applied sciences

MDPI

Article

Prediction of Key Parameters of Wheelset Based on LSTM Neural Network

Duo Ye [1], Jing Wen [1,*], Shubin Zheng [1], Qianwen Zhong [1], Wanrong Pei [2], Hongde Jia [2], Chuanping Zhou [2] and Youping Gong [3]

[1] School of Urban Rail Transportation, Shanghai University of Engineering Science, Shanghai 201620, China
[2] School of Mechanical Engineering, Hangzhou Dianzi University, Hangzhou 310018, China
[3] School of Mechanical Engineering, Zhejiang University, Hangzhou 310030, China
* Correspondence: wenjing@sues.edu.cn

Abstract: As a key component of rail vehicle operation, the running condition of the wheelset significantly affects the operational safety of track vehicles. The wheel diameter, flange thickness, and flange height are key dimensional parameters of the wheelset, which directly influence the correct position of wheelsets on the track, and the train needs to be continuously monitored during the passenger operation. A prediction model for the key parameters of the wheelset is established based on LSTM (long short-term memory) neural network, and real measured data of wheelsets from the Shanghai Metro vehicles are selected. The predicted results of the model are compared and analyzed, and the results show that the LSTM-based prediction model for key parameters of wheelsets performs well, with the mean absolute percentage errors (MAPEs) for wheel diameter, flange thickness, and flange height being 0.08%, 0.42%, and 0.44%, respectively, for the left wheel and 0.07%, 0.35%, and 0.44%, respectively, for the right wheel. The prediction model for the train wheelset parameters established in this paper has a good prediction accuracy. By predicting the key parameters of the wheelset, the faults and causes of the wheelset can be found and determined, which is helpful for engineers to overhaul the wheelset faults, make maintenance plans, and perform preventive maintenance.

Keywords: LSTM neural network; wheelset wear; parameter prediction; deep learning; reprofiling plan

check for updates

Citation: Ye, D.; Wen, J.; Zheng, S.; Zhong, Q.; Pei, W.; Jia, H.; Zhou, C.; Gong, Y. Prediction of Key Parameters of Wheelset Based on LSTM Neural Network. *Appl. Sci.* **2023**, *13*, 11935. https://doi.org/10.3390/app132111935

Academic Editor: Diogo Ribeiro

Received: 11 September 2023
Revised: 21 October 2023
Accepted: 26 October 2023
Published: 31 October 2023

Copyright: © 2023 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

During the operation of a train, the wheelset plays a pivotal role in supporting the vehicle body, facilitating braking, and maintaining contact with the track. However, due to uneven distribution of the track curves, poor performance of the braking system, rail replacement, and lack of lubrication, the wheelsets may gradually experience abnormal conditions, such as wear on the tread surface, wear on the wheel flange, as well as scuffing and peeling of the tread surface [1]. The operational condition of the wheelset has a direct effect on the stability, safety, and comfortable operation of subway vehicles [2]. The wheel diameter, flange height, and flange thickness values, as key parameters of the wheelset, can reflect the potential risks in the operation of the wheelset [3]. Therefore, it is vital to predict the key parameters of the wheelset. With the development of rail transportation technology, an increasing number of methods have been applied to the prediction of wheelset parameters in rail vehicles. For example, Yunguang Ye, Yu Sun, and Dachuan Shi et al. implemented a new prediction model and applied particle swarm optimization algorithm on the basis of Kriging surrogate models, which is abbreviated as KSM-PSO, to achieve automatic adjustment of the scheme [4]. Singh S K, Das A K, and others predicted the wheelset parameters by combining the models of artificial neural network (ANN) and recurrent neural network (RNN) [5]. Because these parameters are

difficult to obtain through experiments, the efficiency is very low if they are calculated by multi-body dynamics theory. Thus, improved real-coded genetic algorithm (IRGA), success history-based adaptive differential evolution (SHADE), bonobo optimizer (BO), and grey wolf optimizer (GWO), the four meta-heuristic techniques algorithms, have been adopted in the research to optimize the artificial neural network and recurrent neural network models. The results confirm that success history-based adaptive differential evolution (SHADE) is the best optimizer among the four meta-heuristic techniques adopted above, whose MAPE in tests is less than 7% for SHADE-ANN and SHADE-RNN. Deng Y, Liu L, and others established a data-driven wheelset wear parameter prediction model based on LM-OMP-NARXNN [6]. LM-OMP-NARXNN is a nonlinear autoregressive model including exogenous input neural networks (NARXNNs), Levenberg Marquardt (LM), and orthogonal matching pursuit (OMP) algorithms. Their experimental results confirm that this method can obtain more effective and compact models with a smaller size. At the same time, compared with other prediction models and training algorithms used in NARXNN, this method has a higher accuracy in predicting the state parameters of the wheelset in the short term. Wang S, Yan H, and others built a dynamic model of high-speed railway vehicle with the SIMPACK software, and they used the Archard wear model based on this SIMPACK dynamic model to predict wheelest wear [7]. At the same time, taking the daily measurement data provided by Beijing Railway Administration as a reference, the wear was verified using BP neural network (BPNN) classification. The results of the combination of the two methods demonstrate that the wheel position in the bogie acquires great influence on the wheelset wear, while the position of the carriage in the train has barely no influence. The aforementioned research mainly analyzes the wheelset based on stable wear data. However, the abrasion prediction of the wheelset discussed in the above research is a unilateral prediction result, which only reflects the operating trend of the wheelset by predicting the wear of a single parameter. Firstly, because the wear change of the wheelset parameter is relatively small, the actual parameters of the wheelset cannot be accurately understood [8]. Secondly, the accuracy of the prediction result needs to be improved. In addition, the above neural network prediction model is only for the ideal wheelset, without considering the influence of wheelset repair on the prediction results. Wheelset parameters will fluctuate after rotational repair, and the wheelset restores other parameters to a safe range by reducing the wheel diameter value. If the prediction accuracy is not enough, the wheel diameter will be wasted, resulting in an increase in operation cost [9]. Moreover, the above prediction model does not select the neural network method from the aspect of wheelset data characteristics. The prediction accuracy is not high and cannot provide accurate guidance and reference for wheelset maintenance planning. Therefore, this study first took into account the wheelset rotational repair situation and verified that the prediction model was still good before and after rotational repair. Secondly, considering that the wheelset parameters are time-based parameters, the LSTM neural network was introduced, the dropout layer was used to optimize the prediction model, and the three most critical parameters of the wheelset were determined by investigating and sorting out the reasons for the actual wheelset rotational repair operation. All of them were modeled and predicted. The prediction of key parameters of the wheelset has a high accuracy, which can improve important support and guidance for future wheelset maintenance arrangement and preventive maintenance, preventing trains not running on time due to too many wheelsets requiring maintenance. Recently, there are too many wheelset maintenance instances in the Shanghai Metro, which leads to the train not going further, so it is necessary to arrange wheelset maintenance in a planned way.

This paper proposes a prediction model for the key parameters of the wheelset based on the LSTM neural network. Through the analysis of the causes of wheelset wear and relating data, the key parameters of the wheelset are determined: wheel diameter, flange thickness, and flange height. Based on the LSTM neural network, through data processing, model fitting, cross-validation, and other machine learning processes [10], a prediction model for the key parameters of the wheelset is constructed. Real measured data of the

wheelset are inputted into the LSTM prediction model and then compared with the BP (back propagation) prediction model and GRU (gate recurrent unit) prediction model. The results confirm that the LSTM model maintains high prediction accuracy even when the data fluctuate significantly after wheelset's reprofiling.

2. LSTM Neural Network

2.1. Recurrent Neural Network

2.1.1. The Basic Theory of Recurrent Neural Network

RNN (recurrent neural network) is one of the deep learning algorithms that has strong capabilities in processing sequential data. In traditional neural network models, data are typically inputted from the input layer, passing through multiple hidden layers, and finally outputted from the output layer. Each layer is fully connected, with no connections within the layer, and they are relatively independent of each other [11]. However, in a recurrent neural network, all input values at the current time 't' are related to the partial state of the previous time step '$t-1$'. The network state of the previous time step will affect the network state at the next time step, allowing the network to learn effectively by incorporating the data from the previous time step. The unfolded diagram of a recurrent neural network is shown in Figure 1.

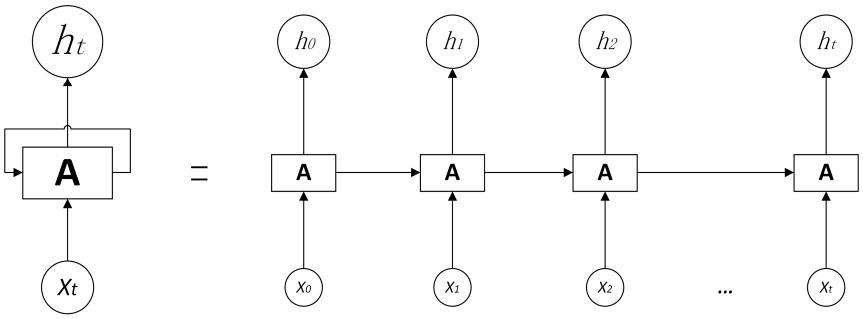


Figure 1. Unfolded diagram of a recurrent neural network.

2.1.2. The Basic Structure of RNN

An RNN network consists of input, output, and neural network units. The input are sequence data, where each sample is a time series. The input values of the same sample before and after the time are related, and the sequence length may be different. During training, forward propagation is performed for each time step of the input, and then parameter gradients are calculated through back propagation to update the parameters [12].

The calculation structure diagram of RNN is shown in Figure 2, where *S*-*W* on the right represents the self-connection of the hidden layer. In RNN, each layer shares parameters U, V, and W, which reduces the parameters that need to be learned in the network and improves learning efficiency. Here, input units are $\{x_0, \cdots, x_{t-1}, x_t, x_{t+1}, \cdots\}$, hidden units are $\{s_0, \cdots, s_{t-1}, s_t, s_{t+1}, \cdots\}$, and output units are $\{o_0, \cdots, o_{t-1}, o_t, o_{t+1}, \cdots\}$. Input layer x_t represents the input of time T. The hidden layer is $s_t = f(U_{x_t} + W_{s_{t-1}})$, and f is the non-linear activation function. Output layer is $o_t = softmax(W_{s_t})$, $softmax$ is a normalized exponential function.

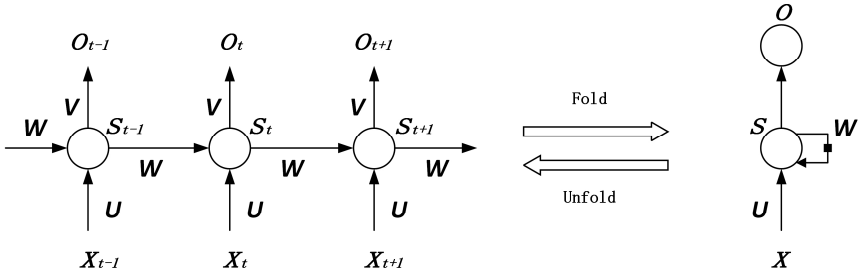


Figure 2. The calculation structure of RNN.

2.2. The Basic Theory of LSTM (Long Short-Term Memory) Network

2.2.1. LSTM Neural Network

The LSTM network is called long short-term memory recurrent neural network [13]. The LSTM neural network also has a chain structure similar to a regular RNN, but it differs in terms of its improved internal structure [14]. Compared to RNN, the LSTM neural network has an additional input and output. These two additional parts are the input and output of LSTM's memory and forgetting mechanism, which is commonly represented by C to denote the cell state [15]. The emergence of the cell state effectively avoids the problems of gradient vanishing and exploding that often occur in an RNN model. The internal structure of the LSTM neural network is shown in Figure 3.

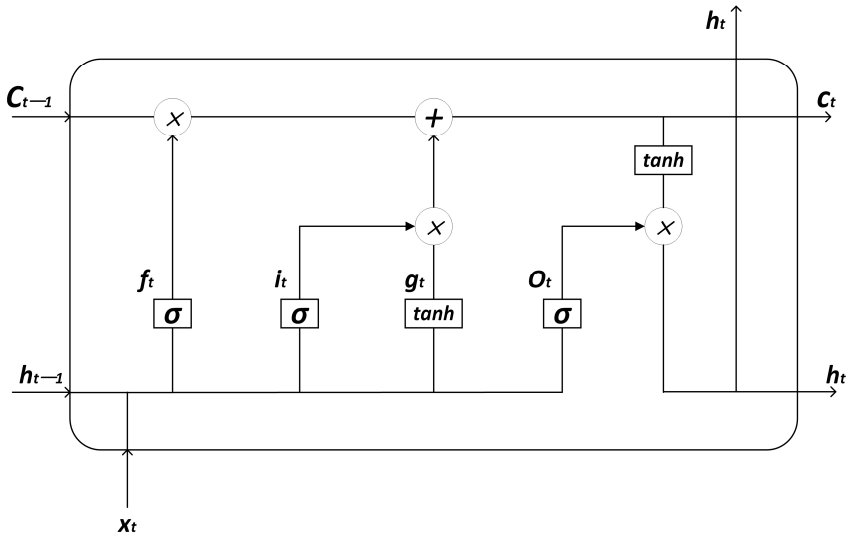


Figure 3. The internal structure of LSTM neural network.

The internal structure of an LSTM neuron consists of four parts: input gate, output gate, forget gate, and memory unit. The forget gate is represented by *sigmoid* layer, while the other two *tanh* layers correspond to the input and output of the memory unit [16]. The LSTM neural network has the ability to remove or add information to the *cell* state, which is controlled and regulated by the *gate* structure. The *gate* structure is composed of a *sigmoid* neural network layer and a pointwise multiplication operation and can selectively pass information. The output of the *sigmoid* layer ranges from 0 to 1 to describe how much content each component allows to pass through. A value of 0 indicates prohibiting content from passing, while a value of 1 indicates allowing everything to pass. An LSTM neural network structure has three such gate structures to protect and control the state of the unit [17]. The functions of the three gates are as follows:

(1) Forget gate: It is used to determine the information discarded from the node state. This determination is made by the *sigmoid* layer, which decides whether to pass or partially pass based on the previous time step's output. Every number output of the *cell* state c_{t-1} between 0 and 1 is determined by inputting h_{t-1} and x_t. The calculation is given in Equation (1):

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

(2) Input gate: It determines the new information stored in the *cell* state and is responsible for the input at the current sequence position. In the first step, the *sigmoid* layer called the input gate layer decides which values to update. At the same time, the *tanh* layer creates a new candidate value vector $\widetilde{c_t}$ and adds it to the *cell* state. Then these two components are combined to update the *cell* state. The calculation is shown in Equations (2) and (3):

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \tag{2}$$

$$\widetilde{c_t} = tanh(W_c \cdot [h_{t-1}, x_t] + b_c) \tag{3}$$

Next is the update of the old *cell* state c_{t-1} to the new *cell* state c_t. This is done by multiplication between the old cell state c_{t-1} and f_t to forget the information decided in the previous step. Then, the new candidate value c_t is added to scale to varying degrees to update the new state value c_t. The calculation is given in Equation (4):

$$c_t = f_t * c_{t-1} + i_t * \widetilde{c_t} \tag{4}$$

(3) Output gate: It determines the output layer. First, determine the output *cell* state by the *sigmoid* layer. Then, the *cell* state is scaled between -1 and 1 by the *tanh* layer. The scaled cell state is pairwise-multiplied with the initial output obtained from the *sigmoid* layer, resulting in the model's output, which contains the desired information. The calculation is shown in Equations (5) and (6):

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \tag{5}$$

$$h_t = o_t * tanh(c_t) \tag{6}$$

2.2.2. Training Method of LSTM Neural Network

The back propagation algorithm for the LSTM neural network is similar to the back propagation algorithm for RNN. It calculates the partial derivatives of the shared parameters based on the loss function and iteratively updates the parameters with gradient descent [18]. The difference is that LSTM neural network has two hidden states h_t and c_t. The error term is defined as Equations (7) and (8):

$$\delta_h^{(t)} = \frac{\partial L}{\partial h^{(t)}} \tag{7}$$

$$\delta_c^{(t)} = \frac{\partial L}{\partial c^{(t)}} \tag{8}$$

obtaining the global loss function $\delta_h^{(\tau)}$ and $\delta_c^{(t)}$ at the last moment of the time series data τ. The calculation is shown in Equations (9) and (10):

$$\delta_h^{(\tau)} = \frac{\partial L}{\partial o^{(\tau)}} \frac{\partial o^{(\tau)}}{\partial h^{(\tau)}} = V^T(y^{*(\tau)} - y^{(\tau)}) \tag{9}$$

$$\delta_c^{(t)} = \frac{\partial L}{\partial h^{(\tau)}} \frac{\partial h^{(\tau)}}{\partial c^{(\tau)}} = \delta_h^{(\tau)} \odot o^{(\tau)} \odot (1 - tanh^2(c^{(\tau)})) \tag{10}$$

The reverse derivation to obtain W_f parameter gradient is shown in Equation (11):

$$\frac{\partial L}{\partial W_f} = \sum_{t=1}^{\tau} \frac{\partial L}{\partial c^{(t)}} \frac{\partial c^{(t)}}{\partial f^{(t)}} \frac{\partial f^{(t)}}{\partial W_f} = \sum_{t=1}^{\tau} [\delta_c^{(t)} \odot c^{(t-1)} \odot f^{(t)}(1 - f^{(t)})](h^{(t-1)})^T \tag{11}$$

3. Wheelset Key Parameter Analysis

3.1. Standards and Requirements for Key Parameters

Figure 4 shows the appearance of the bogie and wheelset, and Figure 5 shows the appearance of the wheelset rim. The bogie frame includes the end beam, side beam, cross beam, and other parts, which mainly support the car body and the acceleration and braking of the train [19]. The wheelset is the only part of the vehicle that has contact with the track and is a very key part as the stress point of the train [20]. The wheelbase is 1353 ± 2 mm and the wheel diameter value running range is 770 mm to 840 mm.

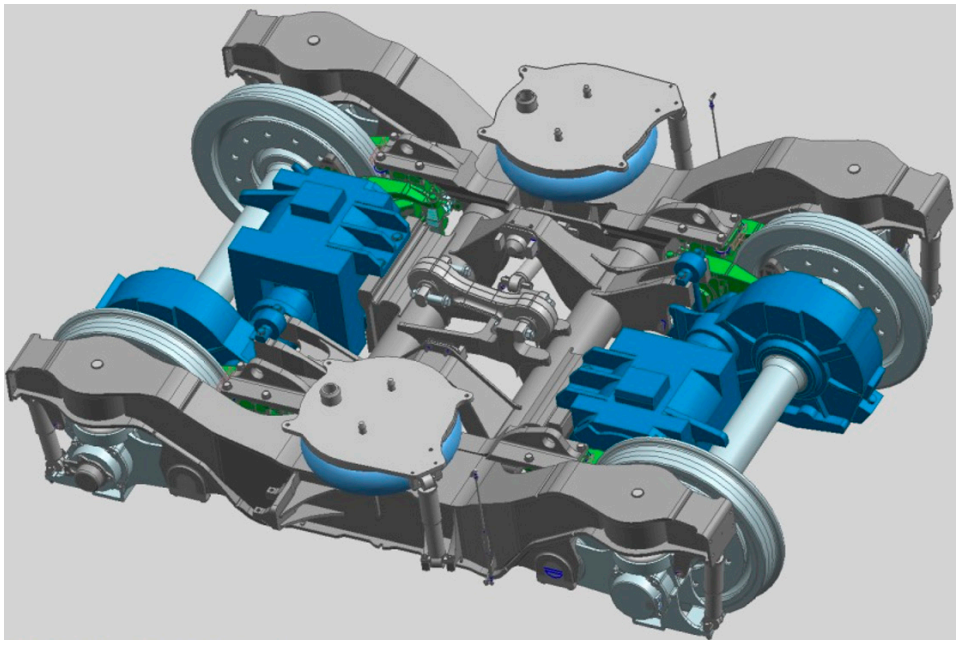


Figure 4. Appearance of train bogie.

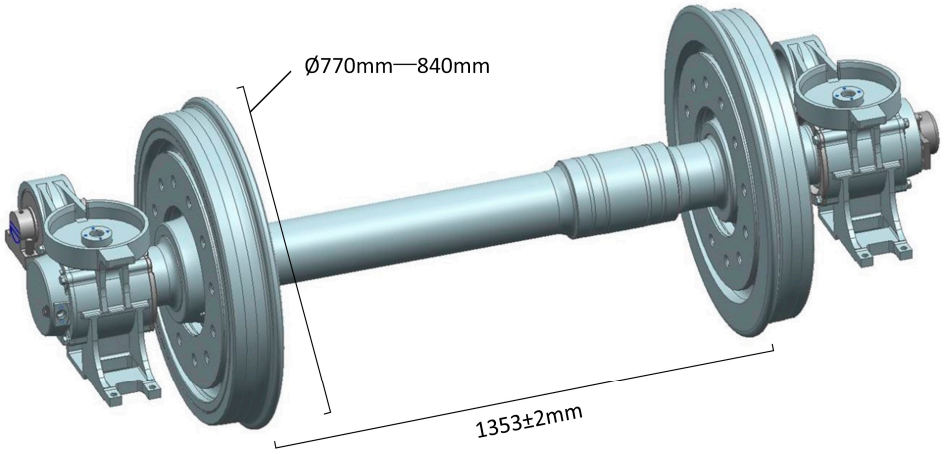


Figure 5. Train wheelset appearance diagram.

The daily inspection of the railway vehicle wheelset is mainly carried out by measuring parameters such as wheel diameter D, flange thickness Sw, flange height Sh, and flange slope Qr using the trackside measurement system. These parameters are evaluated to assess the operation of the wheelset. When the wheel flange thickness value Sw decreases and the flange height value Sh increases, it indicates that the wear on the flange is greater than that on the tread. Conversely, when the flange thickness value Sw increases and the flange height value Sh decreases, it indicates that the wear on the tread is greater than that on the flange. The flange parameters are shown in Figure 6.

Point P_0 is defined at a distance of 70 mm from the wheel back at the tread. Point P_1 is defined 2 mm below the highest point of the flange. Point P_2 is defined at a vertical distance of 10 mm above point P_0 (DIN5573 profile). The flange thickness Sw is the horizontal distance between the wheel back and point P_2. The flange height is the vertical distance between the highest point of the flange and point P_0. The wheel flange slope Qr is the horizontal distance between points P_1 and P_2.

To ensure the safe operation of railway vehicles, there are strict operational requirements for the dimensions of the wheelset. According to the requirements of a certain line in the Shanghai Metro, the range of the wheel diameter D is 770 mm $\leq$ D $\leq$ 840 mm, the

range of the flange thickness *Sw* is 26 mm $\leq$ Sw $\leq$ 33 mm, the range of the flange height *Sh* is 27.5 mm $\leq$ Sh $\leq$ 34 mm, and the range of the flange slope Q*r* is 6.5 $\leq$ Qr $\leq$ 13.5. In addition, for coaxial wheelset, the wheel diameter deviation should not exceed 2 mm; for the same bogie, the wheel diameter deviation should not exceed 4 mm; and for the same car, the wheel diameter deviation should not exceed 7 mm.

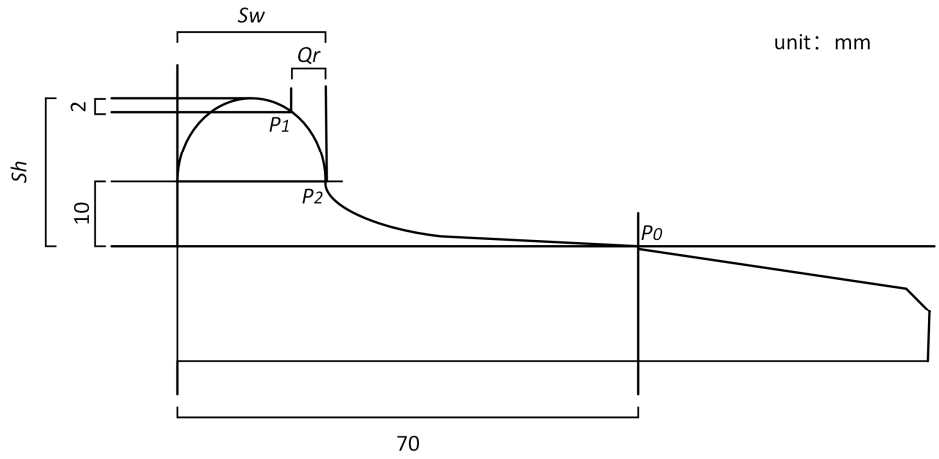


Figure 6. Schematic diagram of flange parameters.

3.2. Analysis of Key Parameter Evolution

The phenomenon of uneven wear may occur during the operation of the wheelset, and this study focuses on both left and right wheels. Under normal wear conditions, as the vehicle runs, the values of flange thickness *Sw* and wheel diameter *D* decrease, while the value of flange height *Sh* increases. When these key parameter values reach their limits, reprofiling is carried out to increase the flange thickness *Sw* value and decrease the flange height value *Sh* by reducing the wheel diameter D This ensures that the wheelset meets the safety requirements. Approximately 8–10 mm of wheel diameter loss is required to restore 1 mm of flange thickness [21]. The changes in wheel diameter, flange thickness, and flange height values of the wheelset on a certain line between 2017 and 2022 are shown below.

As Figure 7 shows, the wheel diameter value decreases continuously as railway vehicles run. The decreasing trends in both left and right wheels are basically the same, and there is no excessive deviation of the wheel diameter between the two sides. However, in June 2019 and October 2020, the decreasing slopes became steeper due to reprofiling for the excessive difference in flange thickness of left and right wheels.

From Figure 8, it can be seen that the flange thickness value decreases continuously as railway vehicles run. The trends in both left and right wheels are basically the same, but the value of the right wheel is significantly higher than that of the left wheel. In June 2019, the difference in flange thickness values between the left and right wheels reached its limit, then reprofiling was carried out. In October 2020, due to the difference in flange thickness values between the left and right wheels approaching the limit again, the wheels underwent a second reprofiling. After the first reprofiling, the flange height values were 29.3 mm for the left wheel and 30.2 mm for the right wheel. Before the second reprofiling, the flange height values were 28 mm for the left wheel and 29.3 mm for the right wheel. The change was 1.3 mm for the left wheel and 0.9 mm for the right wheel, and the wear difference between the left and right wheels was 0.4 mm. Combined with the analysis of the cause of wheelset uneven wear, it is considered that the wheelset has experienced uneven wear.

From Figure 9, it is shown that the trends in flange height values for the left and right wheels are basically the same as railway vehicles run. The flange height values continuously increase before June 2019 and experience a sudden drop around June 2019 due to the excessive difference in flange thickness between the left and right wheels. Reprofiling was performed on the wheelset, and the reprofiling template required a flange height value of 28 $\pm$ 0.5 mm after reprofiling. In October 2020, a second reprofiling was

conducted, and there was not much change in the flange height values. However, they gradually increased during operation after reprofiling.

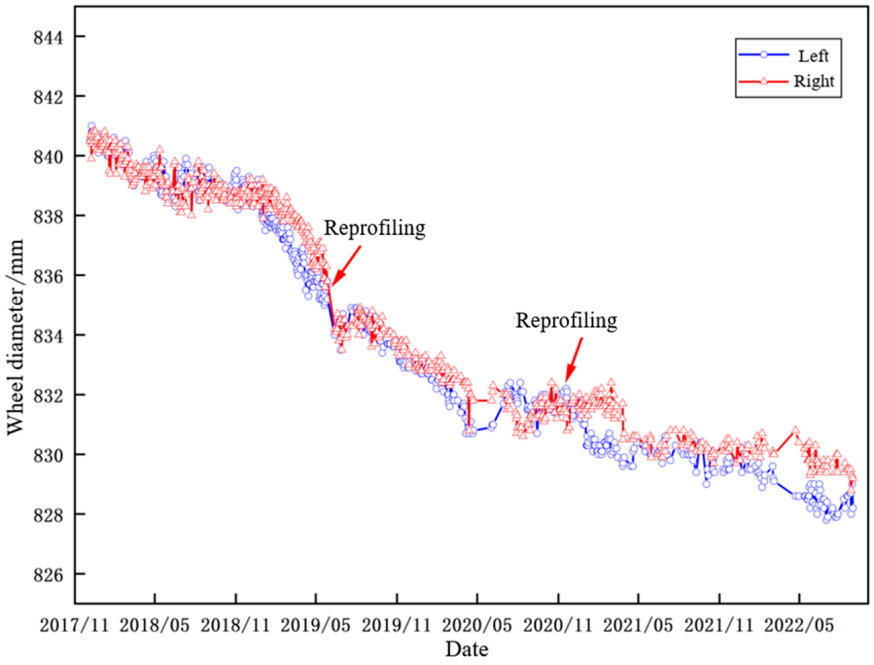


Figure 7. Changes in the wheel diameter of a certain wheelset.

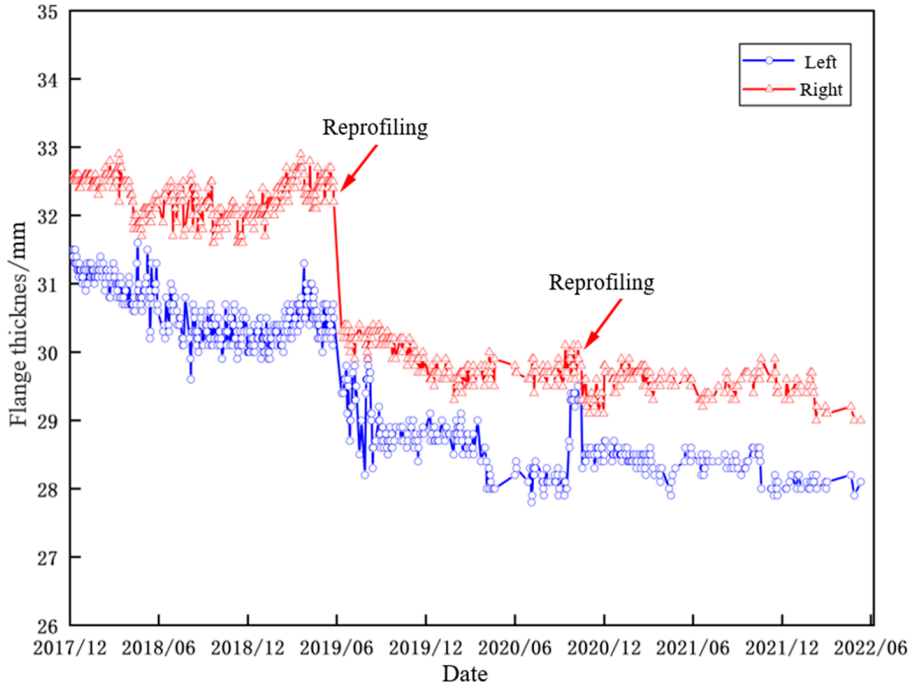


Figure 8. Changes in the flange thickness of a certain wheelset.

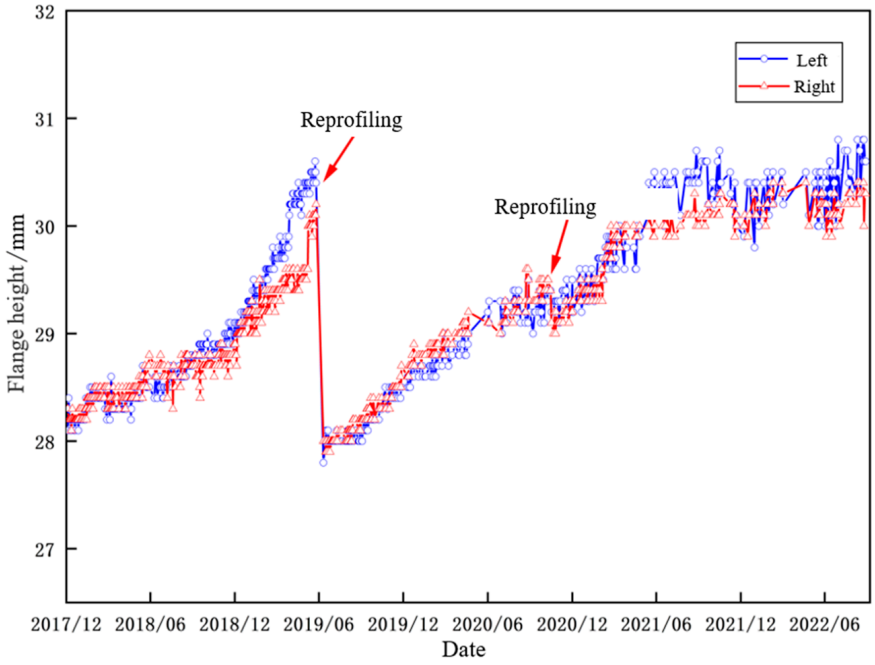


Figure 9. Changes in the flange height of a certain wheelset.

4. LSTM Prediction Model

4.1. Model Construction Process

Based on the measured data of a certain line's trackside system in the Shanghai Metro, a prediction model for the key parameters of wheelset wear is established with an LSTM neural network. The measured key parameters of the wheelset, including wheel diameter, flange thickness, and flange height, are used as inputs to the model to predict the key parameter values for the next year. The process of model construction mainly involves the processing of raw operational data of the wheelset, identification of the main key parameters, and optimization of model parameters. The flowchart of the construction process is shown in Figure 10.

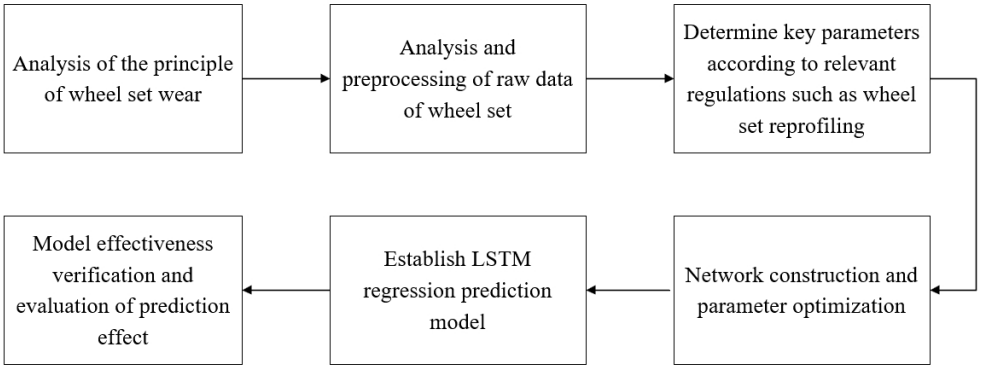


Figure 10. Flowchart of wheelset key parameter prediction model.

4.2. Wheelset Key Parameters and Data Preprocessing

The wheel diameter D, flange thickness Sw, and flange height Sh of the left and right wheels are the focus of this study on wheelset key parameters. The data are sourced from the wheelset measurements obtained by a monitoring system alongside a certain metro line in the Shanghai Metro. The time span of the measured wheelset data ranges from December 2017 to October 2022. Due to the susceptibility of the trackside monitoring system to factors such as the environment at the train base, the wear of the wheelset is influenced by various factors. Therefore, the original data are processed first, followed by handling of outliers

and missing values. Since the measurement intervals are relatively short, the method used for handling outliers involves taking the average of the changing values. The calculation is given in Equation (12):

$$X_t = \frac{\Delta\overline{X_1} + \Delta\overline{X_2}}{2} + X_{t-1} \tag{12}$$

In Equation (12), X_t represents the corrected value for an outlier, $\Delta\overline{X_1}$ represents the average of the changing values measured each time in the month before the outlier point, $\Delta\overline{X_2}$ represents the average of the changing values measured each time in the month after the outlier point, and X_{t-1} represents the value measured at the previous point before the outlier.

4.3. Model Construction and Parameter Optimization

4.3.1. Model Training

After confirming the key parameters and preprocessing the data, 809 sets of data of the left and right wheel diameter, flange thickness, and flange height are extracted. These three parameters will serve as the input values for the prediction model. A total of 80% of the data will be used as the training set, and the remaining 20% will be used as the validation set. To normalize the wheel diameter value, wheel flange thickness value, and wheel flange height value, we will use the Min–Max normalization method, which maps the values to the [0, 1] range. For building the neural network model, we will utilize the sequential model from the Keras deep learning library in Python. The model will consist of three layers. To optimize the model, we will utilize the Adam optimizer and set the learning rate to 0.01. The model will be trained for 100 epochs with a step size of seven.

4.3.2. Dropout Layer Settings

During model training, neural network may often exhibit varying degrees of overfitting, which results in a decreased generalization capability of the data in the training set and reduced practical value [22]. To ensure that the model has stronger generalization ability and to prevent overfitting during the neural network training process [23], the dropout technique can be applied for optimization [24]. In the forward propagation of the model, dropout is generally applied by randomly deactivating a certain range of neurons within the network. The deactivation range for the dropout layer is typically set as [0.2, 0.5] and it is commonly applied in fully connected layers [25]. In this study, parameters were adjusted through comparative research, and a dropout deactivation probability of 0.2 was determined as the final choice. The dropout representation is shown in Figure 11.

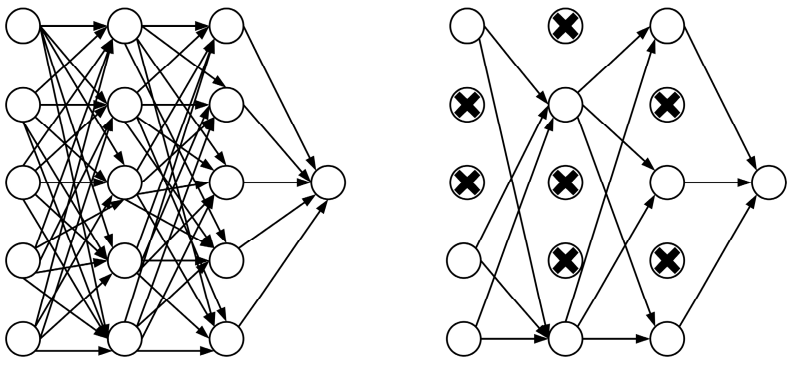


Figure 11. Dropout layer.

The prediction model was built and parameter tuning was performed in Python. The process of establishing the model is shown in Figure 12.

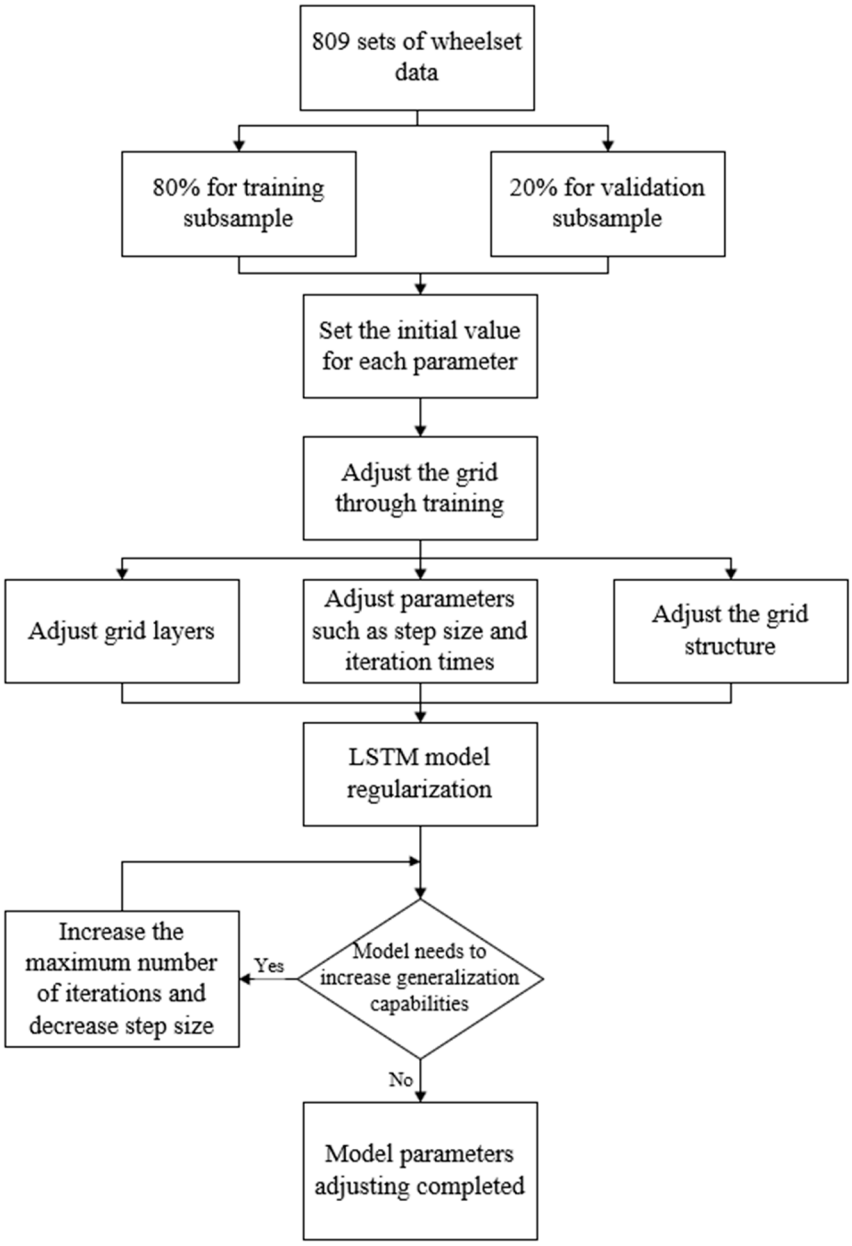


Figure 12. Flowchart of parameter tuning and model establishing.

4.3.3. Model Evaluation Metrics

For evaluating the model, the following metrics are adopted: root mean square error (*RMSE*), mean absolute error (*MAE*), and mean absolute percentage error (*MAPE*). These three metrics can assess the accuracy between the prediction values and the actual values.

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(\hat{y}_i - y_i)^2} \tag{13}$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|\hat{y}_i - y_i| \tag{14}$$

$$MAPE = \frac{100\%}{n}\sum_{i=1}^{n}\left|\frac{\hat{y}_i - y_i}{y_i}\right| \tag{15}$$

After calculation, the wheel diameter values of the validation set results in *RMSE* is 0.093 mm, *MAE* is 0.076 mm, and *MAPE* is 0.13%. The flange thickness values of the

validation set results in $RMSE$ is 0.061 mm, MAE is 0.046 mm, and $MAPE$ is 0.37%. The flange height values of the validation set results in $RMSE$ is 0.104 mm, MAE is 0.085 mm, and $MAPE$ is 0.48%. The predicted results of the wheel diameter, flange thickness, and flange height prediction models respectively established in this study basically meet the expected requirements.

5. Model Prediction Results and Comparative Analysis

To assess the predictive performance of the model, actual measured data from a wheelset of a certain line in the Shanghai Metro are selected. After preprocessing the data with the aforementioned procedure, the data were input into the established prediction model. The wheel diameter, flange thickness, and flange height values were predicted for approximately one year. These predictions are then analyzed in comparison with predictions from the BP (back propagation) neural network prediction model and the GRU (gate recurrent unit) neural network prediction model [26–29]. The comparison between the predicted results and the actual values is shown in Figures 13 and 14.

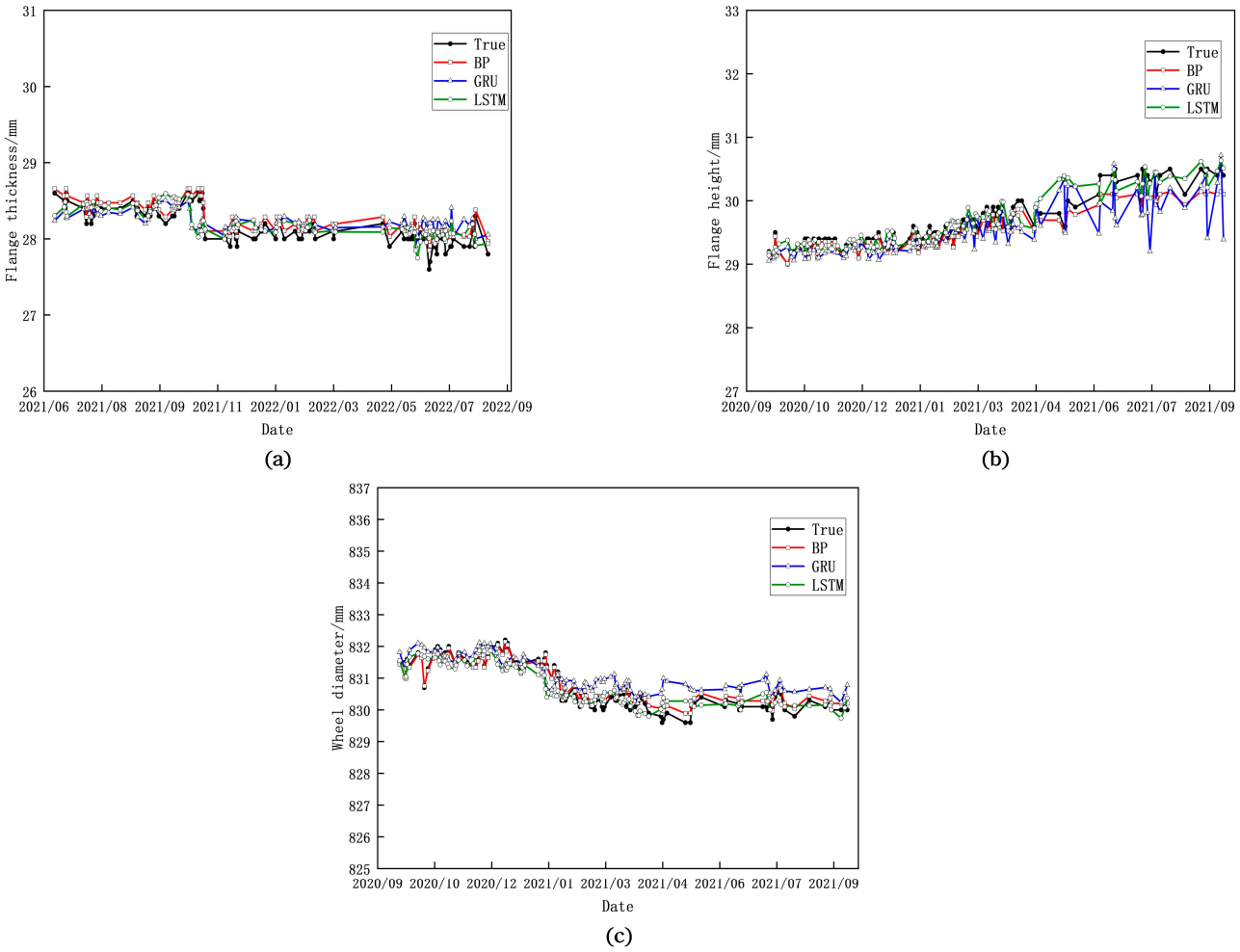


Figure 13. Comparison of left wheel predictions: (**a**) flange thickness value; (**b**) flange height value; (**c**) wheel diameter value.

The prediction time for the left wheel diameter and flange height spans from September 2020 to October 2021, approximately 13 months. The prediction time for the left wheel flange thickness spans from June 2021 to September 2022, approximately 15 months. The prediction time for the right wheel diameter and flange height spans from August 2020 to August 2021, approximately 12 months. The prediction time for the right wheel flange thickness spans from March 2020 to March 2021, approximately 12 months.

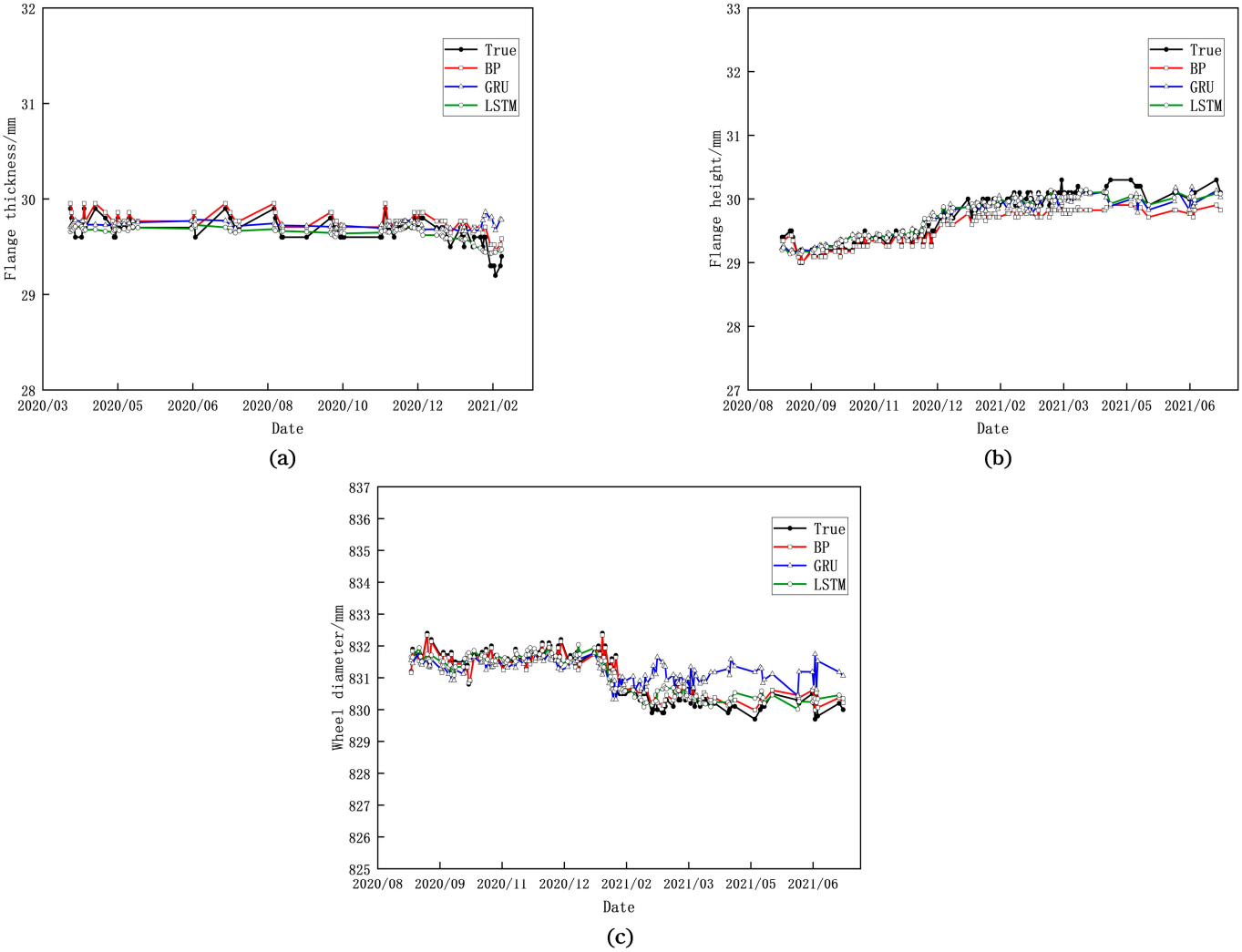


Figure 14. Comparison of right wheel predictions:(**a**) flange thickness value; (**b**) flange height value; (**c**) wheel diameter value.

Figure 13 shows the comparison of left wheel prediction parameters. After the train traveled 500,000 km in April 2021, the prediction errors of both the BP neural network and GRU neural network increase significantly. In addition, the LSTM neural network maintained good prediction accuracy until October 2021, even after the train had traveled 620,000 km. The LSTM neural network is more advantageous in wheel-to-left wheel parameter prediction.

Figure 14 shows the comparison of right wheel prediction parameters. Among the three models compared by the BP neural network, GRU neural network, and LSTM neural network, the predicted value of the LSTM neural network is more consistent with the actual value. In June 2021, after the train had run 560,000 km, the accuracy of the GRU model's prediction of the left wheel decreased significantly. In May 2021, after the train had traveled 550,000 km, the BP model's prediction error for the right wheel increases significantly. In addition, the LSTM neural network model maintained good predictive performance until October 2021, even after the train traveled 620,000 km.

It can be seen from these pictures that, compared with other neural networks, the prediction accuracy of the LSTM neural network has been significantly improved, and it is obviously closer to the true value no matter it is the left wheel or right wheel. The accuracy of the wheel diameter value is significantly higher than that of the flange height value and the flange thickness value, which means that the prediction of these two values requires

higher requirements, and there are many reasons for affecting the flange height value and the thickness value, and it is easy to be affected.

In the wheelset maintenance base, the left and right wheel diameter difference to limit, rim height to limit, and rim thickness to limit are the standards that need to be determined [30]. Table 1 shows the prediction result indexes of three different neural network models based on the train wheelset data of a certain line. RMSE, MAE, and MAPE indexes of the LSTM neural network prediction model based on time series in the table are superior to that of the BP neural network and GRU neural network prediction model, especially for the wheelset diameter value. The MAPE value of the left wheel was 0.08%, and that of the right wheel was 0.07%, which were 0.13% and 0.09% lower than that of the GRU prediction model, respectively. In terms of flange thickness, the RMSE of the left and right wheels of the LSTM neural network prediction model was 0.077 and 0.058, both lower than that of the BP and GRU neural network prediction models. The MAE values of the left and right wheels of the LSTM neural network were 0.114 and 0.085. From the above results, it can be seen that the LSTM neural network based on time series has greater advantages in processing wheelset operation data. The prediction errors fall within an acceptable range, indicating that this model can be applied in the maintenance department of subway systems. In the case of normal wear, the wheelset parameter value changes very little, so better prediction accuracy is required. If the prediction accuracy is not up to standard, it will lead to the rim parameter reaching the limit while the train is still running. Even if the wheelset parameter is extremely small and exceeds the standard, there will be no small accident, which greatly affects the operation safety of the train. In addition, it can accurately judge the wear trend of the train wheelset and help engineers to clearly understand the operating status and trend of the wheelset. For example, it can know when the wheelset needs to be rotated or fails and which parameter of the wheelset is abnormal, and from the parameter changes of the wheelset, it can find problems in other aspects such as route or lubrication, etc. This allows engineers to perform more targeted maintenance, and the prediction model can provide important support for the formulation of maintenance plans and preventive wheelset maintenance [31].

Table 1. Evaluation metrics of prediction results (unit: mm).

Model	Parameter	Left Wheel Evaluation Metrics			Right Wheel Evaluation Metrics		
		RMSE	*MAE*	*MAPE*	*RMSE*	*MAE*	*MAPE*
BP	Wheel diameter *D*	0.145	0.124	0.11%	0.125	0.107	0.10%
	Flange thickness *Sw*	0.103	0.090	0.45%	0.111	0.090	0.51%
	Flange height *Sh*	0.168	0.120	0.49%	0.131	0.117	0.50%
GRU	Wheel diameter *D*	0.262	0.195	0.21%	0.223	0.175	0.16%
	Flange thickness *Sw*	0.201	0.154	0.62%	0.191	0.132	0.55%
	Flange height *Sh*	0.361	0.268	0.86%	0.141	0.114	0.51%
LSTM	Wheel diameter *D*	0.098	0.078	0.08%	0.091	0.074	0.07%
	Flange thickness *Sw*	0.077	0.061	0.42%	0.058	0.044	0.35%
	Flange height *Sh*	0.133	0.114	0.47%	0.105	0.085	0.44%

6. Conclusions

In this study, an analysis of the wheelset wear on a certain line in the Shanghai Metro was conducted to identify key parameters: wheel diameter, flange thickness, and flange height. A prediction model was then established to forecast these three key parameters. The model is based on the LSTM neural network, and through optimization of the network structure, number of layers, and various parameters, the optimal prediction model was obtained. Utilizing the established prediction model, the values for a specific wheelset over one year were predicted, and the predicted values were compared and validated against the actual values. The results showed that the mean absolute MAPE was below 0.5%, indicating the reliability and accuracy of the results.

The main conclusions of this study are as follows:

1. The wheelset data will undergo significant changes after reprofiling. The predictive performance of the current methods is greatly affected by reprofiling. However, the proposed prediction model remains reliable in predicting after reprofiling.

2. Considering the temporal nature of the wheelset data, the introduction of a time series network for wheelset parameter prediction results in higher accuracy compared to traditional neural network models.

3. This study provides a method for maintenance engineers in the railway vehicle industry to understand the trend in wheelset changes, facilitating the prediction of wheel operating conditions and assisting engineers in planning reprofiling work. It has practical value in the rail transit industry.

Since the operation of rail vehicles is also influenced by the running route, such as in the case of circular lines, there may be significant wear on one side of the wheel due to more frequent turns [32]. Therefore, this study predicts both the left and right wheels to facilitate monitoring of wheelset changes on special routes. Subsequent research will incorporate parameters related to the changes in the left and right wheels, as well as the average differences in coaxial wheel diameters, average thicknesses of coaxial flanges, and relevant parameters of the same bogie, to further investigate wheelset wear. This will be followed by optimization of reprofiling strategies and management strategies for wheels in the next step.

Author Contributions: Conceptualization, D.Y. and J.W.; methodology, D.Y.; soft-ware, D.Y.; validation, D.Y. and J.W.; formal analysis, D.Y.; investigation, D.Y.; resources, J.W.; data curation, J.W.; writing—original draft preparation, D.Y.; writing—review and editing, D.Y. and J.W.; visualization, D.Y.; supervision, J.W., S.Z., Q.Z., W.P., H.J., C.Z. and Y.G.; project administration, J.W. and Q.Z. All authors have read and agreed to the published version of the manuscript.

Funding: Supported by National Natural Science Foundation of China (Grant No. 51975347).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available on re-quest from the corresponding author. The data are not publicly available due to data is internal information of the cooperation organization.

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

References

1. Zhao, X.N.; Chen, G.X.; Lv, J.Z.; Zhang, S.; Wu, B.W.; Zhu, Q. Study on the mechanism for the wheel polygonal wear of high-speed trains in terms of the frictional self-excited vibration theory. *Wear* **2019**, *426–427*, 1820–1827. [CrossRef]
2. Kaiser, I.; Strano, S.; Terzo, M.; Tordela, C. Estimation of the railway equivalent conicity under different contact adhesion levels and with no wheelset sensorization. *Veh. Syst. Dyn.* **2023**, *61*, 19–37. [CrossRef]
3. Zhu, W.; Yang, D.; Guo, Z.; Huang, J.; Huang, Y. Data-driven wheel wear modeling and reprofiling strategy optimization for metro systems. *Transp. Res. Rec.* **2015**, *2476*, 67–76. [CrossRef]
4. Ye, Y.; Sun, Y.; Shi, D.; Peng, B.; Hecht, M. A wheel wear prediction model of non-Hertzian wheel-rail contact considering wheelset yaw: Comparison between simulated and field test results. *Wear* **2021**, *474*, 203715. [CrossRef]
5. Singh, S.K.; Das, A.K.; Singh, S.R.; Racherla, V. Prediction of rail-wheel contact parameters for a metro coach using machine learning. *Expert Syst. Appl.* **2023**, *215*, 119343. [CrossRef]
6. Deng, Y.; Liu, L.; Li, M.; Jiang, M.; Peng, B.; Yang, Y. A data-driven wheel wear prediction model for rail train based on LM-OMP-NARXNN. *J. Comput. Inf. Sci. Eng.* **2023**, *23*, 021012. [CrossRef]
7. Wang, S.; Yan, H.; Liu, C.; Fan, N.; Liu, X.; Wang, C. Analysis and prediction of high-speed train wheel wear based on SIMPACK and backpropagation neural networks. *Expert Syst.* **2021**, *38*, e12417. [CrossRef]
8. Mariana de Almeida Costa, M.; Braga, J.P.D.A.P.; Andrade, A.R. Assessing the performance of different devices in railway wheelset inspection. *Measurement* **2020**, *165*, 108145. [CrossRef]
9. Andrade, R.A.; Stow, J. Assessing the potential cost savings of introducing the maintenance option of 'Economic Tyre Turning' in Great Britain railway wheelsets. *Reliab. Eng. Syst. Saf.* **2017**, *168*, 317–325. [CrossRef]

10. Liu, C.H.; Gu, J.C.; Yang, M.T. A Simplified LSTM Neural Networks for One Day-Ahead Solar Power Forecasting. *IEEE Access* **2021**, *9*, 17174–17195. [CrossRef]
11. Kaviani, S.; Sohn, I. Application of complex systems topologies in artificial neural networks optimization: An overview. *Expert Syst. Appl.* **2021**, *180*, 115073. [CrossRef]
12. Liu, R.; Zhang, X. Semi-Implicit Back Propagation. *arXiv* **2020**, arXiv:2002.03516.
13. Chen, Z.; Wu, B.; Li, B.; Ruan, H. Expressway Exit Traffic Flow Prediction for ETC and MTC Charging System Based on Entry Traffic Flows and LSTM Model. *IEEE Access* **2021**, *9*, 54613–54624. [CrossRef]
14. Hailu, B.M.; Assabie, Y.; Sinshaw, Y.B. Semantic Role Labeling for Amharic Text Using Multiple Embeddings and Deep Neural Network. *IEEE Access* **2023**, *11*, 33274–33295. [CrossRef]
15. Landi, F.; Baraldi, L.; Cornia, M. Working memory connections for LSTM. *Neural Netw.* **2021**, *144*, 334–341. [CrossRef]
16. Bicheng, H.; Langxing, T.; Yu, Z. Short-term power generation load forecasting based on LSTM neural network. *J. Phys. Conf. Ser.* **2022**, *2247*, 012033.
17. Song, K.; Jang, J.H.; Shin, S.J.; Moon, I.C. Bivariate Beta-LSTM. In Proceedings of the AAAI Conference on Artificial Intelligence, New York, NY, USA, 7–12 February 2020; pp. 5818–5825. [CrossRef]
18. Shankar, R. Bi-directional LSTM based channel estimation in 5G massive MIMO OFDM systems over TDL-C model with Rayleigh fading distribution. *Int. J. Commun. Syst.* **2023**, *36*, e5585. [CrossRef]
19. Baek, S.; Song, X.; Kim, M.; Kim, J. Multiobjective Optimization of Beam Structure for Bogie Frame Considering Fatigue-Life Extension. *J. Electr. Eng. Technol.* **2021**, *16*, 1709–1719. [CrossRef]
20. Zhang, J.; Zhu, T.; Yang, B.; Wang, X.; Xiao, S.; Yang, G.; Li, B. Influence of wheelset rotational motion on train collision response and wheelset lift mechanism. *Int. J. Rail Transp.* **2023**, *11*, 573–597. [CrossRef]
21. Xu, Z.; Dong, X.; Peng, Z. Mechanism and Control Measures for Abnormal Wear of Small Curve Rims Based on Wheel Rail Matching. *J. Vib. Shock.* **2022**, *41*, 127–133.
22. Karystinos, G.N.; Pados, D.A. On overfitting, generalization, and randomly expanded training sets. *IEEE Trans. Neural Netw.* **2000**, *11*, 1050–1057. [CrossRef] [PubMed]
23. Cheng, G.; Peddinti, V.; Povey, D.; Manohar, V.; Khudanpur, S.; Yan, Y. An exploration of dropout with LSTMs. In Proceedings of the Interspeech 2017, Stockholm, Sweden, 20–24 August 2017; pp. 1586–1590.
24. Xie, J.; Ma, Z.; Lei, J.; Zhang, G.; Xue, J.H.; Tan, Z.H.; Guo, J. Advanced dropout: A model-free methodology for bayesian dropout optimization. *IEEE Trans. Pattern Anal. Mach. Intell.* **2021**, *44*, 4605–4625. [CrossRef] [PubMed]
25. 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.
26. Qi, C.; Ren, J.; Su, J. GRU Neural Network Based on CEEMDAN–Wavelet for Stock Price Prediction. *Appl. Sci.* **2023**, *13*, 7104. [CrossRef]
27. Pan, M.; Zhou, H.; Cao, J.; Liu, Y.; Hao, J.; Li, S.; Chen, C.H. Water Level Prediction Model Based on GRU and CNN. *IEEE Access* **2020**, *8*, 60090–60100. [CrossRef]
28. Yue, X.; Ma, G.; Liu, F.; Gao, X. Research on image classification method of strip steel surface defects based on improved Bat algorithm optimized BP neural network. *J. Intell. Fuzzy Syst.* **2021**, *41*, 1509–1521. [CrossRef]
29. Zou, X. Analysis of consumer online resale behavior measurement based on machine learning and BP neural network. *J. Intell. Fuzzy Syst.* **2021**, *40*, 2121–2132. [CrossRef]
30. Chatzimichailidou, M.M.; Martinetti, A.; Majumdar, A.; Dongen LA, V.; Ochieng, W.Y. Wheel maintenance in rolling stock: Safety challenges in the defect detection process. *Int. J. Syst. Syst. Eng.* **2018**, *8*, 387–397. [CrossRef]
31. Zhang, D.; Hu, H.; Liu, Y.; Dai, L. Railway Train Wheel Maintenance Model and Its Application. *Transp. Res. Rec.* **2014**, *2448*, 28–36. [CrossRef]
32. Wang, S.; Guo, H.; Zhang, S.; Barton, D.; Brooks, P. Analysis and prediction of double-carriage train wheel wear based on SIMPACK and neural networks. *Adv. Mech. Eng.* **2022**, *15*, 1929–2958. [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.