applied sciences

MDPI

Article

Dynamic Response Prediction of Railway Bridges Considering Train Load Duration Using the Deep LSTM Network

Sui Tan [1,2,3,*], Xiandong Ke [4], Zhenhao Pang [4] and Jianxiao Mao [4]

[1] National Engineering Research Center of High-Speed Railway Construction Technology, Changsha 410075, China
[2] China Railway Group Limited, Beijing 100039, China
[3] School of Civil Engineering, Central South University, Changsha 410075, China
[4] Key Laboratory of Concrete and Prestressed Concrete Structures of Ministry of Education, Southeast University, Nanjing 211189, China
* Correspondence: hsrlab_stan@163.com

Abstract: Monitoring and predicting the dynamic responses of railway bridges under moving trains, including displacement and acceleration, are vital for evaluating the safety and serviceability of the train–bridge system. Traditionally, finite element analysis methods with high computational burden are used to predict the train-induced responses according to the given train loads and, hence, cannot easily be integrated as an available structural-health-monitoring strategy. Therefore, this study develops a novel framework, combining the train–bridge coupling mechanism and deep learning algorithms to efficiently predict the train-induced bridge responses while considering train load duration. Initially, the feasibility of using neural networks to calculate the train–bridge coupling vibration is demonstrated by leveraging the nonlinear relationship between train load and bridge responses. Subsequently, the instantaneous multiple moving axial loads of the moving train are regarded as the equivalent node loads that excite adjacent predefined nodes on the bridge. Afterwards, a deep long short-term memory (LSTM) network is established as a surrogate model to predict the train-induced bridge responses. Finally, the prediction accuracy is validated using a numerical case study of a simply supported railway bridge. The factors that may affect the prediction accuracy, such as network structure, training samples, the number of structural units, and noise level, are discussed. Results show that the developed framework can efficiently predict the train-induced bridge responses. The prediction accuracy of the bridge displacement is higher than that of the acceleration. In addition, the robustness of the displacement prediction is proven to be better than that of the acceleration with the variation of carriage number, riding speed, and measurement noise.

Keywords: railway bridges; structural health monitoring; dynamic response prediction; train–bridge coupling; deep learning; LSTM network

check for updates

Citation: Tan, S.; Ke, X.; Pang, Z.; Mao, J. Dynamic Response Prediction of Railway Bridges Considering Train Load Duration Using the Deep LSTM Network. *Appl. Sci.* **2024**, *14*, 9161. https://doi.org/10.3390/app14209161

Academic Editor: Giuseppe Lacidogna

Received: 8 September 2024
Revised: 1 October 2024
Accepted: 3 October 2024
Published: 10 October 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

The total number of railway bridges in China has exceeded 200,000, among which 80% are medium–small-span bridges (porous span is greater than or equal to 8 m but less than 100 m; single span is greater than or equal to 5 m but less than 40 m). These medium–small railway bridges are facing aging problems, which urgently require the improvement and upgrade of the monitoring techniques to ensure the operational safety. The dynamic response of a railway bridge under moving train loads, such as displacement and acceleration, can reflect the overall reliability of the train–bridge system [1]. The accurate and efficient prediction of the dynamic response can provide vital information for evaluating the safety and serviceability of the train–bridge system.

The train-induced dynamic responses of bridges can be obtained through finite element simulation [2] or the sensors included in their structural-health-monitoring (SHM) system, e.g., accelerometers and global positioning system (GPS). In one aspect, the finite

element simulation that considers the train–bridge coupling theory heavily relies on numerical integration, incurring high computational costs and making it difficult to meet the requirement of rapidly predicting bridge responses. In another aspect, the SHM system of railway bridges can collect massive structural response data that are available for in-service assessment; however, it is not realistic and affordable for one to be installed on all medium- and small-span railway bridges [3,4].

Researchers have been focusing their efforts on data-driven structural response prediction methods. Among them, machine learning methods have emerged as a research hotspot [5], gradually becoming applied to the prediction of structural dynamic responses in various contexts such as residential buildings, bridges, and wind turbines [6,7]. Machine learning methods enable the prediction of dynamic responses under different types of loads, including seismic loads, wind loads, and train loads. On the one hand, experts have extensively employed deep learning methods, e.g., fully connected neural networks [8] and feedforward neural networks [9], to predict the nonlinear dynamic time-history responses of bridges and buildings under seismic actions. The prediction errors are then utilized for post-earthquake structural damage assessment and localization [10].

On the other hand, the machine learning methods have also been employed to predict structural dynamic responses under wind excitations. For example, the wind pressure time histories on high-rise buildings were predicted using the backpropagation neural network [11], the aerodynamical responses of long-span bridges were predicted based on the multi-layer feedforward networks [12,13]. Additionally, advancements in predicting the dynamic responses of bridges induced by train loads have been progressing steadily. For example, the GA-TCN-LSTM model proposed by Zhang et al. [14] shows high accuracy in predicting dynamic response at different train speeds. Li et al. [15] employed information such as train speed, train position, and road surface roughness as inputs for the artificial neural network to predict the dynamic response of railway bridge systems. The previous studies proved that the design of the input data is the critical problem during the process of constructing an artificial neural network to predict the train-induced bridge dynamic responses; however, existing studies struggled to accurately incorporate the time history information of train loads.

To address the abovementioned issues, a novel framework for predicting the structural dynamic responses of railway bridges was proposed by combining the deep long short-term memory network, denoted as LSTM, and vehicle–bridge coupling theory. Specifically, the framework considers the load of each node of the bridge structure at any moment as the input information of the neural network to establish a deep LSTM network to predict the train-induced dynamic responses of the railway bridge. A numerical example of a single-supported bridge under a moving train is utilized to validate the proposed methodology and illustrate the potential influence factors.

2. Background Theory and Proposed Framework

2.1. Brief Description of LSTM Network

The long short-term memory (LSTM) network was proposed by Alex Sherstinsky to solve the problems that may exist in training the recurrent neural network (RNN), e.g., long-distance dependence, gradient disappearance, and gradient explosion [16,17]. By introducing the state of neurons, the LSTM network can effectively preserve information over a long duration, avoiding the occurrence of effective information being ignored or forgotten. At the same time, multiple neural network states are added in the LSTM network internally, which ensures the stability and accuracy. In fact, an LSTM network is a special RNN. Compared with the traditional RNN structure, an LSTM network has two transmission states to regulate the information flow between neural network units. Its unit structure consists of three gating units, namely, forget gate, input gate, and output gate, as is shown in Figure 1. The gate essentially takes the known vector as the input, the activated real vector as the output, and uses the value matrix and deviation to adjust the network state and transmit information [18].

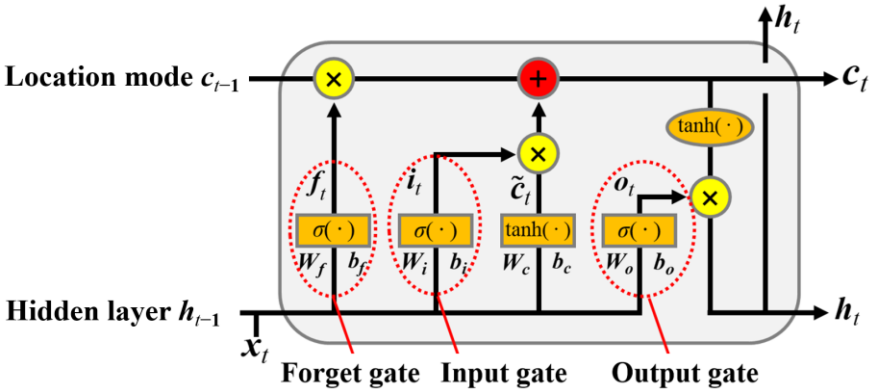


Figure 1. The diagram of LSTM unit structure.

The role of the forget gate is to determine the proportion of the information that the unit state of the previous time step needs to retain to the current time step, according to the transferring and weakening ratio of the previous unit state. The input data of the forget gate are the information input of the current time step and the state output of the previous time step. The degree of forgetting, denoted as f_t, can be calculated according to Equation (1):

$$f_t = \sigma\left(W_f \cdot [h_{t-1};\ x_t] + b_f\right) \tag{1}$$

where $f_t \in \mathbb{R}^{N_h \times 1}$ is the forgetting degree of the forget gate at time t and N_h is the number of hidden layer units of the neural network. $h_{t-1} \in \mathbb{R}^{N_h \times 1}$ is the number of hidden layer states at time $t-1$. $x_t \in \mathbb{R}^{N_x \times 1}$ is the data information input at time t, and N_x is the dimension of the input data. $W_f \in \mathbb{R}^{N_h \times (N_h + N_x)}$ and $b_f \in \mathbb{R}^{N_h \times 1}$ are the weight coefficient matrix and bias value vector of the forget gate, respectively. $\sigma(\cdot) \in (0,\ 1)$ is the Sigmoid operation and can be calculated according to Equation (2):

$$\sigma(x) = \frac{1}{1 + e^{-x}} \tag{2}$$

The functions of the input gate are to receive, correct, and output parameters and to determine the proportion of the input data of the current time step that need to be saved in the unit state. The input data are consistent with the forget gate, as shown in Equations (3) and (4):

$$i_t = \sigma(W_i \cdot [h_{t-1};\ x_t] + b_i) \tag{3}$$

$$\tilde{c}_t = \tanh(W_c \cdot [h_{t-1};\ x_t] + b_c) \tag{4}$$

where $i_t \in \mathbb{R}^{N_h \times 1}$ and $\tilde{c}_t \in \mathbb{R}^{N_h \times 1}$ are the first- and second-part outputs of the input gate at time t. The first-part output is used to update the conveyor information, and the second-part output is the information added to the conveyor. $W_{i/c} \in \mathbb{R}^{N_h \times (N_h + N_x)}$ and $b_{i/c} \in \mathbb{R}^{N_h \times 1}$ are the weight coefficient matrix and bias vector of the input gate, respectively. $\tanh(\cdot) \in (-1,\ 1)$ is a hyperbolic tangent function, as shown in Equation (5):

$$\tanh(x) = \frac{e^{x} - e^{-x}}{e^{x} + e^{-x}} \tag{5}$$

Combined with the output information of the forget gate and the input gate, the unit state of the previous time step is updated, as shown in Equation (6):

$$c_t = f_t \otimes c_{t-1} + i_t \otimes \tilde{c}_t \tag{6}$$

where $c_{t-1} \in \mathbb{R}^{N_h \times 1}$ and $c_t \in \mathbb{R}^{N_h \times 1}$ are the unit states at $t-1$ and t, respectively. $\otimes$ is the Hadamard product, which expresses the mathematical operation of the multiplication of the elements corresponding to the position of the vector. $f_t \otimes c_{t-1}$ is the unit state of

selectively passing or ignoring the previous time step. $\boldsymbol{i}_t \otimes \widetilde{\boldsymbol{c}}_t$ adds new information for the current time step.

The function of the output gate is to control the proportion of the current unit state that needs to be output to the current output value, as shown in Equations (7) and (8):

$$\boldsymbol{o}_t = \sigma(\boldsymbol{W}_{\boldsymbol{o}} \cdot [\boldsymbol{h}_{t-1};\ \boldsymbol{x}_t] + \boldsymbol{b}_{\boldsymbol{o}}) \tag{7}$$

$$\boldsymbol{h}_t = \boldsymbol{o}_t \otimes \tanh(\boldsymbol{c}_t) \tag{8}$$

where $\boldsymbol{o}_t \in \mathbb{R}^{N_h \times 1}$ is the output of the output gate at time t and $\boldsymbol{h}_t \in \mathbb{R}^{N_h \times 1}$ is the hidden layer state at time t. $\boldsymbol{W}_{\boldsymbol{o}} \in \mathbb{R}^{N_h \times (N_h + N_x)}$ and $\boldsymbol{b}_{\boldsymbol{o}} \in \mathbb{R}^{N_h \times 1}$ are the weight coefficient matrix and bias vector of the output gate, respectively.

The mean squared error (MSE) between the predicted and measured data is usually adopted as the loss function when constructing a deep learning framework, as shown in Equation (9):

$$\mathrm{MSE}(\boldsymbol{W}_{\boldsymbol{\alpha}},\ \boldsymbol{b}_{\boldsymbol{\alpha}} | \boldsymbol{D}) = \sum_{i=1}^{N_D} \| \boldsymbol{Y}_i - \hat{\boldsymbol{Y}}_i(\boldsymbol{W}_{\boldsymbol{\alpha}},\ \boldsymbol{b}_{\boldsymbol{\alpha}},\ \boldsymbol{X}_i) \|_2^2 \tag{9}$$

where $\boldsymbol{W}_{\boldsymbol{\alpha}}$ and $\boldsymbol{b}_{\boldsymbol{\alpha}}$ ($\boldsymbol{\alpha} = \boldsymbol{f}, \boldsymbol{i}, \boldsymbol{c}, \boldsymbol{o}$, etc.) are the weight coefficient matrix and bias value vector of each value, respectively. $\boldsymbol{D} = \{(\boldsymbol{X}_I,\ \boldsymbol{Y}_i) \mid i = 1, 2, \ldots, N_D\}$ represents the training set data. $\boldsymbol{X}_i \in \mathbb{R}^{N_x \times M_i}$ and $\boldsymbol{Y}_i \in \mathbb{R}^{N_y \times M_i}$ are the input data and output data, respectively, of the i_{th} group data sample in the dataset. N_y is the dimension of the output data. M_i is the number of time steps of the i_{th} group data sample, and N_D is the number of data samples. $\hat{\boldsymbol{Y}}_i(\boldsymbol{W}_{\boldsymbol{\alpha}},\ \boldsymbol{b}_{\boldsymbol{\alpha}},\ \boldsymbol{X}_i)$ are the neural network prediction output results of given network parameters and input data. For the stacked LSTM network, the unknown parameters to be trained in each layer include the weight matrix and bias vector of the four activation functions, the number of which can be calculated according to Equation (10):

$$N_p = 4 \times [N_h \times (N_h + N_x) + N_h] \tag{10}$$

Combining the advantages of AdaGrad and RMSProp optimization algorithms, Adam optimizer [19] is utilized to train the LSTM network model. The optimization goal is shown in Equation (11):

$$\{\hat{\boldsymbol{W}}_{\boldsymbol{\alpha}}, \hat{\boldsymbol{b}}_{\boldsymbol{\alpha}}\} = \operatorname{argmin} \mathrm{MSE}(\boldsymbol{W}_{\boldsymbol{\alpha}},\ \boldsymbol{b}_{\boldsymbol{\alpha}} | \boldsymbol{D}) \tag{11}$$

where $\hat{\boldsymbol{W}}_{\boldsymbol{\alpha}}$ and $\hat{\boldsymbol{b}}_{\boldsymbol{\alpha}}$ represent the optimal weight coefficient matrix and the optimal bias value vector, respectively.

2.2. Feasibility of Predicting Bridge Responses Using the LSTM Network

This section evaluates the feasibility of using the LSTM network as a surrogate model to predict the vehicle-induced structural responses of a railway bridge. There is a highly nonlinear relationship between the load input and the response output of the railway bridge structure, so the surrogate model can be used to simulate the relevant physical laws. If there is a mathematical model that can accurately represent this complex mapping relationship and the load input is mainly caused by the train load, then the vehicle-induced dynamic response prediction model proposed in this paper is feasible. Therefore, the feasibility of the model depends on the following three conditions: (1) The rationality of the designed LSTM network structure; (2) A specific one-to-one mapping relationship between train load excitation and bridge responses; (3) The train load is much larger than other types of loads on the bridge.

Firstly, the structure-health-monitoring system includes comprehensive sensors to collect a large amount of data; hence, it can provide a good foundation for training the neural networks for bridge response prediction. At the same time, the use of big data for the fitting of nonlinear relationships is the advantage of deep neural networks. The activation function within the LSTM network introduces nonlinear elements to neurons, enabling the neural network to effectively approximate a wide array of nonlinear functions.

Consequently, the LSTM network finds extensive utility in simulating nonlinear systems, thereby facilitating the fulfillment of Condition (1). Furthermore, when viewed through the lens of structural dynamics, the relationship between the excitation and response of a bridge is governed by the vibration equation. Typically, the intrinsic characteristics of a specific railway bridge structure, encompassing the geometric properties, material attributes, boundary conditions, and mechanical properties of railway tracks, remain relatively stable over a certain duration. This implies that the quality, damping, and stiffness of the bridge structure remain constant.

It becomes evident that the physical model governing the dynamic response of the bridge remains constant, establishing a direct correspondence between the load input and the response output, thereby fulfilling Condition (2). Subsequently, the dynamic load exerted on medium- and small-span railway bridges primarily comprises train load, wind load, and additional loads. For such bridges, the wind load is typically negligible when compared to the train load, especially if special loads such as earthquakes, ship impacts, ice pressures, frost heaving forces, and long rail breaking forces are not under consideration. Consequently, Condition (3) can be approximately satisfied. In summary, constructing predictive models for vehicle-induced the dynamic responses of small- and medium-span railway bridges based on LSTM networks appears to be feasible.

2.3. Proposed Framework for Predicting Train-Induced Bridge Responses

This section introduces the proposed framework for predicting train-induced bridge dynamic responses based on the designed LSTM network according to the vehicle–bridge coupling theory. According to the theory of structural dynamics, the vehicle–bridge coupling system is a multi-degree of freedom system. By solving the dynamic system equation [20], the response information of displacement, velocity, and acceleration of each degree of freedom can be obtained. Two types of information are required to solve the dynamic equation. The first one is the physical properties of the vehicle–bridge system, including the mass matrix $\boldsymbol{M}$, the damping matrix $\boldsymbol{C}$, and the stiffness matrix $\boldsymbol{K}$. The second is the load information of the vehicle–bridge system, including the load actioning on the bridge and the vehicle at any time. In short, the vehicle–bridge coupling equation constructs the nonlinear relationship between vehicle load and bridge responses according to the physical law of structural dynamics. Therefore, this study considers the load of each node of the bridge structure at each time as the input information to the neural network, so as to establish an LSTM network that predicts the train-induced bridge responses.

During the process of the train progressing from entering the bridge to leaving the bridge, the sampling numbers and frequency are N and f respectively, the speed is $\boldsymbol{V} = [V_1\ V_2\ V_3 \ldots V_N] \in \mathbb{R}^{1\times N}$, and the bridge span is L. Usually, the axle loads can be identified before the train enters the target bridge using some load identification methods and equipment [21]. Additionally, the initial axle position of the train (the axle position corresponding to the initial value of the response time series) $\boldsymbol{P}^*_{\text{opt}} \in \mathbb{R}^{1\times K^*}$ and the axle weight estimation value $\boldsymbol{A}^*_{\text{opt}} \in \mathbb{R}^{K^*\times 1}$ can be obtained from the structural-health-monitoring system. The bridge structure is divided into N_e unit structures on average, that is, there are $N_e + 1$ structural nodes, and the length of each unit structure is shown in Equation (12):

$$L_e = L/N_e \tag{12}$$

Under the assumption of a singular train traversing the bridge, with pertinent information regarding the train load, bridge structure, and sampling protocol available, it becomes feasible to derive the load value at any longitudinal position on the bridge structure at any given time during the data acquisition process. The neural network's input data encompass the train load, train speed, and the sampling frequency of the sensor, while the output data comprise dynamic response information pertaining to the train. The input data $\boldsymbol{X}_i \in \mathbb{R}^{N_x\times M_i}$ and the output data $\boldsymbol{Y}_i \in \mathbb{R}^{N_y\times M_i}$ of the i_{th} group data sample can be constructed as follows.

Assume the collected data of a train riding through the bridge as the i_{th} input and output pairs. The time step of the i_{th} group of data samples is $M_i = N$, time series $\boldsymbol{T} = [0, \frac{1}{f}, \frac{2}{f}, \ldots, \frac{n-1}{f}, \frac{n}{f}, \ldots, \frac{N-1}{f}] \in \mathbb{R}^{N\times 1}$. At the n_{th} time step $\boldsymbol{T}(n)$, the train position on the bridge, denoted as $\boldsymbol{P}_n \in \mathbb{R}^{1\times K^*}$, can be expressed as Equation (13):

$$\boldsymbol{P}_n = \boldsymbol{P}^*_{\text{opt}} + \sum_{i=1}^{n-1} \frac{1}{f}(V_n + V_{n+1})/2 \tag{13}$$

The train load matrix is defined as $\boldsymbol{\Pi}_w \in \mathbb{R}^{(N_e+1)\times N}$, where the element of row i and column j represents the axle load value at the time step j of the i_{th} structural node. The axle load of the k_{th} (k = 1, 2, 3, …, K^*) axle is $\boldsymbol{A}^*_{\text{opt}}(k)$, and the action position of $\boldsymbol{T}(n)$ is $\boldsymbol{P}_n(k)$. If the k_{th} axle load acts on the a_{th} node, the k_{th} axle load acts on each node at the n_{th} time step, as shown in Equation (14):

$$\boldsymbol{\Pi}^k_w(i,\ 1) = \begin{cases} \boldsymbol{A}^*_{\text{opt}}(k) & i = a \\ 0 & i \neq a \end{cases} \tag{14}$$

If the axle load operates between nodes a and a + 1, following the concept of equivalent load, the load can be partitioned into two components, as illustrated in Figure 2. The two components act independently on nodes a and a + 1, as expressed in Equation (15):

$$\boldsymbol{\Pi}^k_w(i,\ 1) = \begin{cases} m_1 \times \boldsymbol{A}^*_{\text{opt}}(k) & i = a \\ m_2 \times \boldsymbol{A}^*_{\text{opt}}(k) & i = a+1 \\ 0 & i \neq a \text{ and } i \neq a+1 \end{cases} \tag{15}$$

where m_1 and m_2 are axle load distribution coefficients and can be calculated according to Equations (16) and (17):

$$m_2 = \boldsymbol{P}_n(k)/L_e - floor[\boldsymbol{P}_n(k)/L_e] \tag{16}$$

$$m_1 = 1 - m_2 \tag{17}$$

In Equation (16), $floor(\cdot)$ denotes the operation of rounding down a real number to the nearest integer.

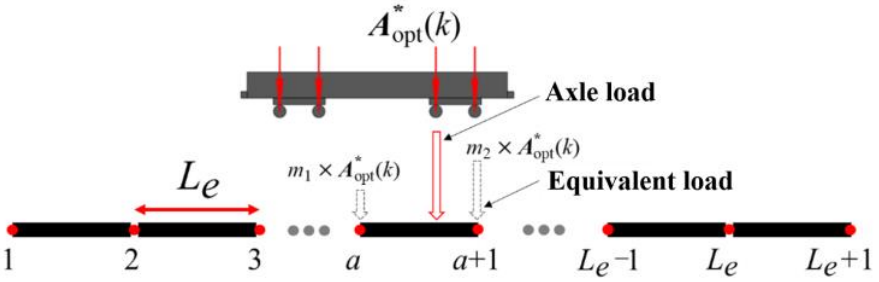


Figure 2. Equivalent decomposition of axial load.

Subsequently, the load exerted on each structural node at the n_{th} time step can be determined, as depicted in Equation (18):

$$\boldsymbol{\Pi}_w(:,\ n) = \sum_{k=1}^{K^*} \boldsymbol{\Pi}^k_w(:,\ 1) \tag{18}$$

After the load on each node at each time step is calculated, the train load matrix $\boldsymbol{\Pi}_w$ can be obtained. In the case of small- and medium-span railway bridges, the dynamic impact is relatively limited. Considering the scenario where the bridge experiences solely train load, the factors influencing the measured response of the bridge encompass bridge

stiffness, bridge damping, bridge mass, track irregularity, bridge damage, train load, train speed, and sensor equipment performance, among others. For a specific railway bridge within a given timeframe, the properties or characteristics of the bridge structure and track irregularity are assumed to remain relatively stable. This stability establishes a consistent one-to-one nonlinear relationship between load input and response output. In this study, the load, train speed, and sampling time series of the train acting on each node are used as the input data of the neural network model $\boldsymbol{X}_i \in \mathbb{R}^{N_x \times N}$, and the bridge structure response is used as the output data $\boldsymbol{Y}_i \in \mathbb{R}^{N_y \times N}$. The dimension N_x of the input data is $N_e + 3$, and the dimension N_y of the output data is 2. The response output includes displacement $\boldsymbol{R}_{\mathrm{Disp}}$ and acceleration $\boldsymbol{R}_{\mathrm{Acc}}$, as shown in Equations (19) and (20):

$$\boldsymbol{X}_i \;=\; [\boldsymbol{\Pi}_w;\; \boldsymbol{V};\; \boldsymbol{T}] \in \mathbb{R}^{(N_e+3)\times N} \tag{19}$$

$$\boldsymbol{Y}_i = [\boldsymbol{R}_{\mathrm{Disp}};\, \boldsymbol{R}_{\mathrm{Acc}}] \in \mathbb{R}^{2\times N} \tag{20}$$

In order to predict the train-induced dynamic responses of a railway bridge, this paper establishes a LSTM deep neural network structure. This network structure includes normalization layer, data input layer, LSTM layer, dropout layer, full connection layer, regression output layer, denormalization layer, etc. Taking the two-layer stacked LSTM network as an example, the neural network model architecture is shown in Figure 3. In order to eliminate the dimensional influence between different input data $\boldsymbol{X}_i$ and avoid the problem of gradient explosion, it is necessary to standardize the data to solve the comparability between data indicators and improve the speed of gradient descent and convergence. After the original data are processed by data standardization, each metric is in the same order of magnitude, that is, normalization. Given the uncertainty surrounding the train load impacting the railway bridge, the resulting vehicle-induced dynamic response data exhibit a degree of discreteness with indistinct boundaries. Consequently, the Z-Score method is employed to normalize the model data. The specific content of this method is to calculate the mean value $\boldsymbol{\mu}$ and standard deviation $\boldsymbol{\sigma}$ of the data sample, and then map the data to the normal distribution. Finally, the standardized data $\hat{\boldsymbol{X}}_i$ and $\hat{\boldsymbol{Y}}_i$ with the same input parameters at each time step conforming to the standard normal distribution (mean value is 0, standard deviation is 1) are obtained, as shown in Equation (21):

$$\left[\hat{\boldsymbol{X}}_i;\; \hat{\boldsymbol{Y}}_i\right] \;=\; \{[\boldsymbol{X}_i;\; \boldsymbol{Y}_i] - \boldsymbol{\mu}\} \oslash \boldsymbol{\sigma} \tag{21}$$

where $\hat{\boldsymbol{X}}_i \;=\; [\boldsymbol{x}_1\; \boldsymbol{x}_2\;\; \boldsymbol{x}_3 \ldots \boldsymbol{x}_{t-1}\;\; \boldsymbol{x}_t \ldots \boldsymbol{x}_N]$, $\boldsymbol{x}_t \in \mathbb{R}^{(N_e+3)\times 1}$ are the input data of the t_{th} time step of the i_{th} training sample and $\oslash$ is a special operation symbol, which represents the mathematical operation of the division of the position elements corresponding to the matrix or vector.

To enhance training efficiency and mitigate overfitting, it is common practice to incorporate a dropout layer following the LSTM layer [22]. The fundamental concept involves randomly disconnecting neuron connections during training and discarding nodes from the connection layer at a specified dropout rate. This process reduces training parameters, accelerates training speed, and prevents excessive sensitivity to particular cue fragments. Consequently, it diminishes the complex cooperative adaptation among neurons and enhances the model's generalization capability. In scenarios with limited data, feature discrimination can be augmented through sparsity. The operational mechanism of dropout technology is illustrated in Figure 4.

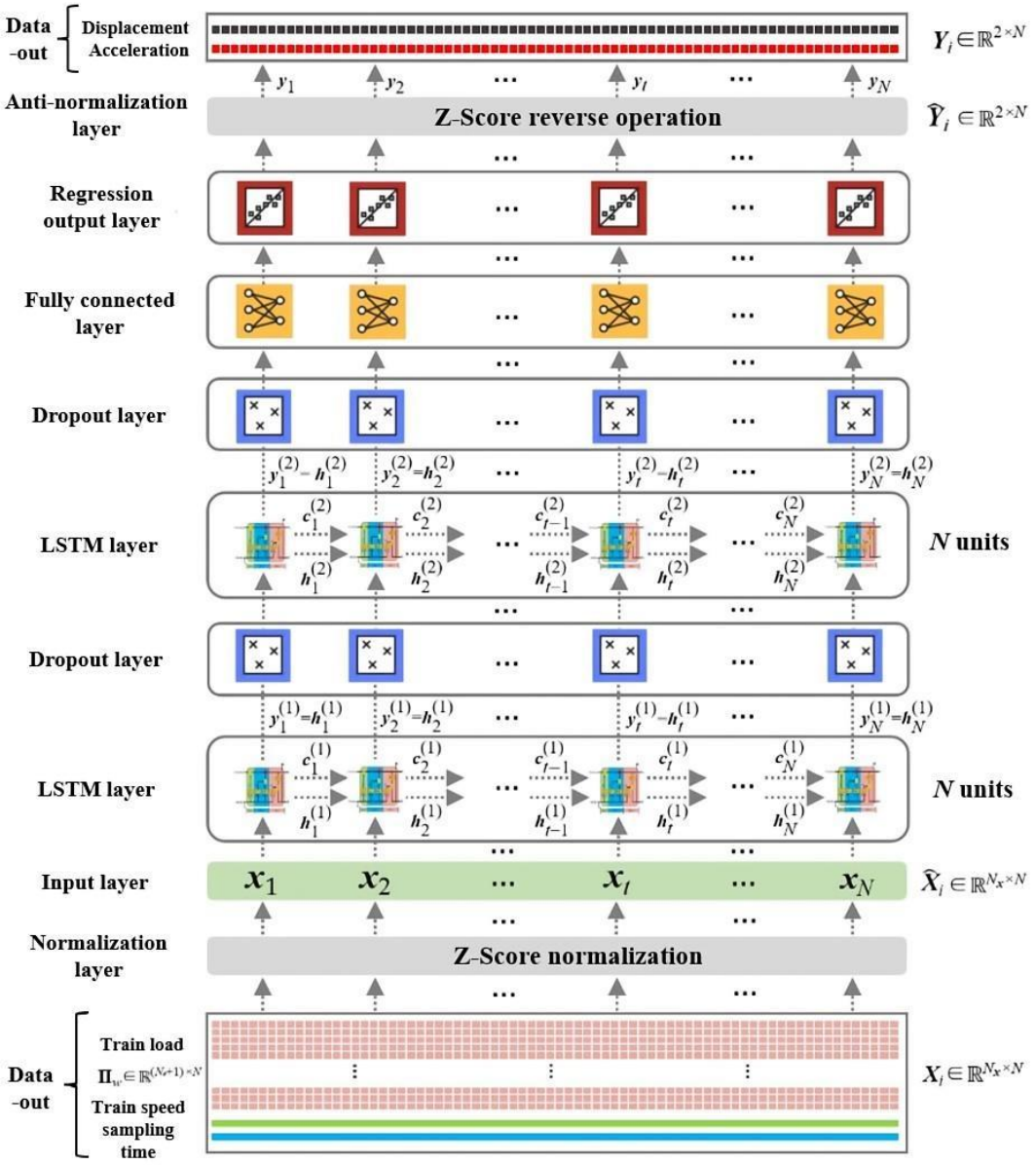
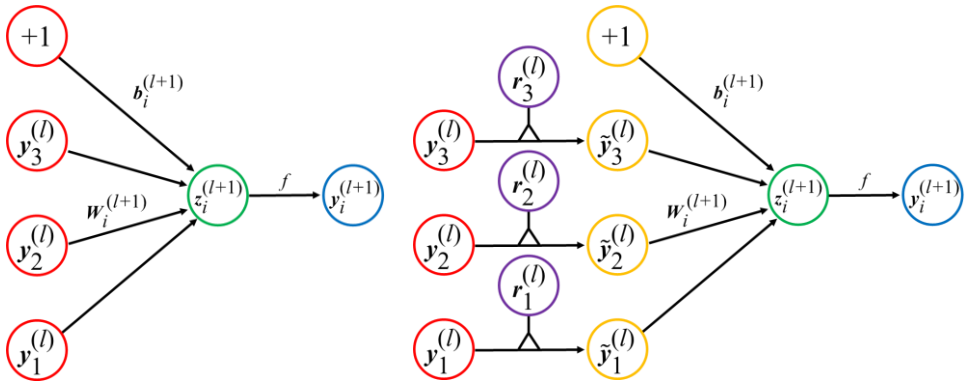


Figure 3. LSTM deep neural network architecture for vehicle-induced dynamic response prediction.



(**a**) standard network layer (**b**) dropout layer

Figure 4. The working principle of dropout layer.

3. Metrics for Evaluating Prediction Performance

The evaluation metrics for assessing the effectiveness of bridge structure response prediction models are primarily categorized into two groups [23]: dimensional and non-dimensional. Dimensional metrics encompass mean absolute error (MAE), mean squared error (MSE), and root mean squared error (RMSE). Non-dimensional metrics include the Pearson correlation coefficient, R-square, and absolute peak value error (APVE). Their specific definitions and expressions are outlined below.

3.1. Dimensional Metrics

The mean absolute error (MAE) signifies the average of the absolute differences between the predicted values and the observed values. A reduction in the MAE value indicates enhanced prediction accuracy, as depicted in Equation (22):

$$\mathrm{MAE}(\hat{\boldsymbol{y}},\ \boldsymbol{y}) \ = \ \frac{1}{N} \times \sum_{i\ =\ 1}^{N} |\hat{y}_i - y_i| \tag{22}$$

where $\hat{\boldsymbol{y}}$ and $\boldsymbol{y}$ are the predicted and true value vectors of structural response, and $\hat{y}_i$ and y_i (i = 1, 2, . . ., N) are the predicted and true values, respectively, of the i_{th} time step in the response sequence.

The mean squared error (MSE) represents the expected value of the squared differences between the predicted parameter value and the true parameter value. A smaller MSE value indicates higher accuracy in the model's predictive performance, as demonstrated in Equation (23):

$$\mathrm{MSE}(\hat{\boldsymbol{y}},\ \boldsymbol{y}) \ = \ \frac{1}{N} \times \sum_{i\ =\ 1}^{N} (\hat{y}_i - y_i)^2 \tag{23}$$

The RMSE is the arithmetic square root of the MSE and can be calculated according to Equation (24):

$$\mathrm{RMSE}(\hat{\boldsymbol{y}},\ \boldsymbol{y}) \ = \ \sqrt{\mathrm{MSE}(\hat{\boldsymbol{y}},\ \boldsymbol{y})} \tag{24}$$

It is evident that dimensional indices encompass both the magnitude of the value and the associated physical unit. While such indices directly reflect differences in prediction accuracy for the comparison of similar data types, they pose challenges when comparing different bridge responses due to the presence of multiple response types. In contrast, non-dimensional indices solely consist of value magnitudes without units, facilitating the comparison of prediction effects across various response types.

3.2. Non-Dimensional Metrics

The Pearson correlation coefficient serves to quantify the degree of linear correlation between two variables. A larger absolute value of the correlation coefficient indicates a stronger correlation. The coefficient approaches 1 or −1 for a stronger positive or negative correlation, respectively, while nearing 0 signifies a weaker correlation, as depicted in Equation (25):

$$\rho_{\hat{\boldsymbol{y}},\ \boldsymbol{y}} \ = \ \frac{\mathrm{cov}(\hat{\boldsymbol{y}},\ \boldsymbol{y})}{\sigma_{\hat{\boldsymbol{y}}}\sigma_{\boldsymbol{y}}} \ = \ \frac{E\left(\hat{\boldsymbol{y}}\ \boldsymbol{y}^{\mathrm{T}}\right) - E(\hat{\boldsymbol{y}})E(\boldsymbol{y})}{\sqrt{E\left(\hat{\boldsymbol{y}}^2\right) - \left(E(\hat{\boldsymbol{y}})\right)^2}\sqrt{E\left(\boldsymbol{y}^2\right) - \left(E(\boldsymbol{y})\right)^2}} \tag{25}$$

where $\mathrm{cov}(\hat{\boldsymbol{y}},\ \boldsymbol{y})$ represents the covariance between the predicted value and the true value and $E(\cdot)$ denotes the mean value. Generally, the strength of correlation between variables can be assessed based on the magnitude of the correlation coefficient.

The coefficient of determination, denoted as R^2, serves as a statistic for assessing the goodness of fit. The fundamental concept underlying R^2 involves employing the least squares method for parameter estimation, calculated as the ratio of the sum of regression squares to the sum of total deviation squares. The maximum value achievable for R^2 is

1, with a value closer to 1 indicating a superior fitting effect. Typically, a model with a goodness of fit exceeding 0.8 is considered relatively high, as expressed in Equation (26):

$$\mathrm{R}^2(\hat{\boldsymbol{y}}, \boldsymbol{y}) = 1 - \frac{\sum_{i=1}^{N}(\hat{y}_i - y_i)^2}{\sum_{i=1}^{N}(\overline{y} - y_i)^2} \tag{26}$$

where $\overline{\boldsymbol{y}} = E(\boldsymbol{y})$ is the mean value of the actual value of the data, which represents the reference value of the model.

The APVE is usually applied to evaluate prediction accuracy of the peak value, as shown in Equation (27):

$$\mathrm{APVE} = \frac{|\max(|\boldsymbol{y}|) - \max(|\hat{\boldsymbol{y}}|)|}{|\max(|\boldsymbol{y}|)|} \tag{27}$$

where $\hat{y}_{\mathrm{peak}}$ represents the response prediction value of the time step corresponding to the peak of the absolute value of the real value of the response data. Because of its normalized nature, the non-dimensional metrics are better suited for comparing prediction effects or accuracy differences across different types of responses.

3.3. Improved Evaluation Metrics

Among the aforementioned evaluation metrics, each time step's prediction error is assigned equal weight (i.e., 1). However, in practical engineering scenarios, the accuracy of peak responses often holds greater significance [24]. For instance, the maximum ground motion acceleration in a seismic event correlates with seismic intensity, and the maximum inter-story displacement angle serves as a crucial parameter in structural design. To enhance the emphasis on peak data values within the evaluation process while mitigating the impact of smaller data values, several weighted evaluation metrics are proposed. The weight of the response data corresponding to the i_{th} time step is shown in Equation (28):

$$w^{(i)} = |y_i|/\max(|\boldsymbol{y}|)\Big) \tag{28}$$

Upon introducing weights, the MAE, MSE, RMSE, and R^2 are, respectively, represented as the weighted absolute error (WMAE), weighted mean squared error (WMSE), weighted root mean squared error (WRMSE), and weighted determination coefficient ($\mathrm{R}^2_{\mathrm{W}}$). The calculation of the weighted metrics can be performed according to Equations (29)–(32):

$$\mathrm{WMAE}(\hat{\boldsymbol{y}}, \boldsymbol{y}) = \frac{\sum_{i=1}^{N} w^{(i)}|\hat{y}_i - y_i|}{\sum_{i=1}^{N} w^{(i)}} \tag{29}$$

$$\mathrm{WMSE}(\hat{\boldsymbol{y}}, \boldsymbol{y}) = \frac{\sum_{i=1}^{N} w^{(i)}(\hat{y}_i - y_i)^2}{\sum_{i=1}^{N} w^{(i)}} \tag{30}$$

$$\mathrm{WRMSE}(\hat{\boldsymbol{y}}, \boldsymbol{y}) = \sqrt{\mathrm{WMSE}(\hat{\boldsymbol{y}}, \boldsymbol{y})} \tag{31}$$

$$\mathrm{R}^2_{\mathrm{W}}(\hat{\boldsymbol{y}}, \boldsymbol{y}) = 1 - \frac{\sum_{i=1}^{N} w^{(i)}(\hat{y}_i - y_i)^2}{\sum_{i=1}^{N} w^{(i)}(\overline{y} - y_i)^2} \tag{32}$$

As per relevant tests [24], it has been demonstrated that the sensitivity of the weighted coefficient of determination to predicted response phase deviation is markedly diminished. Consequently, this enhances the overall scientific evaluation of structural response prediction.

4. Case Study

4.1. Engineering Background

A simply supported concrete box girder bridge with a span of 32 m was taken as the research object. The elastic modulus of the selected reinforced concrete material, denoted as

E, is 3.45×10^{10} N/m^2. The section area, denoted as A, is 8.89 m^2. The moment of inertia of section, denoted as I_y, is 10.95 m^4. Based on the modal analysis, the first and second natural frequencies of the target bridge are obtained, i.e., 5.12 Hz and 17.04 Hz, respectively. The bridge was divided into N_e structural elements; thus, N_e structural nodes can be obtained in the established finite element model.

The German Intercity-Express (ICE) multiple unit train, comprising two power cars and several trailers, is taken as the research object. The power cars are positioned at the front and rear of the train, while the trailers are situated in the middle. The axle spacing dimensions of both the trailers and power cars in each train formation are standardized, ensuring consistent fixed axle spacing l_1, vehicle pitch l_2, and car body length l_3 within the same train. The individual axle loads of the power cars and trailers are 16 tons and 14.6 tons, respectively. The vehicle pitch l_2 is 17.375 m, and the length of the carriage l_3 is 24.775 m. The German low-disruption track vertical irregularity spectrum is employed as the input excitation. The spatial frequency Ω ranges from 0.05 to 3 rad/m. The random spatial frequency is 500. The layout of the train, track, and bridge system is illustrated in detail in Figure 5.

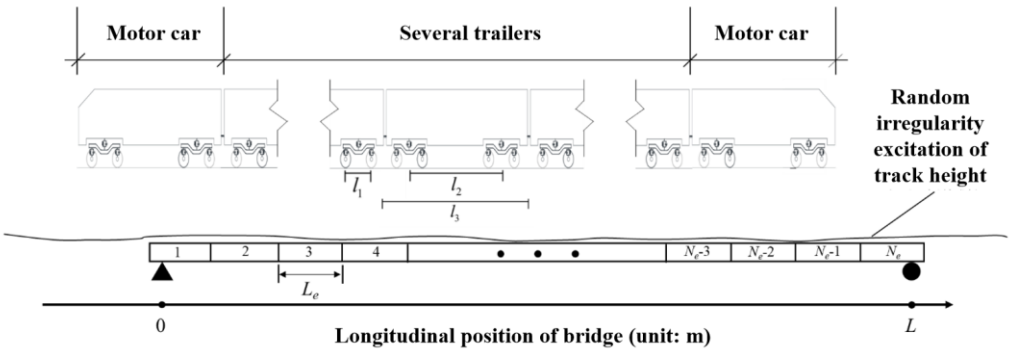


Figure 5. The layout of the train, track, and bridge system.

Series finite element simulations were carried out to obtain the dataset of train loads and structural responses for training the LSTM network. The train loads include the time history of axle loads, axle spacings, and axle positions. The structural responses include the displacement and acceleration of the bridge. The finite element model of the bridge comprises 300 units. The LSTM network takes train loads, train velocity, and sampling time series as inputs and vertical displacement and acceleration responses at the mid-span section of the bridge as outputs.

This section aims to investigate the influencing factors, including the neural network structure, training sample size, number of bridge units, input noise, train speed, and train composition, on the prediction accuracy of train-induced responses utilizing the LSTM network. Initially, this study investigates the impact of the number of LSTM layers and units in hidden layers on the prediction accuracy. The dataset used for training neural network model is defined as dataset C, comprising 100 samples covering various train speeds, train compositions, train spacings, and sampling frequencies, as tabulated in Table 1.

Table 1. Description of the training datasets.

Dataset Name	Train Speed/(km/h)	Number of Carriages	Fixed Wheel Base l_1/m	Sampling Frequency/Hz
A: sample size is 25	20, 40, 60	4, 8, 12	1.8, 2.5	100, 200
B: sample size is 50	20, 40, 60, 80	4, 8, 12, 16	1.8, 2.5	100, 200

Table 1. *Cont.*

Dataset Name	Train Speed/(km/h)	Number of Carriages	Fixed Wheel Base l_1/m	Sampling Frequency/Hz
C: sample size is 100	20, 40, 60, 80, 100	4, 8, 12, 16, 20	1.8, 2.5	100, 200
D: sample size is 150	20, 40, 60, 80, 100	4, 8, 12, 16, 20	1.8, 2.5	100, 150, 200
E: sample size is 200	20, 40, 60, 80, 100	4, 8, 12, 16, 20	1.8, 2.2, 2.5	100, 150, 200

The accuracy of predicted bridge responses using the same neural network structures with different data sizes are calculated. The training dataset sizes for the neural network range from 25 to 200, denoted as A to E, respectively. In addition, the influence of other factors, i.e., train speed, the number of carriages, and axel space, are also investigated. The detailed information for these datasets is provided in Table 1.

4.2. Network Determination and Training

4.2.1. Training Strategies

In this study, the design and training of neural networks are conducted using the Deep Network Designer and Deep Learning Toolbox 14.5 toolboxes within MATLAB. As for computer hardware specifications, the CPU comprises an Intel(R) Core (TM) i7-9750H processor with 6 cores and 12 threads, operating at 2.60 GHz. Additionally, the GPU employed is an NVIDIA GeForce GTX 1650.

The number of parameters of the target neural network can be estimated according to Equation (10). In the case of a two-layer LSTM network with 128 hidden units and an input data dimension of 203, the total number of parameters to be estimated is 339,968, i.e., $2 \times 4 \times (128 \times (128 + 203) + 128)$. The Adam optimizer is employed to update model parameters, leveraging its advantage of adaptive learning rates. It only requires the setting of an initial learning rate, thus alleviating the need for any manual adjustment of learning rates during the training process. However, it is not advisable to set the initial learning rate too high, as this may lead to convergence to local optima.

In this study, the initial learning rate is set to 0.001. The learning rate strategy is configured as piecewise, with a learning rate drop period of 50 and a learning rate drop factor of 0.5. The data shuffling strategy is set to shuffle once per epoch. Additionally, based on existing research experiences, the maximum number of training epochs, denoted as MaxEpochs, is set to 200, and the amount of data used per iteration, denoted as MiniBatchSize, is set to 20. Assuming that there are 100 samples in the dataset, it can be divided into 5 iterations per training epoch, i.e., 100 divided by 20 equals 5. The total number of iterations is 1000. The iterative steps for updating model parameters can be summarized as follows:

Step 1: Build the LSTM model and initialize its parameters. Select 5 samples as the first batch;

Step 2: Gradually propagate neural network information forward and downward through layers to obtain the prediction;

Step 3: Calculate the loss function value based on the prediction and the theoretical value. Perform gradient descent and update iterative model parameters;

Step 4: Input the second batch of sample data. Continue model parameter iteration according to Step 2 and Step 3;

Step 5: Cycle through inputting each batch of sample data until all batches in the dataset have been input and updated. At this point, one epoch or one round of model training is completed;

Step 6: Iterate through Steps 1 to 5 until either the maximum number of training epochs is reached or the training error reaches its minimum value. Upon reaching this criterion, the model training concludes, and the model becomes ready for response prediction.

4.2.2. Network Structures

Previous studies have demonstrated that there exists a neural network with an intermediate layer between the input and output layers that is capable of fitting any function [25]. In general, when the other parameters are properly set, higher numbers of layers and neurons tend to result in greater accuracy. However, an excessive number of layers and neurons often leads to overfitting. Therefore, selecting the appropriate numbers of layers and neurons in the LSTM network is crucial for improving model accuracy and avoiding overfitting. Based on previous relevant studies [26–28], this study selects LSTM network layers of 1, 2, and 3 and neuron numbers of 32, 64, 128, and 256, forming a total of 12 neural network structures, as detailed in Table 2.

Table 2. Network models with different numbers of LSTM layers and neurons.

Model Name	LSTM Layers	Number of Neurons	Model Name	LSTM Layers	Number of Neurons
1-32	1	32	2-128	2	128
1-64	1	64	2-256	2	256
1-128	1	128	3-32	3	32
1-256	1	256	3-64	3	64
2-32	2	32	3-128	3	128
2-64	2	64	3-256	3	256

The dataset C, as shown in Table 2, was utilized for training the LSTM network model. After 200 training epochs, a converged neural network model was obtained, and the representative loss function curve is depicted in Figure 6. Based on Bayesian inference theory [29], the load information of the target train riding on the railway bridge was identified. The driving speed of the target train with 4 carriages and a fixed axle spacing of 2.3 m is 50 km/h. Specific results are presented in Figure 7. Subsequently, vehicle-induced dynamic responses were then predicted using the obtained LSTM network.

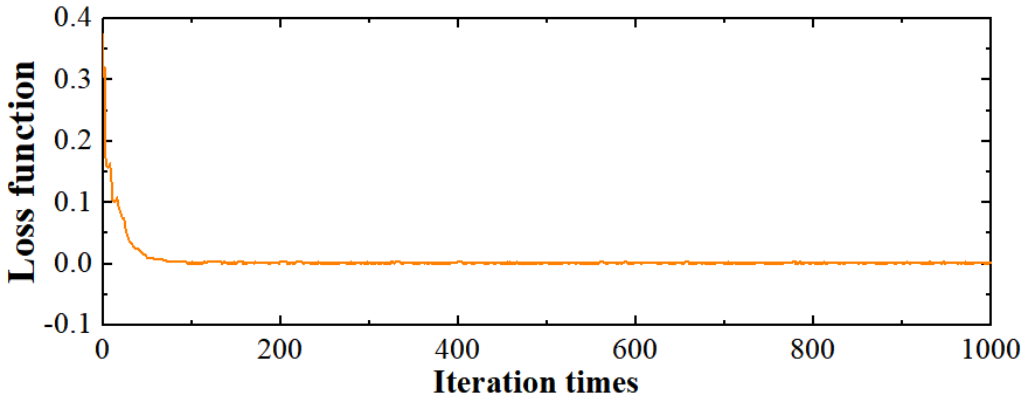


Figure 6. Training loss function curve of the established LSTM network models.

The performance of different LSTM network models, as listed in Table 2, for the prediction of railway bridge displacements are compared. The dimensional and non-dimensional evaluation metrics were applied to assess the prediction accuracy. Results are presented in Figure 8. Non-dimensional metrics include Pearson correlation coefficient, the weighted coefficient of determination R_W^2, and APVE. Dimensional metrics include WMAE, RMSE, and WRMSE.

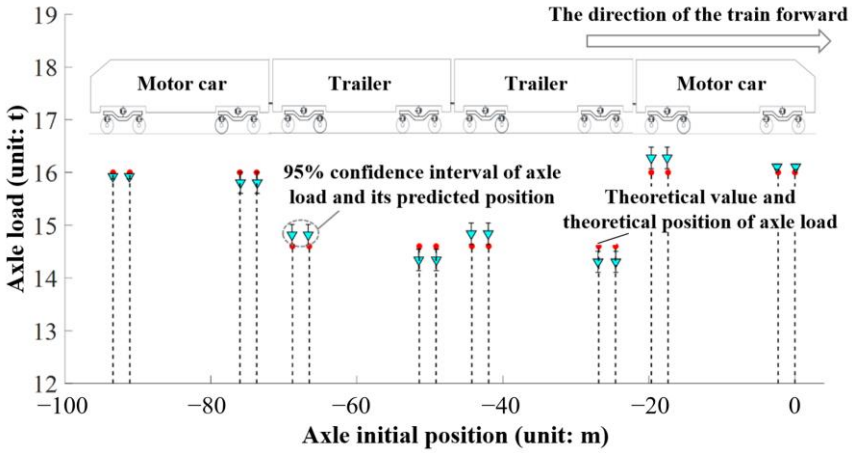
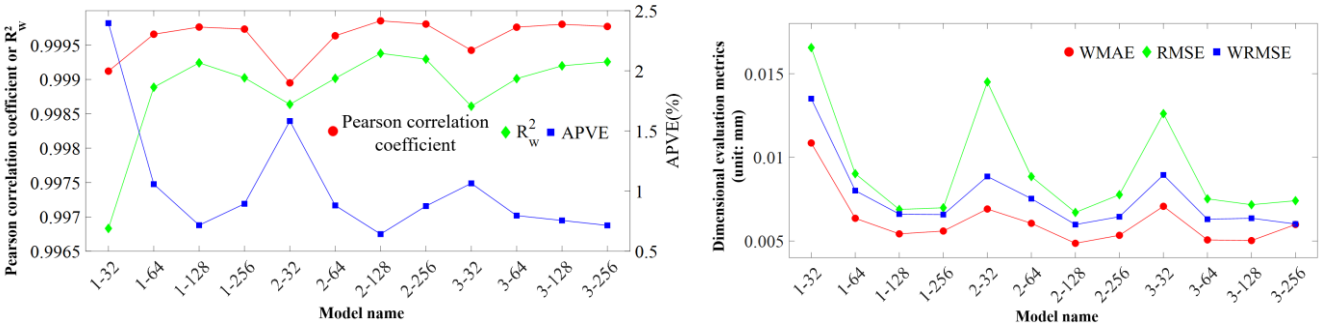


Figure 7. Train load identification results of a railway bridge.

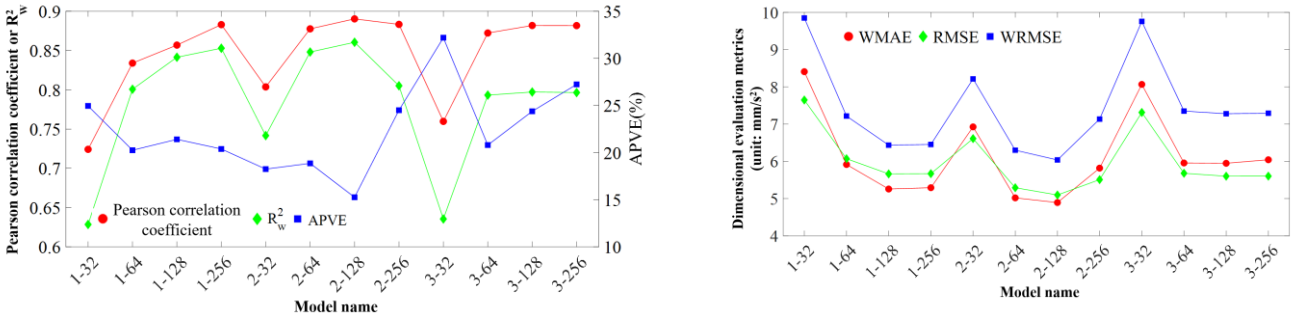


(**a**) Non-dimensional evaluation metrics (**b**) Dimensional evaluation metrics

Figure 8. The influence of LSTM network structure on the prediction accuracy of bridge displacement.

As shown in Figure 8a, the Pearson correlation coefficient and R_W^2 of the aforementioned LSTM network models remain above 0.9965, indicating a significant correlation between the predicted and theoretical values of bridge displacement. The overall APVE values are below 2.5%, suggesting good accuracy in predicting peak displacement values for the bridge. As illustrated in Figure 8b, the dimensional evaluation metrics of the LSTM network models are generally below 0.015 mm, indicating good absolute accuracy in predicting bridge displacement. Overall, for the LSTM network with the same number of hidden layers, the model's prediction accuracy gradually improves with an increased number of neurons. The optimal performance is achieved when the number of neurons reaches 128. However, when the number of neurons reaches 256, there is a slight decrease in prediction accuracy. This may be attributed to overfitting caused by the enhanced nonlinearity of the model, which can be mitigated by employing additional dropout or regularization techniques [30].

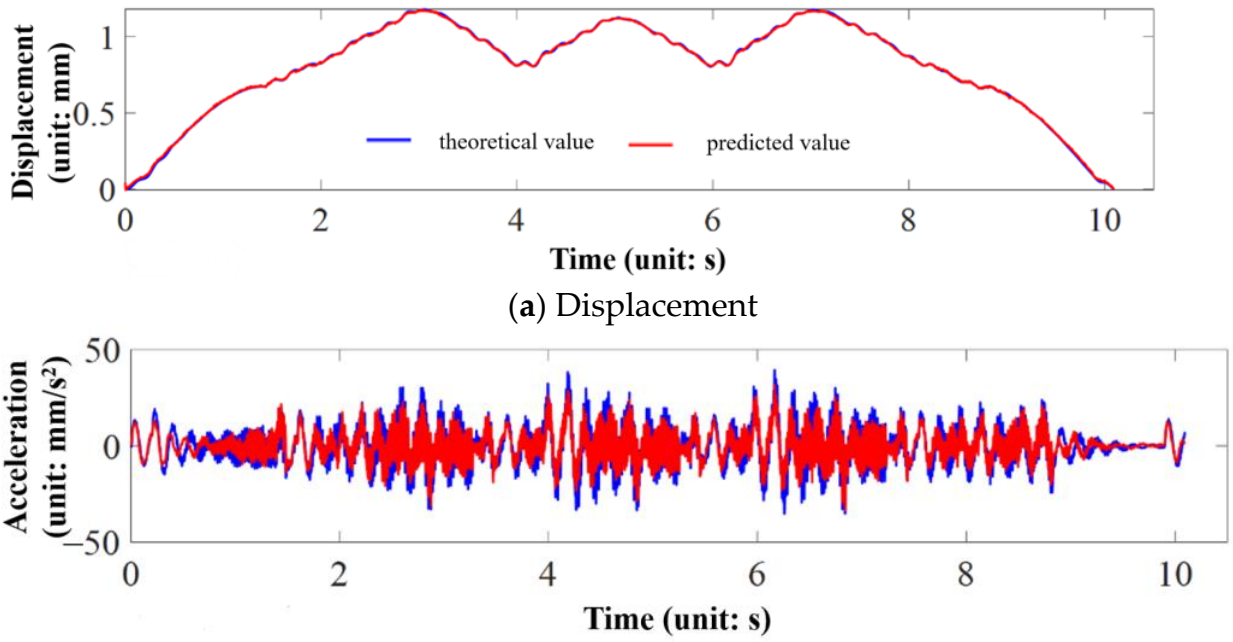
The dimensional and non-dimensional evaluation metrics for assessing the acceleration prediction performance using LSTM network models as listed in Table 2 are calculated and listed in Figure 9. As illustrated in Figure 9a, both the Pearson correlation coefficient and R_W^2 of each model fall within the range of 0.6 to 0.9, while the APVE ranges between 15% and 35%. In terms of dimensional evaluation metrics, values range from 4 to 10 mm/s^2. These fluctuations in metrics suggest that the accuracy of acceleration prediction is lower compared to displacement prediction. Maintaining a constant number of LSTM network layers, increasing the number of neurons gradually enhances acceleration prediction accuracy. However, the trend of APVE changes is not consistently apparent. Specifically, as the number of neurons increases from 128 to 256, certain metrics exhibit a downward trend, indicating a decrease in prediction accuracy.



(**a**) Non-dimensional evaluation metrics (**b**) Dimensional evaluation metrics

Figure 9. The influence of different neural network structures on the accuracy of acceleration prediction.

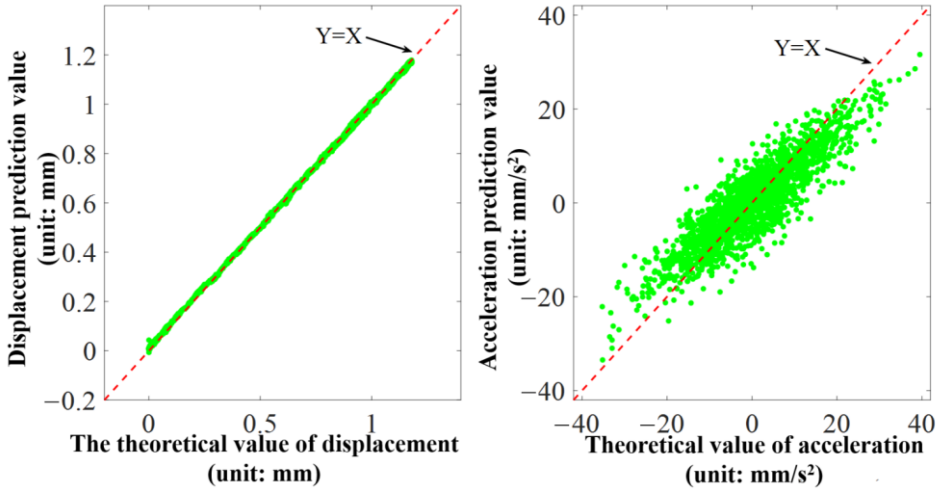
In summary, the "2-128" model demonstrates excellent predictive performance. For displacement prediction, it achieves a Pearson correlation coefficient of 0.99985 and an APVE of 0.64%, while, for acceleration prediction, it yields a Pearson correlation coefficient of 0.89029 and an APVE of 15.26%. Figures 10 and 11 illustrate the theoretical and predicted values of displacement and acceleration under this network structure. It is evident that displacement prediction accuracy notably surpasses that of acceleration. Although the peak values of predicted acceleration are slightly lower than the theoretical values, overall, the predicted results exhibit a strong correlation with the theoretical values. To delve deeper into the prediction results, Figure 12 offers a comparison of the power spectral density of displacement and acceleration. The findings indicate that the principal frequency components of the predicted response in the frequency domain align closely with the theoretical values, particularly in proximity to the quasi-static component near 0 Hz and the first-order resonance frequency, that is, 5.12 Hz.



(**a**) Displacement

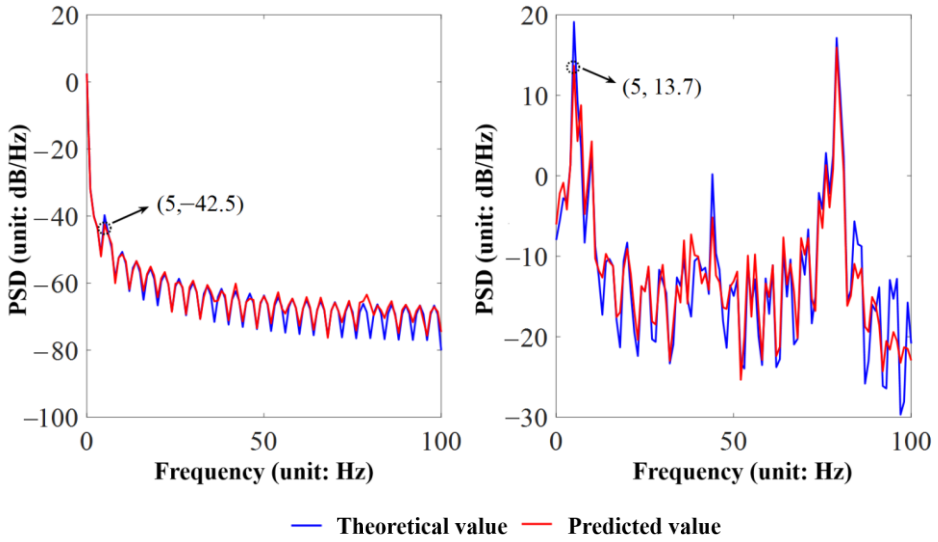
(**b**) Acceleration

Figure 10. Predicted vehicle-induced dynamic response using neural network structure '2-128'.



(**a**) Displacement (**b**) Acceleration

Figure 11. Theoretical and predicted bridge responses using neural network structure '2-128'.



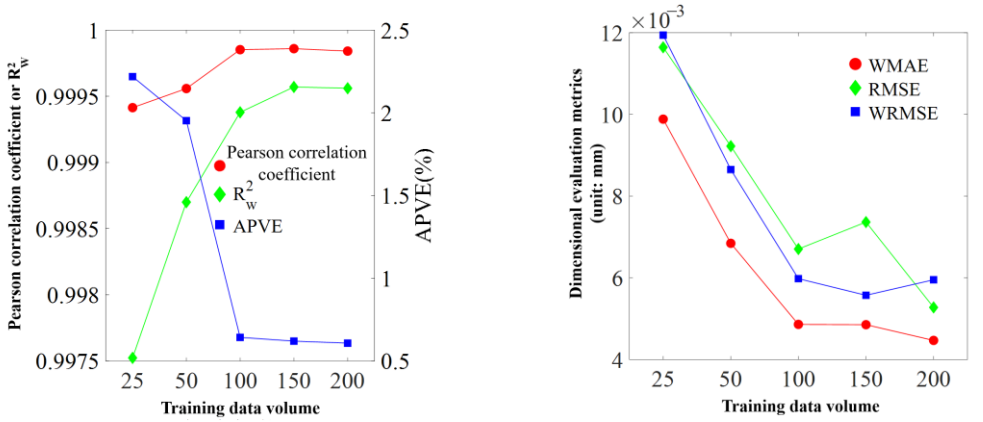
(**a**) Displacement (**b**) Acceleration

Figure 12. The PSD of predicted bridge responses using neural network structure '2-128'.

4.2.3. The Size of Training Dataset

The prediction accuracy of an LSTM network model is usually affected by many factors, including the structure of the neural network, network parameters, training sample size, and so on. As descripted in Section 4.2.2, the neural network model with 2 LSTM layers and 128 neurons can achieve relatively good prediction accuracy. On this basis, this section further discusses the influence of the size of training dataset on the prediction accuracy. The detailed information of the applied datasets, i.e., A to E, can be found in Table 1.

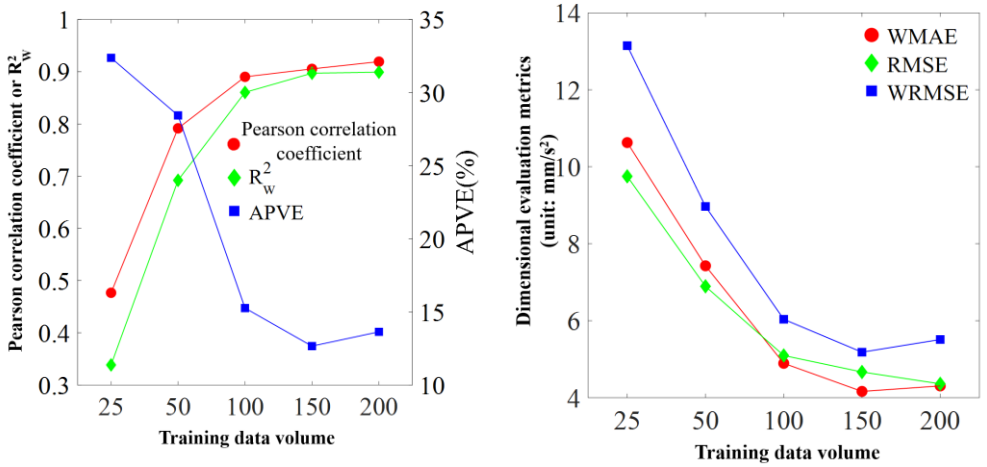
The evaluation metrics of the predicted bridge displacements across datasets A to E are illustrated in Figure 13. The results demonstrate consistently high levels of Pearson correlation coefficient and for all datasets, indicating a robust correlation between predicted displacement and theoretical values. Moreover, the influence of individual samples appears negligible. The APVE is consistently maintained between 0.5% and 2.5%, suggesting low peak error levels. Additionally, all dimensional metrics, including WMAE, RMSE, and WRMSE, range from 0.004 to 0.012 mm, further indicating excellent prediction accuracy.



(**a**) Non-dimensional evaluation metrics (**b**) Dimensional evaluation metrics

Figure 13. The impact of training sample quantity on displacement prediction accuracy.

The evaluation metrics of the predicted bridge acceleration responses across datasets A to E are illustrated in Figure 14. As shown in the figure, the Pearson correlation coefficient and R_W^2 ranged from 0.3 to 0.95. This shows that the prediction accuracy of acceleration is more sensitive than that of displacement. At the same time, the APVE ranges between 10% and 35%, indicating a larger error level of the acceleration than that of the displacement. The dimensional parameters including WMAE, RMSE, and WRMSE range between 4 mm/s^2 and 14 mm/s^2. When the size of the training dataset increases from 25 to 100, the evaluation metrics demonstrate a gradual improvement in prediction accuracy. However, as the size of the training dataset further increases from 100 to 200, the change in each evaluation metrics becomes less apparent, and, in some cases, even tends to increase. This observation suggests that increasing the number of training samples does not consistently enhance the prediction effectiveness of the model.



(**a**) Non-dimensional evaluation metrics (**b**) Dimensional evaluation metrics

Figure 14. The impact of training sample quantity on acceleration prediction accuracy.

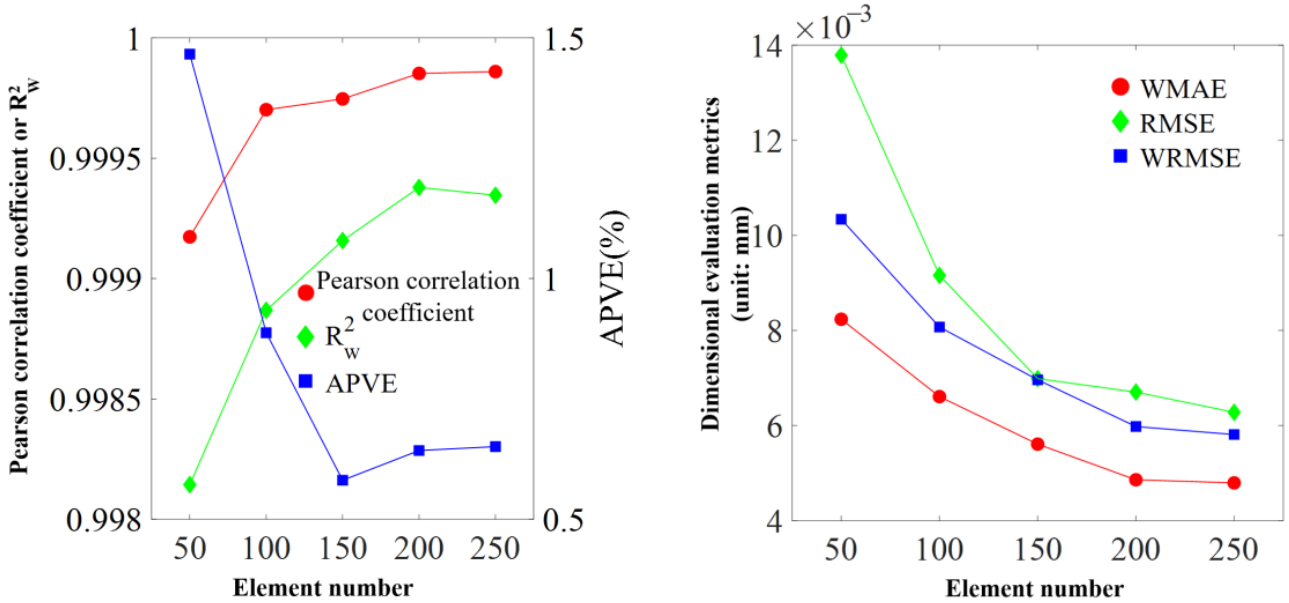
This phenomenon may be attributed to the quality of the dataset. An examination of Table 1 reveals significant differences in working conditions between dataset A (25 samples), dataset B (50 samples), and dataset C (100 samples). However, dataset D (150 samples) and dataset E (200 samples) exhibit relatively minor variations in working conditions compared to dataset C. Specifically, the changes in working conditions for dataset D and dataset E are minimal, with only slight adjustments in sampling frequency and train fixed wheelbase l_1 compared to dataset C.

Indeed, despite the increased size of datasets D and E, the quality enhancement of these datasets is not readily apparent due to the repetition or similarity of working conditions. Consequently, the improvement in model prediction effectiveness is not evident and, in some cases, may even decrease. Therefore, when constructing the dataset, it is crucial to gather a wide range of potential working conditions while avoiding the collection of repeated or similar working conditions. This approach ensures the development of a high-quality training dataset. By doing so, the LSTM network model can undergo comprehensive training and learning, ultimately enhancing the predictive effectiveness of the model.

4.2.4. The Number of Discretized Bridge Elements

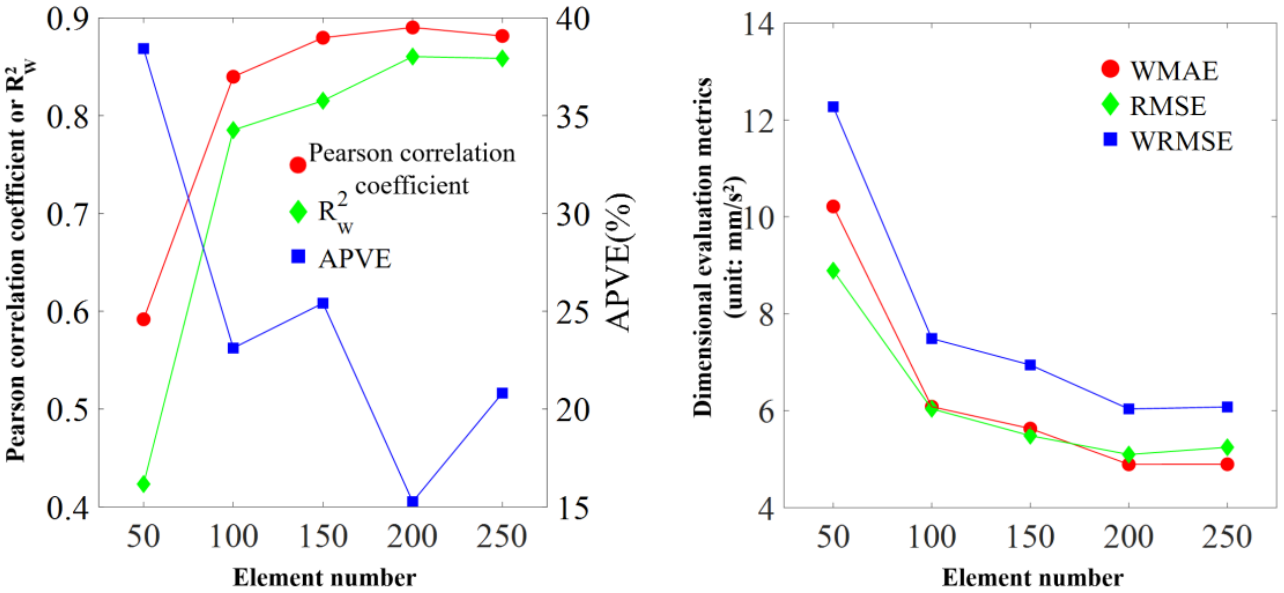
As mentioned in Section 2.3, the dimension of input data for the LSTM model is contingent upon the number of divided elements N_e of the bridge structure. When N_e is too small, accurately reflecting the stress state of the bridge structure becomes challenging. Conversely, a large value of N_e may result in increased input data volume and significantly reduced computational efficiency. Hence, it is essential to ascertain the optimal value of N_e to strike a balance between prediction accuracy and computational efficiency.

The evaluation metrics of predicted displacement and acceleration for N_e, ranging from 50 to 250, are depicted in Figures 15 and 16. The results indicate a gradual increase in both the Pearson correlation coefficient and R_W^2 with the increased N_e, suggesting an enhanced correlation between predicted and theoretical values. Furthermore, APVE demonstrates a downward trend in error values as N_e increases. The dimensional evaluation metrics exhibit a declining trend as N_e increases from 50 to 150. However, the rate of decline becomes slightly more moderate when N_e increases from 150 to 250. Hence, the optimal N_e in this study is determined to be 150, with the corresponding length of the bridge element set at 0.21 m.



(**a**) Non-dimensional evaluation metrics (**b**) Dimensional evaluation metrics

Figure 15. The influence of the number of bridge structure elements on displacement prediction accuracy.



(**a**) Non-dimensional evaluation metrics (**b**) Dimensional evaluation metrics

Figure 16. The influence of the number of bridge structure elements on acceleration prediction accuracy.

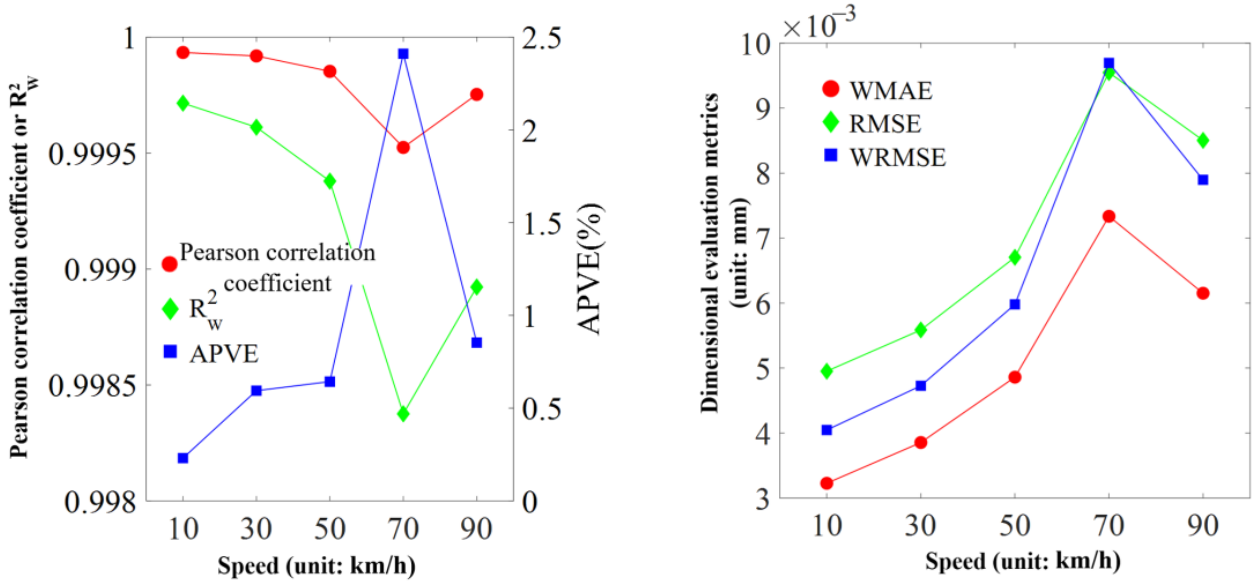
4.3. Effect of the Train System on the Prediction Accuracy

4.3.1. The Riding Speed of the Train

This section aims to investigate the impact of train riding speed on the prediction accuracy of bridge dynamic responses. The study utilizes the German ICE EMU, comprising four carriages. The fixed wheelbase l_1 is 2.3 m, the vehicle spacing l_2 is 17.375 m, and the vehicle body length l_3 is 24.775 m. The bridge response sampling frequency is set at 150 Hz, and the number of bridge units is 200. For the LSTM network model is trained using dataset C with 100 samples. The neural network model structure comprises 2 layers of LSTM with 128 neurons.

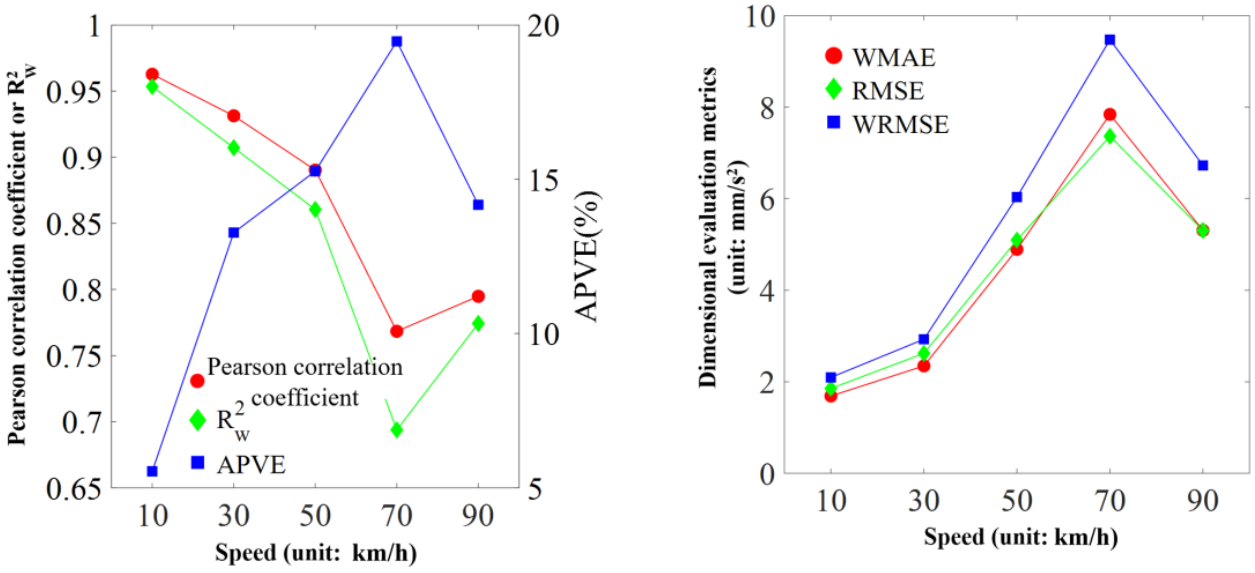
The evaluation metrics of predicted bridge displacement and acceleration with the varying train speed are depicted in Figures 17 and 18. The results indicate that the Pearson correlation coefficient and R_W^2 for displacement prediction remain consistently high, with values exceeding 0.998. For acceleration prediction, the Pearson correlation coefficient and R_W^2 range between 0.7 and 0.95. The APVE values for displacement prediction range from 0.2% to 2.5%, while those for acceleration prediction range from 5% to 20%. In terms of dimensional indices, the displacement ranges from 0.003 mm to 0.01 mm, while the acceleration ranges from 2 mm/s^2 to 10 mm/s^2. Additionally, one can observe that all the evaluation metrics show mutation and obvious deterioration when the train speed reaches 70 km/h. The cause of this phenomenon may be the resonance effect in the vehicle–bridge coupling vibration [31].

Overall, the accuracy of predicted train-induced bridge responses tends to decrease with the increased train speed. The prediction of acceleration is notably sensitive to changes in train speed, potentially resulting in larger prediction errors as speed increases. In contrast, displacement prediction maintains consistent performance across varying train speeds, demonstrating good robustness with train speed.



(**a**) Non-dimensional evaluation metrics (**b**) Dimensional evaluation metrics

Figure 17. The impact of train speed on displacement prediction accuracy.

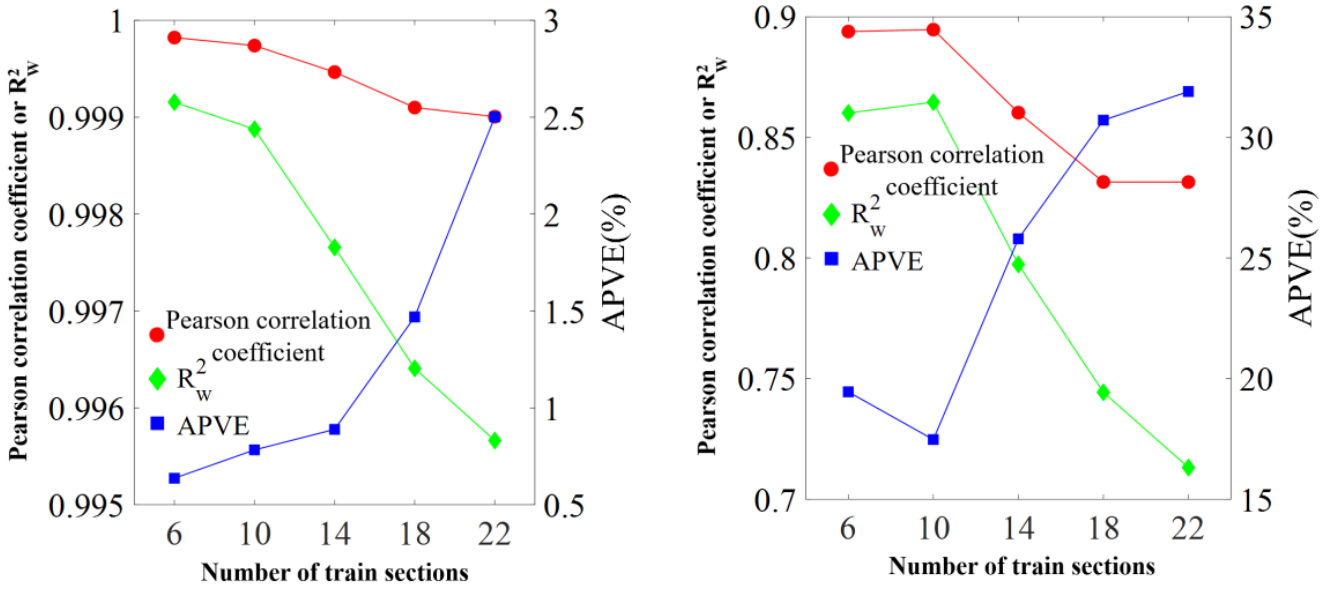


(**a**) Non-dimensional evaluation metrics (**b**) Dimensional evaluation metrics

Figure 18. The impact of train speed on acceleration prediction accuracy.

4.3.2. The Number of Train Carriages

Based on the previous analysis, the performance of predicting the train-induced bridge responses with a varying number of carriages will be discussed in this section. Figure 19 shows the non-dimensional metrics for evaluating the prediction performance of the displacement and acceleration of a railway bridge as the number of carriages varies from 6 to 22. The riding speed of the train is set as 50 km/h. The results show that the increase in the carriage number causes a similar changing trend with the train speed. To be specific, the increase in carriage number leads to a decrease in prediction accuracy. The error level of displacement prediction is relatively stable, and the acceleration prediction is more sensitive to the change in train sections.



(**a**) Displacements (**b**) Acceleration

Figure 19. The influence of the carriage number on prediction performance.

4.4. The Influence of Measurement Noise

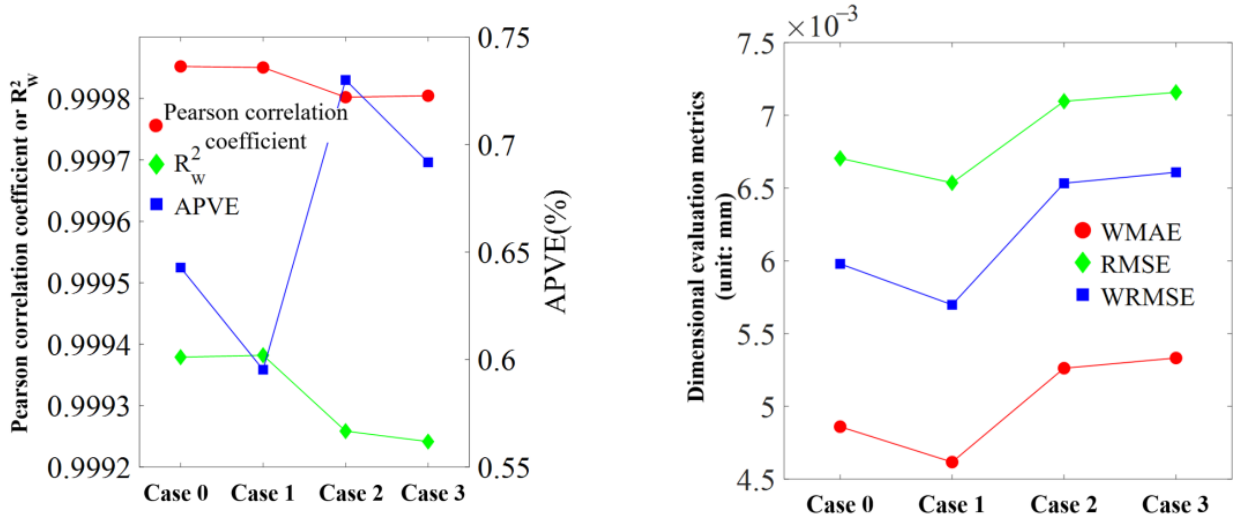
As mentioned in the previous sections, the established LSTM model maps the one-to-one relationship between train loads and bridge responses. The train load information is usually obtained through the installed sensors, which may inevitably include measurement noise [32]. If the established LSTM model is sensitive to the measurement noise, the accuracy and robustness will be weakened. Therefore, this section explores the influence of noise level of monitored axle load and wheelbase on the prediction accuracy. Four testing cases are designed and tabulated in Table 3.

Table 3. The design of testing cases with different noise levels.

Name	Description of the Noise Level
Case 0	Both axle load and axle spacing adopt the optimal values identified by load recognition.
Case 1	Axle spacing adopts the optimal value identified by load recognition, while axle load values are randomly generated based on their uncertainty results.
Case 2	Axle load adopts the optimal value identified by load recognition, while axle spacing adds 5% noise.
Case 3	Axle load values are randomly generated based on their uncertainty results, while axle spacing adds 5% noise.

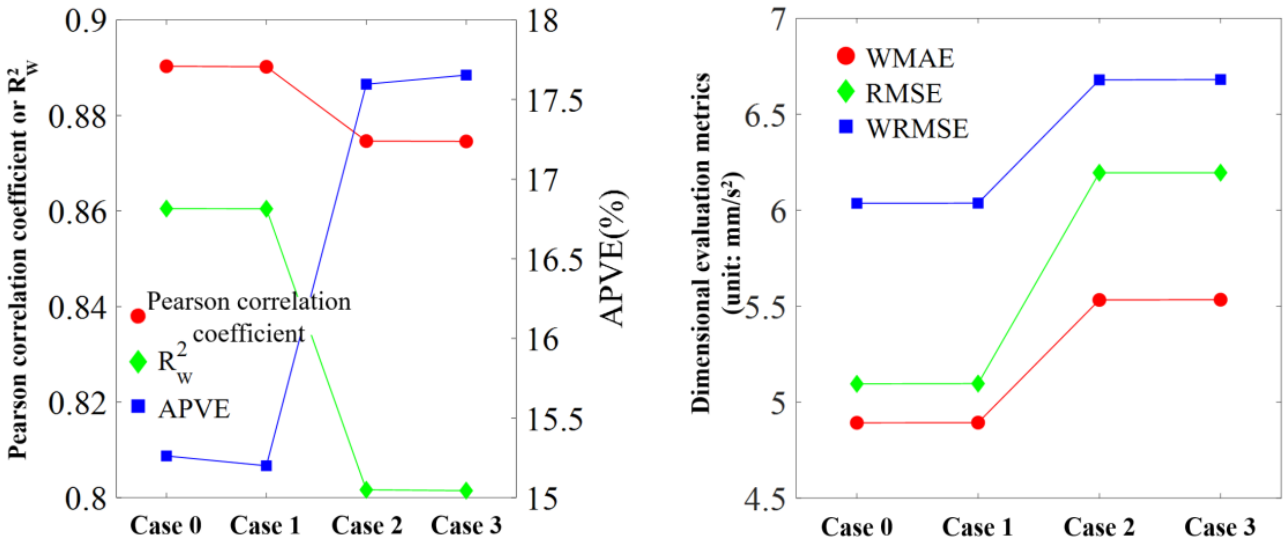
The specific method of adding 5% noise involves adding random errors with a maximum absolute error of 5% based on the optimal axle spacing.

The evaluation metrics of the predicted displacement and acceleration under different noise cases are shown in Figures 20 and 21. The results show that the Pearson correlation coefficient and R_W^2 decrease obviously with the noise included. The decrease of the evaluation metrics of displacement is smaller that of the acceleration. Specifically, the Pearson correlation coefficient and R_W^2 of displacement prediction stay above 0.999, while those of acceleration decreased by a large margin, but still remained above 0.8. This indicates that the predicted values of displacement and acceleration with limited input noise always have a strong correlation with the theoretical values.



(**a**) Non-dimensional evaluation metrics (**b**) Dimensional evaluation metrics

Figure 20. The impact of input noise on displacement prediction accuracy.



(**a**) Non-dimensional evaluation metrics (**b**) Dimensional evaluation metrics

Figure 21. The impact of input noise on acceleration prediction accuracy.

Regarding APVE, both the evaluation metrics for displacement and acceleration increase after adding noise, with the maximum increase being approximately 0.1% and 2%, respectively. In terms of dimensional indicators, each evaluation metric shows an increase after noise addition, but the increase in acceleration evaluation metrics is greater than that of displacement. This suggests that acceleration prediction is more sensitive to noise compared to displacement prediction. Furthermore, transitioning from Case 0 to Case 1, or from Case 2 to Case 3, results in minimal changes in each metric. However, there is a noticeable variation trend in the evaluation indices between Case 2 and Case 3 compared to Case 0 and Case 1.

In summary, the noise associated with axle load has minimal impact on response prediction, whereas noise related to wheelbase has a more pronounced effect. Nonetheless, overall, there remains a strong correlation between predicted response values and theoretical values, with errors consistently falling within an acceptable range. Particu-

larly noteworthy is the robust performance of displacement prediction, which maintains accuracy even when influenced by noise.

5. Conclusions

In this paper, an LSTM modeling method based on physical mechanism guidance was proposed to predict the dynamic response of vehicle-bridge coupling. The principle of LSTM was introduced, and the conditions required for the method to be used in the surrogate model of vehicle-induced dynamic response analysis were explained. Subsequently, the corresponding LSTM deep neural network structure was constructed. Furthermore, the feasibility of the above method was verified by a small- and medium-span concrete box section of a simply supported railway bridge, and the influence of various factors on the response prediction was analyzed. The main conclusions were as follows:

(1) In terms of the prediction accuracy of different types of responses, in general, the accuracy of displacement prediction was obviously better than that of acceleration, and the prediction of displacement was less affected by various unfavorable factors than that of acceleration. The peak value of acceleration prediction value was smaller than that of theoretical value, and the prediction results were generally strongly correlated with the theoretical value. Therefore, displacement prediction had higher reliability than acceleration prediction;

(2) For different dimensional indicators, under the same number of LSTM network layers, as the number of neurons increased, the model evaluation indicators tended to be better, that is, the accuracy of the prediction results tended to increase. However, with the increase in the number of parameters of the neural network model, the nonlinear fitting performance of the model was improved, and it was easy to produce over-fitting phenomena, which led to a decrease in prediction accuracy. When the number of neurons increased from 128 to 256, the trend of each index was different, and some indexes did not change much or even tended to deteriorate;

(3) Considering the influence of the noise of axle load and wheelbase on the response prediction, it was found that the noise of the wheelbase had a more significant influence on the prediction, but, in general, the correlation between the predicted response value and the theoretical value was good and the error was controllable. In particular, the displacement prediction performs well under the influence of noise. In addition, with the increase of train speed, the prediction accuracy decreased slightly—in particular, the acceleration prediction was more sensitive. The increase in the number of train sections also reduced the accuracy of response prediction, and the acceleration prediction was more sensitive to the change in train sections.

Author Contributions: Methodology, S.T. and X.K.; software, S.T. and Z.P.; validation, S.T. and J.M.; writing—original draft preparation, S.T. and X.K.; writing—review and editing, S.T., X.K. and J.M. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Science and Technology Research and Development Project of China State Railway Group Co., Ltd. (L2023G007, L2022G007), Science and Technology Research and Development Program Project of China railway group limited (2020-Special-02, 2021-Special-08), Hunan National Natural Science Foundation (2021JJ30053), and the Open Foundation of National Engineering Laboratory for High-Speed Railway Construction (HSR202003).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available upon request from the corresponding author. The dataset and code cannot be shared due to specific reasons.

Conflicts of Interest: Author Sui Tan was employed by the company China Railway Group Limited. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

1. Li, H.; Wang, T.; Yang, J.P.; Wu, G. Deep learning models for time-history prediction of vehicle-induced bridge responses: A comparative study. *Int. J. Struct. Stab. Dyn.* **2023**, *23*, 2350004. [CrossRef]
2. Xiang, P.; Xu, S.; Zhao, H.; Jiang, L.; Ma, H.; Liu, X. Running safety analysis of a train-bridge coupled system under near-fault ground motions considering rupture directivity effects. *Structures* **2023**, *58*, 105382. [CrossRef]
3. Sun, L.; Shang, Z.; Xia, Y.; Bhowmick, S.; Nagarajaiah, S. Review of bridge structural health monitoring aided by big data and articial intelligence: From condition assessment to damage detection. *J. Struct. Eng.* **2020**, *146*, 04020073. [CrossRef]
4. Mao, J.; Su, X.; Wang, H.; Li, J. Automated Bayesian operational modal analysis of the long-span bridge using machine-learning algorithms. *Eng. Struct.* **2023**, *289*, 116336. [CrossRef]
5. Feng, Y.; Wang, H.; Yang, H.; Wang, F. Time-continuous energy-conservation neural network for structural dynamics analysis. *Neurocomputing* **2021**, *456*, 450–460. [CrossRef]
6. Kim, T.; Song, J.; Kwon, O.S. Pre- and post-earthquake regional loss assessment using deep learning. *Earthq. Eng. Struct. Dyn.* **2022**, *49*, 657–678. [CrossRef]
7. Lagaros, L.D.; Papadrakakis, M. Neural network based prediction schemes of the nonlinear seismic response of 3D buildings. *Adv. Eng. Softw.* **2012**, *44*, 92–115. [CrossRef]
8. Xu, Z.; Wang, H.; Xing, C.; Tao, T.; Mao, J.; Liu, Y. Physics guided wavelet convolutional neural network for wind-induced vibration modeling with application to structural dynamic reliability analysis. *Eng. Struct.* **2023**, *297*, 117027. [CrossRef]
9. Kim, T.; Kwon, O.-S.; Song, J. Response prediction of nonlinear hysteretic systems by deep neural networks. *Neural Netw.* **2019**, *111*, 1–10. [CrossRef]
10. Mao, J.; Zhu, X.; Wang, H.; Guo, X.; Pang, Z. Train Load Identification of the Medium-Small Railway Bridge using Virtual Axle Theory and Bayesian Inference. *Int. J. Struct. Stab. Dyn.* **2024**, *24*, 2450193. [CrossRef]
11. Huang, S.Q.; He, X.H.; Xue, Z. Prediction of wind loads on high-rise building using a BP neural network combined with POD. *J. Wind Eng. Ind. Aerodyn.* **2017**, *170*, 1–17.
12. Abbas, T.; Kavrakov, I.; Morgenthal, G.; Lahmer, T. Prediction of aeroelastic response of bridge decks using articial neural networks. *Comput. Struct.* **2020**, *231*, 106198. [CrossRef]
13. Castellon, D.F.; Fenerci, A.; Øiseth, A. A comparative study of wind-induced dynamic response models of long-span bridges using articial neural networks, support vector regression and buffeting theory. *J. Wind Eng. Ind. Aerodyn.* **2021**, *209*, 104484. [CrossRef]
14. Zhang, Q.; Cai, X.; Zhong, Y.; Tang, X.; Wang, T. Dynamic response prediction of high-speed train on cable-stayed bridge based on genetic algorithm and fused neural networks. *Eng. Struct.* **2024**, *306*, 117869. [CrossRef]
15. Li, H.L.; Wang, T.Y.; Wu, G. Dynamic response prediction of vehicle-bridge interaction system using feedforward neural network and deep long short-term memory network. *Structures* **2021**, *34*, 2415–2431. [CrossRef]
16. Su, Y.H.; Kuo, C.-C.J. On extended long short-term memory and dependent bidirectional recurrent neural network. *Neurocomputing* **2019**, *356*, 151–161. [CrossRef]
17. Sherstinsky, A. Fundamentals of recurrent neural network (RNN) and long short-term memory (LSTM) network. *Phys. D-Nonlinear Phenom.* **2020**, *404*, 132306. [CrossRef]
18. Graves, A.; Schmidhuber, J. Framewise phoneme classification with bidirectional LSTM and other neural network architectures. *Neural Netw.* **2005**, *18*, 602–610. [CrossRef]
19. Chang, Z.H.; Zhang, Y.; Chen, W.B. Electricity price prediction based on hybrid model of adam optimized LSTM neural network and wavelet transform. *Energy* **2019**, *187*, 115804. [CrossRef]
20. Brunton, S.L.; Proctor, J.L.; Kutz, J.N. Discovering governing equations from data by sparse identification of nonlinear dynamical systems. *Appl. Math.* **2016**, *113*, 3932–3937. [CrossRef]
21. Kim, R.E.; Moreu, F.; Spencer, B.F., Jr. System identification of an in-service railroad bridge using wireless smart sensors. *Smart Struct. Syst.* **2015**, *15*, 683–698. [CrossRef]
22. Srivastava, N.; Hinton, G.; Krizhevsky, A.; Sutskever, I.; Salakhutdinov, R. Dropout: A simple way to prevent neural networks from overfitting. *J. Mach. Learn. Res.* **2014**, *15*, 1929–1958.
23. Zhang, R.; Chen, Z.; Chen, S.; Zheng, J.; Büyüköztürk, O.; Sun, H. Deep long short-term memory networks for nonlinear structural seismic response prediction. *Comput. Struct.* **2019**, *220*, 55–68. [CrossRef]
24. Zhao, H.; Wei, B.; Zhang, P.; Guo, P.; Shao, Z.; Xu, S.; Jiang, L.; Hu, H.; Zeng, Y.; Xiang, P. Safety analysis of high-speed trains on bridges under earthquakes using a LSTM-RNN-based surrogate model. *Comput. Struct.* **2024**, *294*, 107274. [CrossRef]
25. Furtado, A.; Rodrigues, H.; Arêde, A.; Varum, H.; Grubišić, M.; Šipoš, T.K. Prediction of the earthquake response of a three-storey infilled RC structure. *Eng. Struct.* **2018**, *171*, 214–235. [CrossRef]
26. Hochreiter, S.; Schmidhuber, J. Long Short-Term Memory. *Neural Comput.* **1997**, *9*, 1735–1780. [CrossRef]
27. Zhang, R.; Liu, Y.; Sun, H. Physics-guided convolutional neural network (PhyCNN) for data-driven seismic response modeling. *Eng. Struct.* **2020**, *215*, 110704. [CrossRef]
28. Joghataie, A.; Farrokh, M. Dynamic analysis of nonlinear frames by Prandtl neural networks. *J. Eng. Mech.* **2008**, *134*, 961–969. [CrossRef]
29. Zhou, S.; Song, G.; Wang, R.; Ren, Z.; Wen, B. Nonlinear dynamic analysis for coupled vehicle-bridge vibration system on nonlinear foundation. *Mech. Syst. Signal Process.* **2017**, *87*, 259–278. [CrossRef]

30. Zhai, W. *Vehicle–Track Coupled Dynamics: Theory and Applications*; Springer Nature: Singapore, 2020.
31. Zhao, H.; Wei, B.; Jiang, L.; Xiang, P.; Zhang, X.; Ma, H.; Xu, S.; Wang, L.; Wu, H.; Xie, X. A velocity-related running safety assessment metrics in seismic design for railway bridge. *Mech. Syst. Signal Process.* **2023**, *198*, 110305. [CrossRef]
32. Li, H.L.; Wang, T.Y.; Wu, G. A Bayesian deep learning approach for random vibration analysis of bridges subjected to vehicle dynamic interaction. *Mech. Syst. Signal Process.* **2022**, *170*, 108799. [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.