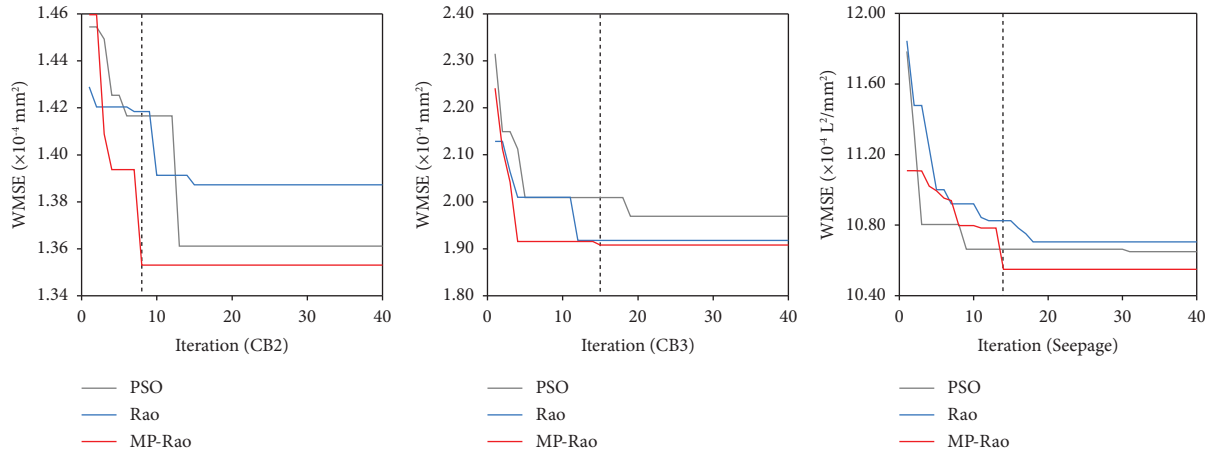
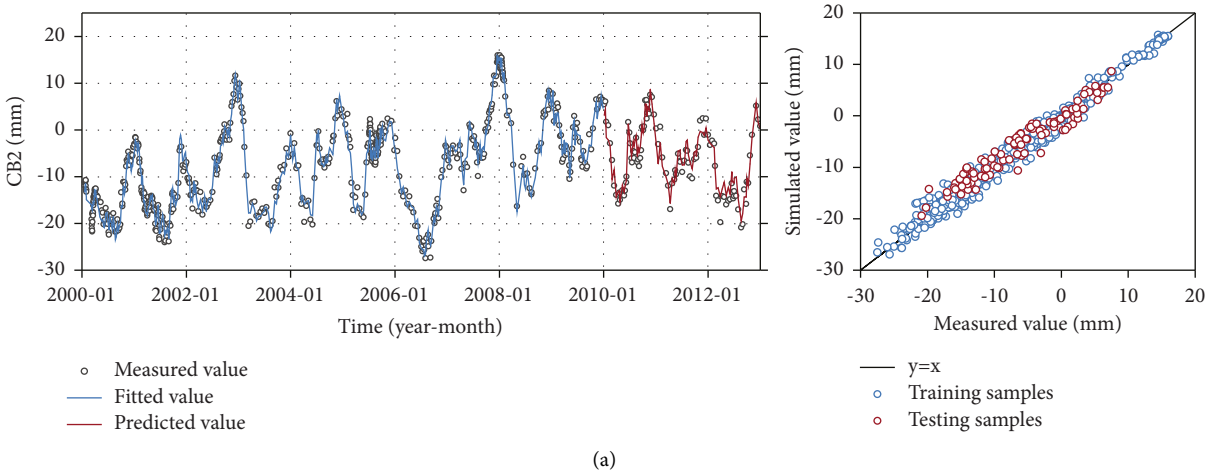


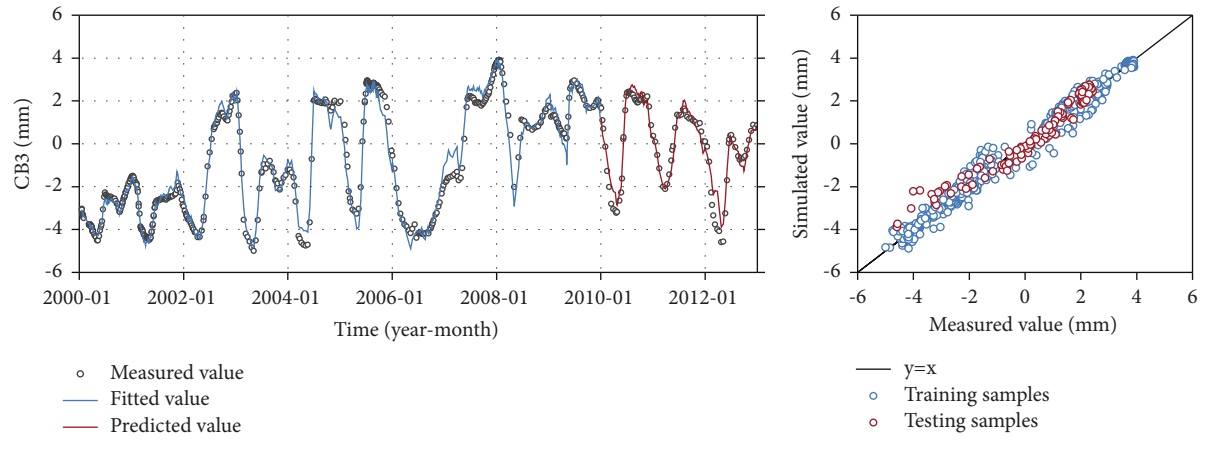
Figure 6: Comparison of convergence curves using different optimization algorithms.

Table 2: Error metrics of simulation results obtained from OSBL.

Monitoring point	Training set			Testing set		
	r	NRMSE	MAE	r	NRMSE	MAE
CB2	0.990	0.032	1.060	0.968	0.069	1.602
CB3	0.992	0.041	0.269	0.973	0.066	0.341
Seepage	0.771	0.101	1.848	0.574	0.178	3.303



(a)



(b)

Figure 7: Continued.

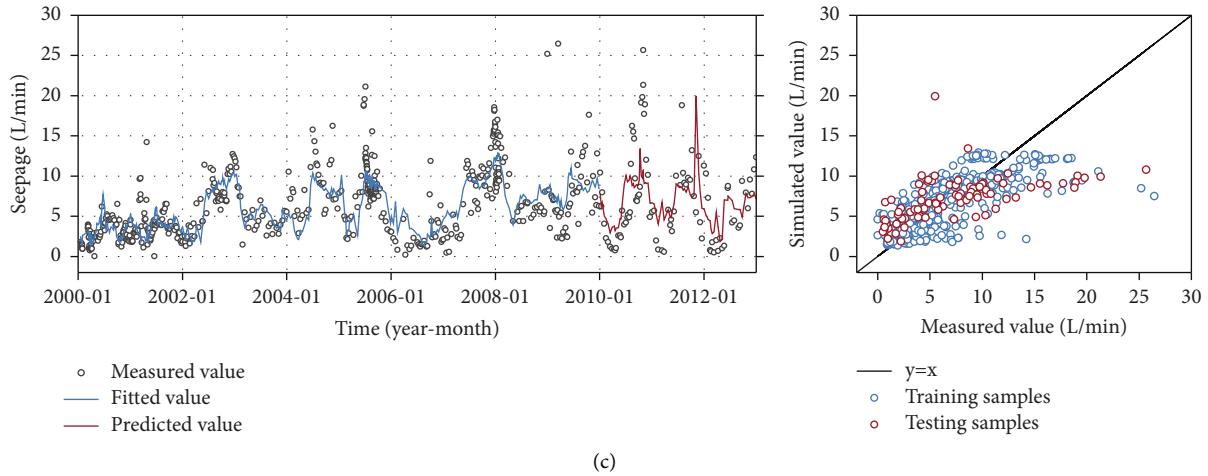


Figure 7: Fitting and prediction results using OSBL: (a) CB2, (b) CB3, and (c) seepage.

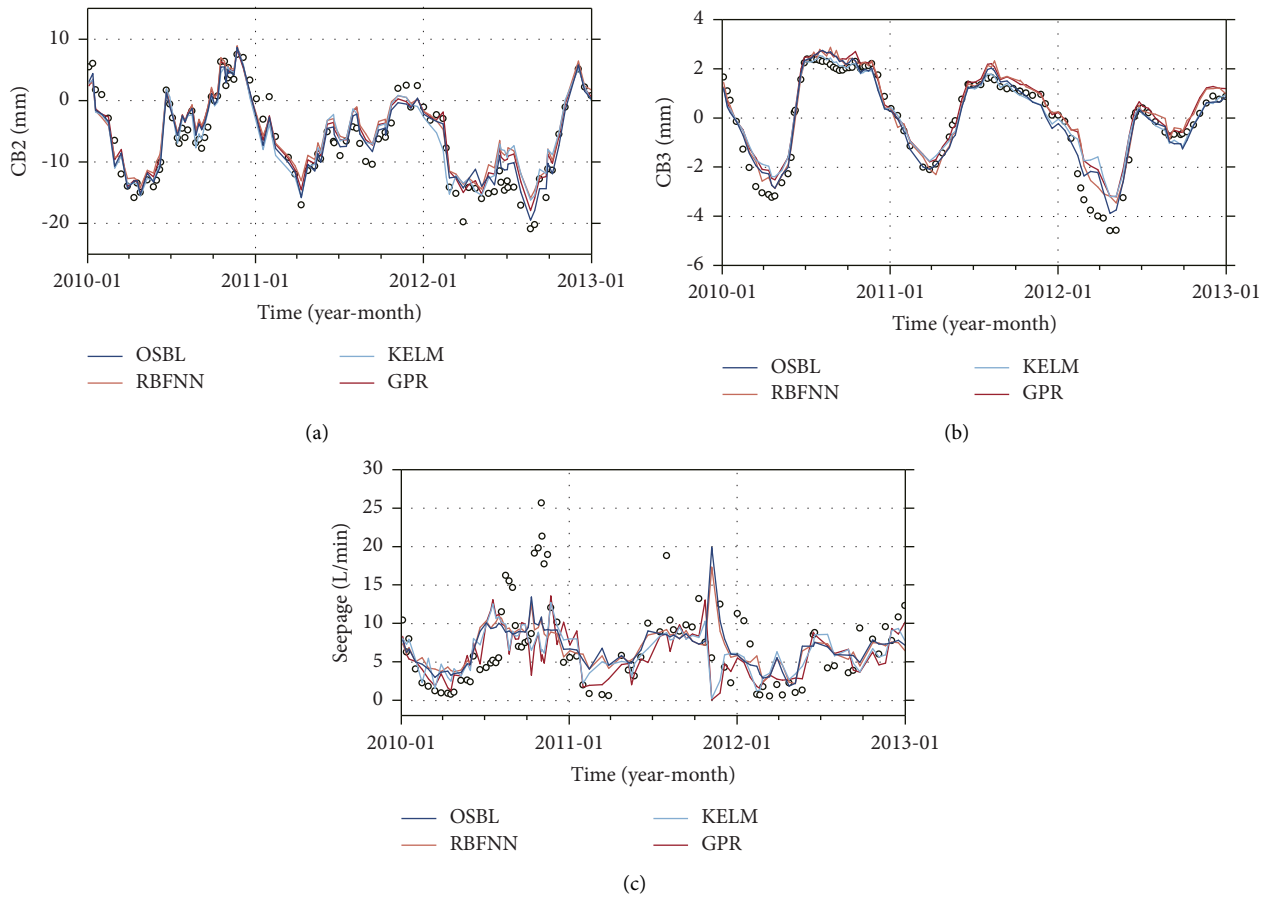


Figure 8: Comparison of point prediction performance between OSBL and other ML-based models on different monitoring points: (a) CB2, (b) CB3, and (c) seepage. The circle marks represent the measured values of dam structural behaviors.

are not within the prediction interval achieved by GPR, which is likely caused by the uncertainties of data noise and model discrepancy. Compared with GPR, the reliability of OSBL is much improved as the PICP value of OSBL is 0.122 larger than that of GPR. It is observed that the peak values (i.e., regarded as potential anomalies) are not within the prediction intervals, indicating that the models can be applied for early warning of dam seepage. In summary, the

Table 3: Error metrics of point prediction results obtained from different models.

Model	CB2			CB3			Seepage		
	r	NRMSE	MAE	r	NRMSE	MAE	r	NRMSE	MAE
OSBL	**0.968**	**0.069**	**1.602**	**0.973**	**0.066**	**0.341**	**0.574**	**0.178**	**3.303**
KELM	0.925	0.095	2.145	0.946	0.086	0.389	0.463	0.193	3.424
RBFNN	0.954	0.098	2.311	**0.973**	0.075	0.380	0.569	0.179	3.390
GPR	0.944	0.083	1.920	0.945	0.087	0.440	0.432	0.202	3.445

Note: the best values are marked in boldface.

Table 4: Quantitative evaluation of interval prediction results obtained from different models.

Model	CB2			CB3			Seepage		
	PICP	NMPIW	CWC	PICP	NMPIW	CWC	PICP	NMPIW	CWC
OSBL	0.951	**0.269**	**0.594**	0.962	**0.259**	**0.557**	0.833	**0.612**	**1.950**
GPR	**0.971**	0.347	0.729	0.894	0.329	0.861	0.711	0.656	3.303
MLR	0.932	0.341	0.797	**0.981**	0.336	0.687	**0.922**	1.809	4.348

Note: the best values are marked in boldface.

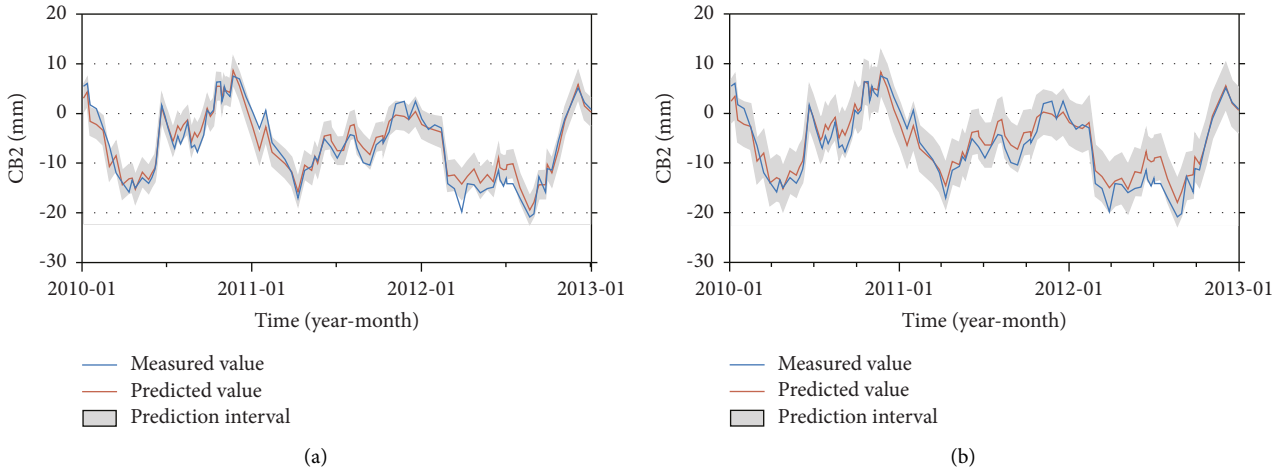


Figure 9: Prediction interval of CB2: (a) OSBL and (b) GPR.

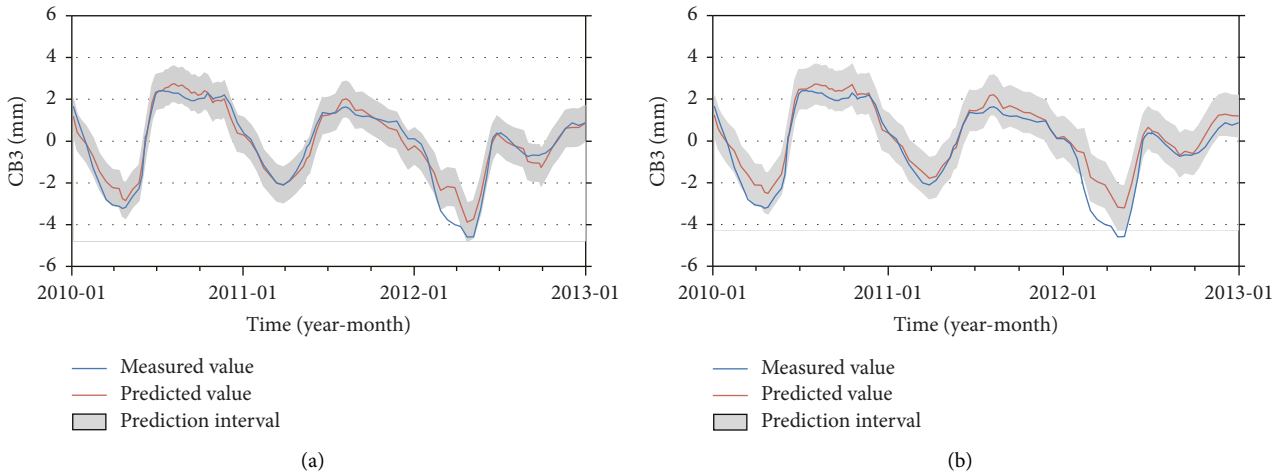


Figure 10: Prediction interval of CB3: (a) OSBL and (b) GPR.

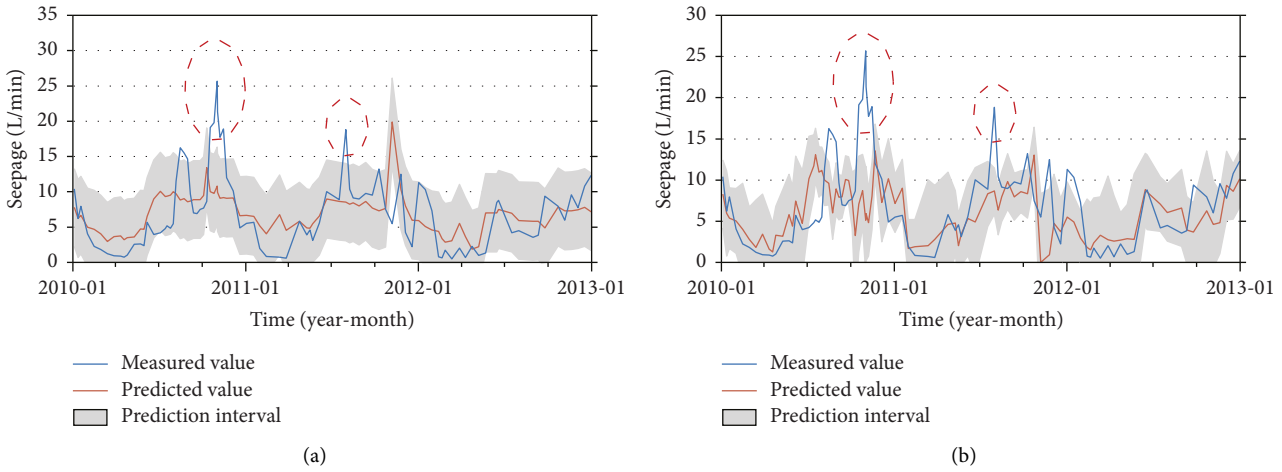


Figure 11: Prediction interval of seepage: (a) OSBL and (b) GPR. The measured value within red ellipses indicates the potential abnormal value.

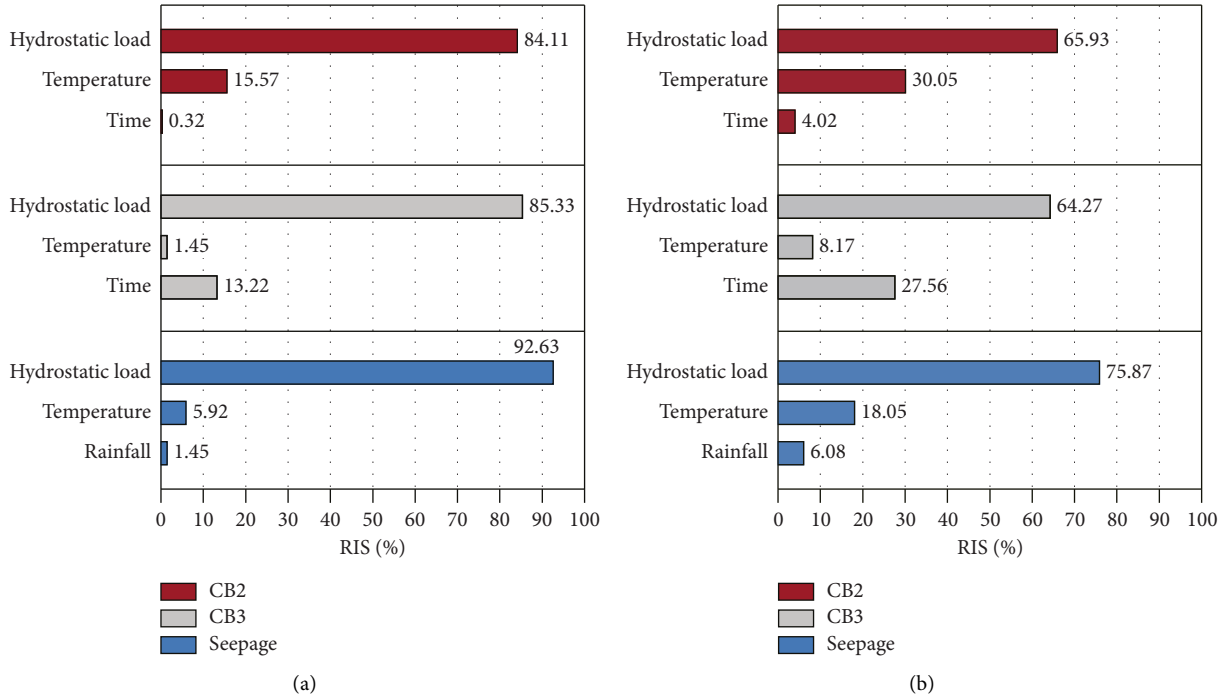


Figure 12: Relative importance score (RIS) of input variables using two sensitivity measures. (a) Variance. (b) Average absolute error.

effectiveness and reliability of the prediction intervals achieved by OSBL are confirmed, and OSBL outperforms GPR due to its smaller CWC values.

3.5. Model Interpretation. Figure 12 shows the RIS of three groups of input variables using two sensitivity measures, with the relative importance of hydrostatic load, temperature, time, and rainfall effects visualized by the bar plot. A larger RIS indicates a greater influence of the load effects on dam structural behavior. However, it is important to note that the RIS cannot quantify the absolute contribution of each group of variables to displacement or seepage.

The RIS results using two sensitivity measures are very similar, revealing that the hydrostatic load has the greatest impact on displacement variation near the dam crest (CB2), followed by the effect of temperature, while the time effect (aging effect) is insignificant. Additionally, the effect of temperature on the displacement variation at the dam toe (CB3) is notably smaller than that on CB2. It is also observed

that seepage is more sensitive to the hydrostatic load compared to the rainfall effect, suggesting that changes in reservoir water height dominate seepage variation.

From Figures 4 and 5, it is found that the radial displacements and reservoir water height have positive correlations, and there are no significant time-varying characteristics in radial displacements. According to the dam engineering mechanics knowledge, temperature-induced deformation at the dam foundation is smaller than that at the dam crest due to the smaller variations in temperature at the dam foundation compared to the dam crest. Similarly, aging-induced deformation at the dam foundation is commonly larger than that at the dam crest due to the foundation creep and restraint effects of the dam abutment. As for seepage, the reservoir is the main source of the leakage and it is reasonable that the reservoir water height is the most influencing factor of the seepage. Therefore, the interpretation results in Figure 12 are generally consistent with engineering experience and knowledge, confirming the effectiveness of the proposed sensitivity analysis.

4. Conclusions and Future Work

In this paper, we proposed an explainable probabilistic model for monitoring concrete dam displacement and seepage, which was validated using long-term monitoring data from an arch dam. The model provided the satisfactory performance of simulation and anomaly detection for dam behaviors, and the conclusions are summarized as follows:

(1) The proposed OSBL model is superior to KELM, RBFNN, and GPR models in terms of goodness of fit and interval prediction results. Meanwhile, OSBL provided better interval prediction results compared to GPR and MLR models.

(2) The OSBL model effectively quantified uncertainty through dynamic prediction intervals, where the widths of the prediction intervals varied at different periods. The upper and lower limits of the intervals served as real-time monitoring indices for dam structural behaviors.

(3) A novel sensitivity analysis-based data mining approach was integrated with the proposed OSBL to enhance the model's explainability. The interpretation results, presented as relative influence scores for different groups of input variables, revealed that reservoir water height had the greatest influence on dam displacement and seepage, while the impact of rainfall was negligible.

The proposed model is proved to be a competitive tool for dam health monitoring. For future work, the following remarks can be considered:

(1) Develop the multi-output SBL model to monitor multi-point displacement or seepage of the dam simultaneously.

(2) Incorporate interval metrics into the model training to improve performance and produce more reliable prediction intervals.

(3) Combine the OSBL model with numerical methods (e.g., finite element modeling) to establish the data-mechanical hybrid-driven model for dam health monitoring.

Data Availability

The monitoring data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Authors' Contributions

Chaoning Lin was responsible for conceptualization, validation, formal analysis, investigation, data curation, original draft preparation, visualization, and funding acquisition. Siyu Chen was responsible for conceptualization, methodology, software, investigation, formal analysis, data curation, review and editing, visualization, and funding acquisition. Mohammad Amin Hariri-Ardebili was responsible for conceptualization and review and editing. Tongchun Li was responsible for review and editing, resources, and funding acquisition.

Acknowledgments

This study was supported by the National Key R&D Program of China (No. 2022YFC3005403), the Fundamental Research Fund for Nanjing Hydraulic Research Institute (No. Y723003), the Open Research Fund of Key Laboratory of Reservoir and Dam Safety Ministry of Water Resources (No. YK323007), and the Excellent Postdoctoral Program of Jiangsu Province.

References

[1] P. Leger and M. Leclerc, "Hydrostatic, temperature, time-displacement model for concrete dams," *Journal of Engineering Mechanics*, vol. 133, no. 3, pp. 267–277, 2007.

[2] F. Salazar, M. A. Toledo, E. Onate, and R. Moran, "An empirical comparison of machine learning techniques for dam behaviour modelling," *Structural Safety*, vol. 56, pp. 9–17, 2015.

[3] F. Salazar, R. Morán, M. A. 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, pp. 1–21, 2015.

[4] M. A. Hariri-Ardebili, S. Y. Chen, and G. Mahdavi, "Machine learning-aided PSDM for dams with stochastic ground motions," *Advanced Engineering Informatics*, vol. 52, Article ID 101615, 2022.

[5] M. A. Hariri-Ardebili and S. Barak, "A series of forecasting models for seismic evaluation of dams based on ground motion meta-features," *Engineering Structures*, vol. 203, 2020.

[6] J. Mata, "Interpretation of concrete dam behaviour with artificial neural network and multiple linear regression models," *Engineering Structures*, vol. 33, no. 3, pp. 903–910, 2011.

[7] S. Y. Chen, C. S. Gu, C. N. Lin, E. F. Zhao, and J. T. Song, "Safety monitoring model of a super-high concrete dam by using RBF neural network coupled with kernel principal component analysis," *Mathematical Problems in Engineering*, vol. 2018, pp. 1–13, 2018.

[8] M. de Granrut, A. Simon, and D. Dias, "Artificial neural networks for the interpretation of piezometric levels at the rock-concrete interface of arch dams," *Engineering Structures*, vol. 178, pp. 616–634, 2019.

[9] B. Dai, H. Gu, Y. T. Zhu, S. Y. Chen, and E. F. Rodriguez, "On the use of an improved artificial fish swarm algorithm-backpropagation neural network for predicting dam deformation behavior," *Complexity*, vol. 2020, Article ID 5463893, 13 pages, 2020.

[10] F. Kang, J. Liu, J. J. Li, and S. J. Li, "Concrete dam deformation prediction model for health monitoring based on extreme learning machine," *Structural Control and Health Monitoring*, vol. 24, no. 10, Article ID 1997, 2017.

[11] S. Y. Chen, C. S. Gu, C. N. Lin, Y. Wang, and M. A. Hariri-Ardebili, "Prediction, monitoring, and interpretation of dam leakage flow via adaptative kernel extreme learning machine," *Measurement*, vol. 166, 2020.

[12] L. Lei, Y. Zhou, H. Huang, and Q. Luo, "Extreme learning machine using improved gradient-based optimizer for dam seepage prediction," *Arabian Journal for Science and Engineering*, vol. 2022, 2022.

[13] W. J. Liu, J. W. Pan, Y. S. Ren, Z. G. Wu, and J. T. Wang, "Coupling prediction model for long-term displacements of arch dams based on long short-term memory network," *Structural Control and Health Monitoring*, vol. 27, no. 7, Article ID 2548, 2020.

[14] Q. B. Ren, M. C. Li, H. Li, and Y. Shen, "A novel deep learning prediction model for concrete dam displacements using interpretable mixed attention mechanism," *Advanced Engineering Informatics*, vol. 50, Article ID 101407, 2021.

[15] J. Wang, Y. S. Zou, P. Lei, R. S. Sherratt, and L. Wang, "Research on recurrent neural network based crack opening prediction of concrete dam," *Journal of Internet Technology*, vol. 21, pp. 1161–1169, 2020.

[16] V. Rankovic, N. Grujovic, D. Divac, and N. Milivojevic, "Development of support vector regression identification model for prediction of dam structural behaviour," *Structural Safety*, vol. 48, pp. 33–39, 2014.

[17] H. Z. Su, Z. X. Chen, and Z. P. Wen, "Performance improvement method of support vector machine-based model monitoring dam safety," *Structural Control and Health Monitoring*, vol. 23, no. 2, pp. 252–266, 2016.

[18] Z. P. Wen, Z. D. Fan, and H. Z. Su, "An APPSO–SVM approach building the monitoring model of dam safety," *Soft Computing*, vol. 26, no. 21, pp. 11451–11459, 2022.

[19] S. W. Wang, C. Xu, Y. Liu, and B. B. Wu, "A spatial association-coupled double objective support vector machine prediction model for diagnosing the deformation behaviour of high arch dams," *Structural Health Monitoring*, vol. 21, no. 3, pp. 945–964, 2022.

[20] C. N. Lin, T. C. Li, S. Y. Chen, X. Q. Liu, C. Lin, and S. L. Liang, "Gaussian process regression-based forecasting model of dam deformation," *Neural Computing & Applications*, vol. 31, no. 12, pp. 8503–8518, 2019.

[21] F. Kang and J. J. Li, "Displacement model for concrete dam safety monitoring via Gaussian process regression considering extreme air temperature," *Journal of Structural Engineering*, vol. 146, no. 1, 2020.

[22] C. N. Lin, T. C. Li, S. Y. Chen, L. Yuan, P. H. A. J. M. van Gelder, and N. Yorke-Smith, "Long-term viscoelastic deformation monitoring of a concrete dam: a multi-output surrogate model approach for parameter identification," *Engineering Structures*, vol. 266, Article ID 114553, 2022.

[23] A. Belmokre, M. K. Mihoubi, and D. Santillan, "Analysis of dam behavior by statistical models: application of the random forest approach," *KSCE Journal of Civil Engineering*, vol. 23, no. 11, pp. 4800–4811, 2019.

[24] F. Salazar, M. A. Toledo, E. Onate, and B. Suarez, "Interpretation of dam deformation and leakage with boosted regression trees," *Engineering Structures*, vol. 119, pp. 230–251, 2016.

[25] H. J. Shi, W. Y. Xu, L. L. Yang, J. R. Xu, and Q. X. Meng, "Investigation of influencing factors for valley deformation of high arch dam using machine learning," *European Journal of Environmental and Civil Engineering*, vol. 27, no. 6, pp. 2399–2410, 2020.

[26] L. H. Nguyen and J. A. Goulet, "Anomaly detection with the Switching Kalman Filter for structural health monitoring," *Structural Control and Health Monitoring*, vol. 25, no. 4, Article ID 2136, 2018.

[27] J. Hu, F. H. Ma, and S. H. Wu, "Anomaly identification of foundation uplift pressures of gravity dams based on DTW and LOF," *Structural Control and Health Monitoring*, vol. 25, no. 5, Article ID 2153, 2018.

[28] C. F. Shao, C. S. Gu, M. Yang, Y. X. Xu, and H. Z. Su, "A novel model of dam displacement based on panel data," *Structural Control and Health Monitoring*, vol. 25, no. 1, Article ID 2037, 2018.

[29] H. M. Li, Z. Y. Li, F. H. Ma, and C. D. Liu, "Similarity analysis of dam behavior characterized by multi-monitoring points based on Cloud model," *International Journal of Distributed Sensor Networks*, vol. 16, no. 5, 2020.

[30] S. Y. Chen, C. S. Gu, C. N. Lin, and M. A. Hariri-Ardebili, "Prediction of arch dam deformation via correlated multi-target stacking," *Applied Mathematical Modelling*, vol. 91, pp. 1175–1193, 2021.

[31] B. Chen, Z. S. Huang, C. D. Liu, and Z. R. Wu, "Spatio-temporal data mining method for joint cracks in concrete dam based on association rules," *Structural Control and Health Monitoring*, vol. 29, no. 1, 2022.

[32] M. A. Hariri-Ardebili and F. Pourkamali-Anaraki, "An automated machine learning engine with inverse analysis for seismic design of dams," *Water*, vol. 14, no. 23, 2022.

[33] M. A. Hariri-Ardebili and F. Salazar, "Engaging soft computing in material and modeling uncertainty quantification of dam engineering problems," *Soft Computing*, vol. 24, 2019.

[34] W. Ge, X. W. Wang, Z. K. Li et al., "Interval analysis of the loss of life caused by dam failure," *Journal of Water Resources Planning and Management*, vol. 147, no. 1, 2021.

[35] Q. B. Ren, M. C. Li, R. Kong, Y. Shen, and S. L. Du, "A hybrid approach for interval prediction of concrete dam displacements under uncertain conditions," *Engineering with Computers*, vol. 39, 2021.

[36] M. A. Hariri-Ardebili, "Risk, Reliability, Resilience (R3) and beyond in dam engineering: a state-of-the-art review," *International Journal of Disaster Risk Reduction*, vol. 31, pp. 806–831, 2018.

[37] G. Prakash and G. P. Balomenos, "A Bayesian approach to model selection and averaging of hydrostatic-season-temperature-time model," *Structures*, vol. 33, pp. 4359–4370, 2021.

[38] Q. B. Ren, M. C. Li, and Y. Shen, "A new interval prediction method for displacement behavior of concrete dams based on gradient boosted quantile regression," *Structural Control and Health Monitoring*, vol. 29, pp. 1–24, 2021.

[39] K. Zhang, C. S. Gu, Y. T. Zhu et al., "A novel seepage behavior prediction and lag process identification method for concrete dams using HGWO-XGBoost model," *IEEE Access*, vol. 9, pp. 23311–23325, 2021.

[40] M. E. Tipping, "Sparse Bayesian learning and the relevance vector machine," *Journal of Machine Learning Research*, vol. 1, pp. 211–244, 2001.

[41] T. Z. Li, Q. Pan, and D. Dias, "Active learning relevant vector machine for reliability analysis," *Applied Mathematical Modelling*, vol. 89, pp. 381–399, 2021.

[42] P. Alocen, M. A. Fernandez-Centeno, and M. A. Toledo, "Prediction of concrete dam deformation through the combination of machine learning models," *Water*, vol. 14, no. 7, p. 1133, 2022.

[43] R. V. Rao, "Rao algorithms: three metaphor-less simple algorithms for solving optimization problems," *International Journal of Industrial Engineering Computations*, vol. 11, pp. 107–130, 2020.

[44] A. Khosravi, S. Nahavandi, and D. Creighton, "Construction of optimal prediction intervals for load forecasting problems," *IEEE Transactions on Power Systems*, vol. 25, no. 3, pp. 1496–1503, 2010.

[45] P. Cortez and M. J. Embrechts, "Using sensitivity analysis and visualization techniques to open black box data mining models," *Information Sciences*, vol. 225, pp. 1–17, 2013.

[46] A. Barredo Arrieta, N. Díaz-Rodríguez, J. Del Ser et al., "Explainable Artificial Intelligence (XAI): concepts, taxonomies, opportunities and challenges toward responsible AI," *Information Fusion*, vol. 58, pp. 82–115, 2020.

[47] M. A. Hariri-Ardebili, G. Mahdavi, A. Abdollahi, and A. Amini, "An RF-PCE hybrid surrogate model for sensitivity analysis of dams," *Water*, vol. 13, no. 3, p. 302, 2021.

[48] F. Kang, J. J. Li, S. Z. Zhao, and Y. J. Wang, "Structural health monitoring of concrete dams using long-term air temperature for thermal effect simulation," *Engineering Structures*, vol. 180, pp. 642–653, 2019.

[49] D. D. Kong, Y. J. Chen, and N. Li, "Gaussian process regression for tool wear prediction," *Mechanical Systems and Signal Processing*, vol. 104, pp. 556–574, 2018.

[50] B. Li, J. Yang, and D. X. Hu, "Dam monitoring data analysis methods: a literature review," *Structural Control and Health Monitoring*, vol. 27, no. 3, 2019.