State-Degradation-Oriented Fault Diagnosis for High-Speed Train Running Gears System

Chao Cheng [1,2,3]**, Weijun Wang** [1] **, Hao Luo** [4,]*****, Bangcheng Zhang** [5]**, Guoli Cheng** [2] **and Wanxiu Teng** [2]

[1] School of Computer Science and Engineering, Changchun University of Technology, Changchun 130012, China; chengchao@mail.tsinghua.edu.cn (C.C.); jinwei312@163.com (W.W.)

[2] CRRC Changchun Railway Vehicles Co., Ltd., Changchun 130062, China; chengguoli@cccar.com.cn (G.C.); tengwanxiu@cccar.com.cn (W.T.)

[3] Department of Automation, Tsinghua University, Beijing 100084, China

[4] Academy of Astronautics, Harbin Institute of Technology, Harbin 150001, China

[5] School of Mechatronic Engineering, Changchun University of Technology, Changchun 130012, China; zhangbangcheng@ccut.edu.cn

***** Correspondence: hao.luo@hit.edu.cn; Tel.:+86-0451-8640-2350

Abstract: As one of the critical components of high-speed trains, the running gears system directly affects the operation performance of the train. This paper proposes a state-degradation-oriented method for fault diagnosis of an actual running gears system based on the Wiener state degradation process and multi-sensor filtering. First of all, for the given measurements of the high-speed train, this paper considers the information acquisition and transfer characteristics of composite sensors, which establish a distributed topology for axle box bearing. Secondly, a distributed filtering is built based on the bilinear system model, and the gain parameters of the filter are designed to minimize the mean square error. For a better presentation of the degradation characteristics in actual operation, this paper constructs an improved nonlinear model. Finally, threshold is determined based on the Chebyshev's inequality for a reliable fault diagnosis. Open datasets of rotating machinery bearings and the real measurements are utilized in the case studies to demonstrate the effectiveness of the proposed method. Results obtained in this paper are consistent with the actual situation, which validate the proposed methods.

Keywords: fault diagnosis; distributed state estimation; state degradation

1. Introduction

Nowadays, railway construction and operation are high-speed, which considerably affects the sustainability and the rapid progress of the national economy. The safety and reliability of railway operation have drawn more and more attention. As a core system that directly affects the smooth running of high-speed trains, the running gears system is a key component of the train in providing dynamics and traction performance and has the functions of buffering, vibration isolation, generating power and supporting the vehicle body. As a matter of fact, the running gears system is a complex system that is coupled with many components. Among those components, the corrosion, shedding and degradation of one component may easily spread the local fault and propagate into a major fault in the system level, which may cause unexpected losses. It is very important to analyze system performance degradation [1,2] and prevent fault propagation. Therefore, the timely diagnosis of a running gears system plays a key role in ensuring the safe operation of trains.

In the past few decades, many work has been done on fault diagnosis technologies in different aspects. Traditionally, the fault diagnosis methods for high-speed train systems mainly consist of

two factors: The model-based one and the signal-based one [3]. The signal-based fault diagnosis methods analyze the measurement data for fault diagnosis without knowing/identifying the system model. The common signal-based methods include multivariate statistical analysis [4,5], machine learning methods [6,7], signal processing methods [8], information fusion methods [9,10], etc. The above methods have high requirements on data volume and type. The measurements of the running gears system of high-speed trains are of high sampling frequency and long measurement cycle. Once the running gears system of the high-speed train is damaged, it will be repaired in a timely manner and regularly maintained. However, the adequate failure data for analysis are very rare. The most important step of the signal-based fault diagnosis method is the accurate extraction of fault characteristics [11]. If the measurements with inconspicuous fault characteristics are used for the learning, accurate feature extraction normally cannot be guaranteed, and the result of fault diagnosis will be directly affected. In contrast, model-based fault diagnosis methods, which are more sensitive to uncertainty/fault, have lower requirements for data quantity and data type. They are more suitable for a reliable fault diagnosis of high-speed train running gears systems.

Among the model-based fault diagnosis methods, a class of improved methods based on the Kalman filter is used to deal with the influence of system/measurement noise. Yan et al. [12] introduced the extended Kalman filter model and the improved support vector machine model into the online classification framework, where the nonlinearity is approximated by the first-order Taylor expansion and the online fault detection of the ventilation and air-conditioning system is later achieved. Ning et al. [13] proposed an enhanced robust Kalman filtering method based on the optimal radial basis function neural network, which reduces the impact of state errors and observation errors on state estimation. Wang et al. [14] reduced the influence of estimated noise of on-board fault diagnosis and fault-tolerant control through the Kalman filter and improved the robustness of fault diagnosis. Finally, the method is verified by an onboard prototype control system. Akai et al. [15] used the particle filter method to realize the state estimation of a target robot in order to realize system filtering under the influence of noise. Wu et al. [16] proposed a fault diagnosis algorithm for early failures based on the generalized estimation. Yin et al. [17] applied a state resonance algorithm-based stochastic resonance method to high-speed train traction systems to achieve optimal extraction and frequency recovery. Cadini et al. [18] used a particle filter to estimate the length of cracks to diagnose the early faults of aeronautical structure cracks. For state estimation with unknown inputs, some researchers use unknown input residual generators and estimation filters to achieve fault diagnosis [19]. Scholars also consider the decomposition of external disturbances and combine unknown input observers to complete composite fault diagnosis [20]. For nonlinear systems, He et al. [21] applied rank particle filtering to diagnosis the fault of underwater vehicles. In recent years, to improve the accuracy of state estimation and reduce the false alarm rate, scholars have carefully studied the measurement noise and unknown disturbance. Zhou et al. [22] applied an improved odorless Kalman filter and a radial basis function based on an adaptive noise factor to the pumping unit for fault diagnosis. Liu et al. [23] used an adaptive controller to process unknown system parameters and disturbances to achieve target state tracking. Gan et al. [24] used an adaptive factor to improve the accuracy of the unscented Kalman filter algorithm, thereby improving the accuracy of fault diagnosis of analog circuits. However, most of the existing methods are based on state estimation and improvement of a single node of the monitoring system. Such methods mostly rely on the reliability of information collection nodes. For an actual high-speed train information collection system, the measurements obtained in different locations of the same monitoring area contain different information, which will indirectly affect the consistency of information collection and reduce the accuracy of fault diagnosis. With the development of a multi-sensor system, distributed estimation has been studied and applied in different scenarios. Jing et al. [25] realized multi-sensor data fusion of the planetary gearbox by fusing vibration signal, sound signal, drive motor current, shaft speed, oil chip and other measurements, and then detected various system faults. Liu et al. [26] proposed a multi-sensor intelligent data fusion method based on the ant colony algorithm to realize fault detection of the gear box, aiming at the

unreliability and instability of fault detection results of single sensor vibration signals. To avoid the disadvantage of combining a single information source with a decision method, a fault diagnosis method combining different types of sensor data sources and multiple classifier decisions is proposed in [27]. Banerjee et al. [28] proposed and investigated a hybrid method for fault signal classification based on sensor data fusion by using the support vector machine and short-term Fourier transform techniques. This method estimates the state of the target node by combining the information of surrounding nodes. This method has been successfully applied in individual formation [29], stability monitoring [30] and output coordination [31,32]. It can be concluded that distributed monitoring has better robustness and stability for single sensor monitoring in the sensor information fusion architecture. Considering multiple sensor information of the target monitoring system can improve the accuracy of state estimation more than considering single sensor information.

While the model-based method has been applied in fault diagnosis for high-speed train, there are still some problems. First, for the area monitored by multiple sensors, the estimation does not fully consider the influence of neighbor nodes which will further affect the later state estimation and fault diagnosis. Secondly, the statistical characteristics and the noise in the measurements collected in high-speed trains are normally unknown. The impacts of initial state and noise will be continuously amplified during the recursive estimation process, which may reduce the accuracy of fault diagnosis. Finally, the degradation of the system performance cannot be avoided. If the degradation is ignored, it will increase the false alarm rate of fault diagnosis and reduce the diagnostic accuracy. Aiming at solving the above problems, this paper proposes a distributed state estimation filter, which is combined with system state degradation characteristics to achieve fault diagnosis. This method is different from fault prognosis, the purpose of which is to analyze the possible faults in the future through the fault state feature [33], and then to provide guidance for future system health management. The purpose of this paper is to diagnose the existing faults of the degraded system. From the perspective of methods, the signal-based fault prognosis method [34] also has the constraint problem of data quantity and data type. The model-based fault prognosis method [35] mainly predicts the future fault generation time by analyzing the evolution trend of the fault. The method proposed in this paper is to diagnose the existing faults through the state evolution trend. They are similar, but still different.

Innovations and contributions of this article are as follows.

(1) Comprehensively consider the measuring point position and information acquisition method of a composite sensor. A distributed topology structure is established by taking the axle box bearing of an actual running gears system as an example. Based on this structure, a bilinear distributed filter is proposed, and the gain parameters of the filter are designed to minimize the mean square error.
(2) Unbiased constraint conditions are used to reduce the impact of the initial unknown information of nodes on state estimation. By constructing the difference to deal with the problem of colored noise in real measurements, estimation accuracy is improved.
(3) A nonlinear degradation model of the Wiener process considering temperature change characteristics is built to describe the state degradation phenomenon during train operation. The solution of nonlinear degradation process parameters is given by maximum likelihood estimation and combined with distributed filters to increase the accuracy of fault diagnosis.

Finally, to demonstrate the effectiveness of the proposed method, this paper conducts simulation experiments on the open dataset of rotating machinery bearings. The fault diagnosis results given are the same as actual experimental results. The proposed method was applied to operational monitoring data of a certain type of high-speed train running gears system, and fault diagnosis results obtained were consistent with a real situation.

This paper is arranged as follows. The second part describes the problem and introduces the preparation work. The third part proposes a method based on the Wiener process and multi-sensor filtering. The fourth part uses two case studies to test and analyze the above method. The final part concludes the article.

2. Preliminaries and Problem Formulations

The running gears system is a complex mechanism, and its state information is an important indicator of traffic safety, as shown in Figure 1a,b. At present, the train fault diagnosis system is mainly composed of the train main engine, vehicle extension, preprocessing and composite sensors. The monitoring area of high-speed train operation consists of three or four sensors. Each sensor is integrated with two or more sensor units to detect different physical quantities and collect different types of data, such as temperature and vibration. Therefore, a single sensor is also called a composite sensor component. However, information acquisition in fault diagnosis based on single sensor information is not comprehensive or accurate. For the bearings of high-speed trains running a gears system, equipment status information from a single sensor is often significantly incomplete, which cannot provide an accurate and detailed data basis for subsequent fault diagnosis. Therefore, the fault diagnosis method based on single sensor information widely used in trains is limited in the accuracy of diagnosis and the completeness of information. To solve the problem, this paper abstracts a high-speed train running gears system into a multi-sensor system. The train control and management system processes the node information and gives full play to the advantages of a multi-node joint operation so that each node in a multi-sensor system can generate a consistent interpretation and description of the monitoring target. The reliability of sensor information is enhanced to achieve the purpose of collaborative monitoring and improve the accuracy of state estimation.
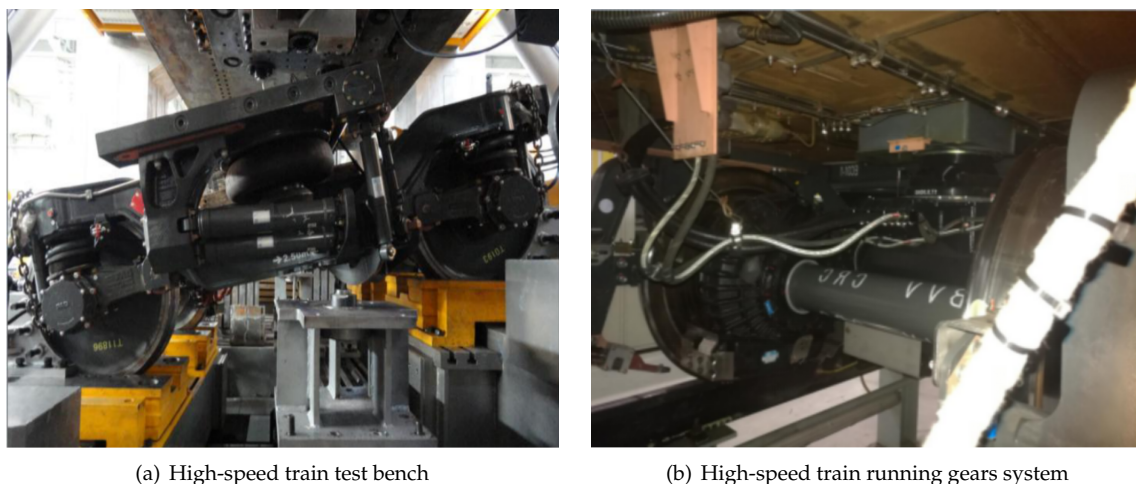

(a) High-speed train test bench

(b) High-speed train running gears system

Figure 1. (**a**) High-speed train test bench; (**b**) high-speed train running gears system.

Eleven composite sensors are installed in the support area of the monitoring bearing, and the installation direction needs to be consistent with the direction of the impact signal. The specific position is shown as follows, A1–A4: Measuring points of the axle box bearing; B1–B3: Measuring points for traction motor bearings; C1–C4: Measuring points for gearbox bearings. Taking the axle box bearing and its composite sensor in the above system as an example, a networked sensor system as shown in Figure 2 is established. The topology directed graph is represented by $G = (V, E, D)$, where the vertex set is represented as $V = \{1, 2, 3, 4\}$, the set of edges is represented as $\varepsilon \subseteq \nu \times \nu$.
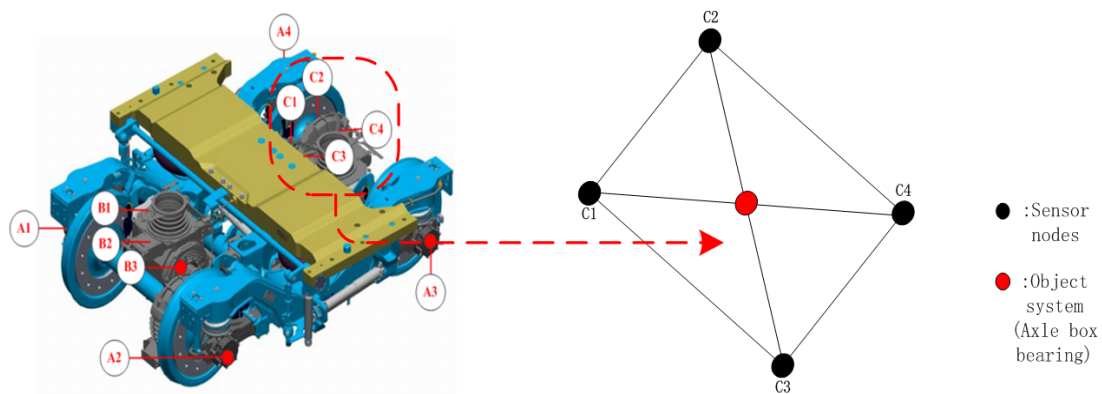
Figure 2. Distribution of measuring points of running gears system and topological structure of axle box bearing nodes.

This paper considers the effects of a sensor's gain attenuation, colored measurement noise and accuracy reduction noise. The model of the target system is as follows.

$$
\begin{aligned}
x(k+1) &= A(k)x(k) + B(k)u(k) + w(k) \\
y_i(k) &= \lambda_i C_i(k)x(k) + D_i(k)u(k) + v_i(k) + s_i(k)
\end{aligned}
\tag{1}
$$

To better conform to the changes in the actual state, the running characteristics of axle box bearings of the running section need to be analyzed. The axle box bearings of the running section are made of metal, and the specific heat capacity of metal is a function of temperature. In general, its specific heat capacity increases with increasing temperature. Therefore, the change of temperature of the axle box bearing will no longer conform to the law of linear change, and a bilinear term will be generated in a system equation. Consider describing the above system as the following bilinear form.

$$
\begin{aligned}
x(k+1) &= A(k)x(k) + B(k)u(k) + N(k)u(k)x(k) + w(k) \\
y_i(k) &= \lambda_i C_i(k)x(k) + D_i(k)u(k) + v_i(k) + s_i(k)
\end{aligned}
\tag{2}
$$

where k is the discrete time index, $u(k)$ is the control input signal, $w(k)$ is the process noise and the mean is 0, the covariance is Q. $v(k)$ represents measurement noise of the i-th sensor. $s(k)$ is precision-degraded noise of the i-th sensor, the mean value is 0 and the covariance is O. λ is gain reduction of the i-th sensor. $A(k)$, $B(k)$, $C(k)$ and $D(k)$ are known coefficient matrices, and $N(k)u(k)x(k)$ is a bilinear term.

In addition, as a key mechanical rotating part of the running gears system, the performance of the bearing plays an essential role in the safe running of the train. Generally speaking, mechanical equipment must go through a series of regular degradation stages from normal operation to complete failure. Therefore, taking the degradation process into account in the system model can more accurately describe the changing process of the actual system and increase the accuracy of state estimation. This paper considers the modeling of system state degradation based on the Wiener process. The more common nonlinear degradation model is as follows.

$$
X(t) = X(0) + \int_0^t \mu(u;\theta)du + \sigma B(t)
\tag{3}
$$

where $\mu(u;\theta)$ is a non-linear function used to characterize the non-linear characteristics of the system degradation. θ is vector of unknown parameters contained in the function. $\sigma B(t)$ is the diffusion term used to describe the uncertainty of the degradation process. When the surface temperature of the friction pair rises to a certain extent for a short time, a secondary quenching layer and a high-temperature tempering layer will be produced on the surface of the bearing part. This burnt layer will cause obvious changes in the structure and performance of the surface of the bearing part,

affecting the performance of the friction pair. This paper considers improving the above-mentioned Wiener process degradation model to achieve state degradation modeling.

3. Main Results and Discussion

The fault diagnosis method proposed in this paper is mainly divided into the following three stages: First, the distributed filtering design of multi-sensor systems. Second, model of the system state degradation based on the Wiener process. Third, fault diagnosis methods for the running gears system, shown in Figure 3.
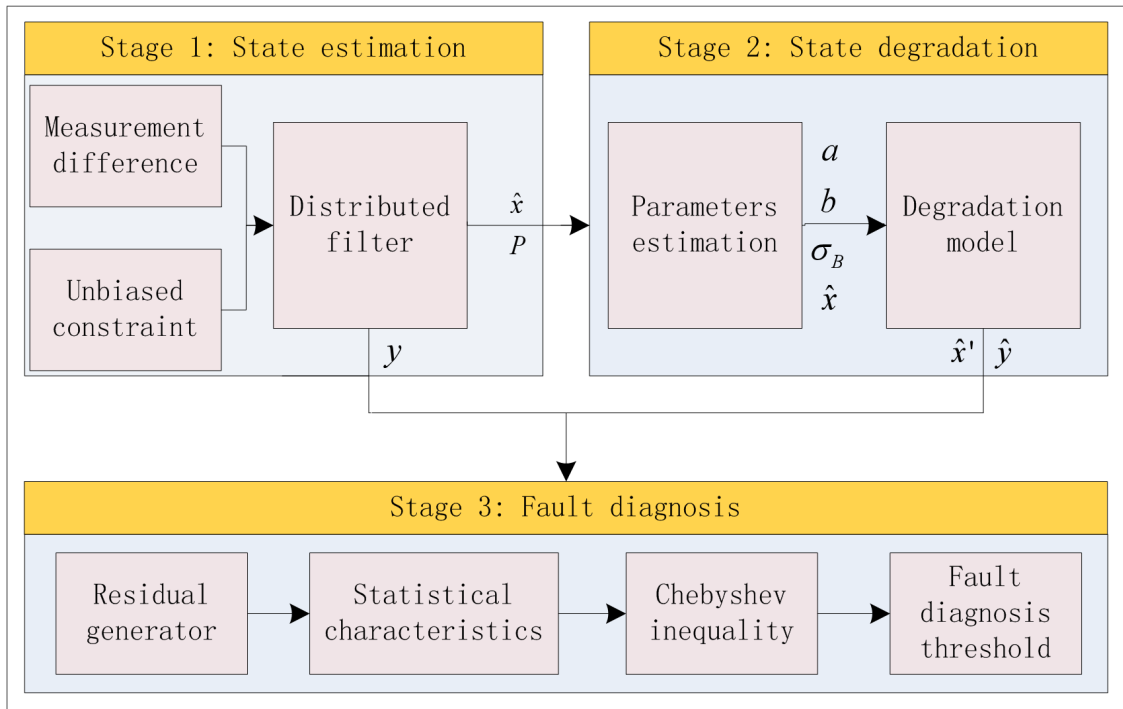

Figure 3. Fault diagnosis method framework of high-speed train running gears system.

3.1. Multi-Sensor Filter

Measurement noise during actual sensor operation is colored noise, the measurement noise in the differential form is constructed as follows.

$$v_i(k) = \psi_i(k-1)v_i(k-1) + \zeta_i(k-1) \tag{4}$$

where ζ represents Gaussian white noise. To process v, an auxiliary signal z similar to the above is constructed.

$$\begin{aligned} z_i(k) &= y_i(k) - \psi_i(k-1)v_i(k-1) \\ &= \lambda_i C_i(k)x(k) + D_i(k)u(k) + v_i(k) + s_i(k) - \psi_i(k-1)v_i(k-1) \\ &= \lambda_i C_i(k)x(k) + D_i(k)u(k) + \zeta_i(k-1) + s_i(k) \end{aligned} \tag{5}$$

The auxiliary signal z thus obtained will no longer contain colored noise terms, which is more convenient for processing. Continuing to use the equation of state of the original system, the measurement equation is described by auxiliary signal z, and the following equivalent equation of target system can be obtained:

$$\begin{aligned} x(k+1) &= A(k)x(k) + B(k)u(k) + N(k)u(k)x(k) + w(k) \\ z_i(k) &= \lambda_i C_i x(k) + D_i(k)u(k) + \zeta_i(k-1) + s_i(k) \end{aligned} \tag{6}$$

Considering the mutual influence between multi-sensor nodes, a distributed state estimation structure for bilinear systems is proposed as follows.

$$\hat{x}_i(k+1) = A(k)\hat{x}_i(k) + L_i(k)u(k) + N(k)u(k)\hat{x}_i(k) + \sum_{j \in Ni} H(k)[z_j(k) - \lambda_i C_i(k)x(k)] \tag{7}$$

where $\hat{x}_i(k+1)$ represent the value of the estimated state of the *i*-th sensor, *Ni* is the set consisting of node *i* and its neighbors and $H(k)$ is the filter gain of the multi-sensor.

In the process of state estimation, the uncertainty of the initial state will be transferred in the process of estimation iteration, leading to the deviation of state estimation. To satisfy the unbiasedness of the filter, the state estimation process is processed as follows, so that the mean of each state estimate is the same as the mean of the states, which is $E[\hat{x}(k)] = E[x(k)]$. The above formula is used as an unbiased constraint, so for the state estimation equation and the state equation at time $k+1$:

$$E[\hat{x}_i(k+1)] = A(k)E[\hat{x}_i(k)] + L_i(k)u(k) + N(k)u(k)E[\hat{x}_i(k)] + \sum_{j \in Ni} H(k)D_i(k)u(k)$$

$$E[x_i(k+1)] = A(k)E[x_i(k)] + B_i(k)u(k) + N(k)u(k)E[x_i(k)] \tag{8}$$

$$E[\hat{x}_i(k+1)] - E[x_i(k+1)] = [L_i(k) - B_i(k) + \sum_{j \in Ni} H(k)D_i(k)]u(k) = 0$$

Therefore, the unbiased constraint can be equivalent to:

$$L_i(k) - B_i(k) + \sum_{j \in Ni} H(k)D_i(k) = 0 \tag{9}$$

Next, the gain parameter H of the filter is obtained by minimizing the trace of estimated error covariance. The estimated error function is established as follows.

$$e_i(k+1) = [\hat{x}_i(k+1) - x_i(k+1)] \tag{10}$$

The gain parameter H of the filter is expressed as:

$$\begin{aligned}
H_B &= \arg\min_H E\{e^T(k+1) \cdot e(k+1)\} \\
&= \arg\min_H E\{trace[e(k+1) \cdot e^T(k+1)]\} \\
&= \arg\min_H E\{trace[(\sum_{j \in Ni} H(k)\zeta_j(k-1) + H(k)S_j(k) - w_j(k)) \cdot (\cdots)^T]\} \\
&= \arg\min_H trace\{\sum_{j \in Ni} [H(k)R_j(k-1)H^T(k) + H(k)O_j(k)H^T(k) - Q_j(k)]\}
\end{aligned} \tag{11}$$

Using the unbiased constraint as the condition for minimizing the above formula, the Lagrange function is constructed.

$$\begin{aligned}
Js(H,\lambda) =& [\sum_{j \in N_i} H(k)R_j(k-1)H^T(k) + H(k)O_j(k)H^T(k) - Q_i(k)] \\
& + \lambda[L_i(k) - B_i(k) + \sum_{j \in Ni} H(k)D_j(k)]
\end{aligned} \tag{12}$$

Then, H can be expressed as:

$$\begin{cases}
H(k) = \sum_{j \in N_i} -\dfrac{1}{2}\lambda[R_j(k-1) + O_j(k)]^{-1} \cdot D_j(k) \\
\sum_{j \in N_i} H(k)D_j(k) = B_i(k) - L_i(k)
\end{cases} \tag{13}$$

In summary, the analytical solution of the gain parameter H of the distributed filter is finally obtained as:

$$H(k) = [B_i(k) - L_i(k)] \sum_{j \in N_i} \{[R_j(k-1) + O_j(k)]^{-1}D_j(k)\}^{-1}D_j^{-1}(k)[R_j(k-1) + O_j(k)]^{-1}D_j(k) \quad (14)$$

Remark 1. *This paper only considers the measurement noise of a system as colored noise, instead of colored process noise. The reason is that the state degradation process is the main reason that directly affects system change. The characteristics of process noise have little effect on the system. If process noise is processed in the same form, it will increase the computational complexity.*

3.2. Parameter Estimation of State Degradation

Considering the effects of random shocks and environmental factors (temperature), this paper builds a nonlinear degradation process based on the Wiener degradation model and provides a solution for parameter estimation.

Construct a nonlinear state degradation model.

$$X(t) = X(0) + \int_0^t \mu(u;\theta)du + \sum_{i=1}^N J_i + u(t)x_{state}(t) + \sigma B(t) \quad (15)$$

where $X(t)$ is the degradation process driven by the standard drift Brownian motion $B(t)$, $\mu(u;\theta)$ is a non-linear function, J is the amplitude of random shocks, N represent the random shock, and $u(t)x(t)$ is the environmental factor on the system state degraded parameters. It can be seen that when $\mu(u;\theta) = \mu$, the degradation process becomes a linear degradation model. In order to describe the above models in detail, this paper gives a parameter estimation method for a class of nonlinear models.

$$\mu(t;\theta) = abt^{b-1} \; , \; X(t) = X(0) + at^b + \sum_{i=1}^N J_i + u(t)x_{state}(t) + \sigma B(t) \quad (16)$$

The unknown parameters of the drift and diffusion terms in the degradation process are estimated, assuming current time is t_k, and the historical detection data of equipment degradation is $x_1, x_2, \cdots, x_k$, where x is obtained by sampling the degradation process at equal intervals. $\tilde{X}_{1:k}$ represents the incremental degradation process.

Theorem 1. *The average value of increment $\tilde{X}_{1:k}$ of the above degradation process is:*

$$\begin{aligned} &aE[t_1^b - t_0^b \quad t_2^b - t_1^b \quad \cdots \quad t_{i+1}^b - t_i^b]^T \\ &- E[u_1 x_{state1} - u_0 x_{state0} \quad u_2 x_{state2} - u_1 x_{state1} \quad \cdots \quad u_{i+1} x_{state\,i+1} - u_{i+1} x_{state\,i+1}]^T \end{aligned} \quad (17)$$

The variance is $\sigma_B^2 \cdot Q$ and follows a normal distribution.

Proof of Theorem 1. Tectonic degradation increment:

$$\begin{aligned} X(i+1) &= x(i+1) - x(i) \\ &= a[t(i+1)]^b + u[t(i+1)]x_{state}[t(i+1)] + \sigma_B B[t(i+1)] \\ &\quad - at(i)^b - u[t(i)]x_{state}[t(i)] - \sigma_B B[t(i)] \\ &= a[t(i+1)^b - t(i)^b] + \Delta + \sigma_B \{B[t(i+1)] - B[t(i)]\} \end{aligned} \quad (18)$$

where Δ can represent the influence of the environment (temperature) on state degradation, and the mean value of the above equation can be obtained:

$$E[X(i+1)] = aE[t(i+1)^b - t(i)^b] - E[\Delta] \quad (19)$$

The covariance can be expressed as:

$$
\begin{aligned}
&\mathrm{var}[X(i+1) \cdot X(j+1)] \\
=&\mathrm{var}\{x[t(i+1)] - x[t(i)]\} \cdot \{x[t(j+1)] - x[t(j)]\} \\
=&E\{x[t(i+1)] - x[t(i)]\}\{x[t(j+1)] - x[t(j)]\} - E\{x[t(i+1)] - x[t(i)]\} \cdot E\{x[t(j+1)] - x[t(j)]\} + \sigma_B(\cdot) \\
=&E\{a^2\{[t(i+1)]^b - t(i)^b][t(j+1)^b - t(j)^b]\} + E\{a[t(i+1)^b - t(i)^b] \cdot \Delta(j)\} \\
&+E\{\Delta(i) \cdot a[t(j+1)^b - t(j)^b]\} + E[\Delta(i)\Delta(j)] - E\{a^2[t(i+1)^b - t(i)^b][t(j+1)^b - t(j)^b]\} \\
&-aE[t(i+1)^b - t(i)^b]\Delta(i) - \Delta(j) \cdot aE[t(j+1)^b - t(j)^b] + \sigma_B(\cdot) \\
=&E\{\sigma_B\{B[t(i+1)] - B[t(i)]\}\{a[t(j+1)^b - t(j)^b] + \Delta(j)\}\} \\
&+E\{\sigma_B\{B[t(j+1)] - B[t(j)]\}\{a[t(i+1)^b - t(i)^b] + \Delta(i)\}\} \\
&+E\{\sigma_B^2\{B[t(i+1)] - B[t(i)]\} \cdot \{B[t(j+1)] - B[t(j)]\}\} \\
=&\sigma_B^2 E\{\{B[t(i+1)] - B[t(i)]\} \cdot \{B[t(j+1)] - B[t(j)]\}\}
\end{aligned}
\tag{20}
$$

where $\sigma_B(\cdot)$ represents the term associated with σ_B.

Set $Q = E\{\{B[t(i+1)] - B[t(i)]\}\{B[t(j+1)] - B[t(j)]\}\}$, the variance of the degenerate increment is $\sigma_B^2 \cdot Q$.

Construct the likelihood function:

$$
l(\theta; \sigma_B | \tilde{X}_{0:k}) = \frac{1}{\sqrt{2\pi} \cdot \sigma_B \sqrt{Q}} \cdot e^{-\frac{(\tilde{X}_{1:k} - a\mu_{1:k} + \Delta_{1:k})^2}{2\sigma_B^2 Q}}
\tag{21}
$$

according to the above likelihood function, the desired a and σ_B can be obtained:

$$
\frac{d\ln(l)}{da} = \frac{\tilde{X}_{1:k} \cdot \mu_{1:k} + \Delta_{1:k} \cdot \mu_{1:k}}{2\sigma_B^2 Q} - \frac{\mu_{1:k}^2 \cdot a}{\sigma_B^2 Q} = 0 \; ; \; \hat{a} = \frac{\tilde{X}_{1:k} + \Delta_{1:k}}{2\mu_{1:k}}
\tag{22}
$$

$$
\frac{d\ln(l)}{d\sigma_B} = -\frac{1}{\sigma_B} + \frac{(\tilde{X}_{1:k} - a\mu_{1:k} + \Delta_{1:k})^2}{\sigma_B^3 Q} = 0 \; ; \; \hat{\sigma}_B = \frac{\tilde{X}_{1:k} - a\mu_{1:k} + \Delta_{1:k}}{\sqrt{Q}}
\tag{23}
$$

The estimated values of a and σ_B are substituted into the log-likelihood function, and the log-likelihood function is maximized by the simplex method to get b. □

Remark 2. *The state degradation model needs to be combined with the state estimation model. The continuous function of the degradation process needs to be discretized.*

3.3. Residual Generator Design

Discretizing the state variables in stage 2 and combining them with the state variables in stage 1:

$$
\begin{aligned}
x'(k) &= x(k) + X(k) \\
\hat{x}'(k) &= \hat{x}(k) + X(k)
\end{aligned}
\tag{24}
$$

Constructing a residual generator:

$$
T_i(k) = y_i(k) - \hat{y}_i(k)
\tag{25}
$$

where $\hat{y}_i(k) = \lambda_i C_i(k)\hat{x}(k) + D_i(k)u(k)$, and then:

$$
\begin{aligned}
T_i(k) &= \lambda_i C_i(k)[x(k) - \hat{x}(k)] + v_i(k) + s_i(k) \\
&= \lambda_i C_i(k) \cdot E^+(k-1)E(k-1) \cdot [x(k) - \hat{x}(k)] + v_i(k) + s_i(k) \\
&= \lambda_i C_i(k) \cdot E^+(k-1)\{A(k-1)x(k-1) + N(k-1)u(k-1)x(k-1) + B(k-1)u(k-1) + w(k-1) \\
&\quad - [A(k-1)\hat{x}_i(k-1) + L_i(k-1)u(k-1) + N(k-1)u(k-1)\hat{x}_i(k-1) \\
&\quad + \sum_{j \in Ni} H(k-1)[Z_j(k-1) - \lambda_j C_j(k-1)x(k-1)]]\} + v_i(k) + s_i(k) \\
&= \lambda_i C_i(k) \cdot E^+(k-1)\{A(k-1)[x_i(k-1) - \hat{x}_i(k-1)] + N(k-1)u(k-1)[x_i(k-1) - \hat{x}_i(k-1)] \\
&\quad + w(k-1) - \sum_{j \in Ni} H(k-1)[\zeta_j(k-2) + s_j(k-1)]\} + v_i(k) + s_i(k)
\end{aligned}
\tag{26}
$$

Theorem 2. *The given residual generator satisfies the following distribution*

$$
Ti(k)\{E[v_i(k)], \ \tilde{P}(k)\}
\tag{27}
$$

where the mean value and variance of v are derived by following equation:

$$
v_i(k) = y_i(k) - \lambda_i C_i(k)\hat{x}_i(k) - D_i(k)u(k) - s_i(k)
\tag{28}
$$

Proof of Theorem 2.

$$
E[Ti(k)] = E[v_i(k)]
\tag{29}
$$

$$
\begin{aligned}
\tilde{P}(k) &= E[Ti(k) \cdot Ti^T(k)] \\
&= [\lambda_i C_i(k)] \cdot E^+(k-1)\{Q(k-1) - \sum_{j \in Ni} H(k-1)[R(k-2) + O(k-1)]H^T(k-1)\} \\
&\quad \cdot \{[\lambda_i C_i(k)] \cdot E^+(k-1)\}^T + E[v_i(k) \cdot v_i^T(k)] + O(k)
\end{aligned}
\tag{30}
$$

The fault diagnosis threshold is obtained by the Chebyshev inequality:

$$
\Pr(\mu(X) - \alpha\sigma(X) \le X \le \mu(X) + \alpha\sigma(X)) \ge 1 - \frac{1}{\alpha^2}
\tag{31}
$$

□

4. Practical Verifications and Discussion

4.1. State Estimation

Firstly, this paper presents the state estimation curve of four nodes in monitoring systems, as shown in Figure 4.

Figure 5a,b, respectively, shows the state estimation error value and state estimation error covariance of a single node. It can be seen that the state estimation error keeps within ±0.5, while the state estimation error covariance converges to a smaller value, proving the effectiveness of the filter.
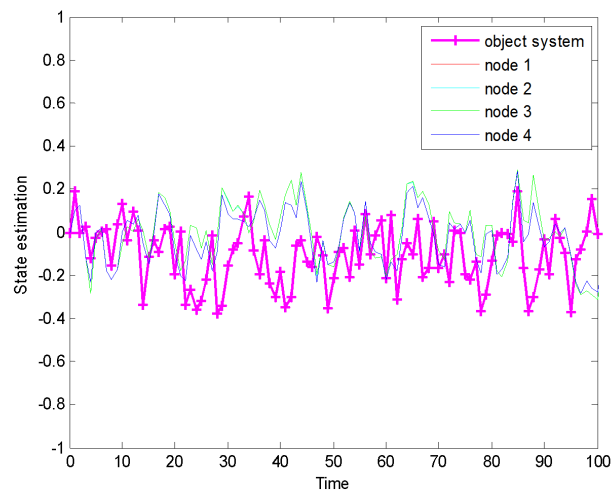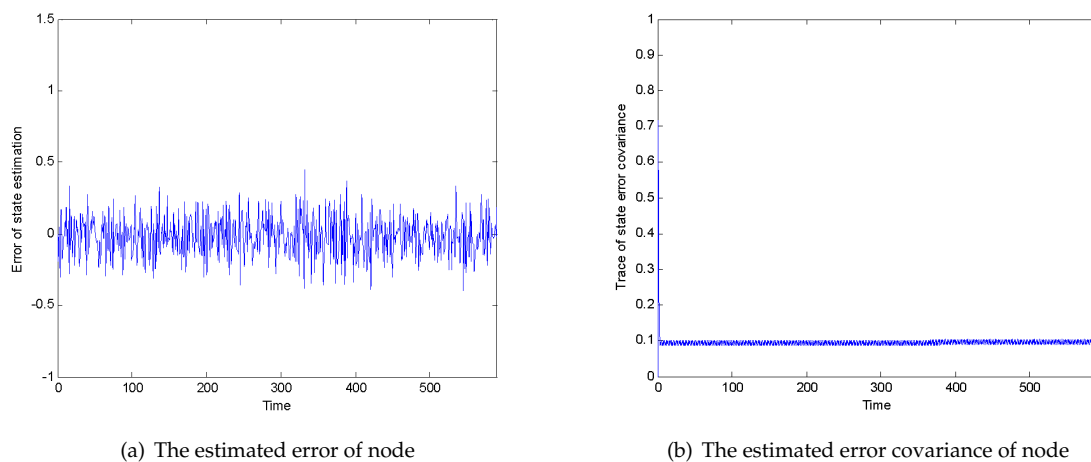
Figure 4. Distributed state estimation.

(a) The estimated error of node

(b) The estimated error covariance of node

Figure 5. (**a**) The estimated error of a node; (**b**) the estimated error covariance of a node.

4.2. Cincinnati Dataset

First, the 2003 Cincinnati test 1 bearing 2 channel 4 data was used for simulation experiments. The sampling time of this dataset is from 12:06:24 on 22 October 2003 to 23:39:56 on 25 November 2003. The test bench device contains four bearings, and the AC motor is coupled to the shaft through a friction band. The speed was kept at 2000 RPM. A 6000-pound radial load was applied to the shaft and bearings. Each dataset consists of a single file of 1 second. The sampling rate is 20 kHz, and 20,480 points are saved at a time.

The degree of bearing damage in the experiment is related to operating time of device, and the corresponding data format will also change. Generally speaking, it can be clearly seen from the amplitude of the data whether the monitoring system has failed. Therefore, to better illustrate the reliability of algorithm, this article selects some data from above data set, including non-faulty data points, early fault data points, and severe fault data points. In this paper, the initial wear state of the bearing is defined as early failure, but it does not affect the continued operation of the whole system. A serious fault may refer to the state of near failure of the bearing in the period before the end of the experiment. At this point, the serious fault has affected the operation of system and even caused some damage to bearing itself. Therefore, it is of greater practical significance to the diagnosis of serious fault, and diagnosis of early fault has certain guiding significance to prediction and maintenance of

the future.To better demonstrate effectiveness of the proposed method, data with small amplitude differences were selected from above three types of data for verification, as shown in Figure 6.
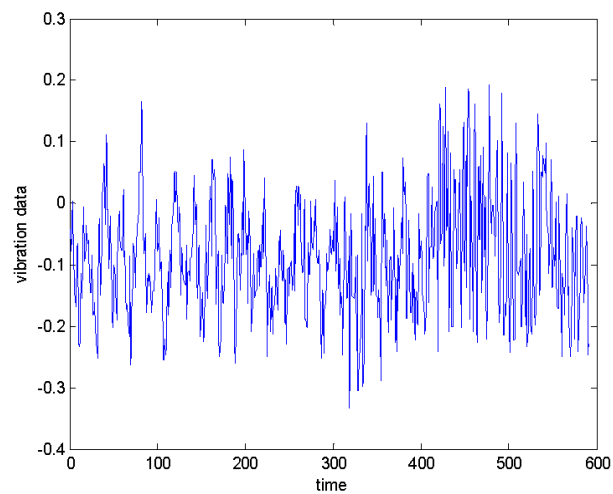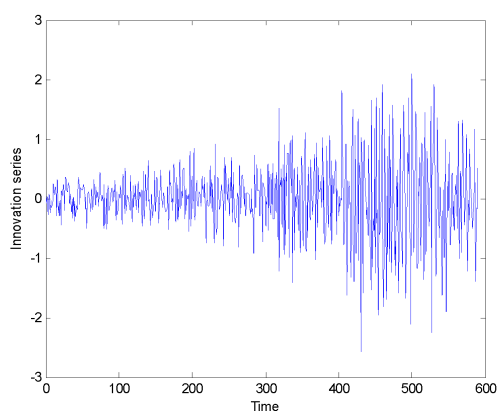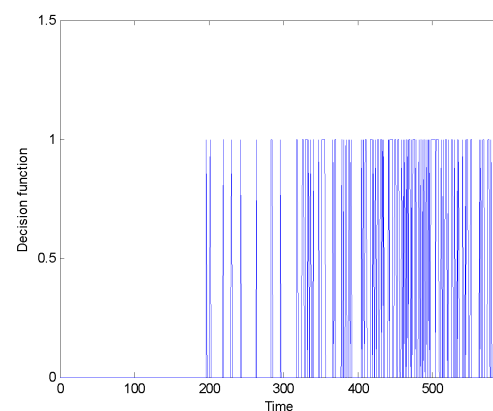

Figure 6. Cincinnati bearing vibration data values.

Among them, time 0–200 is a non-fault data point; time 200–400 is an early fault data point; 400–600 is a severe fault data point. After passing the filter with state degradation, the innovation vector is obtained as shown in Figure 7a. Taking the adjustable parameter $\alpha = 7$ in the Chebyshev inequality, the threshold value obtained are ± 0.53. For the convenience of viewing, a decision function as shown in Figure 7b is given. The threshold exceeded is set to 1, and the threshold not exceeded is set to 0. It can be seen that the fault occurred from 200–400 moments, but because the fault occurred early, the number is relatively sparse, and the fault is a serious fault at 400–600 moments, and the number of faults is relatively large.

(a) Innovation vector of Cincinnati vibration data

(b) Fault decision function of Cincinnati bearing vibration data

Figure 7. (a) Innovation vector of Cincinnati vibration data; **(b)** fault decision function of Cincinnati bearing vibration data.

4.3. High-Speed Train Temperature Dataset

This section uses an actual running gears system to prove the effectiveness of proposed algorithm. Selecting data of the failed train and the data of non-faulted train. The data is derived from temperature value of temperature sensor on the right position of the 2 axis of one-run part of the train, in Figure 8.
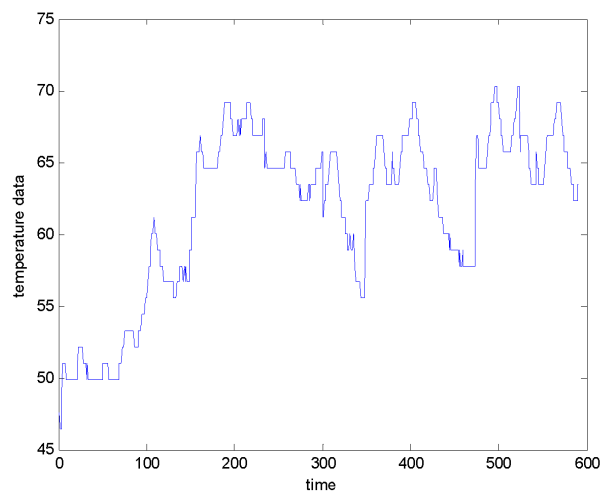
Figure 8. Temperature data of running gears system

Time 0–300 is a non-faulty data point, and time 300–600 is a faulty data point. After passing through the filter with state degradation, the innovation vector is obtained as shown in Figure 9a. Taking the adjustable parameter $\alpha = 7$ of Chebyshev inequality, the obtained threshold is ± 0.56. Similarly, for the convenience of viewing, the decision function shown in Figure 9b is given, and it can be seen that there is a fault in time 300–600, which is consistent with the error information given by actual high-speed train.
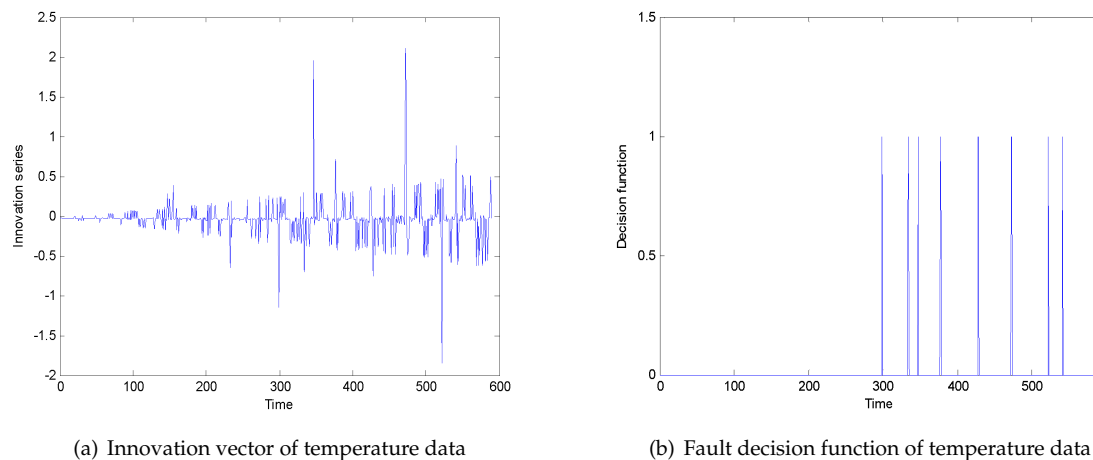

(a) Innovation vector of temperature data

(b) Fault decision function of temperature data

Figure 9. (**a**) Innovation vector of temperature data; (**b**) fault decision function of temperature data.

4.4. Performance Comparison

In this paper, the proposed algorithm is compared with algorithms commonly used in fault diagnosis. Partial data of the Cincinnati public dataset is selected for fault diagnosis. The diagnostic accuracy is shown in following table (see Table 1). Among them, FNR is the rate of false negatives and FPR represents the rate of false positives.

Table 1. Performance comparison.

Methods	FNR	FPR
State-degraded distributed filter	0.26%	0.52%
Kalman filter	1.28%	6.03%
Extended Kalman filter	2.04%	5.46%

5. Conclusions

The stable operation of high-speed trains is closely related to the running gears system. This paper proposes a method based on the Wiener state degradation process and multi-sensor filtering for running gears system diagnosis. First, a distributed topology diagram of running gears system axle box bearings and its composite sensors is established by analyzing a running system of a high-speed train. Secondly, a distributed filter is proposed to the problem of unknown initial information of sensor nodes by the unbiased constraint condition, considering the problems of gain attenuation and precision decline, and the gain parameters of the filter are calculated under the premise of minimum error mean square error. Then, to improve the fault diagnosis accuracy of the running gears system, this paper built a nonlinear degradation model of the Wiener process considering the factor of temperature, and estimated parameters through maximum likelihood estimation. Finally, a fault diagnosis threshold is obtained by using the Chebyshev inequality. The performance of the method is verified by an open dataset of rotating mechanical bearings and temperature data of a certain type of high-speed rail running gears system. The result is as expected.

The method proposed in this paper is more meaningful for practical engineering, and improves the accuracy of fault diagnosis for systems that cannot obtain available data. This method aims to diagnose early faults. For other types of dynamic systems, thresholds can be set based on their actual characteristics and the definition of the corresponding fault level. Future research work starts from the correlation between states, decouples filter states, and considers a non-linear degradation problem of non-Markov processes, making the model closer to the actual process.

Author Contributions: C.C. proposed the idea and designed the experiment. W.W. conducted experiments and completed the manuscript. H.L. provides guidance for data analysis and thesis writing. B.Z. made recommendations and supervised the work. G.C. and W.T. provide the experimental environment/equipment. All authors have read and agreed to the published version of the manuscript.

References

1. Ji, H.; He, X.; Shang, J.; Zhou, D. Incipient fault detection with smoothing techniques in statistical process monitoring. *Control Eng. Pract.* **2017**, *62*, 11–21. [CrossRef]
2. Chen, H.; Jiang, B.; Lu, N. A multi-mode incipient sensor fault detection and diagnosis method for electrical traction systems. *Int. J. Control* **2018**, *16*, 1783–1793. [CrossRef]
3. Su, L.; Ma, L.; Qin, N.; Huang, D.; Andrew, H.K. Fault diagnosis of high-speed rrain bogie by residual-squeeze net. *IEEE Trans. Ind. Inform.* **2019**, *15*, 3856–3863. [CrossRef]
4. Huang, J.; Yan, X. Quality-driven principal component analysis combined with kernel least squares for multivariate statistical process monitoring. *IEEE Trans. Control Syst. Technol.* **2019**, *27*, 2688–2695. [CrossRef]
5. Chen, H.; Jiang, B.; Ding, S.X.; Lu, N.; Chen, W. Probability-relevant incipient fault detection and diagnosis methodology with applications to electric drive systems. *IEEE Trans. Control Syst. Technol.* **2018**, *27*, 2766–2773. [CrossRef]
6. Wang, Z.; Jia, L.; Kou, L.; Qin, Y. Spectral kurtosis entropy and weighted SaE-ELM for bogie fault diagnosis under variable conditions. *Sensors* **2018**, *18*, 1705. [CrossRef]
7. Fu, J.; Chu, J.; Guo, P.; Chen, Z. Condition monitoring of wind turbine gearbox bearing based on deep learning model. *IEEE Access* **2019**, *7*, 57078–57097. [CrossRef]

8. Zhou, R.; Bao, W.; Li, N.; Huang, X.; Yu, D. Mechanical equipment fault diagnosis based on redundant second generation wavelet packet transform. *Digit. Signal Process.* **2010**, *20*, 276–288. [CrossRef]

9. Cai, B.; Liu, Y.; Fan, Q.; Zhang, Y.; Liu, Z.; Yu, S.; Ji, R. Multi-source information fusion based fault diagnosis of ground-source heat pump using Bayesian network. *Appl. Energy* **2014**, *114*, 1–9. [CrossRef]

10. Irhoumah, M.; Pusca, R.; Lefevre, E.; Lefevre, E.; Mercier, D.; Romary, R.; Cristian, D. Information fusion with belief functions for detection of inter-turn short circuit faults in electrical machines using external flux sensors. *IEEE Trans. Ind. Electron.* **2017**, *65*, 2642–2652. [CrossRef]

11. Gao, Z.; Cecati, C.; Ding, S. A survey of fault diagnosis and fault-tolerant techniques-part I: fault diagnosis with model-based and signal-based approaches. *IEEE Trans. Ind. Electron.* **2015**, *62*, 3757–3767. [CrossRef]

12. Yan, K.; Ji, Z.; Shen, W. Online fault detection methods for chillers combining extended kalman filter and recursive one-class SVM. *Neurocomputing* **2016**, *228*, 205–212. [CrossRef]

13. Ning, Y.; Wang, J.; Han, H.; Tan, X.; Liu, T. An optimal radial basis function neural network enhanced adaptive robust kalman filter for GNSS/INS integrated systems in complex urban areas. *Sensors* **2018**, *18*, 3091. [CrossRef]

14. Wang, Y.; Sun, Y.; Chang, C.l.; Hu, Y. Model-based fault detection and fault tolerant control of SCR urea injection system. *IEEE Trans. Veh. Technol.* **2016**, *65*, 4645–4654. [CrossRef]

15. Akai, N.; Morales, Y.; Murase, H. Simultaneous pose and reliability estimation using convolutional neural network and Rao–Blackwellized particle filter. *Adv. Robot.* **2018**, *32*, 930–944. [CrossRef]

16. Wu, Y.; Jiang, B.; Lu, N. A descriptor system approach for estimation of incipient faults with application to high-speed railway traction devices. *IEEE Trans. Syst. Man Cybern. Syst.* **2017**, *49*, 2108–2118. [CrossRef]

17. Yin, J.; Xie, Y.; Chen, Z.; Peng, T. Weak-fault diagnosis using state-transition-algorithm-based adaptive stochastic-resonance method. *J. Cent. South Univ.* **2019**, *26*, 1910–1920. [CrossRef]

18. Cadini, F.; Sbarufatti, C.; Corbetta, M.; Giglio, M. A particle filter-based model selection algorithm for fatigue damage identification on aeronautical structures. *Struct. Control Health Monit.* **2017**, *24*, 2002. [CrossRef]

19. Jain, T.; Yame, J. Fault-tolerant economic model predictive control for wind turbines. *IEEE Trans. Sustain. Energy* **2018**, *10*, 1696–1704. [CrossRef]

20. Bai, W.; Dong, H.; Yao, X.; Ning, B. Robust fault detection for the dynamics of high-speed train with multi-source finite frequency interference. *ISA Trans.* **2018**, *75*, 76–87. [CrossRef] [PubMed]

21. He, J.; Li, Y.; Jiang, Y.; Li, Y.; Li, A. Propeller fault diagnosis based on a rank particle filter for autonomous underwater vehicles. *Brodogr. Brodogr. Theory Pract.* **2018**, *69*, 147–164. [CrossRef]

22. Zhou, W.; Li, X.; Yi, J.; He, H. A novel UKF-RBF method based on adaptive noise factor for fault diagnosis in pumping unit. *IEEE Trans. Ind. Inform.* **2018**, *15*, 1415–1424. [CrossRef]

23. Liu, S.; Jiang, B.; Mao, Z.; Ding, S. Adaptive backstepping based fault-tolerant control for high-speed trains with actuator faults. *Int. J. Control Autom. Syst.* **2019**, *17*, 1408–1420. [CrossRef]

24. Gan, X.; Gao, W.; Dai, Z.; Liu, W. Research on WNN soft fault diagnosis for analog circuit based on adaptive UKF algorithm. *Appl. Soft Comput.* **2016**, *50*, 252–259.

25. Jing, L.; Wang, T.; Zhao, M.; Wang, P. An adaptive multi-sensor data fusion method based on deep convolutional neural networks for fault diagnosis of planetary gearbox. *Sensors* **2017**, *17*, 414. [CrossRef] [PubMed]

26. Liu, Z.; Guo, W.; Tang, Z.; Chen, Y. Multi-sensor data fusion using a relevance vector machine based on an ant colony for gearbox fault detection. *Sensors* **2015**, *15*, 21857–21875. [CrossRef]

27. Niu, G.; Han, T.; Yang, B.; Tan. A. Multi-agent decision fusion for motor fault diagnosis. *Mech. Syst. Signal Process.* **2007**, *21*, 1285–1299. [CrossRef]

28. Banerjee, T.; Das, S. Multi-sensor data fusion using support vector machine for motor fault detection. *Inf. Sci.* **2012**, *217*, 96–107. [CrossRef]

29. Viegas, D.; Batista, P.; Oliveira, P.; Silvestre, C.; Chen, C.L.P. Distributed state estimation for linear multi-agent systems with time-varying measurement topology. *Automatica* **2015**, *54*, 72–79. [CrossRef]

30. Chen, Z.; Cao, Y.; Ding, S.; Zhang, K.; Koenings, T.; Peng, T.; Yang, C.; Gui, W. A Distributed Canonical Correlation Analysis-based Fault Detection Method for Plant-wide Process Monitoring. *IEEE Trans. Ind. Inform.* **2019**, *15*, 2710–2720. [CrossRef]

31. Cai, H.; Lewis, F. L.; Hu, G.; Huang, J. The adaptive distributed observer approach to the cooperative output regulation of linear multi-agent systems. *Automatica* **2017**, *75*, 299–305. [CrossRef]

32. Mohammad, P.; Ehsan, H.; Amir, K.; Baris, F.; Bakhtiar, L.; Shih-ken, C.; Shreyas, S. Cooperative vehicle speed fault diagnosis and correction. *IEEE Trans. Intell. Transp. Syst.* **2018**, *20*, 783–789.

33. Wu, C.; Zhu, C.; Ge, Y. A new fault diagnosis and prognosis technology for high-power lithium-ion battery. *IEEE Trans. Plasma Sci.* **2017**, *45*, 1533–1538. [CrossRef]

34. Wang, X.; McArthur, S.; Strachan, S.; Kirkwood, J.; Paisley, B. A data analytic approach to automatic fault diagnosis and prognosis for distribution automation. *IEEE Trans. Smart Grid* **2018**, *9*, 6265–6273. [CrossRef]

35. Yan, W.; Zhang, B.; Dou, W.; Liu, D.; Peng, Y. Low-cost adaptive lebesgue sampling particle filtering approach for real-time li-ion battery diagnosis and prognosis. *IEEE Trans. Autom. Sci. Eng.* **2017**, *14*, 1601–1611. [CrossRef]