MDPI

A Multi-Point Correlation Model to Predict and Impute Earth-Rock Dam Displacement Data for Deformation Monitoring

Lilang Pi [1,2], **Chunfang Yue** [1,2,*] **and Jiachen Shi** [1,2]

1   College of Hydraulic and Civil Engineering, Xinjiang Agricultural University, Urumqi 830052, China;
     xsyxpll@163.com (L.P.); jiachen1232023@163.com (J.S.)
2   Xinjiang Key Laboratory of Hydraulic Engineering Security and Water Disasters Prevention,
     Urumqi 830052, China
*   Correspondence: yuecf12@xjau.edu.cn

Abstract: Deformation is a critical indicator of structural integrity, and monitoring deformation is essential for ensuring the long-term safety of dams. However, characterizing the spatial correlations among dam deformation sequences and the similarity between displacements at various measurement points poses significant challenges when using single-point measurement models. Considering the limitations inherent in conventional models for processing spatiotemporal data, this paper introduces a novel model for predicting and imputing multi-point displacement monitoring data from earth-rock dams. The model integrates a convolutional neural network (CNN) with a bidirectional long short-term memory neural network (BiLSTM) while also incorporating an attention mechanism (AM). The CNN captures the spatial features of the displacement data, while the BiLSTM extracts temporal features. The AM assigns varying weights to input features, thereby enhancing the predictive accuracy of the model. The proposed model was experimentally validated, demonstrating its robust capabilities in data prediction and the imputation of missing data. The model provides a new strategy for forecasting dam deformation and addressing issues related to incomplete data.

Keywords: dam deformation prediction; data filling; CNN; BiLSTM; attention mechanism

1. Introduction

Dams, as structures designed to retain water, play a crucial role in various applications including flood control, power generation, and irrigation. However, dam failure can lead to catastrophic consequences. To ensure safety and facilitate the real-time monitoring of dams, appropriate monitoring instruments are integrated during the construction phase to monitor a series of established safety parameters. Deformation monitoring is particularly significant due to its intuitive and reliable nature, leading to its widespread adoption as a primary variable in dam safety evaluations worldwide. The data obtained from dam deformation monitoring reflect the alterations in dam structure resulting from loads and environmental influences, thereby providing critical insights into the condition of the dam [1]. Using these monitoring data to develop high-precision monitoring models allows the evaluation of the current stability and structural integrity of the dam. Furthermore, predicting dam deformation is vital for estimating the dam's response to various external factors and loads. Consequently, the prediction of dam deformation is essential for conducting accurate assessments of dam safety [2,3].

Due to the limitations in instrument accuracy, a certain amount of noise is present in dam deformation monitoring data. In addition, factors such as component aging introduce noticeable randomness and non-stationarity into the data [4,5], making it difficult for traditional prediction models to meet the required prediction accuracy. Therefore, a robust and accurate dam safety monitoring and prediction model is needed. Existing dam deformation monitoring models can be divided into three categories: statistical models, deterministic models, and hybrid models [6,7]. Statistical models, which include multiple

regression analysis, stepwise regression analysis, weighted regression, orthogonal polynomial regression, and difference regression models, are simple, easy to implement, and widely used [8]. Statistical models are based on statistical mathematics and have good interpretability. However, for long monitoring sequences and data with strong nonlinearity, deformation prediction models based on statistical methods tend to have large prediction errors. By contrast, deterministic models are related to the actual structural properties of the dam's body and foundation. For deterministic models, the finite element method is used to establish a calculation model, while numerical simulation techniques are applied to compute the effect of loads such as water pressure and temperature on the dam's structure, thereby determining its deformation [9]. Although deterministic models offer high accuracy, they have some limitations in practical application. The correlation between environmental variables significantly affects the performance of a deterministic model, and the inclusion of factors with low correlation can reduce the prediction accuracy [10]. In addition, deterministic models require considerable time for grid-based training, leading to inefficiencies in their real-world use [11]. In hybrid models, the load is concentrated, and the hydraulic component is calculated using finite elements; meanwhile, the temperature and aging factors are calculated using a statistical model [12,13]. The measured values are then optimally fitted, and the statistical equation is solved through stepwise regression [14]. Regardless of the model type, the performance is easily affected by uncertainty when complex nonlinear relationships exist between the influencing factors and the amount of deformation, resulting in reduced accuracy [15].

In recent years, with the rapid advances in computer and artificial intelligence, researchers have increasingly applied intelligent prediction and machine learning models to analyze and process dam safety monitoring data. For instance, Liu et al. [16] integrated a grey model with a backpropagation neural network to predict dam monitoring data, resulting in reduced prediction uncertainty and improved prediction accuracy compared with traditional models. Similarly, Ren et al. [17] employed a support vector machine to predict dam displacement in landslides and validated the model through case studies. Liu et al. [18] introduced a long short-term memory neural network (LSTM) based on the time-series characteristics of dam deformation monitoring data and used the LSTM to construct a prediction model for long-term arch dam deformation. The model demonstrated superior predictive performance compared with traditional models, including a hydrostatic seasonal time model and multilayer perceptron model, particularly for long-sequence deformation data. The LSTM effectively mitigated the issue of gradient vanishing and was able to capture the nonlinear trends, correlations, and temporal characteristics of the data.

Consequently, numerous researchers have applied LSTM, improved LSTM, and hybrid LSTM models to the prediction of dam deformation, yielding promising results [18–20]. The development of deep learning has significantly addressed the limitations of traditional machine learning. Compared with machine learning models, deep learning models, which have multiple network layers, more effectively uncover the hidden relationships between input parameters, more effectively extract the temporal correlations of data features, and more accurately capture the fluctuation characteristics of the data, thereby improving prediction performance.

Dam safety monitoring projects often involve numerous measurement points, each providing abundant monitoring data. The displacements at different locations in the same direction exhibit certain correlations. Therefore, data from a single monitoring point are insufficient to evaluate the safety and stability of the dam. However, the most commonly used dam safety monitoring models primarily analyze data sequences from individual measurement points. These single-point models have high stability requirements for the data and tend to overlook the spatial relationships between monitoring points, resulting in the incomplete extraction of hidden features within the data. This limitation introduces bias: the overall condition of the dam cannot be reliably assessed if the data from one or several observation points show abnormal fluctuations [21]. Consequently, the development of multi-point monitoring models has become a popular direction of research [22,23].

In this paper, we propose a multi-point model for dam deformation prediction and data imputation based on the combination of a convolutional neural network (CNN), bidirectional long short-term memory neural network (BiLSTM), and attention mechanism (AM). The CNN–BiLSTM–AM model combines the strengths of CNN in extracting spatial features from data with the advantages of BiLSTM in capturing temporal features, while the AM assigns weights to the input features. The main steps in predicting dam deformation using this model are as follows: First, the displacement values of some measurement points are predicted through model fitting, a process used to train the model and select appropriate parameters. Second, the trained model is applied to estimate the co-directional displacement values of the remaining measurement points. This new model offers a comprehensive understanding of the overall displacement behavior of the dam and provides a method for addressing missing data in the monitoring system.

2. Theory and Methodology

2.1. Convolutional Neural Network (CNN)

CNNs are a type of feedforward neural network characterized by a deep structure incorporating convolutional operations. Initially proposed by LeCun et al. [24], CNNs have become foundational algorithms in deep learning. A typical CNN architecture comprises three main components: convolutional layers, pooling layers, and fully connected layers. The structural layout of a CNN is illustrated in Figure 1.
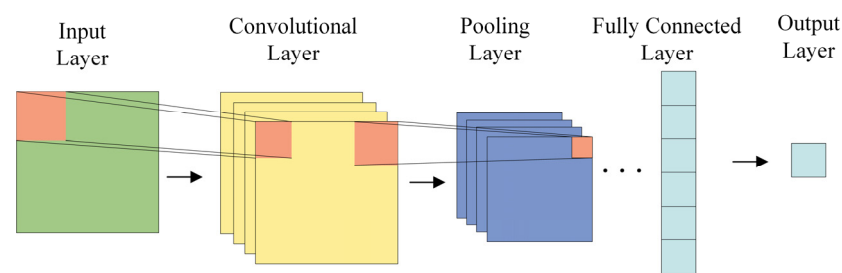

Figure 1. Structural diagram of a CNN.

In CNNs, pooling layers and convolutional layers alternate to facilitate feature extraction and dimensionality reduction. During the convolution process, CNNs adaptively capture implicit features from the data while simultaneously reducing data redundancy and complexity [25]. The extracted features are subsequently fused and passed into the fully connected layer. Non-linearity is introduced into the neuron outputs via activation functions at this stage. Each convolutional layer consists of multiple convolutional kernels that perform convolutions on the input data to extract hidden features, generating feature maps in the process. These feature maps are then processed by nonlinear activation functions to produce the output of the convolutional layer. The convolution process is mathematically described by Equations (1)–(4):

$$c_i = f(w_i * x_i + b_i) \tag{1}$$

where c_i represents the i-th feature map, w_i denotes the weight matrix, x_i is the input to the convolutional layer, $*$ denotes the dot product operation, b_i represents the bias vector, and $f(\cdot)$ refers to the activation function.

The pooling operation reduces the dimensionality of the feature map and decreases the computational complexity. Commonly used pooling methods include average pooling and max pooling. In our model, max pooling is applied to the dam deformation monitoring data to retain the most relevant information. The max pooling process is conducted according to Equations (2) and (3):

$$\gamma(c_i, c_{i-1}) = max(c_i, c_{i-1}) \tag{2}$$

$$p_i = \gamma(c_i, c_{i-1}) + \beta_i \tag{3}$$

where $\gamma(\cdot)$ indicates the maximum pooling downsampling function, β_i represents the deviation, and p_i represents the output of the maximum pooling layer.

Finally, the feature maps obtained from the convolution and pooling operations are passed to the fully connected layer, where the final output vector is computed. The fully connected layer calculates the output as follows:

$$y_i = f(t_i p_i + \delta_i) \tag{4}$$

where y_i represents the final output vector, δ_i represents the deviation, and t_i represents the weight matrix.

2.2. Bidirectional LSTM Network (BiLSTM)

In 1982, John Hopfield, a physicist at the California Institute of Technology, introduced the Hopfield network, a single-layer feedback neural network designed to solve combinatorial optimization problems. This network served as the prototype for the earliest recurrent neural networks (RNNs). A standard RNN is a neural network that includes a self-connected hidden layer. The primary advantage of this network is its ability to "remember" and use past contextual information. However, extensive studies have demonstrated that standard RNNs struggle to retain information over long periods due to limitations in memory and information storage. Bengio et al. [26] found that standard RNNs suffer from gradient vanishing and gradient explosion during the iterative training process. To address these issues, in 1997, Jürgen Schmidhuber proposed the LSTM network, which overcomes the long-term dependency problem in RNNs by introducing memory cells and gating mechanisms. Building on the LSTM, the BiLSTM was introduced by Graves and Schmidhuber [27]. This model extends the LSTM structure by unfolding it bidirectionally along the time axis, allowing the network to capture temporal information in both the forward and backward directions simultaneously. The bidirectional architecture leverages the forward and backward hidden layers to extract context from both directions. A schematic diagram of the BiLSTM structure is shown in Figure 2.
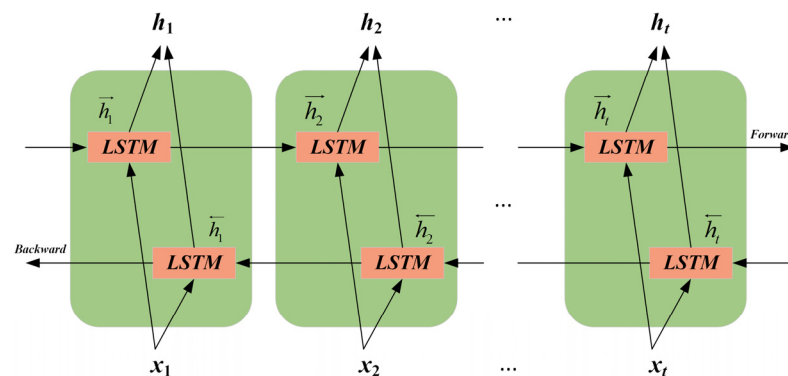

Figure 2. Diagram of the BiLSTM structure.

The forward and backward inputs and outputs of the BiLSTM model can be represented by Equations (5)–(7):

$$\overrightarrow{h_t} = \sigma(\overrightarrow{\omega} \cdot x_t + \overrightarrow{v} \cdot h_{t-1} + \overrightarrow{b}) \tag{5}$$

$$\overleftarrow{h_t} = \sigma(\overleftarrow{\omega} \cdot x_t + \overleftarrow{v} \cdot h_{t-1} + \overleftarrow{b}) \tag{6}$$

$$y_t = \sigma([\overrightarrow{h_t}, \overleftarrow{h_t}] + b) \tag{7}$$

where x_t is the input at time t; $\overrightarrow{h_t}$ represents the outputs of the forward layer; $\overleftarrow{h_t}$ represents the outputs of the backward layer; y_t represents the outputs of the hidden layer at time

t; σ is the LSTM unit function; $\overrightarrow{\omega}$, $\overrightarrow{v}$, $\overleftarrow{\omega}$ and $\overleftarrow{v}$ are weight coefficients; $\overrightarrow{b}$, $\overleftarrow{b}$ and b are bias vectors.

2.3. Attention Mechanism (AM)

Inspired by the human visual system, Treisman and Gelade first applied an AM in the field of visual image processing to simulate the attention mechanism of the human brain [28]. The core idea of the AM can be summarized as weighted averaging with dynamic weighting. In our model, the AM incorporated in the BiLSTM network uses the last cell state of the BiLSTM or aligns the implicit state of the BiLSTM with the unit state of the current input step. Subsequently, the correlation between the output states and these candidate intermediate states is calculated. During the learning process, the AM allocates weights; higher weights are assigned to the most relevant information, while irrelevant data are suppressed, thereby enhancing the accuracy and efficiency of the model's predictions. The output A of the attention layer is determined according to Equations (8)–(10):

$$M = \tanh(Y) \tag{8}$$

$$\alpha = softmax(w_a^T M) \tag{9}$$

$$A = Y\alpha^T \tag{10}$$

where Y is a matrix representing the features captured by the BiLSTM model, α denotes the attention weights of the features, w_a represents the weight coefficient of the attention layer, and T indicates the transpose operation.

3. Proposed CNN–BiLSTM–AM Hybrid Model and Evaluation Metrics

3.1. Framework of the CNN–BiLSTM–AM Hybrid Model

This section introduces the proposed CNN–BiLSTM–AM hybrid model structure and its components. To enhance feature extraction from the deformation monitoring sequences and improve the predictive capabilities, we integrated a CNN, BiLSTM, and AM into a single framework, resulting in a novel CNN–BiLSTM–AM model for predicting dam deformation. As illustrated in Figure 3, the model consists of five fundamental modules: the input module, feature extraction module, sequence learning module, attention module, and prediction module. In the feature extraction module, CNNs are employed to extract spatial features from the input data, which are then fed into the BiLSTM network for sequence learning. The sequence learning module utilizes BiLSTM to capture long-term temporal information, and the resulting outputs serve as inputs to the attention module. Within the attention module, the AM assigns varying weights based on the model's input features, thereby enhancing prediction accuracy. Finally, the prediction module comprises a fully connected layer followed by an output layer, culminating in the model's final prediction.
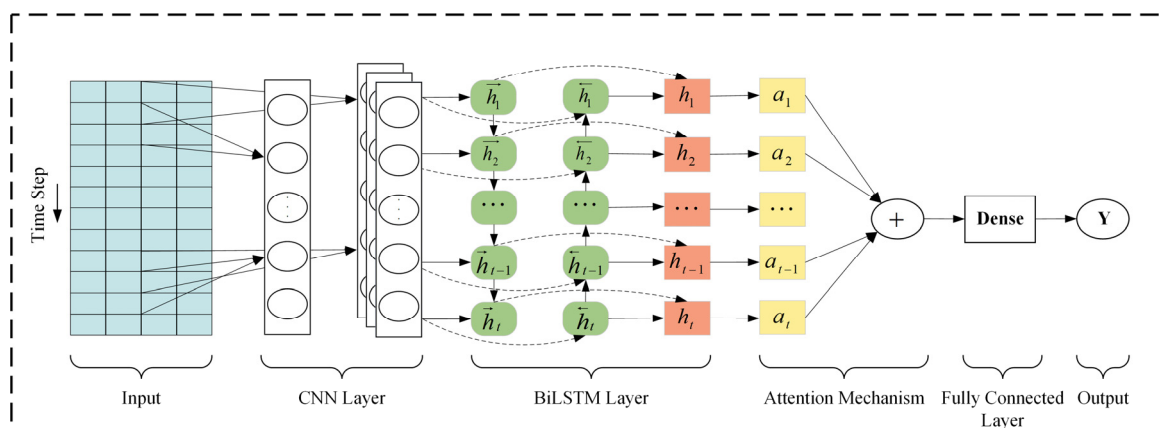

Figure 3. Framework of the CNN–BiLSTM–AM hybrid model.

3.2. Multi-Point Modeling of Dam Displacement Based on the CNN–BiLSTM–AM Model

3.2.1. Basic Idea of Multi-Point Displacement Modeling

The dam as an integral structure, the displacement values of two points in close proximity did not exhibit sudden changes (in the absence of structural cracks), indicating a certain degree of correlation between the displacements at various points in the same section. However, this correlation diminished as the distance between the points increased. As illustrated in Figure 4, points T1 to T10 represent deformation monitoring instruments. Each point (T1 to T10) in the figure indicates where the monitoring points were deployed at different positions within the dam. In the schematic diagram: T1–T4 are at the same elevation, and T5–T10 are at the same elevation. Even at the same elevation, each monitoring point is also arranged at different positions. The main reason for such an arrangement of monitoring instruments is to capture the deformation information of the dam at different heights and locations. These deformation data imply the spatial characteristic relationships between data at different spatial positions. Moreover, these deformation monitoring data provide a basis for establishing multi-point models for displacement data prediction and filling. Based on the correlation between deformation at different points, a displacement model for multiple measurement points was established. At the same elevation, the actual displacement values of several measurement points were treated as independent variables, while the displacement values of other points were considered dependent variables. A multi-measurement-point model for predicting dam deformation was constructed based on the CNN–BiLSTM–AM model to predict the unknown displacement data at the same elevation. For illustration, we consider four adjacent measurement points: T1, T2, T3, and T4. The displacement values of T3 are predicted using the data from T1, T2, and T4. A model is developed, and part of the actual observation data for T3 is used in the training process of the deformation prediction model for T1, T2, and T4. The trained model is then employed to predict the deformation of T3 for the remaining time periods, leveraging the correlation among the homologous measurement point data. Thus, the displacement values of the dam at the same elevation can be generalized from individual points to the entire surface.
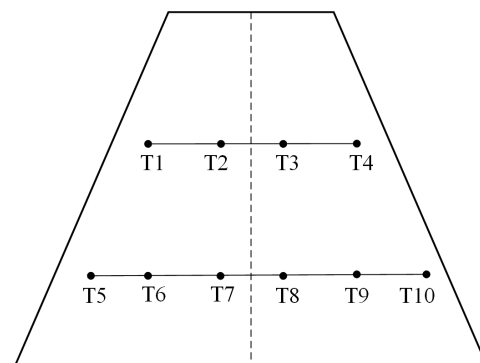

Figure 4. Schematic Diagram of the Deployment of Deformation Monitoring Instruments.

3.2.2. Technical Roadmap for the Multi-Point Displacement Model Proposed in This Paper

Based on the above-mentioned methods, a hybrid model for predicting deformation during the operation and management of earth-rock dams was developed. The flowchart of the model is shown in Figure 5. The implementation process of the model is divided into three main stages: data preparation and preprocessing; model training; and evaluation of prediction results.

The first step in data preparation and preprocessing is to collect the monitoring data from homologous deformation measurement points. The collected data are then preprocessed through the following processes, which are crucial for enhancing the accuracy of data prediction:

1. Data cleaning: The goal of data cleaning is to improve the quality of the deformation monitoring data by eliminating outliers and noise. Data cleaning is typically achieved using scatter plots.
2. Data standardization: Standardization reduces systematic error caused by data being collected at different monitoring points over different time periods. In addition, standardization helps mitigate discrepancies between datasets.
3. Normalization: Normalization plays a significant role in accelerating the training process of neural networks and preventing issues such as gradient explosion [29]. Normalization ensures that the original characteristics of the data are preserved while improving the convergence speed of computations. In our model, input variables are normalized to the [0,1] range, as shown in Equation (11):

$$x_i = \frac{x_o - x_{\min}}{x_{\max} - x_{\min}} \tag{11}$$

where x_i is the normalized value of the variable, x_o is the original value of the variable, $x_{\max}$ is the maximum value of the variable, and $x_{\min}$ is the minimum value of the variable.
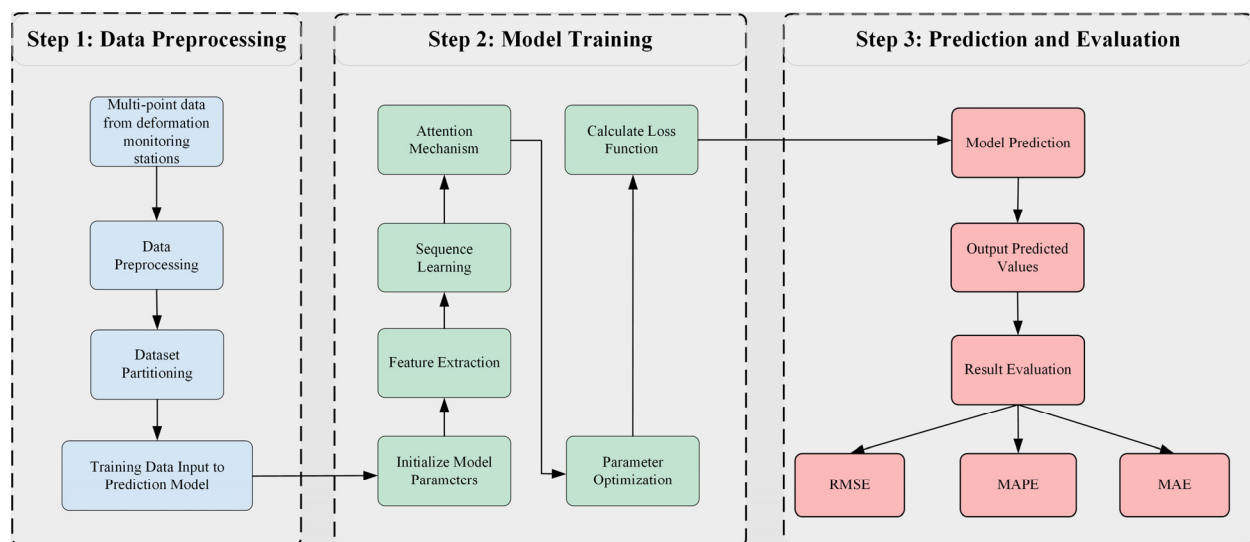

Figure 5. Flowchart of the proposed model.

The second step is model training. First, the initial parameters of the model are determined. Next, the spatial features of the input data are extracted through the feature extraction module and passed to the sequence learning layer, where temporal features are captured. The AM emphasizes features that have greater influence on the prediction results. The training process is terminated once the predefined maximum training period is reached.

3.3. Evaluation Metrics

The prediction model was evaluated based on root-mean-square error (RMSE), mean absolute percentage error (MAPE), and mean absolute error (MAE):

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (y_i - y_p)^2} \tag{12}$$

$$\text{MAPE} = \frac{1}{n} \sum_{i=1}^{n} \frac{|y_i - y_p|}{|y_i|} \times 100\% \tag{13}$$

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^{n} |y_i - y_p| \tag{14}$$

where y_i and y_p represent the observed and predicted values, respectively. Larger RMSE and MAE values indicate greater prediction errors, while lower MAPE values indicate more accurate predictions.

4. Experimental Study

4.1. Study Area

To evaluate the effectiveness of the proposed method for predicting dam deformation, the Uluwati Water Conservancy Project in Xinjiang was selected as a case study. This project is situated in the middle reaches of the Karakashi River, in the Hotan District, and serves as a control project for the river. The dam site is located 71 km from Hotan City. The total reservoir capacity is 347 million m³. The maximum dam height is 133.0 m, and the elevation, length, and width of the dam crest are 1965.80, 365.0, and 8.9 m, respectively. The project provides comprehensive benefits, including irrigation, power generation, ecological preservation, and flood control.

4.2. Dataset

Deformation monitoring data collected from 2015 to 2017 at points H1-3, H1-4, H1-5, and H1-6 on the typical 0 + 190 section of the dam were selected for analysis using the method proposed in this paper. The horizontal deformation data of the dam were analyzed. Figure 6 shows the layout of measurement points on the 0 + 190 section of the dam. The original data were recorded at 7-day intervals, resulting in 100 sets of horizontal deformation time-series data for each monitoring point after preprocessing. The original data series is displayed in Figure 7.
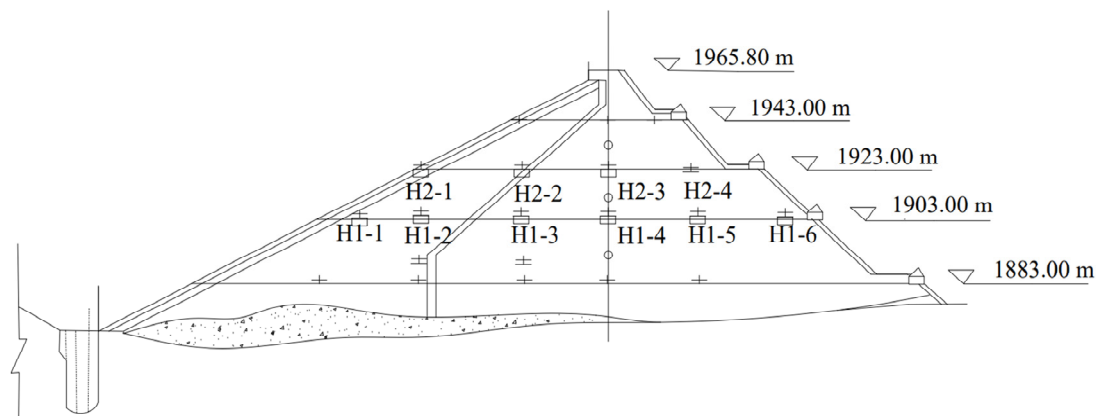

Figure 6. Diagram showing the layout of measurement points on the 0 + 190 section of the dam.
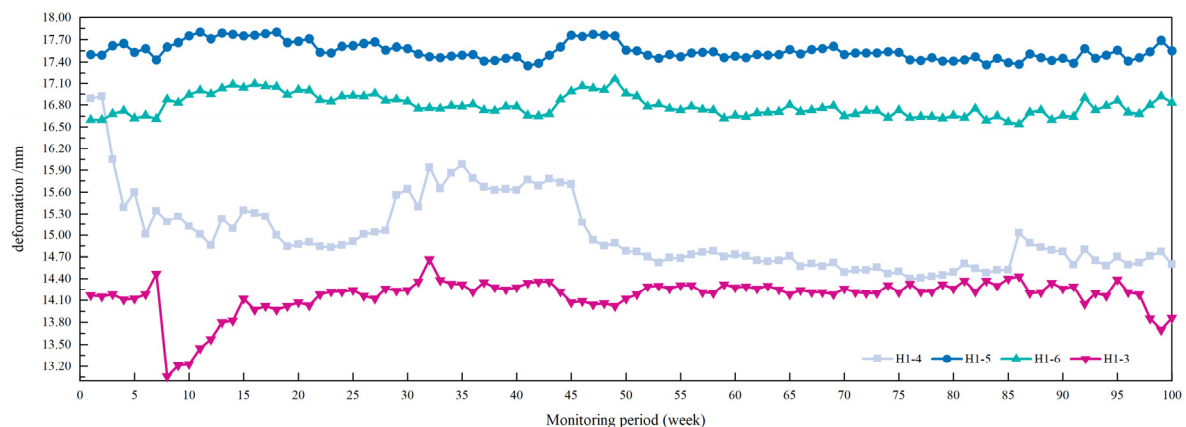

Figure 7. Measured deformation at the H1-3, H1-4, H1-5, and H1-6 deformation measurement points.

As shown in Figure 7, the variations in deformation data at monitoring points along the same stretch exhibit a certain degree of correlation. For instance, the curves representing data from the adjacent measurement points H1-3 and H1-4 as well as H1-5 and H1-6 display similar trends over specific time periods. These similarities can also be observed during rapid increases or decreases in deformation over certain time intervals.

4.3. Model Experiments

The model experiments were divided into two parts: first, the CNN–BiLSTM–AM model was applied to predict deformation; second, the prediction model was used to fill in missing time-series data. The performance of the model was evaluated through these experiments. All model codes were written in Matlab language.

For comparison with the CNN–BiLSTM–AM model, four deep learning models (BiLSTM, LSTM, CNN, and BiLSTM-AM) and one machine learning model (multilayer perceptron, MLP) were used to predict the radial displacement of the H1-3 measurement point in the 0 + 190 section. By ensuring consistent conditions, equitable comparisons can be conducted, thereby facilitating a precise assessment of the performance of each model [30,31]. Table 1 presents the hyperparameters for all deep learning and machine learning models. Epoch represents the number of training epochs; Batch size indicates the batch size; Lr denotes the learning rate; Nh denotes the number of neurons in the hidden layer, while Ni denotes the number of hidden layers; Head represents different attention heads in the multi-head attention mechanism; Keys are vectors associated with each position in the input sequence; and Convolution kernel size refers to the size of the convolutional kernel in a convolutional neural network.

Table 1. Hyperparameters of the six models.

Model Type	Model Name	Parameters
Deep learning	CNN–BiLSTM–AM	Epoch = 70, batch size = 7, l_r = 0.001, head = 1, keys = 2, convolution kernel size = 64, Nh = 64
	BiLSTM–AM	Epoch = 70, batch size = 7, l_r = 0.001, head = 1, keys = 2, Nh = 64
	BiLSTM	Epoch = 70, batch size = 7, l_r = 0.001, Nh = 64
	LSTM	Epoch = 70, batch size = 7, l_r = 0.001, Nh = 64
	CNN	Epoch = 70, batch size = 7, l_r = 0.001, convolution kernel size = 64
Machine learning	MLP	NI = 1, Nh = 64

In this study, the selection of hyperparameters was based on a balance between model performance and computational efficiency. The convolution kernel size was set to 64, determined by balancing model fitting accuracy and computational time. The number of hidden units was configured as 64 to capture temporal dependencies while avoiding insufficient information with fewer units, or overfitting with more. The multi-head attention mechanism employed a single head with two key vectors to strike a balance between feature representation capability and computational cost. The initial learning rate was set to 0.001 to ensure a stable exploration of the loss function space and prevent local oscillations. For the relatively small dataset, a "small batch size + multiple training rounds" strategy was adopted, with the maximum number of training epochs set to 70 and the batch size defined as one-tenth of the epoch size, ensuring an optimal trade-off between model performance and training efficiency. All hyperparameter configurations were experimentally validated to meet the objectives of this study.

4.3.1. Deformation Prediction Experiments

When applying our proposed model to deformation prediction, the feature data corresponding to the predicted data were not used as input during model training. A

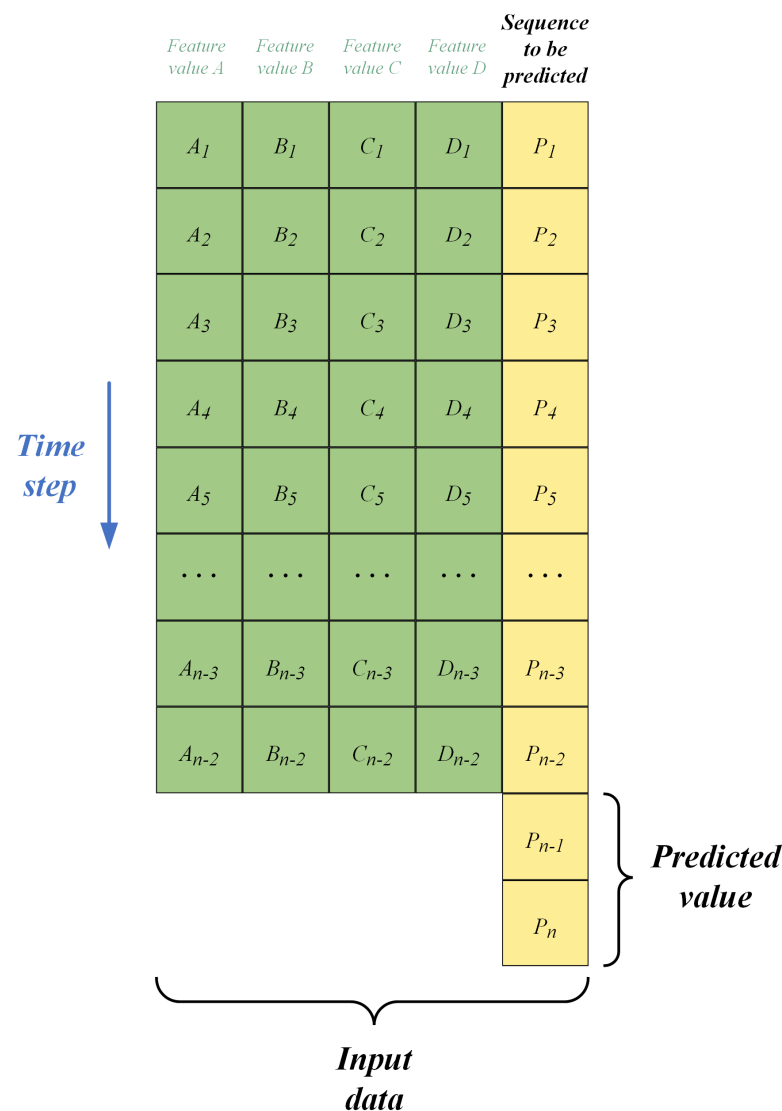schematic diagram showing the selection of the deformation prediction dataset is shown in Figure 8.

Figure 8. Schematic diagram showing the selection of the prediction dataset.

For the experiment, deformation monitoring data from points H1-3, H1-4, H1-5, and H1-6 on the typical 0 + 190, collected from 2015 to 2017, were selected as the study objects. The proposed multi-measurement-point model for predicting earth-rock dam displacement was used to predict the deformation at measurement point H1-3. The dataset consisted of monitoring data from 100 time periods, collected from four homologous measurement points (H1-3, H1-4, H1-5, and H1-6). The dataset was divided into training, validation, and test sets in a ratio of 8:1:1. Specifically, the first 80 datasets were used for training, the next 10 sets were used for validation, and the final 10 sets were used for testing. During deformation prediction, the data from H1-4, H1-5, and H1-6 were used as the feature data to predict future sequences at H1-3. Importantly, some features corresponding to the predicted sequence were not used in model training.

4.3.2. Data Filling Experiment

When the proposed model was applied for data imputation, feature data corresponding to the missing data were used as the input for model training. A schematic diagram showing the process for selecting the imputation dataset is shown in Figure 9.
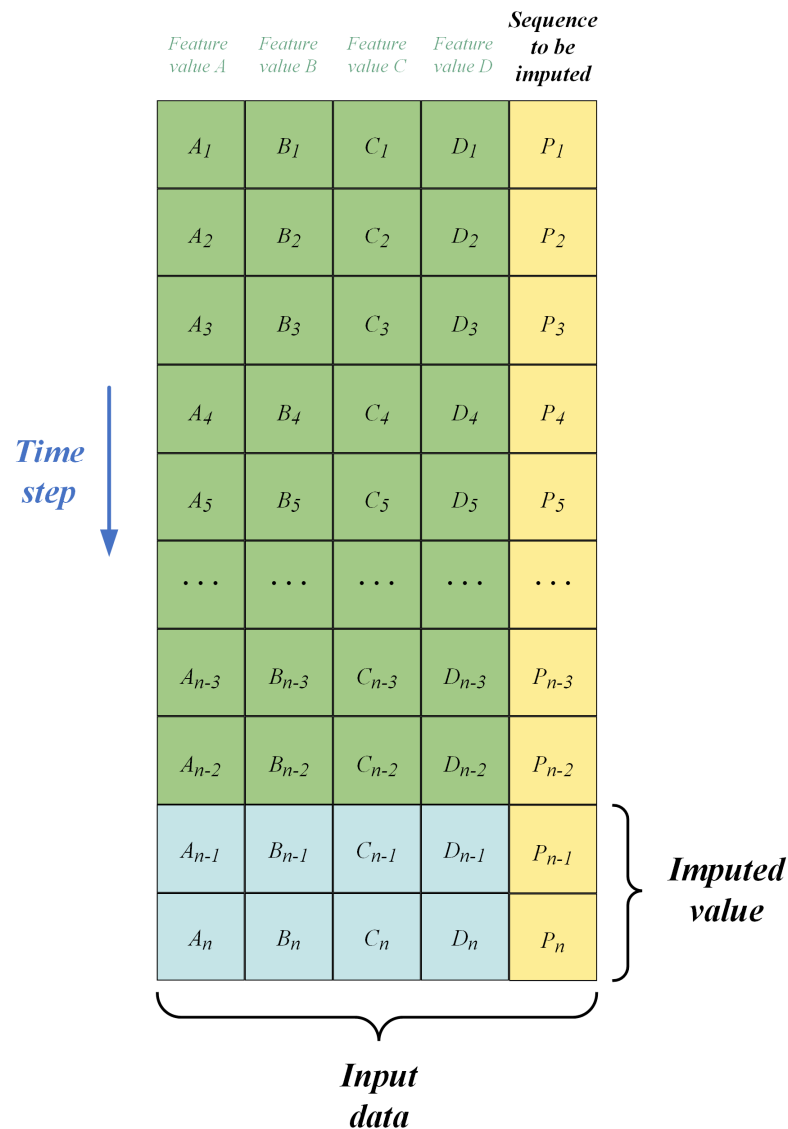
Figure 9. Schematic diagram showing the selection of the filling datasets.

The division of data for the imputation experiment into the training set, validation set, and test set followed the process used for the deformation prediction experiment. However, when the proposed multi-point displacement model was used to impute missing data, the feature data from measurement points H1-4, H1-5, and H1-6, along with the feature data corresponding to the missing segments of the H1-3 sequence, were incorporated into the model for training and data construction. The missing values in the H1-3 sequence were treated as the target data for prediction. Subsequently, the model was used to predict the missing segments, and the predicted values were used to fill in the missing data.

5. Results and Analyses

To assess the effectiveness of the proposed CNN–BiLSTM–AM multi-point model for earth-rock dam displacement monitoring, a series of control experiments were designed to verify the model from two perspectives: data prediction and data imputation. The results are presented below.

5.1. Analysis of the Deformation Experiment Results

In this experiment, six models—CNN–BiLSTM–AM, BiLSTM–AM, BiLSTM, LSTM, CNN, and MLP—were trained to predict deformation data at different monitoring points.

As shown in Figure 10a–f, which show the prediction results for each model on the test set, none of the prediction curves fully capture the fluctuations of the measured data. This discrepancy is mainly attributed to the large fluctuations and strong nonlinearity inherent in the dam monitoring data. For the prediction of dam deformation, the model must effectively reflect the changes in deformation at different monitoring points over specific time periods. This ability allows for a deeper understanding of the internal patterns within the data. Among the tested models, the predictions of the proposed CNN–BiLSTM–AM model were closest to the measured values in the test set. Nevertheless, some discrepancies were observed in the prediction results, particularly at peak points.
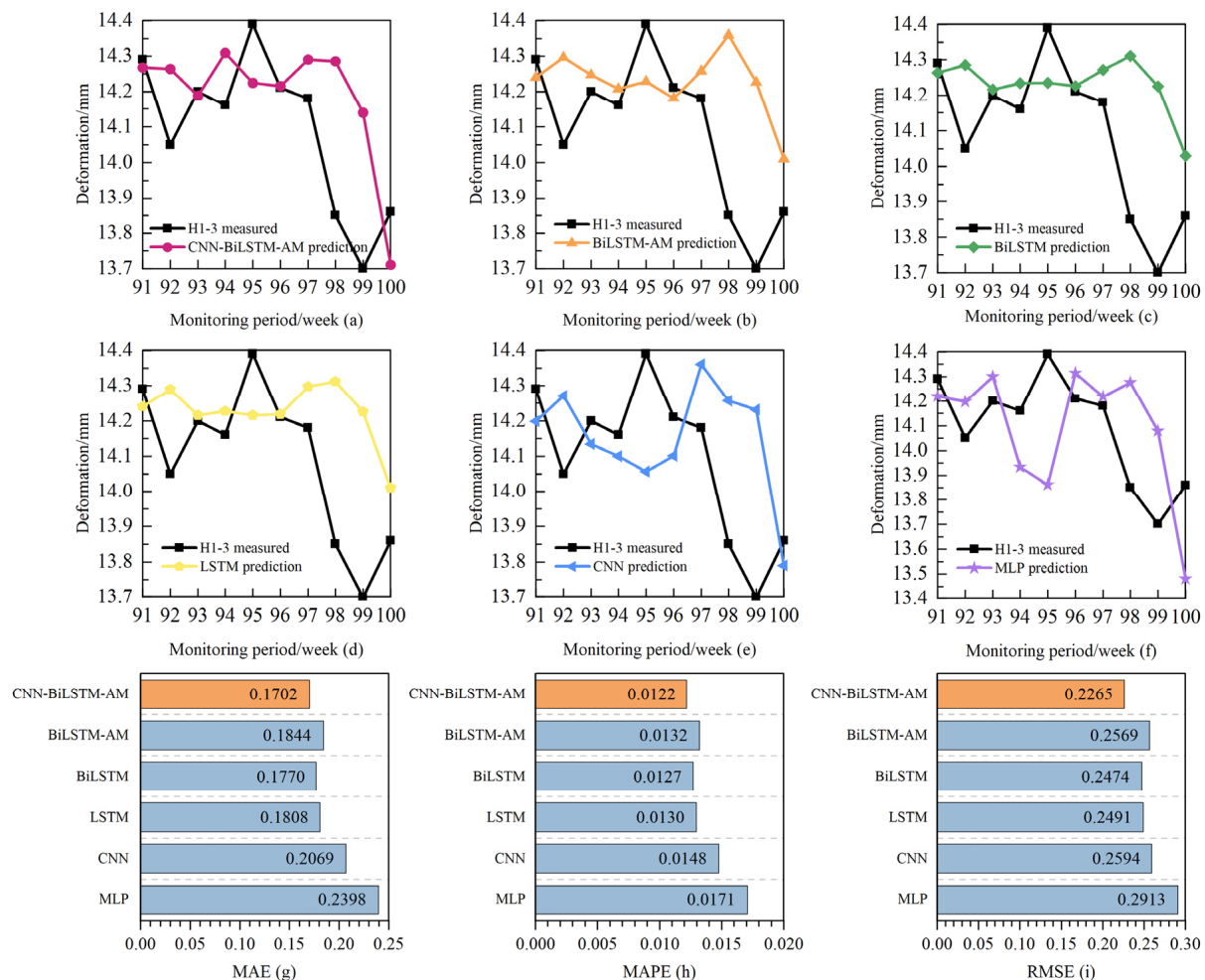

Figure 10. Comparison of the prediction results and evaluation indicators for each model.

The performance of the proposed CNN–BiLSTM–AM model was evaluated using several metrics, as shown in Figure 10g–i. The CNN–BiLSTM–AM model performed better on small datasets than the other models, with lower RMSE, MAE, and MAPE values compared with the other five models. Specifically, the RMSE, MAE, and MAPE values for the CNN–BiLSTM–AM model were 0.1702, 0.0122, and 0.2265, respectively, indicating high computational accuracy. By contrast, Figure 10g–i shows that the MLP model had the highest MAP, MAPE, and RMSE values; although MLP can handle nonlinearities, it struggles to capture long-term correlations within the data series. Furthermore, since the BiLSTM and LSTM models account for temporal dependencies, they improve the data prediction accuracy more effectively than the CNN model. Overall, the prediction performance of the models decreased in the following order: CNN–BiLSTM–AM > BiLSTM > LSTM > BiLSTM–AM > CNN > MLP.

5.2. Analysis of Data Filling Experimental Results

Unlike with data prediction, in the data filling process, the data features corresponding to the test set are included in the model training. The performance of each model on the test set is shown in Figure 11a–f. When the relevant data features of the test set were incorporated into model training, the stability of the models improved, allowing them to better capture the trends in the data. Figure 11g–i shows the evaluation metrics (MAE, MAPE, and RMSE) for each model during the data filling process. The values of these metrics were lower than those for data prediction, suggesting an improvement in prediction accuracy after adding the feature dataset. This highlights the feasibility of using the predicted values to fill in for missing data points. As in the prediction experiment, the CNN–BiLSTM–AM model outperformed the other five models in the data filling experiments, resulting in the smallest MAE, MAPE, and RMSE values (0.1465, 0.0105, and 0.1897, respectively). These MAE, MAPE, and RMSE values represent improvements of 13.92%, 6.25%, and 16.65% compared with the MLP model, respectively. The results demonstrate the effectiveness of deep learning models for filling in missing data based on the correlation characteristics between homologous measurement points.
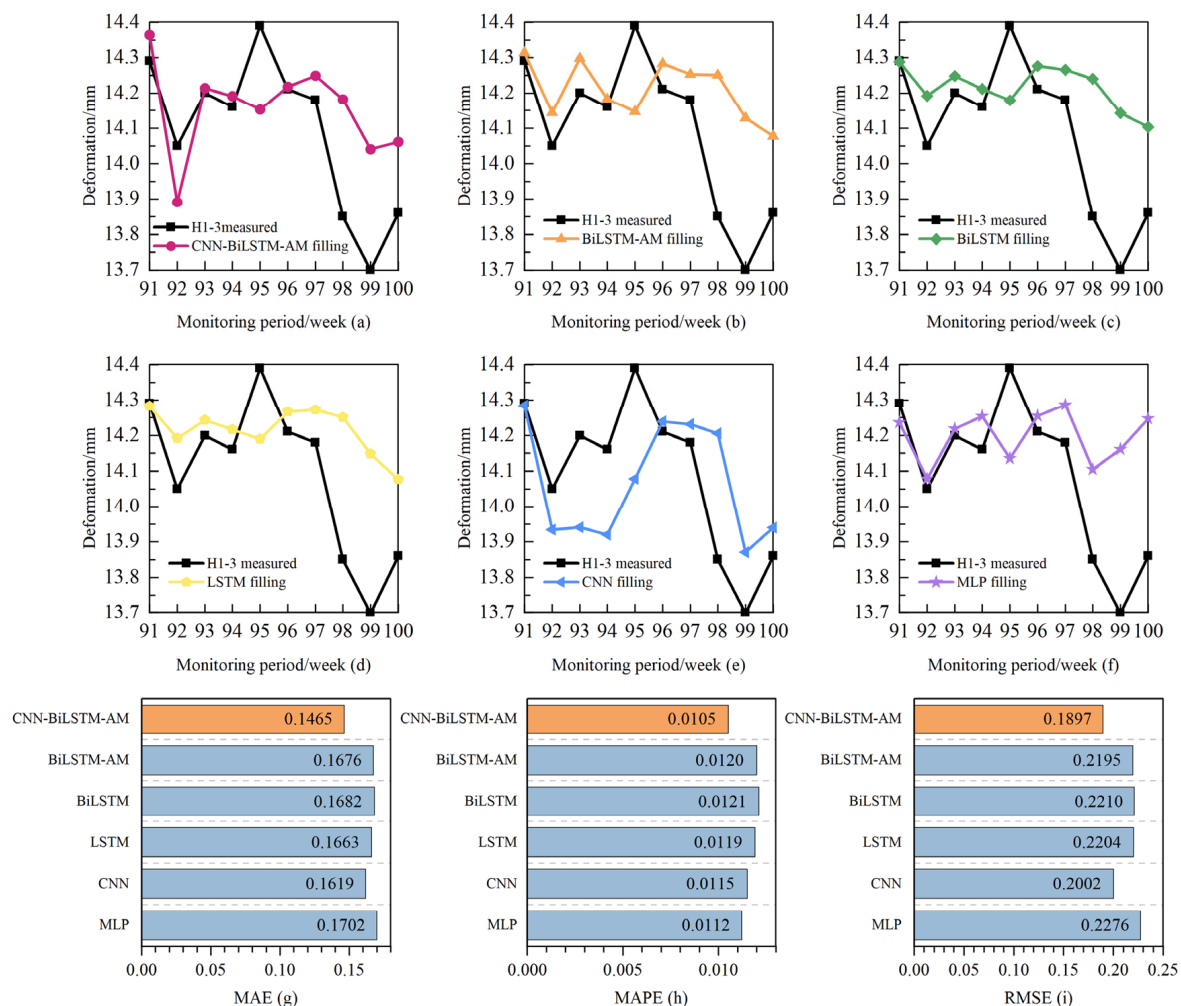

Figure 11. Comparison of the filling results of each model and the effects on evaluation indicators.

5.3. The Limitations of the Model and Considerations in Practical Applications

5.3.1. The Constraints of Model Experiments

From the horizontal comparison of the model results shown in Figures 10 and 11, it can be observed that the error evaluation metrics for all models are lower in the data imputation experiments. This indicates that the models exhibit better performance and

higher predictive accuracy in these experiments. This also suggests that the distribution of the data significantly impacts the model's predictive performance—namely, the more comprehensive the training feature dataset, the better the robustness of the model. However, it should be noted that both the data prediction and data imputation experiments were conducted on datasets with small sample sizes, where predictions or imputations were performed for the last 10 monitoring periods. From the perspective of the curve shapes, all models demonstrated better short-term predictive performance (e.g., for weeks 91–95), with predicted values closely matching the actual values. However, during periods of dramatic temporal changes (e.g., weeks 96–100), all models exhibited certain degrees of deviation and reduced data fitting accuracy. This indicates that there is still room for improvement in the long-term predictive performance of the models.

5.3.2. Considerations for Practical Applications of the Model

Although the CNN–BiLSTM–AM model demonstrates excellent performance in dam deformation monitoring for both data prediction and imputation, it still faces certain limitations and computational challenges. Firstly, the model heavily relies on high-quality and complete monitoring data. Noise or missing data can significantly affect prediction accuracy, particularly when the number of monitoring points is limited. Secondly, the model exhibits limited generalization capability across different scenarios; applying it directly to other structures or regions may require retraining. Additionally, while the model is well-suited for short-term predictions, its performance may degrade in long-term predictions due to error accumulation. Moreover, the model's "black-box" nature poses challenges to its interpretability in engineering applications. This study primarily focuses on short- or medium-term predictions (e.g., deformation trends over a few time steps), but for long-term predictions (e.g., deformation trends spanning years), error accumulation may become a significant issue. Future work could integrate physics-driven models (e.g., finite element analysis) with data-driven models to achieve synergistic optimization for both short- and long-term predictions.

While enhancing prediction performance, the CNN–BiLSTM–AM model also imposes higher demands on computational resources and deployment conditions. The combination of CNN, BiLSTM, and the attention mechanism increases computational complexity, especially when dealing with a large number of monitoring points or longer time series inputs. In terms of computational requirements, the integration of CNN, BiLSTM, and attention mechanisms substantially raises both the computational load and memory demands, making the model unsuitable for resource-constrained devices or real-time deployments. Large-scale training requires significant computational resources, and the inference phase demonstrates relatively low efficiency. Future research could explore lightweight network architectures and incorporate distributed training strategies to improve the model's computational efficiency.

6. Conclusions

To address the inherent nonlinearity and complexity of dam monitoring data, we have introduced a novel CNN–BiLSTM–AM multi-point displacement monitoring model for earth-rock dams. In this model, a CNN is employed to extract features from the input data, BiLSTM is responsible for learning and predicting based on these features, and the AM captures the influence of temporal data feature states on the prediction outcomes. Using the correlation between homologous observation points, the effectiveness of the model was assessed in data prediction and data filling, demonstrating its accuracy and reliability. Key findings from the study include:

- Improved Capture of Spatial Relationships: Unlike traditional single-point displacement models, the multi-point model can capture the spatial relationships between measurement points, overcoming the limitations resulting from relying solely on data from individual points. The experimental results confirm the superior predictive

performance of the proposed model as well as its high reliability in filling missing values in monitoring datasets.

- Superiority of Deep Learning Models: Compared with machine learning models, deep learning models show greater accuracy in both dam deformation prediction and data filling. This is largely attributed to the deeper, more complex network layers of deep learning models, which enable better feature extraction and improved prediction capabilities.
- Hybrid Model Advantages: Achieving high prediction accuracy with a single network is challenging. However, hybrid models—such as the combination of CNNs with other models—can improve accuracy by leveraging a CNN's ability to capture spatiotemporal features in the monitoring data. Although hybrid models increase complexity, the experimental results demonstrate that they outperform single models in predictive accuracy.
- Applications in Future Work: For researchers and engineers aiming to enhance prediction accuracy, the CNN–BiLSTM–AM model is a robust choice. Compared with existing models, the proposed model achieves higher prediction accuracy, making it a valuable tool for dam monitoring data analysis and decision-making.

Author Contributions: Conceptualization, Methodology, Software, Validation, writing—original draft preparation, L.P.; Resources, Data curation, J.S.; Writing—review and editing, Supervision, C.Y. All authors have read and agreed to the published version of the manuscript.

Data Availability Statement: The original contributions presented in the study are included in the article, further inquiries can be directed to the corresponding author.

Conflicts of Interest: The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

1. 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]
2. Kang, F.; Li, J.; Dai, J. Prediction of Long-Term Temperature Effect in Structural Health Monitoring of Concrete Dams Using Support Vector Machines with Jaya Optimizer and Salp Swarm Algorithms. *Adv. Eng. Softw.* **2019**, *131*, 60–76. [CrossRef]
3. Cai, S.; Gao, H.; Zhang, J.; Peng, M. A Self-Attention-LSTM Method for Dam Deformation Prediction Based on CEEMDAN Optimization. *Appl. Soft Comput.* **2024**, *159*, 111615. [CrossRef]
4. Mao, Y.; Li, J.; Qi, Z.; Yuan, J.; Xu, X.; Jin, X.; Du, X. Research on Outlier Detection Methods for Dam Monitoring Data Based on Post-Data Classification. *Buildings* **2024**, *14*, 2758. [CrossRef]
5. Li, F.; Wang, Z.; Liu, G.; Fu, C.; Wang, J. Hydrostatic Seasonal State Model for Monitoring Data Analysis of Concrete Dams. *Struct. Infrastruct. Eng.* **2015**, *11*, 1616–1631. [CrossRef]
6. 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]
7. Liu, M.; Feng, Y.; Yang, S.; Su, H. Dam Deformation Prediction Considering the Seasonal Fluctuations Using Ensemble Learning Algorithm. *Buildings* **2024**, *14*, 2163. [CrossRef]
8. Li, Z.; Jia, H.; Zhang, Y.; Liang, J.; Abdelhafiz, A.E.; Cao, Q.; Lu, W. Deflection Statistical Monitoring Model Identification of the Concrete Gravity Dam Based on Uncertainty Analysis. *Struct. Control Health Monit.* **2022**, *29*, e3026. [CrossRef]
9. Zhou, W.; Hua, J.; Chang, X.; Zhou, C. Settlement Analysis of the Shuibuya Concrete-Face Rockfill Dam. *Comput. Geotech.* **2011**, *38*, 269–280. [CrossRef]
10. Szymanowski, M.; Kryza, M.; Spallek, W. Regression-Based Air Temperature Spatial Prediction Models: An Example from Poland. *Meteorol. Z.* **2013**, *22*, 577–585. [CrossRef]

11. Alalade, M.; Nguyen-Tuan, L.; Wuttke, F.; Lahmer, T. Damage Identification in Gravity Dams Using Dynamic Coupled Hydro-Mechanical XFEM. *Int. J. Mech. Mater. Des.* **2018**, *14*, 157–175. [CrossRef]

12. Li, R.; Jie, Y.; Pengli, Z.; Jiaming, W.; Chunhui, M.; Chao, C.; Lin, C.; Jian'e, W.; Mingjuan, Z. A Hybrid Monitoring Model of Rockfill Dams Considering the Spatial Variability of Rockfill Materials and a Method for Determining the Monitoring Indexes. *J. Civ. Struct. Health Monit.* **2022**, *12*, 817–832. [CrossRef]

13. Zhang, K.; Gu, C.; Zhu, Y.; Li, Y.; Shu, X. A Mathematical-Mechanical Hybrid Driven Approach for Determining the Deformation Monitoring Indexes of Concrete Dam. *Eng. Struct.* **2023**, *277*, 115353. [CrossRef]

14. Xu, C.; Yue, D.; Deng, C. Hybrid GA/SIMPLS as Alternative Regression Model in Dam Deformation Analysis. *Eng. Appl. Artif. Intell.* **2012**, *25*, 468–475. [CrossRef]

15. 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]

16. Liu, H.-F.; Ren, C.; Zheng, Z.-T.; Liang, Y.-J.; Lu, X.-J. Study of a Gray Genetic BP Neural Network Model in Fault Monitoring and a Diagnosis System for Dam Safety. *ISPRS Int. J. GEO-Inf.* **2018**, *7*, 4. [CrossRef]

17. Ren, F.; Wu, X.; Zhang, K.; Niu, R. Application of Wavelet Analysis and a Particle Swarm-Optimized Support Vector Machine to Predict the Displacement of the Shuping Landslide in the Three Gorges, China. *Environ. Earth Sci.* **2015**, *73*, 4791–4804. [CrossRef]

18. Liu, W.; Pan, J.; Ren, Y.; Wu, Z.; Wang, J. Coupling Prediction Model for Long-Term Displacements of Arch Dams Based on Long Short-Term Memory Network. *Struct. Control Health Monit.* **2020**, *27*, e2548. [CrossRef]

19. Wei, Y.; Li, Q.; Hu, Y.; Wang, Y.; Zhu, X.; Tan, Y.; Liu, C.; Pei, L. Deformation Prediction Model Based on an Improved CNN plus LSTM Model for the First Impoundment of Super-High Arch Dams. *J. Civ. Struct. Health Monit.* **2023**, *13*, 431–442. [CrossRef]

20. Jiedeerbieke, M.; Li, T.; Chao, Y.; Qi, H.; Lin, C. Gravity Dam Deformation Prediction Model Based on I-KShape and ZOA-BiLSTM. *IEEE Access* **2024**, *12*, 50710–50722. [CrossRef]

21. Luo, S.; Wei, B.; Chen, L. Multi-Point Deformation Monitoring Model of Concrete Arch Dam Based on MVMD and 3D-CNN. *Appl. Math. Model.* **2024**, *125*, 812–826. [CrossRef]

22. Wei, B.; Liu, B.; Yuan, D.; Mao, Y.; Yao, S. Spatiotemporal Hybrid Model for Concrete Arch Dam Deformation Monitoring Considering Chaotic Effect of Residual Series. *Eng. Struct.* **2021**, *228*, 111488. [CrossRef]

23. Yang, G. Deformation Similarity Characteristics-Considered Hybrid Panel Model for Multi-Point Deformation Monitoring of Super-High Arch Dams in Operating Conditions. *Measurement* **2022**, *192*, 110908. [CrossRef]

24. LeCun, Y.; Boser, B.; Denker, J.S.; Henderson, D.; Howard, R.E.; Hubbard, W.; Jackel, L.D. Backpropagation Applied to Handwritten Zip Code Recognition. *Neural Comput.* **1989**, *1*, 541–551. [CrossRef]

25. Pan, M.; Zhou, H.; Cao, J.; Liu, Y.; Hao, J.; Li, S.; Chen, C.-H. Water Level Prediction Model Based on GRU and CNN. *IEEE Access* **2020**, *8*, 60090–60100. [CrossRef]

26. Bengio, Y.; Simard, P.; Frasconi, P. Learning Long-Term Dependencies with Gradient Descent Is Difficult. *IEEE Trans. Neural Netw.* **1994**, *5*, 157–166. [CrossRef]

27. Graves, A.; Schmidhuber, J. Framewise Phoneme Classification with Bidirectional LSTM and Other Neural Network Architectures. *Neural Netw.* **2005**, *18*, 602–610. [CrossRef]

28. Kavianpour, P.; Kavianpour, M.; Jahani, E.; Ramezani, A. A CNN-BiLSTM Model with Attention Mechanism for Earthquake Prediction. *J. Supercomput.* **2023**, *79*, 19194–19226. [CrossRef]

29. Hrynaszkiewicz, I. A Call for BMC Research Notes Contributions Promoting Best Practice in Data Standardization, Sharing and Publication. *BMC Res. Notes* **2010**, *3*, 235. [CrossRef]

30. Lu, W.; Li, J.; Wang, J.; Qin, L. A CNN-BiLSTM-AM Method for Stock Price Prediction. *Neural Comput. Appl.* **2021**, *33*, 4741–4753. [CrossRef]

31. Kang, X.; Li, Y.; Zhang, Y.; Wen, L.; Sun, X.; Wang, J. PCA-IEM-DARNN: An Enhanced Dual-Stage Deep Learning Prediction Model for Concrete Dam Deformation Based on Feature Decomposition. *Measurement* **2025**, *242*, 115664. [CrossRef]