

Article

Sieve Estimation-Based Data-Driven Fault Detection Method for Linear Time-Varying Systems

Caixin Fu ¹, Changhong Jiang ¹, Zhiwei Wan ¹ , Weijun Wang ^{2,*}  and Shenquan Wang ¹

¹ School of Electrical and Electronic Engineering, Changchun University of Technology, Changchun 130012, China; 1202301017@stu.ccut.edu.cn (C.F.); jiangchanghong@ccut.edu.cn (C.J.); 1202503002@stu.ccut.edu.cn (Z.W.); shenquanwang@126.com (S.W.)

² School of Computer Science and Engineering, Changchun University of Technology, Changchun 130012, China

* Correspondence: jinwei312@163.com

Abstract

The advancement of sensor technology has spurred the emergence of data-driven fault detection (FD) strategies. Particularly for many control systems with time-varying parameter characteristics, data-driven FD offers significant practical value and application value. This paper addresses the FD problem for linear time-varying (LTV) systems. Specifically, within the Sieve estimation framework, the stable kernel representation of LTV systems is designed using a novel data stacking method and B-spline basis functions. Subsequently, the coefficient function is constructed to provide a general solution to online FD tasks. Finally, this method is validated through experiments. The results demonstrate that the sieve estimation-based strategy exhibits excellent estimation and FD performance in both offline and online phases.

Keywords: data-driven designs; fault detection; state-space model; time-varying systems

1. Introduction

Process monitoring and fault detection (FD) of industrial control systems have received considerable attention in recent years [1–4]. In the current FD studies, observer-based schemes designed under linear time-invariant (LTI) system assumptions have emerged as a prevailing solution. However, control systems widely exhibit time-varying parameter characteristics, posing significant limitations and challenges to traditional observer-based schemes [5,6].

To address FD in linear time-varying (LTV) systems, significant methodological advancements have been achieved through observer-based schemes [7]. Foundational work by ref. [8] established an optimization paradigm for FD filter design in discrete-time LTV systems under bounded signals and stochastic noise, proposing an effective method for performance index design. Subsequent studies extended robustness to complex operational environments. Recent innovations focus on structural complexities [9]. Geometric methods leveraging unobservable subspaces were introduced in ref. [10] to generalize time-invariant concepts to LTV systems. This approach derived sufficient conditions for residual generator design using Lyapunov–Krasovskii functionals, with Wirtinger-based integral inequalities reducing conservatism. Despite progress, critical gaps persist in data-driven implementation of observer, particularly for LTV systems. Additionally, a number of studies have concentrated on developing robust FD approaches tailored for time-varying dynamic systems. In ref. [11], an innovative FD scheme by fusing the Kalman filter with the weighted



Academic Editor: Audrius Čereška

Received: 1 June 2026

Revised: 3 July 2026

Accepted: 6 July 2026

Published: 9 July 2026

Copyright: © 2026 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 \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

moving average algorithm is presented, which exhibits favorable performance for systems exposed to intense noise environments. In ref. [12], a robust filtering strategy is established via residual error suppression, thus elevating the precision of both state reconstruction and fault identification. Regarding distributed parameter systems, literature [13] constructs a model-driven sensor fault diagnosis architecture capable of accurately locating fault occurrence instants and quantifying fault amplitudes. Concerning unknown input issues encountered in LTV systems, an observer-based bank is constructed to implement concurrent FD and isolation tasks. As demonstrated in [14], this solution effectively mitigates the critical shortcomings of conventional model-based techniques, especially regarding computational efficiency and robustness. Despite these notable advancements in observer-based and robust FD methodologies for LTV systems, the data-driven implementation of such observers remains an under-explored area, leaving a critical gap in bridging theoretical developments with practical engineering applications for complex time-varying systems. To fill this gap, the present study focuses on advancing data-driven observer designs for LTV system FD, aiming to enhance the applicability and reliability of FD strategies in real-world operational scenarios.

Stable kernel representation (SKR) methods have been widely studied in data-driven process monitoring and FD as an effective solution [15,16]. The most significant benefit of this scheme is that the residual generator for FD can be constructed using process signals of systems without the need for cumbersome parameter identification of the model [17]. In [18], a component-level method is proposed for FD of suspension control box by determining the size of data stacking. To mitigate the impact of external disturbances on algorithm performance, a disturbance decoupling SKR is proposed for FD of traction systems in [19]. This constitutes an effective attempt at robust design. For nonlinear systems, a SKR-based scheme can also be regarded as an available solution [20]. The incrementally exponentially stable image and kernel representation is designed for the Youla parameterization problem [21]. Deep learning-based generalized SKR methods are applicable to FD in nonlinear systems [22]. In [23], deep neural networks are employed to learn the closed-loop SKR and stable image representation for nonlinear systems, which provides a theoretical foundation for constructing residual generators and implementing FD. In ref. [24], a data-driven SKR for closed-loop systems is derived via subspace identification approaches, which further allows the calculation of the corresponding data-driven closed-loop stability margin. In ref. [25], a SKR-based residual generation scheme is developed on the basis of the input–output model to realize the online detection of actuator faults. In addition, a series of recent studies has extended the application of the SKR framework to systems subject to degraded systems [26], distributed systems [27], and interconnected dynamical systems [28].

Despite these advancements, several critical problems remain to be addressed.

- (1) Most studies address the FD problem in LTV systems using model-based strategies in observer design, which escalates the difficulty of system parameter identification.
- (2) Owing to the infeasibility of achieving least-squares optimal estimation via QR/SVD factorization in LTV systems, traditional SKR-based methods cannot be directly implemented.
- (3) In the online application phase, existing data-driven observers struggle to discriminate between parameter variation dynamics and fault dynamics within residual signals.

To solve problems mentioned above, this study constructs a novel FD method for LTV systems. The main contributions and advantages of this study are summarized as follows.

- (1) A novel SKR-based FD framework for LTV systems is proposed. The time-varying kernels of the systems are identified via the data stacking technique and B-spline basis functions.

(2) The large sample properties of the proposed method and the convergence of parameter identification are analyzed.

(3) A coefficient function-based online FD strategy is designed via coefficients of basis function.

The remainder of this paper is organized as follows. The problem formulations are given in Section 2. The proposed FD model is introduced in Section 3. Verifications are shown in Section 4. Discussions are shown in Section 5. The conclusions of this study are summarized in Section 6.

2. Problem Formulation

Considering the LTV systems modeled by

$$\begin{cases} x(k+1) = A(k)x(k) + B(k)u(k) \\ y(k) = C(k)x(k) + D(k)u(k) \end{cases} \quad (1)$$

where $x(k) \in \mathcal{R}^{k_x}$, $u(k) \in \mathcal{R}^{k_u}$, $y(k) \in \mathcal{R}^{k_y}$ are the state, input, and output, respectively. $A(k)$, $B(k)$, $C(k)$, and $D(k)$ are dimensionally matched matrices.

For data vector $\psi(k) \in \mathcal{R}^{\xi}$ and window size s , the data stacking is given as follows.

$$\Psi_{k,s} = \begin{bmatrix} \psi_s(k) & \cdots & \psi_s(k+N-1) \end{bmatrix} = \begin{bmatrix} \Psi_k \\ \vdots \\ \Psi_{k+s} \end{bmatrix} \quad (2)$$

where

$$\begin{aligned} \psi_{k,s} &= \begin{bmatrix} \psi_s^T(k) & \cdots & \psi_s^T(k+N-1) \end{bmatrix}^T \in \mathcal{R}^{(s+1)\xi}, \\ \Psi_k &= \begin{bmatrix} \psi(k) & \cdots & \psi(k+N-1) \end{bmatrix} \in \mathcal{R}^{\xi \times N}. \end{aligned}$$

In linear time-invariant systems, the full-order observer, which is the most important framework for model-based methods mentioned in [1], is given as follows.

$$\begin{aligned} \hat{x}(k+1) &= A\hat{x}(k) + Bu(k) + L(r(k)) \\ \hat{y}(k) &= C\hat{x}(k) + Du(k) \end{aligned} \quad (3)$$

where $r(k)$ denotes the residual signal. Based on the steady Kalman filter, the state estimation can be represented as

$$\hat{x}(k+1) = A_k\hat{x}(k) + B_ku(k) + Ky(k) \quad (4)$$

where $A_k = A - KC$, $B_k = B - KD$. For the sufficiently large s , it has $A_K^s \approx 0$. Therefore, the I/O model has the following form.

$$Y_{k,s} = \Gamma_s L_p Z_p + H_{u,s} U_{k,s} + H_{w,s} W_{k,s} + V_{k,s} \quad (5)$$

where $Z_p = \begin{bmatrix} U_{k-s,s-1}^T & Y_{k-s,s-1}^T \end{bmatrix}^T$, $\Gamma_s L_p$, and $H_{u,s}$ are parameters to be estimated.

Given the kernels of systems, a residual generator can be constructed using estimated parameters in Equation (5) via QR or SVD [2]. The following exposition verifies the

optimality of this well-established scheme by exploring the relationship between matrix decomposition and least-squares. On account of Equation (5), it is obvious that

$$\underbrace{Y_{k,s}}_{\tilde{Y}} = \underbrace{\begin{bmatrix} \Gamma_s L_p & H_{u,s} \end{bmatrix}}_{\Theta} \underbrace{\begin{bmatrix} Z_p \\ U_{k,s} \end{bmatrix}}_{\tilde{U}} + \tilde{N} \quad (6)$$

where Θ represents the parameters need to estimated, and $\tilde{N}$ is consistent with the noises. For least-squares problem $\min \|r\|_2^2 = \|\tilde{Y}^{-1} - \tilde{U}^{-1}\Theta^{-1}\|_2^2$, it holds

$$\begin{aligned} \min \|r\|_2^2 &= \left\| \tilde{Y}^{-1} - Q \begin{bmatrix} R \\ 0 \end{bmatrix} \Theta^{-1} \right\|_2^2 = \left\| Q^T \tilde{Y}^{-1} - \begin{bmatrix} R \\ 0 \end{bmatrix} \Theta^{-1} \right\|_2^2 \\ &= \|c - R\Theta^{-1}\|_2^2 + \|e\|_2^2. \end{aligned} \quad (7)$$

At this time, when $R\Theta^{-1} = c$ is the optimal solution, the condition number is $\kappa_2(R) = \kappa_2(\tilde{U}^{-1})$. There is $\|r\|_2^2 = \|e\|_2^2$, and the error is bounded by $\kappa_2(\tilde{U}^{-1})$. If $\tilde{U}^{-1} = Q \times R$, it has

$$\Theta^{-1} = \tilde{U}\tilde{Y}^{-1} = (R^T Q^T Q R)^{-1} R^T Q^T \tilde{Y}^{-1} = R^{-1} Q^T \tilde{Y}^{-1}. \quad (8)$$

Similarly, the same conclusion can be drawn through the SVD. With regard to Equation (8), the decomposition cannot yield an optimal estimation in LTV systems because of the changes in regression coefficient. Consequently, it becomes a limitation to design the SKR-based residual generator.

3. Sieve Estimation-Based Parameter Identification and FD for LTV Systems

The data-driven FD framework designed by the subspace identification technique is illustrated in the Figure 1. The specific steps are given below.

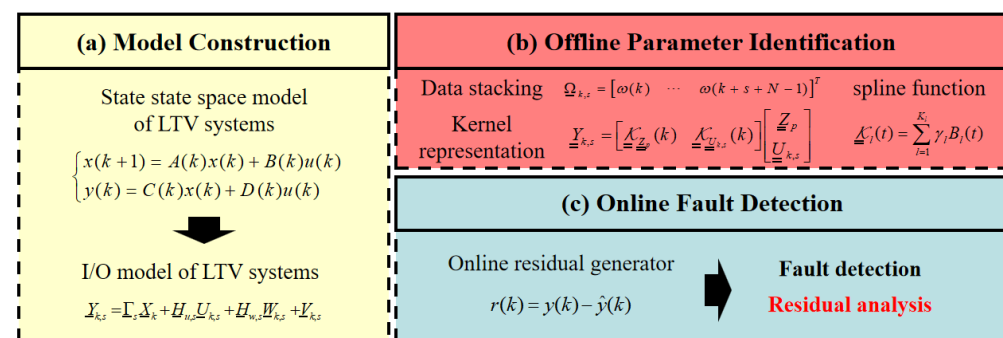


Figure 1. Schematic diagram of the proposed framework.

3.1. Model Construction and Parameter Identification for LTV Systems

For the purpose of FD design of LTV systems, the time-varying I/O model that analyzes the system dynamics is of great importance. In the first instance, the I/O model of LTV systems is

$$Y_{k,s} = \Gamma_s X_k + H_{u,s} U_{k,s} + H_{w,s} W_{k,s} + V_{k,s} \quad (9)$$

where the parameters have the following form.

$$\underline{\Gamma}_s = \begin{bmatrix} C(k) \\ C(k+1)A(k) \\ \vdots \\ C(k+s)A(k-1) \cdots A(k) \end{bmatrix}, \quad \underline{H}_{u,s} = \begin{bmatrix} D(k) & 0 \\ H' & \ddots & \ddots \\ \vdots & \ddots & \ddots \\ H'' & \cdots & H''' & D(k+s) \end{bmatrix}. \quad (10)$$

where $H' = C(k+1)B(k)$, $H'' = C(k+s)A(k+s-1) \cdots A(k+1)B(k)$, and $H''' = C(k+s)B(k+s-1)$.

According to the Kalman filter, it has

$$\hat{x}(k+1) = \underline{A}_K(k)\hat{x}(k) + \underline{B}_K(k)u(k) + \underline{K}(k)y(k), \quad (11)$$

where $\underline{A}_K(k) = \underline{A}(k) - \underline{K}(k)\underline{C}(k)$, $\underline{B}_K(k) = \underline{B}(k) - \underline{K}(k)\underline{D}(k)$. For a large integer s_l , it holds that

$$\hat{x}(k) = \underline{A}_K^{s_l} \hat{x}(k-s_l) + \sum_{i=1}^{s_l} \underline{A}_K^{i-1} \begin{bmatrix} \underline{B}_K & \underline{K} \end{bmatrix} \begin{bmatrix} u(k-i) \\ y(k-i) \end{bmatrix} \quad (12)$$

For control systems, the key required conditions are the observability and controllability of matrices A and C . Therefore, under this condition, it has

$$\underline{A}_K(k) \times \underline{A}_K(k+1) \times \cdots \times \underline{A}_K(k+s_l-1) \approx 0. \quad (13)$$

It turns out

$$\hat{x}(k) \approx \sum_{i=1}^{s_l} \underline{A}_K^{i-1} \begin{bmatrix} \underline{B}_K & \underline{K} \end{bmatrix} \begin{bmatrix} u(k-i) \\ y(k-i) \end{bmatrix}. \quad (14)$$

Therefore, the I/O model of LTV systems can be described as

$$\underline{Y}_{k,s} = \underline{\Gamma}_s \underline{L}_p \underline{Z}_p + \underline{H}_{u,s} \underline{U}_{k,s} + \underline{H}_{e,s} \underline{E}_{k,s}. \quad (15)$$

Then, for data vector $\underline{\omega}(k) \in \mathcal{R}^{\xi}$, the data stacking form designing for Sieve estimation is

$$\underline{\Omega}_{k,s} = \begin{bmatrix} \omega(k) & \cdots & \omega(k+s) & \cdots & \omega(k+s+N-1) \end{bmatrix}^T \in \mathcal{R}^{(s+1)\xi}. \quad (16)$$

Therefore, it implies

$$\begin{aligned} \underline{Y}_{k,s} &= \begin{bmatrix} y(k) & \cdots & y(k+s+N-1) \end{bmatrix}^T \\ \underline{U}_{k,s} &= \begin{bmatrix} u(k) & \cdots & u(k+s+N-1) \end{bmatrix}^T \\ \underline{Z}_p &= \begin{bmatrix} y(k-s-1) & \cdots & y(k+N-2) \\ u(k-s-1) & \cdots & u(k+N-2) \end{bmatrix}^T \end{aligned} \quad (17)$$

According to the SKR theory, the parameter identification problem is equivalent to solving

$$\underline{Y}_{k,s} = \underbrace{\begin{bmatrix} \underline{\Gamma}_s \underline{L}_p & \underline{H}_{u,s} \end{bmatrix}}_{\underline{K}(k)} \begin{bmatrix} \underline{Z}_p \\ \underline{U}_{k,s} \end{bmatrix}. \quad (18)$$

Then, it can be transformed into a regression problem of a varying-coefficient model given below.

$$\underline{Y}_{k,s} = \underbrace{\begin{bmatrix} \underline{K}_{\underline{Z}_p}(k) & \underline{K}_{\underline{U}_{k,s}}(k) \end{bmatrix}}_{\underline{K}(k)} \begin{bmatrix} \underline{Z}_p \\ \underline{U}_{k,s} \end{bmatrix} \quad (19)$$

To solve this problem, the Sieve estimation method is introduced in this section. Suppose that there are basis functions $B_l(t)$ and constants γ_l , after fixing the time point, it satisfies

$$\underline{K}_l(t) = \sum_{l=1}^{K_l} \gamma_l B_l(t), \quad (20)$$

where l represents the order parameter of basis function. By means of the asymptotic theory as $n \rightarrow \infty$ and the approximation property of B-spline bases, the parameter estimates computed at discrete points can converge to the true value in the continuous function space.

Given the framework outlined above, the asymptotic property of kernel estimation for LTV systems becomes amenable to study under the assumptions specified below.

Assumption 1. For random $\underline{K}_0 \in \mathcal{C}^{p+1}([0,1])$, it exists $\underline{K}_{n0}(t) = \sum_{j=1}^{K(n)} \theta_{0j} B_j(t)$ in Sieve space, satisfying

$$\|\underline{K}_{n0} - \underline{K}_0\|_\infty = O(K(n)^{-(p+1)}). \quad (21)$$

Assumption 2. For $n \rightarrow \infty$, the number of basis functions satisfies

$$K(n) \rightarrow \infty, \frac{K(n)}{n} \rightarrow 0. \quad (22)$$

Theorem 1. Under the conditions of Assumptions 1 and 2, the estimator satisfies

$$\|\hat{\underline{K}}_n - \underline{K}_0\|_{L^2} \xrightarrow{P} 0, \quad n \rightarrow \infty. \quad (23)$$

Proof. Decomposition of estimation error via triangle inequality is as follows

$$\|\hat{\underline{K}}_n - \underline{K}_0\|_{L^2} \leq \|\hat{\underline{K}}_n - \underline{K}_{n0}\|_{L^2} + \|\underline{K}_{n0} - \underline{K}_0\|_{L^2}. \quad (24)$$

According to the B-spline approximation theorem, there exists a $\underline{K}_{n0}$ such that

$$\|\underline{K}_{n0} - \underline{K}_0\|_\infty \leq C_1 h^{p+1} = C_1 K(n)^{-(p+1)}, \quad (25)$$

where $h = \frac{1}{m(n)} \approx K(n)^{-1}$. It follows that

$$\|\underline{K}_{n0} - \underline{K}_0\|_{L^2} \leq \|\underline{K}_{n0} - \underline{K}_0\|_\infty = O(K(n)^{-(p+1)}). \quad (26)$$

Given the design matrix, it can be derived that:

$$\mathbb{E} \left[\|\hat{\underline{K}}_n - \underline{K}_{n0}\|_{L^2}^2 \right] = O\left(\frac{K(n)}{n}\right). \quad (27)$$

Thus, when the optimal number of basis functions is selected, it follows that

$$\|\hat{\underline{K}}_n - \underline{K}_0\|_{L^2} \xrightarrow{P} 0, \quad n \rightarrow \infty. \quad (28)$$

□

Theorem 2. Under the conditions of Theorem 1, let the optimal number of basis functions be $K(n) = \frac{1}{2^{p+3}}$. Then

$$\|\hat{\underline{\mathcal{K}}}_n - \underline{\mathcal{K}}_0\|_{L^2} = O(n^{-\frac{p+1}{2p+3}}). \quad (29)$$

Proof. Decomposition the errors:

$$\mathbb{E} \left[\|\hat{\underline{\mathcal{K}}}_n - \underline{\mathcal{K}}_0\|_{L^2}^2 \right] \leq C_1 K(n)^{-1(p+1)} + C_2 \frac{K(n)}{n} \quad (30)$$

where $C_1 K(n)^{-1(p+1)}$ and $C_2 \frac{K(n)}{n}$ represent the bias and variances, respectively.

By differentiating the basis functions and substituting them into the total error, the following can be obtained

$$\mathbb{E} \left[\|\hat{\underline{\mathcal{K}}}_n - \underline{\mathcal{K}}_0\|_{L^2}^2 \right] \leq C_1 n^{\frac{-2(p+1)}{2p+3}} + C_2 n^{\frac{-2(p+1)}{2p+3}} = O\left(n^{\frac{-2(p+1)}{2p+3}}\right). \quad (31)$$

It follows that

$$\|\hat{\underline{\mathcal{K}}}_n - \underline{\mathcal{K}}_0\|_{L^2} = O(n^{-\frac{p+1}{2p+3}}). \quad (32)$$

□

3.2. FD in Online Phase

This section develops a reliable online FD strategy for LTV systems with the assistance of basis function weights obtained in the offline phase. Specifically, the non-parametric estimation method employed in the offline phase cannot discriminate fault dynamics in residual during the online phase, as it is impossible to decouple the fault f from y . The primary objective is therefore to identify a set of fixed parameters in the basis function space, enabling the acquisition of knowledge about the fault-free system using abundant offline data.

Theorem 3. Given the I/O model (5) and Sieve estimation scheme (20), the residual generator in the online phase can be obtained by

$$r(k) = y(k) - \left[\sum_{l_0=1}^{K_{l_0}} \gamma_{l_0} B_{l_0}(k), \sum_{l_1=1}^{K_{l_1}} \gamma_{l_1} B_{l_1}(k), \sum_{l_2=1}^{K_{l_2}} \gamma_{l_2} B_{l_2}(k), \sum_{l_3=1}^{K_{l_3}} \gamma_{l_3} B_{l_3}(k) \right] \begin{bmatrix} 1 \\ y(k-s-1) \\ u(k-s-1) \\ u(k) \end{bmatrix}, \quad (33)$$

where γ_{l_0} , γ_{l_1} , γ_{l_2} , and γ_{l_3} represent the offline projection weights of system dynamics in the spline basis function space.

The proof of the theorem is based on the relation $r(k) = y(k) - \hat{y}(k) = y(k) - \underline{\mathcal{K}}_1(k)y(k-s-1) - \underline{\mathcal{K}}_2(k)u(k-s-1) - \underline{\mathcal{K}}_3(k)u(k) - \underline{\mathcal{K}}_0(k)$.

Following the data-driven method, the online implementation of FD is given in Algorithm 1.

Algorithm 1. Online implementation of FD

- Step 1. Stack the input signals u and output signals y by Equation (17) and construct the I/O model of LTV systems by Equation (9).
 - Step 2. Identify the SKR of LTV systems by Equation (20), and save the offline coefficient γ_l .
 - Step 3. Calculate the online residual $r(k)$ by Equation (33) and setting the threshold according to the statistical characteristics of offline residuals.
 - Step 4. Repeat Step 3 until the system completes the FD tasks.
-

4. Case Studies

4.1. Simulation Examples

To verify the feasibility of the Sieve estimation-based scheme, different numerical examples are considered in this section. Throughout all simulation and industrial experiments in this work, cubic B-spline basis functions with fixed spline order $p = 3$ are utilized. The number of basis functions $K(n)$ is adaptively determined according to the total sample size n following the optimal scaling rule in Theorem 2.

Example 1. Consider that the time-varying parameters in Equation (1) are as follows [29]:

$$\begin{aligned} A(k) &= \begin{bmatrix} 0.4 \cos(k) & 0 \\ 0.5e^{-k} & 0.6 \end{bmatrix}, \quad B(k) = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \\ C(k) &= \begin{bmatrix} 1 & 0.1 + \frac{k}{k+1} \\ 0 & 1 \end{bmatrix}, \quad D(k) = \begin{bmatrix} 0 \\ 0 \end{bmatrix}. \end{aligned} \quad (34)$$

Four different scenarios are considered as: (1) Case 1: $f(k) = 0$; (2) Case 2: $f(k) = 0.001 * k + 0.005$; (3) Case 3: $f(k) = 0.5$; (4) Case 4: Multiplicative fault with parameter 1.02. Figure 2 illustrates the performance of estimation and FD in both offline and online phases. Figure 3 verifies the effectiveness of the proposed method by different fault cases. It illustrates the variations in residual distributions with respect to different cases during the online phase.

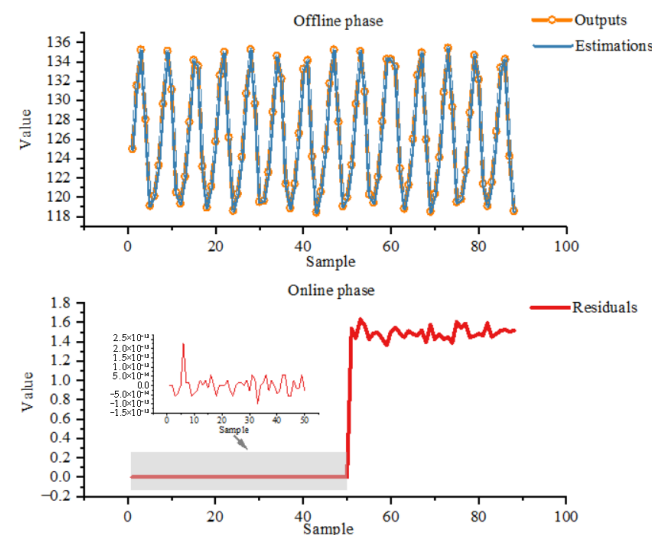


Figure 2. Results of estimation and FD in Example 1.

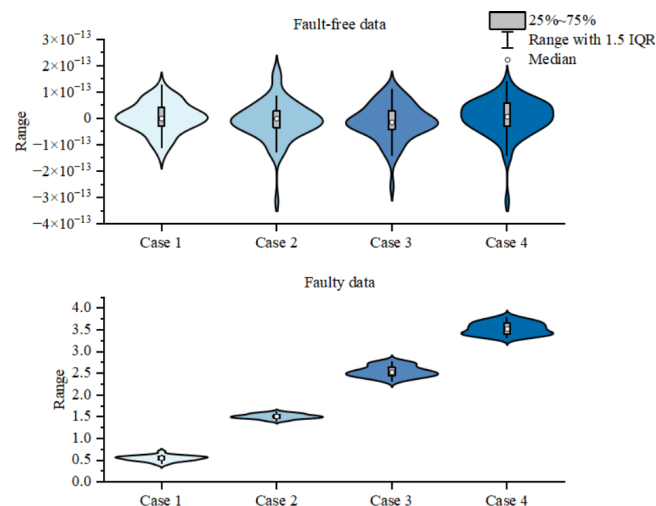


Figure 3. Residual variation of online phase in Example 1.

Example 2. Consider the time-varying parameters in Equation (1) as follows [30]:

$$\begin{aligned} x(k) &= \begin{bmatrix} g_0 \sin(\varphi(k)) \\ g_0 \cos(\varphi(k)) \end{bmatrix}, \quad A(k) = \begin{bmatrix} 1 & hz(k) \\ -hz(k) & 1 \end{bmatrix}, \\ B(k) &= 0, \quad y(k) = \begin{bmatrix} y_1(k) \\ y_2(k) \end{bmatrix}, \quad C(k) = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}. \end{aligned} \quad (35)$$

where φ and g_0 are unknown Toolface and gravitational acceleration, respectively. h corresponds to the sampling interval and z to the measurements. The experiment examines four categories of scenarios: (1) Case 1: $f(k) = 0$; (2) Case 2: $f(k) = 0.2$; (3) Case 3: $f(k) = 0.2(1 + 0.5\sin(k/100))$; (4) Case 4: $f(k) = 0.4(2 + \sin(k/100))$. Figure 4 also presents the estimation and detection performance in both offline and online phases, wherein it can be observed that the residual indices exhibit clear changes upon fault occurrence. Figure 5 provides a comparison under different cases, demonstrating that this scheme achieves satisfactory accuracy.

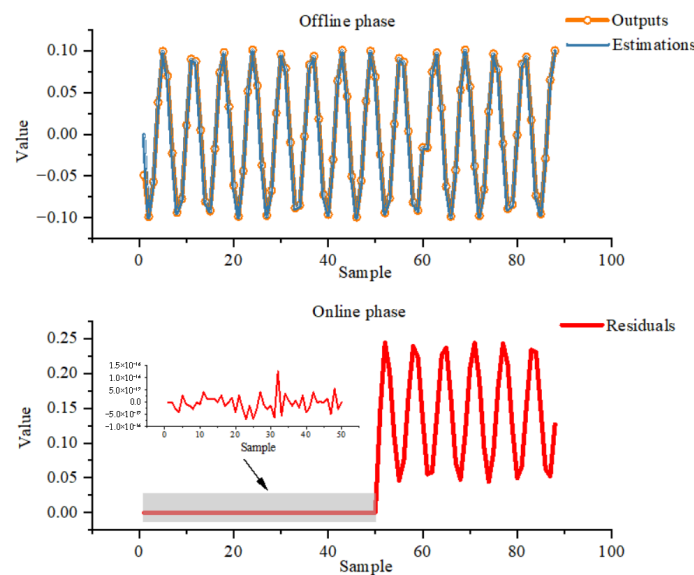


Figure 4. Results of estimation and FD.

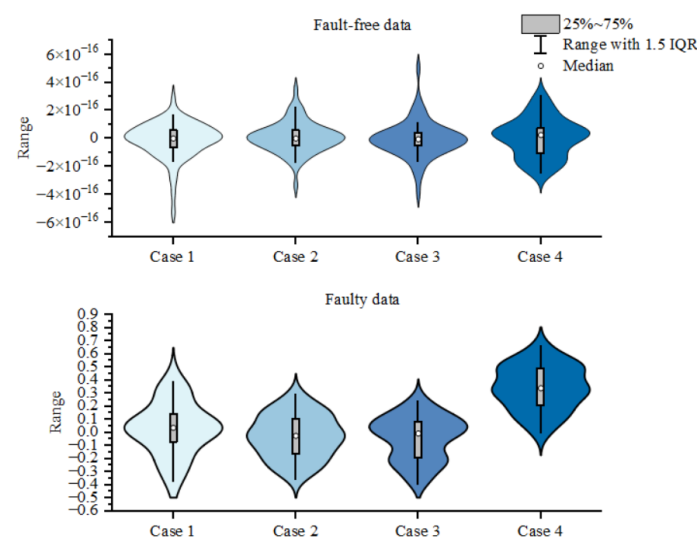


Figure 5. Residual variation in online phase.

In addition to the results shown above, two critical remarks are given as follows. (1) The accuracy of the Sieve estimation-based method is related to the basis functions. It is necessary to choose a suitable basis function to obtain better results. (2) The sieve estimation yields numerical solutions for the estimator, similar to neural networks. The main distinction between them lies in the

offline phase, where neural networks preserve parameters through training while sieve estimation retains the coefficient γ_l of the basis functions.

4.2. Traction System Experiment

A laboratory-based high-voltage motor test bench is employed in this section to conduct the experiment. As illustrated in Figure 6, the traction control system employed in this study is composed of a high-voltage control cabinet, a permanent magnet synchronous motor, a data acquisition card, and a host computer for algorithm implementation. The experimental setup is powered by a 220 V input voltage, with a rated power of 10 kW for the power drive unit, and the power switches are rated at 50 A/1200 V. The sampling time of this study is 1 ms. In the course of case analysis, four distinct time-varying models are taken into account, with the parameters in each model denoting the amplification coefficients of the sequence at various phases.

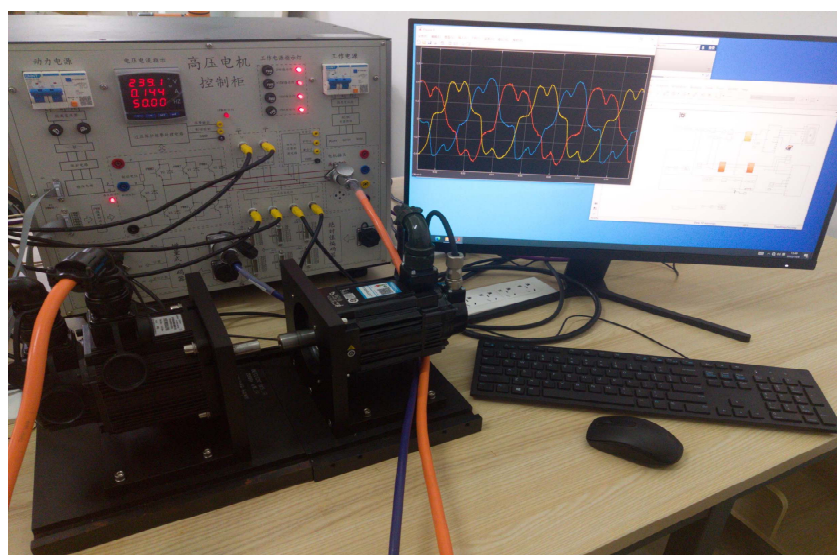
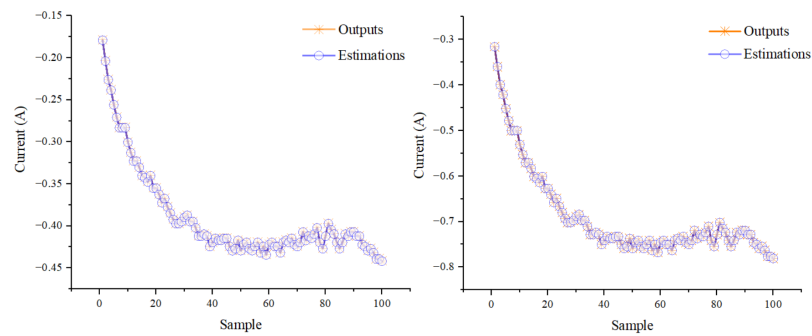


Figure 6. High-voltage motor test bench.

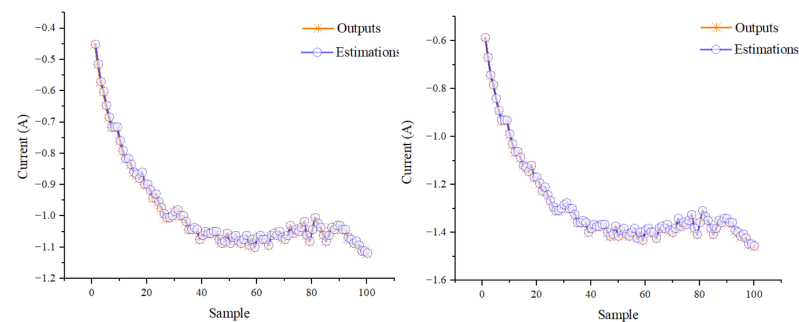
Figures 7 and 8 show the estimation and FD results under four different scenarios of the high-voltage motor experiment. The experimental results indicate that the proposed method is suitable for LTV systems with different time-varying parameters, adapting to the effect of process parameter changes on algorithm performance. Table 1 presents the FD results under different sample sizes and different stacking length parameters. Four scenarios of fault are considered as sensor drift faults, offset faults, stuck faults, and compound faults, respectively. The sample sizes are selected as 1000 and 3000, respectively, and the stacking lengths are 50 and 100. The results include the MSE change results before and after the occurrence of faults and the calculation time consumption results.

Table 1. FD results under different experimental settings.

Scenarios	Samples	Stacking Size	MSE ₁	MSE ₂	Time
I	1000	50	3.82×10^{-15}	1.74×10^2	0.25 s
	3000	100	1.77×10^{-18}	1.61×10^2	0.26 s
II	1000	50	3.83×10^{-15}	1.96×10^2	0.27 s
	3000	100	1.79×10^{-18}	1.41×10^2	0.26 s
III	1000	50	3.83×10^{-15}	1.86×10^2	0.26 s
	3000	100	1.77×10^{-18}	1.49×10^2	0.26 s
IV	1000	50	3.83×10^{-15}	1.84×10^2	0.28 s
	3000	100	1.78×10^{-18}	1.50×10^2	0.25 s

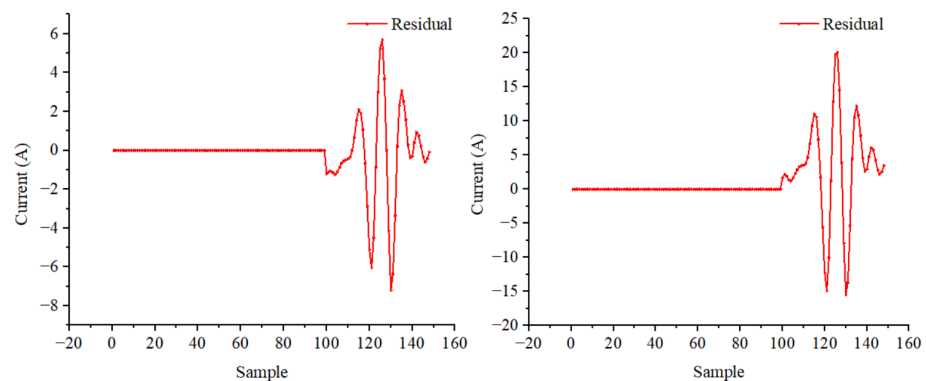


(a) Result of estimation in scenario I (b) Result of estimation in scenario II



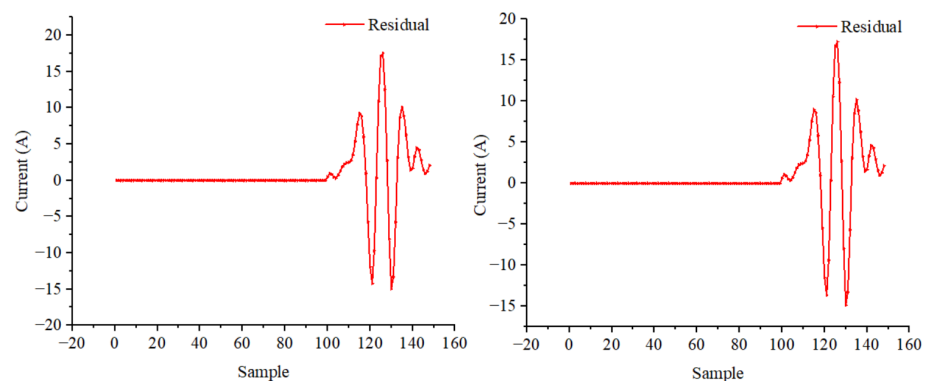
(c) Result of estimation in scenario III (d) Result of estimation in scenario IV

Figure 7. Results of estimations for four scenarios.



(a) Result of FD in scenario I

(b) Result of FD in scenario II



(c) Result of FD in scenario III

(d) Result of FD in scenario IV

Figure 8. Results of FD for four scenarios.

4.3. Comparisons

To verify the superiority of the proposed method, the TE process data [31], which is widely used in the FD field, is adopted in the comparative experiments shown in Table 2. The chemical processes described by the TE process, including separation, circulation, and other reactions and operations, cover data of normal operating states and fault states. The input and output data selected in the experiment are $u1$: Stream A; $u2$: Streams A and C; $y1$: Stream A outlet; $y2$: Streams A and C outlet. The comparative methods considered in the comparative experiments include: LSTM, PF, VAE, GRU, and CNN. The selected comparative methods cover various types of commonly used methods for prediction and FD. LSTM and GRU are good at capturing long-short term dependencies, PF is well-established in state tracking of dynamic systems, VAE performs prominently in feature extraction and fitting of noisy time-series data, and CNN is suitable for time-series scenarios with local fluctuations. The ultra-low normal MSE (10^{-24}) of our algorithm directly brings superior fault detection metrics: nearly 100% FDR, negligible MDR, and far lower FAR than baselines. The significant gap in MSE is not caused by inconsistent evaluation standards or data leakage, but the theoretical unbiased fitting characteristic of the sieve estimator, which separates normal and faulty residuals thoroughly and optimizes three core detection indicators simultaneously. To quantitatively characterize the performance degradation of the model under unseen unmodeled disturbances, multiple validation tests are further carried out in this study, and the averaged MSE is finally obtained as 0.055. It can be observed that unknown disturbances will significantly degrade the algorithm performance. In addition, the proposed method achieves an extremely tiny fault-free MSE and a much lower faulty-state MSE, creating a clear gap between normal and faulty residuals for sensitive fault detection. It consumes far less computation time than all comparative deep learning and filtering methods due to its limited interpretable B-spline coefficients without heavy iterative training. Moreover, the presented approach maintains stable superior accuracy and slight runtime growth as sample number and stacking window increase, showing stronger adaptability to varying industrial data conditions.

Table 2. Comparison results.

Methods	Variables	Samples	Stacking Size	MSE ₁	MSE ₂	Time
LSTM	$u1, y1$	1000	50	2.67×10^{-3}	9.55×10^{-3}	16.11 s
			100	2.65×10^{-3}	9.43×10^{-3}	34.34 s
		3000	50	2.67×10^{-3}	1.49×10^{-1}	64.59 s
			100	2.68×10^{-3}	1.52×10^{-1}	133.98 s
PF	$u1, y1$	1000	50	7.92×10^{-3}	1.30×10^{-2}	16.45 s
			100	5.71×10^{-3}	1.25×10^{-2}	29.73 s
		3000	50	9.92×10^{-3}	1.55×10^{-1}	62.54 s
			100	9.36×10^{-3}	1.59×10^{-1}	119.03 s
VAE	$u1, y1$	1000	50	2.72×10^{-3}	1.32×10^{-2}	0.78 s
			100	2.74×10^{-3}	1.74×10^{-2}	0.82 s
		3000	50	2.95×10^{-3}	1.98×10^{-1}	1.06 s
			100	2.77×10^{-3}	3.03×10^{-1}	1.09 s
GRU	$u1, y1$	1000	50	2.58×10^{-3}	4.49×10^{-3}	41.62 s
			100	2.56×10^{-3}	4.34×10^{-3}	78.73 s
		3000	50	2.65×10^{-3}	1.09×10^{-1}	108.34 s
			100	2.62×10^{-3}	7.97×10^{-2}	198.60 s
CNN	$u1, y1$	1000	50	2.71×10^{-3}	1.03×10^{-2}	1.51 s
			100	2.69×10^{-3}	1.24×10^{-2}	2.58 s
		3000	50	2.72×10^{-3}	4.12×10^{-1}	4.46 s
			100	2.71×10^{-3}	5.73×10^{-1}	8.53 s
The proposed method	$u1, y1$	1000	50	1.83×10^{-24}	8.47×10^{-5}	0.40 s
			100	2.59×10^{-24}	6.31×10^{-5}	0.85 s
		3000	50	3.71×10^{-24}	8.10×10^{-5}	0.83 s
			100	4.11×10^{-24}	6.63×10^{-5}	1.53 s

5. Discussions

Beyond the foregoing consequences, several important remarks and open issues are provided below.

(1) In practical applications, it is necessary to select the basis function type and its parameters adaptively according to the time-varying characteristics of the system: for systems with slow time-varying parameters, low-order B-spline basis functions with a small number of terms can be used to balance computational efficiency and estimation accuracy; for systems with fast and non-smooth time-varying parameters, high-order B-spline basis functions or a combination of multiple basis functions are recommended to improve the approximation ability of time-varying kernels.

(2) The sieve estimation adopted in this study is a numerical approximation method for estimating time-varying functions, which is similar to deep learning methods in that it obtains numerical solutions of system dynamics from data, but there are essential differences in the parameter storage and learning mechanism. This characteristic makes the proposed method have two advantages over deep learning methods: first, the number of parameters to be learned is far less than that of deep learning models, which reduces the risk of overfitting and the demand for offline data; second, the basis function coefficients can be interpreted and adjusted according to the system's time-varying characteristics, which makes the FD system have better interpretability, a key requirement for industrial process monitoring systems.

(3) The proposed method can effectively suppress the Gaussian noise and small random disturbances embedded in the offline fault-free data through the sieve estimation's approximation and the statistical characteristics of the residual signal. However, the robust performance of this method is limited by the types and ranges of disturbances contained in the offline training data: for large external disturbances, abrupt system mode changes, or non-stationary disturbances that do not appear in the offline data, the approximation accuracy of the sieve estimator will decrease, and the residual signal will contain large disturbance components, which may lead to false alarms or missed detections of faults. In practical industrial applications, it is necessary to combine disturbance decoupling techniques with adaptive residual threshold setting methods to improve the robust performance of the method for unknown large disturbances.

6. Conclusions

This paper develops a data-driven fault detection strategy for linear time-varying systems combining sieve estimation and stable kernel representation. A novel data stacking framework and B-spline basis functions are designed to identify time-varying system kernels, and the large-sample convergence and error convergence rate of the estimator are theoretically proven. Relying on offline solved spline coefficients, an online residual generator is constructed to realize real-time fault judgment, and the corresponding complete detection algorithm is summarized. Numerical simulations, high-voltage motor testbed experiments, and TE process comparative tests fully verify the superiority of the proposed method in fitting accuracy, fault sensitivity, and computational efficiency over LSTM, GRU, PF, VAE, and CNN. Finally, the real-time implementation limitations of this approach are analyzed, and future research directions including robust optimization and adaptive data updating are pointed out.

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

Funding: This research was funded by the National Natural Science Foundation of China, grant number 62273058.

Data Availability Statement: The data presented in this study are available from the corresponding author upon reasonable request.

Conflicts of Interest: The authors declare no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

Abbreviations

The following abbreviations are used in this manuscript:

FD	Fault detection
I/O	Input–output
LTI	Linear time-invariant
LTV	Linear time-varying
MSE	Mean squared error
SKR	Stable kernel representation

References

1. Park, Y.J.; Fan, S.K.; Hsu, C.Y. A Review on Fault Detection and Process Diagnostics in Industrial Processes. *Processes* **2020**, *8*, 1123. [\[CrossRef\]](#)
2. Ding, S.X. *Data-Driven Design of Fault Diagnosis and Fault-Tolerant Control Systems*; Springer: London, UK, 2014.
3. Melo, A.; Câmara, M.M.; Pinto, J.C. Data-Driven Process Monitoring and Fault Diagnosis: A Comprehensive Survey. *Processes* **2024**, *12*, 251. [\[CrossRef\]](#)
4. Ji, C.; Sun, W. A Review on Data-Driven Process Monitoring Methods: Characterization and Mining of Industrial Data. *Processes* **2022**, *10*, 335. [\[CrossRef\]](#)
5. Zhang, Y.; Zhang, H. Fault Detection for Time-Varying Processes. *IEEE Trans. Control Syst. Technol.* **2014**, *22*, 1527–1535. [\[CrossRef\]](#)
6. Zhong, M.; Xue, T.; Ding, S.X. A Survey on Model-Based Fault Diagnosis for Linear Discrete Time-Varying Systems. *Neurocomputing* **2018**, *306*, 51–60. [\[CrossRef\]](#)
7. Qiao, X.; Luo, H.; Zhang, K.; Li, K.; Jiang, Y.; Huo, M. Subspace Frequency Estimation under Colored Noise with Application to Fault Diagnosis of Motor Rolling Bearings. *IEEE Trans. Ind. Inform.* **2024**, *20*, 11014–11023. [\[CrossRef\]](#)
8. Zhong, M.; Ding, S.X.; Ding, E. Optimal Fault Detection for Linear Discrete Time-Varying Systems. *Automatica* **2010**, *46*, 1395–1440. [\[CrossRef\]](#)
9. Li, Y.; Karimi, H.; Zhong, M.; Ding, S.X.; Liu, S. Fault Detection for Linear Discrete Time-Varying Systems with Multiplicative Noise: The Finite-Horizon Case. *IEEE Trans. Circuits Syst. I Regul. Pap.* **2018**, *65*, 3492–3505. [\[CrossRef\]](#)
10. Zhang, Z.; He, X. Fault Detection and Isolation for Linear Parameter-Varying Systems with Time-Delays: A Geometric Approach. *Sci. China Inf. Sci.* **2023**, *66*, 172202. [\[CrossRef\]](#)
11. Gao, M.; Huai, W.; Niu, Y.; Sheng, L.; Zhou, D. Event-triggered Fault Detection for Time-varying Stochastic Systems with Strong Noise. *IEEE Trans. Instrum. Meas.* **2025**, *74*, 3513611. [\[CrossRef\]](#)
12. Guo, S.; Luo, S.; Xiang, G.; Huang, D.; Song, J. Consensus Control and Sensor Fault Diagnosis of Heterogeneous Intelligent Connected Vehicles with Time-Varying Parameters and Random Noises under Zero-Trust Architecture. *IEEE Trans. Veh. Technol.* **2025**, *74*, 13713–13727. [\[CrossRef\]](#)
13. Wang, K.; Zhou, J.; Shen, W.; Zhou, Y.; Wei, P.; Mou, X. Robust Sensor Fault Detection and Estimation for Parabolic Distributed Parameter Systems. *IEEE Trans. Instrum. Meas.* **2025**, *74*, 3515710. [\[CrossRef\]](#)
14. Wang, T.; Liu, W.; Hao, Z. A Decentralized Model-Based Fault Detection and Isolation Scheme for MVDC Shipboard Power Systems. *IEEE Trans. Transp. Electr.* **2024**, *10*, 7804–7815. [\[CrossRef\]](#)
15. Li, K.; Luo, H.; Yang, C.; Yin, S. Subspace-Aided Closed-Loop System Identification with Application to DC Motor System. *IEEE Trans. Ind. Electron.* **2020**, *67*, 2304–2313. [\[CrossRef\]](#)
16. Li, L.; Ding, S.; Peng, X. Optimal Observer-Based Fault Detection and Estimation Approaches for T-S Fuzzy Systems. *IEEE Trans. Fuzzy Syst.* **2022**, *30*, 579–590. [\[CrossRef\]](#)
17. Xue, T.; Ding, S.X.; Zhong, M.; Zhou, D. An Integrated Design Scheme for SKR-Based Data-Driven Dynamic Fault Detection Systems. *IEEE Trans. Ind. Inform.* **2022**, *18*, 6828–6839. [\[CrossRef\]](#)
18. Wang, P.; Long, Z.; Xu, Y. Component-Level Fault Detection for Suspension System of Maglev Trains Based on Autocorrelation Length and Stable Kernel Representation. *IEEE Trans. Veh. Technol.* **2021**, *70*, 7594–7604. [\[CrossRef\]](#)
19. Cheng, C.; Wan, Z.; Xue, T.; Peng, Y.; Yin, T.; Chen, H. An Improved Data-Driven Scheme of Robust Fault Detection for Traction Drive Systems. *IEEE Trans. Syst. Man Cybern. Syst.* **2025**, *55*, 4684–4692. [\[CrossRef\]](#)
20. Han, H.; Han, H.; Zhao, D.; Gao, X.; Yang, Y. Takagi-Sugeno Fuzzy Realization of Stability Performance-Based Fault-Tolerant Control for Nonlinear Systems. *IEEE Trans. Transp. Electr.* **2022**, *30*, 4249–4261. [\[CrossRef\]](#)

21. Kawano, Y.; Schaft, A.; Scherpen, J. Youla-Kučera Parameterization in Contraction Framework. *IEEE Trans. Autom. Control* **2025**, *70*, 1667–1682. [[CrossRef](#)]
22. Chen, H.; Liu, Z.; Alippi, C.; Biao, H.; Liu, D. Explainable Intelligent Fault Diagnosis for Nonlinear Dynamic Systems: From Unsupervised to Supervised Learning. *IEEE Trans. Neural Netw. Learn. Syst.* **2024**, *35*, 6166–6179. [[CrossRef](#)] [[PubMed](#)]
23. Wu, S.; Zhu, Y.; Luo, H.; Wang, H.; Zhang, J.; Qiao, X.; Tian, J. A Fault Detection Approach for Nonlinear Systems Based on Deep Learning-Aided Kernel Representations. *IEEE Trans. Ind. Inform.* **2025**, *21*, 6284–6293. [[CrossRef](#)]
24. Qiao, X.; Luo, H.; He, F.; Wang, H.; Wu, S.; Zhang, J. A Data-Driven Quantitative Stability Monitoring Approach for Quadrotor Blade Damage. *IEEE Trans. Ind. Electron.* **2026**, *73*, 1437–1448. [[CrossRef](#)]
25. Huang, D.; Yan, W. Data-Driven Fault Detection for HVAC Control Systems in Pharmaceutical Manufacturing Workshops. *Processes* **2025**, *13*, 2015. [[CrossRef](#)]
26. Li, L.; Ding, S.; Luo, H.; Peng, K.; Yang, Y. Performance-Based Fault-Tolerant Control Approaches for Industrial Processes with Multiplicative Faults. *IEEE Trans. Ind. Inform.* **2020**, *16*, 4759–4768. [[CrossRef](#)]
27. Wang, W.; Xu, S.; Liu, R.; Yin, T.; Chen, H.; Cheng, C. A Joint Observer-Based Distributed Fault Detection Method for Traction Systems of High-Speed Trains. *IEEE Trans. Instrum. Meas.* **2025**, *74*, 3561310. [[CrossRef](#)]
28. Sun, W.; Zhou, Z.; Ma, F.; Wang, J.; Ji, C. Industrial Application of Data-Driven Process Monitoring with an Automatic Selection Strategy for Modeling Data. *Processes* **2023**, *11*, 402. [[CrossRef](#)]
29. Zhong, M.; Ding, S.X.; Han, Q.; Ding, Q. Parity Space-Based Fault Estimation for Linear Discrete Time-Varying Systems. *IEEE Trans. Autom. Control* **2010**, *55*, 1726–1731. [[CrossRef](#)]
30. Niu, Y.; Sheng, L.; Gao, M.; Zhou, D. Optimal Fault Detection for Stochastic Linear Time-Varying Systems by χ^2 Test. *IEEE Trans. Autom. Control* **2024**, *69*, 2294–2307. [[CrossRef](#)]
31. Zou, W.; Xia, Y.; Li, H. Fault Diagnosis of Tennessee-Eastman Process Using Orthogonal Incremental Extreme Learning Machine Based on Driving Amount. *IEEE Trans. Cybern.* **2018**, *48*, 3403–3410. [[CrossRef](#)] [[PubMed](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.