




## Article

# Monitoring Data Fusion Model for Subsoil Layer Deformation Prediction

Huiguo Wu <sup>1,2,3</sup> , Yuedong Wu <sup>1,2,3</sup>, Jian Liu <sup>2,3,4,\*</sup> , Lei Zhang <sup>2,3</sup>, Yongyang Zhu <sup>2,3</sup> and Chuanyang Liang <sup>5</sup> 

<sup>1</sup> College of Artificial Intelligence and Automation, Hohai University, Changzhou 213000, China; hhuwhg@hhu.edu.cn (H.W.); hhuwyd@163.com (Y.W.)

<sup>2</sup> Key Laboratory of Ministry of Education for Geomechanics and Embankment Engineering, Hohai University, Nanjing 210098, China; 210404010010@hhu.edu.cn (L.Z.); 210404010011@hhu.edu.cn (Y.Z.)

<sup>3</sup> Geotechnical Engineering Research Center of Jiangsu Province, Nanjing 210098, China

<sup>4</sup> Engineering Research Center of Dredging Technology of Ministry of Education, Hohai University, Changzhou 213000, China

<sup>5</sup> School of Civil Engineering and Architecture, Anhui University of Technology, Ma'anshan 243002, China; 15950500017@163.com

\* Correspondence: 20170053@hhu.edu.cn; Tel.: +86-1836-296-0323

**Abstract:** Predicting soil deformation is critical for the success of building construction projects. The traditional methods used for this task, which rely on theoretical calculations and numerical simulations, require detailed information on soil characteristics and geological conditions. These essential details are often challenging to obtain in practical engineering, thereby limiting the accuracy of these methods in building construction contexts. Deep learning (DL) provides a direct approach for modeling soil deformation without having a detailed understanding of the soil properties and geological conditions. However, the existing DL algorithms mainly focus on modeling deformation directly. With advancements in monitoring technology, integrating diverse monitoring data has become crucial for accurately predicting deformation, a need often overlooked in current practices. This paper introduces a monitoring data fusion (MDF) model aimed at enhancing the utilization efficiency of diverse monitoring data. Validated against real-world engineering scenarios, this model significantly outperforms traditional single-feature and multi-feature long short-term memory (LSTM) models. It achieves a mean absolute percentage error (MAPE) of approximately 2.12%, representing reductions of 30% and 63%, and a root mean square error (RMSE) of around 12.5 mm, with reductions of 36% and 77%. Additionally, the DL interpretability method, Shapley additive explanations (SHAP), is utilized to elucidate how various model features contribute to generating predictions.

**Keywords:** soil deformation; deep learning; monitoring data; LSTM; SHAP



**Citation:** Wu, H.; Wu, Y.; Liu, J.; Zhang, L.; Zhu, Y.; Liang, C.

Monitoring Data Fusion Model for Subsoil Layer Deformation Prediction.

*Buildings* **2024**, *14*, 2055. <https://doi.org/10.3390/buildings14072055>

Academic Editor: Eugeniusz Koda

Received: 3 June 2024

Revised: 30 June 2024

Accepted: 2 July 2024

Published: 5 July 2024



**Copyright:** © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Soil deformation predictions are crucial in geotechnical engineering to ensure the stability and safety of buildings [1]. The traditional methods used for this task have relied on theoretical calculations and numerical simulations to predict soil behavior under various loading conditions [2,3]. Theoretical calculations typically involve Darcy's law and classical one-dimensional consolidation theory [4,5], employing the soil parameters derived from geotechnical tests and then calculating deformations through analytical solutions. However, these methods incorporate numerous assumptions and uncertainties. In numerical simulations, soil deformation is modeled through finite element software, necessitating detailed knowledge of the soil properties and geological conditions, which are often challenging to ascertain with accuracy [6].

The emergence of machine learning (ML) has significantly transformed soil deformation prediction techniques. ML has the capability to directly learn deformation patterns from data, providing a new avenue for addressing the limitations of traditional methods. ML techniques such as genetic algorithm (GA), support vector machines (SVM),

extreme gradient boosting (XGboost), random forest regression (RFR), and extreme learning machine (ELM) have demonstrated potential in simulating soil behavior [7–10]. These techniques have shown promise as they do not require extensive information on soil properties and geological conditions [11,12].

Despite the successes of ML in modeling soil behavior, traditional ML techniques struggle to capture the temporal dynamics inherent in soil deformation processes [13]. Deep learning (DL) models, particularly long short-term memory (LSTM) networks [14], have emerged as powerful solution for modeling time-series data. LSTM is adept at comprehending the temporal correlation between data points, making it particularly suitable for predicting time-varying soil behaviors such as landslide displacement [15,16], deformation during the excavation of foundation pits [17], and tunnel deformation [18].

A key limitation of earlier DL models was their focus on single-point deformation data, so they overlooked the spatial variability in soil properties and deformation patterns. In practice, monitoring data from multiple points are usually available, and prediction accuracy can be improved by utilizing multipoint monitoring data. Recent developments have focused on this, and authors have proposed integrated models that utilize multipoint monitoring data. For instance, CNN–LSTM hybrid models leverage both convolutional neural network (CNN) and LSTM to analyze data from multiple monitoring points, offering a comprehensive view of soil behavior over time and space [19]. Song et al. further extended this approach with a multipoint recurrent convolutional neural network (RCNN) model, specifically designed to be trained on monitoring data to achieve accurate soil deformation predictions in deep foundations [20].

While DL models are valuable in various engineering applications, they often focus primarily on deformation monitoring data [21]. The rise in advanced monitoring techniques has enabled the collection of multidimensional monitoring data, such as stress data and pore water pressure data [22]. The existing DL methods typically do not use multiple types of monitoring data, thereby limiting further improvements in model accuracy. Moreover, traditional modeling methods usually adopt a standard train–test split (assigning 80% of the initial data as the training set and the remaining 20% as the test set) [20,23]. However, soil deformation data often lack periodicity, and the volume of such data is usually relatively small. In many cases, the deformation observed during the testing period tends to stabilize. Consequently, test sets divided in this manner may not reflect the distribution of the data or adequately demonstrate the model’s performance.

This paper introduces a monitoring data fusion (MDF) model, which was designed to integrate the features from multiple types of monitoring data, aiming to enhance the accuracy of soil deformation predictions. Employing a dynamic training–testing approach, this model adapts to the limited and dynamically updated nature of geotechnical monitoring data. To verify the reliability of the proposed model, an engineering case containing 150 days of monitoring data was used to validate the model. Additionally, the importance of different features in the proposed model was evaluated using the Shapley additive explanations (SHAP) method, which provided insights into how different features contribute to the model’s predictions.

## 2. Methods

### 2.1. Problem Statement

In building construction, monitoring data are crucial for guiding project decisions. For a specific construction project, diverse monitoring data can be obtained. In this study, three monitoring features—total stress, effective stress, and deformation—were selected to predict deformation. The aim was to amalgamate features from multiple data sources, thereby enhancing the model’s predictive accuracy.

The relationship between deformation and its monitoring sequence is dynamic and complex. Additionally, the monitoring data obtained in civil engineering are often limited in volume. As a result, it is imperative for the model to be regularly updated with new data to adeptly adjust to these changes.

### 2.2. Monitoring Data Fusion Model

This paper proposes a deformation prediction model that is based on a DL model. The model excels in extracting features individually from monitoring data. The fundamental framework of the MDF model is illustrated in Figure 1. The MDF model was designed to enhance the application of DL techniques to monitoring data. We utilized all data collected during the construction period, including effective stress  $\sigma$ , total stress  $P$ , and deformation  $S$ . The DL model automatically extracts various features; thus, specific features are not selected: all monitoring data are directly input into the model for comprehensive training.

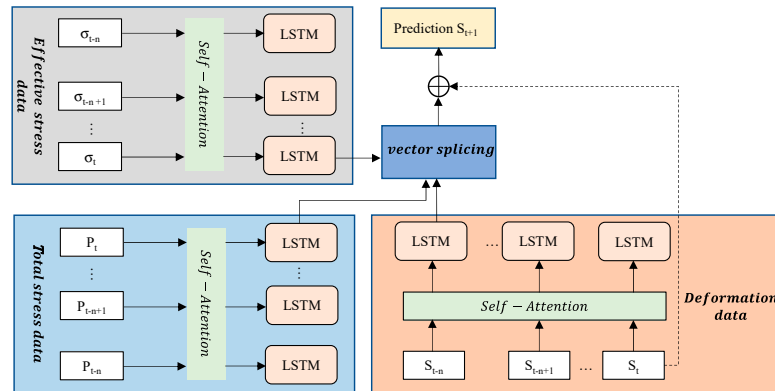


Figure 1. Structural diagram of MDF model.

LSTM networks are used as the feature extraction module for each type of monitoring datum. Prior to entering the LSTM networks, a self-attention mechanism processes the data, enabling the model to assess the significance of various data points. Subsequently, the extracted features are integrated using vector splicing techniques. Additionally, to counteract model forgetting, residual structures are incorporated, allowing the model to utilize deformation data from previous time steps.

### 2.3. Feature Extraction Module of MDF Model

LSTM models have become indispensable tools for analyzing soil deformation [24]. Figure 2 demonstrates the inputs and outputs of the LSTM units. As depicted, each LSTM unit processes a composite input, which includes three critical elements: the memory state from the preceding moment (denoted as  $C_{t-1}$ ), the current input data (represented as  $x_t$ ), and the previous output of the hidden layer (denoted as  $H_{t-1}$ ). The output of an LSTM unit comprises two principal components: the output of the hidden layer at the current time ( $H_t$ ) and the memory state at that moment ( $C_t$ ).

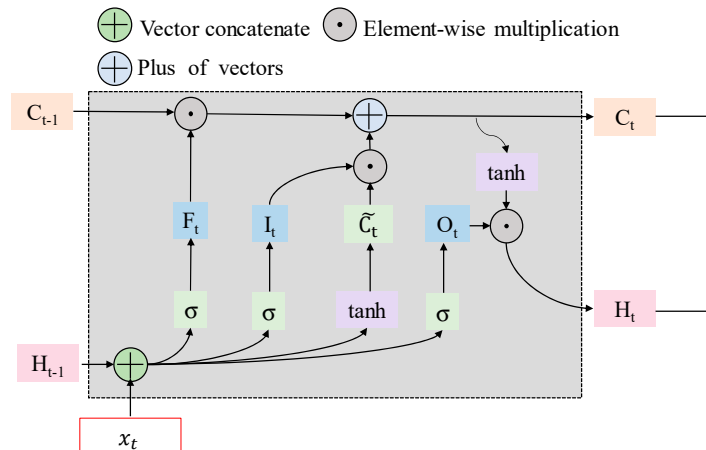


Figure 2. An LSTM unit structure.

Mathematically, the computation procedure of an LSTM unit is described through the following equations:

$$I_t = \sigma(W_i x_t + U_i H_{t-1} + b_i) \quad (1)$$

$$F_t = \sigma(W_f x_t + U_f H_{t-1} + b_f) \quad (2)$$

$$O_t = \sigma(W_o x_t + U_o H_{t-1} + b_o) \quad (3)$$

$$C_t = F_t \odot C_{t-1} + I_t \odot \tilde{C}_t \quad (4)$$

$$H_t = O_t \odot \tanh(C_t) \quad (5)$$

where  $I_t$ ,  $F_t$ , and  $O_t$  denote, respectively, the input gate, the forget gate, and the output gate. The weight matrices for these gates are represented by  $W$ , while  $U$  signifies the weights of the hidden layer in the preceding time step. The bias terms for different gating mechanisms are denoted by  $b$ . The sigmoid activation function is denoted by  $\sigma$ , and  $\tanh$  denotes the hyperbolic tangent activation function. The symbol  $\odot$  represents the element-wise product of vectors, and  $\tilde{C}_t$  denotes the candidate cell state.

#### 2.4. Feature Fusion Module

Our model includes a feature fusion module, which is essential for integrating information from multiple input sources effectively. This module is pivotal for capturing a comprehensive representation of the input data. The extracted feature vectors are concatenated along the feature dimension to form a unified feature representation, denoted as

$$H_{concat} = \{H_\sigma, H_P, H_S\} \quad (6)$$

where  $H_\sigma$ ,  $H_P$ , and  $H_S$  represent the features extracted by the LSTM models from input sources  $\sigma$ ,  $P$ , and  $S$ , respectively. The concatenated feature vector  $H_{concat}$  is then passed through a fully connected layer to perform linear transformation and nonlinear mapping.

#### 2.5. Residual Module

Our model incorporates a residual module to enhance the effects of feature fusion and improve the predictive performance. This module promotes effective information transmission and stabilizes model training by directly adding the output of the model to the input features from the previous timestep. This methodology ensures the preservation of the original input feature information and facilitates smoother gradient propagation, thereby augmenting the overall model performance. Notably, the residual module in our study uses deformation features from the prior timestep as its inputs. This is expressed as

$$O_{final} = O + X_{T-1} \quad (7)$$

where  $O$  represents the output feature vector after processing through a fully connected layer, and  $X_{T-1}$  denotes the value of the input feature in the last time step.

#### 2.6. Attention Mechanism of the Model

In this study, a self-attention mechanism was employed to capture the long-distance dependencies within monitoring data [25]. The architecture of the self-attention mechanism was designed to enhance the correlation between monitoring data sequences at different time points by weighting and summing these sequences [26], as illustrated in Figure 3. As detailed in Equations (8)–(11), the initial step involves the input data undergoing three linear mappings to produce matrices  $Q$  (query),  $K$  (key), and  $V$  (value). The similarity between  $Q$  and  $K$  is determined through matrix multiplication. To ensure the similarity scores do not become overly large, the results are divided by the scaling factor  $D_k$ . The resulting similarity matrix is then normalized using the SoftMax function, which converts

it into an attention weights matrix. The final output is derived through the weighted summation of the  $V$ , based on the attention weight matrix.

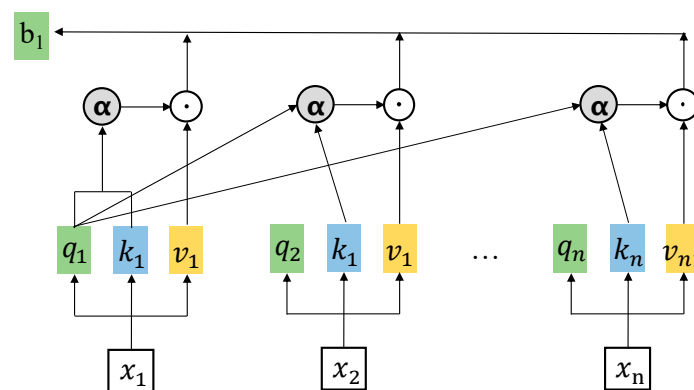
$$Q = W_q X \quad (8)$$

$$K = W_k X \quad (9)$$

$$V = W_v X \quad (10)$$

$$OUTPUT = V softmax\left(\frac{K^T Q}{\sqrt{D_k}}\right) \quad (11)$$

where  $Q$  is the query matrix,  $K$  is the key matrix,  $V$  is the value matrix, and  $D_k$  is the scaling factor.



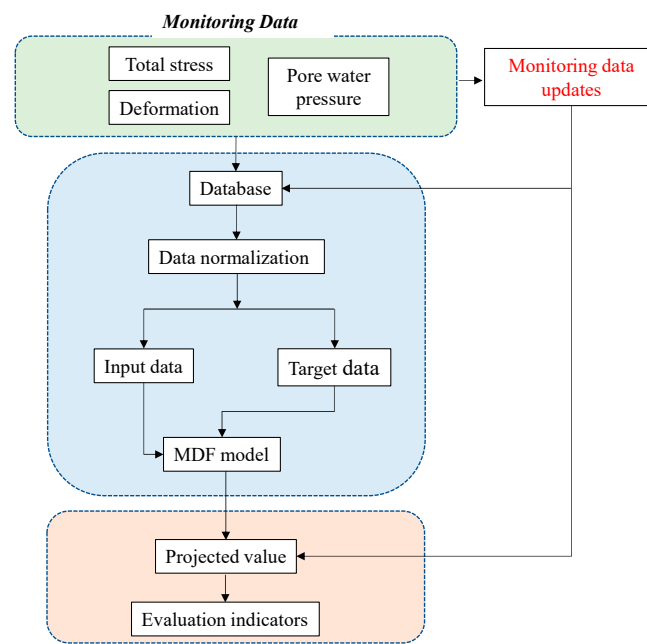
**Figure 3.** Attention mechanism network structure.

### 2.7. Prediction Procedure of the Dynamic Training–Testing Method

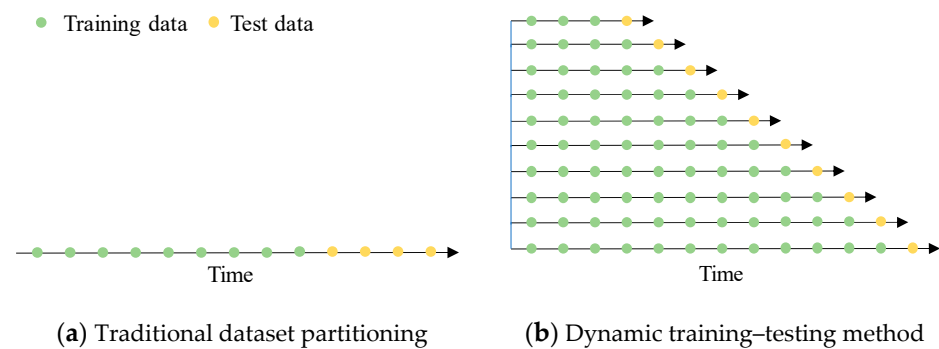
As depicted in Figure 4, the process of predicting soil deformation using a DL model involves several key steps. Initially, monitoring data are collected to create a complete dataset. These data are then normalized using a maximum–minimum normalization technique to ensure consistency across all measurements. Subsequently, the normalized data are trained using the MDF model. It is important to note that the collected data are utilized not only for training but also for evaluating the model’s performance. As the monitoring data are updated, the accuracy of the model’s predictions is reassessed. With continuous updates to the dataset, the training data must be refreshed, necessitating the retraining of the model to accurately predict future soil deformation.

The method proposed in this paper emphasizes the continuity and dynamism of dataset updates in the field of soil monitoring. Figure 5 compares this dynamic updating process with traditional dataset delineation methods. In method (a), which represents the general data division approach, a certain portion of the data is selected for training, followed by another portion for testing. However, in engineering contexts, where the focus is often on monitoring dynamic deformations throughout a construction project, this method may not sufficiently reflect the data distribution. Consequently, using method (a) to assess model accuracy might not yield a true representation of the data.

Conversely, method (b) illustrates a dynamic training–testing approach. As the dataset evolves, the model is continuously trained with newly updated data, with the subsequent day’s data selected for testing. This method more effectively captures the dynamic nature of engineering datasets and provides a more realistic assessment of the model’s performance over time.



**Figure 4.** Implementation flowchart for the proposed method.



**Figure 5.** Different dataset partitioning methods.

## 2.8. SHAP

DL models, which are currently the most popular nonlinear models, have been widely applied in various domains. However, DL models are often considered black-box models with poor interpretability. SHAP [27], which is based on game theory, provides an interpretable method of explaining the output of any DL model [28]. The SHAP method can measure the impact of features on the dependent variable in DL models, assigning an importance value (SHAP value) to each feature. Its expression is

$$\varphi_i = \sum_{S \subseteq N \setminus \{i\}} \frac{|S|!(|N| - |S| - 1)!}{|N|!} [f(S \cup \{i\}) - f(S)] \quad (12)$$

where  $N$  is the set of all features,  $S$  is a subset of  $N$  that does not include feature  $i$ ,  $f(S)$  is the model prediction when only the feature set  $S$  is involved,  $f(S \cup \{i\})$  is the model prediction when feature  $i$  is added to feature set  $S$ , and  $\frac{|S|!(|N| - |S| - 1)!}{|N|!}$  is a combinatorial coefficient used to assign weights to each feature subset.

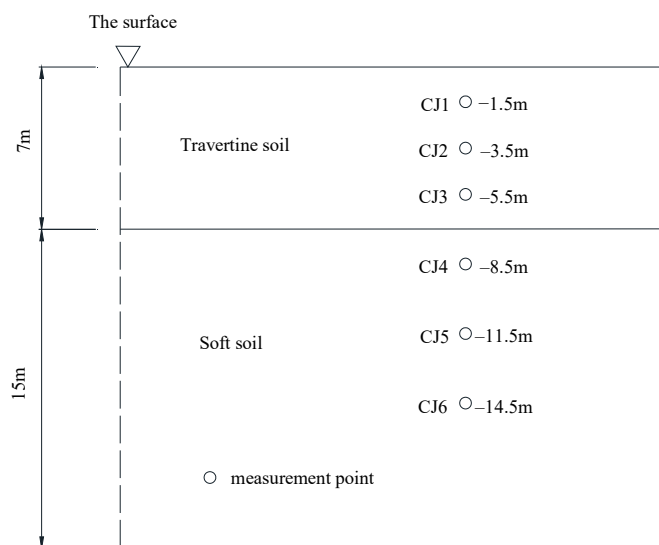
## 3. Experimental Example

### 3.1. Project Overview

This study involved the data measured during the construction of a high-speed railway roadbed, focusing on a test section treated with the vacuum precompression

method to reinforce a soft foundation. The soils in this test section were identified as normally consolidated (NC) based on their stress history. A vacuum gauge was strategically positioned at the center of the reinforced area to measure the vacuum degree beneath the membrane. Concurrently, settlement pipes and magnetic rings were utilized to monitor the layered settlement, and a pore water pressure gauge was embedded to record the pore water pressure.

Figure 6 shows the position of the different measurement points CJ1–CJ6 at various depths in the soil; the depths of the measurement points were 1.5 m, 3.5 m, 5.5 m, 8.5 m, 11.5 m, and 14.5 m, respectively. For the model inputs, deformation was assessed from the stratified settlement data gathered at points CJ1 to CJ6, effective stress was calculated from the pore water pressure recorded at these depths, and total stress was determined using the vacuum levels measured with the vacuum gauges.



**Figure 6.** Monitoring instrument configuration diagram.

### 3.2. Experimental Details

Our model was implemented in PyTorch v2.0, a DL framework written in Python, and ran on an NVIDIA 4070 GPU. The configuration of the MDF model, as detailed in Table 1, was determined through a series of iterative tests. Each hyperparameter was systematically varied within a predetermined range to ascertain the optimal performance of the model. The model's weights were updated using the Adam optimizer, with the batch size, learning rate, and epoch set to 32, 0.01, and 200, respectively. Dropout was employed to mitigate overfitting.

**Table 1.** Details of different models.

Model	MDF	Single-Feature LSTM	Multi-Feature LSTM
Epoch	200	200	200
Batch size	32	32	32
Learning rate	0.01	0.01	0.01
Activation function	Tanh	Tanh	Tanh
Drop out	0.5	0.5	0.5
Optimizer	Adam	Adam	Adam

For comparative analysis, two LSTM models were also tested. One model used a single feature input (deformation), and the other was a multi-feature LSTM incorporating deformation, effective stress, and total stress inputs. Both LSTM models were configured with the same hyperparameters as the MDF model to ensure consistency for comparison.

The model's performance was evaluated using a loss function that measures the agreement between actual and predicted values. Specifically, the mean squared error (MSE) function was employed, which is defined as

$$\text{MSE} = \frac{1}{N} \sum_{i=1}^N (y_i^p - y_i)^2 \quad (13)$$

where  $N$  represents the total number of data sets;  $y_i^p$  and  $y_i$  are the predicted and measured values, respectively.

### 3.3. Model Performance Criteria

The mean absolute percentage error (MAPE) and root mean square error (RMSE) were employed to analyze the performance of the various models. The higher these values, the worse the prediction performance. The definitions of the MAPE and RMSE are as follows:

$$\text{MAPE} = \frac{1}{N} \sum_{i=1}^N \frac{|y_i^p - y_i|}{y_i} \times 100\% \quad (14)$$

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_i^p - y_i)^2} \quad (15)$$

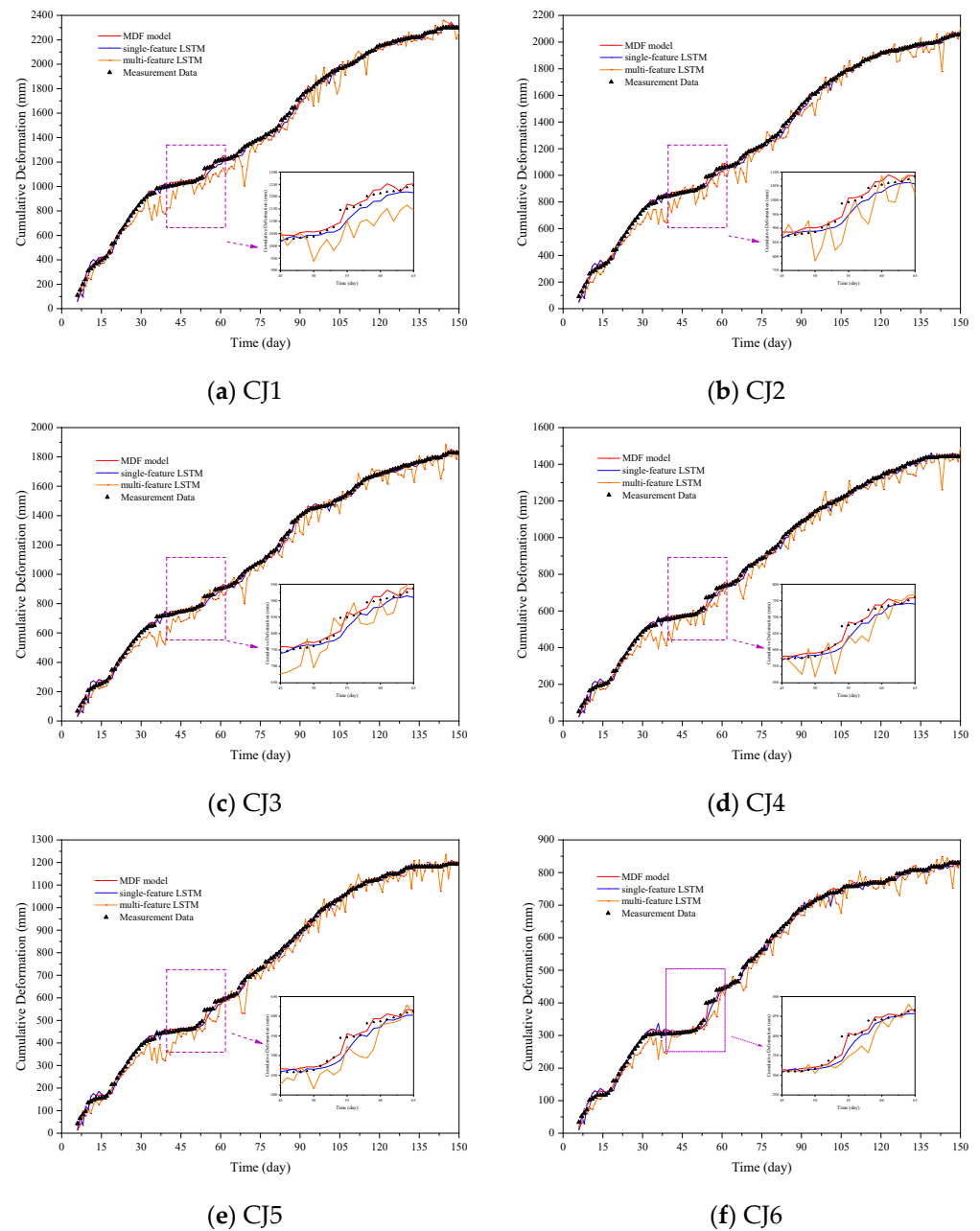
where  $N$  represents the number of samples,  $y_i$  represents the target data of the  $i$ th set of samples, and  $y_i^p$  represents the output of the  $i$ th set of samples.

### 3.4. Results

In this study, the proposed MDF model was employed to predict soil deformation, and its performance was compared with that of the traditional single-feature LSTM and multi-feature LSTM models. Figure 7 shows the results of the comparison of these three models at the CJ1–CJ6 measurement points. The deformation data of six measurement points, CJ1–CJ6, over approximately 150 days, demonstrate different deformation variations, revealing the strong nonlinear characteristics of the deformations. It is evident that all models—MDF, single-feature LSTM, and multi-feature LSTM—adequately captured the fundamental deformation trends. Nonetheless, the MDF model demonstrated superior accuracy in capturing deformations. Although LSTM-based models are adept at feature extraction, their predictive accuracy diminishes significantly at points of abrupt deformation changes. To highlight this, insets showing these deformation changes are incorporated into each plot. The analysis showed that the MDF model surpassed the LSTM models in terms of prediction accuracy at deformation mutation points CJ1–CJ6.

Although the multi-feature LSTM model considers several parameters, including stress features, it failed to deliver the expected superior results. Several factors contributed to this outcome: Firstly, the measured data inevitably contained noise, and the increased dimensionality of the input features exacerbated the issue of data quality, making it more challenging for the model to learn effective patterns, resulting in performance degradation. Secondly, the correlation between the dimensions of multiple input features introduced multicollinearity problems, further deteriorating model accuracy. This results in a significant performance degradation of the LSTM model when multiple features were extracted. In contrast, the MDF model boosts feature extraction capabilities by separately extracting features before fusing them, significantly enhancing model accuracy.





**Figure 7.** Comparison of three prediction models at different measurement points.

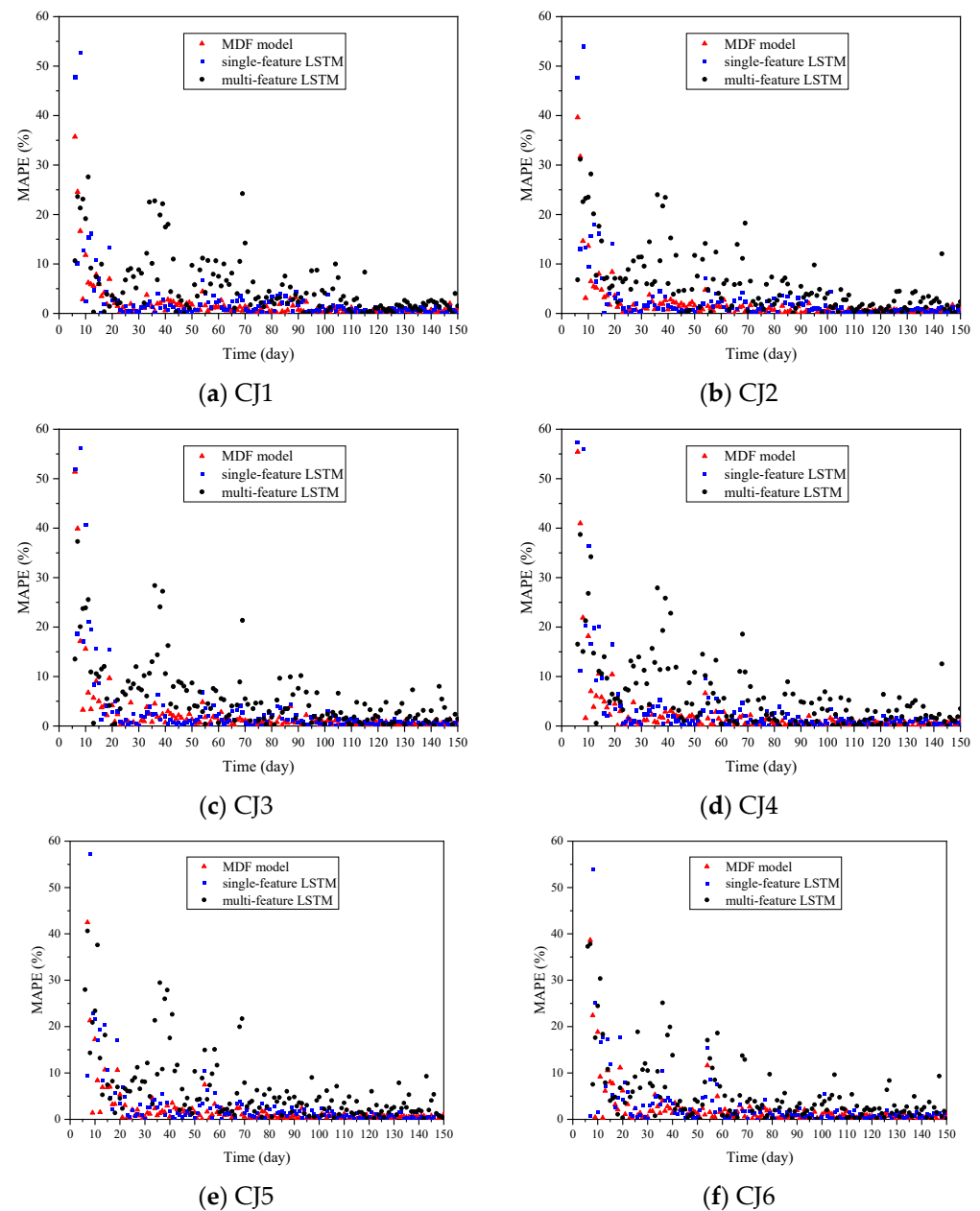
## 4. Discussion

### 4.1. Error Analysis

To more clearly reflect the prediction accuracy of each model, the daily prediction errors were analyzed concurrently. As the models were updated, daily predictions were compared against actual monitoring data to assess errors, as illustrated in Figure 8. The MAPE results for CJ1–CJ6 demonstrated that the MDF model consistently recorded the lowest errors in the scatter plot relative to both the single-feature and multi-feature LSTM models, indicating a substantial enhancement in prediction accuracy.

Moreover, the MDF model consistently exhibited uniform error distribution, in contrast to the single-feature and multi-feature LSTM models, which displayed more volatile error distributions. Particularly, multi-feature LSTM had the most uneven prediction error distribution, with large dispersion, further revealing its limitations in handling multifea-

ture data. Conversely, the MDF model, specifically designed to handle multifeature data, showed more stable and robust prediction performance.



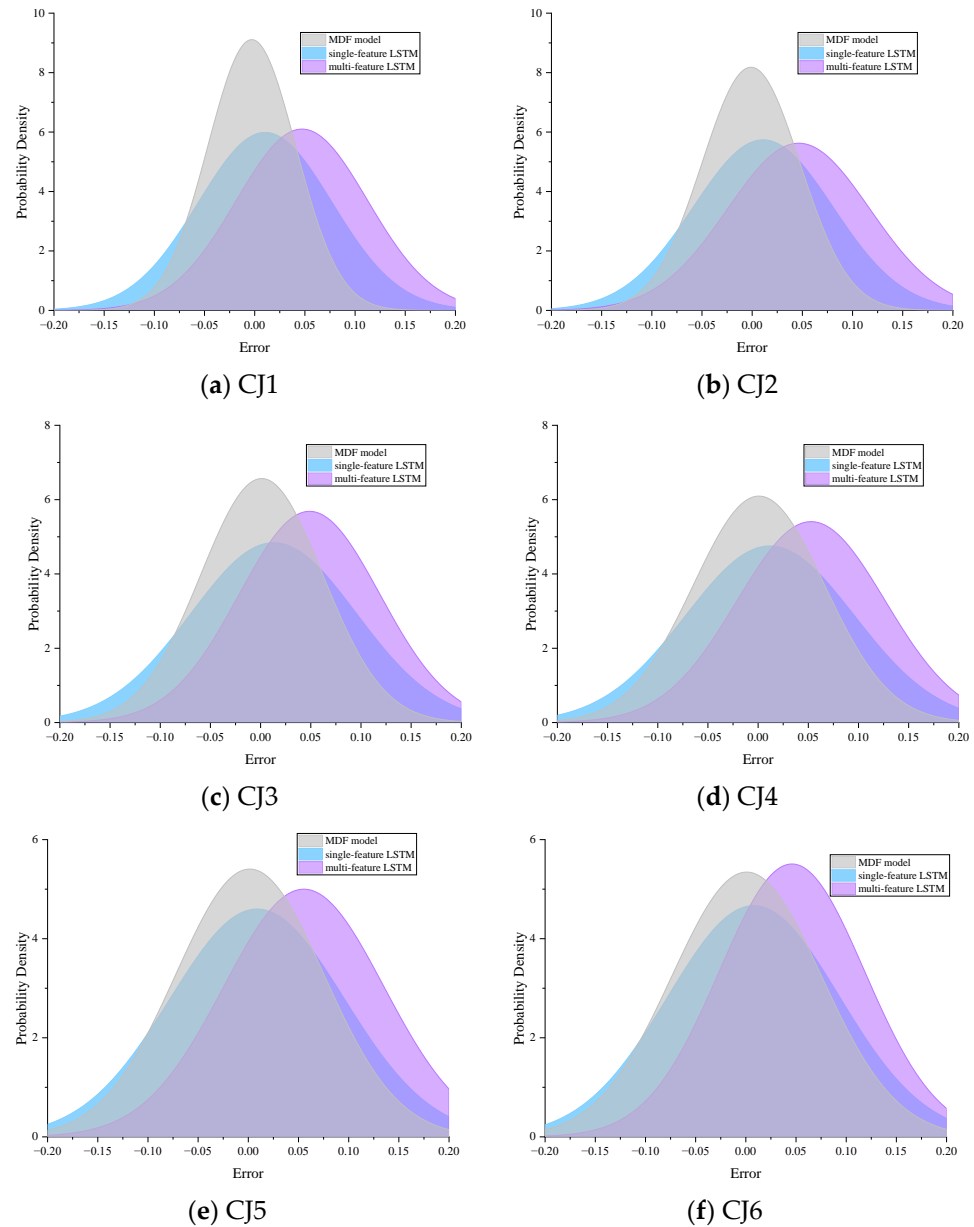
**Figure 8.** Daily prediction errors for three different models.

Figure 8 also reveals that at the onset of the monitoring period, all models—including the MDF, single-feature LSTM, and multi-feature LSTM—exhibited large errors due to the initially limited data available. DL models typically require a substantial amount of data to ensure high generalization performance. As the dataset grew and the models underwent continual updates and training, their performance improved progressively. By the 25th day, the prediction errors of the MDF model were decreased significantly, indicating improved accuracy and stability.

To gain a comprehensive understanding of the prediction errors, the probability density function (PDF) of the normalized error distribution was employed, as outlined in Equation (16). As depicted in Figure 9, the prediction errors generated by the MDF model for all measurement points were predominantly concentrated within a 10% range, with a probability of approximately 90%. In contrast, the probability for errors within

10% for single-feature LSTM was approximately 80%; for multi-feature LSTM, it was approximately 74%. This indicated that the MDF model had a higher probability of having fewer errors. Additionally, the average error value for the MDF model was approximately zero, suggesting that its errors mainly concentrated at zero. In comparison, the average error values of single-feature LSTM and multi-feature LSTM were noticeably far from zero, indicating a more dispersed error distribution.

$$\pi_i = \text{PDF} \left\{ \frac{y_i^p - y_i}{y_i} \right\} \quad (16)$$



**Figure 9.** The PDF of three types of prediction model.

In order to illustrate the performance of the different models more clearly, Table 2 presents the MAPE and RMSE values for each model at various measurement points, along with the probability of each model maintaining an error rate within 10%. The MAPE and RMSE values were calculated by comparing the predictions from the MDF models against actual field measurements, as detailed in Section 3.3. Additionally, the probability of

sustaining an error rate within 10% was computed based on the described error distribution. Specifically, the MAPE of the MDF model was lower by approximately 30% compared with that of single-feature LSTM and about 63% compared with that of the multi-feature LSTM model. Additionally, the RMSE was reduced by approximately 36% compared with that of the single-feature LSTM model and by about 77% compared with that of the multi-feature LSTM model. Furthermore, the probability of error within 10% was higher by approximately 13% than that of the single-feature LSTM model and by about 20% compared with that of the multi-feature LSTM model.

**Table 2.** Indicators of different measurement points on the test set.

Metric		MAPE (%)			RMSE (mm)			Probability of Error within 10% (%)		
Models	MDF	Single-Feature LSTM	Multi-Feature LSTM	MDF	Single-Feature LSTM	Multi-Feature LSTM	MDF	Single-Feature LSTM	Multi-Feature LSTM	
CJ1	1.81	2.52	5.49	17.63	26.98	81.11	98	86	78	
CJ2	1.81	2.68	5.76	14.30	23.49	68.34	96	84	76	
CJ3	2.08	3.16	5.73	13.88	22.03	60.88	90	77	75	
CJ4	2.17	3.27	5.99	11.07	18.13	49.83	87	76	72	
CJ5	2.35	3.31	6.38	9.65	14.49	43.61	82	75	69	
CJ6	2.51	3.28	5.36	8.39	12.19	26.83	82	76	75	

The accuracy improvement primarily arises from the MDF model's ability to effectively integrate diverse monitoring data and accurately capture the complex interdependencies among the factors that influence soil deformation. This comprehensive data integration allows the model to describe soil behavior more precisely, an accomplishment that single-feature or multi-feature LSTM models fail to achieve due to their limited utilization of the available monitoring data.

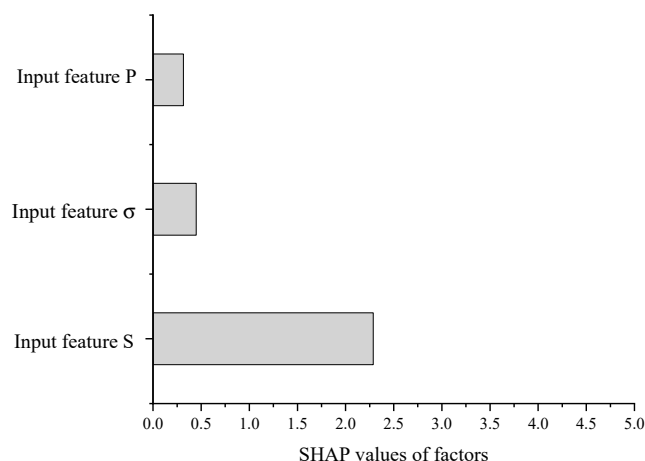
The model proposed in this paper demonstrated high predictive accuracy in the provided examples. Currently, the MDF model requires the availability of all three specified types of monitoring data for training. It is crucial to acknowledge that the MDF model cannot automatically adjust to the absence of any data type. Moreover, the model's performance in other scenarios remains untested, which not only identifies a key limitation of this study but also directs future research efforts.

In subsequent studies, we will address scenarios where one or more data type is missing. The MDF model will be applied across a variety of cases. These enhancements are anticipated to refine our deformation prediction techniques, resulting in more precise risk assessments. The MDF model is projected to offer a comprehensive solution for effectively managing excessive deformation. By proactively predicting significant deformations, timely and appropriate mitigation measures can be implemented. This proactive strategy will significantly enhance our capacity to minimize the impacts of both human-made and natural disasters.

#### 4.2. SHAP Analysis of Feature Importance in the MDF Model

The MDF model proposed in this study integrates various monitoring data features to enhance prediction accuracy. To understand how the model operates, the SHAP method was employed. This method involves calculating the SHAP values for different features, which are subsequently aggregated to determine the outcomes of predictions.

Analyzing these values elucidates how a model derives its predictions from the contributions of distinct features. As presented in Figure 10, the horizontal axis represents the SHAP values, where higher values indicate a greater influence on the dependent variable. Notably, the deformation feature's SHAP value substantially surpasses those of other data features, emphasizing its critical role. This significant impact stems from the MDF model's ability to autonomously learn from historical data, confirming that past deformation data are crucial for accurate future deformation predictions.



**Figure 10.** SHAP values of factors.

Furthermore, the analysis showed that the effective stress feature had a more pronounced impact than the total stress feature. Each data feature provides a unique contribution, showcasing the sophisticated integration of various monitoring data by the proposed model. This intricate integration allows the MDF model to deliver highly accurate predictions, demonstrating its utility in practical applications of deformation monitoring.

## 5. Conclusions

We addressed the inefficiencies of existing DL models in effectively utilizing multiple types of monitoring data for predicting soil deformation by proposing the MDF model. Specifically, three LSTM layers were designed to extract the features from various monitoring data types, employing data feature fusion and residual structures to enhance prediction accuracy. A dynamic training–testing method was also developed to address the shortcomings of traditional dataset partitioning strategies. The effectiveness of this method was validated using a dataset containing 150 days of strain–time history measurements, and a comparative analysis was conducted against traditional single-feature and multi-feature LSTM models.

The results showed that the MAPE of the MDF model were lower by approximately 30% than that of the single-feature LSTM model and by about 63% than that of the multi-feature LSTM model. Additionally, the RMSE was reduced by approximately 36% compared to that of the single-feature LSTM model and by about 77% compared to that of the multi-feature LSTM model. These findings underscored the MDF model’s superior ability in predicting soil deformation. The model accurately forecasted nonlinear soil deformation during construction, enabling the implementation of pre-emptive measures to mitigate risks and prevent catastrophic damage, thus enhancing public safety and reducing economic losses.

Furthermore, SHAP analysis highlighted the contributions of various factors to the predictions. This analysis showed that the MDF model’s predictions primarily stem from three types of monitoring data: deformation characteristics, effective stress characteristics, and total stress characteristics. Deformation characteristics were identified as the most critical, followed by effective stress characteristics, with total stress characteristics being the least influential. This understanding of the relative importance of the different factors not only enhances the model’s interpretability but also provides valuable guidance for future monitoring and data collection strategies.

**Author Contributions:** Conceptualization: J.L. and Y.W.; methodology: H.W. and Y.W.; software: H.W.; validation: L.Z. and Y.W.; formal analysis: Y.Z., L.Z. and Y.W.; investigation: H.W. and Y.Z.; resources: Y.W. and C.L.; data curation: C.L. and H.W.; writing—original draft preparation: H.W.; writing—review and editing: Y.W.; visualization: H.W.; supervision: Y.W.; project administra-

tion: J.L.; funding acquisition: H.W. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was supported by the Postgraduate Research & Practice Innovation Program of Jiangsu Province (No. KYCX23-0686).

**Data Availability Statement:** The data presented in this study are available in this article.

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

## References

1. Qi, Y.; Tian, G.; Bai, M.; Song, L. Study on Construction Deformation Prediction and Disaster Warning of Karst Slopes Based on Grey Theory. *Bull. Eng. Geol. Environ.* **2023**, *82*, 62. [[CrossRef](#)]
2. Cui, J.; Yang, Z.; Azzam, R. Field Measurement and Numerical Study on the Effects of Under-Excavation and Over-Excavation on Ultra-Deep Foundation Pit in Coastal Area. *J. Mar. Sci. Eng.* **2023**, *11*, 219. [[CrossRef](#)]
3. Yin, J.H.; Chen, Z.J.; Feng, W.Q. A General Simple Method for Calculating Consolidation Settlements of Layered Clayey Soils with Vertical Drains under Staged Loadings. *Acta Geotech.* **2022**, *17*, 3647–3674. [[CrossRef](#)]
4. Li, P.L.; Yin, Z.Y.; Song, D.B.; Yin, J.H.; Pan, Y. Axisymmetric Finite Strain Consolidation Model for Soft Soil Consolidation with Vertical Drains under Combined Loading Considering Creep and Non-Darcy Flow. *Geotext. Geomembr.* **2024**, *52*, 241–259. [[CrossRef](#)]
5. Lei, M.; Luo, S.; Chang, J.; Zhang, R.; Kuang, X.; Jiang, J. The Influences of Vacuum-Surcharge Preloading on Pore Water Pressure and the Settlement of a Soft Foundation. *Sustainability* **2023**, *15*, 7669. [[CrossRef](#)]
6. Yin, J.H.; Feng, W.Q. A New Simplified Method and Its Verification for Calculation of Consolidation Settlement of a Clayey Soil with Creep. *Can. Geotech. J.* **2017**, *54*, 333–347. [[CrossRef](#)]
7. Zhang, R.; Wu, C.; Goh, A.T.C.; Böhlke, T.; Zhang, W. Estimation of Diaphragm Wall Deflections for Deep Braced Excavation in Anisotropic Clays Using Ensemble Learning. *Geosci. Front.* **2021**, *12*, 365–373. [[CrossRef](#)]
8. Zheng, G.; Zhang, W.; Zhang, W.; Zhou, H.; Yang, P. Neural Network and Support Vector Machine Models for the Prediction of the Liquefaction-Induced Uplift Displacement of Tunnels. *Undergr. Space* **2021**, *6*, 126–133. [[CrossRef](#)]
9. Park, H.I.; Kim, K.S.; Kim, H.Y. Field Performance of a Genetic Algorithm in the Settlement Prediction of a Thick Soft Clay Deposit in the Southern Part of the Korean Peninsula. *Eng. Geol.* **2015**, *196*, 150–157. [[CrossRef](#)]
10. Kong, F.; Lu, D.; Ma, Y.; Li, J.; Tian, T. Analysis and Intelligent Prediction for Displacement of Stratum and Tunnel Lining by Shield Tunnel Excavation in Complex Geological Conditions: A Case Study. *IEEE Trans. Intell. Transport. Syst.* **2022**, *23*, 22206–22216. [[CrossRef](#)]
11. Chen, J.H.; Cui, D.W. A multi-step prediction model of ROA-ELM dam deformation based on wavelet packet transform. *J. Three Gorges Univ. (Nat. Sci. Ed.)* **2022**, *44*, 21–27. [[CrossRef](#)]
12. Ray, R.; Kumar, D.; Samui, P.; Roy, L.B.; Goh, A.T.C.; Zhang, W. Application of Soft Computing Techniques for Shallow Foundation Reliability in Geotechnical Engineering. *Geosci. Front.* **2021**, *12*, 375–383. [[CrossRef](#)]
13. Zhang, P.; Wu, H.N.; Chen, R.P.; Chan, T.H.T. Hybrid Meta-Heuristic and Machine Learning Algorithms for Tunneling-Induced Settlement Prediction: A Comparative Study. *Tunn. Undergr. Space Technol.* **2020**, *99*, 103383. [[CrossRef](#)]
14. Zhang, N.; Shen, S.L.; Zhou, A.; Jin, Y.F. Application of LSTM Approach for Modelling Stress–Strain Behaviour of Soil. *Appl. Soft Comput.* **2021**, *100*, 106959. [[CrossRef](#)]
15. Xie, P.; Zhou, A.; Chai, B. The Application of Long Short-Term Memory(LSTM) Method on Displacement Prediction of Multifactor-Induced Landslides. *IEEE Access* **2019**, *7*, 54305–54311. [[CrossRef](#)]
16. Yang, B.; Yin, K.; Lacasse, S.; Liu, Z. Time Series Analysis and Long Short-Term Memory Neural Network to Predict Landslide Displacement. *Landslides* **2019**, *16*, 677–694. [[CrossRef](#)]
17. Zhang, S.J.; Tan, Y. Prediction of pit deformation based on LSTM algorithm. *Tunn. Constr.* **2022**, *42*, 113–120. [[CrossRef](#)]
18. Lv, Q.F.; Li, Y.; Niu, R.; Xui, H.X.; Mao, N.; Kang, Q.Y. Deep learning-based prediction of surrounding rock deformation in tunnels with special geotechnical conditions. *J. Appl. Basic Eng. Sci.* **2023**, *31*, 1590–1600. [[CrossRef](#)]
19. Hong, Y.C.; Qian, J.G.; Ye, Y.X.; Cheng, L. Application of spatio-temporal correlation feature-based CNN-LSTM model in deformation prediction of foundation pit engineering. *J. Geotech. Eng.* **2021**, *43*, 108–111. [[CrossRef](#)]
20. Song, F.; Zhong, H.; Li, J.; Zhang, H. Multi-Point RCNN for Predicting Deformation in Deep Excavation Pit Surrounding Soil Mass. *IEEE Access* **2023**, *11*, 124808–124818. [[CrossRef](#)]
21. Zhang, J.; Phoon, K.K.; Zhang, D.; Huang, H.; Tang, C. Deep Learning-Based Evaluation of Factor of Safety with Confidence Interval for Tunnel Deformation in Spatially Variable Soil. *J. Rock Mech. Geotech. Eng.* **2021**, *13*, 1358–1367. [[CrossRef](#)]
22. Chen, R.; Liu, J.; Li, J.H.; Ng, C.W.W. An Integrated High-Capacity Tensiometer for Measuring Water Retention Curves Continuously. *Soil Sci. Soc. Am. J.* **2015**, *79*, 943–947. [[CrossRef](#)]
23. Zhou, H.; Zhang, S.; Peng, J.; Zhang, S.; Li, J.; Xiong, H.; Zhang, W. Informer: Beyond Efficient Transformer for Long Sequence Time-Series Forecasting. *AAAI* **2021**, *35*, 11106–11115. [[CrossRef](#)]
24. Chen, X.X.; Yang, J.; He, G.F.; Huang, L.C. Development of an LSTM-Based Model for Predicting the Long-Term Settlement of Land Reclamation and a GUI-Based Tool. *Acta Geotech.* **2023**, *18*, 3849–3862. [[CrossRef](#)]

25. Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A.N.; Kaiser, Ł.; Polosukhin, I. Attention is all you need. In Proceedings of the 31st International Conference on Neural Information Processing Systems, Long Beach, CA, USA, 4–9 December 2017. [[CrossRef](#)]
26. Feng, J.; Yan, L.; Hang, T. Stream-Flow Forecasting Based on Dynamic Spatio-Temporal Attention. *IEEE Access* **2019**, *7*, 134754–134762. [[CrossRef](#)]
27. 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. [[CrossRef](#)]
28. Baptista, M.L.; Goebel, K.; Henriques, E.M.P. Relation between Prognostics Predictor Evaluation Metrics Andlocal Interpretability SHAP Values. *Artif. Intell.* **2022**, *306*, 103667. [[CrossRef](#)]

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.