Article

Performance Degradation Prognosis Based on Relative Characteristic and Long Short-Term Memory Network for Components of Brake Systems of in-Service Trains

Jingxian Ding 1,2 and Jianyong Zuo 2,3,*

1 Postdoctoral Station of Mechanical Engineering, Tongji University, Shanghai 201804, China
2 Institute of Rail Transit, Tongji University, Shanghai 201804, China
3 Shanghai Key Laboratory of Rail Infrastructure Durability and System Safety, Shanghai 201804, China
* Correspondence: zuojy@tongji.edu.cn; Tel.: +86-138-1615-2980

Abstract: During the service life of brake systems, performance degradation of the components is inevitable. In order to grasp the health status of components of brake systems, and aiming at the problem that the performance degradation trend of the components of the brake system is not completely clear due to signal coupling between components, the influence of variable working conditions, and the long performance degradation cycle, a performance degradation prognosis method of the components of the brake system based on relative characteristic (RC) and the long short-term memory (LSTM) network was proposed. The input and output signals of the components were isolated and fused, the working condition-independent RC was extracted to construct the health indicator (HI), and the validity of the HI was tested by using the monotonicity, correlation, and robustness metrics. Moreover, considering the time memory characteristics, the trend prediction of the HI curve of the components of the brake system was carried out based on the LSTM network. Furthermore, data augmentation for the training and testing sets was performed. Taking the typical component of brake systems as an example, a performance degradation test was carried out. The analysis results of the test data show that the accuracy of the performance degradation prognosis of the intake filter was over 99%, which validates the effectiveness and accuracy of the proposed method. The research results could provide a reference for health management and to improve the active safety protection capability of brake systems of in-service trains.

Keywords: in-service trains; components of brake system; performance degradation prognosis; relative characteristic; long short-term memory

1. Introduction

The special advantages of energy-saving, high speed, and comfort have made rail transit more and more popular in recent years. Furthermore, the characteristics of high speed, long service time, and large passenger capacity determine the importance of safety and reliability for equipment of in-service trains, which have received widespread attention [1,2]. As safety equipment in trains, the brake system plays a very important role in slowing down or stopping trains. However, although the brake system is required to have a high design reliability, due to the harsh operating environment, such as vibration, humidity, and electromagnetic interference, the components of the brake system are subject to temperature stress, mechanical stress, and other multi-source loads. As a result, performance degradation may occur inevitably, which further affects the operational safety, efficiency, quality, and maintenance costs of trains [3]. Fault prognosis, detection, and diagnosis are effective ways to improve system active safety ability and reduce accident risk [4]. In recent years, research on fault detection and diagnosis methods for components of brake systems have been reported [5]. Lu [6], Zuo [7], and Zhou [8] proposed data-driven methods to detect and diagnose faults of sensors, pneumatic units, and brake cylinders, respectively. Seo [9]...
conducted fault diagnosis for the solenoid valve of brake systems with embedded sensor signals and physical interpretation. However, few reports on performance degradation prognosis for components of brake systems of in-service trains have been reported.

Generally, prognosis methods are mainly divided into two categories: model-based prognosis methods, and data-driven prognosis methods [10–12]. As a brake system is typically a mechanic–electric–pneumatic-coupled and time-varying nonlinear system with a complex structure and multiple operating modes [13], it is not easy to establish an accurate degradation mathematical model, while data-driven prognosis methods focus on data characteristics [14,15], and a large amount of operational data is generated during the service period of components of brake system. Therefore, data-driven methods are more suitable for performance degradation prognosis for components of brake systems, and there are two main steps to achieve this process: constructing a health indicator and selecting a suitable prognosis algorithm. To address the problem of health indicator construction, Camci [16] extracted time-domain characteristic parameters of vibration signals as health indicators for bearing performance degradation prediction. Qian [17] extracted the recursive graph entropy characteristics of vibration signals to predict the performance degradation process of bearings. In addition to the direct extraction of feature parameters as health indicators, some scholars have also used indirectly obtained health indicators for degradation prediction by fusing multiple source feature parameters. Qiu [18] fused time domain feature parameters by using a self-organizing mapping method and treated the minimum quantization error obtained as the health indicator of rolling bearings. Huang [19] used the minimum quantization error as a health indicator and proposed a residual life prediction method for ball bearings. Yu [20] used a dynamic PCA method to obtain the principal elements and calculate the equine distance to construct a health indicator.

To address the problem of prognosis algorithm selection, the state variables of components of brake systems form a time series during its performance degradation process, and the time series prognosis algorithms commonly used include traditional probabilistic statistical analysis-based algorithms, and machine learning-based algorithms [21]. Of these, prognosis models based on traditional probabilistic statistical analysis methods mainly include autoregressive models, autoregressive sliding average models, and autoregressive integrated sliding average models [22,23], which are simple and easy to implement. However, they are generally suitable for prognosing stable time series and are not suitable for brake systems with variable operating conditions, continuous non-periodicity, large dynamic range of input and output time series, and strong stochasticity. With the continuous development of machine learning methods, artificial neural networks (ANNs) and support vector regression (SVR) have been widely used in the field of nonlinear time series prediction. Li [24] used radial basis function neural networks to predict the vibration trends of aircraft engines. You [25] proposed a real-time health state prediction method for batteries of electric vehicle by using neural networks. Gao [26] proposed a gyroscopic drift time series prediction method based on phase space reconstruction and SVR. Shen [27] proposed a multivariate support vector machine (SVM)-based method for predicting the remaining life of rolling bearings. Tran [28] used SVM for machine performance degradation assessment and remaining useful life prediction. Although ANN and SVR have a high prediction accuracy for non-linear time series, the algorithms themselves do not adequately consider the temporal memory properties (continuity and correlation) between the data points in the state variable time series during the performance degradation process. In fact, the performance degradation process of components of brake systems is generally slow, varying, and closely related to time. For example, relay valves of brake systems are dominated by wear-and-tear failures and characterized by a gradual performance degradation process. In order to make full use of time series data for prediction, Heimes [29] proposed a recurrent neural network (RNN) to estimate the remaining useful life. The input information of each layer of the RNN depends on the output information of the previous layer, introducing the concept of time series into the design of the network structure and considering the correlation between data points in the time series. However, RNN suffers
from a lack of long-term memory capacity and tends to suffer from gradient disappearance when processing long-term data [30]. To address this problem, Hochreiter [31] proposed a long short-term memory (LSTM) network, which not only takes into account the correlation between data points in a time series, but also effectively improves the long-term dependency of RNN, making it possible to predict performance degradation trends considering historical and real-time operational data. Wang [32] proposed a method for predicting the remaining life of bearings based on convolutional long and short-term memory neural networks. Qi [33] proposed a passenger flow prediction method based on LSTM networks by considering both historical and real-time data. Wu [34] used LSTM networks to estimate the remaining service life of engineering systems and demonstrated the predictive performance of the method by using a case study on an aircraft turbofan engine. However, due to the problem that the signals between the components of the brake system are coupled, the magnitude of the signal is affected by variable operating conditions, and the performance degradation cycle is long, the performance degradation trend is not completely clear.

In this paper, a performance degradation prognosis method for components of brake systems of in-service trains was proposed based on relative characteristic (RC) and the long and short-term memory (LSTM) network. Taking a typical component of brake systems—the intake filter—as an example, RC was extracted based on degradation mechanism analysis. Considering working condition independence, a health indicator based on RC was constructed. Furthermore, considering time dependence, the performance degradation trend of components of the brake system was predicted based on the LSTM network.

The remaining parts of this paper are organized as follows: Section 2 introduces the construction methods for health indicators of components of brake systems. Section 3 presents the prognosis method. Section 4 describes case studies to demonstrate the effectiveness of the proposed method. Finally, conclusions are drawn in Section 5, with some perspectives on research and development.

2. Health Indicator of Components of Brake System

2.1. Brake System of Trains

The microcomputer-controlled straight-through electropneumatic brake system, which mainly consists of a driver controller, a compressed air supply unit (CASU), an electronic brake control unit (EBCU), a pneumatic brake control unit (PBCU), and a basic brake unit (BBU), is widely adopted in trains.

As shown in Figure 1, the driver controller generates brake command signals and transmits them to the EBCU for brake force calculation and distribution. The CASU supplies compressed air to the PBCU. The PBCU generates brake cylinder pressure according to the instructions of the EBCU. The BBU transfers brake cylinder pressure into mechanical brake force [35].

2.2. Health Indicator Construction

One of the key issues in performance degradation prognosis for components of brake systems is the use of information such as service data or characteristics of the components to construct a health indicator (HI). In general, HIs can be divided into two categories: physical health indicator (PHI), and virtual health indicator (VHI). The former are the condition monitoring quantities directly related to the degradation mechanism, while the latter are calculated by fusing multiple condition monitoring quantities or characteristic parameters extracted from them. As brake system has a complex structure and the working medium has continuity, the fault influence of a component will propagate and trigger a chain reaction, making the signals between the components coupled with each other. In addition, the amplitude of the input and output signals of each component will change dynamically with the changing operating conditions, even under normal conditions, and the directly obtained condition monitoring quantity has non-stationary characteristics.
2.2. Health Indicator Construction

In order to effectively characterize the performance degradation state of each component, firstly, it is necessary to monitor the input and output signals of each component separately so that they can be isolated. Secondly, by analyzing the degradation mechanism and integrating the input and output signals of each component and their characteristics, the relative characteristic (RC), which is independent of the working conditions and can accurately and effectively reflect the health status of each component, is extracted to construct a virtual health indicator of the components of the brake system, and the virtual health indicator based on RC is expressed as:

\[ VHI_{RC} = f_h(x_1, x_2, \cdots, x_p) \]  

where \( f_h(\cdot) \) is the constructor of the virtual health indicator, and the specific type of constructor is designed according to the degradation mechanism or signal characteristics of each component; \( x_1, x_2 \ldots, x_p \) are input and output condition monitoring quantities for components of the brake system or their characteristic parameters.

2.3. Health Indicator Evaluation

In order to evaluate the validity of the above RC-based health indicator, a number of evaluation indicators needs to be selected. In general, the performance degradation process of components is irreversible, and therefore the health indicator should have a monotonic degradation trend, i.e., monotonicity. Secondly, the performance degradation process is closely related to time, and there should be a correlation between the health indicator of components and time, i.e., correlation. Finally, the health indicator should be resistant to interference, i.e., robustness. Therefore, the three evaluation indicators of monotonicity, correlation, and robustness were adopted in this paper.

For a time series of component health indicators containing \( n \) samples, i.e., \( RC = (rc_1, rc_2, \ldots, rc_n) \), monotonicity, correlation, and robustness are expressed as:

**Monotonicity**

\[ \text{Mon}(RC) = \frac{\left| \sum_i \epsilon(rc_i - rc_{i-1}) - \sum_i \epsilon(rc_{i-1} - rc_i) \right|}{n - 1} \]  

\[ (2) \]
where \( \varepsilon(x) = \begin{cases} 1, & x \geq 0 \\ 0, & x < 0 \end{cases} \) is a unit step function.

Correlation

\[
\text{Corr}(RC) = \frac{|n \sum rc \cdot t - \sum rc \cdot \sum t|}{\sqrt{n \sum rc^2 - (\sum rc)^2} \sqrt{n \sum t^2 - (\sum t)^2}}
\]  

(3)

where \( T = (t_1, t_2, \ldots, t_n) \) is the sampling time series.

Robustness

\[
\text{Rob}(RC) = \frac{1}{n} \sum e^{-\frac{|\tilde{rc}_i - \tilde{rc}_i|}{n}}
\]  

(4)

where \( \tilde{RC} = (\tilde{rc}_1, \tilde{rc}_2, \ldots, \tilde{rc}_n) \) is the trend term of a time series of health indicators for components.

After constructing accurate and valid health indicators, the design of suitable prediction algorithms directly affects the accuracy of the performance degradation prognosis results. In fact, as the safety stopping devices of trains, the components of brake systems are generally required to have a high design reliability, which means that the performance degradation cycle of the components of brake systems is generally long and belongs to the category of long-term prediction of time series. It is clear from the literature review that the LSTM network with deep learning capabilities takes into account the temporal memory properties among data points in time series and has good long-term prediction performance. Therefore, the LSTM network was used in this paper to predict the performance degradation trend of components of brake systems of in-service trains.

3. LSTM-Based Method for Performance Degradation Prognosis of Components of Brake Systems

3.1. Long Short-Term Memory Network

As a deep learning method, the long short-term memory (LSTM) network is designed to overcome the gradient disappearance problem of the recurrent neural network (RNN) when processing long-term data, and it is suitable for processing time series data with long-term dependencies. In addition, the LSTM network is enhanced by several layers of non-linear transformations and has a long-term information memory function, making it ideal for predicting important events with relatively long intervals and delays in the time series.

Similar to back propagation (BP), RNN, and other neural networks, the LSTM network also consists of three layers, namely, the input layer, the hidden layer, and the output layer, but the hidden layer adds an input gate, a forget gate, and an output gate, which are used to determine whether to retain existing information. Specifically, the hidden layer of the LSTM network consists of a cell, an input gate, an output gate, and a forget gate. The cell is responsible for remembering the value at any time interval, and the three gates are responsible for regulating the flow of information in and out of the cell. This improved design of the LSTM network not only takes into account the correlation between data points in the time series, but it also allows for the discarding of unwanted information and enables good long-distance transmission of useful information from previous data.

The basic structure of the LSTM network is illustrated in Figure 2.

In Figure 2, \( c_{t-1} \) and \( h_{t-1} \) denote the state of the cell and the output of the hidden layer at time \( t - 1 \), respectively; \( x_t, c_t, \) and \( h_t \) respectively denote the input data, the state of the cell, and the output of the hidden layer at time \( t \); \( f_t, i_t, \) and \( o_t \) respectively denote the output of the forget gate, the input gate, and the output gate at time \( t \); \( \sigma \) denotes the sigmoid activation function, and the range of its output is \([0, 1]\); \( \tanh \) denotes the hyperbolic tangent activation function, and the range of its output is \([-1, 1]\); and \( \odot \) denotes the Hadamard product, and the output is 0 or 1, indicating the amount of information allowed to be delivered.
The message forwarding transfer process of the LSTM network consists of four steps, i.e., the forget gate, update input information, update cell state, and network output information, and the details are as follows:

1) Forget gate

Firstly, the \( f_t \) can be obtained based on the output \( h_{t-1} \) of the hidden layer at the previous moment and input data \( x_t \) at the current moment, and the historical state information \( c_{t-1} \) of the cell at the previous moment can be determined to pass or not. The forget gate adopts the sigmoid activation function and outputs the number from the range of \([0, 1]\); among them, “0” denotes completely abandoned, while “1” denotes completely reserved. The calculation formula is as follows:

\[
f_t = \sigma(w_{fh} \cdot h_{t-1} + w_{fx} \cdot x_t + b_f)
\]  

(5)

2) Update input information

Secondly, deciding what new information is stored in the cell is carried out. The input gate uses a sigmoid activation function to generate the data \( i_t \) used to determine which values are to be updated, and \( \tanh \) is used to generate new candidate values \( \tilde{c}_t \). The calculation formula is as follows:

\[
i_t = \sigma(w_{ih} \cdot h_{t-1} + w_{ix} \cdot x_t + b_i)
\]

(6)

\[
\tilde{c}_t = \tanh(w_{ch} \cdot h_{t-1} + w_{cx} \cdot x_t + b_c)
\]

(7)

3) Update cell state

Then, the state of the cell \( c_{t-1} \) from the previous moment multiplies \( f_t \) element by element to discard unwanted information, and then adds \( i_t \odot \tilde{c}_t \) to obtain the candidate value \( c_t \) of the cell to thus achieve the state updating of the cell. The calculation formula is as follows:

\[
c_t = c_{t-1} \odot f_t + i_t \odot \tilde{c}_t
\]

(8)

4) Network output information

Finally, an initial output \( o_t \) is obtained via the sigmoid activation function, and \( \tanh \) is used to scale \( c_t \) to \([-1,1]\) to prevent gradient explosion. The output of the hidden layer \( h_t \) of the LSTM network model can be obtained by multiplying with the initial output \( o_t \) element by element. The calculation formula is as follows:

\[
o_t = \sigma(w_{oh} \cdot h_{t-1} + w_{ox} \cdot x_t + b_o)
\]

(9)

\[
h_t = \tanh(c_t) \odot o_t
\]

(10)
where $w_{fh}$, $w_{ih}$, $w_{ch}$, and $w_{oh}$ respectively denote the connection weights between the forget gate, the input gate, the cell, the output gate at the previous moment, and the output layer; $w_{fx}$, $w_{ix}$, $w_{cx}$, and $w_{ox}$ respectively denote the connection weights between the input layer at the current moment and the forget gate, the input gate, the cell, and the output gate; and $b_f$, $b_i$, $b_c$, and $b_o$ respectively denote the biases of the forget gate, input gate, cell, and output gate. The above parameters can be obtained by training the data from the training sample set.

The aim of training the LSTM network is to obtain the optimal parameters (connection weights and biases) to minimize the loss function, and the sum of the squared errors between the actual and predicted values is used as the loss function in this paper, expressed as:

$$J = \sum_{k=1}^{n} (y_p(k) - y(k))^2$$  \hspace{1cm} (11)

where $n$ is the number of samples, $y_p(k)$ is the predicted value, and $y(k)$ is the actual value.

The LSTM network differs from the standard feed-forward neural network, and it has a feedback connection. The parameters in the network are adjusted and updated by using error reverse transmission, and the parameters to be iterated include the input weight matrix of the forget gate, the input gate, the cell, and the output gate and the hidden layer weight matrix. Since the LSTM network consists of two lines of cells and a hidden layer composed of multiple gates, parameter reverse transmission also needs to be performed in two parts, and reverse transmission of the hidden layer requires a chain derivation. Taking the forget gate as an example, the weight updating process is as follows:

Let loss function at time $t$ be $J_t$; then the gradient of $J_t$ against the $W_f$ of forget gate is expressed as

$$\frac{\partial J_t}{\partial W_f} = \frac{\partial h_t}{\partial f_t} \cdot \frac{\partial f_t}{\partial W_f} = \frac{\partial h_t}{\partial h_t} \cdot \frac{\partial h_t}{\partial c_t} \cdot \frac{\partial c_t}{\partial f_t} \cdot \frac{\partial f_t}{\partial W_f}$$  \hspace{1cm} (12)

Based on Formula (10), the partial derivative of the hidden layer output $h_t$ to cell $c_t$ can be calculated as

$$\frac{\partial h_t}{\partial c_t} = o_t \cdot [1 - \tanh^2(c_t)]$$  \hspace{1cm} (13)

Based on Formula (8), the partial derivative of cell $c_t$ to $f_t$ can be calculated as

$$\frac{\partial c_t}{\partial f_t} = c_{t-1}$$  \hspace{1cm} (14)

Based on Formula (5), the partial derivative of $f_t$ to $W_f$ can be calculated as

$$\frac{\partial f_t}{\partial W_f} = f_t(1 - f_t)S_t$$  \hspace{1cm} (15)

Coupling Formulas (12)–(15), the gradient of $J_t$ to $W_f$ can be calculated as

$$\frac{\partial J_t}{\partial W_f} = \delta h_t \cdot o_t \cdot [1 - \tanh^2(c_t)] \cdot c_{t-1} \cdot f_t(1 - f_t)S_t$$  \hspace{1cm} (16)

Similarly, the parameter gradients in the input gate, cell, and output gate can be obtained. After calculating the gradient of each parameter using error reverse transmission, the parameters are updated iteratively until convergence, and the training of the LSTM network-based prediction model is completed.

3.2. Data Augmentation for the Training and Testing Sets

According to the Takens delay-embedding theorem, there is a certain functional relationship between the future value of a time series and its previous $m_q$ values, i.e., given
a time series with length $n \ X_n = (x_1, x_2, \cdots, x_n)$, the known data at and before time $t$ can be used to predict the data at time $t + 1$, which is expressed as

$$\hat{x}_{t+1} = f(x_t, x_{t-1}, \cdots, x_{t-mq+1}) \tag{17}$$

where $mq$ is the embedding dimension.

Based on the above ideas, in order to make full use of the acquired data for prognosis, data augmentation was carried out for the training and testing sets. Specifically, unlike utilizing the input vector and output vector directly, new input vector $X$ and output vector $Y$ were constructed according to Formulas (18) and (19).

$$X = \begin{bmatrix} x_{m+1} \\ x_{m+2} \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} y_1 & y_2 & \cdots & y_m & x_{m+1} \\ y_2 & y_3 & \cdots & y_{m+1} & x_{m+2} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ y_{n-m} & y_{n-m+1} & \cdots & y_{n-1} & x_n \end{bmatrix} \tag{18}$$

$$Y = \begin{bmatrix} y_{m+1} \\ y_{m+2} \\ \vdots \\ y_n \end{bmatrix} \tag{19}$$

In summary, the flowchart for conducting a prognosis of the performance degradation of the components of brake systems of in-service trains is shown in Figure 3.

![Figure 3. Prognosis flowchart.](image-url)

4. Experiment Verification

4.1. Performance Degradation Test Data Collection

4.1.1. Performance Degradation Test Data Collection

As the source of compressed air for brake systems of trains, the quality of the compressed air produced by the compressed air supply unit (CASU) directly affects the reliability and service life of other components of the brake system due to the continuity of the compressed air. As the CASU draws air directly from the external natural environment, the air is mixed with impurities such as water vapor, sand, and dust, and direct use can cause blockage, stagnation, wear, or corrosion of valves, pipelines, and other components, which can affect the normal operation of the brake system. The intake filter of the CASU is used to remove impurities such as sand and dust from the air to improve the cleanliness of the intake air, and its filtration performance directly affects the quality of compressed air. To verify the effectiveness of the proposed performance degradation prognosis method, the air
intake filter was selected, and the performance degradation test of the CASU of the brake system was carried out. By adding sensors to the CASU, the performance degradation test data of the intake filter were collected. The performance degradation test rig of the compressed air supply unit is shown in Figure 4.

![Performance degradation test rig of compressed air supply unit](image)

**Figure 4.** Performance degradation test rig of compressed air supply unit.

The performance degradation test rig of the compressed air supply unit was mainly composed of a compressed air supply unit (CASU); a total air cylinder; collection equipment; air pressure, temperature, flow rate, and other sensor sets; a temperature box, etc. In order to realistically simulate the stress of the working environment, the inlet air temperature of the CASU was set to the actual ambient temperature, the inlet air pressure was the atmospheric pressure in the actual environment, the load was the same volume of the total air cylinder as what is used on the vehicle, and the total air pressure was controlled at around 900 kPa. By supplying three-phase AC 380 V voltage to the CASU, the CASU was driven to work continuously with load. During the test, the input and output signals of the intake filter as well as the ambient temperature were collected in real time by the sensor sets, as shown in Figure 5. The sampling interval for the performance degradation test of the CASU was set as 600 s, and the sampling frequency was 1 Hz.

![Performance degradation test curves of intake filter](image)

**Figure 5.** Performance degradation test curves of intake filter.
4.1.2. HI of Intake Filter

The degradation mechanism of the intake filter was analyzed as follows: Under the action of the pressure difference between the front and rear end of the intake filter, the mixture flows through the filter medium of the intake filter, and the solid particles or liquid in the mixture are retained. The filtered air is passed through the pores of the filter medium, and thus the impurities such as sand and dust mixed in the air are separated. Impurity particles such as sand and dust are captured by the cartridge fibers and accumulate on the surface of the cartridge. With the accumulation of impurity particles, the gaps between the cartridge fibers are gradually blocked, the effective circulation area of the cartridge becomes smaller, the resistance becomes larger, and the intake filter gradually degrades. The resistance of the intake filter is generally expressed in terms of the pressure loss, and the resistance is related to the speed of air flow and the effective circulation area of the cartridge, expressed as:

\[ \Delta p = P_{\text{in}} - P_{\text{out}} = b \left( \frac{m_{\text{air}}}{S} \right)^n \]

(20)

where \( \Delta p \) is pressure loss of the intake filter, \( P_{\text{in}} \) is inlet pressure, \( P_{\text{out}} \) is outlet pressure, \( S \) is the effective flow area, \( m_{\text{air}} \) is air mass flow rate, and \( b \) and \( n \) are constants.

From the above theoretical analysis of the intake filter degradation mechanism, it is clear that the degradation degree of the intake filter can be assessed by measuring the input and output pressures and calculating the pressure loss, i.e., the health indicator of the intake filter can be expressed as

\[ H I_{jq} = \Delta p \]

(21)

For the raw input and output data of the intake filter obtained in Section 4.1.1, after five-point sliding average filtering of the raw data, the health indicator curve for the intake filter during performance degradation was calculated according to Equations (20) and (21), as shown in Figure 6.

![Figure 6. Health indicator curve of intake filter.](image)

As can be seen in Figure 6, as the performance degradation time increased, the health indicator (HI) of the intake filter fluctuated but showed an overall trend of gradually becoming larger, which can reflect the performance degradation phenomenon of the intake filter. Specifically, \( H I_{jq} \) remained relatively flat in the beginning phase and started to increase with time. The monotonicity of the inlet filter HI was 0.13, the correlation was 0.9, and the robustness was 0.4, which further proves that the HI of the intake filter has a high lifetime correlation.
4.2. Performance Degradation Prognosis Results

In order to train and calibrate the prediction model, the HI time series data were further divided into a training set and a test set. Moreover, the LSTM network descended in the opposite direction of the gradient, so that the prediction value approximated the true value by iterative approximation, i.e., the loss function $J$ was minimized. The normalized pre-processing could make the prediction model smoother in the gradient descent process, which helped converge to the optimal solution and also avoided the problem of poor fitting due to gradient explosion. Therefore, the z-score method was used to normalize the training set data and the test set data.

After normalizing the HI of the intake filter, the first 90% of the test data were selected as the training set and the remaining 10% as the testing set. The LSTM network was trained using the training set, and the trained prediction model was used to predict the trend of the HI of the intake filter. The Adam optimization algorithm was used to create an LSTM network, and the following parameters were set for the LSTM network: the number of nodes in the input layer was 3, the number of nodes in the output layer was 1, the number of implied units was $18 \times 3$, the number of iterations was 250, the gradient threshold was set to 1, the initial learning rate was 0.005, and the learning rate was reduced after 125 rounds of training by multiplying by a factor of 0.2. The prediction results of the HI of the intake filter based on the LSTM network are shown in Figure 7.

![Figure 7. Performance degradation prediction results of the intake filter based on the LSTM network.](image)

As can be seen in Figure 7, the predicted values of the HI of the intake filter based on the LSTM network could follow the measured values very well, and the error between the predicted and measured values was very small, with a maximum error of less than 0.02 kPa, which proved that the performance degradation prognosis method based on the LSTM network had good prediction accuracy.

4.3. Prediction Results of the Comparison Method

To further validate the effectiveness and accuracy of the LSTM network-based performance degradation prognosis method for components of brake systems, a prediction model based on the PSO-SVR method was used to compare the prediction results of the HI of the intake filter.

Specifically, the PSO algorithm was used to optimize the model parameters of the prediction model based on the SVR method, and the training and testing sets were constructed in the same way as those based on the LSTM network. The prediction results of the PSO-SVR-based prediction model are shown in Figure 8.

![Figure 8. Performance degradation prediction results based on the PSO-SVR method.](image)
As can be seen in Figure 7, the predicted values of the HI of the intake filter based on the LSTM network could follow the measured values very well, and the error between the predicted and measured values was very small, with a maximum error of less than 0.02 kPa, which proved that the performance degradation prognosis method based on the LSTM network had good prediction accuracy.

4.3. Prediction Results of the Comparison Method

To further validate the effectiveness and accuracy of the LSTM network-based performance degradation prognosis method for components of brake systems, a prediction model based on the PSO-SVR method was used to compare the prediction results of the HI of the intake filter.

Specifically, the PSO algorithm was used to optimize the model parameters of the prediction model based on the SVR method, and the training and testing sets were constructed in the same way as those based on the LSTM network. The prediction results of the PSO-SVR-based prediction model are shown in Figure 8.

![Figure 8. Performance degradation prediction results based on the PSO-SVR method.](image)

As can be seen from Figures 8 and 9, both the prediction model based on the PSO-SVR method and the prediction model based on the LSTM network could follow the measured value, but the prediction result of the prediction model based on the LSTM network matched the curve of the measured value better. In order to compare the predictive performance of the two aforementioned prediction models more comprehensively and accurately, root mean squared error (RMSE), mean absolute error (MAE), maximum absolute error (MAXERROR), mean absolute percent error (MAPE), mean relative percent error (MRPE), Accuracy (A), and Score were selected as evaluation indicators, and the calculation formulas for each evaluation indicator are expressed as:

\[
RMSE = \sqrt{\frac{1}{N} \sum_{k=1}^{N} \frac{(y_p(k) - y(k))^2}{y(k)^2}}
\]

\[
MAE = \frac{1}{N} \sum_{k=1}^{N} |y_p(k) - y(k)|
\]

\[
MAXERROR = Max |y_p(k) - y(k)|
\]

\[
MAPE = \frac{1}{N} \sum_{k=1}^{N} \frac{|y_p(k) - y(k)|}{y(k)} \times 100\%
\]

\[
MRPE = \frac{1}{N} \sum_{k=1}^{N} \frac{100|y_p(k) - y(k)|}{y(k)}
\]

\[
A = 1 - \sqrt{\frac{1}{N} \sum_{k=1}^{N} \frac{(y_p(k) - y(k))^2}{y(k)^2}} \times 100\%
\]

\[
Score = \begin{cases} \sum_{k=1}^{N} \frac{y(k) - y_p(k)}{y(k)} - 1, & y_p(k) - y(k) < 0 \\ \sum_{k=1}^{N} \frac{y_p(k) - y(k)}{y(k)} - 1, & y_p(k) - y(k) \geq 0 \end{cases}
\]

where \(N\) is the prediction sample size, \(y_p(k)\) is the predicted value, and \(y(k)\) is the measured value.

According to Formulas (22)–(28), the evaluation indicators of the prediction results of the two prediction models were calculated, as shown in Table 1.
It can be seen from the results in Table 1 that the prediction accuracy of the LSTM network-based prediction model and the PSO-SVR method-based prediction model using the same prediction steps were 99.94% and 99.54%, respectively, and the former was much more accurate. In addition, compared to the comparison model, the prediction model using the method proposed in this paper had a 12.2% reduction in RMSE, a 10.7% reduction in MAE, a 10.5% reduction in MAXERROR, a 3.3% reduction in MAPE, a 9.1% reduction in MRPE, a 0.4% improvement in A, and a 0.02% reduction in Score. In summary, the prediction model based on LSTM networks had better long-term prediction performance and higher reliability than the prediction model based on the PSO-SVR method.

![Comparison curves of performance degradation predicted results.](image)

**Figure 9.** Comparison curves of performance degradation predicted results.

### 4.4. The Impact of the Number of Input Nodes on Prediction Performance

The above modeling of the intake filter performance degradation prediction model used the data from the previous three moments as inputs to predict the data for the current moment. In order to investigate the effect of modelling with different numbers of input nodes on the prediction performance of the LSTM network-based prediction model proposed in this paper, a comparison test of the performance degradation prediction with different numbers of input nodes was conducted, and the results are shown in Table 2.

| Number of Input Nodes | RMSE   | MAE    | MAXERROR | MAPE   | MRPE   | A (%) | Score     |
|-----------------------|--------|--------|----------|--------|--------|-------|-----------|
| 1                     | 0.0037 | 0.0026 | 0.0182   | 0.0030 | 0.2955 | 99.85 | 361.0308  |
| 2                     | 0.0036 | 0.0025 | 0.0189   | 0.0028 | 0.2811 | 99.90 | 360.0186  |
| 3                     | 0.0036 | 0.0025 | 0.0197   | 0.0029 | 0.2863 | 99.94 | 359.0060  |
| 4                     | 0.0037 | 0.0026 | 0.0215   | 0.0029 | 0.2941 | 99.87 | 358.0242  |

As can be seen from Table 2, the number of input nodes had an influence on the predicted results of the performance degradation prognosis model. Combining the magnitudes of all indicators in Table 2, the prediction performance was better when the number

---

Table 1. Evaluation indicators for the prediction results of the two prediction models.

| Evaluation Indicators | RMSE   | MAE    | MAXERROR | MAPE   | MRPE   | A (%) | Score     |
|-----------------------|--------|--------|----------|--------|--------|-------|-----------|
| LSTM                  | 0.0036 | 0.0028 | 0.0197   | 0.0029 | 0.2863 | 99.94 | 359.0060  |
| PSO-SVR               | 0.0041 | 0.0026 | 0.0220   | 0.0030 | 0.3151 | 99.54 | 359.0827  |
of input nodes was 3. This shows that it is reasonable to use the data of the first three moments as inputs to predict the state of the current moment in this paper, which provides a reference for other components to choose a reasonable number of input nodes to establish the prediction model.

5. Conclusions

This paper investigated the performance degradation prognosis method for components of brake systems of trains.

(1) In view of the coupling of signals between components and the variable operating conditions, the input and output signals of the components were isolated and fused. The relative characteristic that can effectively characterize the degradation state of the components was extracted as health indicators, and the validity of the health indicators was verified by calculating three evaluation indexes of monotonicity, correlation, and robustness.

(2) Considering the time-memory characteristics of components during the performance degradation process, a method based on LSTM networks for trend prediction of the health indicator curves of the components was proposed.

(3) A performance degradation test of the intake filter was carried out, and the validity of the performance degradation prognosis method was analyzed and verified in detail. Furthermore, the prediction results were compared with those of the prediction model based on the PSO-SVR method, verifying that the prediction model based on the LSTM network had higher prediction accuracy.

By accurately predicting the performance degradation trend of the components of the brake system, the health status of the components can be grasped in advance, which can provide a reference for health management such as performance inspection and maintenance planning in engineering applications and improve the active safety protection capability of brake systems of in-service trains.

Author Contributions: Conceptualization, J.Z.; Data curation, J.D.; Funding acquisition, J.Z.; Methodology, J.D.; Software, J.D.; Validation, J.D.; Writing—original draft, J.D. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by National Natural Science Foundation of China (Grant No. 62273258), and Science and Technology Research and Development Programme Topics of China State Railway Group Co., Ltd. (Grant No. N2022J019).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Some or all data are available from the authors upon reasonable request.

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

References

1. Fumeo, E.; Oneto, L.; Anguita, D. Condition Based Maintenance in Railway Transportation Systems Based on Big Data Streaming Analysis. Procedia Comput. Sci. 2015, 53, 437–446. [CrossRef]

2. Bai, Y.C.; An, C.; Li, M.G.; Lin, J.H.; Li, Q. Tracking Research on Service Performance of CRH3 EMU in Wuhan-Guangzhou Passenger Dedicated Line. Railw. Locomot. EMUs 2018, 1, 37–40, 43.

3. Zuo, J.; Ding, J.; Hu, W.; Han, F.; Zhang, L. Performance degradation monitoring based on data fusion method for in-service train pneumatic brake system. Proc. Inst. Mech. Eng. Part C J. Mech. Eng. Sci. 2019, 233, 1924–1938. [CrossRef]

4. Liu, Z.L.; Pan, D.; Zuo, M.J.; Liang, X.L. A Review on Fault Diagnosis for Rail Vehicles. J. Mech. Eng. 2016, 52, 134–146. [CrossRef]

5. Liu, J.; Li, Y.-F.; Zio, E. A SVM framework for fault detection of the braking system in a high speed train. Mech. Syst. Signal Process. 2017, 87, 401–409. [CrossRef]

6. Lu, Y.; Fan, X.; Gao, D.; Cheng, Y.; Yang, Y.; Zhang, X.; Li, S.; Peng, J. A data-based approach for sensor fault detection and diagnosis of Electro-Pneumatic brake. In Proceedings of the 2019 IEEE International Conference on Prognostics and Health Management, San Francisco, CA, USA, 17–20 June 2019.
7. Zuo, J.; Ding, J.; Feng, F. Latent Leakage Fault Identification and Diagnosis Based on Multi-Source Information Fusion Method for Key Pneumatic Units in Chinese Standard Electric Multiple Units (EMU) Braking System. Appl. Sci. 2019, 9, 300. [CrossRef]

8. Zhou, D.; Ji, H.; He, X.; Shang, J. Fault Detection and Isolation of the Brake Cylinder System for Electric Multiple Units. IEEE Trans. Control Syst. Technol. 2018, 26, 1744–1757. [CrossRef]

9. Seo, B.; Jo, S.H.; Oh, H.; Youn, B.D. Solenoid Valve Diagnosis for Railway Braking Systems with Embedded Sensor Signals and Physical Interpretation. In Proceedings of the Annual Conference of the Prognostics and Health Management Society, Denver, CO, USA, 3–6 October 2016; pp. 337–343.

10. Lei, Y.; Li, N.; Gontarz, S.; Lin, J.; Radkowski, S.; Dybala, J. A Model-Based Method for Remaining Useful Life Prediction of Machinery. IEEE Trans. Reliab. 2016, 65, 1314–1326. [CrossRef]

11. An, D.; Kim, N.H.; Choi, J.-H. Practical options for selecting data-driven or physics-based prognostics algorithms with reviews. Reliab. Eng. Syst. Saf. 2015, 133, 223–256. [CrossRef]

12. Baraldi, P.; Cadini, F.; Mangili, F.; Zio, E. Model-based and data-driven prognostics under different available information. Probabilistic Eng. Mech. 2013, 32, 66–79. [CrossRef]

13. Zuo, J.Y.; Ding, J.X. Research progress on intelligent control and maintenance technology of railway vehicle braking system. J. Traffic Transp. Eng. 2021, 21, 50–62.

14. Li, N.; Lei, Y.; Gebrael, N.; Wang, Z.; Cai, X.; Xu, P.; Wang, B. Multi-Sensor Data-Driven Remaining Usefull Life Prediction of Semi-Observable Systems. IEEE Trans. Ind. Electron. 2021, 68, 11482–11491. [CrossRef]

15. Zhao, S.K.; Jiang, C.; Long, X.Y. Remaining useful life estimation of mechanical systems based on the data-driven method and bayesian theory. J. Mech. Eng. 2018, 54, 115–124. [CrossRef]

16. Camci, F.; Medjaher, K.; Zerhouni, N.; Nectoux, P. Feature Evaluation for Effective Bearing Prognostics. Qual. Reliab. Eng. Int. 2013, 29, 477–486. [CrossRef]

17. Qian, Y.; Hu, S.; Yan, R. Bearing performance degradation evaluation using recurrence quantification analysis and auto-regression model. In Proceedings of the IEEE Instrumentation and Measurement Technology Conference, Minneapolis, MN, USA, 6–9 May 2013; pp. 1713–1716.

18. Qiu, H.; Lee, J.; Lin, J.; Yu, G. Robust performance degradation assessment methods for enhanced rolling element bearing prognostics. Adv. Eng. Inform. 2003, 17, 127–140. [CrossRef]

19. Huang, R.; Xi, L.; Li, X.; Liu, C.R.; Qiu, H.; Lee, J. Residual life predictions for ball bearings based on self-organizing map and back propagation neural network methods. Mech. Syst. Signal Process. 2007, 21, 193–207. [CrossRef]

20. Yu, J. Health Condition Monitoring of Machines based on Hidden Markov Model and Contribution Analysis. IEEE Trans. Instrum. Meas. 2012, 61, 2200–2211. [CrossRef]

21. Pei, H.; Hu, C.H.; Si, X.S.; Zhang, J.X.; Pang, Z.N.; Zhang, P. Review of Machine Learning Based Remaining Usefull Life Prediction Methods for Equipment. J. Mech. Eng. 2019, 55, 1–13. [CrossRef]

22. Zhang, Q.; Wang, B.-D.; He, B.; Peng, Y.; Ren, M.-L. Singular Spectrum Analysis and ARIMA Hybrid Model for Annual Runoff Forecasting. Water Resour. Manag. 2011, 25, 2683–2703. [CrossRef]

23. Partanen, S.V.A. Forecasting electricity price and demand using a hybrid approach based on wavelet transform, ARIMA and neural networks. Int. J. Energy Res. 2014, 38, 626–637.

24. Li, F.X.; Chen, G.M.; Zhou, L.; Liu, X.L.; Li, S.C. Hybrid prediction method based on empirical mode decomposition and RBF neural network. Vib. Meas. Diagn. 2012, 32, 817–822.

25. You, G.-W.; Park, S.; Oh, D. Real-time state-of-health estimation for electric traction motor and batteries: A data-driven approach. Appl. Energy 2016, 176, 92–103. [CrossRef]

26. Gao, Y.; Li, Y.Fault Prediction Model Based on Phase Space Reconstruction and Least Squares Support Vector Machines. In Proceedings of the 2009 Ninth International Conference on Hybrid Intelligent Systems, IEEE Computer Society, Shenyang, China, 12–14 August 2009; pp. 464–467.

27. Shen, Z.J.; Chen, X.F.; He, Z.J.; Sun, C.; Zhang, X.L.; Liu, Z.W. Remaining life predictions of rolling bearing based on relative features and multivariable support vector machine. J. Mech. Eng. 2013, 49, 183–189. [CrossRef]

28. Tran, V.T.; Pham, H.T.; Yang, B.-S.; Nguyen, T.T. Machine performance degradation assessment and remaining useful life prediction using proportional hazard model and support vector machine. Mech. Syst. Signal Process. 2012, 32, 320–330. [CrossRef]

29. Felix, O.; Heimes, B.S. Recurrent neural networks for remaining useful life estimation. In Proceedings of the 2008 International Conference on Prognostics and Health Management, Denver, CO, USA, 6–9 October 2008.

30. Bengio, Y.; Simard, P.; Frasconi, P. Learning long-term dependencies with gradient descent is difficult. IEEE Trans. Neural Netw. 1994, 5, 157–166. [CrossRef] [PubMed]

31. Hochreiter, S.; Schmidhuber, J. Long Short-term Memory. Neural Comput. 1997, 9, 1735–1780. [CrossRef] [PubMed]

32. Wang, J.J.; Yang, S.P.; Liu, Y.Q.; Wen, G.L. A Method of Bearing Remaining Useful Life Estimation Based on Convolutional Long Short-term Memory Neural Network. J. Mech. Eng. 2021, 57, 88–95.

33. Qi, Q.Y.; Lv, Y.B.; Ma, J.H.; Li, J. An LSTM-Based Method Considering History and Real-Time Data for Passenger Flow Prediction. Appl. Sci. 2020, 10, 3788.
34. Wu, Y.; Yuan, M.; Dong, S.; Lin, L.; Liu, Y. Remaining useful life estimation of engineered systems using vanilla LSTM neural networks. *Neurocomputing* **2018**, *275*, 167–179. [CrossRef]

35. Zuo, J.; Ding, J.; Liu, C.; Luo, Z. A virtual prototype for performance analysis of electropneumatic brake on metro trains. *Adv. Mech. Eng.* **2020**, *12*, 1687814020926275. [CrossRef]