Journal of Dynamics, Monitoring and Diagnostics, 2025, 4, 239-249
https://doi.org/10.37965/jdmd.2025.780

ISTP
RESEARCH ARTICLE

Fault Detection in Liquid Rocket Engines Via Segmented Vector Autoregressive Modeling

Tianxiang Teng,[1] Zhirong Zhong,[2] Jimin Liu,[1] Zhongyi Zhang,[1] and Meng Ma[1]
[1]School of Mechanical Engineering, Xi'an Jiaotong University, Xi'an, China
[2]School of Future Technology, Xi'an Jiaotong University, Xi'an, China

(Received 09 April 2025; Revised 08 August 2025; Accepted 02 September 2025; Published online 09 September 2025)

Abstract: Ensuring the safe operation of liquid rocket engine (LRE) systems requires reliable fault diagnosis, yet the scarcity of real fault data limits deep learning applications despite their modeling strengths. We address this by developing an offline detection method based on piecewise stationary vector autoregressive modeling, employing a two-phase approach that first identifies candidate change points through block fused LASSO regularization and subsequently refines them using smoothly clipped absolute deviation regularization to leverage its asymptotic unbiasedness. Validated on a high-fidelity LRE simulation dataset (26 sensors, 2000 time points) with injected faults including turbopump efficiency degradation, hydrogen turbine leakage, and valve failures across 48 scenarios, our method achieves 100% precision (±50-sample tolerance) in fault timing detection without requiring training data, demonstrating superior performance to conventional autoregressive moving average models while overcoming the data dependency of neural networks.

Keywords: LRE health monitoring; SCAD regularization; VAR model

I. INTRODUCTION

The liquid rocket engine (LRE), often referred to as the "heart" of a spacecraft, is a critical component that ensures the flight and safe launch of the rocket. However, LREs are required to operate for extended periods under harsh conditions such as high temperatures, high pressures, strong corrosion, and high energy release, making them prone to failures within the spacecraft system. If an LRE fails during ignition or flight, it can lead to severe accidents such as explosions in a very short time frame (50 ms), causing not only significant economic losses but also endangering the lives of astronauts. Therefore, research on fault diagnosis technology for LREs is particularly important. The detection of abrupt changes in multiple time series is a typical problem in the field of fault diagnosis for LREs.

Recent years, LRE fault detection approaches are divided into three types by some scholars [1]: signal processing-based approaches, model-driven approaches, and data-driven approaches. Data-driven fault detection methods involve the direct processing of sensor-measured operational data from LRE. Alternatively, data-driven techniques—such as correlation functions or autoregressive moving average (ARMA [2]) models—may extract signal features (e.g., variance, frequency) to diagnose potential engine failures. Deng *et al*. [3] developed an ARMA model-based fault detection method for the main-stage fault diagnosis of high-thrust hydrogen–oxygen staged-combustion cycle engines, and successfully validated the reliability of the approach through hardware-in-the-loop simulation. Xue *et al*. [4] developed a real-time fault simulation system based on ARMA models for reusable LREs, conducting simulation tests on typical fault scenarios and validating the algorithm's applicability through hardware-in-the-loop testing. However, since ARMA processes each signal independently, it is prone to overlooking system-level fault characteristics.

The VAR model was proposed by Christopher Sims in 1980 as an innovation to traditional simultaneous equation models [5]. VAR model is widely applied in fields such as economics [6][7][8] and engineering [9]. The VAR model can be seen as a multivariate extension of the univariate AR model. By capturing the dynamic relationships between multiple time series variables, the VAR model can jointly predict the future values of these variables.

However, high-dimensional VAR models suffer from the issues of having a large number of parameters and being computationally challenging to solve. Richard *et al*. [10] proposed a two-stage method to fit sparse VAR (sVAR) models where many AR coefficients are zero. In the first stage, nonzero AR coefficients are selected based on estimates of partial spectral coherence (PSC) and the use of BIC. PSC can be used to quantify the conditional relationships between marginal series in multivariate processes. A second refinement stage is then applied to further reduce the number of parameters. To address the issue of high-dimensional VAR models requiring the estimation of a large number of parameters and potential inference problems, Monica Billio *et al*. [11] proposed a new Bayesian nonparametric (BNP) Lasso prior (BNP-Lasso) for high-dimensional VAR models, which can improve estimation efficiency and prediction accuracy. In 2022, Aramayis *et al*. [12] improved the sparse VAR model (msVAR) by using time series graphical lasso (TSGlasso) for sparse inverse spectral density matrix estimation, avoiding direct high-dimensional matrix inversion, and introducing false discovery rate (FDR) controlled multiple hypothesis testing methods in the model refinement stage, thereby improving the two-stage sparse VAR framework proposed by Davis *et al*.

Despite these advancements, existing methods still face several limitations in the context of LRE fault diagnosis. Signal processing and ARMA-based approaches often overlook system-level dependencies. Traditional VAR

Corresponding author: Meng Ma (e-mail: Meng_Ma@mail.xjtu.edu.cn)

Algorithm 1: FISTA.

Input: Objective function components (smooth part f, non-smooth part g), Lipschitz constant L for $f(x)$, initial solution x_0, maximum iterations K. Function to compute gradient of smooth part ∇f, The tolerance value. ε.

Output: Optimized solution x^*.

1. **Initialize Parameters:** Set $y_1 = x_0$, $t_1 = 1$. Choose step size $t_k = \frac{1}{L}$
2. **For each iteration** k **from** 1 **to** K:
   a. Gradient step:
      I. Compute gradient at y_k:
         $z_k = y_k - \nabla f(y_k)$
   b. Proximal Operator (Shrinkage):
      I. Apply proximal operator for g:
         $x_k = prox_{t_k g}(z_k)$
   c. Momentum Update:
      I. Update step size:
         $t_{k+1} = (1 + sqrt(1 + 4t^2))/2$
      II. Update extrapolation point:
         $y_{k+1} = x_k + (t_k - 1)(x_k - x_{k-1})/t_{k+1}$
3. **Termination:**
   If $\|x_k - x_{k-1}\| < \varepsilon$
      Break
4. **End For**
5. **Return** x^*

methods, while able to capture multivariate dynamics, encounter computational difficulties in high dimensions and strictly require global stationarity—a condition rarely satisfied in real LRE operations, where signals often exhibit piecewise stationarity across different stages (e.g., start-up ignition, steady combustion, and shutdown depressurization). Moreover, most existing VAR-based fault detection methods rely on LASSO regularization, which enforces sparsity but introduces bias by excessively shrinking large coefficients, thereby limiting detection accuracy.

To address these issues, this study introduces a segmented VAR framework that incorporates the piecewise stationarity assumption and improves detection performance by replacing LASSO with SCAD regularization. The SCAD penalty alleviates the bias problem of LASSO, ensuring asymptotic unbiasedness and more accurate identification of significant variables while maintaining sparsity. In this way, our method specifically addresses two major shortcomings of existing approaches: (1) the inability of ARMA or traditional VAR methods to handle nonstationary LRE fault data and (2) the limitations of LASSO-regularized VAR in precisely identifying fault points under sparse conditions.

This paper focuses on an offline fault diagnosis method for LRE operational data based on the data generation mechanism of the VAR model. First, we construct a numerical simulation dataset containing various potential fault types during the LRE startup process. Then, we improve the VAR-based change-point detection method proposed in [1] by replacing the LASSO regularization in the second step with the SCAD regularization method. The improved detection method, employing SCAD regularization, exhibits asymptotic unbiasedness and the Oracle property, enabling more accurate identification of significant variables while maintaining sparsity and avoiding excessive shrinkage of large coefficients. Finally, we apply this method to the fault diagnosis dataset, achieving favorable detection performance. The contributions of this study are summarized as follows:

(1) We propose a novel VAR-based fault point detection algorithm incorporating SCAD regularization.

(2) Through simulations of the LRE system, we generated a synthetic dataset covering both normal and faulty states.

(3) We conducted fault point detection experiments on a simulated LRE dataset using the proposed fault detection method, achieving 100% precision. In addition, the algorithm was validated on real-world LRE fault data, further demonstrating its effectiveness.

The remainder of this paper is organized as follows: Section II introduces the theoretical background of the research and elaborates on the proposed method in detail. Sections III and IV verify the effectiveness and superiority of the proposed model based on simulation experiments and hot-fire testing experiments. Section V summarizes the main content of this paper.

II. METHODOLOGY

A. VAR MODEL

Suppose we have a piecewise stationary time series dataset. These data contain $n + 1$ time points, with m_0 change points denoted as $0 = t_0 < t_1 < \cdots < t_{m_0} < t_{m_0+1} = n$. Then, for any stationary segment where $t_{j-1} < t < t_j$. the VAR model can be expressed as

$$X_t = B'_j X_{t-1} + \varepsilon_t \quad and \quad B_j = L^* + S_j^* \tag{1}$$

where X_t is the p dimensional vector of sensor observed time series at t, B_j is the $p \times p$ transition matrix for the $j - th$ segment that reflects time-varying effects of historical values. Further, each transition matrix is assumed to be a superposition of a stable L^* low-rank component and a time-varying S_j^* sparse component. Finally, we assume that the $p-$dimensional noise process is normally distributed, i.e., $\varepsilon \sim N(0, \sum_\varepsilon)$. Building upon this foundation, we may postulate that the number of nonzero elements in $j - th$ sparse component S_j^* is $\|S_j^*\|_0 = s$ with $s << p^2$ and that low-rank component L^* has rank l with $l << p$. The low-rank component L^* encodes the static cross-autocorrelation structure across all p time series, whereas S_j^* captures dynamic cross-sectional dependencies.

B. FAST ITERATIVE SHRINKAGE-THRESHOLDING ALGORITHM (FISTA)

Fast iterative shrinkage-thresholding algorithm (FISTA) is an accelerated optimization algorithm designed to solve large-scale linear inverse problems, particularly those involving sparse signal recovery and regularized regression. It is an enhanced version of the iterative shrinkage-thresholding algorithm (ISTA), achieving significantly faster convergence through Nesterov's momentum techniques.

FISTA solves convex optimization problems of the form:

$$\min_{x \in \mathbb{R}^n} F(x) = f(x) + g(x) \tag{2}$$

where $f(x)$ is a smooth convex function (e.g., least-squares loss $\frac{1}{2}\|Ax - b\|_2^2$), $g(x)$ is a nonsmooth convex regularizer (e.g., ℓ_1-norm $\lambda\|x\|_1$ for sparsity).

C. The Changing points Detection Procedure

In the changing point detection procedure, we propose mainly two steps: (A) solving a regularized regression problem using the BFL penalty to identify candidate change points and (B) filtering the obtained candidate points through SCAD regularization to compute a new information criterion.

1) STEP 1: IDENTIFY CANDIDATE POINTS. For a piecewise stationary time series dataset containing $n+1$ time points, we first partition it into blocks of size b_n, keeping all model parameters fixed within each block—meaning each block's endpoints serve as candidate change points. It is important to note that to accurately identify true change points, b_n cannot be set too large. Therefore, the selection of b_n follows these criteria:

$$k_n = \lfloor 2\sqrt{n} \rfloor \tag{3}$$

Thus, the time series length for each block is

$$b_n = \left\lfloor \frac{n}{2\sqrt{n}} \right\rfloor \tag{4}$$

For $j-th$ block, we define the following variables:

$$\begin{cases} \mathbf{X}_{r_j} = [X_{r_{j-1}}, \ldots, X_{r_j-1}] \\ \mathbf{Y}_{r_j} = [X_{r_{j-1}+1}, \ldots, X_{r_j}] \\ \boldsymbol{\varepsilon}_{r_j} = [\varepsilon_{r_{j-1}+1}, \ldots, \varepsilon_{r_j}] \end{cases} \tag{5}$$

For global variables, we define:

$$\begin{cases} \mathcal{X} = [\mathbf{X}_{r_1}, \ldots, \mathbf{X}_{r_{k_n+1}}]' \in \mathbb{R}^{n \times p} \\ \mathcal{Y} = [\mathbf{Y}_{r_1}, \ldots, \mathbf{Y}_{r_{k_n+1}}]' \in \mathbb{R}^{n \times p} \\ \boldsymbol{\varepsilon} = [\boldsymbol{\varepsilon}_{r_1}, \ldots, \boldsymbol{\varepsilon}_{r_{k_n}}]' \in \mathbb{R}^{n \times p} \end{cases} \tag{6}$$

Furthermore, we define the variable Z as

$$Z = \begin{bmatrix} \mathbf{X}'_{r_1} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{X}'_{r_2} & \mathbf{X}'_{r_2} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{X}'_{r_{k_n+1}} & \mathbf{X}'_{r_{k_n+1}} & \cdots & \mathbf{X}'_{r_{k_n+1}} \end{bmatrix} \in \mathbb{R}^{n \times p k_n} \tag{7}$$

Based on the above definition, we can formulate the following linear regression problem

$$\mathcal{Y} = \mathcal{X}L^* + Z\Theta + \varepsilon \tag{8}$$

Wherein $\Theta = [\theta'_1, \ldots, \theta'_{k_n}]' \in \mathbb{R}^{pk_n \times p}$, where $\theta'_1 = S_1^*$, then for the subsequent ones

$$\theta_i = \begin{cases} S_{j+1}^* - S_j^* & i = t_j \\ 0 & \text{otherwise} \end{cases} \tag{9}$$

Based on Equation, the model can be expressed as:

$$\begin{bmatrix} \mathbf{Y}_{r_1} \\ \cdots \\ \mathbf{Y}_{r_{k_n+1}} \end{bmatrix} = \begin{bmatrix} \mathbf{X}_{r_1} \\ \cdots \\ \mathbf{X}_{r_{k_n+1}} \end{bmatrix} \bullet L^* + \begin{bmatrix} \mathbf{X}_{r_1} \bullet S_1^* \\ \cdots \\ \mathbf{X}_{r_{k_n+1}} \bullet S_{r_{k_n+1}}^* \end{bmatrix} + \varepsilon \tag{10}$$

The model coefficients Θ and L can be estimated via a composite LASSO-regularized approach, as expressed in the linear regression formulation (8)

$$\begin{aligned} \left\langle \Theta, \hat{L} \right\rangle = & \, arg\,min_{\Theta, L \in \Omega} \frac{1}{n} \|\mathcal{Y} - \mathcal{X}L - Z\Theta\|_2^2 \\ & + \lambda_{1,n}\|L\|_* + \lambda_{2,n}\|\Theta\|_1 + \lambda_{3,n} \sum_{l=1}^{k_n} \left\| \sum_{j=1}^{l} \theta_j \right\|_1 \end{aligned} \tag{11}$$

We employ the FISTA (**Algorithm 1**) for efficient solution.

2) STEP 2: SCREEN CANDIDATE POINTS. In the previous step, we have obtained the candidate change point set A_n, and now we need a new step to filter the existing change point set. Specifically, we already have m change points, namely: $1 = s_0 < s_1 < \cdots < s_m < s_{m_0+1} = n$. Based on this, similar definitions as in Step 1 can be made:

$$\begin{cases} \mathbf{X}_{s_j} = [X_{s_{j-1}}, \ldots, X_{s_j-1}] \\ \mathbf{Y}_{s_j} = [X_{s_{j-1}+1}, \ldots, X_{s_j}] \\ \boldsymbol{\varepsilon}_{s_j} = [\varepsilon_{s_{j-1}+1}, \ldots, \varepsilon_{s_j}] \end{cases}$$

$$\begin{cases} \mathcal{X} = [\mathbf{X}_{s_1}, \ldots, \mathbf{X}_{s_m}]' \in \mathbb{R}^{n \times p} \\ \mathcal{Y} = [\mathbf{Y}_{s_1}, \ldots, \mathbf{Y}_{s_m}]' \in \mathbb{R}^{n \times p} \\ \varepsilon = [\boldsymbol{\varepsilon}_{s_1}, \ldots, \boldsymbol{\varepsilon}_{s_m}]' \in \mathbb{R}^{n \times p} \end{cases} \tag{12}$$

Furthermore, we define the variable $Z_{s_1,\ldots,s_m}$ as:

$$Z_{s_1,\ldots,s_m} = \begin{bmatrix} \mathbf{X}'_{s_1} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{X}'_{s_2} & \mathbf{X}'_{s_2} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{X}'_{s_{m+1}} & \mathbf{X}'_{s_{m+1}} & \cdots & \mathbf{X}'_{s_{m+1}} \end{bmatrix} \tag{13}$$

And the corresponding coefficient matrix is given by $\Theta_{s_1,\ldots,s_m} = [\theta'_1, \ldots, \theta'_m]'$. Then, based on the variables we have defined, the following linear regression expression can be written:

$$\mathcal{Y} = \mathcal{X}L^* + Z_{s_1,\ldots,s_m}\Theta_{s_1,\ldots,s_m} + \varepsilon \tag{14}$$

We then obtain $\Theta_{s_1,\ldots,s_m}$ by solving the following optimization problem incorporating SCAD regularization:

$$\begin{aligned} (\hat{L}, \hat{\Theta}_{s_1,\ldots,s_m}) = \underset{L,\Theta_{s_1,\ldots,s_m}}{\arg\min} \Bigg\{ & \underbrace{\sum_{i=1}^{m+1} \frac{1}{s_i - s_{i-1}} \left\| \mathbf{Y}_{s_i} - \mathbf{X}_{s_i}(\theta_{(s_{i-1},s_i)} + L) \right\|_2^2}_{\text{Prediction Error (Weighted Least Squares)}} + \underbrace{\sum_{i=1}^{m+1} \sum_j P_{\lambda,a}(\theta_{(s_{i-1},s_i),j})}_{\text{Segmented Sparsity Penalty (SCAD)}} \\ & + \underbrace{\sum_k P_{\lambda_L,a_L}(L_k)}_{\text{Low-Rank Matrix Penalty (SCAD)}} + \underbrace{\sum_k P_{\lambda_L,a_L}(L_k)}_{\text{Low-Rank Matrix Penalty (SCAD)}} \Bigg\} \end{aligned} \tag{15}$$

In the equation, the term $P_{\lambda,a}(\bullet)$ represents the smoothly clipped absolute deviation (SCAD) regularization component. The SCAD penalty function is defined as

$$P_{\lambda,a}(x) = \begin{cases} \lambda|x| & \text{if } |x| \le \lambda, \\ \frac{(a\lambda|x| - \frac{x^2+\lambda^2}{2})}{a-1} & \text{if } \lambda < |x| \le a\lambda, \\ \frac{(a+1)\lambda^2}{2} & \text{if } |x| > a\lambda. \end{cases} \tag{16}$$

The total error can be expressed as

$$\varepsilon_{\text{total}}(L,\Theta_{s_1,\ldots,s_m}) = \sum_{i=1}^{m+1} \frac{1}{s_i - s_{i-1}} \left\| \mathbf{Y}_{s_i} - \mathbf{X}_{s_i}(\theta_{(s_{i-1},s_i)} + L) \right\|_2^2 + \sum_{i=1}^{m+1} \sum_{j} P_{\lambda,a}(\theta_{(s_{i-1},s_i),j}) + \sum_{k} P_{\lambda_L,a_L}(L_k) \tag{17}$$

Next, we perform iterative screening of the number of change points. Let the set of change points after the k-th iteration be denoted as $s^{(k)} = (s_1^{(k)}, \ldots, s_{m_k}^{(k)})$. Therefore, the error in the current iteration can be expressed as

$$\varepsilon(s^{(k)}) = \sum_{i=1}^{m_k+1} \frac{1}{s_i^{(k)} - s_{i-1}^{(k)}} \cdot \left\| \mathbf{Y}_{[s_{i-1}^{(k)}, s_i^{(k)})} - \mathbf{X}_{[s_{i-1}^{(k)}, s_i^{(k)})}(\theta_{(s_{i-1}^{(k)}, s_i^{(k)})} + L) \right\|_2^2 + \sum_{i=1}^{m_k+1} \sum_{j} P_{\lambda,a}(\theta_{(s_{i-1}^{(k)}, s_i^{(k)}),j}) + \sum_{\ell} P_{\lambda_L,a_L}(L_\ell) \tag{18}$$

We evaluate change points by sequentially testing the impact of removing each candidate point through an information criterion (IC) framework. First, the IC is defined as:

$$IC(s^{(k)}) = \varepsilon(s^{(k)}) + m_k \omega_n \tag{19}$$

In the equation, m_k represents the current number of change points (which dynamically changes during iterations), and$\omega_n$denotes the sample-size-dependent penalty weight. Based on the above equation, the change in information criterion (ΔIC_r) resulting from removing the r-th change point can be expressed as

$$\Delta IC_r = \varepsilon(s_{-r}^{(k)}) - \varepsilon(s^{(k)}) - \omega_n \tag{20}$$

A change point can be removed when ΔIC_r. By sequentially evaluating and removing each candidate point in the current change point set while computing ΔIC, we achieve effective change point screening.

III. SIMULATION EXPERIMENTS

The experimental validation utilizes three types of fault simulation data from a LRE. The model's effectiveness is verified by evaluating the accuracy in detecting the timing of simulated fault occurrences.

This experiment utilizes a LRE simulation dataset, where numerical simulation was achieved by establishing component-level simulation modules using Amesim software.

The simulation model outputs synthetic readings from 26 sensors, with each sample consists of 2000 data points collected at a 1 kHz sampling rate during steady-state operation, representing a 2-second duration. Faults were injected at three specific time points: the 500th time point (0.5 seconds after simulation start), the 1000th time point (1.0 seconds after simulation start), and the 1500th time point (1.5 seconds after simulation start). The fault magnitudes were set to 0.8, 0.85, 0.9, and 0.95 of nominal values, respectively, generating a total of 48 distinct fault scenarios.

D. SIMULATION SYSTEM CONSTRUCTION

This study simulates an LRE system to generate operational data under both normal and fault conditions.

The LRE system, comprising turbopump assemblies, thrust chambers, and auxiliary subsystems (Fig. 1), was simulated under fault conditions by perturbing parameters in the nominal model. Four failure modes were implemented: turbopump efficiency degradation, hydrogen turbine leakage, valve actuation faults, and cooling jacket breaches.

The fault simulation of the LRE is based on a simulation model, employing a fault-injection approach to conduct a systematic and dynamic comprehensive analysis of various potential failure modes in a high-thrust hydrogen–oxygen engine. Table I summarizes common fault modes in LREs. The investigation implements fault injection on three critical failure modes: turbopump efficiency degradation, hydrogen turbine module leakage, and valve actuation failure, representing common yet high-impact scenarios in LRE operation.

3) TURBINE EFFICIENCY DECREASE. During operation, turbomachinery components may experience various failure modes including rotor rubbing/jamming, shaft fracture, turbine blade detachment, pump blade fracture, and turbopump cavitation, all of which can lead to varying degrees of efficiency degradation in the turbine assembly. A detailed analysis of the failure mechanisms reveals that when rotor rubbing and partial turbine blade detachment occur during operation, the former increases the torque demand on the turbopump assembly while the latter enhances the flow resistance between propellant and blades. Both effects contribute to reduced turbine actuation capability, ultimately manifesting as decreased turbine assembly efficiency [15]. To simulate these fault conditions, an efficiency correction factor f is introduced to modify the operational efficiency of both the turbine and centrifugal pump. This approach models the rotational speed reduction caused by power loss and the consequent decrease in pump work output, achieving accurate fault simulation as demonstrated in Fault Mode 1 of Figure 1. The mathematical representation of this relationship is as follow:

$$P_{turbine} = dpQ\eta_{turbine}f = n_{turbine}T \tag{21}$$

where $P_{turbine}$ is the power, dp is the pressure difference across the turbine, Q is the volumetric flow rate, $\eta_{turbine}$ is the turbopump efficiency, f is the correction factor, $n_{turbine}$ is the common rotational speed of both the turbine and centrifugal pump, and T is the torque.

4) HYDROGEN TURBINE MODULE LEAKAGE. Hydrogen, as a fuel, is relatively prone to leakage due to its small molecular weight. Furthermore, hydrogen turbopumps operate at extremely high rotational speeds reaching tens of thousands of revolutions per minute. The coaxial design of these turbopumps, combined with higher pressure at the turbine end compared to the pump end, creates conditions conducive to hydrogen leakage into the pump and surrounding environment [16].

In this failure scenario, liquid hydrogen leaking directly into the pump and environment essentially introduces two additional flow paths to the engine system. To model this condition, a valve assembly with a maximum flow area A is added to each flow path. The opening size of these valves is controlled by external signals to simulate varying degrees of leakage severity. The mathematical representation of this leakage model is as follows:

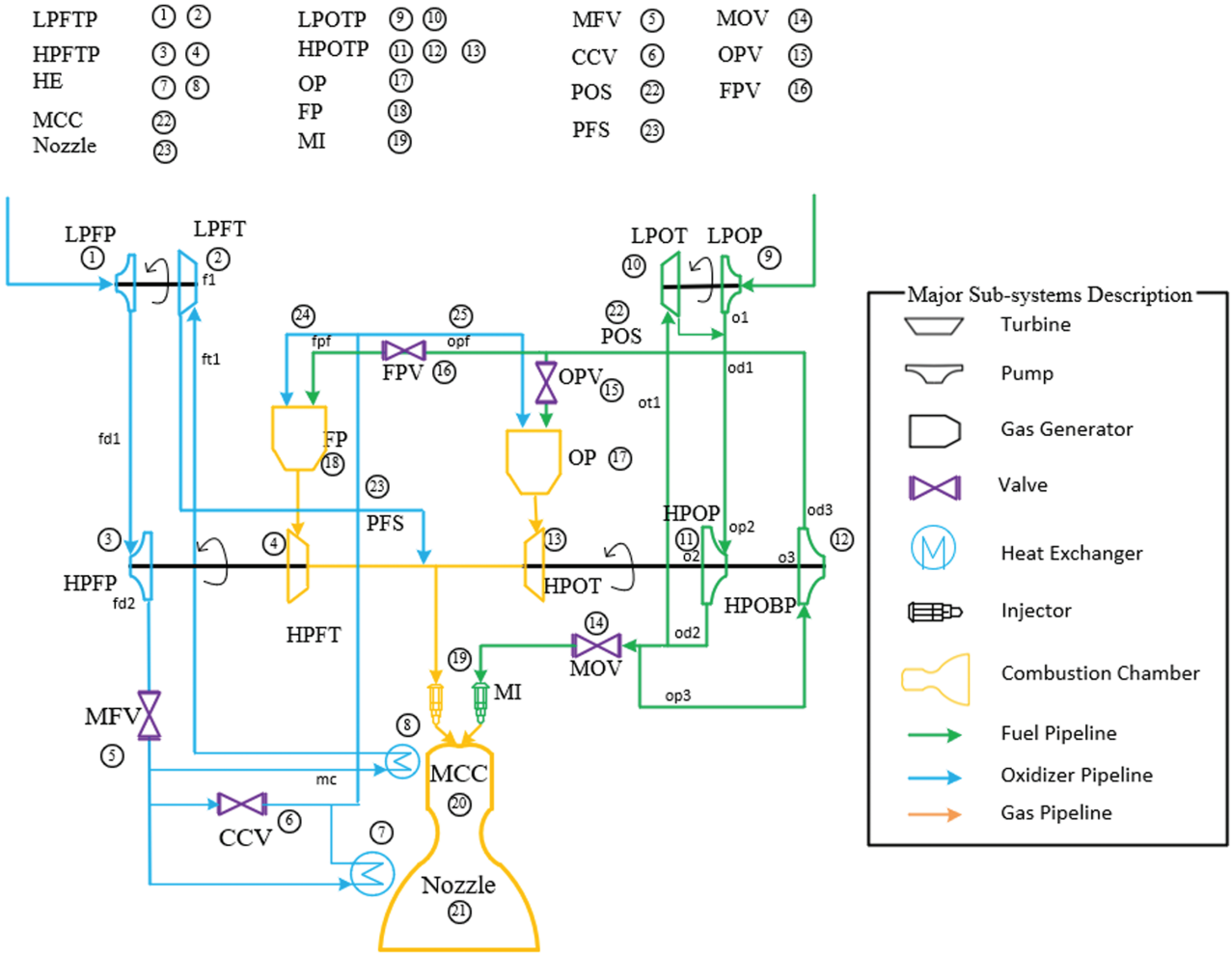

Fig. 1. Schematic of LRE and selected failure modes.

$$\dot{m}_3 = \dot{m}_1 + \dot{m}_2 \tag{22}$$

$$\dot{m}_2 = c_q A \tau \sqrt{2\rho\Delta p} \tag{23}$$

where $\dot{m}_3$ and $\dot{m}_1$ are the flow rate through the primary valve, $\dot{m}_2$ is the flow rate through the leakage valve, and A is the maximum flow area of the leakage passage.

5) VALVE OPENING FAILURE. Valve control constitutes a critical factor for normal engine startup. The simulation of valve failures—including failure to open, slow opening, and blockage—can be achieved by adjusting the timing and response speed of five main valves: the main oxidizer valve, main fuel valve, fuel preburner oxidizer valve, combustion chamber coolant valve, and oxidizer preburner oxidizer valve. The following equations primarily regulate flow through control functions to simulate these valve failure modes:

$$\dot{m} = c_q A \tau \sqrt{2\rho\Delta p} \tag{24}$$

where $\dot{m}$ is the flow rate through the valve, c_q is the flow coefficient, A is the maximum flow area, τ is the control function, ρ is the average density of the fluid flowing through the valve, and Δp is the pressure difference between the two ports of the valve.

E. DATA PREPROCESSING

The data preprocessing module consists of two main steps: data normalization and noise generation.

6) DATA NORMALIZATION. Since the variables calculated by the LRE's mathematical model exhibit significant differences in magnitude, normalization is essential to ensure consistency across different parameters. To address this, all simulation-derived data are uniformly scaled using the Min-Max normalization method, which transforms the values into a common range. This technique involves identifying the global minimum and maximum values of the dataset before applying the normalization formula:

$$X_{nomalized} = \frac{X - X_{\min}}{X_{\max} - X_{\min}} \tag{25}$$

where X is the original data value, $X_{\min}$ is the minimum value in the dataset, and $X_{\max}$ is the maximum value in the dataset. This transformation scales all features to a fixed range, ensuring consistent treatment of variables regardless of their original measurement scales.

The normalized simulation data (without noise injection) for all three types of faults are illustrated in the figure.

7) NOISE GENERATION. In the process of sensor measurements for LREs, environmental interference and

TABLE I. LRE failure modes

Components	Classification	Fault mode	Fault performance
Turbopump	Centrifugal pump	(1) Impeller damage (2) Bearing wear or damage (3) Pump cavitation	Pump efficiency decrease
	Pump efficiency decrease	(1) Blade detachment (2) Bearing wear or damage (3) Turbine blade erosion (4) Gas flow obstruction (5) Turbine inlet flow leakage	Turbine efficiency decrease Downstream stream flow rate decrease
Pipeline	Gas pipeline	(1) Pipeline blockage (2) Pipeline leakage	Increased flow resistance
	Liquid pipeline		Downstream flow rate decrease
Thrust Chamber	Combustion chamber	Combustion deterioration	Combustion efficiency decrease
	Gas generator	Combustion deterioration	
	Cooling jacket	Cooling jacket blockage	Increased flow resistance
		Cooling jacket leakage	Downstream flow rate decrease
	Nozzle	(1) Nozzle deformation (2) Large nozzle detachment	Nozzle efficiency decrease
Others	Regulating valve	Stuck during switching	Reduced flow area
	Cavitation tube	Cavitation tube blockage	Increased flow resistance
	Sonic nozzle	Sonic nozzle blockage	

measurement errors inevitably introduce random noise into the actual data. To accurately simulate this phenomenon, this study employs a Gaussian noise model to enhance the normalized data.

In this study, Gaussian noise was injected into the normalized signals to realistically simulate sensor uncertainty in LRE measurements. Specifically, three noise intensity levels were considered: 2.5%, 5%, and 7.5% of the signal amplitude. Among these, the 5% noise level closely reflects the signal-to-noise ratio observed in actual LRE sensor data, while the 2.5% and 7.5% levels were introduced as additional robustness tests to evaluate the stability of the proposed method under lower and higher noise conditions.

The mathematical expression of the Gaussian noise model is as follows:

$$X_{noisy} = X_{normalized} + \varepsilon, \quad \varepsilon \sim N(0,\sigma^2) \tag{26}$$

where $X_{normalized}$ is the normalized value within [0,1], ε is the Gaussian-distributed noise term, and σ determines the noise magnitude as a given percentage (2.5%, 5%, or 7.5%) of signal amplitude. This formulation allows us to evaluate robustness under different noise intensities that approximate varying levels of sensor uncertainty. By doing so, the additive noise model retains the primary signal trends while superimposing controlled random fluctuations. This design allows us to approximate sensor measurement disturbances under different noise levels without completely distorting the essential dynamics of the simulated engine data. The preprocessed simulation data are shown in Fig. 2.

F. IMPLEMENTATION DETAILS

8) EXPERIMENT DESIGN. The simulation model outputs synthetic readings from 26 sensors, with each sample consists of 2000 data points collected at a 1kHz sampling rate during steady-state operation, representing a 2-second duration. Faults were injected at three specific time points: the 500th time point (0.5 seconds after simulation start), 1000th time point (1.0 seconds after simulation start), and 1500th time point (1.5 seconds after simulation start). The fault magnitudes were set to 0.8, 0.85, 0.9, and 0.95 of nominal values, respectively, generating a total of 48 distinct fault scenarios. The running environment is described as follows: the CPU is a Core i5-13600KF @ 3.50GHz, the memory is 32GB, the GPU is a GTX 4070, and the programming language is R 4.4.0.

9) PARAMETER SELECTION. In our change-point detection method, several important tuning parameters are involved, including the regularization parameters $\lambda_{1,n}$, $\lambda_{2,n}$, $\lambda_{3,n}$ in Step 1, the penalty term ω_n in the information criterion, and the tolerance level *tol* used in the FISTA optimization. The following provides the approach for selecting them:

$\lambda_{1,n}$: In this study, we manually select $\lambda_{1,n}$ from the range $[\sqrt{\frac{\log p}{n}}, 10\sqrt{\frac{\log p}{n}}]$. This choice is motivated by the nonasymptotic analysis of [17]. Within this interval, we perform a grid search over a set of evenly spaced candidate values. The optimal $\lambda_{1,n}$ is then selected as the one that minimizes the validation error across the candidate grid.

$\lambda_{2,n}$: To select the regularization parameter $\lambda_{2,n}$, we adopt a block-based cross-validation approach. In the simulation study, 20% of the time series blocks are randomly selected as the validation set, with the blocks spaced evenly from a random initial point. The last time points of these selected blocks form the set T. We then estimate Θ for a range of candidate $\lambda_{2,n}$ values using only the data excluding T. These estimates are subsequently used to predict the time series at the time points in T. The value of $\lambda_{2,n}$ that minimizes the mean squared prediction error over T is selected as the optimal parameter via cross-validation.

$\lambda_{3,n}$: The regularization parameter $\lambda_{3,n}$, which controls the inter-block sparsity structure, is theoretically required to satisfy $\lambda_{3,n} = o((nd_n^*)^{-1})$as stated in Assumption H3 of [14], implying that its value should vanish as the sample size n increases. For simplicity, we set $\lambda_{3,n} = 0$ throughout this study.

ω_n: The penalty weight controlling the number of change points (the larger the value, the fewer the change

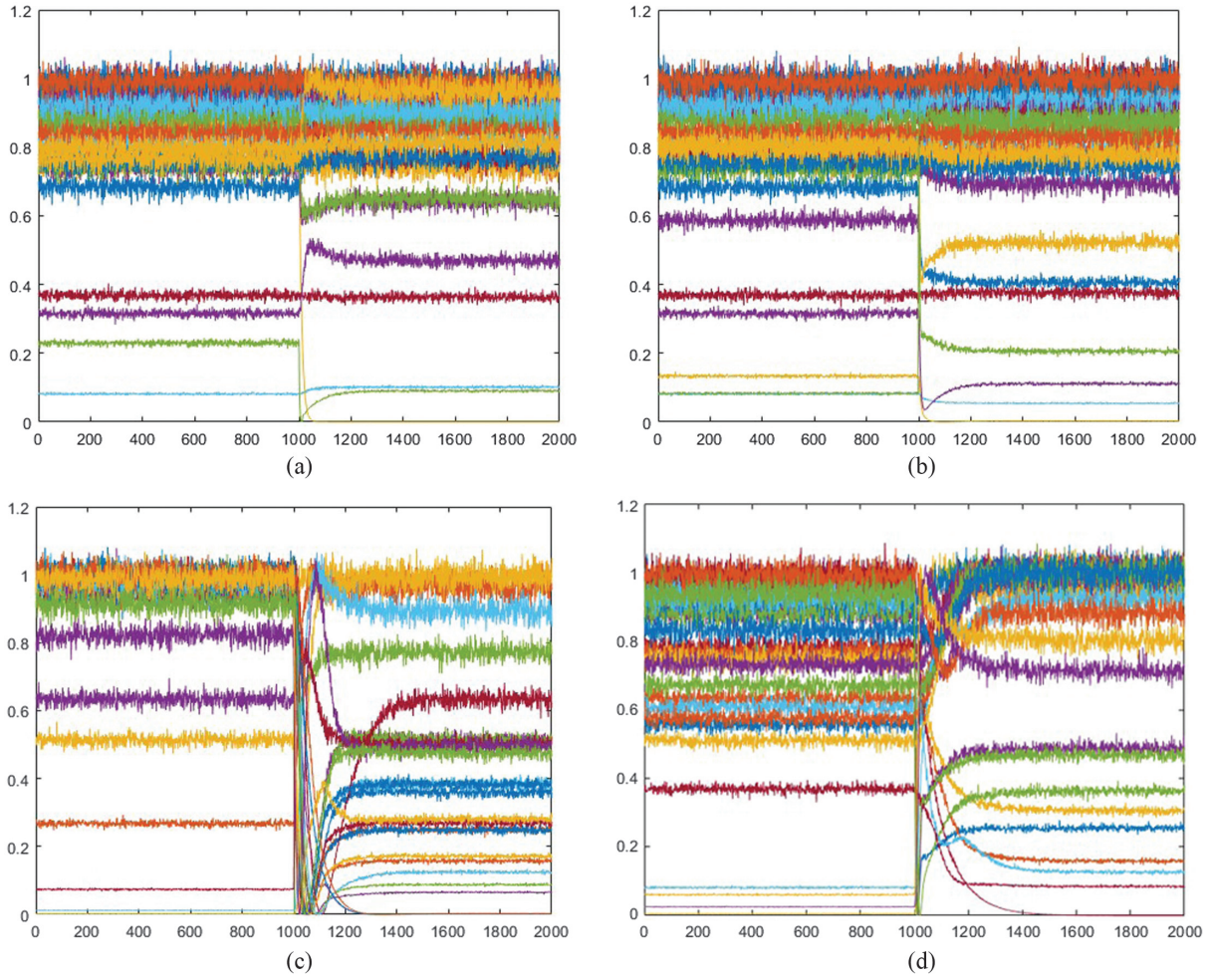

Fig. 2. The preprocessed simulation data for four types of faults. (a) Low-pressure fuel turbine efficiency degradation fault. (b) Abnormal oxygen main valve opening. (c) High-pressure oxygen turbine efficiency degradation fault. (d) High-pressure fuel turbine efficiency degradation fault.

points). In this study, it is manually selected from the range $[\frac{1}{30}\ln(T-1)\ln(p), \frac{1}{2}\ln(T-1)\ln(p)]$, which is selected based on the theoretical range suggested by Assumption H4 in [14]. Guided by this theoretical interval, we further perform a grid search within the range to determine the optimal ω_n.

tol: The convergence tolerance ϵ in Algorithm 1 is manually set to 10^{-3} here.

The regularization parameter in Step 2 is selected with reference to that in Step 1.

10) EVALUATION INDEX. The classification performance is evaluated based on the proportion of true positives (TP), which focuses solely on the ratio between TP and false positives (FP):

$$P = \frac{TP}{TP + FP} \tag{27}$$

TP is counted when any detected break point $\widehat{s}_j$ falls within $\pm p$ lags of a true change point s_j (accounting for VAR lag effects). Precision P measures the proportion of correctly identified breaks among all detections. The tolerance window is set to $p = 50$ timepoints, accommodating both the model's autoregressive structure.

G. RESULT AND DISCUSSION

For the 48 sets of fault scenario data generated by the LRE simulation model, we conducted fault point detection using three approaches: the baseline method from [14], our SCAD-regularized variant, and the LASSO-regularized variant as an additional reference. Experiments were performed under three Gaussian noise levels: 2.5%, 5%, and 7.5% of the signal amplitude. The experimental results are summarized in Table II. While all methods achieved 100% precision at the 2.5% noise level due to the clarity of fault signatures, under higher noise conditions (5% and 7.5%), the SCAD-based method demonstrated noticeably more robust detection performance compared to the LASSO-based approach. This agrees with the theoretical findings of [18], who showed that SCAD's nonconvex penalty avoids the over-shrinkage of large coefficients that is common with LASSO, thereby preserving essential structural signals and enhancing robustness. In addition, we also conducted comparative experiments with the model in [19] to further validate the effectiveness of our method.

Figure 3 illustrates the detection results of one of the fault datasets using the improved method proposed in this paper.

TABLE II. Detection results of liquid rocket engine fault simulation data using three models under different noise levels

Noise level	LASSO-VAR	SCAD-VAR (Proposed)	BSS-VAR
2.5%	Accuracy = 100%	Accuracy = 100%	Accuracy = 95.83%
5.0%	Accuracy = 100%	Accuracy = 100%	Accuracy = 87.5%
7.5%	Accuracy = 91.66%	Accuracy = 95.83%	Accuracy = 83.33%

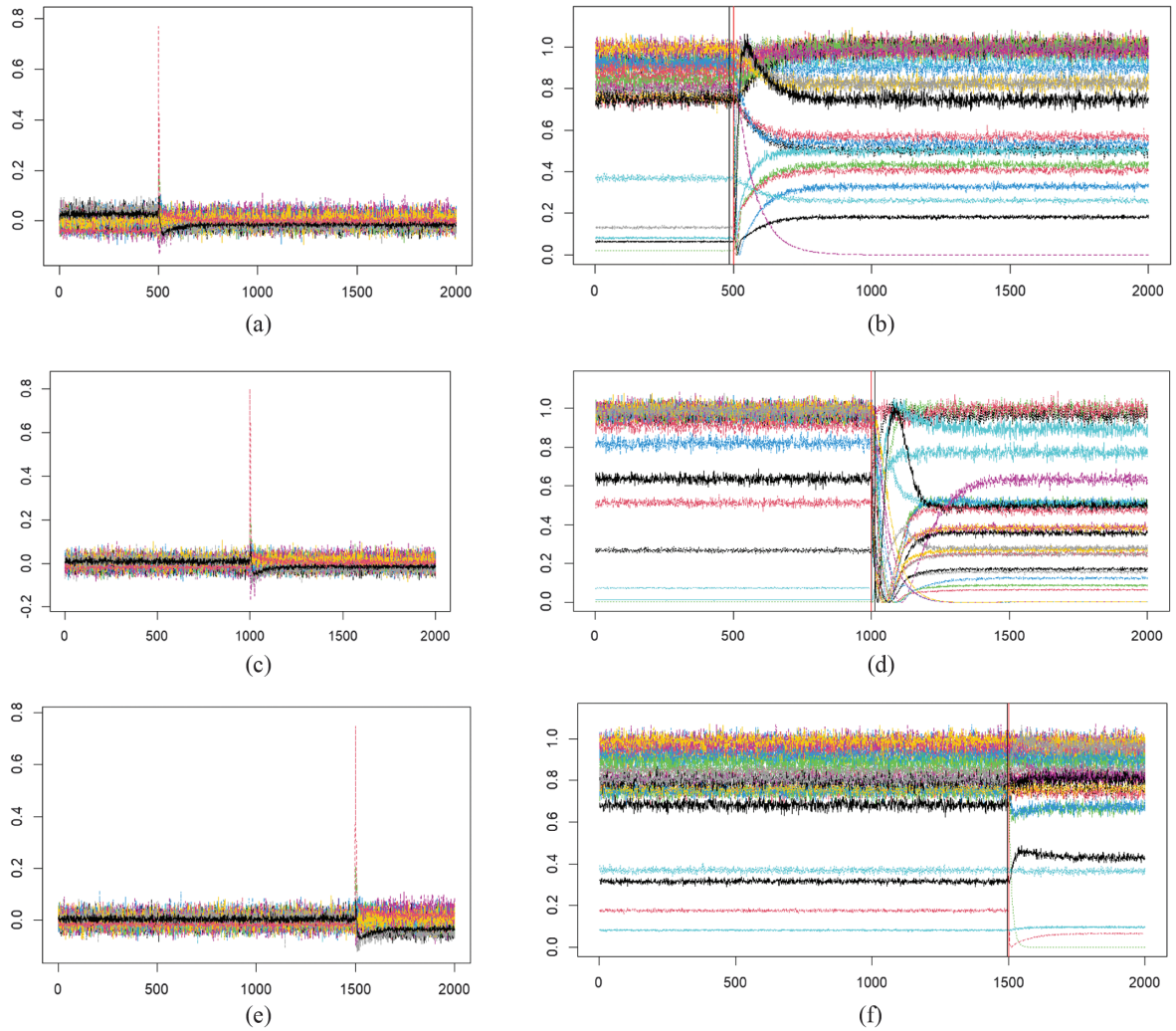

Fig. 3. Fault point detection results for 3 scenarios. (a)(c)(e) The calculation results of $(X_{t-1} - L^* X_t)$. (b)(d)(f) Results of fault point detection. The red vertical lines indicate the actual fault injection points at 500, 1000, and 1500, respectively. The black vertical lines denote the detected fault points at 484, 1012, and 1496, respectively.

Figure 3 shows the change-point detection results for three different fault scenarios:

1. At the 1000th data point, an abnormal fault in the oxygen main valve opening is injected with a fault factor of 0.9.
2. At the 500th data point, a fault in the high-pressure oxygen turbine efficiency is injected with a fault factor of 0.85.
3. At the 1500th data point, a fault in the low-pressure fuel turbine efficiency is injected with a fault factor of 0.8.

In (a)(c)(e), the calculation results of $(X_{t-1} - L^* X_t)$are presented, reflecting the extraction of the low-rank component L^* of transition matrix in step 1. It is evident that the parameter estimation of L^* in step 1 is highly accurate, with the residuals containing only Gaussian noise and a distinct pulse signal at the fault injection time point.

IV. HOT-FIRE TESTING EXPERIMENTS

To further verify the effectiveness of the proposed change-point detection method, this paper validates it using real-world data from a hot-fire test of a certain model of LRE. During the test, over 100 sensors were deployed on the engine to monitor its health status. A leakage fault was

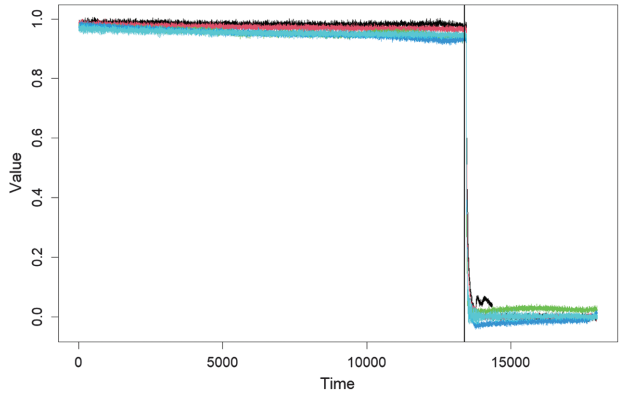

Fig. 4. Fault point detection result for hot-fire testing experiments.

detected during the experiment. However, unlike in simulation scenarios, the actual fault occurrence time in the hot-fire test is unknown. The change-point detection results of the proposed method applied to the hot-fire test time series data are shown in Fig. 4.

Considering that the original data contain confidential information, de-identification measures were applied. Specifically, the data were standardized, and noise was injected. Only processed features and fault labels are published to ensure compliance with data security regulations.

In contrast to simulated environments, the method also proves effective when applied to real-world fault scenarios in LRE testing.

V. CONCLUSION

This paper proposes an offline change-point detection method based on the VAR model and applies it to fault diagnosis in sLREs. Fault point detection experiments were conducted on an LRE system simulation dataset, achieving a precision of 100%, which demonstrates the effectiveness of the method in the context of LRE fault diagnosis. In addition, the proposed method was further validated on real-world fault data from hot-fire testing experiments, confirming its practical applicability and robustness. Compared to the change-point detection method in [14], the proposed approach introduces the SCAD regularization technique, leveraging its nonconvex penalty mechanism to reduce the shrinkage of large coefficients, thereby minimizing estimation bias and achieving a more accurate sparse solution. Furthermore, unlike artificial neural network methods, the proposed method does not require extensive training data and can localize fault points with relatively high precision.

However, the proposed method still has several limitations. First, although FISTA accelerates the optimization process, the block-wise segmentation strategy requires sequential processing of high-dimensional VAR models for each block. As the data scale increases —such as longer time series or more sensors—the number of blocks grows linearly, leading to significantly increased computational costs and challenges in meeting real-time requirements. Second, key parameters such as the regularization coefficients ($\lambda_{1,n}$, $\lambda_{2,n}$) and the convergence tolerance (ε) must be manually tuned, which relies on expert knowledge or trial and error. This parameter selection process is not only time consuming but also introduces subjectivity, potentially resulting in unstable outcomes. Finally, the method divides the time series into fixed-length blocks, implicitly assuming a uniform distribution of change points. In practice, however, failures in LREs often concentrate in specific periods (e.g., the ignition phase). Fixed block sizes lack the flexibility to adapt to such temporal variations, potentially causing insufficient resolution in change-point-dense regions and redundant computation in sparse ones.

Moreover, the proposed framework is built on the assumption of piecewise stationarity. This assumption is supported by the operational characteristics of LREs, where the working process can be divided into several distinct phases such as ignition, steady combustion, and shutdown depressurization. Each phase typically corresponds to different control logics, fuel flow adjustments, or structural responses, which lead to clear differences between transient and steady states. While transient signals exhibit strong nonstationary behaviors (e.g., rapid thrust rise, chamber pressure oscillation, or temperature jumps), steady phases usually demonstrate small fluctuations and relatively stable inter-variable relationships, thus presenting local stationarity suitable for VAR modeling. Nevertheless, under conditions with strong high-frequency disturbances, this assumption may no longer hold, which imposes limitations on the model's generality.

We hope to further enhance the method's real-time applicability and robustness through adaptive segmentation and automated parameter selection in future work.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest.

REFERENCES

[1] T. Wang, L. Ding, and H. Yu, "Research and development of fault diagnosis methods for liquid rocket engines", *Aerospace*., vol. 9, no. 9, pp. 481, 2022.

[2] G. E. Box *et al*., *Time series analysis: Forecasting and control*, 5th ed. Hoboken, NJ, USA: John Wiley & Sons, pp. 1–712. 2015.

[3] C. Deng *et al*., "Study on real time diagnosis method of the main stage working condition of rocket engine based on improved ARMA model," *Comput. Meas. Control.*, vol. 28, pp. 33–38, 2020.

[4] W. Xue, Q. Zhang, and X. P. Wu, "Based on the ARMA model for the liquid rocket propulsion fault detection," *Comput. Meas. Control.*, vol. 27, pp. 4–7, 2019.

[5] C. A. Sims, "Macroeconomics and reality," *Econometrica: Journal of the Econometric Society*, vol. 48, no. 1, pp. 1–48, 1980.

[6] H. Lütkepohl, "Vector autoregressive models," in Handbook of research methods and applications in empirical macroeconomics, Cheltenham, UK and Northampton, MA, USA: Edward Elgar Publishing, 2013, pp. 139–164.

[7] D. F. Ahelegbey, P. Giudici, and S. Q. Hashem, "Network VAR models to measure financial contagion," *North Am J. Econ. Finance.*, vol. 55, pp. 101318, 2021.

[8] L. Barbaglia, C. Croux, and I. Wilms, "Volatility spillovers in commodity markets: A large t-vector autoregressive approach," *Energy Econ.*, vol. 85, pp. 104555, 2020.

[9] R. N. Hasanah, R. R. OMP, and H. Suyono, "Comparison analysis of electricity load demand prediction using recurrent neural network (rnn) and vector autoregressive model (var)," in 2020 12th International Conference on Electrical Engineering (ICEENG), IEEE, 2020, pp. 23–29.

[10] R. A. Davis, P. Zang, and T. Zheng, "Sparse vector autoregressive modeling," *J. Comput. Graph. Statistics.*, vol. 25, no. 4, pp. 1077–1096, 2016.

[11] M. Billio, R. Casarin, and L. Rossini, "Bayesian nonparametric sparse VAR models," *J. Econom.*, vol. 212, no. 1, pp. 97–115, 2019.

[12] A. Dallakyan, R. Kim, and M. Pourahmadi, "Time series graphical lasso and sparse VAR estimation," *Comput. Statistics Data Analysis.*, vol. 176, pp. 107557, 2022.

[13] S. Basu, X. Li, and G. Michailidis, "Low rank and structured modeling of high-dimensional vector autoregressions," *IEEE Trans. Signal Process.*, vol. 67, no. 5, pp. 1207–1222, 2019.

[14] P. Bai, A. Safikhani, and G. Michailidis, "Multiple change points detection in low rank and sparse high dimensional vector autoregressive models," *IEEE Trans. Signal Process.*, vol. 68, pp. 3074–3089, 2020.

[15] L. H. L. Whitacker, J. T. Tomita, and C. Bringhenti, "An evaluation of the tip clearance effects on turbine efficiency for space propulsion applications considering liquid rocket engine using turbopumps," *Aerosp. Sci. Technol.*, vol. 70, pp. 55–65, 2017.

[16] X. Zhang *et al*., "Intelligent fault diagnosis of liquid rocket engine via interpretable LSTM with multisensory data," *Sensors*, vol. 23, no. 12, pp. 5636, 2023.

[17] P. J. Bickel, Y. Ritov, and A. B. Tsybakov. "Simultaneous analysis of Lasso and Dantzig selector," 1705–1732, 2009.

[18] J. Fan and R. Li, "Variable selection via nonconcave penalized likelihood and its oracle properties," *J. Am. Stat. Assoc.*, vol. 96, no. 456, pp. 1348–1360, 2001.

[19] A. Safikhani, Y. Bai, and G. Michailidis, "Fast and scalable algorithm for detection of structural breaks in big VAR models." *J. Comput. Graph. Statistics.*, vol. 31, no. 1, pp. 176–89, 2021.

Appendix A

Table AI summarizes the symbols used in this paper.

TABLE AI. List of symbols used in this paper

Symbol	Meaning
X_t	Multivariate time series observation at time t
B_j	Transition matrix for the $j-th$ segment that reflects time-varying effects of historical values
L^*	Low-rank component
S_j^*	Sparse component
p	Number of variables (sensor channels)
b_n	Intercept vector for the $n-th$ segment
k_n	Length of the $n-th$ segment
$\mathbf{X}_{r_j}$	Multivariate time series observations in block j
Y_{r_j}	Multivariate time series observations in block j
$\boldsymbol{\varepsilon}_{r_j}$	Error (innovation) terms in block j
$\mathcal{X}$	Collection of all historical inputs across all blocks
$\mathcal{Y}$	Collection of all current outputs across all blocks
ε	Collection of all error terms.
Z	Block-structured matrix built from segment predictors $\mathbf{X}_{r_j}$, describing segment-specific contributions (Eq. 7)
Θ	Parameter matrix capturing structural changes across segments (Eq. 9)
$\lambda_{1,n}$, $\lambda_{2,n}$, $\lambda_{3,n}$	Regularization parameters controlling penalties.
s_i	End index of the $i-th$ segment
$P_{\lambda,a}(x)$	SCAD (smoothly clipped absolute deviation) penalty function with parameters λ and a.
m	Number of segments (or subintervals) in the time series.
$P_{turbine}$	Turbine output power
Q	Mass flow rate through the turbine
$\eta_{turbine}$	Turbine efficiency
f	Rotational speed (frequency) of turbine shaft
$n_{turbine}$	Rotational speed of turbine (rpm)
T	Torque produced by turbine
$\dot{m}_3$	The total flow rate through the valve system
$\dot{m}_2$	The flow rate through the leakage valve.
$\dot{m}_1$	The flow rate through the primary valve.
c_q	Flow coefficient of the leakage passage.
A	Maximum flow area of the leakage passage.
ρ	Fluid density.
Δp	Pressure drop across the leakage valve.