Predicting the Deformation of a Concrete Dam Using an Integration of Long Short-Term Memory (LSTM) Networks and Kolmogorov–Arnold Networks (KANs) with a Dual-Stage Attention Mechanism

Rui Xu [1], Xingyang Liu [1,2,3,*], Jiahao Wei [4], Xingxing Ai [1], Zhanchao Li [1] and Hairui He [5]

[1] College of Hydraulic Science and Engineering, Yangzhou University, Yangzhou 225009, China; 233201332@stu.yzu.edu.cn (R.X.); 15879932012@163.com (X.A.); 006520@yzu.edu.cn (Z.L.)
[2] National Dam Safety Research Center, Wuhan 430019, China
[3] State Key Laboratory of Coastal and Offshore Engineering, Dalian University of Technology, Dalian 116024, China
[4] School of Hydraulic Engineering, Dalian University of Technology, Dalian 116024, China; wjhao@mail.dlut.edu.cn
[5] Ningbo Reservoir Management Center, Ningbo 315020, China; hhhhhhairui@sina.com
* Correspondence: 007975@yzu.edu.cn

Abstract: An accurate prediction model for dam deformation is crucial for ensuring the safety and operational integrity of dam structures. This study introduces a hybrid modeling approach that integrates long short-term memory (LSTM) networks with Kolmogorov–Arnold networks (KANs). Additionally, the model incorporates a dual-stage attention mechanism (DA) that includes both factor and temporal attention components, enhancing the model's precision and interpretability. The effectiveness of the DA-LSTM-KAN model was validated through a case study involving a concrete gravity dam. A comparative analysis with traditional models, including multiple linear regression and various LSTM variants, demonstrated that the DA-LSTM-KAN model significantly outperformed these alternatives in predicting dam deformation. An interpretability analysis further revealed that the seasonal and hydrostatic components contributed significantly to the horizontal displacement, while the irreversible component had the least impact. This importance ranking was qualitatively consistent with the results obtained from the Shapley Additive Explanations (SHAP) method and the relative weight method. The enhancement of the model's predictive and explanatory capabilities underscores the hybrid model's utility in providing detailed and actionable intelligence for dam safety monitoring.

Keywords: monitoring; dam deformation; LSTM; KAN; attention; SHAP

1. Introduction

The reservoir dam is a critical infrastructure in irrigation, flood control, water resource distribution, and hydroelectric power generation. However, despite its valuable utility, dam failures can lead to significant property damage and loss of life in surrounding areas. Therefore, accurately identifying the operational status of dams and issuing timely alerts for any abnormal behavior are crucial for ensuring the long-term operation safety of dams [1]. Concrete dams are often the preferred option for dams that need to be taller than 60 m. The concrete dam responds to variations in load and environmental factors by exhibiting changes such as deformation, seepage, cracking, and other structural responses. To effectively monitor dam behavior, a wide array of instruments are strategically placed both within and around the structures to assess environmental variables and structural responses. This strategic deployment results in the accumulation of detailed databases that are essential for the ongoing surveillance of dam performance. Notably, among the

various monitoring efforts, deformation monitoring is distinguished by its early adoption, exceptional reliability, extended operational lifespan, and high frequency of data collection. It provides reliable insights for analyzing the deformation behavior [2]. Moreover, the deformation behavior intuitively and comprehensively reflects the working status of dams [3]. Thus, developing a predictive monitoring model for deformation based on prototype monitoring data is crucial for evaluating dam operational safety.

The most commonly used monitoring model of dams is the statistical model [4]. Its basic principle involves categorizing factors influencing concrete dam deformation into three main parts: water pressure, temperature, and aging. These factors are then represented as polynomials. The Hydrostatic–Seasonal–Time (HST) model is a benchmark statistical approach that quantitatively analyzes the factors influencing dam deformation based on mechanical theory assumptions. The thermal effect in the HST model is described by harmonic sinusoidal functions. Building upon the HST framework, various adaptations of the HST model have been introduced [5]. For instance, Penot et al. [6] developed the Hydraulic–Seasonal–Thermal–Time (HSTT) model by incorporating a thermal correction factor into the original HST model to account for air temperature influences. Similarly, Léger and Leclerc [7] derived the Hydrostatic–Temperature–Time (HTT) statistical model, wherein the seasonal component of the HST model is substituted with actual temperature recordings. Compared to the HST model, the HTT model more accurately captures the thermal effect on dam behavior when sufficient measured temperature data are available [3]. Furthermore, Tatin et al. [8] proposed a physico-statistical model (HST-Grad) that incorporates the water temperature profile, allowing both the mean temperature and the temperature gradient within the structure to be properly accounted for. Hu and Ma [9] developed a statistical model to enhance the accuracy of dam displacement estimations during initial impoundment phases, specifically by refining the assessment of thermal and time-dependent effects. Statistical approaches are widely utilized in practical hydraulic engineering projects because they offer simple formulas and rapid response times. Nonetheless, dams are inherently complex and dynamic systems. Given the variety of structural forms and the complexity of external environmental factors, dams are characterized by both uncertainty and diversity, with a typically non-linear mapping relationship between dam behavior and its underlying causes. Additionally, there is also multi-collinearity among the independent variables used to predict dam deformation, potentially resulting in unreasonable interpretation of regression coefficients. To overcome this issue, methodologies such as principal component regression [3] and partial least-squares regression [10] have been employed.

With ongoing advancements in dam monitoring theory and artificial intelligence, machine learning (ML)-based models, such as artificial neural networks (ANNs) [11,12], random forest (RF) [13–15], support vector machine (SVR) [16–18], and extreme learning machine (ELM) [19–21], have been proposed. These models have demonstrated unique advantages in addressing the issues of uncertainty and nonlinearity associated with the monitoring model factors and often yield more precise predictions than traditional statistical models in various scenarios. Furthermore, a variety of intelligent optimization algorithms, such as genetic algorithm [22], ant lion algorithm [23], particle swarm algorithm [24], artificial fish swarm algorithm [25], salp swarm algorithm [26], were utilized for hyperparameter tuning in ML-based models, thereby enhancing their predictive capabilities. ML models do not typically incorporate time dependencies. Instead, they approach the prediction of dam deformation as a static regression task, thereby neglecting any potential temporal relationships within the deformation data during the model development process [27]. However, dam deformation undergoes real-time changes that are influenced by temporal variations and environmental factors. Advancements in deep learning have greatly addressed the shortcomings of traditional machine learning approaches. By utilizing multiple layers within their networks, deep learning models achieve improved prediction accuracy and better handle nonlinear mappings compared to standard ML-based methods. Techniques such as recurrent neural networks (RNNs) [28] and long short-term memory

(LSTM) networks [29] have shown exceptional performance in managing multivariate time series data. RNNs are capable of effectively transmitting temporal information from the input sequence throughout the prediction process. Additionally, the specialized three-gate architecture of LSTM networks resolves the common RNN challenges of gradient vanishing and explosion [30]. Recently, several researchers have employed LSTM networks for predicting dam deformation, achieving significant and satisfactory outcomes. Qu et al. [31] focused on developing prediction models for concrete dam deformation by integrating the rough set (RS) theory with LSTM networks. Liu et al. [32] integrated principal component analysis (PCA) and the moving average (MA) method with the LSTM to achieve two integrated prediction models (i.e., LSTM-PCA and LSTM-MA), aimed at predicting the long-term deformation of Lijiaxia arch dam. Yang et al. [33] compared various models for predicting the deformation of concrete dams. The results indicate that LSTM-based models are potentially more effective for analyzing time-dependent data and capturing temporal correlations.

Although deep learning models have made significant advancements, there remain inherent limitations associated with these methodologies. For instance, many models exhibit significant memory demands and computational complexities, alongside the constraint of single-step prediction and extended training times [34]. Moreover, the limited interpretability of black box models often renders them impractical for engineering applications, as it is challenging to explain the relationship between input and output [35]. Recently, LSTM-based models were modified with attention mechanisms [36] to enhance interpretability. This mechanism, modeled after the way visual attention is distributed in the human brain, dynamically filters and prioritizes information from a wide range of input features. It uses attention weights to emphasize the significance of different time steps within the sequence. Yang et al. [37] enhanced the LSTM model with an attention mechanism to identify information that significantly influences deformation. Shu et al. [38] introduced a forecasting model that integrates a variational autoencoder (VAE) with a temporal attention-enhanced LSTM network. Ren et al. [39] proposed a prediction model for concrete dam deformation by integrating encoder–decoder architecture, attention mechanism, and LSTM neural network. Cai et al. [40] utilized an innovative decomposition-based algorithm to enhance the LSTM network by incorporating a self-attention mechanism. On the other hand, numerous studies have explored the application of model-agnostic ML interpretation methods such as Local Interpretable Model-Agnostic Explanation (LIME) and Shapley Additive Explanations (SHAP) in geotechnical engineering [41,42], earthquake engineering [43,44], and structural engineering [45,46]. These methods provide valuable insights for knowledge discovery, model debugging, and justification of predictions. However, the application of LIME or SHAP in dam deformation monitoring has been relatively limited. More recently, Li et al. [47] used SHAP to enhance the interpretability of the light gradient boosting model for high arch dam stress analysis.

In summary, researchers have extensively investigated the use of LSTM in predicting dam deformation, thereby achieving notable results. Several studies also explored the integration of LSTM with other machine learning or deep learning models to enhance prediction accuracy. However, efforts to simultaneously enhance the predictive capability and interpretability of dam monitoring models remain constrained. More recently, a novel neural network architecture, Kolmogorov–Arnold Networks (KANs), was proposed by a MIT team [48]. This architecture was developed as an alternative to the conventional Multi-Layer Perceptron (MLP) and has quickly attracted global attention within the AI community. Drawing inspiration from the Kolmogorov–Arnold representation theorem [49], the KAN replaces traditional linear weights with spline-parametrized univariate functions. This modification enables the dynamic learning of activation patterns and significantly improves interpretability [50,51]. To the best of the authors' knowledge, there is limited research on the application of KANs for predicting dam deformation, particularly concerning model construction, parameter optimization, and performance evaluation. Therefore, this study aims to develop and implement a hybrid model for dam deformation by combining

LSTM with KAN. The proposed model integrates the memory capabilities of LSTM with the nonlinear expressive strengths of the KAN, thereby overcoming the limitations of single models in handling complex time series data associated with dam deformation.

2. Models for the Deformation Monitoring of Concrete Dams

2.1. Hydrostatic–Seasonal–Time (HST)

In the classic HST model expressed in Equation (1), dam displacement $\delta(H, S, T)$ is segmented into three distinct components: the hydrostatic component $\delta(H)$, which accounts for the effects of reservoir water; the seasonal component $\delta(S)$, associated with temperature variations; and the irreversible component $\delta(T)$, encompassing factors such as heat of hydration dissipation, creep, and alkali–aggregate reactions.

$$\delta(H, S, T) = \delta(H) + \delta(S) + \delta(T) + \varepsilon \tag{1}$$

where ε is the residuals.

The hydrostatic component $\delta(H)$ includes the deformation of the dam body under hydrostatic pressure, the deformation of the dam body caused by the dam bedrock deformation, and the deformation of the dam body resulting from the rotation of the dam bedrock due to the water weight. According to the mechanical analysis, $\delta(H)$ is defined as a polynomial function of water height:

$$\delta(H) = \sum_{i=1}^{n} a_i H^i \tag{2}$$

where H is the upstream water height, a_i is the fitting coefficient, and i is the power exponent. In general, n is 4 or 5 for an arch dam and 3 for a gravity dam.

The seasonal component arises from temperature variations within the dam and its foundation rock, which are influenced by the ambient air temperature. For a concrete dam undergoing long-term operation, the heat of cement hydration has already dissipated, and the dam body's temperature primarily changes seasonally due to variations in ambient temperature. Consequently, the seasonal component is determined as the superposition of harmonic sinusoidal functions:

$$\delta(S) = \sum_{i=1}^{m} \left(b_{1i} \sin \frac{2\pi it}{365} + b_{2i} \cos \frac{2\pi it}{365} \right) \tag{3}$$

where i denotes the period, m denotes the cycle, with $m = 1$ for the annual cycle and $m = 2$ for a half cycle, b_{1i} and b_{2i} represent the fitting coefficients, and t is the cumulative number of days from the initial survey date to the current survey date.

The time-dependent irreversible component is intricate and thoroughly encompasses the creep and plastic deformations of both the dam structure and the bedrock, as well as the compressive deformation of the bedrock's geological formations. Changes in the irreversible component generally follow an increasing trend at the beginning and slow down over time, which can be described as:

$$\delta(T) = c_1 t + c_2 \ln(t) \tag{4}$$

where c_1 and c_2 are two fitting coefficients.

2.2. Long Short-Term Memory (LSTM)

LSTM, a variant of RNNs, addresses the vanishing gradient problem by incorporating a gating mechanism [52]. Figure 1 depicts the architecture of an LSTM. At each time step t, the LSTM layer retains a hidden memory cell $\tilde{c}_t$ along with three gating mechanisms (i.e., input gate, forget gate, and output gate). The LSTM cell takes the current input x_t, the previous output h_{t-1}, and the previous cell state c_{t-1} as inputs. These gates work together

to update or discard information. The input, forget, and output gates, along with the LSTM memory cell, are expressed as follows:

$$f_t = \sigma_g(W_f \cdot x_t + U_f \cdot h_{t-1} + b_f) \tag{5}$$

$$i_t = \sigma_g(W_i \cdot x_i + U_i \cdot h_{t-1} + b_i) \tag{6}$$

$$o_t = \sigma_g(W_o \cdot x_t + U_o \cdot h_{t-1} + b_o) \tag{7}$$

$$\widetilde{c}_t = \tanh(W_c \cdot x_t + U_c \cdot h_{t-1} + b_c) \tag{8}$$

where W_f, W_i, W_o, and W_c represent the weight matrices of four layers, each of which is connected to the input vector. U_f, U_i, U_o, and U_c represent the weight matrices of four layers, each of which is connected to the previous hidden state. b_f, b_i, b_o, and b_c are four bias vectors. σ_g and $\tanh(\cdot)$ are sigmoid and hyperbolic tangent activation functions, respectively. Then, the cell state vector state c_t and the memory cell output vector h_t can be expressed as:

$$c_t = f_t \odot c_{t-1} + i_t \odot \widetilde{c}_t \tag{9}$$

$$h_t = o_t \odot \tanh(c_t) \tag{10}$$
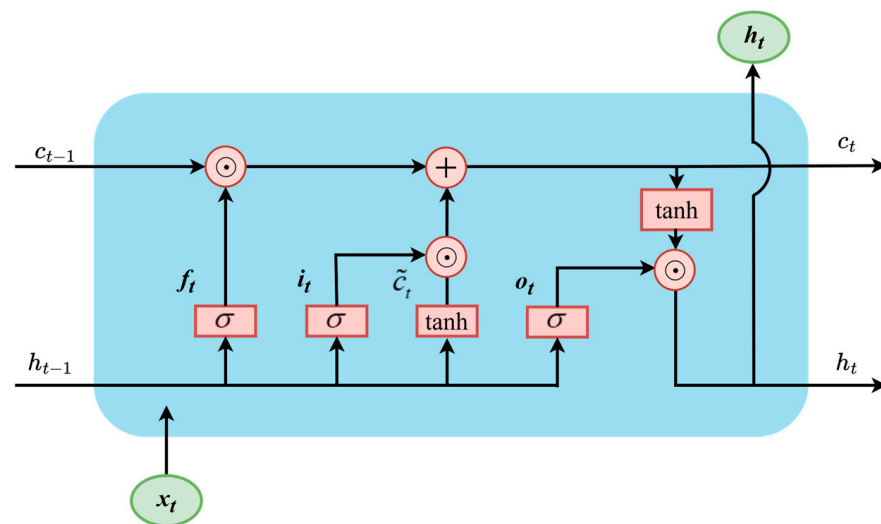
where $\odot$ represents multiply by element.

Figure 1. The structure diagram of LSTM.

2.3. Attention Mechanism

The attention mechanism [53] is a specialized architecture integrated into machine learning models that automatically determines the contribution of each input to the output. This facilitates decision-making from large datasets, reducing the computation costs, and enhances both learning efficiency and accuracy. The attention mechanism serves as an advancement to the encoder–decoder architecture by focusing on and accumulating information from inputs at each time step, thereby building a more effective memory representation. The framework of the scaled dot-product attention is described in Figure 2. To implement the attention mechanism, the raw input data are expressed as ⟨*key,value*⟩ pairs, and the similarity coefficient between the *Key* and the *Query* can be computed. The corresponding weight coefficient for the *Value* can be obtained using the *Query* from the specified task within the target. The weight coefficient *W* is then multiplied by the *Value* to produce the output *a*. When *Query*, *Key*, and *Value* are, respectively, denoted by *Q*, *K*, and *V*, the formulas for calculating *W* and *a* can be expressed as follows:

$$W = \text{softmax}(QK^T) \tag{11}$$

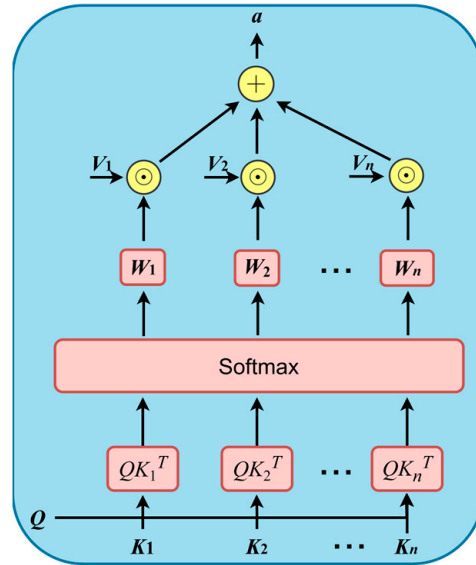$$a = attention(Q, K, V) = W \odot V \tag{12}$$

Figure 2. Illustration of the scaled dot-product attention.

2.4. Kolmogorov–Arnold Networks (KANs)

The Kolmogorov–Arnold Networks (KANs) present a neural network architecture that leverages learnable spline-based functions to enhance the approximation of complex nonlinear relationships. KANs rely on the Kolmogorov–Arnold representation theorem [49]. The theorem states that any multivariate continuous function *f*, which depends on $x = [x_1, x_2, \ldots, x_n]$, on a bounded domain, can be represented as a composition of univariate functions and the addition operation:

$$f(x_1, \ldots, x_n) = \sum_{q=1}^{2n+1} \Phi_q \left(\sum_{p=1}^{n} \phi_{q,p}(x_p) \right) \tag{13}$$

where $\phi_{q,p}$ indicates univariate functions that map each input variable x_p, and Φ_q indicates continuous functions.

In conventional MLPs, which are based on the universal approximation theorem, weight parameters are typically assigned to the edges of the network, whereas neurons are equipped with predefined activation functions. Different from this traditional approach, KANs incorporate learnable activation functions on the edges (i.e., "weights") between nodes. These activation functions dynamically adapt during network training, effectively replacing the network weight parameters with univariate spline functions. This innovative approach enables the network to maintain flexibility while achieving a precise fitting of intricate network structures.

A KAN layer is defined by a matrix Φ composed of univariate functions $\{\phi_{q,p}(\cdot)\}$ with $p = 1, \ldots, N_{in}$ and $q = 1, \ldots, N_{out}$, where N_{in} and N_{out} denote the number of inputs and the number of outputs, respectively, and $\phi_{q,p}$ indicates the trainable spline functions described above. This architectural innovation allows KANs to better capture complex, nonlinear relationships more effectively than traditional MLPs. Note that in the Kolmogorov–Arnold theorem, the inner functions constitute a KAN layer with $N_{in} = n$ and $N_{out} = 2n + 1$, while the external functions constitute a KAN layer with $N_{in} = 2n + 1$ and $N_{out} = 1$. To extend the capabilities of KANs, deeper network architectures have been developed. A deeper KAN is essentially a composition of multiple KAN layers, with enhanced ability to model more complex functions. A general KAN can be expressed by the composition *L* layers:

$$y = \text{KAN}(x) = (\Phi_L \circ \Phi_{L-1} \circ \cdots \circ \Phi_1 \circ \Phi_0)x \tag{14}$$

2.5. LSTM-KAN with a Dual-Stage Attention Mechanism

This study introduces a hybrid model that merges LSTM and KAN, leveraging the ability of LSTM to process sequential data and the ability of KANs in capturing complex nonlinear relationships. During the initialization of the model, the LSTM layer is established, and the KAN model is innovatively incorporated as a subsequent processing layer for the output of the LSTM, replacing the traditional fully connected layer. In the forward propagation function, the final hidden state of the LSTM serves as the input to the KAN layer. Additionally, both a factor attention mechanism and a temporal attention mechanism (i.e., a dual-stage attention mechanism) are introduced into the LSTM network, where the LSTM serves as both encoder and decoder [39]. The factor attention mechanism is devised to adaptively identify the most influential factors at each time step. Furthermore, the temporal attention mechanism accurately extracts crucial time segments by identifying the relevant hidden states across all time steps [54]. The proposed LSTM-KAN network, featuring a dual-stage attention mechanism, is named DA-LSTM-KAN and is specifically designed for predicting concrete dam displacement. Figure 3 illustrates the overall architecture of the DA-LSTM-KAN model.
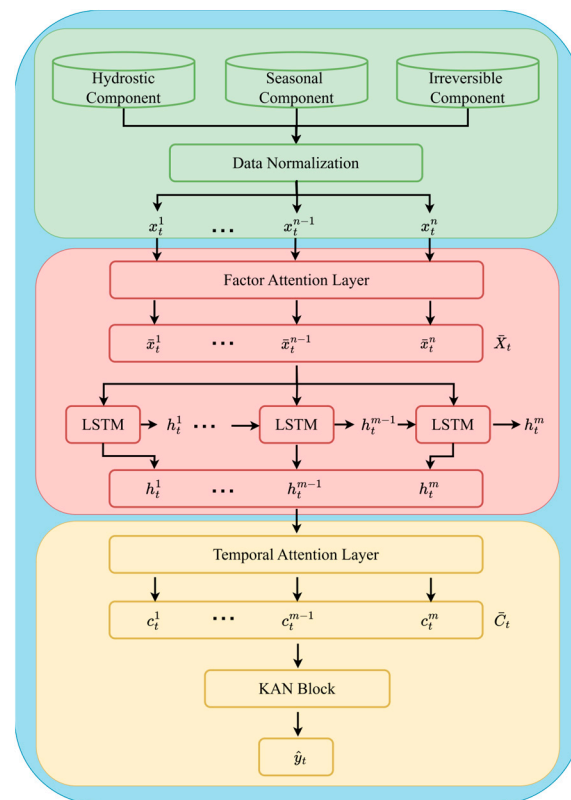

Figure 3. Architecture of the proposed DA-LSTM-KAN model.

The detailed procedure of the proposed method is outlined as follows:

Step 1: Identify the input variable x and the output variable y based on the HST model.
Step 2: Normalize the data to a range between 0 and 1, then split the data into training and test sets in a 7:3 ratio.
Step 3: Configure the DA-LSTM-KAN model parameters. The root-mean-square error (RMSE) is selected as the loss function.
Step 4: Train the model using the training dataset and optimize the parameters until the loss value converges.
Step 5: Evaluate the robustness of the trained model using the test dataset.
Step 6: Apply the finalized model for predicting the displacement of the concrete dam.

To evaluate the prediction performance, four evaluation metrics, including maximum absolute error ($AE_{\max}$), root-mean-square error ($RMSE$), mean absolute error (AE_{mean}), and forecast qualification rate (QR), were employed. The following formulas define the evaluation metrics:

$$AE_{\max} = \max(y_D(i) - y(i)), i = 1, 2, \ldots, N \tag{15}$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (y_D(i) - y(i))^2} \tag{16}$$

$$AE_{\mathrm{mean}} = \frac{1}{N} \sum_{i=1}^{N} |y_D(i) - y(i)| \tag{17}$$

$$QR = \frac{m}{n} \times 100\% \tag{18}$$

where y denotes the monitored displacement, y_D represents the model prediction results, and N denotes the number of samples; m and n, respectively, represent the number of qualified predictions and the number of total samples, where a qualified prediction is defined as one whose error is less than $\pm 5\%$ of the actual value.

3. Case Study

The effectiveness of the proposed DA-LSTM-KAN was validated through its application to a concrete gravity dam located on the Minjiang River in Fujian Province, China. The dam consists of 21 dam sections, and the maximum dam height is 71.0 m. Horizontal displacement monitoring data along the river were selected and analyzed from a measurement point on the No. 12 dam section (Figure 4). Due to the lack of thermometer monitoring data, the input environmental variables included the reservoir water levels and the annual and semiannual harmonic variations recorded from the start of the monitoring to the survey date. The time series of the reservoir level is shown in Figure 5. The variation in horizontal displacement is presented in Figure 6. Note that the monitoring data of the measurement points selected in this case were relatively complete. However, when there are missing data or abnormal data, it becomes imperative to adopt cluster analysis [55], mode decomposition [56], transfer learning [57], and other data processing and reconstruction techniques prior to establishing the monitoring model.
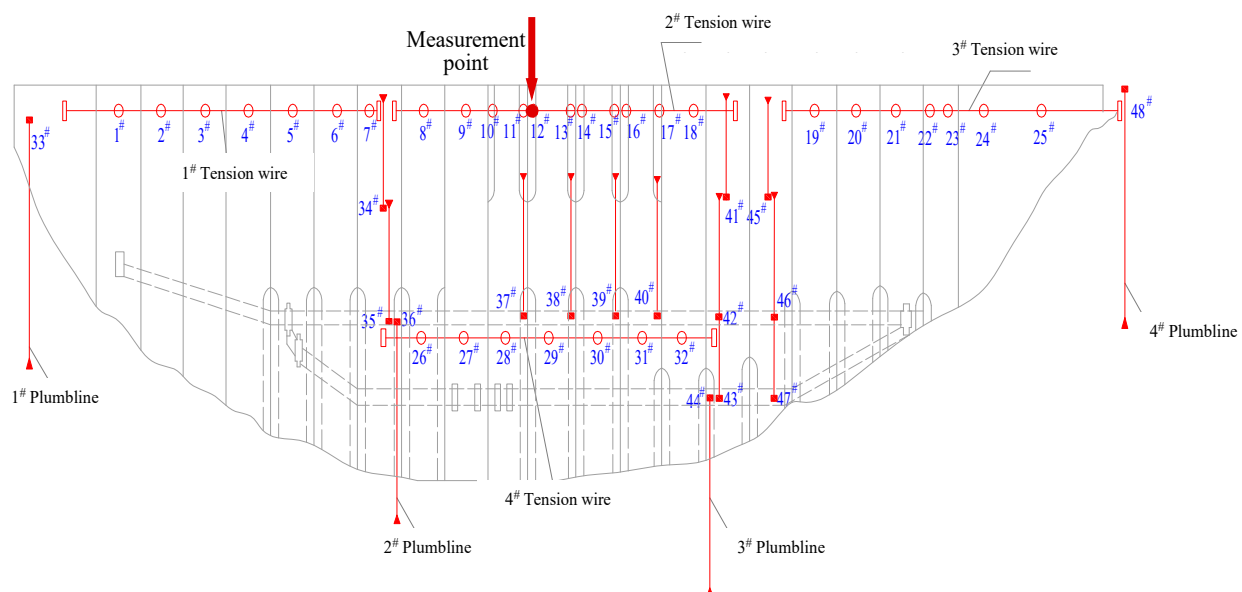

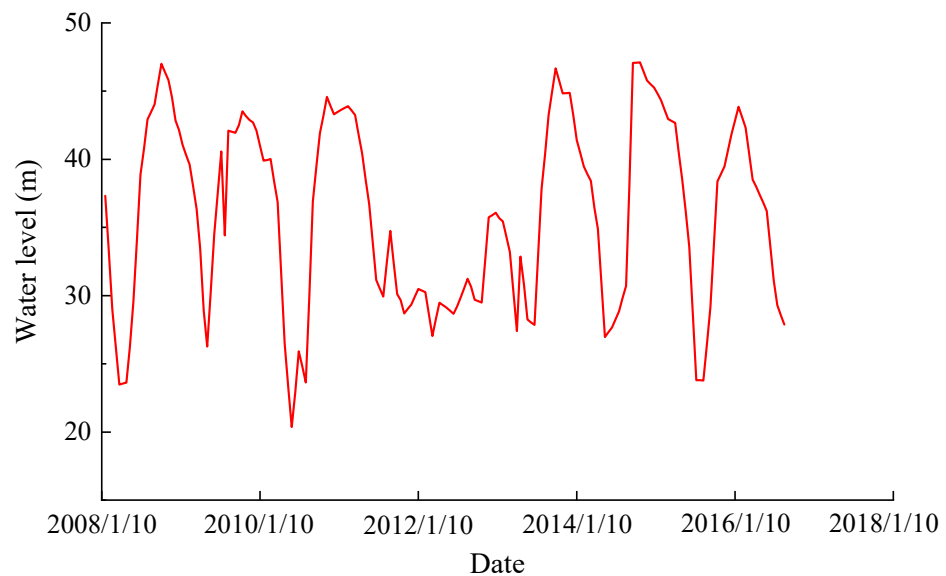**Figure 4.** The general view of the dam section and the selected measurement point.

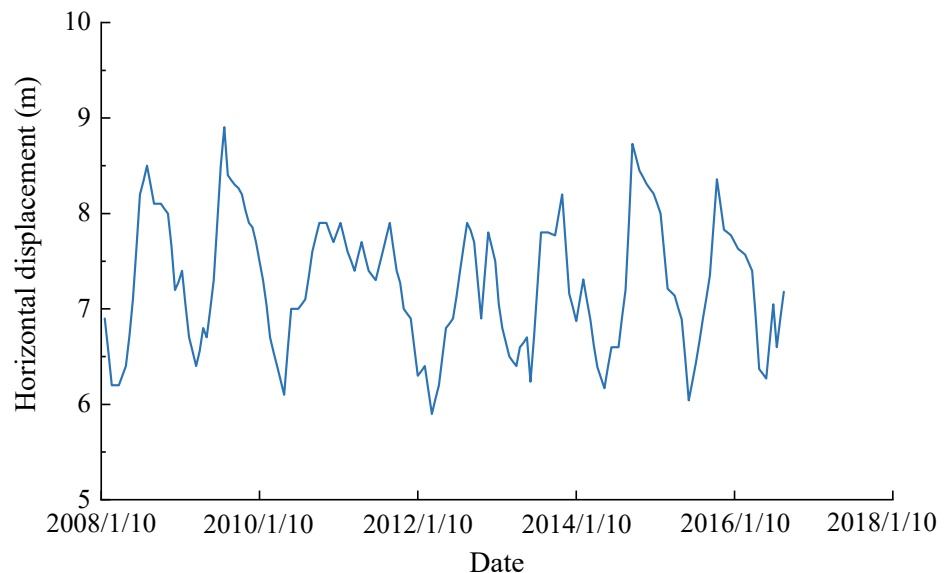**Figure 5.** Measured reservoir water level.

Figure 6. Measured horizontal displacements of the measurement point.

Furthermore, the performance of the proposed model was compared with those of multiple linear regression (MLR), LSTM, LSTM with only a factor attention mechanism (FA-LSTM), LSTM with only a temporal attention mechanism (TA-LSTM), and LSTM with a dual-stage attention mechanism (DA-LSTM), using the monitoring data from a period of 3019 days, from 10 January 2008 to 20 April 2016. All the models were trained with the same training set and verified with the same test set. The input variables for all models were selected as $x = \{x_1, x_2, \ldots, x_8, x_9\} = \{H^1, H^2, H^3, \sin(s), \cos(s), \sin2(s), \cos2(s), t, \ln t\}$, where $s = 2\pi t/365$.

4. Results and Discussion

4.1. Hyperparameter Adjustment

A careful adjustment of the parameter settings is crucial for training an optimal model. The learning rate, a crucial hyperparameter, dictates the speed at which the LSTM network learns. A learning rate that is too low will cause the model to converge very slowly, while a rate that is too high can lead to oscillations, preventing the model from reaching a stable solution. In this study, based on preliminary experiments, the initial learning was set at

0.01, and the number of hidden layers in LSTM was set at 2. Similarly, the hidden sizes of the encoder and decoder are critical for capturing the complexity of the data and were selected to range from 16 to 256, allowing for flexible model capacity. The batch size, which affects memory consumption and training stability, was set between 8 and 64. The time step value, impacting how the model processes sequential data, was set between 2 and 8 to balance computational efficiency and temporal resolution. Finally, 200 training epochs were implemented to ensure sufficient learning without excessive computation. Using the same procedure, the optimal hyperparameters for the other models were identified through a trial-and-error approach, as presented in Table 1.

In order to further quantify the influence of hyperparameters on the model output, Figure 7a–d show the relationships between the performance of the LSTM model and epoch, batch size, time step, and hidden size. The hyperparameters used in this paper and those optimized by an automated method called OPTUNA [58] are marked in the figure. The results indicate that hyperparameter selection does have a certain impact on the model performance, and the hyperparameters in Table 1 closely align with those obtained through OPTUNA optimization, leading to favorable model outcomes. However, as hyperparameter optimization is not the focus of this paper, it will not be discussed further.
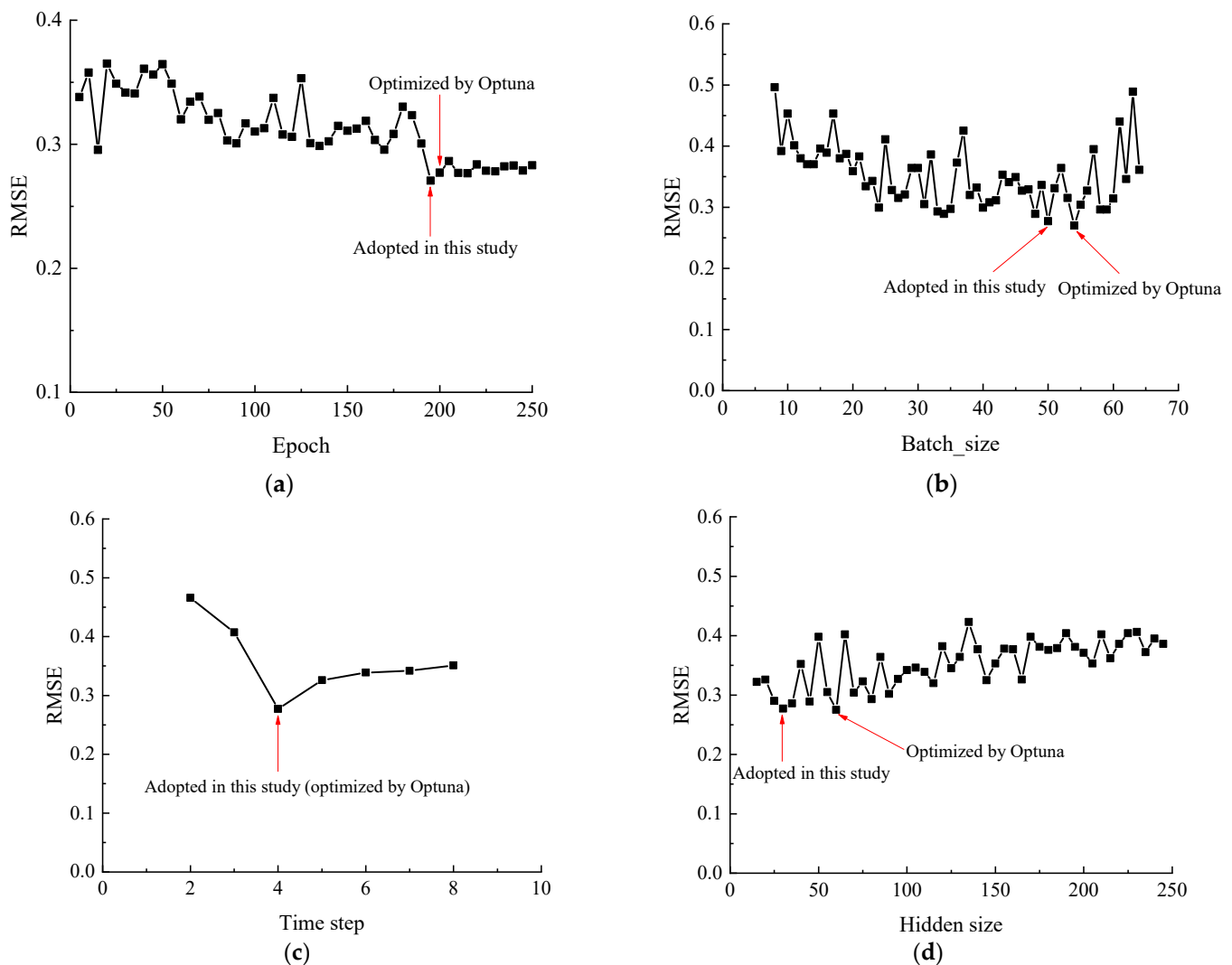
Figure 7. Relationships between the performance of the LSTM model and (**a**) epoch, (**b**) batch_size, (**c**) time step, and (**d**) hidden size.

Table 1. Basic information of model hyperparameter.

Model	Hyperparameter Values
LSTM	epoch = 200, batch_size = 50, time step = 4, hidden size = 30
FA-LSTM	epoch = 200, batch_size = 45, time step = 7, attention dimension = 9, hidden size = 210, dropout rate for attention = 0.1
TA-LSTM	epoch = 200, batch_size = 64, time step = 3, hidden size = 150, attention dimension = 150, dropout rate for attention = 0.1
DA-LSTM	epoch = 200, batch_size = 24, time step = 4, factor attention dimension = 9, hidden size = 130, time attention dimension = 130, dropout rate for attention = 0.1
DA-LSTM-KAN	epoch = 200, batch_size = 32, time step = 5, factor attention dimension = 9, hidden size = 100, time attention dimension = 100, dropout rate for attention = 0.1, number of hidden layers in KAN = 3

4.2. Performance Comparison

This section provides a comprehensive overview of the prediction results, aiming to assess and compare the performance of baseline models with that of the proposed model. To visually demonstrate predictive performance, Figure 8 presents the predicted displacements from different models alongside the measured displacements from the training and test sets at the measurement point. The figure shows that the DA-LSTM-KAN model predicted the measured data better than other models. The prediction accuracies on the training and test sets of all models are shown in Table 2. The table indicates that the proposed model outperformed all comparative models across all four performance metrics. Among the models evaluated, the MLR model showed the weakest performance, while the LSTM-based model with attention achieved better results than the model that relied solely on LSTM. Compared to DA-LSTM, the elevated predictive power of DA-LSTM-KAN is potentially attributed to the replacement of the fully connected layer processing the LSTM output with the KAN. In order to further verify the robustness of the proposed model, we split the data into training and test sets in an 8:2 ratio and compared all models again. The results, shown in Table 3, indicated that the DA-LSTM-KAN model performed well across all four metrics compared to the other models. In the subsequent analyses, we reverted to the original data split in a 7:3 ratio for consistency.
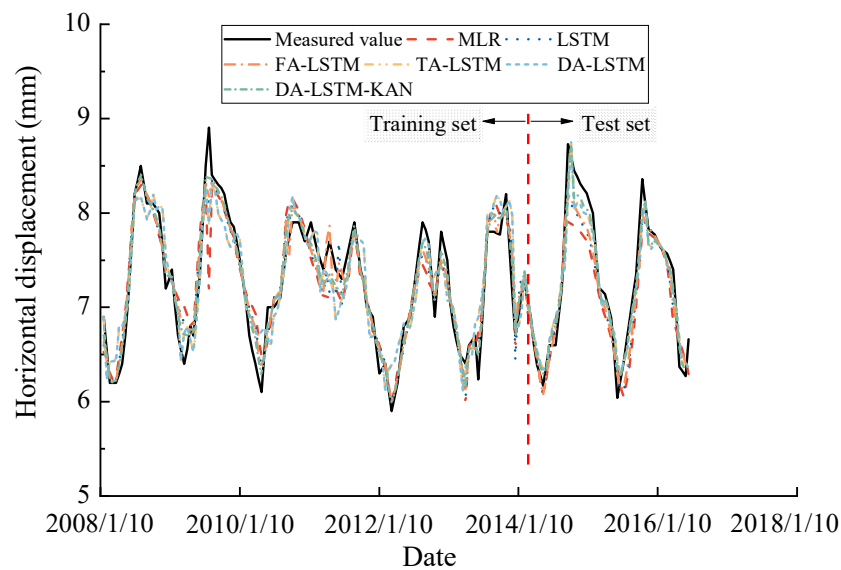

Figure 8. Prediction results of different models.

Table 2. Evaluation metrics of different models (7:3 ratio).

Model	Training Set				Test Set			
	AE_{max}	*RMSE*	AE_{mean}	*QR* (%)	AE_{max}	*RMSE*	AE_{mean}	*QR* (%)
MLR	0.3547	0.2558	0.2017	87.59	0.1277	0.3458	0.2749	61.40
LSTM	0.3015	0.2175	0.1714	91.24	0.1022	0.2766	0.2200	80.70
FA-LATM	0.2713	0.1957	0.1543	96.35	0.0920	0.2490	0.1980	89.47
TA-LATM	0.2442	0.1761	0.1389	96.35	0.0900	0.2223	0.1831	91.23
DA-LATM	0.2198	0.1585	0.1250	97.81	0.0318	0.2023	0.1611	89.47
DA-LSTM-KAN	0.1954	0.1409	0.1111	99.27	0.0255	0.1618	0.1289	100

Table 3. Evaluation metrics of different models (8:2 ratio).

Model	Training Set				Test Set			
	AE_{max}	*RMSE*	AE_{mean}	*QR* (%)	AE_{max}	*RMSE*	AE_{mean}	*QR* (%)
MLR	0.1715	0.2161	0.2146	93.51	0.2312	0.2017	0.2009	89.47
LSTM	0.3933	0.1967	0.1532	96.75	0.2089	0.1824	0.1817	97.37
FA-LATM	0.1454	0.1803	0.1792	100	0.1689	0.1646	0.1633	100
TA-LATM	0.1715	0.1792	0.1783	100	0.1761	0.1658	0.1647	100
DA-LATM	0.1603	0.1656	0.1647	100	0.1723	0.1609	0.1603	100
DA-LSTM-KAN	0.1254	0.1524	0.1516	100	0.1518	0.1453	0.1446	100

To gain deeper insights into the method's effectiveness, residual box plots were utilized to compare the predictive performance across different models. The results of these residual box plots are displayed in Figure 9. It is evident from the figure that the proposed method exhibited significantly lower volatility in deformation prediction compared to the other comparison methods. Meanwhile, the residuals were more concentrated for DA-LSTM-KAN, indicating superior deformation prediction performance. Additionally, the model incorporating solely the factor attention mechanism (i.e., FA-LSTM) exhibited a slightly lower predictive performance compared to the models utilizing only the temporal attention mechanism (i.e., TA-LSTM) as well as the models integrating both temporal and factor attention mechanisms (i.e., DA-LSTM).
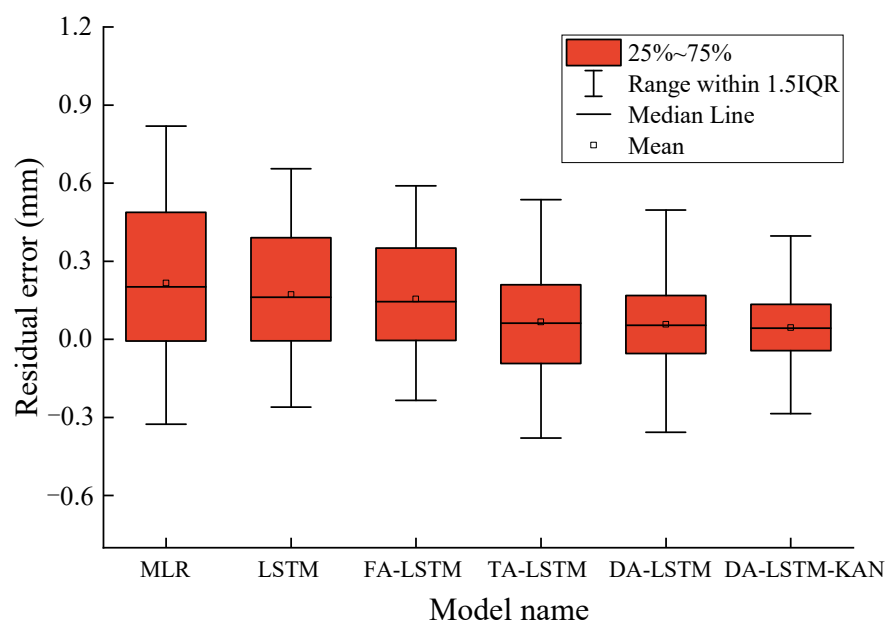

Figure 9. Box plot of residuals for different models.

4.3. Interpretability Analysis

The attention mechanism automatically enables a model to focus on the most significant factors along the time dimension. Figure 10 is the visualization of the temporal attention weights of each factor of the DA-LSTM-KAN model. Among the nine input factors, the weights of the hydrostatic and seasonal factors were relatively larger, while the weights of irreversible factors were relatively smaller. Figure 11 shows the average attention weight of each factor derived from Figure 10. SHAP, introduced by Lundberg and Lee [59], offers a unified approach to interpreting machine learning models. This framework interprets each feature as a contributor to the model, quantifying its average marginal contribution to ascertain its influence on the model's output. In order to further validate the attention mechanism's ability to enhance model interpretability, we conducted a feature importance analysis of the LSTM black box model using SHAP, and the results are shown in Figure 12. From Figures 11 and 12, it becomes evident that despite some discrepancies, the feature importance rankings derived from both methods exhibit a significant degree of similarity. In particular, two of the first three important features involve two hydrostatic components (i.e., $x_1 = H$ and $x_2 = H^2$), while at least one irreversible component is included in the least significant three variables.
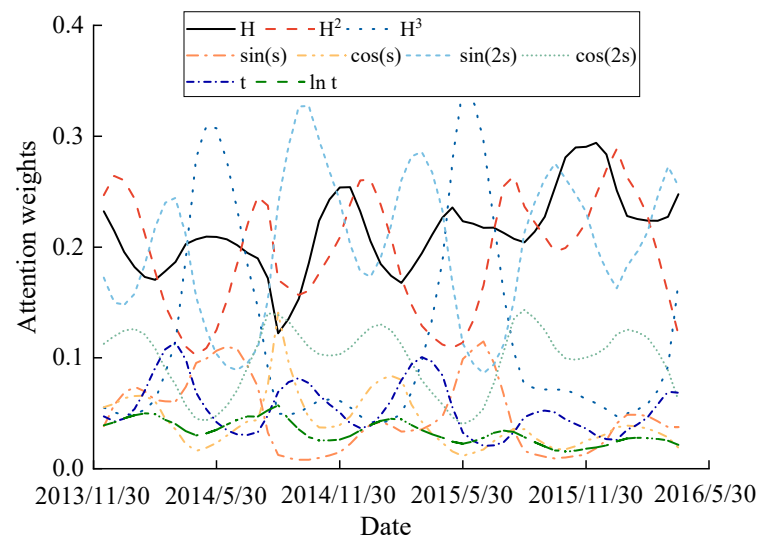

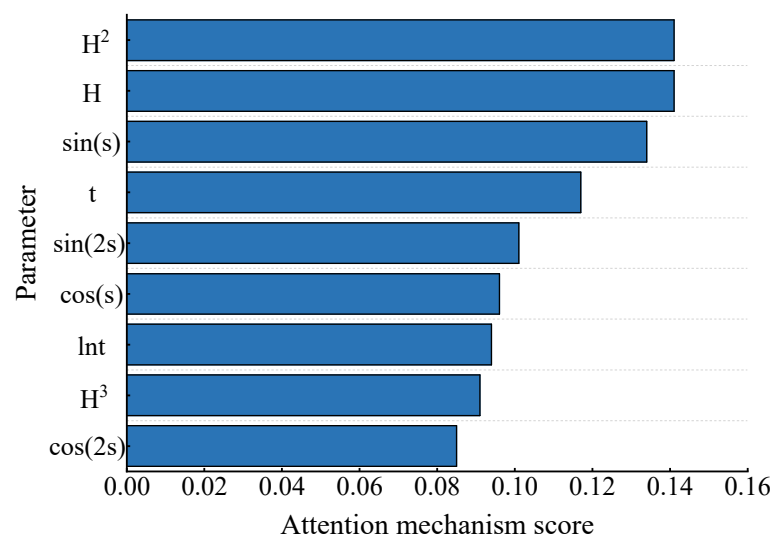**Figure 10.** Variation in attention weight of each factor with time.

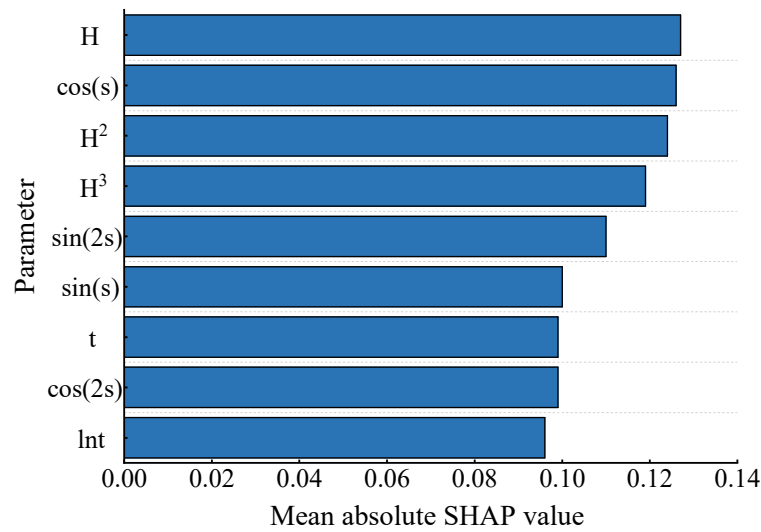**Figure 11.** Average attention weight of each factor.

Figure 12. Importance value of each factor of the LSTM model.

Figure 13a,b show the average attention weights (or average importance value) of the hydrostatic, seasonal, and time-dependent irreversible components for DA-LSTM-KAN and LSTM-SHAP, respectively. It is encouraging to note that the quantitative results for three components obtained through the two methods exhibit a high degree of consistency. Specifically, the figure indicates that the seasonal component is the largest contributor to displacement, with a contribution of 41.6% for DA-LSTM-KAN and 43.5% for LSTM-SHAP. The hydrostatic component follows, accounting for 37.3% of the displacement in DA-LSTM-KAN and 37.0% in LSTM-SHAP. Lastly, the irreversible component contributes 21.1% of the displacement for DA-LSTM-KAN and 19.5% for LSTM-SHAP. Multiple linear regression (MLR) models often stand out as the most straightforward approach, offering inherent interpretability. For a more comprehensive comparison, Figure 14 displays the process lines of each component at the measurement point using the traditional MLR model, while Figure 15 illustrates the relative importance of each component as determined by the relative weight method [60]. In this method, the relative weight of each variable represents the proportion of the predictable variance it accounts for. As shown in Figures 14 and 15, the importance ranking of the three components derived from MLR aligns with that identified by DA-LSTM-KAN. Furthermore, the attention and relative weight techniques offer a more intuitive evaluation than the process line method, particularly when two components have similar levels of importance.
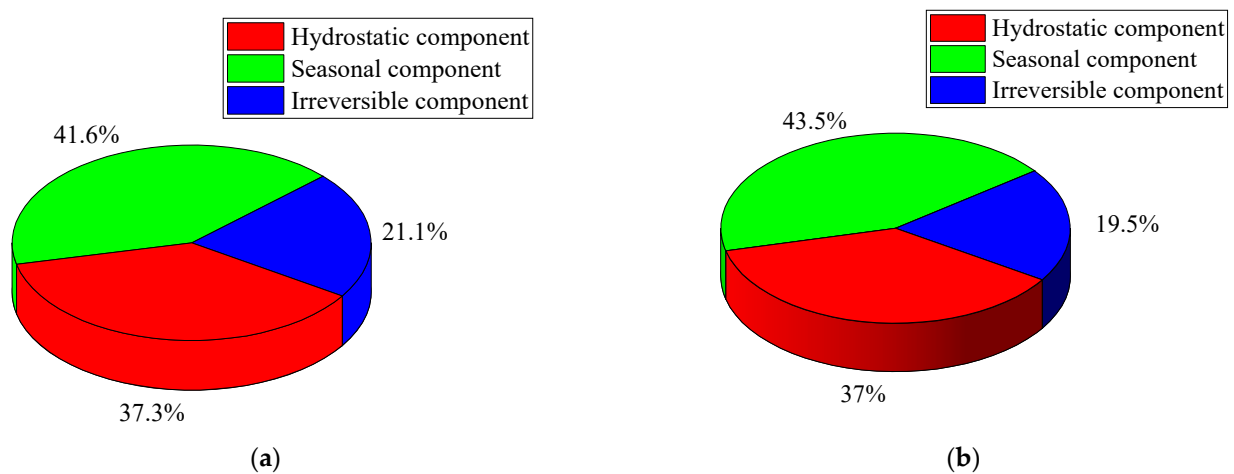

(**a**)

(**b**)

Figure 13. Feature importance ranking of the three components using (**a**) the attention mechanism approach and (**b**) the SHAP approach.
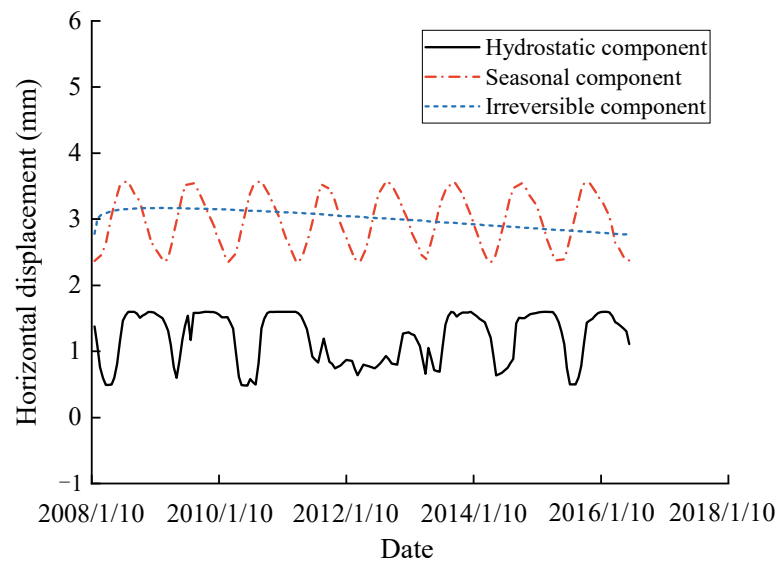
Figure 14. The process line of each component of horizontal displacement.
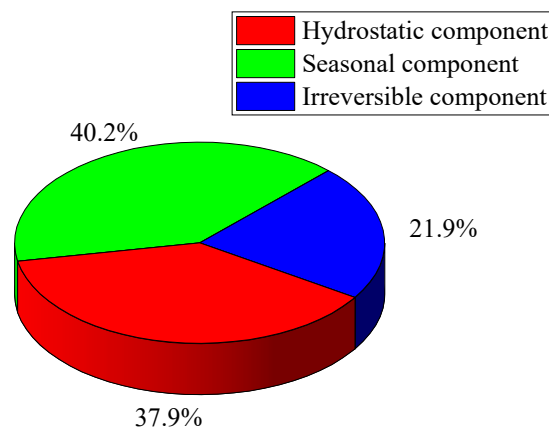

Figure 15. Relative importance of each component determined by MLR.

5. Conclusions

This study developed the DA-LSTM-KAN model, a sophisticated hybrid predictive model for dam deformation that integrates the sequential data handling capabilities of LSTM with the complex nonlinear relationship modeling of KANs, enhanced by a dual-stage attention mechanism. The model was tested using prototype monitoring data from a concrete gravity dam constructed on the Minjiang River in China. The results revealed that the DA-LSTM-KAN model demonstrated superior predictive performance compared to other models such as multiple linear regression and various LSTM variants, and the LSTM-based architecture integrated with an attention mechanism outperformed the model that relies exclusively on LSTM. Moreover, the interpretability analysis of DA-LSTM-KAN revealed that for the investigated measurement point, the seasonal component was the most significant contributor to dam displacement, followed by the hydrostatic and irreversible components. This finding is qualitatively consistent with the results obtained from the SHAP method and the relative weight method, highlighting the potential benefits of applying deep learning models in dam deformation monitoring.

The hybrid model presented in this study is designed to enhance both the prediction accuracy and the interpretability of the model concurrently. The results revealed that it effectively discovers and interprets the time-varying trends in the deformation sequence, overcoming the limitations of information mining inherent in single models. Future research will focus on a more comprehensive comparative analysis between the hybrid model proposed in this study and the existing models when they are used in more practical

projects. Additionally, after thoroughly evaluating the model complexity and computational efficiency with large datasets, achieving greater scalability in real-time monitoring emerges as the subsequent challenge that needs to be tackled.

Author Contributions: Conceptualization, X.L. and Z.L.; methodology, R.X. and X.L.; software, R.X. and X.A.; validation, J.W. and X.A.; formal analysis, J.W. and H.H.; writing—original draft, R.X. and X.L. All authors have read and agreed to the published version of the manuscript.

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

Conflicts of Interest: The authors declare no conflicts of interest.

References

1.　Zhang, Y.; Zhong, W.; Li, Y.; Wen, L. A deep learning prediction model of DenseNet-LSTM for concrete gravity dam deformation based on feature selection. *Eng. Struct.* **2023**, *295*, 116827. [CrossRef]
2.　Wei, B.; Chen, L.; Li, H.; Yuan, D.; Wang, G. Optimized prediction model for concrete dam displacement based on signal residual amendment. *Appl. Math. Model.* **2020**, *78*, 20–36. [CrossRef]
3.　Mata, J.; Tavares de Castro, A.; Sá da Costa, J. Constructing statistical models for arch dam deformation. *Struct. Control Health Monit.* **2014**, *21*, 423–437. [CrossRef]
4.　Wang, S.; Gu, C.; Liu, Y.; Gu, H.; Xu, B.; Wu, B. Displacement observation data-based structural health monitoring of concrete dams: A state-of-art review. *Structures* **2024**, *68*, 107072. [CrossRef]
5.　Liu, X.; Li, Z.; Sun, L.; Khailah, E.Y.; Wang, J.; Lu, W. A critical review of statistical model of dam monitoring data. *J. Build. Eng.* **2023**, *80*, 108106. [CrossRef]
6.　Penot, I.; Daumas, B.; Fabre, J.P. Monitoring behaviour. *Int. Water Power Dam Constr.* **2005**, *57*, 24–27.
7.　Léger, P.; Leclerc, M. Hydrostatic, temperature, time-displacement model for concrete dams. *J. Eng. Mech.* **2007**, *133*, 267–277. [CrossRef]
8.　Tatin, M.; Briffaut, M.; Dufour, F.; Simon, A.; Fabre, J.P. Statistical modelling of thermal displacements for concrete dams: Influence of water temperature profile and dam thickness profile. *Eng. Struct.* **2018**, *165*, 63–75. [CrossRef]
9.　Hu, J.; Ma, F. Statistical modelling for high arch dam deformation during the initial impoundment period. *Struct. Control Health Monit.* **2020**, *27*, e2638. [CrossRef]
10.　Deng, N.; Wang, J.G.; Szostak-Chrzanowski, A. Dam deformation analysis using the partial least squares method. In Proceedings of the 13th FIG Int. Symp. on Deformation Measurements and Analysis & 4th IAG Symp. on Geodesy for Geotechnical and Structural Engineering, Lisbon, Portugal, 12–15 May 2008.
11.　Mata, J. Interpretation of concrete dam behaviour with artificial neural network and multiple linear regression models. *Eng. Struct.* **2011**, *33*, 903–910. [CrossRef]
12.　Kao, C.; Loh, C. Monitoring of long-term static deformation data of Fei-Tsui arch dam using artificial neural network-based approaches. *Struct. Control Health Monit.* **2013**, *20*, 282–303. [CrossRef]
13.　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]
14.　Li, X.; Wen, Z.; Su, H. An approach using random forest intelligent algorithm to construct a monitoring model for dam safety. *Eng. Comput.* **2021**, *37*, 39–56. [CrossRef]
15.　Liu, M.; Wen, Z.; Zhou, R.; Su, H. Bayesian optimization and ensemble learning algorithm combined method for deformation prediction of concrete dam. In *Structures*; Elsevier: Amsterdam, The Netherlands, 2023; Volume 54, pp. 981–993.
16.　Su, H.; Li, X.; Yang, B.; Wen, Z. Wavelet support vector machine-based prediction model of dam deformation. *Mech. Syst. Signal Process.* **2018**, *110*, 412–427. [CrossRef]
17.　Lin, C.; Li, T.; Chen, S.; Liu, X.; Lin, C.; Liang, S. Gaussian process regression-based forecasting model of dam deformation. *Neural Comput. Appl.* **2019**, *31*, 8503–8518. [CrossRef]
18.　Wang, S.; Xu, C.; Liu, Y.; Wu, B. A spatial association-coupled double objective support vector machine prediction model for diagnosing the deformation behaviour of high arch dams. *Struct. Health Monit.* **2022**, *21*, 945–964. [CrossRef]
19.　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]
20.　Zhang, Y.; Zhang, W.; Li, Y.; Wen, L.; Sun, X. AF-OS-ELM-MVE: A new online sequential extreme learning machine of dam safety monitoring model for structure deformation estimation. *Adv. Eng. Inform.* **2024**, *60*, 102345. [CrossRef]

21. Cai, Z.; Yu, J.; Chen, W.; Wang, J.; Wang, X.; Guo, H. Improved extreme learning machine-based dam deformation prediction considering the physical and hysteresis characteristics of the deformation sequence. *J. Civ. Struct. Health Monit.* **2022**, *12*, 1173–1190. [CrossRef]

22. Zhang, H.; Xu, S.F. Multi-scale dam deformation prediction based on empirical mode decomposition and genetic algorithm for support vector machines (GA-SVM). *Chin. J. Rock Mech. Eng.* **2011**, *30* (Suppl. S2), 3681–3688.

23. Chen, Y.; Gu, C.; Shao, C.; Gu, H.; Zheng, D.; Wu, Z.; Fu, X. An approach using adaptive weighted least squares support vector machines coupled with modified ant lion optimizer for dam deformation prediction. *Math. Probl. Eng.* **2020**, *2020*, 9434065. [CrossRef]

24. Su, H.; Wen, Z.; Sun, X.; Li, H. Rough set-support vector machine-based real-time monitoring model of safety status during dangerous dam reinforcement. *Int. J. Damage Mech.* **2017**, *26*, 501–522. [CrossRef]

25. Dai, B.; Gu, H.; Zhu, Y.; Chen, S.; Rodriguez, E.F. On the Use of an Improved Artificial Fish Swarm Algorithm-Backpropagation Neural Network for Predicting Dam Deformation Behavior. *Complexity* **2020**, *2020*, 5463893. [CrossRef]

26. 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]

27. Li, M.; Shen, Y.; Ren, Q.; Li, H. A new distributed time series evolution prediction model for dam deformation based on constituent elements. *Adv. Eng. Inform.* **2019**, *39*, 41–52. [CrossRef]

28. Wen, Z.; Zhou, R.; Su, H. MR and stacked GRUs neural network combined model and its application for deformation prediction of concrete dam. *Expert Syst. Appl.* **2022**, *201*, 117272. [CrossRef]

29. Sun, S.; Wang, S.; Wei, Y. A new ensemble deep learning approach for exchange rates forecasting and trading. *Adv. Eng. Inform.* **2020**, *46*, 101160. [CrossRef]

30. Xiang, S.; Qin, Y.; Zhu, C.; Wang, Y.; Chen, H. Long short-term memory neural network with weight amplification and its application into gear remaining useful life prediction. *Eng. Appl. Artif. Intell.* **2020**, *91*, 103587. [CrossRef]

31. Qu, X.; Yang, J.; Chang, M. A Deep Learning Model for Concrete Dam Deformation Prediction Based on RS-LSTM. *J. Sens.* **2019**, *2019*, 4581672. [CrossRef]

32. 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]

33. Yang, S.; Han, X.; Kuang, C.; Fang, W.; Zhang, J.; Yu, T. Comparative Study on Deformation Prediction Models of Wuqiangxi Concrete Gravity Dam Based on Monitoring Data. *CMES-Comput. Model. Eng. Sci.* **2022**, *131*, 49–72. [CrossRef]

34. Tian, K.; Yang, J.; Cheng, L. Deep learning model for the deformation prediction of concrete dams under multistep and multifeature inputs based on an improved autoformer. *Eng. Appl. Artif. Intell.* **2024**, *137*, 109109. [CrossRef]

35. Molnar, C.; Casalicchio, G.; Bischl, B. Interpretable machine learning—A brief history, state-of-the-art and challenges. In *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*; Springer International Publishing: Cham, Switzerland, 2020; pp. 417–431.

36. Zhu, Y.; Gao, Y.; Wang, Z.; Cao, G.; Wang, R.; Lu, S.; Li, W.; Nie, W.; Zhang, Z. A tailings dam long-term deformation prediction method based on empirical mode decomposition and LSTM model combined with attention mechanism. *Water* **2022**, *14*, 1229. [CrossRef]

37. Yang, D.; Gu, C.; Zhu, Y.; Dai, B.; Zhang, K.; Zhang, Z.; Li, B. A concrete dam deformation prediction method based on LSTM with attention mechanism. *IEEE Access* **2020**, *8*, 185177–185186. [CrossRef]

38. Shu, X.; Bao, T.; Li, Y.; Gong, J.; Zhang, K. VAE-TALSTM: A temporal attention and variational autoencoder-based long short-term memory framework for dam displacement prediction. *Eng. Comput.* **2022**, *38*, 3497–3512. [CrossRef]

39. Ren, Q.; Li, M.; Li, H.; Shen, Y. A novel deep learning prediction model for concrete dam displacements using interpretable mixed attention mechanism. *Adv. Eng. Inform.* **2021**, *50*, 101407. [CrossRef]

40. 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]

41. Jas, K.; Dodagoudar, G.R. Explainable machine learning model for liquefaction potential assessment of soils using XGBoost-SHAP. *Soil Dyn. Earthq. Eng.* **2023**, *165*, 107662. [CrossRef]

42. Goudjil, K.; Boukhatem, G.; Boulifa, R.; Bekkouche, S.R.; Djihane, D. Prediction and interpretation of limit pressure of clayey soils using ensemble machine learning methods and shapely additive explanations. *Stud. Eng. Exact Sci.* **2024**, *5*, e5567. [CrossRef]

43. Somala, S.N.; Chanda, S.; AlHamaydeh, M.; Mangalathu, S. Explainable XGBoost–SHAP Machine-Learning Model for Prediction of Ground Motion Duration in New Zealand. *Nat. Hazards Rev.* **2024**, *25*, 04024005. [CrossRef]

44. Matin, S.S.; Pradhan, B. Earthquake-induced building-damage mapping using Explainable AI (XAI). *Sensors* **2021**, *21*, 4489. [CrossRef] [PubMed]

45. Feng, D.C.; Wang, W.J.; Mangalathu, S.; Taciroglu, E. Interpretable XGBoost-SHAP machine-learning model for shear strength prediction of squat RC walls. *J. Struct. Eng.* **2021**, *147*, 04021173. [CrossRef]

46. Barkhordari, M.S.; Fattahi, H.; Armaghani, D.J.; Khan, N.M.; Afrazi, M.; Asteris, P.G. Predictive Failure Mode Identification in Reinforced Concrete Flat Slabs Using Advanced Ensemble Neural Networks. *Preprint* **2024**. [CrossRef]

47. Li, B.; Ning, J.; Yang, S.; Zhang, L. Prediction model for high arch dam stress during the operation period using LightGBM with MSSA and SHAP. *Adv. Eng. Softw.* **2024**, *192*, 103635. [CrossRef]

48. Liu, Z.; Wang, Y.; Vaidya, S.; Ruehle, F.; Halverson, J.; Soljačić, M.; Hou, T.; Tegmark, M. Kan: Kolmogorov-arnold networks. *arXiv* **2024**, arXiv:2404.19756.

49. Kolmogorov, A.N. On the representation of continuous functions of many variables by superposition of continuous functions of one variable and addition. In *Doklady Akademii Nauk*; Russian Academy of Sciences: Moscow, Russia, 1957; Volume 114, pp. 953–956.

50. Xu, K.; Chen, L.; Wang, S. Kolmogorov-Arnold Networks for Time Series: Bridging Predictive Power and Interpretability. *arXiv* **2024**, arXiv:2406.02496.

51. Hassan, M.M. Bayesian Kolmogorov Arnold Networks (Bayesian_KANs): A Probabilistic Approach to Enhance Accuracy and Interpretability. *arXiv* **2024**, arXiv:2408.02706.

52. Hochreiter, S. *Long Short-Term Memory*; Neural Computation MIT-Press: Cambridge, MA, USA, 1997.

53. Treisman, A.M.; Gelade, G. A feature-integration theory of attention. *Cogn. Psychol.* **1980**, *12*, 97–136. [CrossRef]

54. Huang, B.; Kang, F.; Li, J.; Wang, F. Displacement prediction model for high arch dams using long short-term memory based encoder-decoder with dual-stage attention considering measured dam temperature. *Eng. Struct.* **2023**, *280*, 115686. [CrossRef]

55. Su, Y.; Weng, K.; Lin, C.; Chen, Z. Dam deformation interpretation and prediction based on a long short-term memory model coupled with an attention mechanism. *Appl. Sci.* **2021**, *11*, 6625. [CrossRef]

56. Song, J.; Yang, Z.; Li, X. Missing data imputation model for dam health monitoring based on mode decomposition and deep learning. *J. Civ. Struct. Health Monit.* **2024**, *14*, 1111–1124. [CrossRef]

57. Li, Y.; Bao, T.; Chen, H.; Zhang, K.; Shu, X.; Chen, Z.; Hu, Y. A large-scale sensor missing data imputation framework for dams using deep learning and transfer learning strategy. *Measurement* **2021**, *178*, 109377. [CrossRef]

58. Akiba, T.; Sano, S.; Yanase, T.; Ohta, T.; Koyama, M. Optuna: A next-generation hyperparameter optimization framework. In Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, Anchorage, AK, USA, 4–8 August 2019; pp. 2623–2631.

59. Lundberg, S.M.; Lee, S.I. A Unified Approach to Interpreting Model Predictions. In Proceedings of the 31st International Conference on Neural Information Processing Systems, Long Beach, CA, USA, 4–9 December 2017; pp. 4768–4777.

60. Johnson, J.W. A heuristic method for estimating the relative weight of predictor variables in multiple regression. *Multivar. Behav. Res.* **2000**, *35*, 1–19. [CrossRef]