ENGINEERING Information Technology & Electronic Engineering
https://www.academax.com/EITEE; https://www.engineering.org.cn
https://ieeexplore.ieee.org/xpl/RecentIssue.jsp?punumber=11344065
E-mail: eng.itee@zju.edu.cn

Engineering Journals
Available on IEEE Xplore®

Research Article
https://doi.org/10.1631/ENG.ITEE.2026.0060

Dynamic trend perception for fine-grained unmanned aerial vehicle anomaly detection in flight data

Yuanyuan YIN, Yongfeng YIN✉, Yuecen WEI, Qingran SU, Jiahui LI

School of Software, Beihang University, Beijing 100191, China

Abstract: With the widespread application of UAVs, ensuring their safety and reliability is increasingly crucial. Data-driven anomaly detection methods demonstrate outstanding performance in unmanned aerial vehicle (UAV) health monitoring. Previous studies extract spatiotemporal features from high-dimensional flight data, but emphasize intricate multidimensional correlations, neglecting the intrinsic characteristics of single-dimensional data. This oversight leads to insufficient supervision of intrinsic univariate information. Consequently, subtle deviations in individual flight variables may be masked by dominant multidimensionality, limiting the model's ability to distinguish fine-grained anomalies from normal flight states. To address this problem, we propose F-UAD, a novel dynamic trend perception framework for fine-grained UAV anomaly detection. F-UAD is designed to coordinate univariate feature preservation, multidimensional coupling extraction, and dynamic trend perception, enabling fine-grained anomaly detection of subtle deviations in UAV flight data. During training, the dynamic trend perception module incorporates (1) multiscale spatiotemporal feature reconstruction to balance independent univariate feature preservation with comprehensive multidimensional coupling extraction and (2) dynamic trend forecasting based on decomposed flight trends to improve the perception of complex temporal evolution. During testing, reconstruction consistency and dynamic trend consistency are jointly used to compute anomaly scores, thereby enhancing the separation between normal and abnormal flight states. Experiments demonstrate that our F-UAD framework outperforms state-of-the-art methods across six AirLab failure and anomaly (ALFA) real flight dataset subsets, improving UAV reliability.

Key words: Anomaly detection; Multivariate time series; Unsupervised anomaly detection; Unmanned aerial vehicle (UAV); Flight data

1 Introduction

In recent years, unmanned aerial vehicles (UAVs) have garnered significant attention due to their rapid technological advancements and expanding applications across critical domains, including environmental monitoring (Ullah et al., 2019; Radoglou-Grammatikis et al., 2020), military operations (Baskaya et al., 2017; Chen MZ et al., 2021), and transportation (Khosravi and Samadi, 2021; Maddikunta et al., 2021). The increasing adoption of UAVs has highlighted the urgent need for safe and reliable autonomous flight operations (Chen YF et al., 2017). According to the European Aviation Safety Agency (EASA) 2023 safety review, two of the 12 reported serious UAV incidents involved collisions with manned aircraft, reflecting an 83% accident rate primarily attributed to technical malfunctions. Consequently, the timely detection of anomalous UAV behavior and precise identification of faults are critical prerequisites for enhancing flight safety, minimizing economic losses, and protecting personnel (Yang L et al., 2023). Extensive research shows that many deep learning models overlook mutual cancellation effects in feature extraction. This leads to poor spatiotemporal feature capture, limited complex data perception, and reduced anomaly detection precision (Harrou et al., 2015; Pang et al., 2015).

Current research has introduced various improvements, but limitations remain in effectively capturing the spatiotemporal features and dynamic characteristics of flight data, which impacts the accuracy and robustness of anomaly detection.

✉ Yongfeng YIN, yyf@buaa.edu.cn
Yuanyuan YIN, https://orcid.org/0009-0008-0967-7794
Yongfeng YIN, https://orcid.org/0000-0001-9432-3051
Yuecen WEI, https://orcid.org/0000-0002-0218-948X
Qingran SU, https://orcid.org/0000-0003-1488-900X
Jiahui LI, https://orcid.org/0009-0005-5272-8698
CLC number: TP277

Received: Mar. 9, 2026; Revision accepted: July 16, 2026;
Crosschecked: Aug. 3, 2026; Published online: Sept. 9, 2026

© The Authors 2026. Published by Zhejiang University Press Co., Ltd. This is an open access article distributed under the terms of the CC BY-NC-ND license (https://creativecommons.org/licenses/by-nc-nd/4.0/)

Data-driven methods (Baskaya et al., 2017; Keipour et al., 2019) build anomaly detection models by analyzing sensor data from flight operations, capturing correlations among internal features without requiring detailed knowledge of system dynamics (Yang L et al., 2023). For example, support vector machines (SVMs) (Liu DT et al., 2017, 2018) and Gaussian process regression (Pang et al., 2015) learn patterns from normal sensor data using nonlinear functions to detect anomalies. Reconstruction models, notably autoencoders (An and Cho, 2015; Wang et al., 2019), capture spatiotemporal correlations in multidimensional flight data, making them particularly effective in detecting complex failure modes unsupervised. Prediction-based methods, such as those employing long short-term memory (LSTM) networks (Bontemps et al., 2016; Park JH et al., 2022), model temporal dependencies to forecast future states, enabling proactive anomaly identification.

UAV flight data are typically represented as a multidimensional time-series (e.g., acceleration, attitude angles, and voltage), exhibiting significant cross-variable physical coupling (Fig. 1) due to the multi-degree-of-freedom dynamics, such as nonlinear interactions between roll angles, pitch angles, and yaw angles. However, each single-dimensional data stream also possesses intrinsic characteristics that are independent of other variables. These characteristics arise from the specific physical constraints (e.g., range and rate of change) that they must adhere to. Therefore, UAV flight data are a combination of these coupled and intrinsic characteristics. Deep learning has demonstrated advantages in multivariate time-series anomaly detection (MTSAD) (Homayouni et al., 2021; Liu PF et al., 2022), offering the potential to advance data-driven UAV anomaly detection. Notably, variational autoencoders (VAEs) (Lin et al., 2020) and Transformer-based models (Xiong et al., 2024) have proven effective in capturing spatiotemporal correlations, enhancing detection performance in complex dynamical systems (Wu et al., 2021). Additionally, hybrid methodologies, such as generative adversarial network (GAN) models (Zhou B et al., 2019), integrate reconstruction and prediction techniques to synergistically model multidimensional features (Li et al., 2019). However, this starkly contrasts with the single-distribution or weakly coupled characteristics of other multivariate time-series domains, like financial fraud and smart traffic data. While existing methods capture spatiotemporal correlations, they often prioritize multidimensional coupling at the expense of intrinsic univariate features, resulting in inadequate spatiotemporal feature extraction. This imbalance reduces sensitivity to the complex dynamics inherent in UAV operations. The primary challenge lies in effectively coordinating independent modeling of univariate features with comprehensive extraction of multidimensional associations to improve the perception of these dynamics.

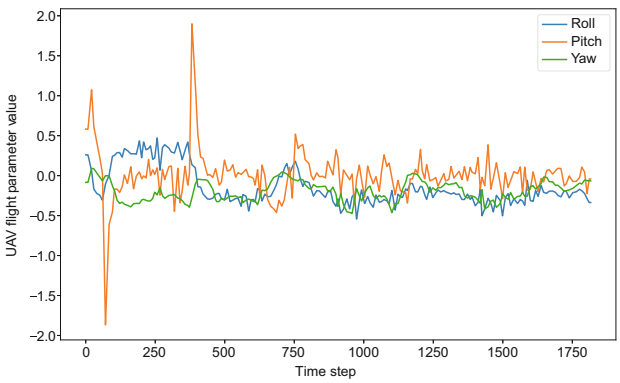

Fig. 1 Visualization of three UAV flight parameters (ALFA dataset)

Drawing on current research, we investigate multiscale feature extraction and trend decomposition strategies to improve the representational capacity of UAV flight data, addressing two key challenges: (1) Dynamic trend perception. UAV flight patterns exhibit complex and unpredictable dynamics, resulting in data characterized by intricate fluctuations and high-dimensional modal heterogeneity. Existing anomaly detection methods for UAVs typically exploit spatiotemporal, temporal, or spatial correlations to capture flight data features, often at the expense of suppressing univariate data characteristics due to strong inter-data coupling. The first challenge is to capture robust associative relationships among high-dimensional data while preserving the distinct attributes of individual dimensions. (2) Fine-grained anomaly detection (FAD). Current UAV anomaly score calculations rely on direct processing of univariate data, overlooking the persistent dynamics and contextual relationships of anomalies. The second challenge involves developing a comprehensive method for computing anomaly scores to effectively capture the dynamic interaction characteristics of multidimensional flight data.

To address these challenges, we propose a dynamic trend perception for fine-grained UAV anomaly detection (F-UAD). The main contribution of F-UAD lies in jointly considering univariate feature preservation, multidimensional coupling extraction, and dynamic trend perception for UAV flight data. Specifically, multiscale feature reconstruction is used to retain subtle single-variable characteristics, while dynamic trend forecasting provides temporal evolution information for fine-grained anomaly scoring. Unlike traditional coarse-grained methods that focus only on macroscopic anomalies, F-UAD distinguishes subtle abnormal deviations from normal flight states, significantly enhancing the accuracy and comprehensiveness of anomaly detection. Empirical evaluations on real flight datasets demonstrate the framework's superior effectiveness in detecting anomalies. Our contributions are summarized as follows:

1. Dynamic trend perception. We design a multiscale spatiotemporal feature reconstruction method that effectively captures complex dynamics in multidimensional time-series data while preserving univariate characteristics. Additionally, we propose a dynamic trend forecasting model to enhance adaptability to evolving flight data dynamics.

2. FAD. We propose a wide-ranging anomaly score calculation method that leverages joint optimization to capture temporal interactions in multidimensional flight data, improving anomaly detection accuracy and addressing the limitations of single-point computations.

3. Experimental evaluation. The F-UAD framework is evaluated on six subsets of the AirLab failure and anomaly (ALFA) real flight dataset, benchmarked against 10 state-of-the-art (SOTA) deep learning models. The results demonstrate superior performance, confirming the framework's effectiveness in detecting anomalies in UAV flight data.

2 Related works

This section presents existing works relevant to our paper, grouped into two categories: UAV anomaly detection methods and multivariate time-series anomaly detection methods.

2.1 Anomaly detection in UAVs

UAV anomaly detection methods are primarily categorized into model-based (Narasimhan and Brownston, 2007; D'Amato et al., 2018) and data-driven (Chandola, 2009; Qin et al., 2021) approaches. Model-based methods construct physical models of UAVs using flight dynamics (Yi and Zhang, 2017). Guo et al. (2018) employed the observability rank criterion based on Lie derivatives for validation and a cumulative sum detector to identify faults. Marcos et al. (2022) proposed a parameterized linear parameter-varying (LPV) system to detect abnormal flight data. However, these methods rely on expert knowledge to build physical models, which is challenging to acquire and limited to specific systems (Ahmad et al., 2022), thus restricting their application. By contrast, data-driven methodologies do not necessitate precise physical models for system description and do not depend on a profound comprehension of internal mechanisms (Oussous et al., 2018; Zhang QC et al., 2018). Recently, deep learning has garnered significant scholarly interest for its capacity to automatically extract and learn features for constructing anomaly detection models, primarily encompassing reconstruction and prediction models. Zhong et al. (2022) used artificial neural network correlation analysis (ANNCA) to extract spatiotemporal correlations and applied the 3σ principle (σ is the standard deviation) for detection and recovery prediction. Zhou X et al. (2023) employed the maximum information coefficient (MIC) and an attention mechanism LSTM for fault detection. Yang L et al. (2024) introduced a spatiotemporal correlation multivariate regression framework (STC-1D CBiAM-RF) to extract spatiotemporal features and mitigate noise using filtering techniques. Park KH et al. (2021) introduced an unsupervised fault detection model using Stacked-VAE in the reconstruction method. This model is trained on normal data and categorizes states through reconstruction loss. Jiang et al. (2024) employed the maximum correntropy criteria (MCC) to bolster the resilience of the detection model. Regarding hybrid models, Bell et al. (2022) combined LSTM, deep autoencoders, and dynamic threshold algorithms to enhance performance. Zhang YX et al. (2023) developed a deep convolutional self-encoding memory network (CAE-M). This network uses convolutional autoencoders to extract low-dimensional features, incorporates the maximum mean difference (MMD) to mitigate noise, employs the attention mechanism bi-directional LSTM (BiLSTM) to capture time dependencies, and optimizes the network for precise detection. Although existing UAV anomaly detection methods improve performance through correlation modeling, reconstruction, or prediction-reconstruction fusion, single-variable deviations may still be weakened during multivariate joint learning, latent representation compression, or global anomaly score aggregation. Therefore, preserving intrinsic univariate characteristics while modeling multidimensional coupling and dynamic persistence remains a key challenge for fine-grained UAV anomaly detection.

2.2 Anomaly detection in multivariate time series

The complexity of real-world interaction data has led to a growing interest in MTSAD, which focuses on detecting behaviors that deviate significantly from normal patterns in multidimensional correlation data (such as sensor readings and performance indicators) (Vaswani et al., 2017; Choi et al., 2021). In the realm of deep learning, two prominent strategies for MTSAD are prediction (Zhao et al., 2020; Deng and Hooi, 2021) and reconstruction methods (Audibert et al., 2020; Garg et al., 2022; Nam et al., 2024). Nevertheless, the effectiveness of prediction or reconstruction models in anomaly detection is constrained due to the influence of numerous factors on multivariate time-series variables. To address this challenge, researchers have explored the integration of prediction and reconstruction models (Nguyen et al., 2021; Tuli et al., 2022) and contrastive learning (Yang YY et al., 2023; Darban et al., 2025a). Yin et al. (2022) combined a convolutional neural network (CNN) with recurrent neural network (RNN)-based autoencoders to extract spatiotemporal features from Internet of Things (IoT) time-series data, improving anomaly detection performance. Models based on GANs (Tuli et al., 2022) have successfully integrated multidimensional feature modeling by merging reconstruction and prediction tasks. Braei and Wagner (2020) used LSTM to enhance anomaly detection capabilities and streamline GAN training through the joint training of the encoder, generator, and discriminator. Tuli et al. (2022) devised a Transformer-based temporal anomaly detection architecture that incorporates a multi-attention mechanism and a two-stage training strategy (adversarial learning and meta-learning) to calculate anomaly scores based on reconstruction and prediction errors, significantly improving the modeling of multidimensional temporal dynamics. Yang YY et al. (2023) developed a dual-attention asymmetric approach using contrastive learning, achieving a permutation-invariant representation. Current research on high-dimensional heterogeneous data overlooks comprehensive feature representation while lacking robust modeling of complex dynamic characteristics, resulting in insufficient anomaly detection distinctiveness.

3 Preliminaries

3.1 Time-series decomposition

Time-series decomposition is a statistical technique that disentangles time-series data into distinct components to unveil underlying structure and dynamic features. Mathematically, it can be represented as follows for a given time series $\boldsymbol{\mathcal{X}}_t$:

$$\boldsymbol{\mathcal{X}}_t = \boldsymbol{\mathcal{T}}_t + \boldsymbol{\mathcal{S}}_t + \boldsymbol{\mathcal{R}}_t, \tag{1}$$

where $\boldsymbol{\mathcal{T}}_t$ represents the trend component at time t, which reflects the long-term trend of the response series, $\boldsymbol{\mathcal{S}}_t$ represents the seasonal component, accounting for systematic periodic fluctuations, and $\boldsymbol{\mathcal{R}}_t$ represents the residual component, encompassing random noise that is not accounted for by trends and seasonality.

To quantitatively assess the impact of trend and seasonal components on time series, we define trend intensity $\mathcal{F}_{\mathrm{t}}$ and

seasonal intensity $\mathcal{F}_{\rm s}$, which are calculated as follows:

$$\mathcal{F}_{\rm t} = \max\left(0, 1 - \frac{\mathrm{Var}(\boldsymbol{\mathcal{R}}_t)}{\mathrm{Var}(\boldsymbol{\mathcal{T}}_t + \boldsymbol{\mathcal{R}}_t)}\right), \tag{2}$$

$$\mathcal{F}_{\rm s} = \max\left(0, 1 - \frac{\mathrm{Var}(\boldsymbol{\mathcal{R}}_t)}{\mathrm{Var}(\boldsymbol{\mathcal{S}}_t + \boldsymbol{\mathcal{R}}_t)}\right), \tag{3}$$

where the variance operator denoted as $\mathrm{Var}(\cdot)$ is used to calculate the variance of a variable. In this context, $\mathcal{F}_{\rm t}$, and $\mathcal{F}_{\rm s}$ are values within the range of $[0, 1]$. The ratio $\mathrm{Var}(\boldsymbol{\mathcal{R}}_t)/\mathrm{Var}(\boldsymbol{\mathcal{T}}_t + \boldsymbol{\mathcal{R}}_t)$ quantifies the proportion of residual variance in the combined variance of trend and variance. A value less than 1 ($\mathrm{Var}(\boldsymbol{\mathcal{R}}_t)/\mathrm{Var}(\boldsymbol{\mathcal{T}}_t + \boldsymbol{\mathcal{R}}_t) < 1$) suggests a strong influence of the trend component on the overall variability, particularly when the $\mathcal{F}_{\rm t}$ value is high (e.g., $\mathcal{F}_{\rm t} > 0.9$). Similarly, the ratio $\mathrm{Var}(\boldsymbol{\mathcal{R}}_t)/\mathrm{Var}(\boldsymbol{\mathcal{S}}_t + \boldsymbol{\mathcal{R}}_t)$ assesses the contribution of residual variance to both seasonality and the combined residual variance. A high $\mathcal{F}_{\rm s}$ value indicates the significance of the seasonal component in the overall variance.

In this paper, the components in Eq. (1) are derived using seasonal and trend decomposition using LOESS (STL) method (LOESS is locally estimated scatterplot smoothing). Subsequently, the ALFA dataset is assessed using Eq. (2), yielding trend strength $\mathcal{F}_{\rm t} = 0.98$ and seasonal strength $\mathcal{F}_{\rm s} = 0.2$. These findings suggest that the ALFA dataset is primarily influenced by trend components. Additionally, an analysis of the residual component's distribution $\boldsymbol{\mathcal{R}}_t$ reveals its approximate adherence to the standard normal distribution (Fig. 2). This conformity implies that the residual primarily comprises random noise devoid of systematic patterns, affirming the appropriateness of the decomposition method employed. This attribute signifies that the ALFA dataset constitutes a stationary time series primarily characterized by trends, thereby allowing trend components to effectively encapsulate the series' overall characteristics.

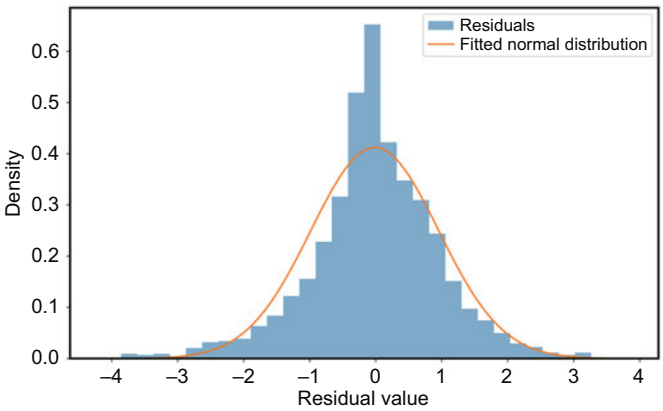

Fig. 2 Histogram of residual distribution based on the ALFA dataset

3.2 Problem definition

This paper investigates unsupervised anomaly detection in multivariate time series for UAV systems. UAV flight data are typically gathered from various sensors and stored in either the flight controller or the ground control station (GCS). Due to the significant imbalance between normal and abnormal data in real-world scenarios, we propose an unsupervised training method based on normal flight data to detect anomalies in test data that may contain anomalies. The training and test datasets are denoted as $\boldsymbol{\mathcal{X}}$ and $\boldsymbol{\mathcal{X}}'$, respectively. During the training phase, a collection of normal multivariate time-series flight data is provided from m sensors of the UAV over T time steps:

$$\boldsymbol{\mathcal{X}} = \Big[\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_T\Big] := \{\boldsymbol{x}_t\}_{t=1}^{T}, \tag{4}$$

where $\boldsymbol{x}_t \in \mathbb{R}^m$ denotes the m-dimensional data vector obtained at time t. A labeled set of T' time steps of test data $\{\boldsymbol{\mathcal{X}}', \boldsymbol{\mathcal{Y}}'\} = \{(\boldsymbol{x}'_t, y'_t)\}_{t=1}^{T'}$ is collected from m sensors. $y'_t = 1$ signifies an anomalous observation and $y'_t = 0$ denotes a normal observation. To forecast binary labels $y'_t \in \{0, 1\}$, the proposed method evaluates abnormal states by computing an abnormal score $\mathbf{sc}_{2,t}$ for each time step and subsequently comparing it against a predefined threshold $\boldsymbol{\tau}$. Here, the subscript 2 denotes the abnormal score during the testing.

4 Methodology

This section outlines the workflow of the proposed F-UAD framework designed for unsupervised anomaly detection in intricate multidimensional flight data. This framework effectively captures data dynamics and possible anomaly patterns through a multi-stage collaborative process that integrates data preprocessing, dynamic trend perception, and FAD.

4.1 Model architecture

We introduce a novel dynamic trend perception for the F-UAD framework to tackle the complex characteristics of flight data, which are high-dimensional, non-stationary, and multimodal. The architecture is depicted in Fig. 3, which extracts critical features from raw data via a perception process, infers future states from historical trends through a prediction process, and generates outputs closely aligned with the original data using a reconstruction process. These interconnected processes collaboratively form a robust and precise anomaly detection pipeline. Algorithms 1 and 2 illustrate a comprehensive framework from data input to anomaly detection output. The framework comprises three core modules: (1) Flight data preprocessing module. This module encompasses a workflow from correlation analysis to aerodynamic modeling, transforming raw flight data into a structured format and laying the groundwork for subsequent analysis. (2) Dynamic trend perception module. This module employs reconstruction and prediction strategies to extract features from multidimensional flight data, capturing both local and global dynamic patterns in time series and generating dimensional feature representations. (3) FAD module. This module calculates anomaly scores for each dimension based on the acquired features and determines anomaly states of each dimension's data using predefined thresholds. These modules are specifically designed to address the intricate nature of UAV flight data, enabling models to dynamically perceive flight behaviors and detect anomalies effectively. The structured design of the framework is transparent, demonstrating its systematic unsupervised anomaly detection approach through the visualization of data flow and module interactions.

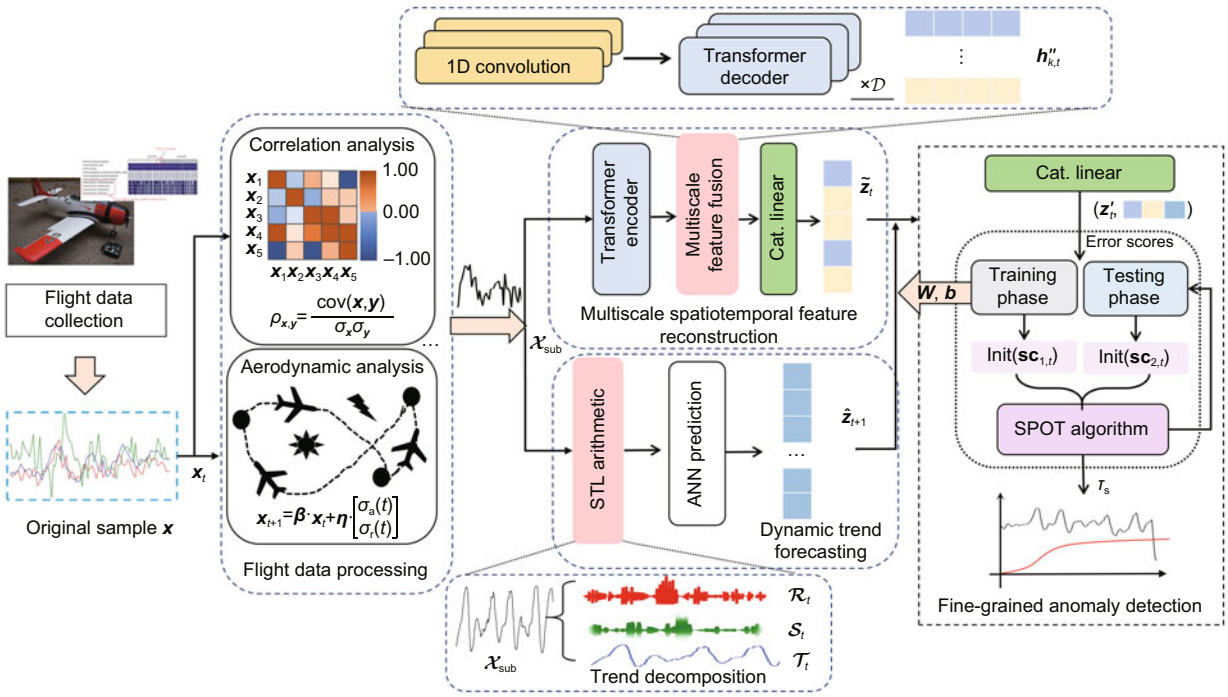

Fig. 3 The F-UAD framework, with input $\mathcal{X}$, comprises three core components designed to enhance anomaly detection in UAV flight data. (1) Flight data preprocessing: we first process the raw flight data $\mathcal{X}$ using feature selection, data space reconstruction, etc. (2) Dynamic trend perception: the multiscale spatiotemporal feature reconstruction model, along with the dynamic trend forecasting model, is used to obtain the refined representation $\tilde{z}_t$ and $\hat{z}_{t+1}$. (3) FAD: $\tilde{z}_t$ and $\hat{z}_{t+1}$ are injected into the fine-grained fusion network, and the finalized test results are the final output

Algorithm 1 Training procedure of F-UAD

Input: normal UAV's flight data $\mathcal{X} = [\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_T]$; number of training epochs E.
Output: weight matrix $\boldsymbol{W}$ and bias matrix $\boldsymbol{b}$ in neural networks, and error sequence $\mathbf{sc}_{1,t}$.
preprocess each sample $\mathcal{X} \in \mathbb{R}^{m\times T}$ into $\mathcal{X}_{\text{sub}} \in \mathbb{R}^{\mathcal{P}\times m\times n}$
let $t = 0$
randomly initialize parameters $\boldsymbol{W}$ and $\boldsymbol{b}$
while the training epoch does not converge **do**
  calculate low-dimensional representation $\boldsymbol{h}_t^i$ and multiscale feature representations $\boldsymbol{h}_{k,t}^{\prime(j_c)}$ at each time step using Eqs. (9)–(11)
  compute the multiscale feature reconstruction $\boldsymbol{h}_{k,t}^{\prime\prime i}$ and reconstruction fusion $\tilde{z}_t$ at each time point using Eqs. (12) and (13)
  extract the trend component of the sequence $\mathcal{X}_{\text{sub}}$ and use the values $\boldsymbol{\mathcal{T}}_t$ in the past window to predict the value $\hat{z}_{t+1}$ at the next time point by feed-forward neural network (FNN) using Eqs. (1)–(17)
  denote the fusion result of $\tilde{z}_t$ and $\hat{z}_{t+1}$ at each sample point as $\boldsymbol{z}_t'$ using Eq. (18)
  update $\boldsymbol{W}$ and $\boldsymbol{b}$ by minimizing the compound objective function using Eq. (19)
end while
calculate the sequence error $\mathbf{sc}_{1,t}$ by the training samples under the optimal $\boldsymbol{W}$ and $\boldsymbol{b}$ using Eq. (19)
return optimal $\boldsymbol{W}$, $\boldsymbol{b}$ and error sequence $\mathbf{sc}_{1,t}$

Algorithm 2 Inference procedure of F-UAD

Input: anomalous dataset $\mathcal{X}' = [\boldsymbol{x}_1', \boldsymbol{x}_2', \ldots, \boldsymbol{x}_{T'}']$, error sequence $\mathbf{sc}_{1,t}$, and model parameters $\boldsymbol{W}$ and $\boldsymbol{b}$.
Output: labels of all y_t''s.
preprocess each sample $\mathcal{X}' \in \mathbb{R}^{m\times T'}$ into $\mathcal{X}'_{\text{sub}} \in \mathbb{R}^{\mathcal{P}'\times m\times n}$
for all $\boldsymbol{x}_t'$ **do**
  calculate anomaly score $\mathbf{sc}_{2,t}$ using Eq. (19)
  calculate fixed thresholds $\boldsymbol{\tau}_{\text{s}}$ for each dimension using SPOT methods from sequence errors $\mathbf{sc}_{1,t}$ and anomaly score $\mathbf{sc}_{2,t}$, using Eq. (20)
  if $\mathbf{sc}_{2,t} > \boldsymbol{\tau}_{\text{s}}$ **then**
    $y_t' = 1$ //abnormal
  else
    $y_t' = 0$ //normal
  end if
end for
return labels of all y_t''s

4.2 Flight data preprocessing

The data preprocessing module systematically processes the original multidimensional flight data $\mathcal{X} \in \mathbb{R}^{m\times T}$ and $\mathcal{X}' \in \mathbb{R}^{m\times T'}$ to enhance the performance of subsequent modules. The preprocessing includes the following key essential steps. Following common practice in multivariate time-series anomaly detection (Audibert et al., 2020; Tuli et al., 2022; Darban et al., 2025b), the preprocessing pipeline includes timestamp alignment, resampling, feature-wise normalization, feature selection, and data space reconstruction.

4.2.1 Feature selection

The UAV dataset comprises control signals, inertial measurements, timestamps, and redundant features. The accuracy of model reconstruction and prediction is hindered by irrelevant variable interference, necessitating careful selection of input features. In this paper, input features for the UAV

fault detection model are initially chosen based on Pearson correlation coefficients, with features exhibiting lower correlations being eliminated. Subsequently, residual characteristics are selected using the decoupled state equation of the transverse system:

$$\boldsymbol{x}_{t+1} = \boldsymbol{\beta} \cdot \boldsymbol{x}_t + \boldsymbol{\eta} \cdot \begin{bmatrix} \sigma_{\mathrm{a}}(t) \\ \sigma_{\mathrm{r}}(t) \end{bmatrix}, \tag{5}$$

where $\begin{bmatrix} \sigma_{\mathrm{a}}(t) \\ \sigma_{\mathrm{r}}(t) \end{bmatrix}$ denotes the control instructions. The matrix $\boldsymbol{\beta}$ serves as a state transition matrix, delineating the dynamic mapping relationship among the internal states of the system. The values of its components are contingent upon the dynamic characteristics and kinematic parameters of the UAV, encompassing factors like the body structure and aerodynamic coefficients. Additionally, the matrix $\boldsymbol{\eta}$ functions as the input matrix, specifying the impact of external control inputs on the system state evolution.

4.2.2 Data space reconstruction

Data space reconstruction involves segmenting the original multidimensional time-series data into fixed-length subsequences using a sliding window strategy. In the case of the training set sequence, the data are divided into $\mathcal{P} = T - n + 1$ sub-sequences by employing a sliding window of length n and a step size of 1:

$$\boldsymbol{\mathcal{X}}_{\mathrm{sub}} = [\boldsymbol{\mathcal{X}}_1, \boldsymbol{\mathcal{X}}_2, \ldots, \boldsymbol{\mathcal{X}}_{\mathcal{P}}], \tag{6}$$

$$\boldsymbol{\mathcal{X}}_t = [\boldsymbol{x}_{t-n+1}, \boldsymbol{x}_{t-n+2}, \ldots, \boldsymbol{x}_t] \in \mathbb{R}^{m \times n}, \tag{7}$$

where $\boldsymbol{x}_t \in \mathbb{R}^m$ represents a multidimensional observation at time t. The test set sequence $\boldsymbol{\mathcal{X}}'$ follows a similar sliding window approach to produce $\mathcal{P}' = T' - n + 1$ subsequences:

$$\boldsymbol{\mathcal{X}}'_{\mathrm{sub}} = [\boldsymbol{\mathcal{X}}'_1, \boldsymbol{\mathcal{X}}'_2, \ldots, \boldsymbol{\mathcal{X}}'_{\mathcal{P}'}]. \tag{8}$$

This reconstruction method transforms the initial time series into a sequence of high-dimensional subsequence matrices. These matrices retain the local dynamic attributes of the time series and are compatible with the model's input dimensions. This process establishes a structured data basis for subsequent feature extraction and anomaly detection.

4.3 Dynamic trend perception

The dynamic trend perception module is a pivotal component of the F-UAD framework, designed to discern multiscale and perspective dynamic patterns embedded within multidimensional flight data. This module first perceives and extracts key features and underlying trend components from the raw data. It then reconstructs the normal of the original data and predicts the expected future behaviors based on the perceived key features and trend components. To this end, we introduce a dual-channel learning framework comprising two synergistic submodules: (1) a multiscale spatiotemporal feature reconstruction module that reconstructs low-dimensional representations of multivariate data to enhance feature expressiveness, and (2) a dynamic trend forecasting module using STL to isolate and forecast long-term trend sequences from flight data. The following subsections rigorously elucidate the mathematical formulations and implementation details of these submodules.

4.3.1 Multiscale spatiotemporal feature reconstruction

To efficiently encode and reconstruct preprocessed input time series $\boldsymbol{\mathcal{X}}_{\mathrm{sub}} \in \mathbb{R}^{\mathcal{P} \times m \times n}$, we define the iteration range for feature extraction. Our approach involves a fine-grained spatiotemporal reconstruction strategy that integrates a Transformer and one-dimensional (1D) CNNs for multiscale feature extraction and reconstruction. Each subsequence $\boldsymbol{\mathcal{X}}_t$ undergoes independent processing through the modules to produce multiscale feature representations and reconstructed sequences. Specifically, as shown in Fig. 3, we first use an $\mathcal{N}$-layer Transformer encoder with a multi-head self-attention mechanism to encode the original input data $\boldsymbol{\mathcal{X}}_t$ to capture their long-term temporal dependencies. We then use $\mathcal{D}$ down-sampled 1D CNNs with different convolutional kernels in parallel to encode the encoded features at a more fine-grained level, thereby obtaining latent features at different scales. These parallel CNNs are crucial for capturing the unique intrinsic characteristics of every single variable. Finally, we use an $\mathcal{N}$-layer Transformer decoder to process the latent features at each scale in parallel and fuse all decoding results to generate the final reconstructed output $\tilde{\boldsymbol{z}}_t$. This design effectively balances the inherent characteristics of a single variable with the coupling features between multiple variables. It captures the complex interactions between variables through the global modeling capabilities of the Transformer, while leveraging the local feature extraction capabilities of CNN to preserve the unique patterns of a single variable, thereby achieving more accurate reconstruction and more robust performance in anomaly detection tasks. Thus, we initially adhere to the encoder's information propagation mechanism and define the ith layer's embedding learning $\boldsymbol{h}_t^i$ as follows:

$$\begin{aligned} \boldsymbol{h}_t^1 = & \Big\{ \sigma_{\mathrm{en}}(\boldsymbol{W}_{\mathrm{e1}}^1 \boldsymbol{\mathcal{X}}_t + \boldsymbol{b}_{\mathrm{e1}}^1), \sigma_{\mathrm{en}}(\boldsymbol{W}_{\mathrm{e2}}^1 \boldsymbol{\mathcal{X}}_t + \boldsymbol{b}_{\mathrm{e2}}^1), \ldots, \\ & \sigma_{\mathrm{en}}(\boldsymbol{W}_{en}^1 \boldsymbol{\mathcal{X}}_t + \boldsymbol{b}_{en}^1) \Big\}, \end{aligned} \tag{9}$$

$$\boldsymbol{h}_t^i = \left\{ \sum_{j=1}^{n} \sigma_{\mathrm{en}} \left(\boldsymbol{W}_{\mathrm{e}j}^i \boldsymbol{h}_t^{i-1} + \boldsymbol{b}_{\mathrm{e}j}^i \right) \right\}, \tag{10}$$

where the layer $i \geq 2$, $\boldsymbol{W}_{\mathrm{e}j}^i$ is a training weight matrix for the jth neuron in the ith layer, tailored to the input embedding, $\boldsymbol{b}_{\mathrm{e}j}^i$ is an enhanced bias term, and σ_{en} is a nonlinear activation function aimed at improving feature interaction functionality. $\boldsymbol{h}_t^i \in \mathbb{R}^{m \times n}$ represents the output of the Transformer encoder, where m denotes its dimensions. Each layer is designed to learn the dependency relationships among data at each timestamp in the input sequence through a multi-head self-attention mechanism, extracting data features. It is noteworthy that, unlike Transformer encoders for natural language processing, a multi-headed self-attention mechanism with masks is employed in our approach to characterize current features without incorporating future time information. Specifically, this approach ensures that the encoder considers only past and current timestamps during its operations by setting the upper triangular element of the attention matrix to $-\infty$.

To capture multidimensional temporal flight data features across various time scales, we employ $\mathcal{D}$ parallel branches of 1D CNNs. Each branch uses a 1D CNN with distinct kernel sizes l_k to extract multiscale temporal features through point-to-point convolution. The operation for the kth convolution

branch is specified as follows:

$$\boldsymbol{h}_{k,t}^{\prime(j_c)} = \sum_{c=1}^{\mathcal{D}} \sum_{o=1}^{l_k-1} \left(\boldsymbol{W}_{ck}^{(j_c)}[c,o] \cdot \boldsymbol{h}_t^N[c,t+o] + \boldsymbol{b}_{ck}^{(j_c)} \right), \tag{11}$$

where $\boldsymbol{b}_{ck}^{(j_c)} \in \mathbb{R}^{m\times n}$ represents bias, m is the number of output channels, N in $\boldsymbol{h}_t^N[c,t+o]$ is the final layer, j_c is the output channel for CNN, and $\boldsymbol{W}_{ck}^{(j_c)}[c,o]$ signifies the weight of the kth convolution kernel at input channel c and time offset o.

In contrast to the conventional Transformer approach, which relies on the encoder for final feature extraction and direct reconstruction, this paper augments feature expression by conducting multiscale downsampling feature extraction on the encoder's final feature. After performing $\mathcal{D}$ convolution operations to generate feature results, we use an $\mathcal{N}$-layer Transformer decoder that integrates features $\boldsymbol{h}_{k,t}^{(j_c)}$ from $\mathcal{D}$ CNN branches. The goal is to reconstruct a multiscale feature representation, combining multiple sets of information to match the input sequence dimensions. Following a similar information propagation mechanism to the encoder, we initially focus on the ith layer's embedding learning $\boldsymbol{h}_{k,t}^{\prime\prime i}$ of the kth branch. The process equation is defined as follows:

$$\boldsymbol{h}_{k,t}^{\prime\prime i} = \mathrm{TD}^i\Big(\mathrm{LN}^i \sum_{w=1}^{n} \mathrm{MH}^i\big(\boldsymbol{W}_w^i \boldsymbol{h}_{k,t}^{\prime} + \boldsymbol{b}_w^i\big)\Big), \tag{12}$$

where $\mathrm{TD}(\cdot)$ denotes the activation function of the Transformer decoder, $\boldsymbol{W}_w^i$ denotes the weights of layer i, and $\boldsymbol{b}_w^i$ represents the bias of layer i. A fully-connected linear transformation $\mathrm{LN}(\cdot)$ is applied to enhance training stability. Additionally, a multi-head attention mechanism with masks $\mathrm{MH}(\cdot)$ is used to capture dependencies across various representation subspaces. The $\mathcal{N}$-layer decoder produces $\mathcal{D}$ sets of the decoding results $\boldsymbol{h}_t^{\prime\prime} \in \mathbb{R}^{\mathcal{D}\times m\times n}$. To leverage diverse fine-grained feature representations from each set, these representations are concatenated to define the decoder output. This results in a sequence of m-dimensional vectors $\tilde{\boldsymbol{z}}_t$ matching the size of the input sequence, achieved through a full join operation. These vectors facilitate subsequent outlier score predictions. The detailed calculation process is as follows:

$$\tilde{\boldsymbol{z}}_t = \mathrm{FC}\left[\boldsymbol{h}_{1,t}^{\prime\prime}, \boldsymbol{h}_{2,t}^{\prime\prime}, \ldots, \boldsymbol{h}_{k,t}^{\prime\prime}\right], \tag{13}$$

where the concatenation operation function and the linear mapping function $\mathrm{FC}(\cdot)$ are used to concatenate decoding results $\boldsymbol{h}_k^{\prime\prime}$ from $\mathcal{D}$ sets in a row-wise manner. Subsequently, the linear mapping process is applied to generate the final reconstructed sequence $\tilde{\boldsymbol{z}}_t \in \mathbb{R}^{m\times n}$.

4.3.2 Dynamic trend forecasting

The strategy of multiscale spatiotemporal feature reconstruction effectively balances independent modeling of single-dimensional features and comprehensive extraction of multi-dimensional coupled features through a collaborative architecture of a Transformer and a 1D CNN. However, the high complexity and non-deterministic nature of UAV flight data present significant challenges. This type of data typically demonstrates strong fluctuations and high sensitivity, which significantly hinders the model's ability to perceive complex dynamics. To address this limitation, this paper introduces the STL method, which is based on local weighted regression and neural network prediction. Given that UAV flight data typically contain long-term trends, seasonality, and residual components, we focus on long-term trends as the key prediction target. Specifically, long-term trends refer to macroscopic changes within a single flight mission, such as a continuous drop in battery voltage; seasonality denotes periodic fluctuations within a single mission caused by operational patterns or mechanical motion, such as the periodic circling of a UAV on a specific route; residual components capture irregular fluctuations not attributed to trends or seasonality, such as random changes due to environmental noise or sensor errors. In our implementation, only the extracted trend component is used for subsequent prediction, while the seasonal and residual components are not fed into the reconstruction module. This improves the model's ability to perceive complex dynamics and enables it to more accurately model normal trend patterns in the data, ultimately improving the sensitivity and accuracy of anomaly detection.

Initially, the preprocessed time series $\boldsymbol{\mathcal{X}}_t$ undergoes decomposition into trend component $\boldsymbol{\mathcal{T}}_t \in \mathbb{R}^{m\times n}$, seasonal component $\boldsymbol{\mathcal{S}}_t \in \mathbb{R}^{m\times n}$, and residual components $\boldsymbol{\mathcal{R}}_t \in \mathbb{R}^{m\times n}$ using the STL method, which can be expressed as in Eq. (1).

The method employs alternating inner and outer loops to accurately decompose time-series data. The inner loop focuses on extracting trend and seasonal components by applying multiple local weighted regression smoothing techniques. It iteratively refines the estimation of trend and seasonal components through repeated smoothing, de-seasonalizing, and de-trending operations on the original series. Conversely, the outer loop is used to fine-tune the seasonal periodic variations to align with the authentic seasonal patterns within the data. The detailed computational procedure or local weighted regression iteration during the decomposition process is outlined as follows:

$$\boldsymbol{\mathcal{S}}_t^{(u)} = \mathrm{LOESS}\left(\boldsymbol{\mathcal{X}}_t - \boldsymbol{\mathcal{T}}_t^{(u-1)}, q_{\mathrm{s}}, s\right), \tag{14}$$

$$\boldsymbol{\mathcal{T}}_t^{(u)} = \mathrm{LOESS}\left(\boldsymbol{\mathcal{X}}_t - \boldsymbol{\mathcal{S}}_t^{(u)}, q_{\mathrm{t}}\right), \tag{15}$$

$$\boldsymbol{\mathcal{R}}_t^{(k)} = \mathrm{LOESS}\left(\boldsymbol{\mathcal{X}}_t - \boldsymbol{\mathcal{T}}_t^{(u)}, \boldsymbol{\mathcal{S}}_t^{(u)}\right), \tag{16}$$

where $\mathrm{LOESS}(\cdot)$ denotes locally estimated scatterplot smoothing. The seasonal smoothing window parameters and trend smoothing window parameters are denoted as q_{s} and q_{t}, respectively. Here, s signifies the seasonal period, $\boldsymbol{\mathcal{T}}_t^{(u)}$ denotes the trend component from the previous iteration, and u represents the iteration number. Through iterative updates $\boldsymbol{\mathcal{S}}_t^{(u)}$ and $\boldsymbol{\mathcal{T}}_t^{(u)}$ until the convergence condition is met, a stable trend component $\boldsymbol{\mathcal{T}}_t$ is achieved, effectively removing the long-term patterns in the data while minimizing the impact of noise and short-term fluctuations. To further extract dynamic insights, the decomposed trend data $\boldsymbol{\mathcal{T}}_t$ are fed into an $\mathcal{L}$-layer feed-forward neural network (FNN) with the same output dimension as the original data dimension, enabling feature extraction and mapping of the trend data. This process aligns the trend data with the dimension of the original input sequence while generating the single-step predicted trend result sequence:

$$\hat{\boldsymbol{z}}_{t+1} = f_{\mathrm{pred}}\left(\boldsymbol{\mathcal{T}}_t; \boldsymbol{W}_{\mathrm{pred}}\right), \tag{17}$$

where the model comprises a linear transformation layer (f_{pred}) followed by an activation function of $\mathcal{L}$ layers using hyperbolic

tangent (tanh). $\boldsymbol{W}_{\text{pred}}$ is the weighting matrix. The output layer produces an m-dimensional prediction value $\hat{\boldsymbol{z}}_{t+1}$ that signifies the anticipated trend change at the subsequent time step.

4.4 Fine-grained anomaly detection

Anomalous behaviors in UAV flight data typically exhibit temporal continuity, where anomalous points demonstrate persistence within the time series. To address this, we propose an FAD framework. This framework generates anomaly scores through the simultaneous optimization of reconstruction and dynamic trend forecasting models, aimed at enhancing anomaly detection capabilities. Specifically, the proposed FAD framework incorporates two feature representations from the dynamic trend perception modules (reconstructed and predicted features) and integrates the streaming peaks over threshold (SPOT) method to determine thresholds for each dimension. The subsequent subsections provide a comprehensive explanation of the mathematical formulation and implementation specifics of the multimodal interaction network and the inference phase.

4.4.1 Fine-grained fusion network

The dynamic trend perception module produces reconstructed $\tilde{\boldsymbol{z}}_t$ and predicted features $\hat{\boldsymbol{z}}_{t+1}$ by employing multi-scale spatiotemporal feature reconstruction and dynamic trend forecasting. To effectively use these features, a fine-grained network is devised to create a cohesive representation by integrating maps and calculating error loss. Initially, $\tilde{\boldsymbol{z}}_t$ and $\hat{\boldsymbol{z}}_{t+1}$ are combined to produce $\boldsymbol{z}_t'$ with a dimension of m. Here, $\tilde{\boldsymbol{z}}_t$ denotes the reconstructed representation of the current time step, while $\hat{\boldsymbol{z}}_{t+1}$ is used as the forward dynamic trend context inferred from the historical window to construct the enhanced current representation $\boldsymbol{z}_t'$. The precise computational formula is outlined as follows:

$$\boldsymbol{z}_t' = \boldsymbol{W}_{\text{fuse}} \cdot \text{Concat}\left(\tilde{\boldsymbol{z}}_t, \hat{\boldsymbol{z}}_{t+1}\right) + \boldsymbol{b}_{\text{fuse}}, \tag{18}$$

where $\boldsymbol{z}_t' \in \mathbb{R}^m$, $\boldsymbol{W}_{\text{fuse}} \in \mathbb{R}^{m\times 2m}$, and $\boldsymbol{b}_{\text{fuse}} \in \mathbb{R}^m$. This fusion is motivated by the continuous evolution of UAV flight dynamics, where an abnormal state may cause both current reconstruction deviation and subsequent short-term trend inconsistency. By integrating reconstruction information with the forward trend context, the anomaly score can jointly reflect the current state deviation and short-term dynamic deviation. The mean square error $\mathbf{sc}_t$ of the fused representation $\boldsymbol{z}_t'$ and $\boldsymbol{z}_t$ is calculated by the mean square error loss function:

$$\mathbf{sc}_t = (\boldsymbol{z}_t' - \boldsymbol{z}_t) \odot (\boldsymbol{z}_t' - \boldsymbol{z}_t), \tag{19}$$

where the measure assesses the deviation of the model output $\mathbf{sc}_t \in \mathbb{R}^m$ from the normal data pattern, and $\odot$ denotes element-wise multiplication. In the training phase, the error sequence $\mathbf{sc}_{1,t}$ forms the basis of anomaly detection loss, which is used to refine the network parameters. Subsequently, during the testing phase, the error sequence $\mathbf{sc}_{2,t}$ functions as the outlier score for direct comparison with a threshold in the inference phase.

4.4.2 Inference phase

The inference phase involves threshold calculation and anomaly detection to promptly identify continuous anomalies in UAV flight data. We employ the SPOT algorithm, which is based on the sequential probability ratio test. The algorithm combines the peaks over threshold (POT) method from extreme value theory (EVT) with the generalized Pareto distribution (GPD) model to robustly model the tail of the anomaly score sequence. The core process is as follows: First, the entire anomaly score sequence is decomposed into three components—trend, seasonal, and residual—with potential anomaly information primarily residing in the residual term. Subsequently, we use the 99th percentile of the training set's anomaly scores as a preset threshold to filter the residuals, selecting all data points that exceed this value. These points are referred to as POT. Next, based on EVT, the SPOT algorithm assumes that these POT data points follow a GPD and fits its shape and scale parameters using maximum likelihood estimation. Finally, based on the fitted GPD model and a preset confidence level, the algorithm dynamically calculates the final anomaly threshold. This threshold defines the probability of an anomaly score exceeding it, providing a quantifiable and interpretable basis for anomaly detection. This approach is capable of accurately identifying rare events that significantly deviate from normal patterns while effectively ignoring the interference of regular data fluctuations. The threshold calculation formula is outlined as follows:

$$\boldsymbol{\tau}_{\text{s}} = \boldsymbol{\mu} + \boldsymbol{\gamma} g\left(F_t(\boldsymbol{x}_t; \boldsymbol{z}_t')\right), \tag{20}$$

where $\boldsymbol{\mu}$ and $\boldsymbol{\gamma}$ represent the mean and scale parameters of the distribution, respectively, $F_t(\cdot)$ represents the current distribution characteristics of the streaming data series, $g(\cdot)$ adjusts the tail of the distribution in EVT, and a fixed threshold $\boldsymbol{\tau}_{\text{s}} \in \mathbb{R}^m$ is generated for each dimension. Anomaly detection assesses the state by comparing the anomaly score $\mathbf{sc}_{2,t}$ with the fixed threshold $\boldsymbol{\tau}_{\text{s}}$ for each dimension. The calculation is as follows:

$$y_t' = \begin{cases} 1, & \text{if } ||\mathbf{sc}_{2,t}|| > ||\boldsymbol{\tau}_{\text{s}}||, \\ 0, & \text{otherwise}, \end{cases}$$

where $y_t' = 1$ and $y_t' = 0$ indicate that the UAV is in an abnormal or normal state, respectively.

5 Experiments

This section discusses the effectiveness of the proposed anomaly detection method using the real UAV dataset ALFA. Comparative experiments with existing methods are conducted to demonstrate the superiority of F-UAD in extracting single-dimensional and coupled features. This leads to enhanced detection accuracy and reduced false alarm rates.

5.1 Data description

The ALFA dataset is used in the experiments, collected in the real world by researchers at Carnegie Mellon University. It is available at https://theairlab.org/alfa-dataset/. This dataset uses customized Carbon-Z data from a T-28 fixed-wing UAV designed specifically for UAV failure and anomaly detection research (Fig. 4). Its flight environment is affected by

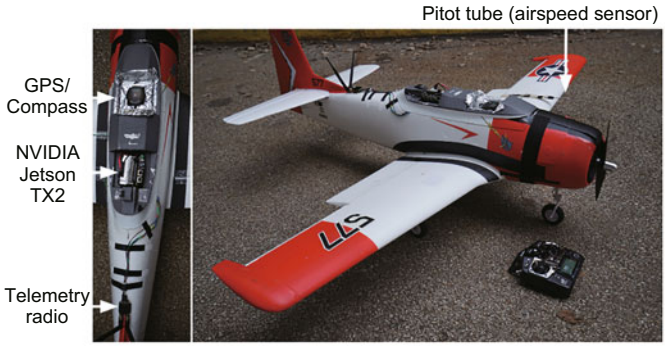

Fig. 4 ALFA AirLab failure dataset, developed by researchers at Carnegie Mellon University, using a customized Carbon-Z T-28 UAV

various environmental factors such as wind and air currents. The dataset comprises 47 autonomous flights, of which 10 are classified as normal and 37 as abnormal. Abnormal flights are caused by common optical control anomaly failures (engine, rudder, elevator, and aileron failures). During autonomous flight, the safety pilot can initiate a failure at any time by sending a command from the ground control station (GCS) to disable the control surfaces. For safety reasons, the safety pilot can also take over control at any time to restore the normal operation of all actuators and engines. The dataset includes various fault types, as detailed in Table 1. Each data point consists of 350 parameters, including timestamps, fault detection logs, and control signals, with a sampling frequency of no less than 5 Hz. Six datasets featuring rudder failures (stuck left and stuck right) and aileron failures (left zero and right zero) are designated as the test set for the experiments. Since this study focuses on anomaly detection related to the UAV lateral control system, we select rudder and aileron faults as the main test cases. These faults are directly associated with roll, yaw, angular velocity, and lateral control variables, making them suitable for evaluating the proposed dynamic trend perception and fine-grained anomaly scoring mechanisms. The training set comprises normal flight data, while other autonomous flight instances are allocated to the validation set. Following the feature selection methodology outlined in Section 4.2.1, 12 key parameters related to the lateral control system are identified as monitoring variables. These parameters encompass roll, pitch, and yaw angles, three-axis angular velocities, throttle commands, vertical velocity, and heading (Table 2). These selected features effectively capture the UAV's dynamic behaviors, serving as a robust foundation for anomaly detection.

5.2 Experimental setup

5.2.1 Performance evaluation

Anomaly detection refers to a model's ability to identify and classify anomalies in a dataset, where anomalies typically manifest as patterns or behaviors significantly deviating from normal data. This paper employs standard statistical metrics, e.g., Accuracy (Acc), Recall, and F1 score (Nguyen et al., 2021), to evaluate the proposed model's anomaly detection performance. These metrics comprehensively assess the model's ability to balance correct identification of anomalies

Table 1 Summary of fault types and time in the ALFA dataset

Fault type	Number of test cases	Fault time (s)
Engine full power loss	23	362
Rudder stuck to left	1	9
Rudder stuck to right	2	32
Elevator stuck at zero	2	23
Left aileron stuck at zero	3	183
Right aileron stuck at zero	4	231
Both ailerons stuck at zero	1	36
Rudder & aileron at zero	1	27
No-fault	10	–
Total	47	903

– denotes no fault time

Table 2 Input flight parameters for anomaly detection

Symbol	Parameter	Unit
ϕ	Roll angle	rad
θ	Pitch angle	rad
ψ	Yaw angle	rad
p	Roll rate	rad/s
q	Pitch rate	rad/s
r	Yaw rate	rad/s
v_u	Vertical velocity	m/s
h	Heading	°
a_x	X-axis acceleration	m/s^2
a_y	Y-axis acceleration	m/s^2
a_z	Z-axis acceleration	m/s^2
$\mathcal{T}$	Throttle	%

(true positives) with minimizing false positives. Their calculation formulae are as follows:

$$\mathrm{Acc} = \frac{\mathrm{TP} + \mathrm{TN}}{\mathrm{TP} + \mathrm{FP} + \mathrm{TN} + \mathrm{FN}}, \tag{21}$$

$$\mathrm{Recall} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}}, \tag{22}$$

$$\mathrm{F1} = \frac{2 \cdot \mathrm{Precision} \cdot \mathrm{Recall}}{\mathrm{Precision} + \mathrm{Recall}}, \tag{23}$$

where TP represents the number of correctly identified fault samples, TN represents the number of correctly identified normal samples, FP represents the number of normal samples identified as fault samples, FN represents the number of fault samples identified as normal samples, and Precision $=$ TP/(TP + FP). Acc serves as a comprehensive measure of model performance.

5.2.2 Baseline methods

To demonstrate the effectiveness of our proposed method, we conduct comparative analyses between F-UAD and various approaches, including (1) reconstruction-based models LSTM-AE+DT+DW (Bell et al., 2022), RSTAE (Jiang et al., 2024), Stacked-VAE (Park KH et al., 2021), Transformer (Vaswani et al., 2017), TranAD (Tuli et al., 2022), and Dual-TF (Nam et al., 2024), (2) the prediction-based model CAE-M (Zhang YX et al., 2023), (3) the autoregressive-based models MIC-LSTM (Zhou X et al., 2023) and STC-LSTM (Zhong et al., 2022), and (4) the contrastive learning model DCdetector (Yang YY et al., 2023).

5.2.3 Model design details

We implement the proposed method and all its variants using Python version 3.8.0 in conjunction with CUDA 12.2. All experiments are conducted on a computing platform with three NVIDIA GeForce RTX 3090 GPUs. For the F-UAD model, we configure a historical sliding window of 100 frames. The encoder and decoder of the multiscale spatiotemporal feature reconstruction model are each equipped with four contextual modeling units, with the multi-head attention mechanism employing eight attention heads. CNN uses three different kernel sizes: 3, 33, and 53. To enhance the model's expressive capacity, we incorporate an FNN with an input dimension expanded to 24 and an output dimension reduced to 12. A dropout regularization strategy is applied with a uniform dropout rate of 0.05 to mitigate overfitting. Given that the input data are normalized within the range $[-1, 1]$, the tanh activation function is employed to ensure that the output of each module remains within this range. During training, the batch size is set to 512, and the AdamW optimizer is used with an initial learning rate of 0.01, complemented by a learning rate scheduling strategy. The model is trained for a maximum of 400 epochs, with an early stopping mechanism adopted. Model parameters are configured for comparative experiments as specified in the original articles. Additionally, the same dataset is used, maintaining identical sample sizes and splits. Publicly available hyperparameters, including training strategies and learning rates, are consistently applied to ensure unbiased performance evaluation. It is noteworthy that, unlike the results of Dual-TF (Nam et al., 2024) and DCdetector (Yang YY et al., 2023), which incorporate a point adjustment (PA) strategy, our algorithm validation process excludes this approach to ensure a distinct evaluation framework.

5.3 Model comparison

The performance of the F-UAD model is systematically evaluated for anomaly detection. Comparative results are presented in Table 3, reporting Accuracy, F1 score, and Recall. The results demonstrate that the F-UAD model achieves SOTA-level performance on the F1 score, indicating its robustness and adaptability in handling normal and abnormal time-series representations. Beyond the overall performance improvement, the gains of F-UAD vary across different fault subsets due to differences in failure mechanisms and data characteristics. As shown in Table 1, rudder-related faults directly affect the yaw angle, yaw rate, and lateral-control variables; therefore, their abnormal patterns tend to produce clearer dynamic deviations in the selected flight parameters. In contrast, some aileron-related faults may be influenced by flight-control compensation, normal maneuvering behavior, or the fault evolution process, making the boundary between abnormal patterns and normal flight states less distinct. As a result, the improvement on these datasets is relatively small, although F-UAD still outperforms most baseline methods. This superior performance is primarily attributed to the fine-grained design of the F-UAD framework, which offers significant advantages over coarse-grained methods when processing UAV data. Unlike coarse-grained methods (such as LSTM-AE+DT+DW), which rely overly on macroscopic time windows and emphasize multidimensional coupling, suppressing the intrinsic characteristics of single-dimensional data (for example, achieving a recall rate of only 0.086 on the rudder 2 dataset), our fine-grained approach is primarily embodied in two core modules: multiscale feature modeling and FAD. The former effectively alleviates the suppression problem of single-dimensional data features by integrating multi-dimensional interactions, enabling the model to capture subtle changes that are difficult to detect with traditional methods. The latter collaboratively optimizes the reconstruction model and the dynamic trend prediction model design to generate anomaly scores, preserving the continuous nature of UAV anomaly data and accurately capturing subtle temporal interactions within the data. This design enables our method to capture deeper levels of scale and details in the data, identifying subtle anomalies that are often overlooked by coarse-grained methods. Therefore, our fine-grained approach surpasses traditional coarse-grained methods in terms of macro-performance indicators, showing significant advantages.

It is noteworthy that prediction and regression models generally exhibit lower performance in anomaly detection tasks compared to models based on reconstruction and contrast learning. Although prediction models can be optimized

Table 3 Comparison of Accuracy (Acc), F1 score, and Recall between our method and baseline methods for different anomaly detection tasks

Method	Rudder 1			Rudder 2			Rudder 3			Aileron 1			Aileron 2			Aileron 3		
	Acc	F1	Recall	Acc	F1	Recall	Acc	F1	Recall	Acc	F1	Recall	Acc	F1	Recall	Acc	F1	Recall
Transformer	0.445	0.429	0.955	0.397	0.453	0.977	0.200	0.334	0.802	0.155	0.205	0.686	0.483	0.339	0.919	0.261	0.414	0.676
LSTM-AE+DT+DW	0.171	0.244	0.638	0.935	0.444	0.086	0.081	0.149	0.323	0.589	0.285	0.516	0.141	0.236	0.956	0.259	0.406	0.655
Stacked-VAE	0.958	0.905	0.965	0.946	0.902	0.993	0.851	0.763	0.963	0.156	0.270	**0.979**	0.980	0.921	0.853	0.354	0.523	0.916
STC-LSTM	0.924	0.828	0.877	0.922	0.861	0.963	0.911	0.182	0.642	0.147	0.256	0.922	0.016	0.032	0.121	0.386	0.557	**0.999**
TranAD	0.951	0.894	0.918	0.994	0.998	0.997	0.897	0.739	0.586	0.481	0.174	0.141	0.685	0.098	0.108	0.481	0.174	0.141
MIC-LSTM	0.941	0.875	**0.981**	0.935	0.884	**0.999**	0.899	0.747	0.596	**0.833**	0.074	0.042	0.967	0.859	0.753	0.564	0.265	0.204
CAE-M	0.975	0.941	0.932	0.943	0.896	0.993	0.806	0.362	0.221	0.713	0.117	0.119	0.979	0.915	0.844	0.332	0.499	0.860
RSTAE	0.952	0.893	0.965	0.914	0.853	0.998	0.838	0.520	0.351	0.621	0.283	0.471	0.985	0.940	0.887	0.538	0.558	0.756
DCdetector	0.962	0.844	0.779	0.962	0.894	0.844	0.960	0.573	0.469	0.677	**0.687**	0.934	0.967	0.949	0.921	0.722	0.709	0.814
Dual-TF	0.979	0.953	0.959	0.950	0.912	0.988	**0.980**	0.933	**0.999**	0.164	0.282	0.164	0.614	0.659	0.919	0.681	0.651	0.986
F-UAD	**0.995**	**0.988**	0.975	**0.999**	**0.999**	**0.999**	0.967	**0.934**	0.928	0.753	0.511	0.809	**0.995**	**0.980**	**0.961**	**0.739**	**0.719**	0.863

The best results are in bold, and the sub-optimal results are underlined

through preprocessing or postprocessing, their performance may be more suitable for application scenarios where the highest accuracy is not required. In summary, the F-UAD model improves the expression of data features while preserving the data's inherent characteristics, significantly enhancing anomaly detection performance.

In addition to comparing the models' anomaly detection performance, we conduct a comprehensive comparison of the F-UAD model with other models using complexity metrics, including training time per epoch, GPU memory usage, and inference time. The results show that F-UAD exhibits certain differences in performance and complexity compared to other models, but remains within the same order of magnitude. As shown in Table 4, F-UAD has a training time of 139.45 s, GPU memory usage of 1792 MiB, and an inference time of 0.01 s. Compared to Transformer and TranAD, F-UAD exhibits slightly higher complexity. However, its training time aligns closely with DCdetector and Dual-TF, demonstrating a balance between complexity and performance. Although F-UAD is slightly more complex than the models we compare, the considerable performance gains it delivers for more comprehensive and accurate anomaly detection fully justify the computational overhead.

Table 4 Training time per epoch, GPU memory usage, and inference time of the model complexity comparison

Model	Training time (s)	GPU memory usage (MiB)	Inference time (s)
F-UAD	139.45	1792	0.0100
Transformer	23.63	1550	0.0039
TranAD	38.25	1554	0.0029
CAE-M	82.83	16 812	0.0050
RSTAE	51.99	1684	0.0012
DCdetector	113.39	1056	0.0330
Dual-TF	160.74	2241	0.0056

5.4 Anomaly score visualization comparison and analysis

The detection results of nine anomaly detection models on ALFA's rudder 1 dataset are visually presented to illustrate the detection capabilities of our model. An anomaly is identified by an anomaly score exceeding a predefined threshold, indicating a deviation between the model's output and the actual data. As depicted in Fig. 5, the anomaly score remains consistently low during normal periods (represented by the white region), while exhibiting significant fluctuations with higher values during abnormal periods (shown as the blue region), particularly evident in Fig. 5i. The data points below the threshold represent stationary phases, whereas those above it signify fluctuating phases. Unlike other models that produce false positives, our F-UAD method demonstrates effectiveness through multiscale spatiotemporal feature reconstruction, dynamic trend prediction mechanisms, and fine-grained anomaly score calculations. Particularly, F-UAD's multiscale feature modeling module captures subtle spatiotemporal variations in high-dimensional UAV data, overcoming the traditional suppression of the intrinsic characteristics of single-dimensional data; the dynamic trend prediction and fine-grained anomaly score calculation, through time-series decomposition and collaborative optimization, reduce fluctuations and precisely quantify deviations.

Furthermore, we employ t-distributed stochastic neighbor embedding (t-SNE) to further visualize the detection results on the rudder 1 dataset. As depicted in Fig. 6, the performances of four anomaly detection algorithms are compared. Purple indicates accurately identified normal data, red represents detected anomalous data, dark blue denotes false negatives (missed anomalies), and green signifies false positives (misclassified normal data). An ideal algorithm would exhibit a predominance of purple and red regions, with minimal presence of dark blue and green. Our analysis reveals that the Transformer model exhibits a relatively high false positive rate and a minor incidence of false negatives. In contrast, RSTAE demonstrates low rates of both false negatives and false positives. TranAD, while exhibiting minimal false positives, shows a notable incidence of false negatives, potentially due to its transformation-based prediction strategy, which may be limited in responding to rapid changes in dynamic anomaly scenarios. However, our F-UAD model significantly reduces false negatives while eliminating false positives, a result of its multiscale spatiotemporal feature reconstruction and dynamic trend prediction mechanisms, which enable a comprehensive capture of subtle data variations and optimize detection accuracy through fine-grained anomaly scoring. Collectively, the quantitative evaluations and visualization results substantiate that F-UAD outperforms the other models in terms of robustness and precision, providing a critical scientific basis for selecting appropriate algorithms tailored to specific application contexts.

5.5 Ablation study

In this subsection, we perform an ablation study on F-UAD to assess the model's effectiveness in incorporating multiscale feature reconstruction and dynamic trend data forecasting for spatiotemporal feature analysis.

5.5.1 Multiscale spatiotemporal feature reconstruction

To evaluate the significance of multiscale spatiotemporal feature reconstruction (MFC) in capturing data characteristics and impacting detection outcomes, we substitute them with single-scale models and conduct training from inception on each ALFA dataset. As shown in Fig. 7, the simplified model without the multiscale component exhibits significant performance degradation across all sub-datasets, with an average F1 score drop of 38.1%. Specifically, on the rudder 1 sub-dataset, the model without (w/o) the MFC module achieves an F1 score of only 0.553, while our full model (blue bars) achieves an F1 score of 0.988, nearly doubling the performance. On the aileron 1 sub-dataset, performance drops even more dramatically, from 0.511 to 0.197, intuitively demonstrating the critical role of the MFC module in processing these data. This significant performance drop directly demonstrates the indispensable role of the multiscale feature reconstruction module in handling the complexity and dynamic interactions of UAV flight data. The core of this improvement lies in the use of a parallel multiscale CNN architecture, which extracts features

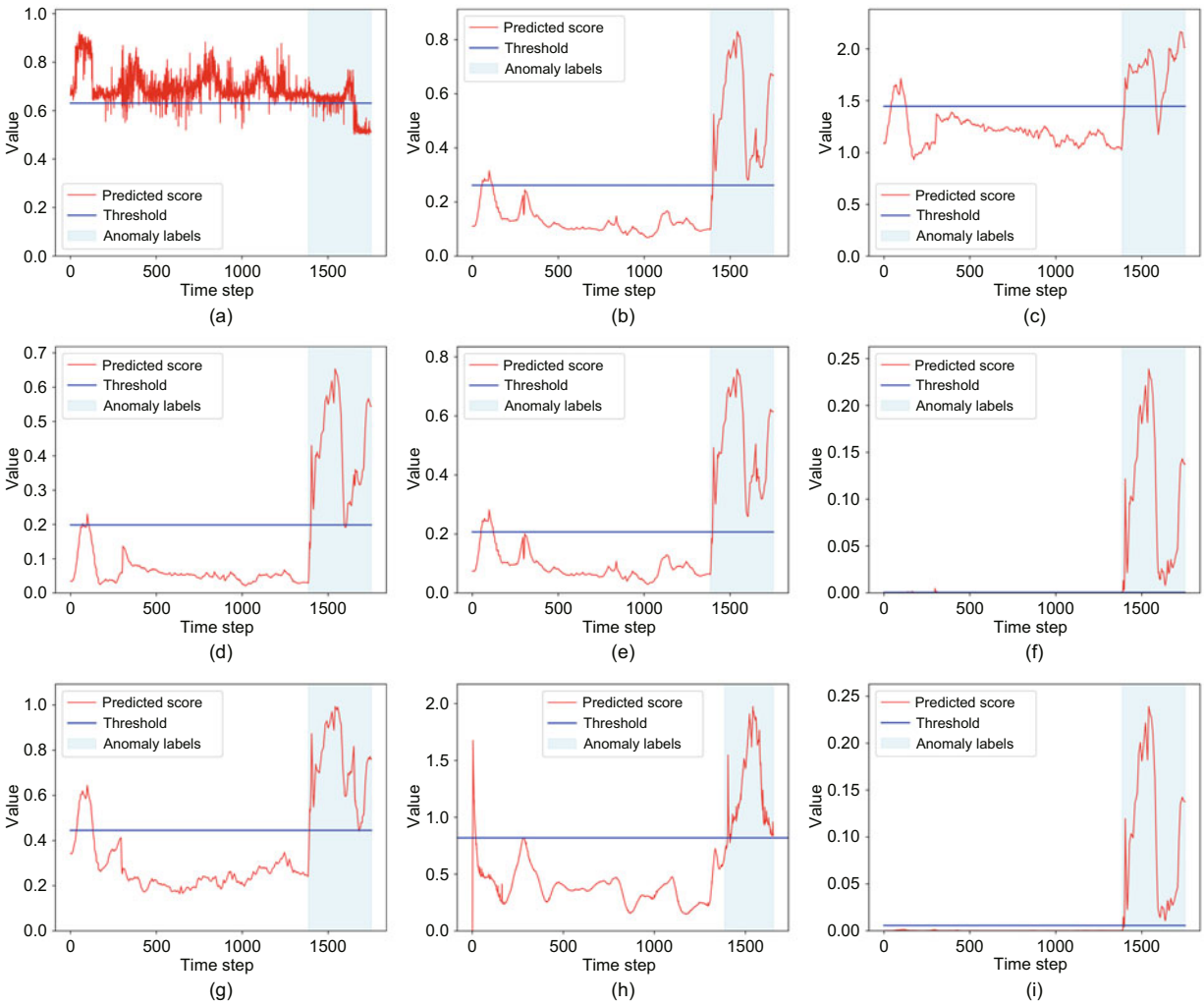

Fig. 5 Case studies of anomaly score visualization across rudder 1 dataset: (a) LSTM-AE+DT+DW; (b) Stacked-VAE; (c) STC-LSTM; (d) CAE-M; (e) RSTAE; (f) TranAD; (g) MIC-LSTM; (h) Dual-TF; (i) F-UAD

at different scales in parallel in the Transformer encoder and feeds the fused results into the decoder for final reconstruction. It is through this approach that our model not only captures fine-grained data features but also effectively integrates macro-trends and micro-fluctuations. This design enables the model to simultaneously preserve and model the inherent characteristics of each variable, avoiding information loss during cross-variable fusion. This approach constructs a more comprehensive spatiotemporal representation, significantly improving the efficiency and accuracy of anomaly detection, especially when processing highly complex and multiscale flight data.

5.5.2 Dynamic trend forecasting

To validate the accuracy of dynamic trend forecasting (DTF), we substitute the trend data derived from STL with direct predictions made from the original data. Correlation modules are trained from scratch on each ALFA dataset. As shown in Fig. 7, the simplified model that directly uses the original data for subsequent tasks experiences a significant drop in the F1 score on all sub-datasets, with an average drop of 23.9%. For example, on the aileron 1 sub-dataset, the model lacking the trend decomposition module achieves an F1 score of only 0.462, while our complete model achieves an F1 score of 0.511. This intuitively demonstrates the indispensability of trend-based prediction methods for such data. This analysis demonstrates that trend decomposition is crucial for accurate data prediction. The potential component plays a crucial role in a dataset. The core of this improvement lies in focusing on the long-term trend as the key prediction target, allowing the model to focus more on predicting smooth trends while effectively mitigating confounding factors such as seasonality and residuals. Not only does this approach enhance the model's ability to perceive complex dynamics, but also the decomposed trend component assists in reconstructing the module, significantly improving anomaly detection accuracy.

5.5.3 Fine-grained anomaly detection

As shown in Fig. 7, the simplified model relying solely on reconstruction loss for anomaly score calculation exhibits significant performance declines across all sub-datasets with an average F1 score drop of 30.7%. Specifically, on the rudder 1 sub-dataset, our model's F1 score is 0.114 higher than that of the simplified model without FAD, indicating a clear performance improvement; on the aileron 1 sub-dataset, this

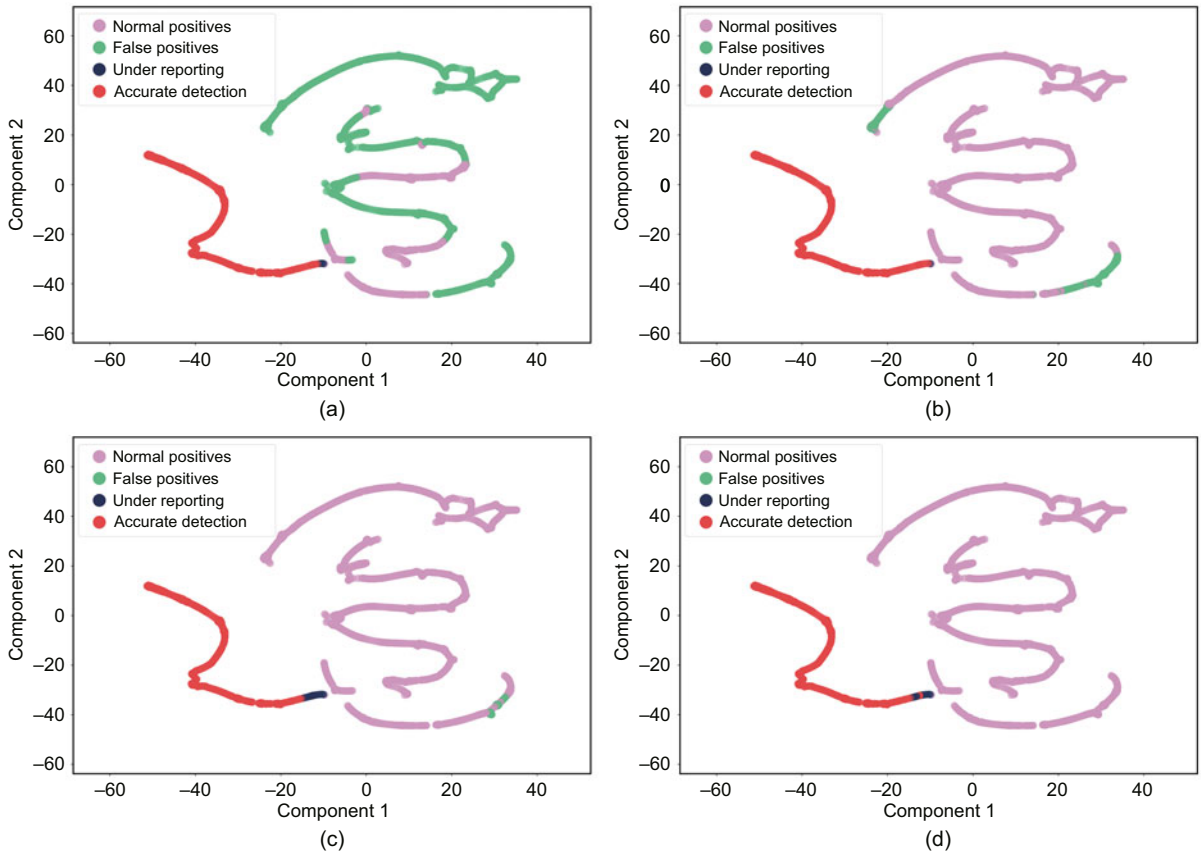

Fig. 6 ***t*-SNE visualization of a comparative experiment on the rudder 1 dataset: (a) Transformer; (b) TranAD; (c) RSTAE; (d) F-UAD**

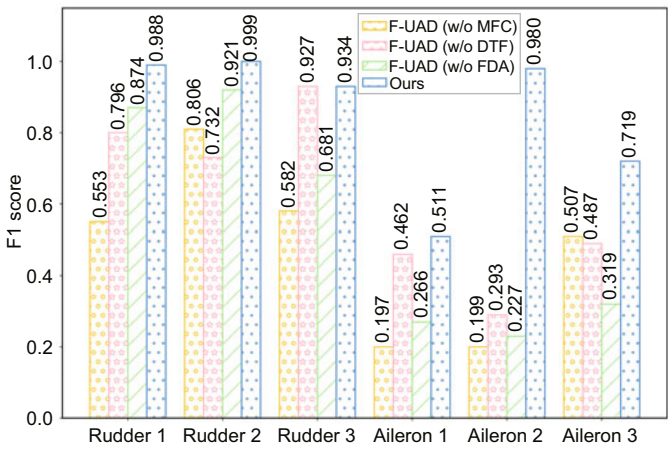

Fig. 7 Ablation study of multiscale spatiotemporal feature reconstruction (MFC), dynamic trend forecasting (DTF), and fine-grained anomaly detection (FAD)

advantage is even more pronounced with an F1 score increase of 0.245. This clearly demonstrates that the FAD approach plays a critical role in handling the contextual continuity of anomalies in UAV data. The core of this improvement lies in introducing the FAD calculation method, which integrates current moment reconstruction data with future moment prediction data to comprehensively compute anomaly perception scores, thus generating more forward-looking anomaly scores. Its innovation enhances the distinguishability between anomalous and normal data and adapts to complex spatiotemporal data changes. The model not only focuses on the current and past states but also predicts anomaly continuity, aligning with the practical needs of UAV anomaly occurrences. Compared to the traditional method that relies solely on reconstruction loss, FAD constructs a more comprehensive and differentiated anomaly score system through the joint optimization of reconstruction and prediction. This approach markedly enhances accuracy and reliability, exhibits strong adaptability in complex dynamic environments, and offers robust technical support for UAV monitoring.

5.6 Hyperparameter experiments

5.6.1 Window size

To verify the impact of window size on the performance of the F-UAD model, Fig. 8a presents the experimental results for three sub-datasets from the aileron dataset under varying window sizes. The results indicate that both small and excessively large window sizes lead to suboptimal performance across all three datasets. An excessively large window introduces significant noise, while a smaller window lacks sufficient relevant information. This observation aligns with the common practice in the field, where the initial window size is often set to around 60 or 100, based on the contextual nature of the data. These window sizes strike a balance between providing enough context for accurate anomaly detection and avoiding the inclusion of irrelevant information.

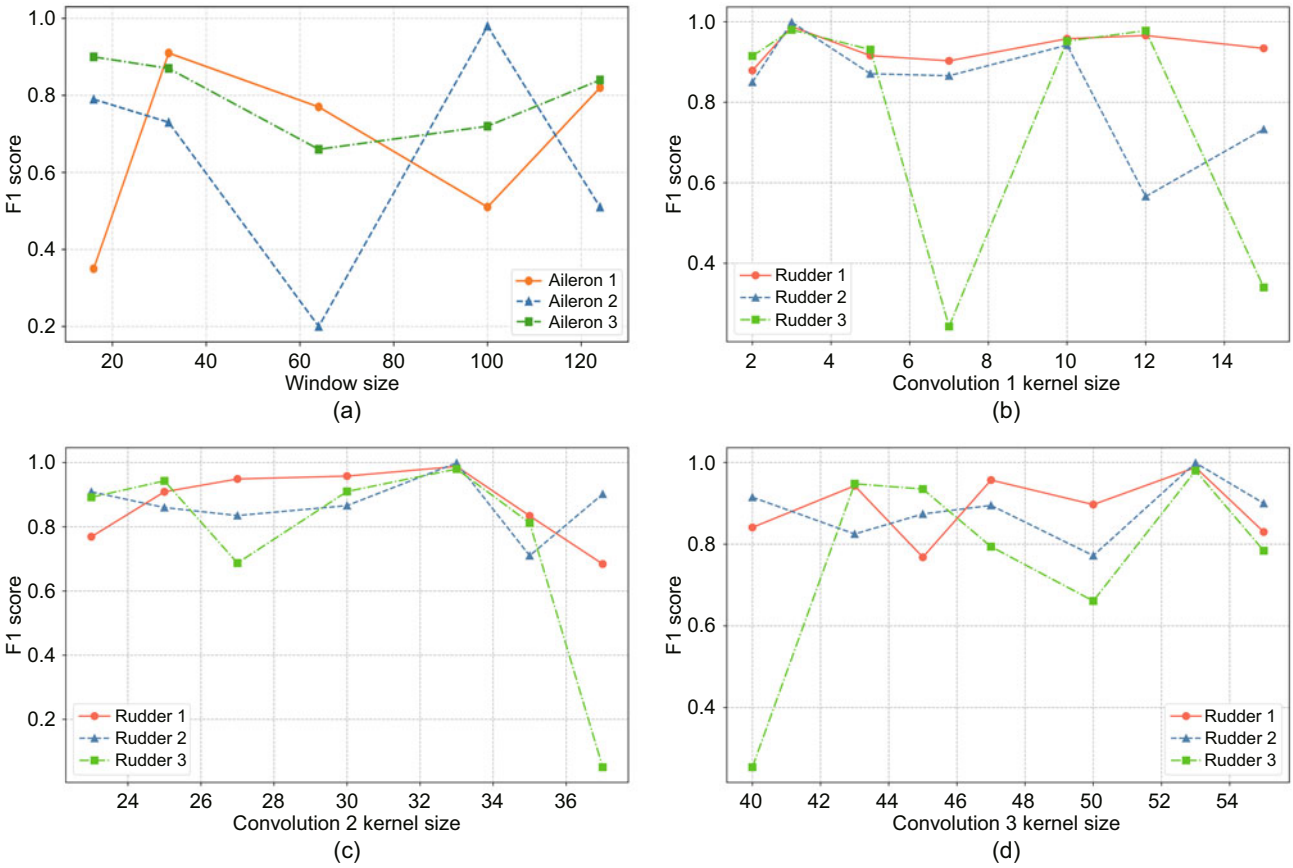

Fig. 8 Results of the hyperparameter experiments: (a) sliding window size; (b) convolution 1 kernel size; (c) convolution 2 kernel size; (d) convolution 3 kernel size

5.6.2 CNN kernel size

To assess the impact of three different CNN kernel sizes on the performance of the F-UAD model, Figs. 8b, 8c, and 8d display the experimental results of the three rudder subsets in the ALFA dataset. The F1 score serves as a crucial metric reflecting differences in detection performance. The findings indicate that the convolution kernel size significantly influences model performance. Specifically, intermediate-size kernels exhibit superior performance in anomaly detection tasks, yielding higher F1 scores and accuracy. Conversely, excessively small or too large kernel sizes may lead to inadequate feature capture or overfitting issues, resulting in a performance decline.

6 Conclusions

This paper proposes a dynamic trend perception framework for fine-grained UAV anomaly detection, named F-UAD. Through fine-grained design and dynamic trend prediction strategies, the framework effectively integrates the independent features of univariate flight data with the spatiotemporal coupling features of multidimensional data, significantly enhancing the model's ability to capture complex dynamics. Experiments on the ALFA real flight dataset show that F-UAD outperforms 10 SOTA models, demonstrating its effectiveness in UAV health monitoring and anomaly detection. Despite its effectiveness, this framework still has several limitations. First, its adaptability to different flight phases, such as takeoff, cruising, and landing, needs further improvement. Second, its robustness to real-time noise remains to be enhanced. Third, the current evaluation focuses mainly on lateral control faults in the ALFA dataset, such as rudder and aileron faults. In contrast, its effectiveness on longitudinal control faults and power system faults requires further validation. Future work will focus on improving flight-phase adaptability, extending the evaluation to broader UAV fault types, and enhancing real-time noise robustness.

Acknowledgments

This work was supported by the National Science and Technology Major Project on Intelligent Manufacturing Systems and Robots (No. 2025ZD1605703).

Author contributions

Yuanyuan YIN and Yuecen WEI designed the research. Jiahui LI processed the data. Qingran SU drafted the paper. Yongfeng YIN helped organize the paper. All the authors revised and finalized the paper.

Conflict of interest

All the authors declare that they have no conflict of interest.

Data availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Declaration on the use of generative AI tools

During the preparation of this work, the authors used ChatGPT to improve language. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the content of the published article.

References

Ahmad MW, Akram MU, Ahmad R, et al., 2022. Intelligent framework for automated failure prediction, detection, and classification of mission critical autonomous flights. *ISA Trans*, 129:355-371. https://doi.org/10.1016/j.isatra.2022.01.014

An J, Cho S, 2015. Variational autoencoder based anomaly detection using reconstruction probability. *Spec Lect IE*, 2(1):1-18.

Audibert J, Michiardi P, Guyard F, et al., 2020. USAD: unsupervised anomaly detection on multivariate time series. Proc 26th ACM SIGKDD Int Conf on Knowledge Discovery & Data Mining, p.3395-3404. https://doi.org/10.1145/3394486.340339

Baskaya E, Bronz M, Delahaye D, 2017. Fault detection & diagnosis for small UAVs via machine learning. Proc IEEE/AIAA 36th Digital Avionics Systems Conf, p.1-6. https://doi.org/10.1109/DASC.2017.8102037

Bell V, Rengasamy D, Rothwell B, et al., 2022. Anomaly detection for unmanned aerial vehicle sensor data using a stacked recurrent autoencoder method with dynamic thresholding. https://arxiv.org/abs/2203.04734

Bontemps L, Cao VL, McDermott J, et al., 2016. Collective anomaly detection based on long short-term memory recurrent neural networks. Proc 3rd Int Conf on Future Data and Security Engineering, p.141-152. https://doi.org/10.1007/978-3-319-48057-2_9

Braei M, Wagner S, 2020. Anomaly detection in univariate time-series: a survey on the state-of-the-art. https://arxiv.org/abs/2004.00433

Chandola V, 2009. Anomaly Detection for Symbolic Sequences and Time Series Data. PhD Dissertation, University of Minnesota, Minneapolis, USA.

Chen MZ, Zhang XC, Xiong XS, et al., 2021. Transformer: a multifunctional fast unmanned aerial vehicles–unmanned surface vehicles coupling system. *Machines*, 9(8):146. https://doi.org/10.3390/machines9080146

Chen YF, Wang BK, Liu W, et al., 2017. On-line and non-invasive anomaly detection system for unmanned aerial vehicle. Proc Prognostics and System Health Management Conf, p.1-7. https://doi.org/10.1109/PHM.2017.8079160

Choi K, Yi JH, Park C, et al., 2021. Deep learning for anomaly detection in time-series data: review, analysis, and guidelines. *IEEE Access*, 9:120043-120065. https://doi.org/10.1109/ACCESS.2021.3107975

D'Amato E, Mattei M, Notaro I, et al., 2018. UAV sensor FDI in duplex attitude estimation architectures using a set-based approach. *IEEE Trans Instrum Meas*, 67(10):2465-2475. https://doi.org/10.1109/TIM.2018.2838718

Darban ZZ, Webb GI, Pan SR, et al., 2025a. CARLA: self-supervised contrastive representation learning for time series anomaly detection. *Patt Recogn*, 157:110874. https://doi.org/10.1016/j.patcog.2024.110874

Darban ZZ, Webb GI, Pan SR, et al., 2025b. Deep learning for time series anomaly detection: a survey. *ACM Comput Surv*, 57(1):15. https://doi.org/10.1145/3691338

Deng AL, Hooi B, 2021. Graph neural network-based anomaly detection in multivariate time series. Proc 35th AAAI Conf on Artificial Intelligence, p.4027-4035. https://doi.org/10.1609/aaai.v35i5.16523

Garg A, Zhang WY, Samaran J, et al., 2022. An evaluation of anomaly detection and diagnosis in multivariate time series. *IEEE Trans Neur Netw Learn Syst*, 33(6):2508-2517. https://doi.org/10.1109/TNNLS.2021.3105827

Guo DF, Zhong MY, Zhou DH, 2018. Multisensor data-fusion-based approach to airspeed measurement fault detection for unmanned aerial vehicles. *IEEE Trans Instrum Meas*, 67(2):317-327. https://doi.org/10.1109/TIM.2017.2735663

Harrou F, Sun Y, Kadri F, et al., 2015. Early detection of abnormal patient arrivals at hospital emergency department. Proc Int Conf on Industrial Engineering and Systems Management, p.221-227. https://doi.org/10.1109/IESM.2015.7380162

Homayouni H, Ray I, Ghosh S, et al., 2021. Anomaly detection in COVID-19 time-series data. *SN Comput Sci*, 2(4):279. https://doi.org/10.1007/s42979-021-00658-w

Jiang GQ, Nan PC, Zhang JC, et al., 2024. Robust spatial-temporal autoencoder for unsupervised anomaly detection of unmanned aerial vehicle with flight data. *IEEE Trans Instrum Meas*, 73:3526014. https://doi.org/10.1109/TIM.2024.3428649

Keipour A, Mousaei M, Scherer S, 2019. Automatic real-time anomaly detection for autonomous aerial vehicles. Proc Int Conf on Robotics and Automation, p.5679-5685. https://doi.org/10.1109/ICRA.2019.8794286

Khosravi MR, Samadi S, 2021. BL-ALM: a blind scalable edge-guided reconstruction filter for smart environmental monitoring through green IoMT-UAV networks. *IEEE Trans Green Commun Netw*, 5(2):727-736. https://doi.org/10.1109/TGCN.2021.3067555

Li D, Chen DC, Jin BH, et al., 2019. MAD-GAN: multivariate anomaly detection for time series data with generative adversarial networks. Proc 28th Int Conf on Artificial Neural Networks, p.703-716. https://doi.org/10.1007/978-3-030-30490-4_56

Lin S, Clark R, Birke R, et al., 2020. Anomaly detection for time series using VAE-LSTM hybrid model. Proc IEEE Int Conf on Acoustics, Speech and Signal Processing, p.4322-4326. https://doi.org/10.1109/ICASSP40776.2020.9053558

Liu DT, Pang JY, Song G, et al., 2017. Fragment anomaly detection with prediction and statistical analysis for satellite telemetry. *IEEE Access*, 5:19269-19281. https://doi.org/10.1109/ACCESS.2017.2754447

Liu DT, Song YC, Li L, et al., 2018. Online life cycle health assessment for lithium-ion battery in electric vehicles. *J Cleaner Prod*, 199:1050-1065. https://doi.org/10.1016/j.jclepro.2018.06.182

Liu PF, Sun XM, Han Y, et al., 2022. Arrhythmia classification of LSTM autoencoder based on time series anomaly detection. *Biomed Signal Process Contr*, 71:103228. https://doi.org/10.1016/j.bspc.2021.103228

Maddikunta PKR, Hakak S, Alazab M, et al., 2021. Unmanned aerial vehicles in smart agriculture: applications, requirements, and challenges. *IEEE Sens J*, 21(16):17608-17619. https://doi.org/10.1109/JSEN.2021.3049471

Marcos A, Waitman S, Sato M, 2022. Fault tolerant linear parameter varying flight control design, verification and validation. *J Franklin Inst*, 359(2):653-676. https://doi.org/10.1016/j.jfranklin.2021.02.040

Nam Y, Yoon S, Shin Y, et al., 2024. Breaking the time-frequency granularity discrepancy in time-series anomaly detection. Proc ACM Web Conf, p.4204-4215. https://doi.org/10.1145/3589334.3645556

Narasimhan S, Brownston L, 2007. HyDE: a general framework for stochastic and hybrid model-based diagnosis. Proc 18th Int Workshop on Principles of Diagnosis, p.186-193.

Nguyen HD, Tran KP, Thomassey S, et al., 2021. Forecasting and anomaly detection approaches using LSTM and LSTM autoencoder techniques with the applications in supply chain management. *Int J Inform Manag*, 57:102282. https://doi.org/10.1016/j.ijinfomgt.2020.102282

Oussous A, Benjelloun FZ, Ait Lahcen A, et al., 2018. Big data technologies: a survey. *J King Saud Univ-Comput Inform Sci*, 30(4):431-448. https://doi.org/10.1016/j.jksuci.2017.06.001

Pang JY, Liu DT, Peng Y, et al., 2015. Detecting continual anomalies in monitoring data stream based on sampling GPR algorithm. Proc IEEE Conf on Prognostics and Health Management, p.1-7. https://doi.org/10.1109/ICPHM.2015.7245071

Park JH, Shanbhag S, Chang DE, 2022. Model-free unsupervised anomaly detection of a general robotic system using a stacked LSTM and its application to a fixed-wing unmanned aerial vehicle. Proc IEEE/RSJ Int Conf on Intelligent Robots and Systems, p.4287-4293. https://doi.org/10.1109/IROS47612.2022.9981950

Park KH, Park EJ, Kim HK, 2021. Unsupervised fault detection on unmanned aerial vehicles: encoding and thresholding approach. *Sensors*, 21(6):2208. https://doi.org/10.3390/s21062208

Qin HM, Yan MR, Ji HJ, 2021. Application of controller area network (CAN) bus anomaly detection based on time series prediction. *Veh Commun*, 27:100291. https://doi.org/10.1016/j.vehcom.2020.100291

Radoglou-Grammatikis P, Sarigiannidis P, Lagkas T, et al., 2020. A compilation of UAV applications for precision agriculture. *Comput Netw*, 172:107148. https://doi.org/10.1016/j.comnet.2020.107148

Tuli S, Casale G, Jennings NR, 2022. TranAD: deep Transformer networks for anomaly detection in multivariate time series data. *Proc VLDB Endowm*, 15(6):1201-1214. https://doi.org/10.14778/3514061.3514067

Ullah S, Kim KI, Kim KH, et al., 2019. UAV-enabled healthcare architecture: issues and challenges. *Fut Gener Comput Syst*, 97:425-432. https://doi.org/10.1016/j.future.2019.01.028

Vaswani A, Shazeer N, Parmar N, et al., 2017. Attention is all you need. Proc 31st Int Conf on Neural Information Processing Systems, p.6000-6010.

Wang T, Qiao MN, Lin ZW, et al., 2019. Generative neural networks for anomaly detection in crowded scenes. *IEEE Trans Inform Forens Secur*, 14(5):1390-1399. https://doi.org/10.1109/TIFS.2018.2878538

Wu HX, Xu JH, Wang JM, et al., 2021. Autoformer: decomposition Transformers with auto-correlation for long-term series forecasting. Proc 35th Int Conf on Neural Information Processing Systems, Article 1717.

Xiong WX, Wang P, Sun XC, et al., 2024. SiET: spatial information enhanced Transformer for multivariate time series anomaly detection. *Knowl-Based Syst*, 296:111928. https://doi.org/10.1016/j.knosys.2024.111928

Yang L, Li SB, Li CJ, et al., 2023. A survey of unmanned aerial vehicle flight data anomaly detection: technologies, applications, and future directions. *Sci China Technol Sci*, 66(4):901-919. https://doi.org/10.1007/s11431-022-2213-8

Yang L, Li SB, Zhu CC, et al., 2024. Spatio-temporal correlation-based multiple regression for anomaly detection and recovery of unmanned aerial vehicle flight data. *Adv Eng Inform*, 60:102440. https://doi.org/10.1016/j.aei.2024.102440

Yang YY, Zhang CL, Zhou T, et al., 2023. DCdetector: dual attention contrastive representation learning for time series anomaly detection. Proc 29th ACM SIGKDD Conf on Knowledge Discovery and Data Mining, p.3033-3045. https://doi.org/10.1145/3580305.3599295

Yi YM, Zhang YM, 2017. Fault diagnosis of an unmanned quadrotor helicopter based on particle filter. Proc Int Conf on Unmanned Aircraft Systems, p.1432-1437. https://doi.org/10.1109/ICUAS.2017.7991322

Yin CY, Zhang S, Wang J, et al., 2022. Anomaly detection based on convolutional recurrent autoencoder for IoT time series. *IEEE Trans Syst Man Cybern Syst*, 52(1):112-122. https://doi.org/10.1109/TSMC.2020.2968516

Zhang QC, Yang LT, Chen ZK, et al., 2018. A survey on deep learning for big data. *Inform Fus*, 42:146-157. https://doi.org/10.1016/j.inffus.2017.10.006

Zhang YX, Chen YQ, Wang JD, et al., 2023. Unsupervised deep anomaly detection for multi-sensor time-series signals. *IEEE Trans Knowl Data Eng*, 35(2):2118-2132. https://doi.org/10.1109/TKDE.2021.3102110

Zhao H, Wang YJ, Duan JY, et al., 2020. Multivariate time-series anomaly detection via graph attention network. Proc IEEE Int Conf on Data Mining, p.841-850. https://doi.org/10.1109/ICDM50108.2020.00093

Zhong J, Zhang YJ, Wang JY, et al., 2022. Unmanned aerial vehicle flight data anomaly detection and recovery prediction based on spatio-temporal correlation. *IEEE Trans Rel*, 71(1):457-468. https://doi.org/10.1109/TR.2021.3134369

Zhou B, Liu SH, Hooi B, et al., 2019. BeatGAN: anomalous rhythm detection using adversarially generated time series. Proc 28th Int Joint Conf on Artificial Intelligence, p.4433-4439. https://doi.org/10.24963/ijcai.2019/616

Zhou X, Chu XY, Zou YQ, 2023. Sensor fault detection for UAVs based on MIC-LSTM with attention mechanism. Proc CAA Symp on Fault Detection, Supervision and Safety for Technical Processes, p.1-6. https://doi.org/10.1109/SAFEPROCESS58597.2023.10295739