Chinese Journal of Electronics
Vol.32, No.5, Sept. 2023

Sigma-Mixed Unscented Kalman Filter-Based Fault Detection for Traction Systems in High-Speed Trains

CHENG Chao[1], WANG Weijun[2], MENG Xiangxi[3], SHAO Haidong[4], and CHEN Hongtian[5]

(1. *Department of Computer Science and Technology, Changchun University of Technology, Changchun 130012, China*)

(2. *Department of Mathematics and Statistics, Changchun University of Technology, Changchun 130012, China*)

(3. *Institute of System Research, China Industrial Control Systems Cyber Emergency Response Team, Beijing 100040, China*)

(4. *Department of Mechanical and Vehicle Engineering, Hunan University, Changsha 410082, China*)

(5. *Department of Chemical and Materials Engineering, University of Alberta, Edmonton AB T6G 2V4, Canada*)

Abstract — Fault detection (FD) for traction systems is one of the active topics in the railway and academia because it is the initial step for the running reliability and safety of high-speed trains. Heterogeneity of data and complexity of systems have brought new challenges to the traditional FD methods. For addressing these challenges, this paper designs an FD algorithm based on the improved unscented Kalman filter (UKF) with consideration of performance degradation. It is derived by incorporating a degradation process into the state-space model. The network topology of traction systems is taken into consideration for improving the performance of state estimation. We first obtain the mixture distribution by the mixture of sigma points in UKF. Then, the Lévy process with jump points is introduced to construct the degradation model. Finally, the moving average interstate standard deviation (MAISD) is designed for detecting faults. Verifying the proposed methods via a traction systems in a certain type of trains obtains satisfactory results.

Key words — Fault detection, Unscented Kalman filter, Traction systems, State degradation.

I. Introduction

The emergence of high-speed trains has brought great convenience to our life. It not only reduces travel time for passengers, but also contributes to the growth of social benefits [1]–[3]. Meanwhile, the safety of high-speed trains has been paid much attention accompanied by the increasing of railway industry. Until now, railway accidents are still not be avoided completely. According to the records of cases, more than seven accidents per year occur around the world, and the fault of system units is the primary cause of catastrophic accidents [4], [5]. In 2012, a severe rail disaster caused by a circuit fault took place in Buenos Aires. Fifty-one people were killed and more than 700 were injured in this accident. In addition, the train collision occurred on the Yongtaiwen railway line in the suburbs of Wenzhou. It is a fatal crash caused by Chinese high-speed trains, and is the third-deadliest high-speed train accident in history. Therefore, the railway is faced with the pressure of rapid transit that calls for significant improvements in reliability and rail safety [6]–[8]. Fault detection (FD), as the primary study for system reliability, is becoming increasingly important in current researches.

As a powerful tool for FD in high-speed trains, the model-based method improves the system reliability by utilizing redundancy information [9]–[11]. It differs from the signal analysis-based methods, which have limited efficiency for FD tasks in dynamic cases. And the online implementation is also a challenging problem. Many improved methods cannot get satisfactory results for real-time performance [12]. Another kind of method favored by researchers is the data-driven method which receives significant attention in recent years [13]–[15]. The main procedure of this type of method concludes

Manuscript Received May 29, 2022; Accepted Oct. 8, 2022. This work was supported by the National Natural Science Foundation of China (61903047, U20A20186), the Science and Technology Department of Jilin Province (20210201113GX), and the Education Department of Jilin Province (JJKH20210754KJ).

 DOI:10.23919/cje.2022.00.154

off-line modeling and online detection. Although the existing methods can achieve the enhanced performance, some limitations are still evident in practical applications. The FD performance rely on both the quality and size of data collected from sensors unavoidably. However, for some systems with high safety requirements, the results will be affected when fault data are not enough for analysis. Consequently, the model-based method shows its own advantages than data-driven ones.

A preliminary attempt introduced a state-space model of high-speed trains for detecting the faults [16]. For improving the dynamic performance, it considered the nonlinear characteristics and the disturbance attenuation of models. Up to now, the state-space model has been a favored solution for FD [17], [18], fault prognostic [19], [20], and fault isolation [21], [22]. For traction control systems, the improved methods based on state-space models have been used such as the classical filter [23], the complementary filter [24], event-driven frameworks [25], [26], and the residual-based estimations [27]. Among them, particle filter and unscented Kalman filter (UKF) are adopted commonly due to the good tracing performance for nonlinear systems. Their advantages lie in accuracy and efficiency of calculations. Considering the studies above, the accurate mathematical model and statistics are critical for fault detection based on the state-space model. In addition, few studies have considered the multi-sensor monitoring of the traction systems. It may be attributed to the fact that there is no communication between nodes in this scene. It is not appropriate to force a distributed approach to detect faults. Different from the electrical parameters, the temperature of target point will be affected by the environment and other measured points in the same monitoring system, even if there is no node interconnection. It will result in the complex distribution of states as shown in Fig.1.

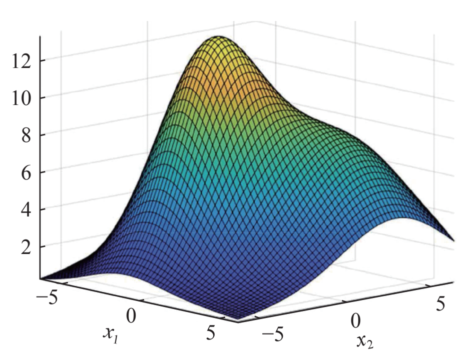

Fig. 1. The complex distribution of systems in actual scenarios.

Besides, state degradation commonly occurs in traction systems. As given in studies of prognostics and health management, the degradation model based on Wiener process is highly focused. It is usually considered to be applicable to systems with random initial values. In fact, most existing stochastics models, such as Gamma process [28], inverse Gaussian process [29], and Lévy process [30], [31], are also useful to avoid the initial unpredictability by addressing deterioration increments or percentage drops in performance. Regrettably, for traction systems, all these models are difficult to apply directly, because of the existence of jump points.

Motivated by the above discussions, this study are exploiting the state estimation theory to finish the FD task of traction systems. Considering the real-time performance of systems with high sampling frequency, this paper focuses on the improvement strategy of UKF. The contributions of the work are summarized as follows.

1) A sigma-mixed UKF is proposed for estimating the states with coupling distribution. The parameters in a mixture model are estimated by the expectation-maximization algorithm.

2) A degradation model with jumps is constructed based on the Lévy process that is a general form for most stochastic processes. The decomposability of the degradation model is given by two lemmas and realized via wavelet transform.

3) A test statistic based on moving average interstate standard deviation (MAISD) is put forward for FD. The sufficiency and completeness of statistics are considered based on the factorization theorem and the properties of exponential families.

The remaining part of this article is organized as follows: Section II introduces the mechanism of traction systems, the general degradation model, and the important steps of linear filtering in statistical derivation. Section III proposes the improved UKF and the degradation model based on Lévy process with jumps, followed by the proposed FD method. Section IV provides the verification by using the temperature of traction systems in a certain type of high-speed train. Finally, Section V concludes the study.

II. Preliminaries

In this section, the traction systems of high-speed trains, the general degradation model, and the statistical derivation of linear filtering are introduced. These bases are the fundamental research of the following sections.

1. Traction systems of high-speed trains

Different to DC motors driven systems, high-speed trains are equipped with three-phase asynchronous motors that are tiny and light. It has a straightforward structure and strong tractability. The three-phase asynchronous motor is made up of two primary components. The stator is the fixed part, and the rotor is the rotating part, and a little gap between them known as the air gap. Three measuring points called the driving terminal point, non-driving terminal point, and stator point are installed on each traction motor. The specific monitoring systems in traction systems are is shown in Fig.2.

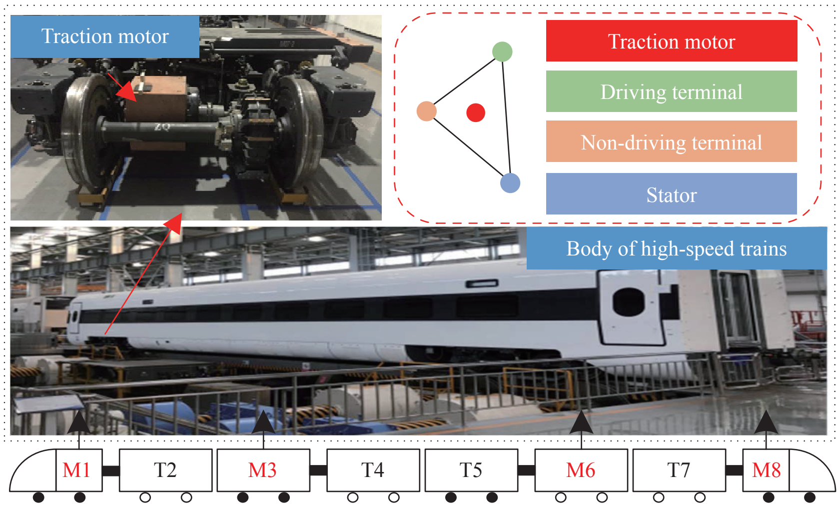

Fig. 2. The structure of traction systems and multi-sensor monitoring.

Relying on the mechanism of traction systems, the temperature change can be expressed in the following form:

$$T_{k+1} = T_k + \frac{(I_k - I_{k-1})^2 R + \xi \cdot (T_k - T_{\text{out}})}{c \cdot m} \tag{1}$$

where T stands for the temperature collected by sensors. The I and R represent the current and resistance of traction motor, respectively. T_{out} is the external temperature. In addition, c stands for the specific heat capacity, and m represents the quality of the motor.

2. General degradation process

Up to now, the most commonly used degradation model is constructed based on Wiener process which is driven by brownian motion. The Wiener process can be constructed as the scaling limit of a random walk, or other discrete-time stochastic processes with stationary independent increments. Therefore, the mechanical degradation with stochastic characteristic can be described by this special Markov process whose incremental change obeys a normal distribution. The general form is shown as

$$X_t = \mu' t + \sigma' B(t) \tag{2}$$

where μ' and σ' are drift and diffusion coefficient, respectively.

It is characterised by the following propoerties:

1) X_t has independent increments: for every $t > 0$, the increments $X_{t+d} - X_t$, $d \geq 0$, are independent of the past values X_s, $s \leq t$.

2) X_t has Gaussian increments: $W_{t+d} - W_t$ is normally distributed with a mean of 0 and a variance of d, $W_{t+d} - W_t \sim \mathcal{N}(0, d)$.

3) X_t has continuous paths: X_t is continuous in t.

Considering the external disturbance of systems, this study uses the jump points to describe the swift changes in actual process. It should be noted that such temperature jumps in the actual environment are not common in the laboratory equipment.

3. Statistical derivations of linear Kalman filter

The focus of linear filtering is to update a priori estimation using the measurement at the time instant k. Therefore, the posterior state of the basic linear process can be described by

$$\hat{x}^+(k) = Ky(k) + b \tag{3}$$

where b is unknown vector. The defect of unknown prior information can be solved by calculating the mean value of posteriori estimation. The unbiased constraint can be expressed as

$$b = \bar{x} - K\bar{y} \tag{4}$$

where K stands for the gain parameter. The posterior covariance is given as follows.

$$P^+(k) = E\{[x(k) - \hat{x}^+(k)][x(k) - \hat{x}^+(k)]^{\mathrm{T}}\} \tag{5}$$

For the error of estimation $e(k) = x(k) - \hat{x}^+(k)$, the trace of the error covariance is

$$\begin{aligned}\mathrm{tr}(P^+(k)) = &\,\mathrm{tr}[P^-(k) - KP_{yx} - P_{xy}K^{\mathrm{T}} + KP_yK^{\mathrm{T}}]\\ &+ \mathrm{tr}[(\bar{x} - K\bar{y} - b)(\bar{x} - K\bar{y} - b)^{\mathrm{T}}]\end{aligned} \tag{6}$$

After minimizing the above equation, it can be obtained respectively that

$$\begin{aligned}&K = P_{xy}P_y^{-1}\\ &\hat{x}^+(k) = \hat{x}^-(k) + K[y(k) - \hat{y}(k)]\\ &P^+(k) = P^-(k) - KP_yK^{\mathrm{T}}\end{aligned} \tag{7}$$

Different from the conventional derivation method, this form focuses on updating the system state according to the measurement value at the time instant k.

Remark 1 The traction systems are monitored by different type of sensors in high-speed trains. Therefore, the traditional data mixture method connot be used for analysis and calculation. To solve this problem, this study achieves state estimation of nonlinear systems through mixed the distribution of states in UKF.

III. The Proposed Method Based on UKF

In this section, the sigma-mixed UKF is first presented, then the model of degradation and detailed FD strategy will be followed.

1. Generation of sigma points in UKF

The UKF has a good effect on state estimation of nonlinear systems, which is more accurate in propagation of means and covariance than linearized extended Kalman filter (EKF). A set of sigma points to approximate the probability distribution are required in unscented transformation. For traction systems with state coupling characteristics, the mean of the sigma points cannot be calculated through a single distribution. Therefore, this study mixes the sigma points by means of mixed distributions for obtaining the reliable result.

Consider the discrete-time random nonlinear system as follows:

$$\begin{aligned}&x(k+1) = f[x(k), u(k)] + w(k)\\ &y(k) = h[x(k), u(k)] + v(k)\end{aligned} \tag{8}$$

where $x \in \mathcal{R}^{k_x}$ represents the state of systems. $u \in \mathcal{R}^{k_u}$ and $y \in \mathcal{R}^{k_y}$ are the input and output of systems. $w \in \mathcal{R}^{k_x}$ and $v \in \mathcal{R}^{k_y}$ are white noises with covariances Q and R, respectively.

Assuming that the state distribution of the monitoring system is determined by n different distributions, the probability density function is

$$p(x|\theta) = \sum_{k=1}^{n} p(k)p(x\,|k,\theta\,) = \sum_{k=1}^{n} \alpha_k \mathcal{N}(x\,|\mu_k, \varSigma_k\,) \tag{9}$$

where α_k is the mixing coefficients. The condition below is required.

$$\sum_{k=1}^{n} \alpha_k = 1 \tag{10}$$

In (9), every single distribution can represented by

$$f(x|\theta) = \frac{1}{\sqrt{2\pi\sigma^2}} \mathrm{e}^{-\frac{(x-\mu)^2}{2\sigma^2}} \tag{11}$$

Define implicit variable γ that reflects state x_s, $s = 1, 2, \dots, N$ from the distribution. A likelihood function that combines the complete data of the hidden variable is

$$\begin{aligned}p(x, \gamma|\theta) &= \prod_{s=1}^{N} p(x_s, \gamma_{s1}, \gamma_{s2}, \dots, \gamma_{sn}|\theta)\\ &= \prod_{i=1}^{n} \alpha_i^{\gamma_{si}} \prod_{s=1}^{N} \left[\frac{1}{\sqrt{2\pi\sigma_i^2}} \mathrm{e}^{-\frac{(x_s-\mu_i)^2}{2\sigma_i^2}}\right]^{\gamma_{si}}\end{aligned} \tag{12}$$

where the value of γ is 1 or 0, which represents the s-th measurement from the i-th distribution. Then, the logarithmic likelihood function of complete data is expressed as

$$\begin{aligned}\log p(x, \gamma|\theta) = &\sum_{i=1}^{n} \left\{\gamma_{si} \log \alpha_i + \sum_{s=1}^{N} \gamma_{si} \left[\log\left(\frac{1}{\sqrt{2\pi}}\right)\right.\right.\\ &\left.\left. - \log \sigma_i - \frac{1}{2\sigma_i^2}(x_s - \mu_i)^2\right]\right\}\end{aligned} \tag{13}$$

The EM algorithm is used to estimate the parameters of the mixed distribution. Establish Q function as follows:

$$\begin{aligned}Q(\theta, \theta^{(t)}) &= E[\log p(x, \gamma|\theta)|x, \theta^{(t)}]\\ &= E\left\{\sum_{i=1}^{n} \left\{\gamma_{si} \log \alpha_i + \sum_{s=1}^{N} \gamma_{si} \left[\log\left(\frac{1}{\sqrt{2\pi}}\right)\right.\right.\right.\\ &\left.\left.\left. - \log \sigma_i - \frac{1}{2\sigma_i^2}(x_s - \mu_i)^2\right]\right\}\right\}\end{aligned} \tag{14}$$

Then the probability of the s-th observation from the i-th distribution is calculated as

$$\hat{\gamma}_{si} = E(\gamma_{si}|x, \theta) = \frac{\alpha_k \phi(x_s|\theta_i)}{\sum_{i=1}^{n} \alpha_k \phi(x_s|\theta_i)} \tag{15}$$

The Q function below can be obtained via (14) and (15) as

$$Q\left(\theta,\theta^{(t)}\right)=\sum_{i=1}^{n}\left\{\gamma_{si}\log\alpha_i+\sum_{s=1}^{N}\widehat{\gamma}_{si}\left[\log\left(\frac{1}{\sqrt{2\pi}}\right)-\log\sigma_i-\frac{1}{2\sigma_i^2}\left(x_s-\mu_i\right)^2\right]\right\} \tag{16}$$

Next, the maximum value of θ is sought in the new round.

$$\theta^{(t+1)}=\arg\max_{\theta} Q(\theta,\theta^{(t)}) \tag{17}$$

Then, the parameters can be estimated via

$$\hat{\alpha}_i=\frac{\sum_{s=1}^{N}\hat{\gamma}_{si}}{N},\quad \hat{\mu}_i=\frac{\sum_{s=1}^{N}\hat{\gamma}_{si}x_s}{\sum_{s=1}^{N}\hat{\gamma}_{si}},\quad \hat{\sigma}_i^2=\frac{\sum_{s=1}^{N}\hat{\gamma}_{si}(x_s-\mu_i)^2}{\sum_{s=1}^{N}\hat{\gamma}_{si}} \tag{18}$$

The mean value of mixed distribution can be regarded as the initial priori information of the unscented transformation, and $2n+1$ sigma points are selected as follows.

$$\begin{aligned}
&x^{(0)}=\widehat{x}\\
&x^{(i)}=\widehat{x}+\tilde{x}^{(i)},\ \ i=1,2,\dots,2n\\
&\dot{x}^{(i)}=(\sqrt{(n+\kappa)P})_i^{\mathrm{T}},\ \ i=1,2,\dots,n\\
&\dot{x}^{(n+i)}=-(\sqrt{(n+\kappa)P})_i^{\mathrm{T}},\ \ i=1,2,\dots,n
\end{aligned} \tag{19}$$

The weighting factors are

$$\begin{aligned}
W^{(0)}&=\frac{\kappa}{n+\kappa}\\
W^{(i)}&=\frac{\kappa}{2(n+\kappa)},\ \ i=1,2,\dots,2n
\end{aligned} \tag{20}$$

After the obtained local solution of EM, the statistical property of the mixed distribution is

$$\widehat{x}=\bar{x}=\frac{1}{n}(x_1+x_2+\cdots+x_n) \tag{21}$$

Remark 2 The sigma-mixed UKF is appropriate for dynamic systems with a coupling distribution. It is more convenient to obtain a consistent result in multi-sensor monitoring systems.

2. State degradation based on Lévy process with jump points

The essence of a degradation model based on Wiener process is to simulate the randomness through the drift and diffusion coefficients. However, it is hard to describe the degradation process in traction systems only using these coefficient. Based on the temperature data collected from sensors, Brownian motion is not a good description of how the system changes because of the existence of jump points. Therefore, this study adopts a jump diffusion model based on the Lévy process.

The degradation model constructed in this paper is given as follows.

$$X(t)=X(0)+\mu' t+\sigma' B_t+Z_t \tag{22}$$

where Z represents the jump property of a random process. The separability of the jump points in Lévy process is illustrated by the following lemmas.

Lemma 1 [32] The Borel probability measure β of the random variable $X\in\mathcal{R}_k$ is infinitely divisible if and only if there is a triplet (μ,A,v) for all $\theta\in\mathcal{R}_k$, its characteristic function satisfies

$$E[\mathrm{e}^{i\theta X_t}]=\exp\left\{t\left(i\mu\theta-\frac{1}{2}\sigma^2\theta^2+int_{R_k}(\mathrm{e}^{i\theta x}-1-i\theta x I_{|x|\le 1})v(dx)\right)\right\} \tag{23}$$

where v is a Lévy measure, representing the number of jumps in unit time.

Lemma 2 [33] The triplet (μ,A,v) is considered, where $\mu\in\mathcal{R}_k$, A is a semi-positive definite matrix with the order k, and v is a Lévy measure of $\mathcal{R}_k$. According to the characteristic function, the Lévy process can be divided into four parts: constant drift, Brown motion, compound Poisson, and pure hop martingale. The characteristic functions are denoted as

$$\begin{aligned}
\psi_1(\theta)&=\exp(t\cdot i\mu\theta)\\
\psi_2(\theta)&=\exp\left(-\frac{t}{2}\sigma^2\theta^2\right)\\
\psi_3(\theta)&=\exp\left(t\cdot\int_{|x|\ge 1}(\mathrm{e}^{i\theta x}-1)v(dx)\right)\\
\psi_4(\theta)&=\exp\left(t\cdot\int_{|x|<1}(\mathrm{e}^{i\theta x}-1-i\theta x)v(dx)\right)
\end{aligned} \tag{24}$$

Lemma 1 suggests that each Lévy process is the sum of Brownian motion with a drift and another independent random variable. Lemma 2 describes the latter as a (stochastic) sum of independent Poisson random variables. Based on them, this study consideres the decomposition of the Lévy process into Gaussian and compound Poisson processes. The jump points are extracted by discrete wavelet transform (DWT) [34] as

$$C(2^j,b)=\frac{1}{\sqrt{2^j}}\int_R f(t)\psi^*\left(\frac{t-b}{2^j}\right)dt \tag{25}$$

The parameters of stochastic process above can be estimated by the maximum likelihood. Therefore, the difference vector of the measurements can be constructed based on

$$\tilde{y}_{1:k} = [y_1 - y_0, y_2 - y_1, \dots, y_k - y_{k-1}] \quad (26)$$

Considering $\gamma(t_i) = \tau$, the expectations m are $\mu' \cdot [\gamma(t_1)\gamma(t_2)\cdots\gamma(t_k)]^{\mathrm{T}}$. Then, the covariance of the increment is $C = Var[(y_{i+1} - y_i), (y_{j+1} - y_j)] = E[(y_{i+1} - y_i), (y_{j+1} - y_j)] - E(y_{i+1} - y_i)E(y_{j+1} - y_j)$. The estimation result of parameters is

$$\hat{\mu}' = \frac{\tilde{y}_{1:k}}{2m_{1:k}}, \quad \hat{\sigma}' = \frac{\tilde{y}_{1:k} - \mu' \cdot m_{1:k}}{\sqrt{C}} \quad (27)$$

Remark 3 The degradation model considered in this paper is a common form. In fact, the parameters are difficult to estimate when the drift term is an exponential or power function. The Nelder-Mead method can be used to maximize the logarithmic likelihood function to obtain the estimation in this situation.

3. State estimation and fault detection systems

The classical Kalman filter equations are used to complete the measurement-update phase for the system states and the covariance matrix as follows:

$$\begin{aligned} &K' = P_{xy}P_y^{-1} \\ &\widehat{x}^{+}(k) = A\widehat{x}^{-}(k) + Bu(k) + K'[y(k) - \widehat{y}(k)] \\ &\widehat{P}^{+}(k) = \widehat{P}^{-}(k) - KP_yK^{\mathrm{T}} \end{aligned} \quad (28)$$

The real state at the time instant k is calculated based on the estimated state and the degraded state, laying the groundwork for FD. In this paper, a test statistic called MAISD is presented for performing FD. For m monitoring systems in actual scenarios, the estimated states are $\hat{x}^i, i = 1, 2, \dots, m$. The moving average technique is applied for monitoring the dynamic characteristics of systems. The moving average vector is defined as follows.

$$\hat{x}_s(k) = \frac{\hat{x}(k) + \hat{x}(k+1) + \cdots + \hat{x}(k+s-1)}{s} \quad (29)$$

where s represents the size of the sliding window. Therefore, for all states estimated by the proposed method, it has the following form

$$\hat{x}_s^i(k) = \begin{bmatrix} \hat{x}_s^1(1) & \hat{x}_s^1(2) & \cdots & \hat{x}_s^1(k) \\ \hat{x}_s^2(1) & \hat{x}_s^2(2) & \cdots & \hat{x}_s^2(k) \\ \vdots & \vdots & \ddots & \vdots \\ \hat{x}_s^m(1) & \hat{x}_s^m(2) & \cdots & \hat{x}_s^m(k) \end{bmatrix} \quad (30)$$

The proposed statistic index is

$$T^i_{\text{MAISD}}(k) = \sqrt{\frac{[x^1(k) - \mu_i]^2 + \cdots + [x^i(k) - \mu_i]^2}{i}} \quad (31)$$

Remark 4 According to the central limit theorem, the samples from traction systems resemble the Gaussian distribution when the system is fault-free. Based on the factorization theorem and the properties of exponential families, MAISD is sufficient and complete.

4. Implementation of FD strategy

The implementation of state estimations and FD for traction systems is shown as Algorithms 1 and 2.

Algorithm 1 State estimation through sigma-mixed unscented kalman filter

Input: The state of systems x.
Output: The state of sigma-mixed unscented Kalman filter $\widehat{x}$.

1: Calculate the logarithmic likelihood function $\log p(x, \gamma|\theta)$ by (12);
2: Establish Q function $Q(\theta, \theta^{(t)})$ by (14);
3: Calculate $\hat{\gamma}_{si}$ by (15);
4: Estimate the parameters $\hat{\alpha}_i$, $\hat{\mu}_i$, and $\hat{\sigma}_i$ by (18);
5: Obtain the mixed state $\widehat{x}$ by (21).

Algorithm 2 Fault detection via statistic index MAISD

Input: The system input u and system output y.
Output: The estimated state $\hat{x}_s(k)$ and statistic $T^i_{\text{MAISD}}(k)$.

1: Load the data $y_i(k)$ from actual systems;
2: Calculate the difference vector $\tilde{y}_{1:k}$ by (26);
3: Estimate the parameters of degradation model μ' and σ' by (27);
4: Obtain the estimated state $\hat{x}(k)$ by (29);
5: Calculate the moving average vector $\hat{x}_s(k)$ and the statistic index $T^i_{\text{MAISD}}(k)$ by (30);

IV. Experiments and Discussions

In this section, the results of the novel FD are verified by actual scenarios. All of the experimental data are collected from the monitoring nodes in traction systems, as shown in Fig.3. The main object of this experiment is the traction motor which plays an important role in traction systems. The parameters involved in this method are shown in Table 1.

1. State estimation of sigma-mixed UKF

For illustrating the effectiveness of sigma-mixed UKF, the results of state estimation are shown as follows. The four sub-figures in Fig.4 represent the different states which contain three sets of measured values and a set of mixed values. x_1, x_2, and x_3 are the temperatures collected from the stator, driving terminal, and the non-driving terminal of motors, respectively. $x_{\text{sigma-mixed}}$ is the mixed state of three temperatures

Fig. 3. Traction motors of high-speed trains.

Table 1. Parameters of traction motors

Parameter	Description	Value
P	Rated power	350 kW
U	Rated voltage	1443 V
I	Rated current	170 A
F	Rated frequency	140 Hz
m	Weight	850 kg

mentioned above. Fig.5 shows the errors of estimation of different states.

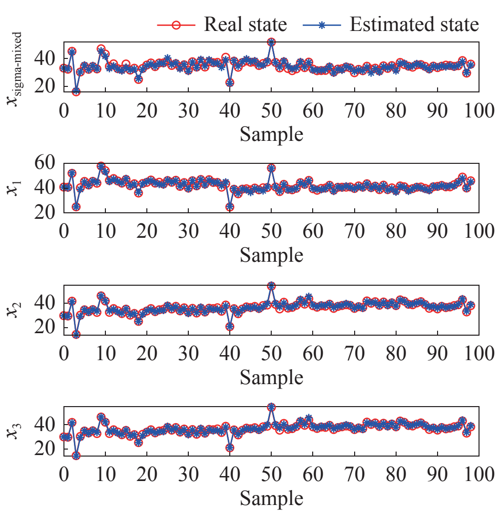

Fig. 4. State estimation results x_1, x_2, x_3 of three sets of measured values of the temperatures collected from the stator, the driving terminal, and the non-driving terminal of motors, respectively, and the mixed state $x_{\text{sigma-mixed}}$ of the three temperatures.

2. Fault detection results of traction systems

The high-speed trains considered in this studies feature a complex construction that includes four motor coaches and four trailer coaches. The four motor coaches are equipped with traction motors to provide the traction power. The states of four motors are the same when the trains are running.

Fig.6 shows the detection results for jump points. Fig.6(a) illustrates the relationship between scale parameters and time using the wavelet. Fig.6(b) shows the detailed signals after transformation. The jump points of the system state can be observed in Fig.6(b).

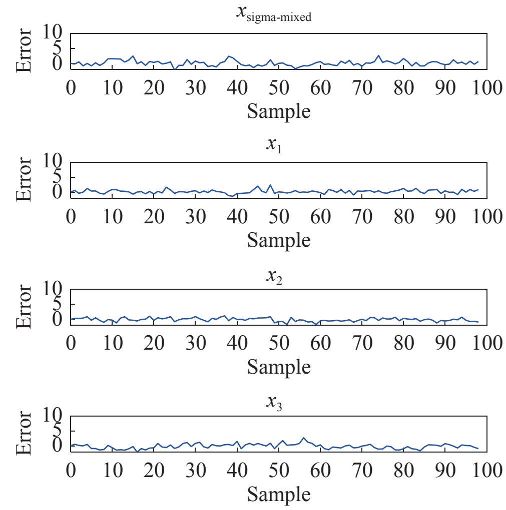

Fig. 5. Estimation error of different states.

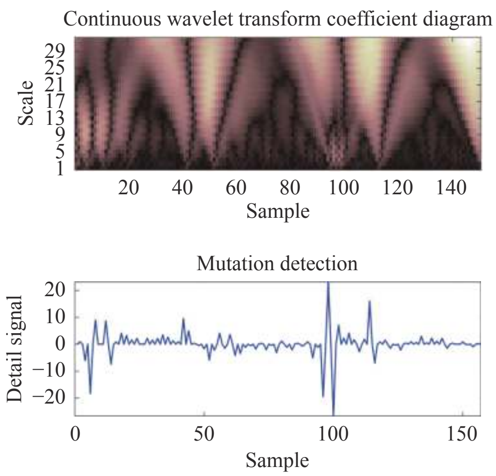

Fig. 6. Detection of jump points.

Fig.7 illustrates the estimated states of motors in different coaches and results of FD using MAISD. It should be noticed that Fig.7(b) has three parts representing the MAISDs between different states of systems. Different from others, the motor in Coach No.4 has a

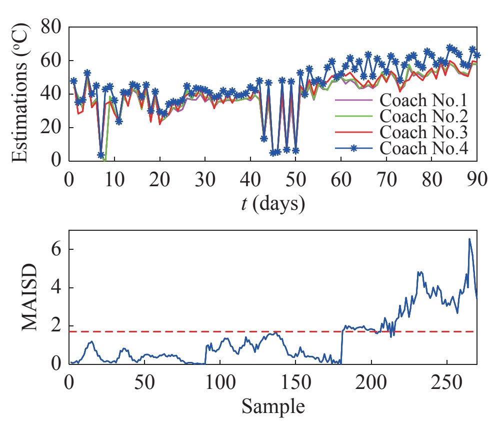

Fig. 7. State estimation of motors in different coaches.

fault in 50th day (see Fig.7(a)). Observed from the results, in Fig.7(b), the fault can be detected in samples from 180th to 270th.

3. Discussions

For demonstrating the better performance of the sigma-mixed UKF, comparison results using five methods are given in Fig.8 and Table 2. These methods are representations of the mainstream approaches. The dynamic principal component analysis (DPCA), long short-term memory (LSTM), just-in-time learning (JITL), deep slow feature analysis (DSFA), and EKF are used to detect faults in traction systems. They are widely considered as representatives to detect faults in different type of methods. In comparison to multivariate statistical analysis, this study uses different statistical indices to improve its persuasiveness of FD.

From the experimental results in Fig.8 and Table 2, the proposed method can effectively deal with the detection task of traction systems. The detecting accuracy, false positive rate (FPR), false negative rate (FNR), and total computing time are considered in the comparisons. In general, the method proposed in this paper is suitable for dynamic systems with complex distribution and nonlinear degradation. Model uncertainty and design complexity are still challenges in current research.

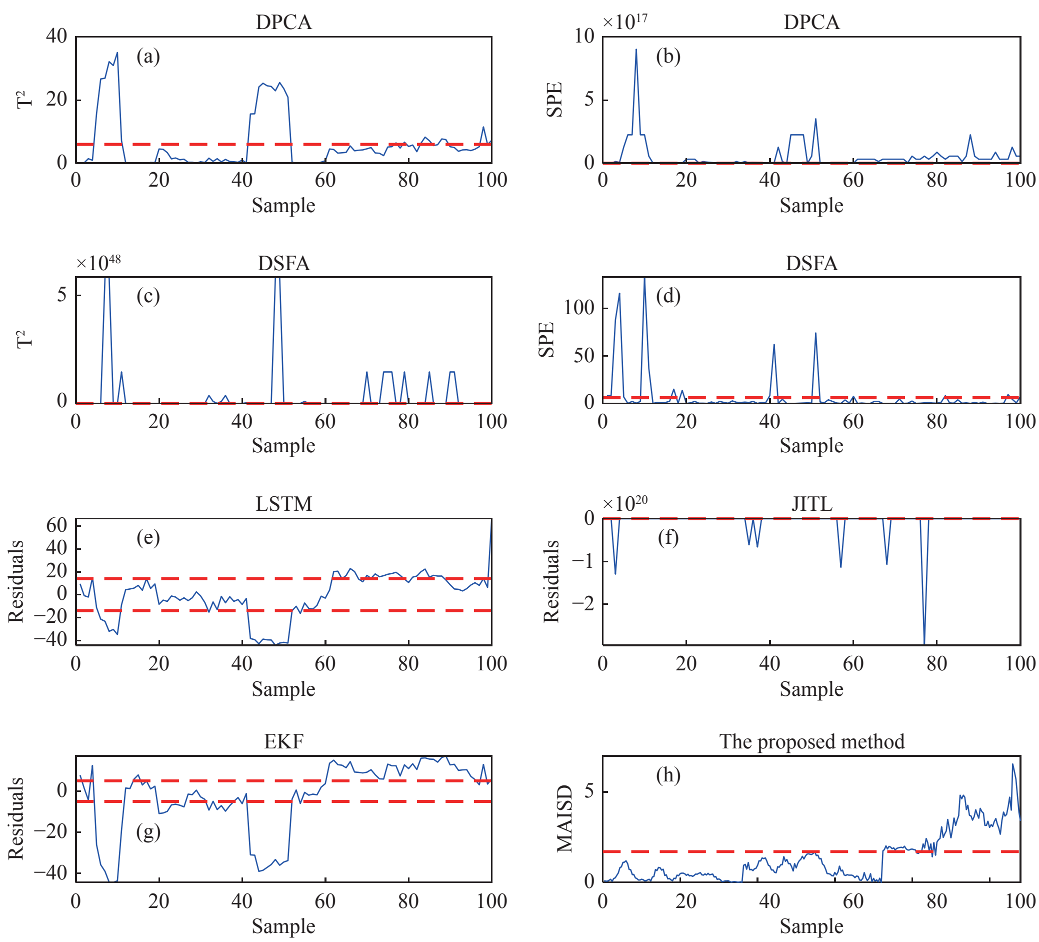

Fig. 8. Detection results using six methods. (a) and (b) DPCA; (c) and (d) DSFA; (e) LSTM; (f) JITL; (g) EKF; (h) The proposed method.

Table 2. Performance comparison using six methods

Method	Detecting accuracy	FPR	FNR	Total computing time
DPCA	71.7%	48%	31%	1.34 s
DSFA	66.7%	49%	37%	1.16 s
LSTM	90%	44%	4%	2.42 s
JITL	55%	42%	18%	3.28 s
EKF	77.5%	57%	13%	0.91 s
The proposed method	94.4%	2.75%	4.44%	1.24 s

V. Conclusions and Future Works

In this paper, a novel FD method, which is based on the state estimation and state degradation, is introduced. The main purpose of this study is to explore an effective FD strategy, which can be used appropriately in traction systems. Based on the sigma-mixed UKF and Lévy stochastic process, the problems of complex distribution and nonlinear degradation are addressed via a state-space model. The performance of the proposed method is evaluated by an actual scenario in a certain type of high-speed train. Benefiting from the established framework in this study, possible research directions such as fault isolation and identification in degraded systems can be considered.

References

[1] Y. Cao, Z. C. Wang, F. Liu, *et al.*, "Bio-inspired speed curve optimization and sliding mode tracking control for subway trains," *IEEE Transactions on Vehicular Technology*, vol.68, no.7, pp.6331–6342, 2019.

[2] H. T. Chen and B. Jiang, "A review of fault detection and diagnosis for the traction system in high-speed trains," *IEEE Transactions on Intelligent Transportation Systems*, vol.21, no.2, pp.450–465, 2020.

[3] T. Wen, G. Xie, Y. Cao, *et al.*, "A DNN-based channel model for network planning in train control systems," *IEEE Transactions on Intelligent Transportation Systems*, vol.23, no.3, pp.2392–2399, 2022.

[4] H. X. Hu, B. Tang, X. J. Gong, *et al.*, "Intelligent fault diagnosis of the high-speed train with big data based on deep neural networks," *IEEE Transactions on Industrial Informatics*, vol.13, no.4, pp.2106–2116, 2017.

[5] Y. P. Wu, W. D. Jin, Y. Li, *et al.*, "Detecting unexpected faults of high-speed train bogie based on Bayesian deep learning," *IEEE Transactions on Vehicular Technology*, vol.70, no.1, pp.158–172, 2021.

[6] H. T. Chen, B. Jiang, W. Chen, *et al.*, "Data-driven detection and diagnosis of incipient faults in electrical drives of high-speed trains," *IEEE Transactions on Industrial Electronics*, vol.66, no.6, pp.4716–4725, 2019.

[7] Y. K. Sun, Y. Cao, G. Xie, *et al.*, "Sound based fault diagnosis for RPMs based on multi-scale fractional permutation entropy and two-scale algorithm," *IEEE Transactions on Vehicular Technology*, vol.70, no.11, pp.11184–11192, 2021.

[8] Y. L. Chen, Z. B. Tian, C. Roberts, *et al.*, "Reliability and life evaluation of a DC traction power supply system considering load characteristics," *IEEE Transactions on Transportation Electrification*, vol.7, no.3, pp.958–968, 2021.

[9] X. H. Sun, C. L. Wen, and T. Wen, "Maximum correntropy high-order extended Kalman filter," *Chinese Journal of Electronics*, vol.31, no.1, pp.190–198, 2022.

[10] Y. S. Lu, X. Peng, D. Yang, *et al.*, "Model-agnostic meta-learning with optimal alternative scaling value and its application to industrial soft sensing," *IEEE Transactions on Industrial Informatics*, vol.17, no.12, pp.8003–8013, 2021.

[11] L. Yang, C. L. Wen, and T. Wen, "Multilevel fine fingerprint authentication method for key operating equipment identification in cyber-physical systems," *IEEE Transactions on Industrial Informatics*, vol.19, no.2, pp.1217–1226, 2023.

[12] H. T. Chen, B. Jiang, S. X. Ding, *et al.*, "Data-driven fault diagnosis for traction systems in high-speed trains: A survey, challenges, and perspectives," *IEEE Transactions on Intelligent Transportation Systems*, vol.23, no.3, pp.1700–1716, 2022.

[13] R. Pandit, D. Infield, and T. Dodwell, "Operational variables for improving industrial wind turbine yaw misalignment early fault detection capabilities using data-driven techniques," *IEEE Transactions on Instrumentation and Measurement*, vol.70, article no.2508108, 2021.

[14] Y. C. Zhang, Y. Xu, Z. Y. Dong, *et al.*, "Real-time assessment of fault-induced delayed voltage recovery: A probabilistic self-adaptive data-driven method," *IEEE Transactions on Smart Grid*, vol.10, no.3, pp.2485–2494, 2019.

[15] C. Cheng, J. H. Wang, Z. J. Zhou, *et al.*, "A BRB-based effective fault diagnosis model for high-speed trains running gear systems," *IEEE Transactions on Intelligent Transportation Systems*, vol.23, no.1, pp.110–121, 2022.

[16] W. Q. Bai, X. M. Yao, H. R. Dong, *et al.*, "Mixed $H\_/H_\infty$ fault detection filter design for the dynamics of high speed train," *Science China Information Sciences*, vol.60, no.4, article no.048201, 2017.

[17] W. Q. Bai, H. R. Dong, X. M. Yao, *et al.*, "Robust fault detection for the dynamics of high-speed train with multi-source finite frequency interference," *ISA Transactions*, vol.75, pp.76–87, 2018.

[18] H. T. Chen, B. Jiang, W. Chen, *et al.*, "Edge computing-aided framework of fault detection for traction control systems in high-speed trains," *IEEE Transactions on Vehicular Technology*, vol.69, no.2, pp.1309–1318, 2020.

[19] L. B. Cosme, W. M. Caminhas, M. F. S. V. D'Angelo, *et al.*, "A novel fault-prognostic approach based on interacting multiple model filters and fuzzy systems," *IEEE Transactions on Industrial Electronics*, vol.66, no.1, pp.519–528, 2019.

[20] J. I. Aizpurua, V. M. Catterson, I. F. Abdulhadi, *et al.*, "A model-based hybrid approach for circuit breaker prognostics encompassing dynamic reliability and uncertainty," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol.48, no.9, pp.1637–1648, 2018.

[21] D. H. Zhou, H. Q. Ji, X. He, *et al.*, "Fault detection and isolation of the brake cylinder system for electric multiple units," *IEEE Transactions on Control Systems Technology*, vol.26, no.5, pp.1744–1757, 2018.

[22] T. Wang, L. L. Y. Liang, S. K. Gurumurthy, *et al.*, "Model-based fault detection and isolation in DC microgrids using optimal observers," *IEEE Journal of Emerging and Selected Topics in Power Electronics*, vol.9, no.5, pp.5613–5630, 2021.

[23] F. Mwasilu and J. W. Jung, "Enhanced fault-tolerant control of interior PMSMs based on an adaptive EKF for EV traction applications," *IEEE Transactions on Power Electronics*, vol.31, no.8, pp.5746–5758, 2016.

[24] M. Corno, G. Panzani, and S. M. Savaresi, "Traction-control-oriented state estimation for motorcycles," *IEEE*

Transactions on Control Systems Technology, vol.21, no.6, pp.2400–2407, 2013.

[25] J. H. Ju, Z. M. Zhao, B. C. Shi, *et al.*, "Motor-oriented discrete state event-driven method for multitime-scale simulation of power traction systems," *IEEE Transactions on Transportation Electrification*, vol.7, no.3, pp.1652–1661, 2021.

[26] L. Zou, T. Wen, Z. D. Wang, *et al.*, "State estimation for communication-based train control systems with CSMA protocol," *IEEE Transactions on Intelligent Transportation Systems*, vol.20, no.3, pp.843–854, 2019.

[27] H. T. Chen, Z. Chai, B. Jiang, *et al.*, "Data-driven fault detection for dynamic systems with performance degradation: A unified transfer learning framework," *IEEE Transactions on Instrumentation and Measurement*, vol.70, article no.3504712, 2021.

[28] T. R. Tsai, W. Y. Sung, Y. L. Lio, *et al.*, "Optimal two-variable accelerated degradation test plan for gamma degradation processes," *IEEE Transactions on Reliability*, vol.65, no.1, pp.459–468, 2016.

[29] X. D. Chen, X. L. Sun, X. S. Si, *et al.*, "Remaining useful life prediction based on an adaptive inverse Gaussian degradation process with measurement errors," *IEEE Access*, vol.8, pp.3498–3510, 2020.

[30] W. W. Peng, Y. F. Li, Y. J. Yang, *et al.*, "Bivariate analysis of incomplete degradation observations based on inverse Gaussian processes and copulas," *IEEE Transactions on Reliability*, vol.65, no.2, pp.624–639, 2016.

[31] P. Wang, R. X. Gao, and W. A. Woyczynski, "Lévy process-based stochastic modeling for machine performance degradation prognosis," *IEEE Transactions on Industrial Electronics*, vol.68, no.12, pp.12760–12770, 2021.

[32] V. N. Kolokoltsov, "The Lévy-Khintchine type operators with variable Lipschitz continuous coefficients generate linear or nonlinear Markov processes and semigroups," *Probability Theory and Related Fields*, vol.151, no.1, pp.95–123, 2011.

[33] J. L. Bretagnolle and P. Ouwehand, "The Lévy-Itô decomposition theorem,"*arXiv preprint*, arXiv: 1506.06624, 2015.

[34] N. Mukherjee, A. Chattopadhyay, S. Chattopadhyay, *et al.*, "Discrete-wavelet-transform and stockwell-transform-based statistical parameters estimation for fault analysis in grid-connected wind power system," *IEEE Systems Journal*, vol.14, no.3, pp.4320–4328, 2020.

CHENG Chao received the M.E. and Ph.D. degrees from Jilin University, Changchun, China, in 2011 and 2014, respectively. He is currently an Associated Professor with the Changchun University of Technology, Changchun. He has been a Postdoctoral Fellow in process control engineering with the Department of Automation, Tsinghua University, Beijing, China, since 2018.

He has also been a Postdoctoral Fellow with the National Engineering Laboratory, CRRC Changchun Railway Vehicles Co., Ltd., China, since 2018. His research interests include dynamic system fault diagnosis and predictive maintenance, wireless sensor network, artificial intelligence, and data-driven method.
(Email: chengx415@163.com)

WANG Weijun received the B.S. and M.S. degrees from the School of Computer Science and Engineering, Changchun University of Technology, Changchun, China, in 2018 and 2021, respectively, where he is currently pursuing the Ph.D. degree with the School of Mathematics and Statistics. His main research interests include state estimation, distributed sensor networks, and fault diagnosis.

MENG Xiangxi received the Ph.D. degree in School of Computer Science and Engineering, Beihang University. He is currently a Researcher with the Institute of System Research, China Industrial Control Systems Cyber Emergency Response Team. His research interests include industrial control network, industrial internet platform, software test, and fault detection.

SHAO Haidong received the B.S. degree in electrical engineering and automation and the Ph.D. degree in vehicle operation engineering from School of Aeronautics, Northwestern Polytechnical University, Xi'an, China, in 2013 and 2018, respectively. He is currently an Associate Professor in College of Mechanical and Vehicle Engineering at Hunan University, Changsha, China. From 2019 to 2021, he was a Postdoctoral Fellow with the Division of Operation and Maintenance Engineering, Luleå University of Technology, Luleå, Sweden. His current research interests include fault diagnosis, intelligent prognosis, maintenance decision, and information fusion.

CHEN Hongtian received the B.S. and M.S. degrees in School of Electrical and Automation Engineering from Nanjing Normal University, China, in 2012 and 2015, respectively; and he received the Ph.D. degree in College of Automation Engineering from Nanjing University of Aeronautics and Astronautics, China, in 2019. He had ever been a Visiting Scholar at the Institute for Automatic Control and Complex Systems, University of Duisburg-Essen, Germany, in 2018. Now he is a Postdoctoral Fellow with the Department of Chemical and Materials Engineering, University of Alberta, Canada. His research interests include process monitoring and fault diagnosis, data mining and analytics, machine leaning, and quantum computation; and their applications in high-speed trains, new energy systems, and industrial processes.

Dr. Chen was a recipient of the Grand Prize of Innovation Award of Ministry of Industry and Information Technology of China in 2019, the Excellent Ph.D. Thesis Award of Jiangsu Province in 2020, and the Excellent Doctoral Dissertation Award from Chinese Association of Automation (CAA) in 2020. He currently serves as Associate Editors and Guest Editors of a number of scholarly journals such as *IEEE Transactions on Instrumentation and Measurement*, *IEEE Transactions on Neural Networks and Learning Systems* and *IEEE Transactions on Artificial Intelligence*. (Email: hongtian.chen@ieee.org)