

Article

# Fault Diagnosis of Brake Train Based on Multi-Sensor Data Fusion

Yongze Jin, Guo Xie \*, Yankai Li, Xiaohui Zhang, Ning Han, Anqi Shangguan and Wenbin Chen

Key Laboratory of Shaanxi Province for Complex System Control and Intelligent Information Processing, Xi'an University of Technology, Xi'an 710048, China; 1190313032@stu.xaut.edu.cn (Y.J.); liyankai@xaut.edu.cn (Y.L.); xhzhang@xaut.edu.cn (X.Z.); 2200320062@stu.xaut.edu.cn (N.H.); 1200313023@stu.xaut.edu.cn (A.S.); 2190321310@stu.xaut.edu.cn (W.C.)

\* Correspondence: guoxie@xaut.edu.cn

**Abstract:** In this paper, a fault diagnosis method is proposed based on multi-sensor fusion information for a single fault and composite fault of train braking systems. Firstly, the single mass model of the train brake is established based on operating environment. Then, the pre-allocation and linear-weighted summation criterion are proposed to fuse the monitoring data. Finally, based on the improved expectation maximization, the braking modes and braking parameters are identified, and the braking faults are diagnosed in real time. The simulation results show that the braking parameters of systems can be effectively identified, and the braking faults can be diagnosed accurately based on the identification results. Even if the monitoring data are missing or abnormal, compared with the maximum fusion, the accuracies of parameter identifications and fault diagnoses can still meet the needs of the actual systems, and the effectiveness and robustness of the method can be verified.

**Keywords:** high-speed train; information fusion; fault diagnosis; parameter identification; unscented Kalman filter (UKF); expectation maximization (EM)



**Citation:** Jin, Y.; Xie, G.; Li, Y.; Zhang, X.; Han, N.; Shangguan, A.; Chen, W. Fault Diagnosis of Brake Train Based on Multi-Sensor Data Fusion. *Sensors* **2021**, *21*, 4370. <https://doi.org/10.3390/s21134370>

Academic Editors: Hamed Badihi, Ningyun Lu and Tao Chen

Received: 12 May 2021  
Accepted: 23 June 2021  
Published: 25 June 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 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

With the increase of speed, the reliability and safety of the train system are put forward with higher requirements. However, influenced by the potential technique abnormalities and component failures, the train system still fails frequently [1–3]. While these failures may not be serious in the early stages, the performance of the system has indeed been degraded [4,5]. Therefore, the early detections and identifications of any potential anomalies and failures are essential, as they avoid dangers for high-speed train operation [6,7].

The high-speed train is composed of multiple subsystems interworking with each other. A stable and reliable braking system is indispensable, to ensure a safe and comfortable operation environment. It can slow down or stop smoothly and timely when needed. Over past few decades, a large number of monitoring, diagnosis, and prediction techniques have been applied to train systems [8]. To be more specific, the fault diagnosis methods based on feature extraction, feature selection, and feature fusion have been studied in [9–13], and the accuracies of fault diagnosis are improved greatly by these methods. The intermittent fault detection, isolation, and diagnosis of train multi-axis speed sensors are addressed in [14–16], and the composite fault diagnosis of rolling equipment such as train bearings has been proposed in [17–19]. These above technologies have greatly improved the level of intelligence in ensuring the safe and reliable operation of trains.

In recent years, with the development of sensors, train monitoring data have become diverse. Therefore, fully mining the value of multi-source data and realizing the fault diagnosis are of great significance based on multi-sensor information fusion [20,21]. In multi-sensor fusion, centralized and distributed fusion methods are mainly used to process the monitoring data. In centralized fusion, all measured data from multiple sensors are stacked into a sensor measurement (with higher dimensions), and the specific fusion

rules are not required [22,23]. However, the application of centralized fusion is limited due to some defects of its own. Therefore, various distributed fusion methods have been proposed. In distributed fusion, the state of each sensor is estimated, and then the estimations are sent to the fusion center by certain fusion criteria. For example, based on linear minimum variance, the matrix-weighted multi-sensor optimal information fusion criterion is investigated in [24,25], and distributed fusion estimations are obtained. On this basis, the multi-sensor optimal data fusion based on adaptive fading unscented Kalman filter (UKF) and unscented transform is presented in [26]. During the train operation, the integrated navigation estimation based on the GPS/INS and tachometer is considered in [27–29]. According to the above literature, it is not difficult to find that, compared with using only a single source of monitoring data, the fusion estimations based on multi-source data can effectively reduce the influences of various adverse factors on the analysis results.

The train-braking process is also affected by many interrelated and coupled factors, which means train braking faults cannot be accurately described and located by the existing methods and models. Hence, in order to ensure the safety of train system, the operating efficiency must be sacrificed. Even so, in extreme weather such as rainstorms and blizzards, there are still hidden dangers to the train safety. Therefore, in order to further improve the operation safety and efficiency, aiming at various brake failures under complex conditions, a train braking model is established, and a fault diagnosis strategy based on the multi-source information fusion is proposed to detect and locate train braking faults in a timely manner. The serious failures will be prevented, and useful guidance can be provided for the component maintenances and replacements.

The paper is organized as follows: In Section 2, the train air braking system is introduced, the train braking model is established, and the specific faults to be diagnosed are expounded. In Section 3, the multi-sensor data fusion strategy, multi-model state prediction, and fault diagnosis method are presented. In Section 4, the detailed experiments are carried out, and a conclusion is drawn in Section 5.

## 2. System Description and Problem Statement

In order to diagnose the faults of the train braking system quickly and accurately, in this section, a train braking model is firstly established based on the mechanism analysis, then combined with the model parameters, the train faults are analyzed, and finally the speed sensor models of the train are built.

### 2.1. Braking Modeling: Air Braking and Adhesion Braking

The actual operation of the train will be affected by the external environments. Regarding the train with complex conditions or sudden emergencies, the air emergency braking will be the last barrier to the safety of the train. The on-line detections and diagnoses of faults in the braking are helpful to prevent more serious faults and provide useful guidance information for the maintenances and replacements of components.

Considering the different characteristics of the train running on dry rail and wet rail, the influences of the wheel-rail adhesion coefficient and adhesion braking force on the train braking performance are analyzed, and the single mass point model of the train braking is established under different operation environments [30]:

$$\begin{cases} v_{t+1} = \begin{cases} v_t - 3.6T \frac{F(\mu) + R_F(v,C)}{M_A} + w_t, F(\mu) < B \\ v_t - 3.6T \frac{B + R_F(v,C)}{M_A} + w_t, F(\mu) \geq B \end{cases} \\ y_t = v_t + e_t \end{cases} \quad (1)$$

where  $v$  is the train braking speed,  $C$  is the train current position,  $T$  and  $M_A$  are the train sampling time and occupancy weight,  $F(\mu)$  is the adhesion braking force,  $B$  is the air braking force,  $R_F(v, C)$  is the operation resistance,  $w_t$  is the random disturbance of speed caused by the external factors, and  $e_t$  is the measurement error.  $F(\mu)$  is jointly affected by the wheel-rail adhesion coefficient  $\mu$  and the weight of the train  $M_A$ , that is:

$$F(\mu) = \mu M_A g \quad (2)$$

where  $g$  is the gravitational acceleration coefficient  $9.8\text{m/s}^2$ . The wheel-rail adhesion coefficient is influenced by the train operation states and external environments, and its value will decrease with the increase of speed and will change with the different rail adhesion conditions. For example, as shown in Figure 1, when the train moves from dry rail to wet rail, the adhesion coefficient will decrease significantly. This will lead to the train adhesion failure, and the train system could easily lock and skid. The air braking force applied to the wheel will be invalid. As a result of this fault, the air brake of the train will shortly fail.



Figure 1. Low wheel-rail adhesion state of the train.

The air braking force  $B$  is affected by several braking parameters, namely:

$$B = \frac{\pi}{4} \frac{r}{R_c} d^2 \times Pre \times \eta \times \gamma_B \times \mu_A \times N \times 10^{-6} \quad (3)$$

where  $d$  is the diameter of brake cylinder,  $r$  is the friction radius of the brake disc,  $R_c$  is the diameter of the wheel, and  $N$  is the total number of brake pads. These parameters can be directly measured by conventional methods and are basically unchanged during the train operation, so they are treated as constants.  $Pre$  is the air pressure of the brake cylinder. During the actual emergency braking, all air pressures will be released.  $\eta$  is the transmission efficiency of the foundation brake device. It reflects the friction between the piston and the cylinder wall, the reverse force of the brake cylinder to release the spring, and the mechanical friction of the transmission mechanism.  $\gamma_B$  is the braking ratio. It represents the ideal expansion multiple of the brake cylinder piston rod when the force is transferred to the brake disc.  $\mu_A$  is the friction coefficient of the brake disc. It changes with the train running state and the brake disc material. As shown in Figure 2, when the train brake disc faults occur, such as degradations or even friction failures, it will lead to the performance decline of the train brake. Even if the brake cylinder applied the maximum air pressures, the train still cannot stop in line with the expected braking distance. Therefore, the friction coefficient has a decisive influence on the braking performance of the train, but its characteristic hidden variables mean it is difficult to directly measure with conventional methods. If the train braking failures are generated by degradations of the brake discs and the adhesion brake failures, the diagnoses of the train braking failures will become difficult.

$R_F(v, C)$  is the sum of the basic resistance  $R_1(v)$  and additional resistance  $R_2(C)$ :

$$R_F(v, C) = R_1(v) + R_2(C) \quad (4)$$

where  $R_1(v)$  is positively correlated with the train running speed  $v$ :

$$R_1(v) = M_A \times (c_0 + c_1 v + c_2 v^2) \times g \times 10^{-3} \quad (5)$$

where  $c_0$  is the rolling resistance coefficient,  $c_1$  is another mechanical resistance coefficient, which is directly proportional to the train speed  $v$ , and  $c_2$  is the air resistance coefficient, which is directly proportional to the  $v^2$ . These values are usually assigned by accumulated experience.

$R_2(C)$  is the combination of additional resistances for ramps, curves, and tunnels:

$$R_2(C) = R_r(C) + R_c(C) + R_t(C) \quad (6)$$

$R_r(C)$  is the additional resistance of the ramp at C:

$$R_r(C) = M_A g \sin(\arcsin(\frac{h_e}{l})) \quad (7)$$

where  $h_e$  and  $l$  represent the height and length of ramps, respectively.

$R_c(C)$  represents the additional resistance of the curve at C:

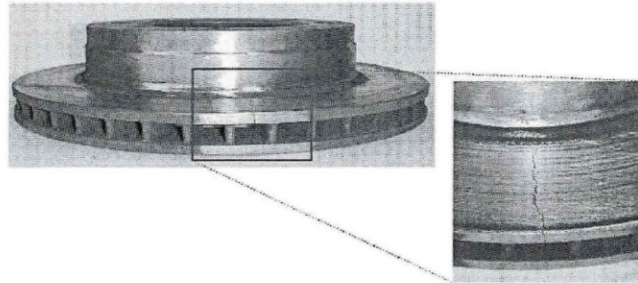
$$R_c(C) = \begin{cases} \frac{0.6M_A g}{R}, & L_0 < L_c \\ \frac{0.0105M_A g \alpha_{L_c}}{L_0}, & L_0 \geq L_c \end{cases} \quad (8)$$

where  $R$  represents the rail radius of the curve,  $L_0$  and  $L_c$  are the length of the train and the curve, respectively, and  $\alpha_{L_c}$  is the deflection angle.

$R_t(C)$  represents the additional tunnel resistance at C:

$$R_t(C) = 0.00013M_A g L \quad (9)$$

where  $L$  represents the length of the tunnel.



**Figure 2.** Severe wear and cracks of brake disc after long-term use.

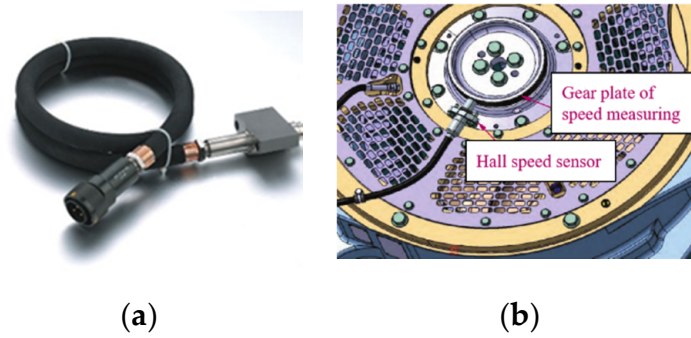
## 2.2. Speed Sensors Modeling

The train speed measurement system is composed of a variety of sensors, including an inertial navigation device (INS), a tachometer, Doppler radar, GPS, etc. Taking the CTCS3-300T train operation control system as an example, the current multi-sensor system architecture is a combination of two tachometers and two Doppler radars.

The tachometer adopts Hall's principle. The wheel Hall speed sensor and its installation example are shown in Figure 3. When the gear disc rotates at each pitch, a counting pulse will be generated. The gear disc contains  $M_{tacho}$  total of pitch, and the wheel diameter is  $R_c$ . If a total of  $p_{tacho}$  pulses are received within the time interval  $T_{tacho}$ , the speed measurement can be expressed as follows:

$$v_{tacho} = \frac{\pi R_c p_{tacho}}{M_{tacho} T_{tacho}} \quad (10)$$

It should be noted that the tachometer will be disturbed by the adhesion state of the rail surface, and the measurement accuracy will be greatly interfered when the train idling or skidding.



**Figure 3.** Wheel Hall speed sensor and installation example: (a) wheel Hall speed sensor; (b) installation example.

The Doppler effect is employed in the Doppler radar speed sensor. The Doppler radar speed sensor and its installation example are shown in Figure 4. The relative speed of train and rail is calculated by the frequency difference between the transmitted wave and ground reflection wave. The number of pulses per kilometer of the radar is recorded as  $N_{dopp}$ . If  $p_{dopp}$  pulses are received in time interval  $T_{dopp}$ , the measured speed can be expressed as follows:

$$v_{dopp} = \frac{1000p_{dopp}}{N_{dopp}T_{dopp}} \quad (11)$$

One characteristic of Doppler radar is that it is not affected by idling wheel skid and wheel diameter change. However, it requires higher installation accuracy and is more susceptible to interference than the tachometer in general.

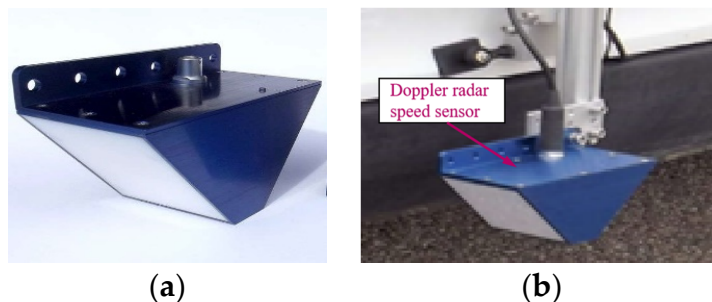
Combined with the train braking model shown in Equation (1), the state space model of wheel speed Hall sensor is:

$$\begin{cases} v_{t+1} = \begin{cases} v_t - 3.6T \frac{F(\mu) + R_F(v,C)}{M_A} + w_t, F(\mu) < B \\ v_t - 3.6T \frac{B + R_F(v,C)}{M_A} + w_t, F(\mu) \geq B \end{cases} \\ p_{tacho_t} = \frac{M_{tacho} T_{tacho}}{\pi R_c} v_t + e_{1t} \end{cases} \quad (12)$$

Similarly, the state space mode of Doppler radar speed sensor is:

$$\begin{cases} v_{t+1} = \begin{cases} v_t - 3.6T \frac{F(\mu) + R_F(v,C)}{M_A} + w_t, F(\mu) < B \\ v_t - 3.6T \frac{B + R_F(v,C)}{M_A} + w_t, F(\mu) \geq B \end{cases} \\ p_{dopp_t} = \frac{N_{dopp} T_{dopp}}{1000} v_t + e_{2t} \end{cases} \quad (13)$$

where  $e_1$  and  $e_2$  are the sensor measurement errors, respectively.



**Figure 4.** Doppler radar speed sensor and installation example: (a) doppler radar speed sensor; (b) installation example.

### 3. Fault Diagnosis Based on Multi-Sensor Data Fusion

In this section, the train braking faults are diagnosed based on the multi-sensor monitoring data. The block diagram of the proposed method is shown in Figure 5. As shown in Figure 5, the multi-sensor data are first fused by the fusion weight pre-allocation and the adaptive fading UKF. Then, the brake parameters are identified based on the improved EM (expectation maximization). Finally, the faults of train braking are diagnosed, according to the analysis of the identification results.

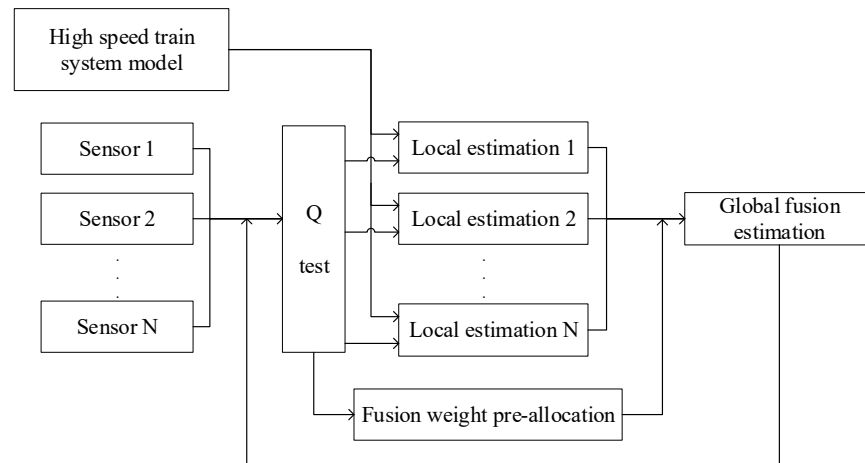


Figure 5. The block diagram of the proposed method.

#### 3.1. Multi-Sensor Monitoring Data Fusion

Assuming that the estimation result obtained by the  $j$ -th ( $j = 1, 2, \dots, N_s$ ) local filter at  $t$  is  $\hat{x}_t^j$ , combined with the fusion weight pre-allocation and the linear weighted fusion criterion, the global state estimation  $\hat{x}_t^*$  at time  $t$  of the system can be described as the following weighted sum:

$$\hat{x}_t^* = \frac{\sum_{j=1}^{N_s} \kappa_t^j \beta_t^j \hat{x}_t^j}{\sum_{j=1}^{N_s} \kappa_t^j \beta_t^j} \quad (14)$$

$$\begin{aligned} \kappa_t &= [\kappa_t^1, \kappa_t^2, \dots, \kappa_t^{N_s}] \\ \sum_{j=1}^{N_s} \kappa_t^j &= 1, \quad \sum_{j=1}^{N_s} \beta_t^j = 1 \end{aligned} \quad (15)$$

where  $\beta_t^j$  is the pre-allocated weight of the  $j$ -th local filter at  $t$ ;  $\hat{x}_t^j$  and  $\kappa_t^j$  are the  $j$ -th local filter result and its fusion weight at  $t$ , respectively. The  $\beta_t^j$ ,  $\hat{x}_t^j$ , and  $\kappa_t^j$  will be inferred in the following sections.

#### 3.2. Qualitative Analysis of Monitoring Data and Pre-Allocation of Fusion Weights

Considering the interferences of the external environments and the accuracies of the sensors it is inevitable that there will be abnormal monitoring data, missing data, and great differences in monitoring results of different types of sensors. The fusion results are affected by these abnormalities seriously, thus the accuracies of diagnoses are reduced. Therefore, through in-depth mining of the information hidden under these abnormalities, the faults can be diagnosed more accurately and efficiently.



In order to detect the abnormal data rapidly and accurately, the Q test is selected. Taking the speed monitoring data of sensors at  $t$  as an example, the  $v_t^i$  is denoted as the speed measurement at  $t$  of the  $i$ -th sensor, the sequence  $V_t$  is defined as  $V_t = [v_t^1, v_t^2, \dots, v_t^{M_s}]$ , and  $M_s$  is the number of sensors. The monitoring data are arranged in an increasing order to get  $V_t^s = [v_t^{1,s}, v_t^{2,s}, \dots, v_t^{M_s,s}]$  ( $3 \leq M_s \leq 10$ ), then the  $Q_1$  is formulated as follows:

$$Q_1 = \frac{|v_t^{i,s} - v_t^{i\pm 1,s}|}{v_t^{M_s,s} - v_t^{1,s}} \quad (16)$$

where  $v_t^{M_s,s}$  and  $v_t^{1,s}$  are the maximum and minimum of monitoring data, respectively, and  $v_t^{i,s}$  and  $v_t^{i\pm 1,s}$  are the measurements of the  $j$ -th sensor and nearest measurements, respectively. The  $Q_2$  is determined according to the number of measurement times and the specified confidence level (e.g., 90%). If  $Q_1 > Q_2$ ,  $v_t^{i,s}$  will be regarded as an abnormality and discarded, otherwise it should be retained. The above steps are repeated for the monitoring data until all the abnormalities are eliminated and the sequence  $V_t^f = [v_t^{1,f}, v_t^{2,f}, \dots, v_t^{N_s,f}] \subseteq V_t$  is obtained.

Although there is no abnormality in  $V_t^f$ , the dispersion of measurements by sensors is different, which means that the external interferences are different. Therefore, in order to improve the accuracies of fusion results, a sensor real-time fusion weight pre-allocation based on distance criterion is proposed as follows:

$$\beta_t^i = 1 - \frac{|v_t^{i,f} - \bar{V}_t^f|}{\sum_{j=1}^{N_s} |v_t^{j,f} - \bar{V}_t^f|} \quad (17)$$

$$\bar{V}_t^f = \frac{1}{N_s} \sum_{k=1}^{N_s} v_t^{k,f} \quad (18)$$

where  $\beta_t^i$  is the pre-allocated fusion weight of the  $i$ -th sensor at  $t$ , and  $\bar{V}_t^f$  is the mean measurement of sensors.

It can be inferred from Equations (17) and (18) that the number of sensors selected at each moment and the pre-allocated fusion weights are constantly changing. Compared with the existing fusion strategies with fixed weights, according to the adaptive weights determinations, the implicit value of data can be fully mined, the adverse influences on fusion filtering results can be eliminated, which contains modeling errors, measurement errors, sensor accuracies, and other adverse factors, and the real attributes of the objects to be tested can be restored as accurately as possible.

### 3.3. Multi-Sensor Data Fusion Based on Adaptive Fading UKF

The filter method and fusion strategy are the focus in this section. The fusion approach is composed of a two-layer fusion structure. In the bottom layer, the adaptive fading UKF based on Mahalanobis distance is used as the local filter [26], and the robustness of local state estimation interfered by the modeling errors is improved. At the top layer, the global state estimation based on the linear minimum variance [25] is deduced.

#### 3.3.1. Local State Estimation

Considering the Markov characteristics of Equations (12) and (13), the local filter estimator of the  $j$ -th model can be expressed as:

$$\begin{cases} x_{t+1} = f(x_t) + w_t \\ z_t = h(x_t) + e_t \end{cases} \quad (19)$$

where  $x_t$  is the system state vector,  $f(\cdot)$  is the nonlinear state function,  $w_t$  is the state noise, which is usually regarded as the zero-mean Gaussian white noise of variance  $Q \geq 0$ ,  $z_t$  is the measurement of the  $j$ -th local filter,  $h(\cdot)$  is the nonlinear measurement function of the  $j$ -th local filter, and  $e_t$  is the measurement noise of the  $j$ -th local filter, which is usually assumed to be the zero-mean Gaussian white noise of variance  $R \geq 0$ .

### 3.3.2. Adaptive Fading UKF Based on Mahalanobis Distance

The Mahalanobis distance of the innovation vector, which is based on hypothesis testing, is commonly adopted to identify system modeling errors for Gaussian systems. In this section, the adaptive fading UKF based on Mahalanobis distance is selected as the local state estimator. By assessing the Mahalanobis distance square of the innovation vector, corresponding measures are taken to improve the adaptability and robustness of UKF for multi-sensor nonlinear stochastic systems modeling. The estimation  $\hat{x}_t^j$  of  $j$ -th sensor at  $t$  is described as follows:

**Step 1.** The initial mean square error matrix and the prior mean value of the state vector are defined as:

$$\hat{x}_0 = E[x_0] \quad (20)$$

$$P_0 = E[(x_0 - \hat{x}_0)(x_0 - \hat{x}_0)^T] \quad (21)$$

**Step 2.** Sigma sampling points of untraced transform are calculated based on sampling strategy:

$$\begin{cases} \tilde{\zeta}_{0,t-1} = \hat{x}_{t-1} & i = 0 \\ \tilde{\zeta}_{i,t-1} = \hat{x}_{t-1} + (\sqrt{(n+\rho)P_{t-1}})_i & i = 1, 2, \dots, n \\ \tilde{\zeta}_{i,t-1} = \hat{x}_{t-1} - (\sqrt{(n+\rho)P_{t-1}})_i & i = n+1, n+2, \dots, 2n \end{cases} \quad (22)$$

where  $n$  is the dimension of the system state vector,  $\rho$  is used to adjust the distance between the sampling points and the original sample points,  $P$  is the covariance matrix of state variables, and  $(\sqrt{(n+\rho)P_{t-1}})_i$  represents the  $i$ -th principal diagonal element of the square-root matrix. The weight coefficients  $\omega_i^1$  and  $\omega_i^2$  of the first-order and second-order statistical characteristics of the sampling points can be expressed by Equation (23):

$$\omega_i^1 = \omega_i^2 = \begin{cases} \frac{\rho}{n+\rho} & i = 0 \\ \frac{1}{2(n+\rho)} & i \neq 0 \end{cases} \quad (23)$$

**Step 3.** The one-step prediction matrix and covariance matrix based on the Sigma sampling points are formulated as:

$$\tilde{\zeta}_{i,t|t-1} = f(\tilde{\zeta}_{i,t-1}) (i = 0, 1, \dots, 2n) \quad (24)$$

$$\begin{cases} \hat{x}_{t|t-1} = \sum_{i=0}^{2n} \omega_i^1 \tilde{\zeta}_{i,t|t-1} + w_t \\ P_{t|t-1} = \sum_{i=0}^{2n} \omega_i^2 [\tilde{\zeta}_{i,t|t-1} - \hat{x}_{t|t-1}] [\tilde{\zeta}_{i,t|t-1} - \hat{x}_{t|t-1}]^T + Q_t \end{cases} \quad (25)$$

**Step 4.** UT transformation is carried out again for the state prediction in Equation (24), and a new Sigma points set is shown in Equation (26):

$$\begin{cases} \tilde{\zeta}'_{0,t|t-1} = \hat{x}_{t|t-1} & i = 0 \\ \tilde{\zeta}'_{i,t|t-1} = \hat{x}_{t|t-1} + (\sqrt{(n+\rho)P_{t|t-1}})_i & i = 1, 2, \dots, n \\ \tilde{\zeta}'_{i,t|t-1} = \hat{x}_{t|t-1} - (\sqrt{(n+\rho)P_{t|t-1}})_i & i = n+1, n+2, \dots, 2n \end{cases} \quad (26)$$



**Step 5.** The points set is substituted into the measurement equation, and the mean of observations is formulated as the weighted sum of the prediction.

$$z_{i,t|t-1} = h(\xi'_{i,t|t-1}) \quad (27)$$

$$\begin{cases} \hat{z}_{t|t-1} = \sum_{i=0}^{2n} \omega_i^1 z_{i,t|t-1} + e_t = \sum_{i=0}^{2n} \omega_i^1 h(\xi'_{i,t|t-1}) + e_t \\ P_{\hat{z}_{t|t-1}} = \sum_{i=0}^{2n} \omega_i^2 [h(\xi'_{i,t|t-1}) - \hat{z}_{t|t-1}][h(\xi'_{i,t|t-1}) - \hat{z}_{t|t-1}]^T + R_t \end{cases} \quad (28)$$

Next, the state estimation covariance matrix is updated as:

$$P_{\hat{x}_t, \hat{z}_t} = \sum_{i=0}^{2n} \omega_i^2 [\xi_{i,t|t-1} - \hat{x}_{t|t-1}] [z_{i,t|t-1} - \hat{z}_{t|t-1}]^T \quad (29)$$

**Step 6.** The Kalman gain is calculated, and the state and covariance are updated as follows:

$$K_t = P_{\hat{x}_t, \hat{z}_t} P_{\hat{z}_{t|t-1}}^{-1} \quad (30)$$

$$\hat{x}_t = \hat{x}_{t|t-1} + K_t (z_t - \hat{z}_{t|t-1}) \quad (31)$$

$$P_t = P_{t|t-1} - K_t P_{\hat{z}_t} K_t^T \quad (32)$$

The above Steps 1–6 are the state estimation processes of classical UKF. The influences of modeling errors, measurement errors, sensor accuracies, and other adverse factors on the filtering results are not taken into account. Therefore, in this paper, the time-varying adaptive fading factor is introduced into the predicted state covariance matrix, and the influence is suppressed, which is generated by prior knowledges on the current state estimation results.

The innovation vector of the  $j$ -th local filter is defined as:

$$\tilde{z}_t = z_t - \hat{z}_{t|t-1} \quad (33)$$

For the nonlinear Gaussian system given in Equation (19),  $\tilde{z}_t$  shall obey the zero-mean Gaussian distribution with the variance as follows:

$$P_{\hat{z}_{t|t-1}} = \sum_{i=0}^{2n} \omega_i^2 [h(\xi'_{i,t|t-1}) - \hat{z}_{t|t-1}][h(\xi'_{i,t|t-1}) - \hat{z}_{t|t-1}]^T + R_t \quad (34)$$

The square of the Maobanobis distance of the innovation vector should obey the  $\chi^2$  distribution of  $m$  degrees of freedom, namely:

$$M_t^2 = \tilde{z}_t^T (P_{\hat{z}_{t|t-1}})^{-1} \tilde{z}_t \sim \chi_m^2 \quad (35)$$

where  $m$  is the measurement dimension of the  $j$ -th local filter.

In accordance with the hypothesis testing theory, for a given significance level  $\alpha$ , the probability is denoted as:

$$P(M_t^2 \leq \chi_{m,\alpha}^2) = 1 - \alpha (0 < \alpha \leq 1) \quad (36)$$

where  $P(\cdot)$  is the probability of the random event.

If Equation (36) is satisfied, it means that when the systems work under the optimal conditions, the nonlinear multi-sensor systems described in Equation (19) have no modeling errors, and the filtering results obtained through UKF can be identified as the local state estimations directly. Otherwise, it is determined that there are the modeling errors in the multi-sensor system (19). The time-varying adaptive fading factor is introduced into the predicted state covariance matrix, namely:

$$P_{t|t-1}^* = \lambda_t \left( \sum_{i=0}^{2n} \omega_i^2 [\tilde{\zeta}_{i,t|t-1} - \hat{x}_{t|t-1}] [\tilde{\zeta}_{i,t|t-1} - \hat{x}_{t|t-1}]^T + Q_t \right) \quad (37)$$

Submitting the adaptive fading factor into the predicted state covariance matrix, Equation (36) can be rewritten as:

$$g(\lambda_t) = \tilde{z}_t^T (P_{\tilde{z}_{t|t-1}}^*)^{-1} \tilde{z}_t - \chi_{m,\alpha}^2 = 0 \quad (38)$$

where  $P_{\tilde{z}_{t|t-1}}^*$  is the measurement prediction covariance matrix calculated by  $P_{t|t-1}^*$ .

It can be seen from Equation (38) that the solution of  $\lambda_t$  is a problem of solving the nonlinear equation. Since the derivative of  $g(\lambda_t)$  with respect to  $\lambda_t$  is hard to solve, the traditional Newton method is no longer applicable. In this paper, the chord secant method is used to determine the adaptive fading factor  $\lambda_t$  by solving the nonlinear Equation (38) iteratively. The classical UKF is then completed by replacing the predicted state covariance matrix  $P_{t|t-1}$  with the revised  $P_{t|t-1}^*$  to update the local state estimation. Then, the estimation  $\hat{x}_t^j$  of the  $j$ -th sensor at  $t$  is deduced as follows:

$$\hat{x}_t^j = \hat{x}_t \quad (39)$$

The local state estimation  $\hat{x}_t^j$  and the corresponding error covariance  $P_t^j$  ( $j = 1, 2, \dots, N$ ) are generated in parallel for each local filter by the above method, and then the global fusion state estimations are obtained by further fusion based on UT transform.

### 3.3.3. Global State Estimation

Without considering the pre-allocated weights, the fusion result  $\hat{x}_t^\Delta$  of the system's global state at  $t$  can be expressed as:

$$\hat{x}_t^\Delta = \sum_{j=1}^{N_s} \kappa_t^j \hat{x}_t^j \quad (40)$$

Define and minimize the loss function shown in Equation (41):

$$\Phi(\kappa) = \operatorname{argmin}_{\kappa} \sum_{t=1}^N (||\hat{x}_t^\Delta - \sum_{j=1}^{N_s} \kappa_t^j \hat{x}_t^j||_2^2), \text{ s.t. } \sum_{j=1}^{N_s} \kappa_t^j = 1 \quad (41)$$

Let  $\hat{X} = [\hat{x}_1^\Delta, \hat{x}_2^\Delta, \dots, \hat{x}_{N_s}^\Delta]$ , and the loss function can be described as follows:

$$\begin{aligned} \Phi(\kappa) &= \sum_{t=1}^N ||\hat{x}_t^\Delta - \sum_{j=1}^{N_s} \kappa_t^j \hat{x}_t^j||^2 \\ &= \sum_{t=1}^N ||\sum_{j=1}^{N_s} (\hat{x}_t^\Delta - \hat{x}_t^j) \kappa_t^j||^2 \\ &= \sum_{t=1}^N ||(\hat{X}_t - N_t) \kappa_t^j||^2, \hat{X}_t = [\hat{x}_t^\Delta, \dots, \hat{x}_t^\Delta], N_t = [\hat{x}_t^1, \hat{x}_t^2, \dots, \hat{x}_t^{N_s}] \\ &= \sum_{t=1}^N \kappa_t^T (\hat{X}_t - N_t)^T (\hat{X}_t - N_t) \kappa_t \end{aligned} \quad (42)$$

The  $S_t$  is defined as the local covariance matrix:

$$S_t = (\hat{X}_t - N_t)^T (\hat{X}_t - N_t) \quad (43)$$

$$\Phi(\kappa) = \sum_{t=1}^N \kappa_t^T S_t \kappa_t \quad (44)$$

The Lagrange multiplier is submitted into  $\Phi(\kappa)$ :

$$L_{ag}(\kappa_t) = \sum_{t=1}^N \kappa_t^T S_t \kappa_t + \lambda^L (\kappa_t^T I_{N_s} - 1) \quad (45)$$

where the dimension of the unit vector  $I_{N_s}$  is  $N_s \times 1$ .

The derivative of Equation (45) with respect to  $\kappa$  can be formulated as follows:

$$\frac{\partial L_{ag}(\kappa_t)}{\partial \kappa_t} = 2S_t \kappa_t + \lambda^L I_{N_s} = 0 \quad (46)$$

$$\kappa_t = \frac{S_t^{-1} I_{N_s}}{I_{N_s}^T S_t^{-1} I_{N_s}} \quad (47)$$

Thus, the parameters  $\beta_t^j$ ,  $\hat{x}_t^j$ , and  $\kappa_t^j$  in Equation (14) are solved. In the rest of the section, the train-braking parameters will be identified based on the fusion data, and the train-braking faults will be diagnosed.

### 3.3.4. Fault Diagnosis Based on Fusion Data

In the previous section, the fusion filtering results of speed measurements are obtained. Based on the fusion results, the parameters of the model are identified, and the faults of the braking systems are diagnosed. Considering the parameters' characteristics of time-varying hidden variables to be identified, an online improved expectation maximization (EM) [31–33] parameter identification method is proposed based on sliding window [34–36].

#### (1) Construction of conditional expectation

The train model (1) was re-expressed as follows:

$$\begin{cases} v_{t+1} = F(v_t, \theta) + \omega_t \\ y_t = H(v_t) + e_t \end{cases} \quad (48)$$

$$\theta = \begin{bmatrix} \mu \\ \mu_A \end{bmatrix} \quad (49)$$

where  $y_t$  represents the sensor fusion result, that is,  $y_t = \hat{x}_t^*$ , and  $\theta$  is the braking parameter to be identified. Based on the Markov probability characteristics of the model, it can be inferred as follows:

$$v_{t+1} \sim p_\theta(v_{t+1}|v_t), y_t \sim p_\theta(y_t|v_t) \quad (50)$$

where  $p_\theta(v_{t+1}|v_t)$  represents the probability density of  $v_{t+1}$  when the train state  $v_t$  is given;  $p_\theta(y_t|v_t)$  is the probability density of  $y_t$  given  $v_t$ .

The measurements at  $t_1$  and  $t_2$  are set as the head and end of the window, respectively, and the sliding length is  $L$ . The data between  $t_1$  and  $t_2$  are taken as the research objects, and the braking parameter  $\theta$  is initialized as the estimated result  $\hat{\theta}_{m-1}$  of the  $(m-1)$ -th interval, that is:

$$\theta_m = \hat{\theta}_{m-1} \quad (51)$$

The current window output vector sequence and its likelihood function are defined as:

$$Y_{t_1:t_2} = \{y_{t_1}, \dots, y_{t_2}\} \quad (52)$$

$$p_\theta(Y_{t_1:t_2}) = p_\theta(y_{t_1}, \dots, y_{t_2}) = p_\theta(y_{t_1}) \prod_{t=(t_1+1)}^{t_2} p_\theta(y_t|Y_{t_1:(t-1)}) \quad (53)$$

where  $p_\theta(y_t|Y_{t_1:(t-1)})$  represents the probability density of  $y_t$  when the measured sequence  $Y_{t_1:(t-1)} = \{y_{t_1}, \dots, y_{t-1}\}$  is known.

Take the logarithm of both sides of Equation (53), the logarithmic of the output vectors sequence can be expressed as:

$$L_{\theta}(Y_{t_1:t_2}) = \ln p_{\theta}(Y_{t_1:t_2}) = \ln p_{\theta}(y_{t_1}) + \sum_{t=(t_1+1)}^{t_2} \ln p_{\theta}(y_t | Y_{t_1:(t-1)}) \quad (54)$$

$$p_{\theta}(y_t | Y_{t_1:(t-1)}) = \int p_{\theta}(y_t | v_t) p_{\theta}(v_t | Y_{t_1:(t-1)}) dv_t \quad (55)$$

where  $p_{\theta}(y_t | v_t)$  represents the probability density of  $y_t$  when the  $v_t$  is given. Due to the existence of unmeasurable parameters to be identified in the model, the train state sequence  $V_{t_1:t_2} = \{v_{t_1}, \dots, v_{t_2}\}$  is regarded as an incomplete measurement. The joint probability density function of all data is derived from the  $V_{t_1:t_2}$  and train output observations  $Y_{t_1:t_2}$ .

$$\begin{aligned} L_{\theta}(V_{t_1:t_2}, Y_{t_1:t_2}) &= \ln p_{\theta}(V_{t_1:t_2}, Y_{t_1:t_2}) = \ln p_{\theta}(V_{t_1:t_2} | Y_{t_1:t_2}) + \ln p_{\theta}(Y_{t_1:t_2}) \\ &= \ln p_{\theta}(v_{t_1}) + \sum_{t=t_1}^{t_2-1} \ln p_{\theta}(v_{t+1} | v_t) + \sum_{t=t_1}^{t_2} \ln p_{\theta}(y_t | v_t) \end{aligned} \quad (56)$$

It is assumed that the current estimation of train brake coefficient is  $\theta_k$ , and the expectation  $E(\theta, \theta_k)$  of  $L_{\theta}(V_{t_1:t_2}, Y_{t_1:t_2})$  is described as follows:

$$E(\theta, \theta_k) = \int \ln p_{\theta}(V_{t_1:t_2}, Y_{t_1:t_2}) \times p_{\theta_k}(V_{t_1:t_2} | Y_{t_1:t_2}) dV_{t_1:t_2} = I_1 + I_2 + I_3 \quad (57)$$

$$\begin{aligned} I_1 &= \int \ln p_{\theta}(v_{t_1}) p_{\theta_k}(v_{t_1} | Y_{t_1:t_2}) dv_{t_1} \\ I_2 &= \sum_{t=t_1}^{t_2-1} \iint \ln p_{\theta}(v_{t+1} | v_t) p_{\theta_k}(v_{t+1}, v_t | Y_{t_1:t_2}) dv_t dv_{t+1} \\ I_3 &= \sum_{t=t_1}^{t_2} \int \ln p_{\theta}(y_t | v_t) p_{\theta_k}(v_t | Y_{t_1:t_2}) dv_t \end{aligned} \quad (58)$$

where  $I_1, I_2$ , and  $I_3$  are related to conditionally smooth densities  $p_{\theta_k}(v_{t_1} | Y_{t_1:t_2})$ ,  $p_{\theta_k}(v_{t+1}, v_t | Y_{t_1:t_2})$ , and  $p_{\theta_k}(v_t | Y_{t_1:t_2})$ , respectively. Considering that there are unmeasurable variables in the train state, the analytical solution by integral method is invalid, so the integration is solved by the particle filter and particle smoothing.

## (2) Maximization of conditional expectation

From Equations (57) and (58), it can be seen that the relationship is nonlinear and non-convex between the estimation of conditional expectation  $\hat{E}(\theta, \theta_k)$  and train braking parameters, and its closed solution is hard to solve. Therefore, the gradient optimization is adopted to find  $\theta$ , which makes  $\hat{E}(\theta, \theta_k)$  the maximum [37]. The partial derivative of  $\hat{E}(\theta, \theta_k)$  with respect to  $\theta$  is as follows:

$$\frac{\partial}{\partial \theta} \hat{E}(\theta, \theta_k) = \frac{\partial \hat{I}_1}{\partial \theta} + \frac{\partial \hat{I}_2}{\partial \theta} + \frac{\partial \hat{I}_3}{\partial \theta} \quad (59)$$

The gradient method is adopted for parameter iterative optimization. When the estimation accuracy of braking parameter meets  $\|\theta_{k+1} - \theta_k\| \leq \varepsilon$ , the iteration terminates, and  $\theta_{k+1}$  is regarded as the final estimation result  $\hat{\theta}_m$  of the  $m$ -th window. Otherwise, the iteration continues.

$$\hat{\theta}_m = \theta_{k+1} \quad (60)$$

$\hat{\theta}_m$  is regarded as the parameter identification result at the midpoint of  $t_1$  and  $t_2$ .

$$\hat{\theta}_{(t_1+t_2)/2} = \hat{\theta}_m \quad (61)$$

#### 4. Analysis of Simulation Results

In order to verify the validity of the diagnostic method presented in this paper, a CRH3 high-speed train with four motions and four tows is selected for the simulation experiment.

##### 4.1. Train Parameters Descriptions

In the train braking simulation, the main parameters of the braking device are shown in Table 1.

The main parameters of the speed sensors are shown in Table 2.

The wheel-rail adhesion coefficient is selected by the actual experience, which can be divided into dry rail and wet rail adhesion coefficient by the different rail states, as follows:

Dry rail:

$$\mu = 0.06 + \frac{46.6}{260 + v} \quad (62)$$

Wet rail:

$$\mu = 0.04 + \frac{13.7}{120 + v} \quad (63)$$

As can be seen from the above equation, the adhesion coefficient of different rail states is different, and it is inversely proportional to the train speed. When the train is running on a dry rail, the system fault that results in poor braking performance can be easily diagnosed. However, when the train is in emergency braking on wet rail, the adhesion coefficient is difficult to be accurately measured, which makes it difficult to distinguish whether the braking performance degradations are caused by braking system failures, train adhesion failures, or both. Therefore, the following four typical cases of the train braking will be taken as examples for simulations.

**Table 1.** Braking parameters of CRH3.

Braking Parameters	Value
Total Weight of Train/(t)	536
Maximum Operating Speed/(km/h)	350
Continuous Operating Speed/(km/h)	300
Brake Cylinder Diameter/(mm)	203
Air Pressure of Brake Cylinder/(kpa)	410
Transmission Efficiency	0.85
Braking Ratio	2.55
Friction Coefficient	0.28
Brake Disc Friction Radius	297.6
Wheel Rolling Radius/(mm)	460

**Table 2.** Speed sensors parameters.

Wheel Hall Speed Sensor	Value	Doppler Radar Speed Sensor	Value
Pitch	7.85	Range Resolution/(mm/pulse)	4
Gear module	2.5	Number of Pulses/(km)	250,000
Operation Temperature/(°C)	20~80	Operation Temperature/(°C)	−20~70

##### 4.2. Four Typical Cases

###### 4.2.1. Adhesion Normal and Braking Normal

In case 1, the braking characteristics of the train are studied with normal adhesion. The braking speed measurements are shown in Figure 6. The red curve is the true braking speed, and the other four curves represent the four sensors' speed measurements. As can be seen from the Figure 6, in the emergency braking, the train speed decreases from 300 km/h to 50 km/h, the braking time is 73.3 s, and the average deceleration speed is 0.9474 m/s<sup>2</sup>. The train speed fusion results are shown in Figure 7. The red curve is the real speed of the train, the blue curve is the fusion result of the four speed sensors, the green curve is the fusion result of the three speed sensors in the case of abnormal or lost data

caused by sensor failures or network transmission failures, the magenta curve is similarly the fusion result of the two speed sensors, and the black curve is the maximum speed fusion result equipped in the train's on-board ATP. It can be seen from the Figure 7 that the multiple speed measurements can be fused effectively by the method proposed in this paper. Affected by the missing or abnormal sensors measurements, in the case of meeting the actual accuracy requirements of the system, the fusion accuracy of four speed sensors is slightly higher than the other two cases. However, when comparing with the maximum fusion strategy equipped in the train's on-board ATP, the fusion accuracy of the proposed method is obviously higher.

The real-time identification results of the friction coefficient of the brake disc are shown in Figure 8. The red curve is the real value of the friction coefficient, and the other curves correspond to the friction coefficient identification values based on the speed fusion results in Figure 7. It can be seen from Figure 8, based on the fusion speed, the unobservable time-varying friction coefficient can be identified accurately, then the real-time braking performance of the train brake disc can be grasped timely. The identification result of the black curve differs greatly from the real value universally, because the speed obtained based on the maximum fusion strategy is higher than the real speed. The identification relative errors of friction coefficient are shown in Figure 9.

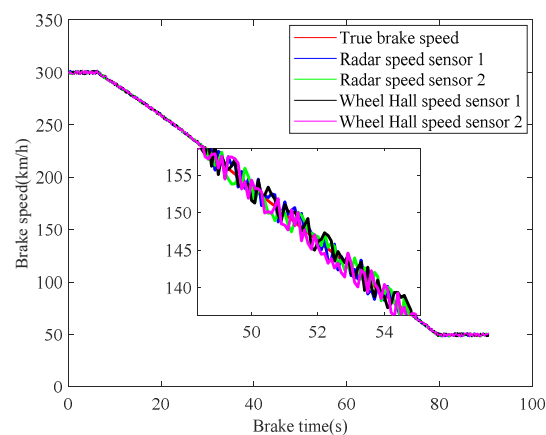


Figure 6. Train braking speed measurements.

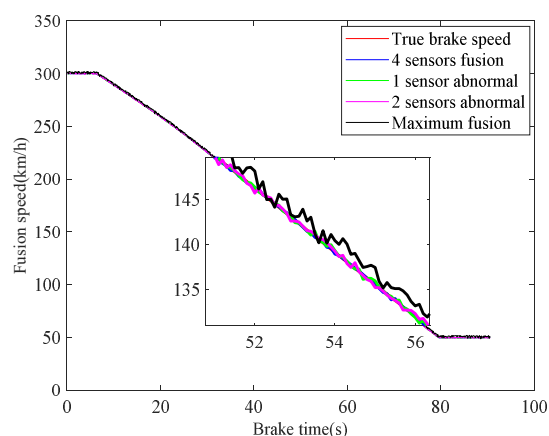
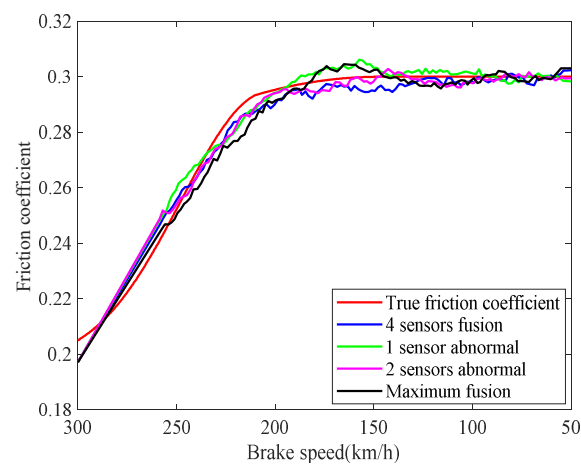
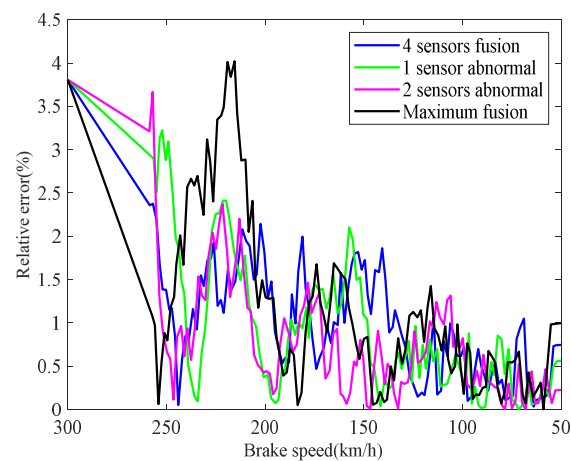


Figure 7. Fusion of speed sensors.





**Figure 8.** Identifications of friction coefficient.



**Figure 9.** Identification relative errors of friction coefficient.

The speed fusion errors and parameter identification errors are shown in Table 3. It can be seen from the table that the speed fusion error and friction coefficient identification error of the method proposed in this paper are  $\pm 1.7472\%$  and  $\pm 2.4891\%$ , respectively, and the speed fusion error and friction coefficient identification error of the maximum fusion data with large error, that the real braking performance of the brake disc can still be obtained effectively and accurately by the fusion identification framework proposed in this paper.

**Table 3.** Speed fusion errors and parameter identification errors.

Fusion Method	Fusion Method			
	4 Sensors Fusion	3 Sensors Fusion	2 Sensors Fusion	Maximum Fusion
Relative Error				
Brake Speed Fusion	$\pm 0.5313\%$	$\pm 1.3412\%$	$\pm 1.7472\%$	$\pm 2.5627\%$
Identification of Friction Coefficient	$\pm 2.2684\%$	$\pm 2.4832\%$	$\pm 2.4891\%$	$\pm 2.4963\%$

#### 4.2.2. Adhesion Normal but Brake Degradation

In case 2, the braking characteristics of the train under normal rail adhesion are studied. Different from case 1, the braking performance of the train declined, and the change of braking characteristics of the brake disc is paid more attention in case 2. The braking speed measurements are shown in Figure 10. The red curve is the true braking speed, and

the other four curves represent the sensors speed measurements. It can be seen from the Figure 10 that the train speed is reduced from 300 km/h to 50 km/h, the braking time is 97.1 s, and the average deceleration is  $0.7152 \text{ m/s}^2$ . Compared with the normal braking in case 1, the braking time increases by 23.8 s and 32.47%, and the average deceleration speed decreases by  $0.2322 \text{ m/s}^2$  and 24.51% in case 2. Therefore, it is considered that the braking performance of the train decreases significantly. Similar to Figure 7, the fusion results of train speed are shown in Figure 11, which are based on the proposed method and the maximum fusion. It can be seen from Figure 11 that the multi-sensor train speed measurements can be effectively fused by the proposed method. Under meeting the actual accuracy requirements of the system, the fusion accuracy of four sensors speed measurements is slightly higher than that of the other two cases; however, compared with the maximum fusion results, the fusion accuracy of the proposed method is higher.

The real-time identification results of the friction coefficient of the brake disc are shown in Figure 12. The red curve is the real value of the friction coefficient, and the other curves correspond to the friction coefficient identifications based on the speed fusion results in Figure 11. Compared with the normal braking parameter identification curves in Figure 8, it can be seen that the friction coefficient of the brake disc in case 2 significantly decreased. In the train braking, the speed decreases from 300 km/h to 50 km/h, and the friction coefficient of the brake disc gradually stabilizes at 0.22, which decreases by 0.08 and 26.67%, compared with 0.3 in Figure 8. Combined with the current braking capacity of the brake disc and the train braking standard, the continued use of the brake disc will pose a threat to the train operation safety. Therefore, when the brake disc is diagnosed with a brake fault, it should be replaced in a timely manner. The identification relative errors of friction coefficient are shown in Figure 13.

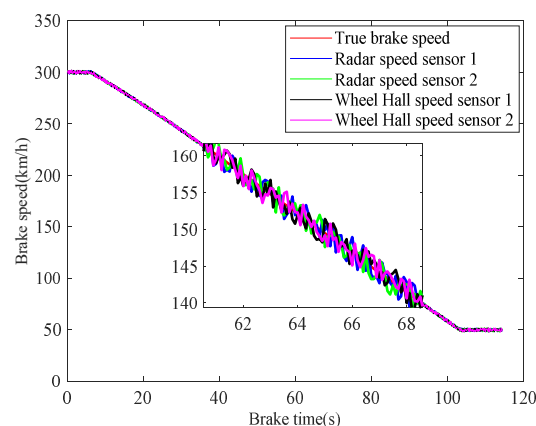


Figure 10. Train braking speed measurements.

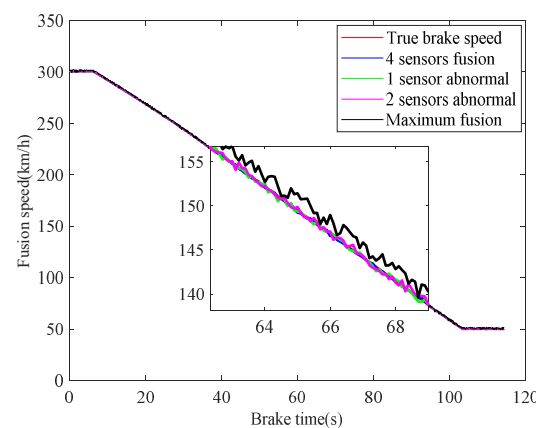


Figure 11. Fusion of speed sensors.

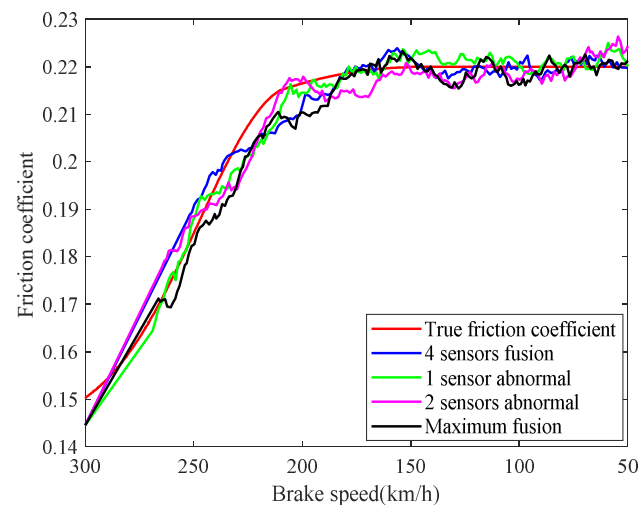


Figure 12. Identifications of friction coefficient.

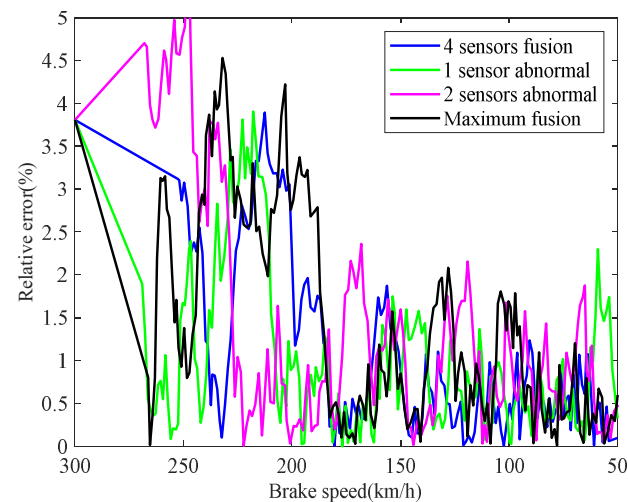


Figure 13. Identification relative errors of friction coefficient.

The speed fusion errors and parameter identification errors are shown in Table 4. It can be seen from the table that the speed fusion error and friction coefficient identification error of the proposed method are  $\pm 1.5152\%$  and  $\pm 2.4937\%$ , respectively. The speed fusion error and friction coefficient identification error of maximum fusion are  $\pm 2.3187\%$  and  $\pm 2.4987\%$ , respectively. Even for the maximum fusion data with large errors, the real braking performance of the brake disc can be still identified effectively and accurately, and the robustness is verified of the proposed fusion identification framework.

Table 4. Speed fusion errors and parameter identification errors.

Fusion Method	4 Sensors	3 Sensors	2 Sensors	Maximum
	Fusion	Fusion	Fusion	Fusion
Relative Error				
Brake Speed Fusion	$\pm 0.6049\%$	$\pm 1.2527\%$	$\pm 1.5152\%$	$\pm 2.3187\%$
Identification of Friction Coefficient	$\pm 2.2258\%$	$\pm 2.4863\%$	$\pm 2.4937\%$	$\pm 2.4987\%$

#### 4.2.3. Adhesion Failure but Braking Normal

In case 3, the braking characteristics of trains with rail adhesion failure are studied. The braking speed measurements are shown in Figure 14. The red curve represents the

actual braking speed, the blue and green curves represent the radar speed measurements, and the black and magenta curves represent the wheel-rail Hall sensors speed measurements. As can be seen from Figure 14, during the whole braking process, the radar speed measurement curves are close to the real value at all times, while the wheel Hall speed measurement curves plummet in 19.0–61.6 s (261.60 km/h–124.74 km/h), and the two Hall speed measurement results are similar. Combined with the Hall speed measuring principle, it can be inferred that the braking process from 300 km/h to 50 km/h is a composite braking, in which 261.60 km/h–124.74 km/h is adhesive braking, and 300 km/h–261.60 km/h and 124.74 km/h–50 km/h are air emergency braking. The braking time of the whole braking process is 77.7 s, and the average deceleration is  $0.8938 \text{ m/s}^2$ . The fusion results of speed are shown in Figure 15, which are fused based on the proposed fusion method and the maximum fusion. It can be seen from Figure 15 that the speed measurements of multiple trains can be effectively fused by the proposed paper. Compared with the maximum fusion results, the fusion accuracy of the proposed fusion method is still higher.

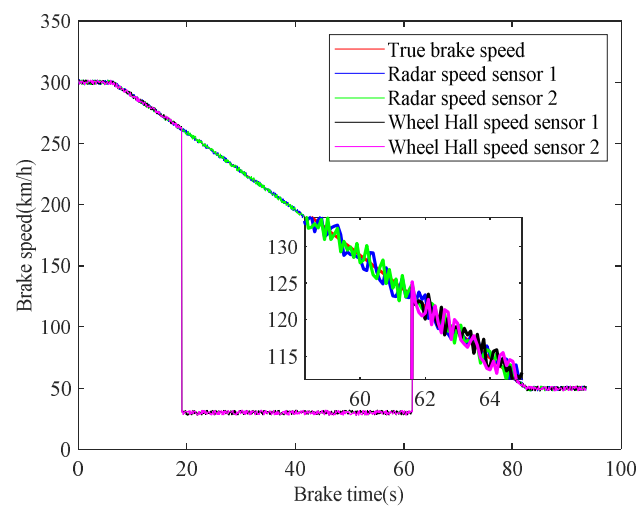


Figure 14. Train braking speed measurements.

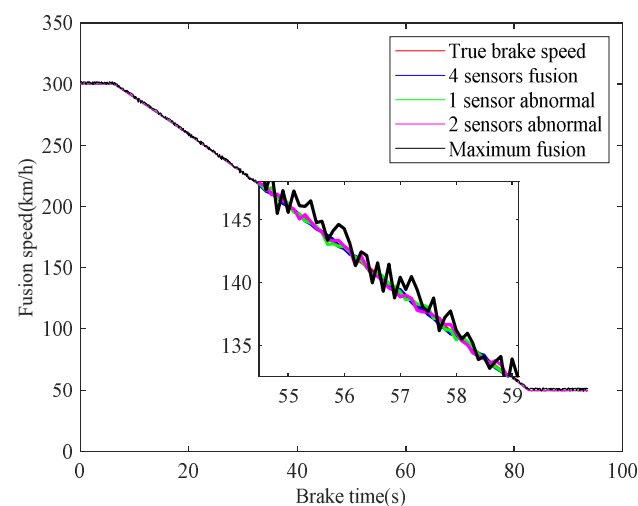
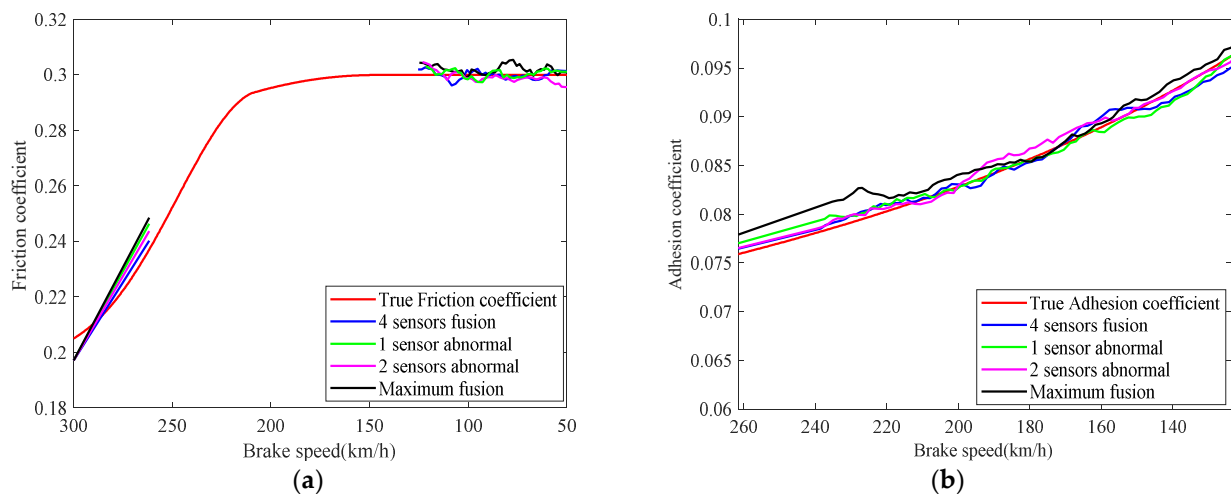


Figure 15. Fusion of speed sensors.

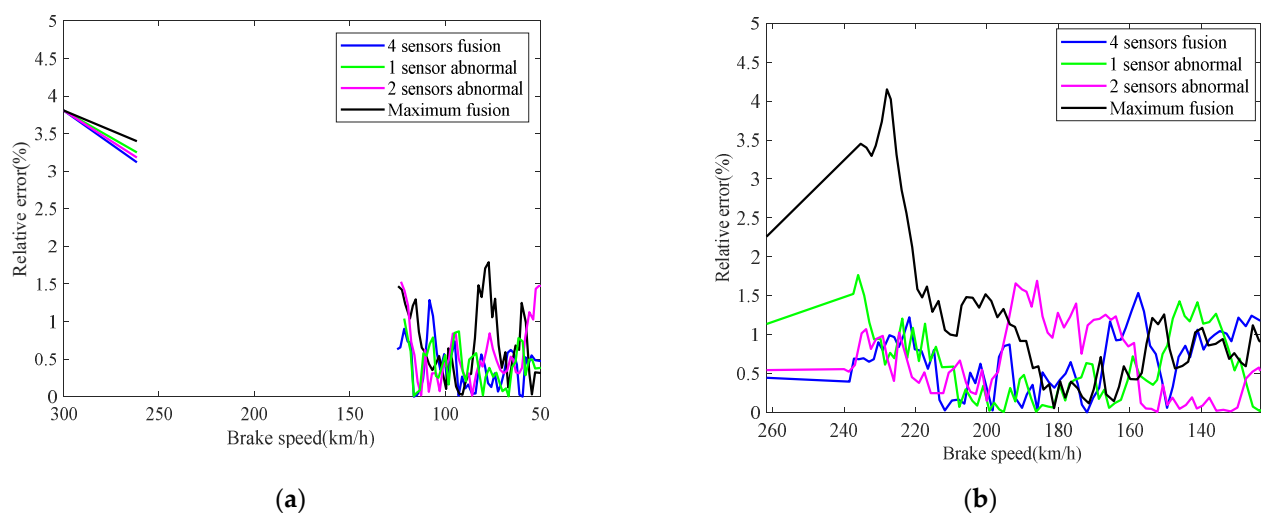
Considering the model switching in the train composite braking, the friction coefficient of the brake disc is identified in the 300 km/h–261.60 km/h and 124.74 km/h–50 km/h range, and the wheel-rail adhesion coefficient is identified in the 261.60 km/h–124.74 km/h range. The real-time identification results of the friction coefficient and the adhesion coefficient are shown in Figure 16. The red lines are the real values of the friction coefficient and

adhesion coefficient, respectively. The other curves correspond to the identifications of the friction coefficient and adhesion coefficient in Figure 15, respectively. As can be seen from Figure 16, the hidden parameters in the train jump model can be accurately estimated by the fusion identification framework proposed in this paper. Then, it can be concluded that the intermittent braking failure is not caused by the decrease of friction performance of brake disc fault, but the adhesion failure of adhesion coefficient between wheel and rail on the wet rail. The identification of the relative errors of the friction coefficient and the adhesion coefficient are shown in Figure 17.

The speed fusion errors and parameter identification errors are shown in Table 5. It can be seen from the table that the speed fusion error, friction coefficient, and adhesion coefficient identification error of the proposed method are  $\pm 1.6742\%$ ,  $\pm 2.2473\%$ , and  $\pm 1.7267\%$ , respectively, and the speed fusion error, friction coefficient, and adhesion coefficient identification error of the maximum fusion are  $\pm 2.4659\%$ ,  $\pm 2.3682\%$ , and  $\pm 1.7431\%$ , respectively. Even for the maximum fusion data with large errors, the braking performance of the brake disc and the real-time wheel-rail adhesion state can be more effectively and accurately obtained.



**Figure 16.** Identifications of friction coefficient and adhesion coefficient: (a) friction coefficient; (b) adhesion coefficient.



**Figure 17.** Identification relative errors of friction coefficient and adhesion coefficient: (a) friction coefficient; (b) adhesion coefficient.

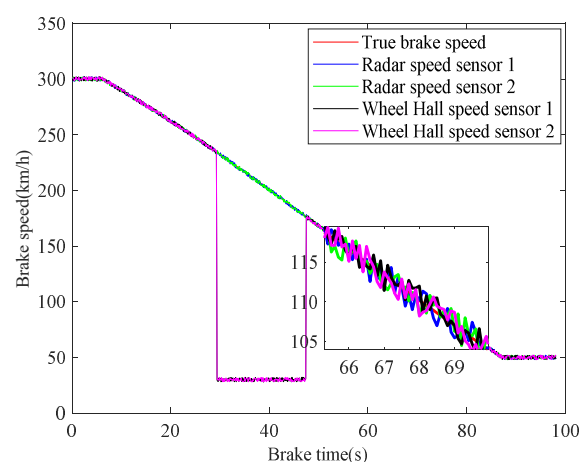
**Table 5.** Speed fusion errors and parameter identification errors.

Fusion Method	4 Sensors	3 Sensors	2 Sensors	Maximum Fusion
	Fusion	Fusion	Fusion	
Brake Speed Fusion	$\pm 0.5657\%$	$\pm 1.5629\%$	$\pm 1.6742\%$	$\pm 2.4659\%$
Identification of Friction Coefficient	$\pm 1.8477\%$	$\pm 1.9337\%$	$\pm 2.2473\%$	$\pm 2.3682\%$
Identification of Adhesion Coefficient	$\pm 1.5391\%$	$\pm 1.6538\%$	$\pm 1.7267\%$	$\pm 1.7431\%$

#### 4.2.4. Adhesion Failure and Brake Degradation

In case 4, the braking characteristics with adhesion failures are also studied. The braking speed measurements are shown in Figure 18. The red curve represents the true braking speed of the train, the blue and green curves represent the radar speed measurement curves, and the black and magenta curves represent the wheel-rail Hall speed measurement curves. Similar to Figure 14, there is also a sharp drop in the two Hall speed measurements in Figure 18, which indicates that the braking process of the train is composite braking, 234.47 km/h–185.38 km/h is adhesive braking, and 300 km/h–234.47 km/h and 185.38 km/h–50 km/h are air emergency braking. In the whole braking process, the braking duration is 80.9 s and the average deceleration is  $0.8544 \text{ m/s}^2$ . Compared with case 3, in case 4, the braking time increases by 3.2 s and 3.96%, and the average deceleration decreases by  $0.0354 \text{ s}$  and 3.96%. Therefore, it is considered that the train braking performance similarly decreases. The fusion results of train speed are shown in Figure 19, which are fused based on the proposed fusion method and the maximum fusion. It can be seen from Figure 19 that the measurements of multiple sensors can be fused effectively. Compared with the maximum fusion, the fusion accuracy of the proposed fusion method is still higher.

The real-time identification results of the friction coefficient and the wheel-rail adhesion coefficient are shown in Figure 20. The red lines are the true values of the friction coefficient and the adhesion coefficient, respectively. The other curves correspond to the identifications of the friction coefficient and the adhesion coefficient based on the speed fusion results in Figure 19. As can be seen from Figure 20, the wheel-rail adhesion characteristics remain unchanged, but comparing with the normal friction coefficient in case 1, the friction performance of the brake discs in case 4 decreases significantly. The identification of relative errors of the friction coefficient and adhesion coefficient are shown in Figure 21.

**Figure 18.** Train braking speed measurements.



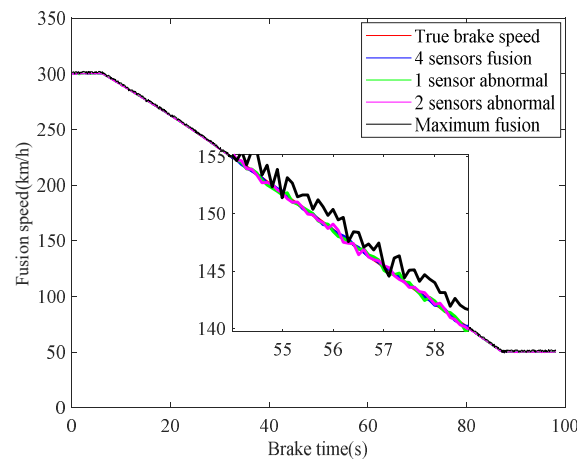


Figure 19. Fusion of speed sensors.

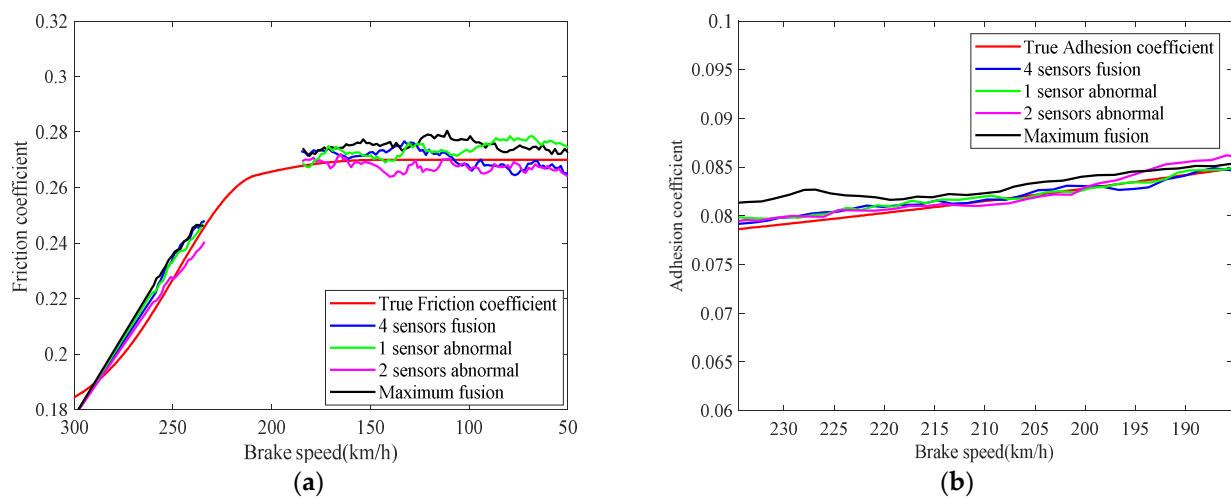


Figure 20. Identifications of friction coefficient and adhesion coefficient: (a) friction coefficient; (b) adhesion coefficient.

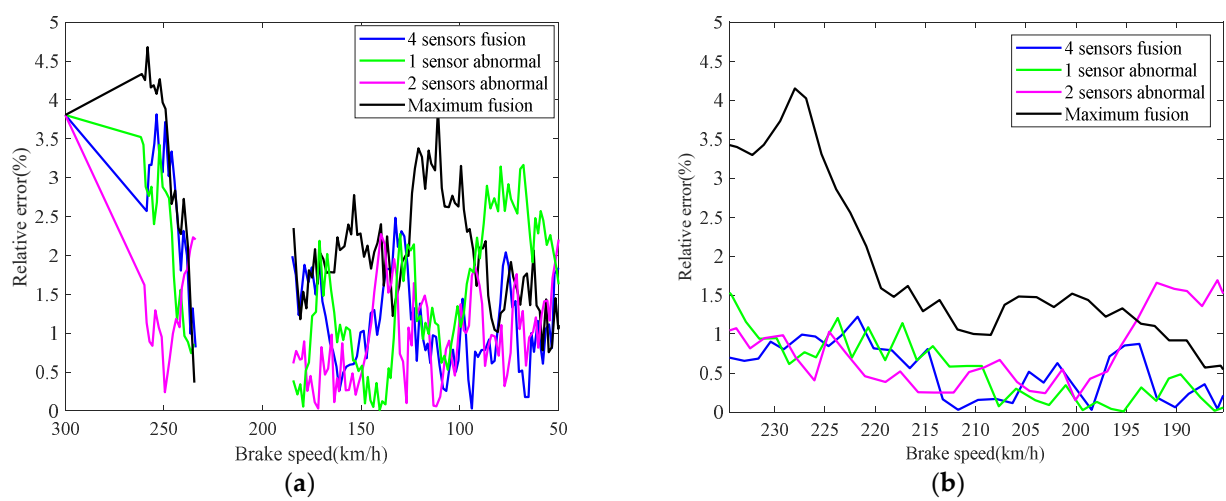


Figure 21. Identification relative errors of friction coefficient and adhesion coefficient: (a) friction coefficient; (b) adhesion coefficient.

The speed fusion errors and parameter identification errors are shown in Table 6. It can be seen from the table that the speed fusion error, identification errors of the friction coefficient, and the adhesion coefficient of the proposed method are  $\pm 1.4364\%$ ,  $\pm 2.4967\%$ ,

and  $\pm 1.6733\%$ , respectively. The speed fusion error, identification errors of the friction coefficient, and the adhesion coefficient of the maximum fusion are  $\pm 2.3751\%$ ,  $\pm 2.4989\%$ , and  $\pm 1.7875\%$ , respectively. Even for the maximum fusion data with large errors, the braking performance of the brake disc and the real-time wheel-rail adhesion state can be more effectively and accurately obtained.

**Table 6.** Speed fusion errors and parameter identification errors.

Fusion Method	4 Sensors Fusion	3 Sensors Fusion	2 Sensors Fusion	Maximum Fusion
Brake Speed Fusion	$\pm 0.6116\%$	$\pm 1.2220\%$	$\pm 1.4364\%$	$\pm 2.3751\%$
Identification of Friction Coefficient	$\pm 2.1127\%$	$\pm 2.3431\%$	$\pm 2.4967\%$	$\pm 2.4989\%$
Identification of Adhesion Coefficient	$\pm 1.5087\%$	$\pm 1.6325\%$	$\pm 1.6733\%$	$\pm 1.7875\%$

## 5. Conclusions and Prospect

Based on the train braking mechanism and the actual operating environments, the train emergency braking model based on the operation environment is established. Aiming at the composite braking faults of the train, the fusion weights pre-allocation and linear weighted fusion criteria are constructed, and the braking modes and braking parameters are identified based on the improved online maximum expectation, and the faults of the train braking system are diagnosed. The simulation results show that the braking parameters of systems can be effectively identified, the identification relative errors of the speed, friction coefficient, and adhesion coefficient are no more than 1.8%, 2.5%, and 1.75%, respectively, which meet the actual requirements of the brake system. Furthermore, the single fault and composite fault of the train braking system can be accurately diagnosed, even if the monitoring data are disturbed or missing. The effectiveness and robustness of the proposed method are verified. It can be concluded that the application of this method will greatly improve the accuracy of train operation monitoring and fault diagnosis, reduce the daily maintenance cost, and ensure train safety.

In future research, the algorithm will be improved and optimized to adjust the size of the sliding window adaptively, and the speed and accuracy of parameter estimation will be further improved. In the future, more types of sensor data will be fused, and a faster fusion algorithm will also be developed. Different trains and types of faults will also be considered.

**Author Contributions:** Conceptualization, G.X. and Y.J.; methodology, Y.J.; software, Y.J.; validation, Y.J., Y.L. and X.Z.; formal analysis, Y.L. and X.Z.; investigation, A.S. and N.H.; resources, Y.L. and N.H.; data curation, Y.L. and W.C.; writing—original draft preparation, Y.J.; writing—review and editing, G.X. and Y.J.; visualization, Y.J. and W.C.; supervision, G.X.; project administration, G.X.; funding acquisition, G.X. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was supported by the National Key R&D Program of China (No.2018YFB1201500), the National Science Foundation of China (No. 61873201, No.U2034209, and No.U1934222), the Natural Science Foundation of Shaanxi Province of China (2021JC-42), the Natural Science Foundation of Shaanxi Provincial Department of Education (19JS051), and the Key Laboratory of Complex System Intelligent Control and Decision, Beijing Institute of Technology.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The data presented in this study are available on request from the corresponding author. The data are not publicly available due to the confidentiality.

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

## References

1. Muñoz-Paniagua, J.; García, J. Aerodynamic drag optimization of a high-speed train. *J. Wind. Eng. Ind. Aerod.* **2020**, *204*, 104215. [[CrossRef](#)]
2. Peixer, M.A.; Montenegro, P.A.; Carvalho, H.; Ribeiro, D.; Bittencourt, T.N.; Calçada, R. Running safety evaluation of a train moving over a high-speed railway viaduct under different track conditions. *Eng. Fail. Anal.* **2021**, *121*, 105133. [[CrossRef](#)]
3. Montenegro, P.A.; Carvalho, H.; Ribeiro, D.; Tokunaga, M.; Tanabe, M.; Zhaif, W.M. Assessment of train running safety on bridges: A literature review. *Eng. Struct.* **2021**, *241*, 112425. [[CrossRef](#)]
4. Ji, H.Q.; Zhou, D.H. Incipient fault detection of the high-speed train air brake system with a combined index. *Control Eng. Pract.* **2020**, *100*, 104425. [[CrossRef](#)]
5. Zhong, M.Y.; Ding, S.X.; Zhou, D.H.; He, X. An  $H_i/H_\infty$  optimization approach to event-triggered fault detection for linear discrete time systems. *IEEE Trans. Automat. Contr.* **2020**, *65*, 4464–4471. [[CrossRef](#)]
6. Safaeipour, H.; Forouzanfar, M.; Casavola, A. A survey and classification of incipient fault diagnosis approaches. *J. Process Control.* **2021**, *97*, 1–16. [[CrossRef](#)]
7. Arunthavanathan, R.; Khan, F.; Ahmed, S.; Imtiaz, S. An analysis of process fault diagnosis methods from safety perspectives. *Comput. Chem. Eng.* **2021**, *145*, 107197. [[CrossRef](#)]
8. Chen, H.T.; Jiang, B. A review of fault detection and diagnosis for the traction system in high-speed trains. *IEEE Trans. Intell. Transp. Syst.* **2020**, *21*, 450–465. [[CrossRef](#)]
9. Chen, H.T.; Jiang, B.; Lu, N.Y. A newly robust fault detection and diagnosis method for high-speed trains. *IEEE Trans. Intell. Transp. Syst.* **2019**, *20*, 2198–2208. [[CrossRef](#)]
10. Chen, H.T.; Jiang, B.; Chen, W.; Yi, H. Data-driven detection and diagnosis of incipient faults in electrical drives of high-speed trains. *IEEE Trans. Ind. Electron.* **2019**, *66*, 4716–4725. [[CrossRef](#)]
11. Liu, Z.; Zhang, M.; Liu, F.; Zhang, B. Multidimensional feature fusion and ensemble learning-based fault diagnosis for the braking system of heavy-haul train. *IEEE Trans. Ind. Inform.* **2021**, *17*, 41–51. [[CrossRef](#)]
12. Glowacz, A.; Tadeusiewicz, R.; Legutko, S.; Caesarendra, W.; Irfan, M.; Liu, H.; Brumercik, F.; Gutten, M.; Sulowicz, M.; Xiang, J.; et al. Fault diagnosis of angle grinders and electric impact drills using acoustic signals. *Appl. Acoust.* **2021**, *179*, 108070. [[CrossRef](#)]
13. Glowacz, A. Fault diagnosis of electric impact drills using thermal imaging. *Measurement* **2021**, *171*, 108815. [[CrossRef](#)]
14. Niu, G.; Xiong, L.J.; Qin, X.X.; Michael, P. Fault detection isolation and diagnosis of multi-axle speed sensors for high-speed trains. *Mech. Syst. Signal Proc.* **2019**, *131*, 183–198. [[CrossRef](#)]
15. Yang, J.; Xie, G.; Yang, Y.X.; Zhang, Y.M.; Liu, W. Deep model integrated with data correlation analysis for multiple intermittent faults diagnosis. *ISA Trans.* **2019**, *95*, 306–319. [[CrossRef](#)]
16. Zhao, Y.H.; He, X.; Zhang, J.F.; Ji, H.Q.; Zhou, D.H.; Pecht, M.G. Detection of intermittent faults based on an optimally weighted moving average  $T^2$  control chart with stationary observations. *Automatica* **2021**, *123*, 109298. [[CrossRef](#)]
17. Dibaj, A.; Etefagh, M.M.; Hassannejad, R.; Ehghaghi, M.B. A hybrid fine-tuned VMD and CNN scheme for untrained compound fault diagnosis of rotating machinery with unequal-severity faults. *Expert Syst. Appl.* **2021**, *167*, 114094. [[CrossRef](#)]
18. Moshrefzadeh, A. Condition monitoring and intelligent diagnosis of rolling element bearings under constant/variable load and speed conditions. *Mech. Syst. Signal Process.* **2021**, *149*, 107153. [[CrossRef](#)]
19. Sohaib, M.; Kim, J. Fault Diagnosis of Rotary Machine Bearings under Inconsistent Working Conditions. *IEEE Trans. Instrum. Meas.* **2020**, *69*, 3334–3347. [[CrossRef](#)]
20. Meng, L.H.; Wang, Y.H.; Zhou, J.; Ding, X.J.; Li, X.B. Multi information fusion and fault diagnosis system for motor drive system in high speed train. In Proceedings of the 18th Prognostics and System Health Management Conference, Chongqing, China, 26–28 October 2018.
21. Luo, H.L.; Bo, L.; Peng, C.; Hou, D.M. Fault diagnosis for high-speed train axle-box bearing using simplified shallow information fusion convolutional neural network. *Sensors* **2020**, *20*, 4930. [[CrossRef](#)] [[PubMed](#)]
22. Hounkpevi, F.O.; Yaz, E.E. Robust minimum variance linear state estimators for multiple sensors with different failure rates. *Automatica* **2007**, *43*, 1274–1280. [[CrossRef](#)]
23. Caballero-Águila, R.; Hermoso-Carazo, A.; Linares-Pérez, J. Linear and quadratic estimation using uncertain observations from multiple sensors with correlated uncertainty. *Signal Process.* **2011**, *91*, 330–3376. [[CrossRef](#)]
24. Caballero-Águila, R.; García-Garrido, I.; Linares-Pérez, J. Information fusion algorithms for state estimation in multi-sensor systems with correlated missing measurements. *Appl. Math. Comput.* **2014**, *226*, 548–563. [[CrossRef](#)]
25. Sun, S.L.; Deng, Z.L. Multi-sensor optimal information fusion Kalman filter. *Automatica* **2004**, *40*, 1017–1023. [[CrossRef](#)]
26. Gao, B.B.; Hu, G.G.; Gao, S.S.; Zhong, Y.M.; Gu, C.F. Multi-sensor optimal data fusion based on the adaptive fading unscented Kalman filter. *Sensors* **2018**, *18*, 488. [[CrossRef](#)] [[PubMed](#)]
27. Wang, Y.; Cong, J.L.; Wang, P.; Liu, X.; Tang, H.Y. A data-fusion approach for speed estimation and location calibration of a metro train based on low-cost sensors in smartphones. *IEEE Sens. J.* **2019**, *19*, 10744–10752. [[CrossRef](#)]
28. Kim, K.H.; Kong, S.H.; Jeon, S.Y. Slip and slide detection and adaptive information sharing algorithms for high-speed train navigation systems. *IEEE Trans. Intell. Transp. Syst.* **2015**, *16*, 3193–3203. [[CrossRef](#)]
29. Hu, G.G.; Gao, S.S.; Zhong, Y.M. A derivative UKF for tightly coupled INS/GPS integrated navigation. *ISA Trans.* **2015**, *56*, 135–144. [[CrossRef](#)]

30. Xie, G.; Jin, Y.Z.; Hei, X.H.; Ji, W.J.; Gao, S.G.; Takahashi, S.; Mochizuki, H. Adaptive identification of time-varying environmental parameters in train dynamics model. *Acta Autom. Sin.* **2019**, *45*, 2268–2280.
31. Noh, S.J.; Shim, D.Y.; Jeon, M.G. Adaptive sliding-window strategy for vehicle detection in highway environments. *IEEE Trans. Intell. Transp. Syst.* **2016**, *17*, 323–335. [[CrossRef](#)]
32. Shakil, S.; Billings, J.C.; Gong, X.J.; Keilholz, S.D. Parametric dependencies of sliding window correlation. *IEEE Trans. Biomed. Eng.* **2018**, *65*, 254–263. [[CrossRef](#)]
33. Gasmi, N.; Boutayeb, M.; Thabet, A.; Aoun, M. Sliding Window Based Nonlinear  $H_\infty$  Filtering: Design and Experimental Results. *IEEE Trans. Circuits Syst. II Express Briefs* **2019**, *66*, 302–306. [[CrossRef](#)]
34. Schön, T.B.; Wills, A.; Ninness, B. System identification of nonlinear state-space models. *Automatica* **2011**, *47*, 39–49. [[CrossRef](#)]
35. Sammaknejad, N.; Zhao, Y.J.; Huang, B. A review of the Expectation Maximization algorithm in data-driven process identification. *J. Process Control* **2019**, *73*, 123–136. [[CrossRef](#)]
36. Efendi, E.; Dulek, B. Online EM-Based Ensemble Classification with Correlated Agents. *IEEE Signal Process. Lett.* **2021**, *28*, 294–298. [[CrossRef](#)]
37. Zhang, Y.Q.; Lou, Y.C.; Hong, Y.G. An approximate gradient algorithm for constrained distributed convex optimization. *IEEE CAA J. Autom. Sin.* **2014**, *1*, 61–67.