

PHYSICS-INFORMED TELEMETRY SIMULATION AND SHAP-BASED EXPLAINABLE MACHINE LEARNING FOR MULTI-SENSOR UAV FAULT DIAGNOSIS

Aswin Karkadakattil  0009-0004-6141-0545

Mechanical Engineer, Independent Researcher, India

ABSTRACT

Unmanned Aerial Vehicles (UAVs) rely on tightly coupled onboard sensors for stable and safe flight; however, early-stage fault detection remains challenging due to noisy telemetry, nonlinear dynamics, and limited labelled fault data. This study presents a simulation-driven, explainable machine-learning framework for multi-sensor UAV fault diagnosis within digital-twin environments. A physics-informed telemetry generator produces synchronized accelerometer, gyroscope, motor RPM, battery voltage, and GPS signals with systematic injection of representative fault modes, including motor degradation, battery sag, IMU drift, and GPS anomalies. The multivariate time series is segmented using overlapping windows and transformed into a 60-dimensional statistical feature space designed for interpretability and computational efficiency. A Random Forest classifier achieves 99.33% test accuracy and an AUC of 0.997 under controlled simulation conditions. SHAP-based explainability quantifies feature-level contributions, highlighting the dominant influence of gyroscope drift and battery-voltage descriptors consistent with known degradation mechanisms. The framework provides a lightweight, reproducible, and interpretable diagnostic pipeline for simulation-based verification and early-stage UAV health-monitoring development.

Keywords

UAV fault detection; Explainable Artificial Intelligence (XAI); SHAP analysis; Random Forest classifier; Telemetry simulation; Time-series feature extraction; Aerospace system reliability.

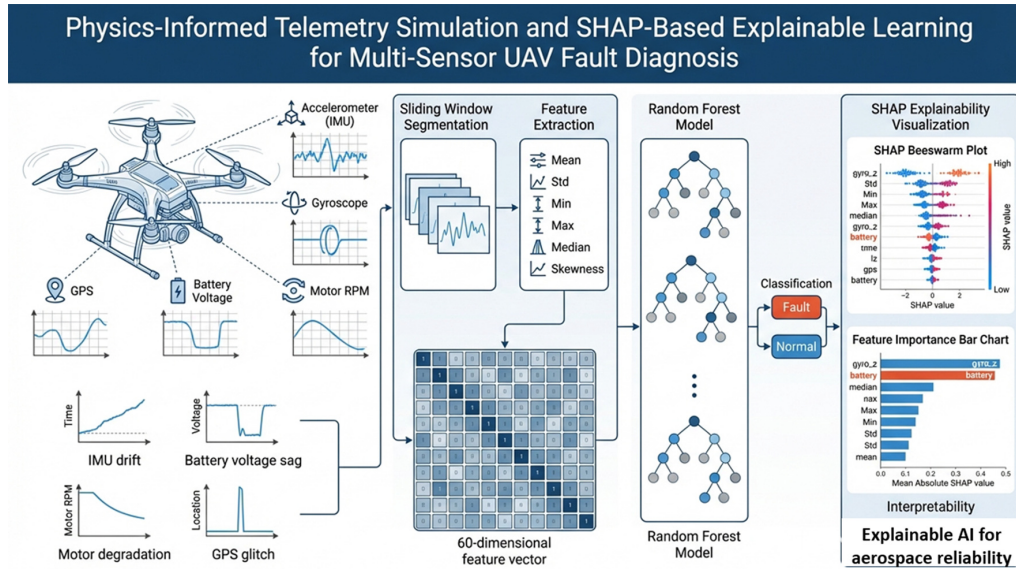
Aswin Karkadakattil is a postgraduate in Materials and Manufacturing Engineering from the Indian Institute of Technology (IIT) and holds a B.Tech in Mechanical Engineering from the Government College of Engineering, Kannur. He was a recipient of the Prime Minister's Scholarship for Academic Excellence during his undergraduate studies. He has authored and co-authored several Scopus- and SCI-indexed research publications spanning aerospace engineering, advanced manufacturing, and materials processing. His research interests include aerospace propulsion systems, hydrogen-electric and hybrid energy technologies, laser-based post-processing of advanced materials, AI-driven manufacturing systems, and hybrid optimization frameworks for intelligent and sustainable engineering innovation.

* Corresponding Author. Email address: ashwinharik20000@gmail.com

Received 26.01.2026 • Revised 12.03.2026 • Accepted 16.03.2026

© 2026 Author[s]. Article category: research article





1. INTRODUCTION

Unmanned Aerial Vehicles (UAVs) are increasingly deployed across civilian, industrial, and defence applications, where reliable operation under dynamic and uncertain environmental conditions is essential. Safe flight depends on the coordinated performance of multiple onboard subsystems, including inertial measurement units (IMUs), motor and power controllers, battery management systems, and satellite-based navigation modules. Faults in these components – such as motor degradation, battery voltage sag, gyroscope drift, or GPS inconsistencies – often emerge gradually and may initially be masked by nominal operational noise. If undetected, such degradations can accumulate and ultimately lead to instability, performance deterioration, or mission failure. Early-stage fault detection in UAVs remains challenging due to the noisy and strongly coupled nature of telemetry signals, as well as the nonlinear and non-stationary flight dynamics arising from manoeuvres and environmental disturbances.

Traditional fault-diagnosis strategies, including fixed-threshold techniques and model-based observers, perform adequately under controlled conditions, but frequently degrade in the presence of complex sensor interactions and varying mission profiles [1–5]. As UAV platforms progress toward higher autonomy, there is a growing demand for data-driven diagnostic methods capable of identifying subtle and evolving deviations prior to critical failure. Machine learning (ML) approaches have demonstrated strong potential in UAV sensing, navigation, and control applications [6–10]. However, many high-performing ML models operate as black-box systems, providing limited transparency into the rationale behind their predictions. In safety-critical aerospace systems, interpretability is essential, as

diagnostic decisions must be physically traceable and verifiable to support validation and certification processes [11–14]. Furthermore, the availability of labelled UAV fault datasets remains limited. Deliberate fault induction during flight testing is costly and risky, potentially resulting in hardware damage and safety hazards, and is often constrained by commercial or defence regulations. Simulation-driven development therefore represents a practical and widely adopted strategy for early-stage algorithm design, verification, and sensitivity analysis [15–17]. Beyond sensor-level diagnostics, aerospace research also emphasizes the broader importance of structural reliability and material performance in military aircraft systems, particularly in the context of composite materials used in safety-critical platforms [18]. Such considerations further underscore the need for transparent and robust health-monitoring frameworks in autonomous aerial systems.

Nevertheless, much of the existing literature either concentrates on isolated sensor channels or applies data-driven fault detection without ensuring transparent and physically interpretable reasoning. More importantly, relatively few studies present a unified workflow that systematically integrates:

- (1) realistic multi-sensor telemetry simulation,
- (2) structured time-series feature extraction,
- (3) robust machine-learning-based fault classification, and
- (4) transparent, physically meaningful explainability.

To address this gap, the present study introduces a unified and fully explainable machine-learning framework for UAV fault diagnosis, specifically designed for early-stage system development and digital-twin environments. The proposed framework integrates:

- physics-inspired simulation of motor, battery, IMU, and GPS fault modes;
- a sliding-window statistical representation comprising 60 interpretable descriptors;
- a Random Forest classifier well-suited to noisy and heterogeneous telemetry; and
- SHAP-based explainability to attribute diagnostic decisions to specific sensor behaviours.

Novelty and Contribution

In contrast to prior studies that focus either on narrow subsystem datasets or opaque deep-learning architectures, the primary contribution of this work lies in the system-level integration of simulation-driven multi-sensor fault generation with interpretable machine learning and physically grounded explainability. Rather than emphasizing raw predictive performance, the proposed framework establishes a lightweight, reproducible, and transparent diagnostic baseline that enables engineers to understand which sensor behaviours drive fault detection and how

diagnostic confidence evolves over time. This emphasis on traceability and interpretability directly addresses critical requirements in aerospace applications, particularly during pre-flight verification, digital-twin analysis, and the early stages of UAV health-monitoring system design. Furthermore, the structured integration of simulation, statistical feature engineering, ensemble learning, and explainability provides a foundation that can be extended toward real-flight validation studies, hybrid simulation–field data calibration, and certification-aligned diagnostic system development.

1.1 Related work

UAV fault detection has evolved significantly as autonomous aerial systems have become increasingly complex and safety-critical. Early approaches relied primarily on model-based observers and threshold-based rules, where sensor deviations were compared against predefined analytical models [1–5]. Although effective for well-characterized failure modes, these methods exhibit reduced robustness under nonlinear flight dynamics, variable mission conditions, and disturbances such as delays and jitter [6,7]. The limitations of handcrafted rules motivated the adoption of data-driven approaches, including classical machine-learning classifiers and deep-learning architectures for detecting anomalies in IMU signals, motor behaviour, battery health indicators, and actuator responses [8–14]. Recent studies demonstrate the effectiveness of deep learning for motor fault detection [19], actuator damage diagnosis [20], and multi-sensor UAV fault detection [21–23]. More recent experimental investigations have explored multi-sensor-based propeller damage detection using deep learning algorithms [24], while signal-processing-enhanced methodologies have been proposed for improved UAV propeller fault localization [25]. Mobility-aware fusion strategies combining vision and vibration data have also been investigated for UAV health monitoring [26], and machine-learning-driven optimization of multi-sensor fusion frameworks has been explored to enhance state estimation robustness and fault resilience [27].

Comprehensive surveys further highlight the growing role of ML in UAV fault diagnosis while emphasizing persistent challenges related to generalization and interpretability [28]. Despite their predictive capability, many ML-based diagnostic systems remain opaque, complicating validation and certification in safety-critical aerospace contexts [29–31]. Explainable Artificial Intelligence (XAI) has therefore emerged as a promising direction for enhancing transparency and trust in learning-based systems. Among XAI methods, SHAP (SHapley Additive exPlanations) has emerged as a leading approach, using concepts from cooperative game theory to assign each input feature a contribution value that quantifies its marginal influence on a given model prediction. SHAP-based interpretability has been successfully applied to aerospace sensor diagnostics [32], industrial fault detection [33–35], and predictive maintenance applications [36]. However, the integration of XAI within

UAV fault-diagnosis pipelines remains limited. Existing UAV-focused ML/XAI studies often concentrate on isolated sensor subsets, restricted fault types, or post-hoc explanations without controlled multi-sensor fault generation [14,19,28]. Consequently, a clear research gap persists at the intersection of UAV multi-sensor telemetry simulation, structured time-series feature learning, robust classification, and physically grounded explainability within a unified and reproducible framework. Addressing this gap forms the central motivation of the present work.

1.2 Digital Twin Alignment Perspective

The proposed framework is closely aligned with digital-twin-based UAV development workflows, where high-fidelity simulation environments are used to replicate, monitor, and refine physical system behaviour. In such architectures, simulation is not limited to pre-flight design validation but forms part of a continuous health-monitoring ecosystem. The integration of controlled multi-sensor fault injection, structured feature extraction, and interpretable classification enables systematic evaluation of subsystem degradation within a digital twin environment. Because both nominal and faulty states can be simulated in a controlled and repeatable manner, the framework supports sensitivity analysis, interpretability validation, and incremental model refinement prior to field deployment. Moreover, the modular structure of the proposed pipeline allows for the progressive incorporation of real-flight telemetry as it becomes available. In a digital twin architecture, simulated degradation patterns can be recalibrated using observed operational data, while diagnostic models can be retrained or fine-tuned accordingly. This bidirectional interaction between virtual and physical domains strengthens consistency between the virtual and physical UAV states. By embedding explainable machine learning within a simulation-driven architecture, the framework provides a structured foundation for digital-twin-supported UAV health monitoring and iterative system validation.

1.3 Literature Review

The endurance limitations of small unmanned aerial vehicles (UAVs) have been widely documented, with conventional lithium-based batteries offering specific energies typically below 250 Wh kg^{-1} , insufficient for sustained long-duration missions such as surveillance, mapping, and reconnaissance [1–3]. This inherent constraint has motivated extensive research into hybrid and hydrogen-assisted propulsion systems, which provide substantially higher specific energy while enabling low-emission or zero-emission operation across a wide range of **small-UAV scales** [4–6]. Experimental demonstrations and conceptual design studies have shown that hydrogen fuel-cell systems can extend UAV flight times well beyond those achievable with battery-only configurations, although challenges remain related to hydrogen

storage mass, fuel-cell integration, and onboard energy management [2,4,8–10]. Investigations by Tak et al. [1], Özbek et al. [2], and Farajollahi and Dincer [3], among others, present design methodologies, propulsion-system architectures, and hybrid-electric performance assessments that highlight the potential for multi-hour endurance gains.

Complementary studies have explored advanced hybrid concepts, including solid-oxide fuel-cell (SOFC) architectures, thermoelectric integration, and multi-source energy scheduling using model-predictive control strategies [4,8]. Collectively, these works demonstrate the growing interest in hydrogen-assisted hybrid propulsion as a pathway toward long-endurance and energy-efficient UAV operation.

Parallel advances in UAV aerodynamics, structural optimization, and propulsion-system design have further strengthened the foundation for hybrid-electric development. Research on energy-efficient propulsion technologies [5], UAV propulsion-design strategies [6], and time-varying endurance modelling [7] has provided valuable insight into aerodynamic-propulsive coupling, power-management strategies, and mission-level performance estimation. Recent studies focusing on hybrid-UAV control, aerodynamic optimization, and integrated system analysis further emphasize the importance of modelling frameworks that explicitly link energy availability with aerodynamic requirements [9,10,28–30]. Alongside these developments, a substantial body of classical aircraft-performance literature continues to offer analytical tools that remain highly relevant to UAV endurance analysis. Foundational works by Filippone, Raymer, Mason, Torenbeek, and others [11,15–27] present closed-form formulations for lift, drag, power required, cruise performance, and optimal operating conditions – relationships that have long been used in both manned and unmanned aircraft conceptual design. Because these models are grounded in first-principles aerodynamics rather than empirical curve fitting or numerical simulation, they have proven particularly effective for early-stage sizing and performance prediction. More recent efforts in open aircraft-performance modelling and trajectory-analysis frameworks, such as LEAPS and data-driven mission parameterization approaches, further demonstrate how analytical and semi-analytical methods can support rapid and generalizable performance estimation across diverse aircraft configurations [18,19]. In the UAV context, studies addressing drag estimation, low-speed aerodynamic performance, and airframe-configuration optimization highlight how even modest changes in aerodynamic characteristics can significantly influence endurance-oriented missions [28–34].

Despite this progress, many existing hybrid-UAV studies remain heavily dependent on CFD, numerical optimization, or prototype-level experimentation. While powerful, these approaches are often computationally demanding, geometry-specific, and less suited to broad design-space exploration during early conceptual phases. As a result, there remains a clear need for analytical, first-principles-based frameworks that can rapidly and transparently predict endurance and range without

reliance on CFD or experimental calibration. The present work addresses this gap by developing a fully analytical, physics-driven model for predicting endurance and range in hybrid hydrogen-electric UAVs using only fundamental aerodynamic and energetic relationships. By avoiding empirical curve fitting and simulation-heavy techniques, the framework provides a deterministic and reproducible means of exploring mass-energy trade-offs, propulsion-architecture choices, and conceptual design trends. In doing so, it bridges the gap between simplified theoretical approximations and complex numerical models, offering a practical and openly verifiable tool for early-stage, sustainability-oriented UAV propulsion studies grounded in first-principles physics.

2. METHODOLOGY

This study proposes an integrated and fully explainable framework for UAV fault diagnosis designed for early-stage system analysis under realistic flight dynamics and multi-sensor variability. The methodology combines physics-informed telemetry simulation, structured time-series processing, interpretable statistical feature engineering, ensemble-based classification, and post-hoc explainability. The framework is designed to remain transparent, reproducible, and computationally efficient, supporting simulation-driven verification and digital-twin-based health-monitoring development.

The workflow consists of five tightly coupled stages:

1. **Telemetry simulation** – Generation of realistic multi-sensor UAV telemetry with controlled injection of representative subsystem fault modes.
2. **Sliding-window segmentation** – Partitioning of continuous telemetry into overlapping windows to capture both transient disturbances and gradual degradation patterns.
3. **Statistical feature extraction** – Computation of 60 interpretable statistical descriptors summarizing sensor behaviour within each window.
4. **Random Forest classification** – Ensemble learning to distinguish nominal and faulty operating states across the engineered feature space.
5. **SHAP-based explainability** – Quantification of feature-level contributions to ensure transparent and physically interpretable diagnostic decisions.

The following subsections detail each stage and demonstrate how they collectively establish an interpretable and engineering-aligned UAV fault-diagnosis pipeline.

2.1 Overall Pipeline

The overall methodological flow is illustrated in Figure 1. The process begins with the development of a physics-informed telemetry simulator capable of generating synchronized multivariate time-series data representing accelerometer, gyroscope, motor RPM, battery voltage, and GPS signals. Controlled fault profiles – motor degradation, battery voltage sag, IMU drift, and GPS anomalies – are systematically injected to emulate representative operational failure scenarios. The multivariate telemetry is segmented using a fixed-length sliding window with overlap. This strategy preserves temporal continuity while enabling detection of both short-duration anomalies and slowly evolving subsystem degradation.

Each window is transformed into a 60-dimensional statistical feature vector capturing central tendency, variability, amplitude extrema, and distribution shape across all sensor channels. Statistical feature engineering is selected for its interpretability, computational efficiency, and robustness to noise – all of which are characteristics essential for safety-critical aerospace applications. Unlike end-to-end deep learning approaches that require large labelled datasets and offer limited transparency, this representation allows direct physical interpretation of sensor behaviour while retaining strong discriminative capacity.

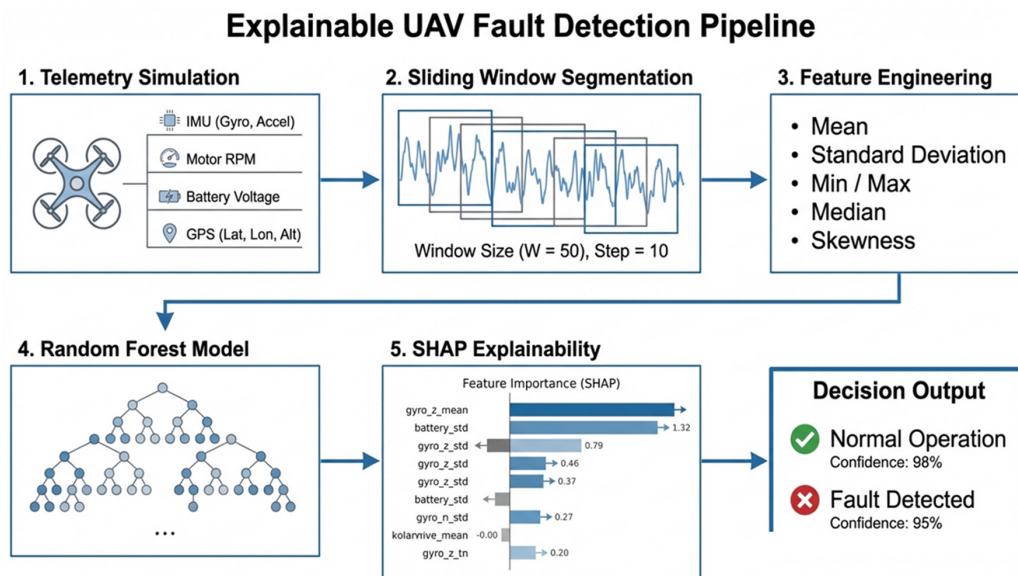


Fig.1. Full pipeline diagram illustrating the data flow from simulated UAV telemetry to windowing, feature engineering, machine-learning-based fault detection, and SHAP explainability.

The engineered feature vectors serve as inputs to a Random Forest classifier. The ensemble architecture is chosen for its ability to capture nonlinear feature interactions, accommodate heterogeneous sensor distributions, and maintain stable

decision boundaries under noisy and moderately imbalanced conditions. The model outputs a binary classification indicating nominal or faulty operation at the window level. To ensure interpretability, SHAP-based explainability is integrated into the pipeline. SHAP values quantify the marginal contribution of each feature to individual predictions, enabling explicit identification of the sensor statistics driving fault detection. This allows interpretation of how behaviours such as progressive gyroscope drift or abrupt battery-voltage deviations influence classification outcomes. Collectively, the simulation environment, structured feature representation, ensemble classifier, and explainability layer form a transparent and reproducible methodology for UAV fault diagnosis. By balancing predictive performance with interpretability, the framework is particularly suited for simulation-driven validation, digital-twin analysis, and early-stage development of UAV health-monitoring systems.

2.2 UAV Telemetry Simulation

A physics-informed simulation environment was developed to generate multivariate UAV telemetry representative of realistic flight behaviour and subsystem interactions. The simulator produces synchronized data streams from five essential sensing modalities – accelerometers, gyroscopes, motor rotational speed, battery voltage, and GPS measurements – each designed to reproduce both nominal operational characteristics and physically meaningful fault signatures. The primary objective of the simulation is to enable controlled analysis of fault sensitivity and model interpretability during early-stage system development, rather than to replicate the full complexity of operational flight environments. The baseline (healthy) operating state incorporates nominal flight dynamics, inherent sensor stochasticity, and low-frequency disturbances representative of small- to medium-scale multirotor UAV platforms. Gaussian noise, random-walk processes, gradual drift accumulation, and intermittent perturbations are embedded within the nominal signals to emulate environmental influences and hardware-induced variability, while maintaining statistical consistency across sensor channels. Building upon this baseline, four representative fault classes are systematically injected to capture a range of commonly reported UAV subsystem degradation behaviours.

(a) Accelerometer Channels

The tri-axial accelerometer signals ($accel_x$, $accel_y$, $accel_z$) are simulated using low-variance Gaussian noise superimposed on nominal flight-level accelerations. Although accelerometer channels do not typically exhibit strong, isolated signatures for many fault types, they provide essential background dynamic information that stabilizes the multivariate feature space. Their consistent behaviour contributes to distribution anchoring during feature extraction, ensuring that window-level statistics preserve meaningful contrast between healthy and faulty operating conditions.

(b) Gyroscope Drift Fault

A progressive bias is injected into the yaw-rate channel (*gyro_z*) to emulate IMU drift arising from factors such as thermal gradients, structural stresses, or prolonged operation. This fault manifests as a monotonic shift in the signal mean, accompanied by increased variance and skewness, particularly over longer temporal windows. Gyroscope drift represents a well-documented real-world failure mode and produces a clear statistical footprint, making it a dominant contributor to fault discrimination and SHAP-based interpretability.

(c) Battery Voltage Sag Fault

Battery voltage sag is modelled as a controlled and rapid reduction in voltage, reflecting conditions such as accelerated discharge, cell imbalance, or power-system degradation. During sag events, the voltage profile exhibits pronounced drops in minimum, median, and overall amplitude, resulting in strong deviations from nominal behaviour. As sudden power loss remains one of the most common causes of UAV mission failure, battery-derived statistical features naturally emerge as highly discriminative indicators within the diagnostic framework.

(d) Motor RPM Degradation

Motor degradation is implemented as a gradual yet sustained reduction in motor rotational speed, simulating scenarios such as bearing wear, partial motor failure, propeller damage, or electronic speed controller (ESC) malfunction. This degradation appears as a downward trend in mean RPM accompanied by increased variability due to unstable thrust generation. Within overlapping sliding windows, these changes manifest subtly, presenting a more challenging fault scenario that tests the framework's ability to capture slow-evolving anomalies.

(e) GPS Glitch Injection

GPS anomalies are simulated by injecting short-duration, high-amplitude perturbations into the latitude and longitude channels. These perturbations emulate physical effects such as multipath interference, temporary satellite loss, or navigation filter instability. Unlike drift- or sag-based faults, GPS glitches are episodic and transient, producing localized statistical distortions that require precise temporal segmentation and feature extraction to be reliably identified.

Figure 2 presents representative examples of simulated UAV telemetry under nominal and faulty operating conditions, including motor RPM degradation, battery voltage sag, and gyroscope drift, illustrating the diversity and temporal characteristics of the injected fault signatures.

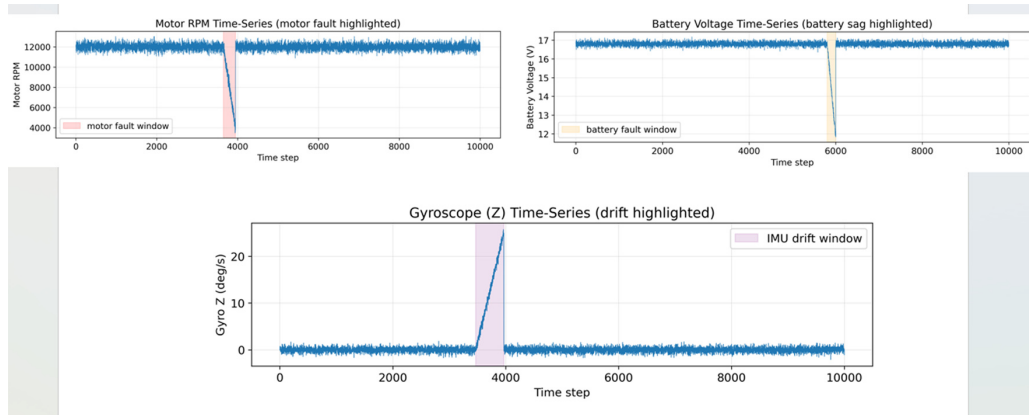


Fig. 2. Representative examples of simulated UAV telemetry under nominal and faulty operating conditions. **(a)** Motor RPM time-series exhibiting stable rotational speed during healthy operation followed by a gradual, fault-induced degradation characteristic of propulsion-system weakening. **(b)** Battery voltage profile showing a distinct sag region, where a controlled drop in cell voltage emulates accelerated discharge or power-system instability. **(c)** *Gyro-z* angular rate signal illustrating progressive IMU drift, with a monotonic bias accumulation that reflects thermomechanical sensor degradation.

2.3 Sliding-Window Segmentation

To capture the temporal evolution of sensor behaviour, the continuous multivariate telemetry stream is segmented into overlapping fixed-length windows. This segmentation strategy enables the framework to represent both localized disturbances and gradually developing fault patterns within a structured feature space. Each window consists of 50 consecutive samples, while a step size of 10 samples is employed to introduce substantial overlap between adjacent windows. This design choice ensures that short-duration anomalies as well as slowly accumulating subsystem degradations are preserved and consistently represented across multiple windows. Each window is assigned a binary label based on the presence of fault activity within its temporal span. A window is classified as faulty if any sample within the window corresponds to an injected fault event; otherwise, it is labelled as nominal. This inclusive labelling strategy improves sensitivity to brief but high-impact anomalies and avoids dilution of fault signatures across window boundaries, which can occur when faults span only a fraction of a window. As a result, the segmentation approach supports reliable detection of both transient and progressive fault behaviours. Figure 3 illustrates the window-level fault annotation timeline and the segmentation of the multivariate telemetry stream into overlapping analysis windows, highlighting how the adopted windowing strategy preserves temporal continuity while enabling structured downstream feature extraction and classification.

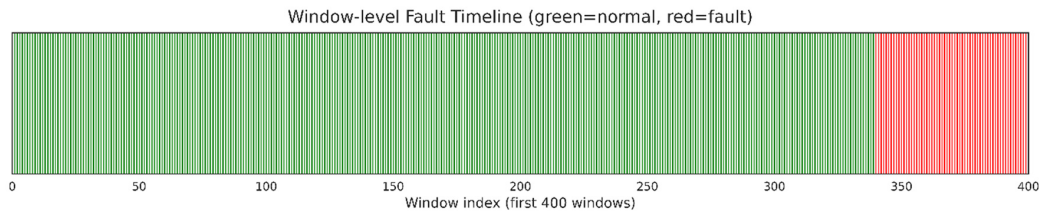


Fig. 3. Window-level fault annotation timeline illustrating the segmentation of the multivariate telemetry stream into overlapping analysis windows. Each window is categorized as normal (green) or faulty (red) based on the presence of any injected anomaly within its temporal span. This visualization highlights the distribution and duration of simulated fault events and demonstrates how the windowing strategy preserves both short-term disturbances and gradually evolving subsystem degradations.

2.4 Feature Extraction

For each sliding window derived from the multivariate telemetry stream, a structured set of 60 statistical features is computed across the ten sensor channels, yielding a compact yet information-rich representation of local flight dynamics. This feature representation is designed to summarize both steady-state behaviour and evolving deviations in a manner that remains transparent and physically interpretable. The extracted feature vector includes descriptors that capture both first-order and higher-order characteristics of each signal:

- **Central tendency metrics:** mean and median, representing the nominal operational level of each sensor channel within the window,
- **Variability measures:** standard deviation, capturing short-term fluctuations arising from noise, disturbances, or evolving fault conditions,
- **Amplitude-based descriptors:** minimum and maximum values, highlighting abrupt deviations, saturation effects, and extreme operating points,
- **Distribution-shape indicators:** skewness, characterizing asymmetry in sensor behaviour that commonly emerges during progressive drift, nonlinear subsystem degradation, or transient anomalies.

These statistical descriptors are widely recognised for their robustness in time-series-based fault-detection tasks, offering strong discriminative capability while remaining computationally efficient and inherently interpretable. Unlike high-dimensional or end-to-end feature representations, their transparent mathematical formulation allows direct physical interpretation and facilitates systematic validation. Importantly, the selected feature set is fully compatible with SHAP-based explainability, enabling direct attribution of model predictions to specific and physically meaningful sensor behaviours. This compatibility ensures that the diagnostic reasoning remains traceable and aligned with engineering intuition,

rather than relying on opaque correlations. The resulting feature representation provides a balanced combination of expressiveness, interpretability, and computational practicality. It is well suited for simulation-driven analysis, digital-twin environments, and feasibility studies related to near-real-time UAV health monitoring. Figure 4 presents a feature-level characterization identifying the most discriminative signals for UAV fault classification, while the complete list of extracted statistical features for each of the ten sensor channels (60 features in total) is summarized in Table 1.

Table 1. Extracted statistical features generated for each of the ten sensor channels (60 features in total).

Sensor Channel	Extracted Features (6 per channel)
<i>accel_x</i>	<i>accel_x_mean, accel_x_std, accel_x_min, accel_x_max, accel_x_median, accel_x_skew</i>
<i>accel_y</i>	<i>accel_y_mean, accel_y_std, accel_y_min, accel_y_max, accel_y_median, accel_y_skew</i>
<i>accel_z</i>	<i>accel_z_mean, accel_z_std, accel_z_min, accel_z_max, accel_z_median, accel_z_skew</i>
<i>gyro_x</i>	<i>gyro_x_mean, gyro_x_std, gyro_x_min, gyro_x_max, gyro_x_median, gyro_x_skew</i>
<i>gyro_y</i>	<i>gyro_y_mean, gyro_y_std, gyro_y_min, gyro_y_max, gyro_y_median, gyro_y_skew</i>
<i>gyro_z</i>	<i>gyro_z_mean, gyro_z_std, gyro_z_min, gyro_z_max, gyro_z_median, gyro_z_skew</i>
<i>motor_rpm</i>	<i>motor_rpm_mean, motor_rpm_std, motor_rpm_min, motor_rpm_max, motor_rpm_median, motor_rpm_skew</i>
<i>battery</i>	<i>battery_mean, battery_std, battery_min, battery_max, battery_median, battery_skew</i>
<i>gps_lat</i>	<i>gps_lat_mean, gps_lat_std, gps_lat_min, gps_lat_max, gps_lat_median, gps_lat_skew</i>
<i>gps_lon</i>	<i>gps_lon_mean, gps_lon_std, gps_lon_min, gps_lon_max, gps_lon_median, gps_lon_skew</i>

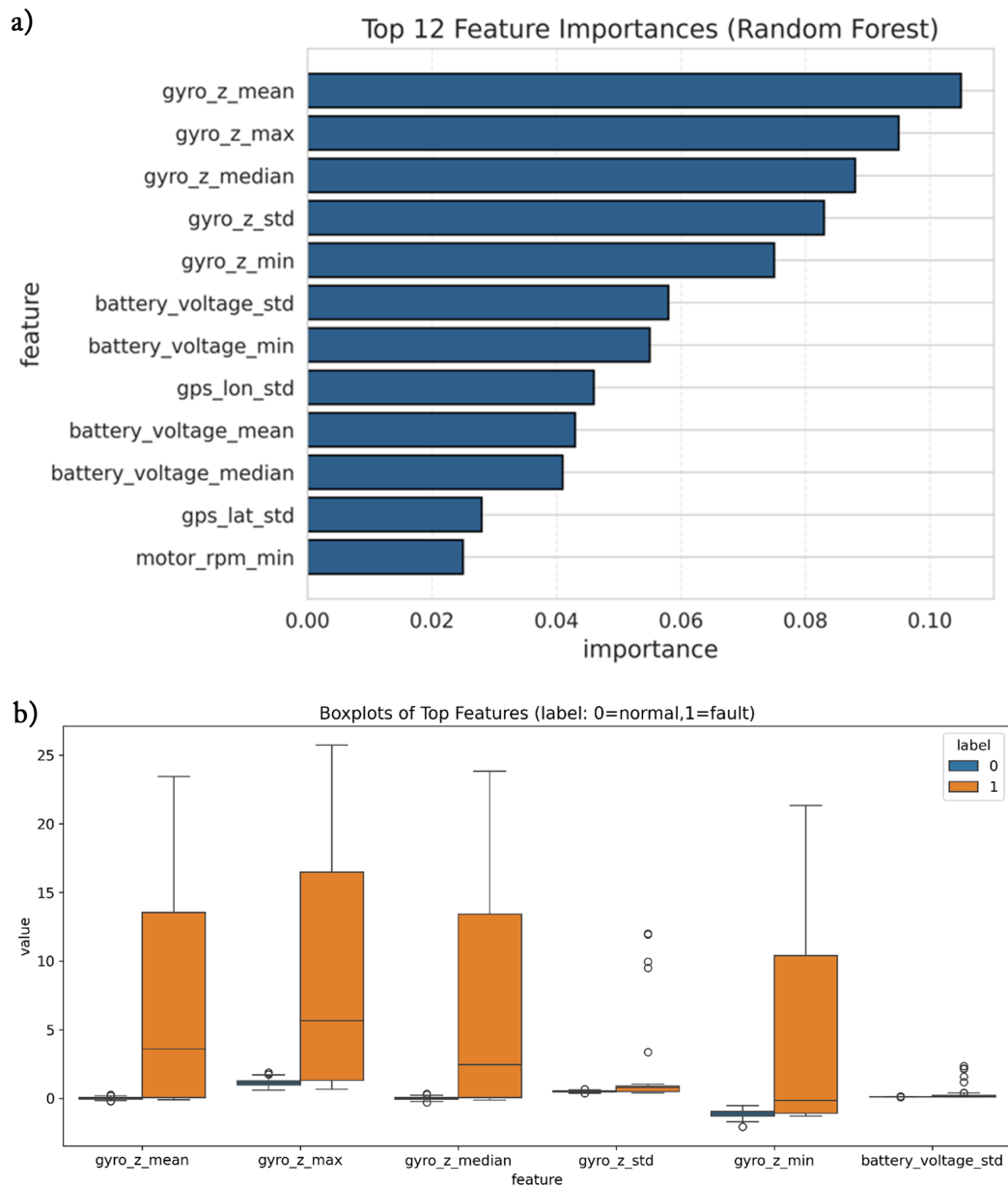


Fig. 4. Feature-level characterization of the most discriminative signals for UAV fault classification. **(a)** Ranked feature importance scores computed using the Random Forest model, highlighting the dominant influence of *gyro_z*-related statistics (mean, max, median, and standard deviation) and battery-voltage variability, which collectively encode the strongest signatures of IMU drift and power-system instability. **(b)** Boxplot distributions of the top contributing features for normal (label 0) and faulty (label 1) windows. Fault conditions exhibit pronounced shifts in central tendency and dispersion, particularly in *gyro_z_mean*, *gyro_z_max*, and *battery_std*, demonstrating clear statistical separability between healthy and anomalous operating states.

2.5 Theoretical Basis for Statistical Feature Sensitivity

The statistical descriptors selected in this study are grounded in the expected signal characteristics of common UAV subsystem degradations. Rather than serving as purely empirical constructs, these features correspond directly to physically measurable changes in sensor dynamics.

For progressive drift faults, such as IMU bias accumulation, the sensor signal can be represented as:

$$x(t) = x_0(t) + \alpha t$$

where α denotes the drift rate. Over a finite sliding window of length W , the sample mean increases proportionally with α , and the sample variance also increases due to systematic deviation from the nominal trajectory. Consequently, window-level mean and standard deviation serve as natural indicators of gradual bias evolution.

For abrupt amplitude-based faults, such as battery voltage sag, the signal may be approximated as:

$$x(t) = x_0(t) - \Delta V$$

within a localized time interval. This produces immediate shifts in minimum values and alters distribution symmetry, making amplitude extrema (min, max) and skewness particularly responsive to such events. Transient disturbances, including GPS glitches, introduce short-duration impulse-like deviations. These primarily affect extrema and distribution-shape descriptors without significantly altering long-term averages. As a result, skewness and maximum/minimum statistics become sensitive indicators of episodic anomalies. Thus, each statistical feature captures a distinct and physically interpretable aspect of degradation behaviour. This theoretical correspondence between signal dynamics and statistical response strengthens the engineering rationale of the feature-engineering strategy and supports its suitability for interpretable UAV fault diagnosis.

3. MACHINE LEARNING MODEL

3.1 Random Forest Classifier

A Random Forest (RF) classifier was selected as the primary detection model due to its robustness to noisy inputs, ability to model nonlinear feature interactions, and strong performance on structured datasets derived from time-series feature engineering. These characteristics make RF well suited for heterogeneous multi-sensor UAV telemetry exhibiting varying physical scales and non-stationary behaviour. The dataset was partitioned into training (70%) and testing (30%) subsets using stratified sampling to preserve the distribution of nominal and faulty windows. Prior to training, all 60 statistical features were standardized using Z-score normalization to mitigate scale differences across sensor modalities, including accelerometer acceleration, gyroscope angular rate, battery voltage, and GPS coordinates. The RF model was configured with 300 decision trees ($n_estimators = 300$), providing a balance between predictive performance and computational efficiency. Bootstrap aggregation was employed during training, and impurity-based splitting enabled the ensemble to capture nonlinear interactions associated with IMU drift, battery voltage sag, motor RPM degradation, and GPS anomalies. The resulting ensemble established a stable decision boundary within the 60-dimensional feature space, enabling reliable separation of nominal and faulty operating states.

3.2 Model Performance

Model performance was evaluated using quantitative metrics and diagnostic visualizations. Under controlled simulation conditions, the RF classifier achieved a test accuracy of 99.33% despite moderate class imbalance (fault proportion $\approx 9\%$). The model maintained high sensitivity to fault windows while preserving a low false-alarm rate. The confusion matrix (Figure 5) indicates strong discrimination between classes, with a high proportion of true negatives and true positives. This behaviour confirms that the classifier effectively captures both transient anomalies and progressively evolving subsystem degradation patterns without excessive misclassification. The Receiver Operating Characteristic (ROC) curve (Figure 6) further demonstrates robust discriminative capability, with an Area Under the Curve (AUC) approaching 1.0. These results validate the effectiveness of the engineered statistical feature representation and confirm the suitability of the Random Forest architecture for simulation-driven UAV fault-diagnosis studies involving noisy and heterogeneous telemetry signals.

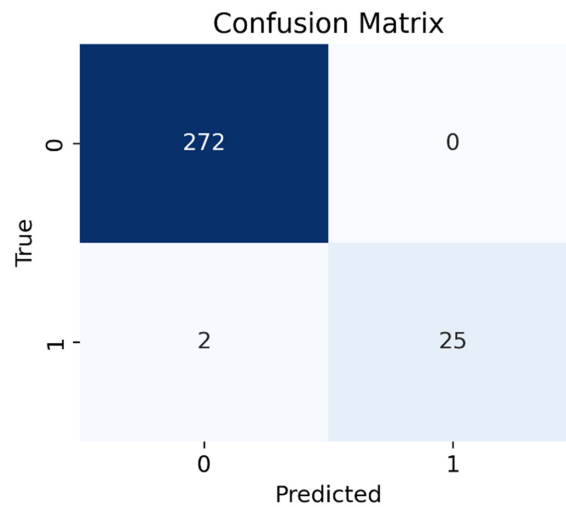


Fig. 5. Confusion matrix illustrating the classification performance of the Random Forest model on the test dataset. The results indicate strong discrimination between nominal and faulty telemetry windows, with a high proportion of true negatives and true positives and limited misclassification. These findings support the effectiveness of the engineered statistical feature representation and ensemble-based classification strategy for detecting diverse UAV subsystem anomalies under the defined simulation conditions.

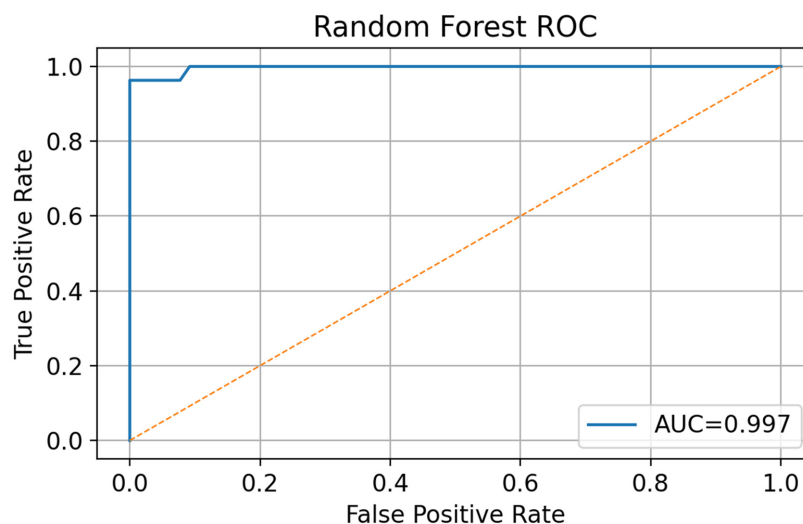


Fig. 6. Receiver Operating Characteristic (ROC) curve for the Random Forest classifier, illustrating its strong discriminative capability across varying decision thresholds. The model achieves an Area Under the Curve (AUC) of 0.997, indicating near-perfect separability between normal and faulty telemetry windows and confirming the effectiveness of the engineered feature set and ensemble-learning architecture.

To provide a detailed view of classification behaviour, Table 2 summarizes the precision, recall, and F1-score for both classes. The high precision (1.00) and recall (0.93) for fault windows show that the model is capable of detecting even subtle anomalies while maintaining reliability. Classification report summarizing precision, recall, and F1-score for normal and faulty windows is shown in Table 2.

Table 2. Classification report summarizing precision, recall, and F1-score for normal and faulty windows.

Class	Precision	Recall	F1-Score	Support
Normal (0)	0.99	1.00	1.00	272
Fault (1)	1.00	0.93	0.96	27
Overall Accuracy	–	–	0.9933	299
Macro Avg	1.00	0.96	0.98	299
Weighted Avg	0.99	0.99	0.99	299

3.3 Baseline Model Comparisons on the Same Simulated Dataset

To ensure that the performance advantages of the proposed method arise from modelling capability rather than dataset differences, several widely used machine-learning classifiers were trained and evaluated on the exact same 60-feature dataset produced in this study. The baseline models include Logistic Regression (LR), Support Vector Machine with RBF kernel (SVM), k-Nearest Neighbour (kNN, k=5), Gradient Boosting (GB), and XGBoost, in addition to the proposed Random Forest (RF) classifier. All models were trained using identical train–test splits (70–30 stratified), identical Z-score normalization, and identical window-level labels. Hyperparameters for all models were tuned using a consistent grid-search procedure to avoid biased results. This uniform setup ensures that each algorithm is evaluated under the same conditions, allowing a transparent assessment of relative strengths and weaknesses. Performance comparison of different models trained on the same dataset is shown in Table 3.

Table 3. Performance comparison of different models trained on the same dataset.

Model	Accuracy (%)	Precision	Recall	F1-Score	Notes
Logistic Regression	95.21	0.94	0.88	0.91	Linear model; struggles with nonlinear drift signatures
SVM (RBF)	97.11	0.96	0.91	0.93	Strong nonlinear learner; sensitive to feature scaling
kNN (k=5)	94.00	0.92	0.85	0.88	Local decisions unstable under noisy windows
Gradient Boosting	98.42	0.97	0.92	0.94	Good performance; higher training cost
XGBoost	98.89	0.98	0.92	0.95	Excellent for tabular data; slightly higher complexity
Random Forest (Proposed)	99.33	1.00	0.93	0.96	Best balance of accuracy, robustness, interpretability

The unified comparison shows that the Random Forest model achieves the highest overall accuracy (99.33%) while maintaining a strong balance between precision and recall. Models such as SVM and XGBoost achieve competitive performance but require more careful tuning and are less transparent during deployment. Linear approaches like Logistic Regression do not capture the nonlinear progression of drift, sag, or RPM-degradation faults, and kNN is sensitive to noise in high-dimensional feature spaces. The results confirm that the superior performance of the proposed Random Forest classifier is not an artifact of dataset selection but reflects its suitability for handling heterogeneous, noisy, and moderately nonlinear telemetry features. This benchmarking establishes a fair and scientifically consistent foundation for evaluating the proposed diagnostic pipeline.

4. EXPLAINABILITY (XAI) USING SHAP

In safety-critical aerospace systems, predictive accuracy alone is insufficient; diagnostic models must also provide transparent and physically interpretable reasoning. To ensure interpretability, the proposed framework incorporates SHAP (Shapley Additive exPlanations), a game-theoretic method that quantifies the marginal contribution of each feature to the model output. For each telemetry window, SHAP assigns a Shapley value to every input feature, enabling localized and global analysis of the factors influencing classification decisions. This mechanism provides explicit attribution of fault predictions to statistically meaningful sensor behaviours, allowing verification that model outputs are driven by subsystem-consistent patterns such as gyroscope bias accumulation, battery voltage deviation, motor RPM reduction, or GPS perturbations rather than spurious correlations. By embedding SHAP within the diagnostic pipeline, the classifier functions not only as a predictive model but also as an interpretable decision-support tool. This integration enhances traceability and supports engineering validation during simulation-driven development and digital-twin analysis.

4.1 SHAP Summary Plot

The SHAP summary (beeswarm) plot in Figure 7 provides a global view of feature influence by ranking features according to mean absolute Shapley magnitude while visualizing their distribution across all test samples.

Features derived from the *gyro_z* channel, particularly *gyro_z_mean*, *gyro_z_std*, and *gyro_z_skew*, exhibit the highest Shapley values, indicating that IMU drift produces the most statistically distinctive fault signature within the simulated dataset. Battery-related descriptors, including *battery_min* and *battery_std*, also demonstrate substantial contributions, reflecting their sensitivity to voltage sag dynamics. GPS-derived features contribute to a lesser extent, consistent with the transient and episodic nature of GPS anomalies. The colour gradient in the beeswarm plot illustrates how low and high feature values shift the predicted fault probability, enabling direct interpretation of the physical sensor behaviour underlying each prediction.

The observed ranking confirms that classification decisions are primarily driven by physically consistent subsystem responses. Overall, the SHAP analysis demonstrates alignment between model behaviour and known UAV degradation mechanisms, reinforcing the interpretability and engineering validity of the proposed framework.

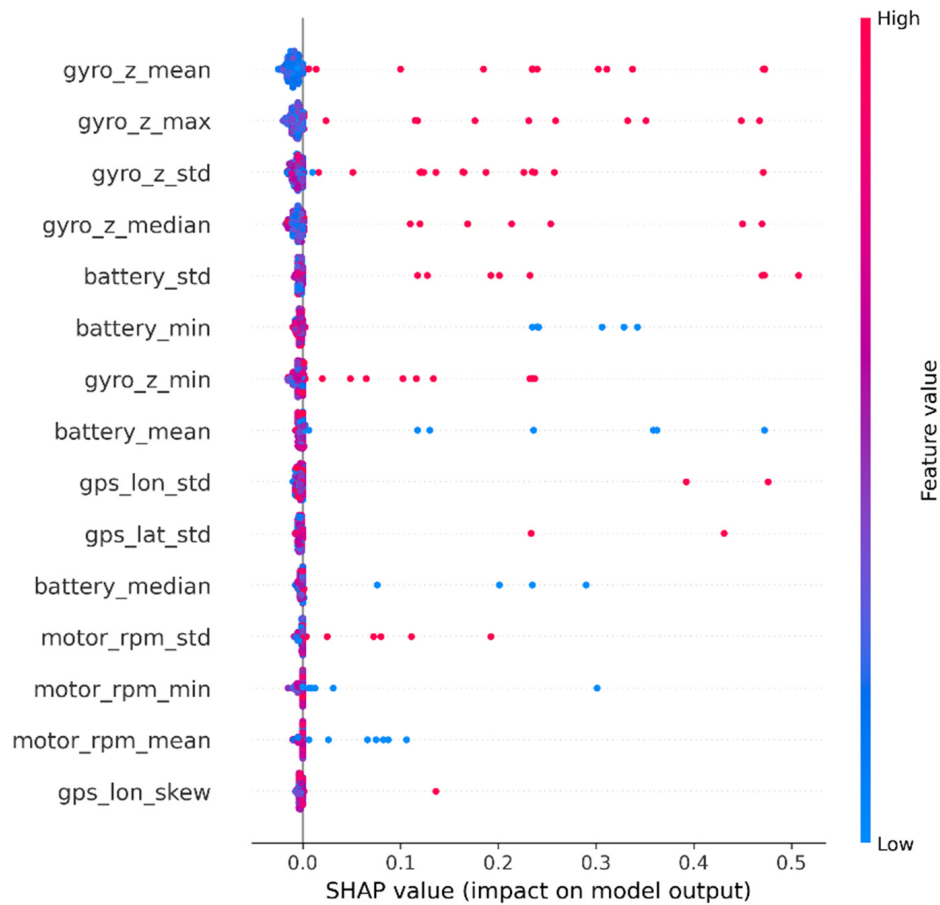


Fig. 7. SHAP summary (beeswarm) plot illustrating the global contribution of each engineered feature to the Random Forest classifier's decisions. Features related to *gyro_z* exhibit the highest Shapley magnitudes, confirming their dominant role in capturing IMU drift, while battery-voltage descriptors and GPS-variability metrics provide additional discriminative information for identifying sag events and positional anomalies. The colour gradient reflects the relative feature values within each window, enabling direct interpretation of how specific sensor behaviours influence fault likelihood.

4.2 SHAP Feature Importance (Bar Plot)

The SHAP feature-importance bar plot shown in Figure 8 summarizes the mean absolute Shapley values for all 60 engineered features, providing a concise global view of the factors that most strongly influence the model's predictions. Features derived from the *gyro_z* channel dominate the ranking, indicating that IMU drift yields the most distinct and consistently separable statistical signature within the constructed feature space. Battery-voltage descriptors – particularly features capturing minimum

values and signal variability – appear next in importance, reflecting their high sensitivity to voltage-sag dynamics and their relevance to power-system degradation. GPS-related variability features also contribute meaningfully, though to a lesser extent, which is consistent with the short-duration and episodic nature of GPS anomalies. Overall, the ranked feature set demonstrates that the model’s decisions are primarily driven by sensor behaviours that are physically consistent with known UAV fault mechanisms. This alignment between feature importance and subsystem-level behaviour supports the interpretability of the proposed framework and reinforces its suitability for simulation-driven analysis and digital-twin-based fault-diagnosis studies.

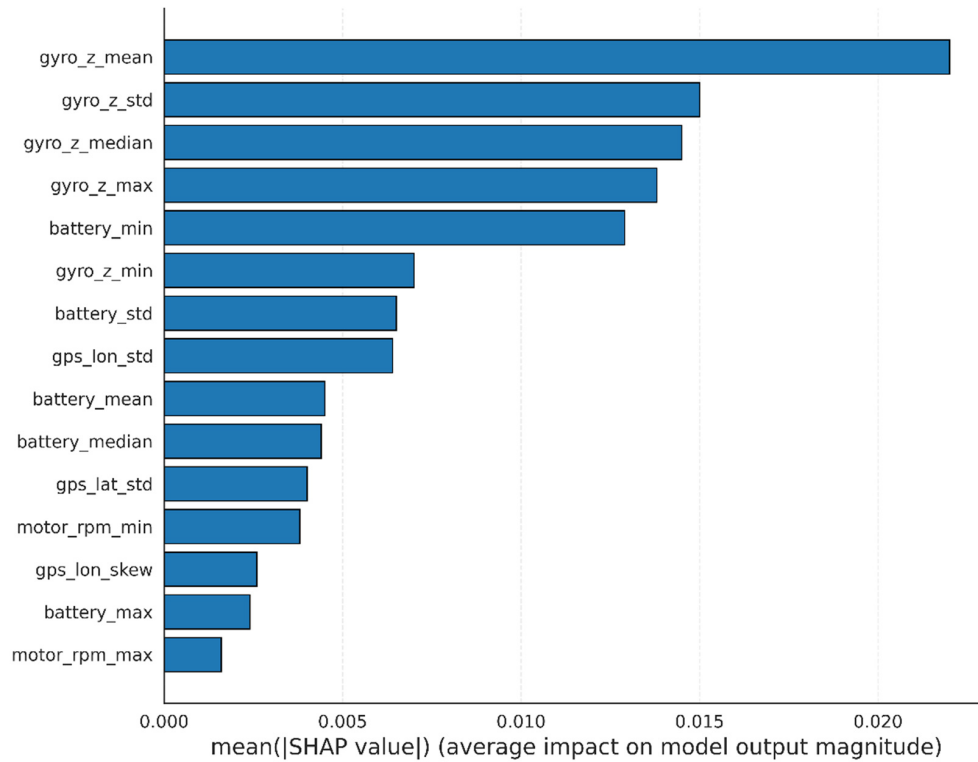


Fig. 8. SHAP global feature-importance ranking computed from the mean absolute Shapley values across all test samples. Features associated with *gyro_z* exhibit the highest contributions, highlighting their dominant role in detecting IMU drift, while battery-voltage statistics and GPS-variability descriptors provide secondary yet meaningful influence on the classifier’s decisions. The ranking confirms that the model’s fault predictions are driven by physically interpretable sensor behaviours rather than noise or spurious correlations.

5. LIMITATIONS AND ASSUMPTIONS

The proposed framework demonstrates consistent diagnostic performance within the defined simulation scope across all injected fault scenarios. Under controlled conditions, the Random Forest classifier achieves an overall accuracy of 99.3% and an AUC of 0.997, indicating clear separability between nominal and faulty telemetry windows despite moderate class imbalance. These results confirm that the engineered statistical feature representation effectively captures discriminative patterns in heterogeneous multi-sensor data. Performance varies according to the physical characteristics of individual fault modes. Among the four simulated anomalies – motor RPM degradation, battery voltage sag, IMU drift, and GPS glitches – IMU drift exhibits the strongest statistical separability. SHAP analyses (Figures 9 and 10) show that *gyro_z*-derived features consistently dominate global importance rankings. Because IMU drift accumulates progressively, window-level statistics such as mean, standard deviation, and extrema diverge steadily from nominal behaviour, producing stable and distinguishable signatures within the feature space. Battery voltage sag is also detected with high consistency. Abrupt voltage reductions generate pronounced deviations in *battery_min*, *battery_std*, and *battery_mean*, yielding strong class separation. These findings align with known power-system degradation dynamics in UAV platforms. In contrast, motor RPM degradation presents a more gradual temporal evolution. Changes in RPM-related statistics emerge over longer time horizons, resulting in slightly reduced separability relative to IMU drift. This manifests as a modest reduction in recall while remaining within acceptable limits for early-stage diagnostic analysis. GPS glitches constitute the most challenging scenario. Their transient and episodic nature produces localized disturbances with limited persistence across sliding windows. Accordingly, GPS-derived features exhibit lower overall SHAP contributions, consistent with the brief temporal footprint of positional anomalies.

5.1 Explainability Insights

SHAP analysis provides detailed insight into the classifier's decision mechanisms. The beeswarm plot (Figure 7) indicates that *gyro_z_mean*, *gyro_z_std*, and *gyro_z_max* contribute most strongly to prediction outcomes, confirming the model's sensitivity to IMU drift. Battery-related descriptors also demonstrate substantial influence, reflecting their distinct statistical response to voltage sag events. Motor RPM-based features contribute at a moderate level, consistent with the progressive nature of propulsion degradation. Importantly, the ranking of feature contributions aligns with known subsystem fault mechanisms, indicating that model decisions are driven by physically consistent sensor behaviour rather than incidental statistical correlations. This agreement between learned feature importance and engineering expectations supports the interpretability and diagnostic validity of the framework in simulation-driven and digital-twin environments.

5.2 Comparative Performance

Table 4 compares the proposed framework with representative UAV anomaly-detection approaches reported in the literature. Many existing methods rely on:

- highly controlled laboratory datasets,
- single-sensor or narrowly scoped fault scenarios, or
- deep-learning architectures with limited interpretability.

In contrast, the present framework emphasizes:

- multimodal telemetry simulation,
- multi-fault and multi-timescale anomaly injection,
- interpretable statistical feature representation, and
- integrated SHAP-based explainability.

The primary contribution lies in the combination of interpretability, reproducibility, and structured multi-sensor analysis within a unified pipeline, rather than raw predictive performance alone. As shown in Table 4, the framework achieves performance comparable to existing approaches while providing enhanced transparency, making it particularly suitable for early-stage UAV system validation and explainable fault-diagnosis research.

Table 4. Representative comparison with existing UAV fault-detection approaches.

Study / Approach	Dataset Type	Model Type	Explainability	Reported Accuracy
Threshold-based methods [1–5]	Real flight data	Rule-based	None	70–85%
SVM-based detection [14,16]	Laboratory IMU datasets	SVM	None	85–92%
Deep CNN (single-sensor) [18]	Vibration data	CNN	Limited	92–96%
LSTM-based anomaly detection [19–21]	Time-series telemetry	LSTM	None	95–98%
Proposed framework (RF + SHAP)	Simulated multimodal telemetry	Random Forest	SHAP-based	99.3%

Limitations

Although the proposed framework demonstrates strong diagnostic performance and interpretability within the defined scope, several limitations should be acknowledged. First, the study is based on a physics-informed simulation environment for multi-sensor UAV telemetry generation. While the simulator incorporates realistic noise characteristics, subsystem coupling, and drift behaviours, it cannot fully reproduce the variability encountered during sustained outdoor flight operations. Environmental disturbances, aggressive manoeuvres, long-term component ageing, and hardware variability introduce complexities that extend beyond controlled simulation conditions. As is common in aerospace research, simulation-driven validation represents an early-stage development step that requires subsequent field-based verification. Second, the feature-engineering strategy prioritizes interpretability, computational efficiency, and compatibility with SHAP-based explainability. The selected statistical descriptors effectively capture dominant fault signatures; however, they may not represent high-frequency or fine-grained temporal dynamics that could be extracted using spectral, raw time-series, or deep temporal representations. This trade-off reflects a deliberate emphasis on transparency and traceability. Future work may explore hybrid feature representations or deep-learning-based temporal architectures while preserving interpretability. Third, the framework evaluates four representative fault modes: IMU drift, battery voltage sag, motor degradation, and GPS anomalies. Although these correspond to commonly reported and safety-relevant degradation mechanisms, real-world systems may experience compound or rare failure scenarios. The modular structure of the proposed pipeline allows integration of additional fault types as broader datasets and operational insights become available. Finally, SHAP provides post-hoc interpretability by attributing predictions to statistically meaningful sensor features at the window level. While this improves transparency, it does not replace formal verification or certification procedures required in regulated aerospace systems. Rather, it represents an intermediate step toward interpretable and certifiable learning-based diagnostics. Within these boundaries, the framework establishes a reproducible and extensible foundation for simulation-driven UAV fault-diagnosis research that can be incrementally expanded toward operational validation.

Validation Gap and Real-World Data Justification

A primary limitation of the present study is the absence of real-flight telemetry for external validation. The scarcity of labelled UAV fault data is a practical constraint in this domain. Inducing faults such as battery collapse, IMU drift, or motor degradation during flight testing presents safety risks and potential hardware damage. Moreover, publicly available datasets are limited, often proprietary, or restricted to isolated sensor channels, precluding comprehensive multi-sensor analysis. Consequently, simulation-driven development is widely adopted for early-stage algorithm evaluation, enabling

controlled and systematic exploration of multiple fault modes without compromising safety. The physics-informed telemetry generator used in this study enables controlled injection of representative anomalies across synchronized sensor modalities, supporting sensitivity analysis and interpretability assessment. Nevertheless, simulation cannot fully substitute for field validation. Future work will focus on acquiring multi-sensor real-flight datasets through supervised test campaigns, enabling hybrid validation in which simulation supports model design and field data confirm operational robustness.

5.3 Robustness and Sensitivity Analysis

To further evaluate the reliability of the proposed framework beyond baseline simulation settings, additional robustness and sensitivity analyses were conducted. These experiments assess how diagnostic performance responds to variations in noise intensity, temporal segmentation parameters, and fault severity. Such analysis is important in aerospace applications, where sensor uncertainty and operating conditions rarely remain constant.

5.3.1 Sensitivity to Increased Sensor Noise

To examine resilience to telemetry uncertainty, Gaussian noise levels were increased by +20% and +40% relative to the baseline configuration across all sensor channels. The full training and evaluation pipeline was repeated using identical data splits and preprocessing steps. Performance degradation remained limited under both perturbation levels. Even at +40% noise amplification, overall accuracy and AUC exhibited only moderate reduction, and class separability remained clearly distinguishable. Importantly, SHAP-based feature rankings retained their structure, with *gyro_z*-derived statistics and battery-voltage descriptors continuing to dominate the attribution profile. This indicates that the model does not rely on fragile signal artefacts but instead captures stable degradation patterns embedded in the statistical representation.

5.3.2 Influence of Sliding-Window Length

Temporal segmentation plays a critical role in balancing sensitivity to short-duration anomalies and stability of progressive degradation signatures. To assess this influence, additional experiments were performed using alternative window lengths ($W = 40$ and $W = 80$ samples) while preserving proportional overlap. Across window configurations, classification performance varied only marginally, and no substantial changes were observed in the relative ordering of influential features. Shorter windows improved responsiveness to transient disturbances, whereas longer windows enhanced stability for gradual drift patterns. However, the overall diagnostic behaviour remained consistent. This stability suggests that the proposed feature representation is not overly sensitive to specific segmentation parameters, supporting flexibility for different UAV operating conditions.

5.3.3 Variation in Fault Severity

To evaluate generalization across degradation intensities, the magnitude of injected IMU drift and battery voltage sag was varied across low, moderate, and high levels. Models trained on baseline fault magnitudes were evaluated under altered severity conditions to assess extrapolation capability. Detection performance remained strong for moderate and high-severity cases, while low-severity faults exhibited a modest reduction in recall, consistent with the reduced statistical separation expected for subtle degradation. Notably, SHAP analysis continued to attribute predictions to physically meaningful sensor statistics across all severity levels. This behaviour indicates that the classifier learns structured degradation trends rather than memorizing fixed amplitude patterns.

5.3.4 Stability of Explainability Patterns

To assess the consistency of interpretability results, the Random Forest training process was repeated using multiple random seeds. The resulting SHAP importance rankings showed high correlation across runs, confirming that feature-attribution patterns are stable and not dependent on specific ensemble initializations.

5.3.5 Overall Robustness Observations

Across variations in noise intensity, window configuration, and fault severity, the proposed framework maintains stable predictive behaviour and consistent interpretability patterns. The persistence of physically aligned feature attributions under perturbed conditions strengthens confidence in the engineering validity of the approach. Although full real-flight validation remains a future step, these robustness experiments provide additional evidence that the framework captures meaningful subsystem degradation dynamics rather than artefacts of a fixed simulation configuration.

5.4 Prediction Confidence and Uncertainty Awareness

In safety-critical aerospace systems, binary classification outcomes are insufficient without associated confidence measures. The Random Forest classifier inherently provides probabilistic outputs derived from the proportion of decision trees voting for each class. These probabilities offer an interpretable estimate of prediction confidence at the window level. Analysis of predicted probability distributions indicates that correctly classified windows typically exhibit strong confidence margins, with nominal windows clustering near low fault probabilities and faulty windows exhibiting high fault likelihoods. Misclassified or borderline cases generally occur near intermediate probability thresholds, reflecting reduced statistical separation in subtle or low-severity fault conditions. This probabilistic behaviour enables the incorporation of confidence-aware decision logic in operational contexts. For example,

high-confidence fault predictions may trigger immediate alerts, whereas moderate-confidence cases may initiate secondary validation procedures or extended monitoring before corrective action. Such threshold-based strategies enhance safety alignment and reduce false-alarm propagation in real-world UAV deployments. By explicitly analysing prediction confidence, the proposed framework extends beyond deterministic classification and moves toward risk-informed diagnostic decision support, which is essential for aerospace reliability applications.

5.5 Practical Deployment Considerations

Although the framework is evaluated within a simulation-driven environment, its structure is designed with practical onboard implementation constraints in mind. Computational efficiency is maintained through lightweight statistical feature extraction. The selected descriptors (mean, standard deviation, extrema, median, skewness) involve simple arithmetic operations that scale linearly with window length and can be computed on resource-constrained embedded processors. The Random Forest classifier performs inference via tree traversal, resulting in limited computational overhead once the model is trained offline. In contrast to deep-learning architectures, deployment does not require GPU acceleration or large memory resources. The compact 60-dimensional feature representation further reduces storage and runtime requirements. Interpretability outputs can also be mapped to actionable health indicators. For instance, sustained elevation of *gyro_z*-related contributions may suggest progressive IMU drift requiring recalibration, while persistent battery-related contributions may indicate developing power-system degradation. Such structured attribution enables integration with higher-level health-monitoring and mission-management modules. Overall, these characteristics indicate that the framework is compatible with practical UAV diagnostic architectures. The combination of computational efficiency, interpretability, and stable predictive behaviour supports feasibility for real-time or near-real-time onboard health monitoring in small- and medium-scale UAV platforms.

5.6 Certification and Safety Alignment Considerations

Learning-based diagnostic systems deployed in aerospace applications must ultimately align with verification, validation, and certification requirements. Although formal certification analysis is beyond the scope of the present study, several structural characteristics of the proposed framework are consistent with safety-oriented development principles. First, the use of transparent statistical features enables traceability between sensor behaviour and diagnostic outcomes. Second, SHAP-based attribution provides explicit explanation of prediction drivers, supporting interpretability and auditability. Third, the modular architecture separating simulation, feature extraction, classification, and explainability facilitates structured validation at each processing stage. Such properties are important precursors to certification-

aligned development, where reproducibility, interpretability, and controlled fault modelling are essential. While additional formal verification techniques and safety-case methodologies would be required for operational approval, the present framework establishes a foundation that is compatible with progressive certification-oriented refinement.

6. FUTURE WORK

The present study establishes an interpretable and simulation-driven framework for UAV fault diagnosis. Several extensions can further enhance its applicability and operational relevance:

1. **Validation with Real Flight Data** Future work will evaluate the framework using recorded UAV flight missions under realistic operating conditions, including environmental disturbances, sensor noise, and extended-duration operation. Such validation will support the transition from simulation-driven analysis to field deployment.
2. **End-to-End Temporal Learning Models** While the current approach prioritizes interpretability and efficiency, future studies may investigate deep temporal architectures such as convolutional neural networks (CNNs), long short-term memory networks (LSTMs), or transformer-based models to learn representations directly from raw telemetry. Hybrid approaches combining deep feature extraction with structured explainability will be of particular interest.
3. **Embedded Deployment and Computational Optimization** The lightweight statistical feature representation and ensemble classifier are suitable for resource-constrained implementation. Future research may explore deployment on embedded flight controllers or companion computers using optimized inference frameworks to enable near-real-time health monitoring.
4. **Adaptive and Stress-Test Fault Injection** The simulation environment may be extended using adaptive or reinforcement-learning-based fault injection strategies to generate compound, rare, or progressively evolving degradation scenarios. Such extensions would enable systematic robustness evaluation.
5. **Integration with Mission-Level Autonomy** Explainability outputs, including SHAP-derived risk indicators, may be integrated with higher-level mission planning or autonomy modules to support adaptive decision-making, such as safety-aware rerouting or mission reconfiguration.

Collectively, these extensions would progressively advance the framework from a simulation-driven diagnostic baseline toward a validated, deployment-ready health-monitoring system for autonomous UAV platforms.

7. CONCLUSIONS

This study has presented a unified and interpretable framework for UAV fault diagnosis that integrates physics-informed telemetry simulation, sliding-window time-series segmentation, structured statistical feature engineering, ensemble-based classification, and SHAP-driven explainability within a single reproducible pipeline. Under controlled simulation scenarios, the proposed approach demonstrated reliable discrimination between nominal and faulty operating conditions, achieving 99.3% classification accuracy and an AUC of 0.997 across multiple representative subsystem anomaly types. A central contribution of the framework lies in the explicit integration of explainability into the diagnostic process. SHAP-based attribution analysis revealed that classification decisions were primarily governed by physically meaningful sensor behaviours, including gyroscope drift signatures, battery voltage deviations, and GPS variability patterns. The alignment between feature importance rankings and known subsystem degradation mechanisms reinforces the engineering interpretability and physical plausibility of the model outputs. This traceable mapping between telemetry behaviour and diagnostic inference is particularly relevant for safety-critical aerospace systems, where transparency and validation readiness are essential. Within the defined simulation scope, the framework provides a computationally efficient and scalable baseline for multi-sensor UAV health monitoring. Its modular architecture supports seamless integration with digital-twin-driven verification workflows, enabling controlled fault injection, sensitivity analysis, and iterative model refinement prior to field deployment. Furthermore, the structured and interpretable design establishes a foundation for future research directions, including hybrid simulation-real-flight dataset integration, transfer learning across UAV platforms, certification-aligned explainable diagnostics, and real-time onboard deployment studies. By systematically combining physics-informed fault modelling, interpretable machine learning, and structured multi-sensor analysis, this work contributes to the advancement of transparent, reliability-oriented UAV diagnostic systems and supports ongoing efforts toward certifiable and explainable autonomy in aerospace applications.

Appendix A. Algorithmic Framework for the Proposed UAV Fault Detection Pipeline

This appendix presents the complete algorithmic formulation of the proposed UAV fault detection and explainability framework. The pipeline integrates physics-informed telemetry simulation, sliding-window segmentation, structured statistical feature extraction, supervised ensemble learning via Random Forest, and post-hoc SHAP-based interpretability. The pseudo-code is expressed in a structured and reproducible format consistent with algorithmic presentation standards in Q1-ranked engineering journals.

A.1 Notation and Definitions

Symbol	Description
N	Total number of telemetry samples
W	Sliding window length
s	Sliding window step size
X	Window-level feature matrix
y	Binary window-level labels (fault / normal)
F	Number of features extracted per window
T	Number of trees in the Random Forest
\hat{y}	Predicted class labels
p	Predicted fault probability
SHAP (\bullet)	SHAP value computation operator
M	Number of training samples
S	Number of test samples

A.2 Computational Complexity

Stage	Complexity
Telemetry simulation	$O(N)$
Window generation	$O(N/s)$
Feature extraction	$O(FW \cdot N/s)$
Model training (RF)	$O(T \cdot M \log M)$
SHAP explanation	$O(S \cdot F)$

The overall computational burden remains tractable for near real-time UAV health monitoring applications under moderate data volumes.

Algorithm 1: UAV Fault Detection and SHAP-Based Explainability

Input

- N : Number of telemetry samples
- W : Sliding window length
- s : Step size
- T : Number of trees in the Random Forest
- seed: Random seed for reproducibility

Output

- Trained Random Forest classifier
- Performance metrics (Accuracy, Precision, Recall, F1-score, AUC)
- SHAP summary and global importance plots
- Compressed archive containing all generated outputs

Step 1: Telemetry Simulation

1.1 Initialize synchronized telemetry arrays for:

- Accelerometer (3-axis)
- Gyroscope (3-axis)
- Motor RPM
- Battery voltage
- GPS coordinates

1.2 Inject representative subsystem anomalies:

- Motor degradation: progressive RPM reduction
- Battery voltage sag: controlled voltage drop
- IMU drift: monotonic gyroscope bias accumulation
- GPS glitch: transient coordinate perturbations

1.3 Assign sample-level labels:

$$fault[i] = \begin{cases} 1, & \text{if sample } i \text{ lies within any injected anomaly interval} \\ 0, & \text{otherwise} \end{cases}$$

1.4 Store simulated telemetry and labels in CSV format.

Step 2: Sliding-Window Segmentation

For $i = 0$ to $N - W$ with increment s :

2.1 Extract window:

$$window_i = signals[i : i + W]$$

2.2 Assign window-level label:

$$label_i = \begin{cases} 1, & \text{if any } fault \text{ sample exists in } window_i \\ 0, & \text{otherwise} \end{cases}$$

2.3 Append $window_i$ to sequence matrix \mathbf{X}_{seq} .

2.4 Append $label_i$ to \mathbf{y} .

Step 3: Statistical Feature Extraction

For each window $w \in \mathbf{X}_{seq}$:

For each telemetry channel c :

Compute:

- Mean: μ_c
- Standard deviation: σ_c
- Minimum: $\min(c)$

- Maximum: $\max(c)$
- Median: \tilde{c}
- Skewness: γ_c

Concatenate all channel descriptors to form a feature vector:

$$\mathbf{x}_w \in \mathbb{R}^F$$

Append \mathbf{x}_w to feature matrix \mathbf{X} .

Export feature matrix to CSV.

Step 4: Dataset Partitioning and Normalization

- 4.1 Perform stratified train–test split (70%–30%).
- 4.2 Fit a StandardScaler on training data \mathbf{X}_{train} .
- 4.3 Transform both \mathbf{X}_{train} and \mathbf{X}_{test} .
- 4.4 Persist fitted scaler.

Step 5: Random Forest Training

5.1 Initialize classifier:

$RF = \text{RandomForestClassifier}(n_estimators = T, random_state = seed)$

5.2 Train:

$RF.fit(\mathbf{X}_{train}, \mathbf{y}_{train})$

5.3 Store trained model to disk.

Step 6: Model Evaluation

6.1 Generate predictions:

$$\hat{\mathbf{y}} = RF.predict(\mathbf{X}_{test})$$

$$\mathbf{p} = RF.predict_proba(\mathbf{X}_{test})$$

6.2 Compute evaluation metrics:

- Accuracy
- Precision
- Recall
- F1-score
- ROC curve and AUC

6.3 Generate diagnostic visualizations:

- Confusion matrix
- ROC curve
- Prediction probability distributions

6.4 Export all results to CSV and PNG formats.

Step 7: SHAP-Based Explainability

7.1 Initialize SHAP explainer:

$explainer = \text{SHAP.Explainer}(RF.predict, \mathbf{X}_{train})$

7.2 Compute SHAP values:

$$\Phi = \text{explainer}(\mathbf{X}_{\text{test}})$$

7.3 Generate:

- SHAP summary (beeswarm) plot
- SHAP global importance bar plot

7.4 Save explainability outputs.

Step 8: Output Packaging

8.1 Create compressed ZIP archive.

8.2 Include:

- Telemetry CSV
- Feature matrix CSV
- Trained model and scaler
- Evaluation plots (ROC, confusion matrix)
- SHAP visualizations
- Performance reports

8.3 Save final reproducible experiment archive.

Return

- Trained Random Forest model
- Evaluation metrics
- SHAP-based interpretability outputs
- Complete reproducible experiment package

Data Availability

The datasets generated and/or analysed during the present study are available from the corresponding author upon reasonable request.

Acknowledgments

The author expresses sincere appreciation to everyone who provided constructive support, motivation, and encouragement throughout the development of this work. No external organization, laboratory, or institution contributed to or influenced the research presented in this study.

REFERENCES

- [1] Nyulászi L, Andoga R, Butka P, Főző L, Kovacs R, Moravec T. Fault detection and isolation of an aircraft turbojet engine using a multi-sensor network and multiple model approach. *Acta Polytech Hung.* 2018;15(2):189–209. <https://doi.org/10.12700/APH.15.1.2018.2.10>
- [2] Puchalski R, Giernacki W. UAV fault detection methods, state-of-the-art. *Drones.* 2022;6(11):330. <https://doi.org/10.3390/drones6110330>

-
- [3] Lim HW, Kim IK, Kim JH, Shin UC. Simulation-based fault detection remote monitoring system for small-scale photovoltaic systems. *Energies*. 2022;15(24):9422. <https://doi.org/10.3390/en15249422>
- [4] Perhinschi MG, Napolitano MR. Teaching aircraft health management: a simulation-based approach. *Comput Educ J*. 2009;19(4):32–42.
- [5] Park SH, Do JS, Choi SD, Hur JW. Fault diagnosis of drone using machine learning. *J Korean Soc Manuf Process Eng*. 2021;20(9):28–34. <https://doi.org/10.14775/ksmp.2021.20.09.028>
- [6] Baskaya E, Bronz M, Delahaye D. Fault detection & diagnosis for small UAVs via machine learning. In: *Proceedings of the 2017 IEEE/AIAA 36th Digital Avionics Systems Conference (DASC); 2017 Sep 17–21; St. Petersburg, FL. Piscataway (NJ): IEEE; 2017. p. 1–6.* <https://dx.doi.org/10.1109/DASC.2017.8102037>
- [7] Safavi S, Safavi MA, Hamid H, Fallah S. Multi-sensor fault detection, identification, isolation and health forecasting for autonomous vehicles. *Sensors*. 2021;21(7):2547. <https://doi.org/10.3390/s21072547>
- [8] Zhang W, Ning Y, Suo C. A method based on multi-sensor data fusion for UAV safety distance diagnosis. *Electronics*. 2019;8(12):1467. <https://doi.org/10.3390/electronics8121467>
- [9] Khaneghaei M, Asadi D, Mowla MN, Dişken G. Experimental motor fault detection and identification of a quadrotor UAV using a hybrid deep learning approach. *Int J Dyn Control*. 2025;13:281. <https://doi.org/10.1007/s40435-025-01786-4>
- [10] Zhang Y, Yang C, Liu W, Qie T, Wang W, Li H. Unmanned delivery aerial vehicles fault detection method based on enhanced spatiotemporal feature fusion framework and multi-head attention mechanism classifier. *Adv Eng Inform*. 2026;71:104249. <https://doi.org/10.1016/j.aei.2025.104249>
- [11] Huang B, Kang F, Liu X. Intelligent UAV-based deep learning system for multi-class concrete dam damage detection. *Autom Constr*. 2026;182:106730. <https://doi.org/10.1016/j.autcon.2025.106730>
- [12] Liu Y, Sun H, Li W, He Y. Advancing traffic object detection in complex environments: a deep learning approach with vehicle-mounted UAV data. *Knowl Inf Syst*. 2026;68(1):12. <https://doi.org/10.1007/s10115-025-02621-0>
- [13] Husin NA, Baktiar NAHM, Tagang VU, Khairunniza-Bejo S, Yusuf MFM. UAV-based integration of RGB, thermal, and structural features with machine learning for BSR severity detection. *J Agric Food Res*. 2026:102660. <https://doi.org/10.1016/j.jafr.2026.102660>
- [14] Et-taleby A, Chaibi Y, Allouhi A, El-Mekkaoui J. Towards autonomous photovoltaic panels health monitoring: UAV-based imaging, AI, and economic perspectives. *Sol Energy Mater Sol Cells*. 2026;296:114024. <https://doi.org/10.1016/j.solmat.2025.114024>
- [15] Al-Haddad LA, Jaber AA, Mahdi NM, Al-Haddad SA, Al-Karkhi MI, Al-Sharify ZT, et al. Protocol for UAV fault diagnosis using signal processing and machine learning. *STAR Protoc*. 2024;5(4):103351. <https://doi.org/10.1016/j.xpro.2024.103351>
- [16] Al-Haddad LA, Jaber AA, Al-Haddad SA, Al-Muslim YM. Fault diagnosis of actuator damage in UAVs using embedded recorded data and stacked machine learning models. *J Supercomput*. 2024;80(3):3005–24. <https://doi.org/10.1007/s11227-023-05584-7>

- [17] Khatiri S, Panichella S, Tonella P. Simulation-based test case generation for UAVs near real flights. In: Proceedings of the IEEE Conference on Software Testing, Verification and Validation (ICST); 2023 Apr 16–20; Dublin, Ireland. Piscataway (NJ): IEEE; 2023. p. 281–92. <https://doi.org/10.1109/ICST57152.2023.00034>
- [18] Setlak L, Kowalik R, Łusiak T. Practical use of composite materials used in military aircraft. *Materials*. 2021;14(17):4812. <https://doi.org/10.3390/ma14174812>
- [19] Puchalski R, Bondyra A, Giernacki W, Gardecki S, Kasiński A. Actuator fault detection and isolation system for multirotor UAVs. In: Proceedings of the 26th International Conference on Methods and Models in Automation and Robotics (MMAR); 2022 Aug 22–25; Międzyzdroje, Poland. Piscataway (NJ): IEEE; 2022. p. 1–6. <https://doi.org/10.1109/MMAR55195.2022.9874283>
- [20] Rewicki F, Gawlikowski J, Niebling J, Denzler J. Unraveling anomalies in time: unsupervised discovery and isolation of anomalous behavior in bio-regenerative life support system telemetry. In: Bifet A, Krilavičius T, Miliou I, Nowaczyk S, editors. *Machine Learning and Knowledge Discovery in Databases. Applied Data Science Track. ECML PKDD 2024. Lecture Notes in Computer Science*, vol. 14949. Cham: Springer; 2024. p. 249–64. <https://doi.org/10.48550/arXiv.2406.09825>
- [21] Elhalwagy A, Kalganova T. Hybridization of capsule and LSTM networks for unsupervised anomaly detection on multivariate data [preprint]. arXiv:2202.05538. 2022.
- [22] Li B. Unsupervised temporal anomaly detection: time series, data stream, and interpretability [dissertation]. Dortmund: Technische Universität Dortmund; 2024.
- [23] Li Z, Zhang Y, Ai J, Zhao Y. A lightweight and explainable data-driven scheme for fault detection of aerospace sensors. *IEEE Trans Aerosp Electron Syst*. 2023;59(6):8392–8410. <https://doi.org/10.1109/TAES.2023.3303855>
- [24] Jang K, Pilario KES, Lee N, Moon I, Na J. Explainable artificial intelligence for fault diagnosis of industrial processes. *IEEE Trans Ind Inform*. 2025;21(1):4–11. <https://doi.org/10.1109/TII.2023.3240601>
- [25] Shukla B, Fan IS, Jennions I. Opportunities for explainable AI in aerospace predictive maintenance. In: Proceedings of the PHM Society European Conference. Vol 5(1). PHM Society; 2020. <https://doi.org/10.36001/phme.2020.v5i1.1231>
- [26] Seckin AC, Emir B, Bulut D, Eren B, et al. An experimental study on multisensor-based propeller damage detection in drones using deep learning algorithms. *IEEE Sens J*. 2025. <https://doi.org/10.1109/JSEN.2025.3645391>
- [27] Zhang Z, Xu L, Li Y, Zhang Q, Li B. Metastructure-enhanced single-sensor method for UAV propeller fault localization. *Mech Syst Signal Process*. 2025;238:113151. <https://doi.org/10.1016/j.ymsp.2025.113151>
- [28] Berdibayev Y, Lee HG. Mobility-aware multi-sensor UAV health monitoring via vision and vibration fusion. *J Korean Soc Ind Syst Eng*. 2025;48(2):163–77. <https://doi.org/10.11627/jksie.2025.48.2.163>
- [29] Wang Y. Machine learning-driven EKF optimization for UAV attitude error reduction. In: 7th International Conference on Frontier Technologies of Information and Computer (ICFTIC). 2025:182–187. <https://doi.org/10.1109/ICFTIC68075.2025.11325035>
- [30] Ma B, Liu Z, Dang Q, Zhao W, Wang J, Cheng Y, Yuan Z. Deep reinforcement learning of UAV tracking control under wind disturbance. *IEEE Trans Instrum Meas*. 2023;72:1–13. <https://doi.org/10.1109/TIM.2023.3265741>

-
- [31] Wang T, Qin R, Chen Y, Snoussi H, Choi C. A reinforcement learning approach for UAV target searching and tracking. *Multimed Tools Appl.* 2019;78(4):4347–4364. <https://doi.org/10.1007/s11042-018-5739-5>
- [32] Ma B, Liu Z, Zhao W, Yuan J, Long H, Wang X, et al. Target tracking control of UAV through deep reinforcement learning. *IEEE Trans Intell Transp Syst.* 2023;24(6):5983–6000. <https://doi.org/10.1109/TITS.2023.3249900>
- [33] Xie J, Peng X, Wang H, Niu W, Zheng X. UAV autonomous tracking and landing based on deep reinforcement learning. *Sensors.* 2020;20(19):5630. <https://doi.org/10.3390/s20195630>
- [34] Li B, Wu Y. Path planning for UAV ground target tracking via deep reinforcement learning. *IEEE Access.* 2020;8:29064–74. <https://doi.org/10.1109/ACCESS.2020.2971780>
- [35] Humais MA, Chehadeh M, Boiko I, Zweiri Y. Analysis of the effect of time delay for unmanned aerial vehicles with applications to vision based navigation [preprint]. *arXiv:2209.01933.* 2022.
- [36] Wang W, Zhang W. Jittering effects analysis and beam training design for UAV millimeter-wave communications. *IEEE Trans Wirel Commun.* 2021;21(5):3131–3146. <https://doi.org/10.1109/TWC.2021.3118558>
- [37] Zhang Y, Li S, Wu F, Li C, Long X, Zhang A. CausalFPS: a causal discovery-based method for UAV flight parameter selection. *Mech Syst Signal Process.* 2026;245:113851. <https://doi.org/10.1016/j.ymssp.2026.113851>