Date of publication xxxx 00, 0000, date of current version xxxx 00, 0000.

*Digital Object Identifier 10.1109/ACCESS.2017.DOI*

# Anomaly Management in Unmanned Aerial Vehicles: A Systematic Literature Review

**IVAN TAN[1], CHRISTOPHER M. POSKITT[1], LINGXIAO JIANG[1], LWIN KHIN SHAR[1]**
[1]School of Computing and Information Systems, Singapore Management University, Singapore (e-mail: ivantan@smu.edu.sg; cposkitt@smu.edu.sg; lxjiang@smu.edu.sg; lkshar@smu.edu.sg)

This work received no specific funding.

**ABSTRACT** Unmanned Aerial Vehicles (UAVs) are increasingly deployed in safety-critical applications such as logistics, surveillance, disaster response, and urban air mobility. While their autonomy enables powerful capabilities, it also introduces vulnerabilities due to hardware faults, software defects, communication failures, and adversarial interference. This survey presents a comprehensive review of research studies closely related to UAV anomalies published between 2015 and 2025, covering 111 papers from academic and industrial sources. We introduce a unified five-pillar taxonomy—anomaly generation, prevention, detection, recovery, and analysis—that organizes existing work across the full anomaly management lifecycle. In contrast to prior surveys that focus primarily on detection algorithms, our review integrates operational and regulatory perspectives, explicitly linking technical anomalies to policy enforcement and compliance requirements. We systematically compare detection techniques, datasets, simulators, and evaluation practices, revealing significant fragmentation in datasets, limited real-world validation, and a lack of standardized real-time benchmarks. Our synthesis highlights key research challenges, including the sim-to-real gap, limited interpretability of learning-based detectors, and the scarcity of policy-aware anomaly management frameworks. Based on these findings, we outline emerging research opportunities for adaptive, explainable, and benchmark-driven anomaly management systems that support safe, transparent, and reliable UAV operations. We release our metadata for all the papers reviewed, as well as filters for easy sorting at [1].

**INDEX TERMS** Unmanned aerial vehicles (UAVs), anomaly detection, cyber-physical systems, fault diagnosis, machine learning, safety-critical systems

## I. INTRODUCTION

Unmanned Aerial Vehicles (UAVs), commonly referred to as drones, are increasingly deployed in domains such as logistics, traffic monitoring, agriculture, environmental monitoring, disaster response, and defense [2], [3]. As these systems become more autonomous and operate in complex environments, ensuring their safe and reliable behavior becomes increasingly challenging. Unlike traditional aircraft, UAVs rely heavily on lightweight sensors, onboard computation, and wireless communication links, which are susceptible to hardware faults, environmental disturbances, and adversarial interference [4], [5]. These vulnerabilities can lead to anomalous behavior that threatens operational safety and regulatory compliance, making reliable monitoring and anomaly management essential to prevent accidents, airspace violations, and system misuse.

### *MOTIVATION*

One of the pressing challenges in drone operations is the detection of anomalies, unusual patterns or behaviors that deviate from expected norms. Anomalies may arise from hardware malfunctions (e.g., sensor drift, motor faults), environmental factors (e.g., GPS multipath, strong winds), or malicious interference (e.g., GPS spoofing, cyber intrusions) [6]. If unmanaged, these anomalies can lead to policy violations, such as breaching geofenced boundaries, exceeding altitude restrictions, or intruding into no-fly zones. Such incidents are not merely hypothetical: unlawful drone operations in protected areas have resulted in criminal charges in Singa-

pore[1], drone incursions have disrupted airport operations[2], and authorities have pursued enforcement actions for flights inside temporary flight restriction (TFR) airspace and at unsafe altitudes[3]. Beyond safety risks, such violations can undermine public trust in drone technologies and expose operators to legal or regulatory consequences. As UAV usage expands in urban and shared airspaces, the ability to generate (simulate) anomalies, prevent, detect, analyze, and recover from anomalies becomes not only a technical necessity but also a regulatory mandate.

While several surveys have examined anomaly detection in UAVs and cyber-physical systems (e.g., [4], [10]–[13]), existing reviews predominantly focus on detection techniques alone, often emphasizing machine learning or deep learning models and their performance metrics. Such detection-centric perspectives, although valuable, provide only a partial view of the anomaly management problem in UAV systems. In practice, effective anomaly management requires a broader lifecycle perspective that encompasses anomaly generation (and simulation) for testing and benchmarking, preventive mechanisms, runtime detection, post-incident analysis, and recovery or mitigation strategies.

Moreover, existing surveys tend to treat anomalies either as abstract statistical deviations or as isolated safety violations, without explicitly linking technical anomaly detection to operational policies and regulatory constraints governing UAV flights. As a result, key questions remain insufficiently addressed, such as how different anomaly types propagate into policy violations, how prevention and recovery mechanisms interact with detection models, and how datasets and simulators reflect real-world regulatory scenarios. This fragmentation hinders the development of standardized benchmarks and limits the transferability of research outcomes to real-world, safety-critical drone operations [14].

### UNIQUENESS OF THIS SLR

To address these gaps, this systematic literature review (SLR) adopts a holistic, lifecycle-oriented perspective on UAV anomalies, organizing prior work along five pillars: generation, prevention, detection, recovery, and analysis. By connecting anomaly sources, detection approaches, datasets, and application contexts such as policy-aware considerations, this SLR provides a unified view that goes beyond algorithmic performance. In doing so, it complements existing detection-focused surveys and offers insights into emerging trends, open challenges, and research gaps that must be addressed to support safe, compliant, and trustworthy UAV deployments.

The main contributions of this SLR are as follows:

- **Lifecycle-Oriented Taxonomy:** We introduce a five-pillar taxonomy for UAV anomaly management covering anomaly generation, prevention, detection, recovery/mitigation, and analysis/learning. Unlike prior UAV surveys that focus mainly on detection algorithms, this taxonomy captures the broader anomaly-management lifecycle and provides a structured basis for comparing existing work.
- **Systematic Synthesis of UAV Anomaly Research:** We systematically review 111 studies published between 2015 and 2025 and synthesize them across multiple dimensions, including anomaly source, detection technique, data source, simulator usage, evaluation practice, and application context. This provides a consolidated view of how UAV anomaly research has evolved across the lifecycle rather than within detection alone.
- **Identification of Structural Gaps and Future Needs:** Through quantitative and qualitative synthesis, we identify persistent gaps in the literature, including the dominance of detection-centric approaches, fragmented datasets and benchmarks, limited real-world validation, weak lifecycle integration, and the underdevelopment of governance- and policy-aware perspectives. These findings motivate the need for more integrated, deployment-oriented, and safety-conscious UAV anomaly-management frameworks.

To support transparency and reuse, in [1], we provide a companion metadata resource containing the reviewed papers and their extracted attributes, together with searchable filters for exploration.

### KEY FINDINGS FROM THIS SLR

- **Detection dominates the lifecycle:** the literature is overwhelmingly detection-centric, with comparatively little work that treats generation, prevention, and especially recovery/mitigation as first-class research contributions.
- **Operational and policy-aware anomalies are underrepresented:** despite their real-world safety and compliance impact, mission- and rule-violation–level anomalies are modeled far less often than telemetry-visible faults.
- **Dataset realism remains a bottleneck:** many studies rely on simulation or private flight logs, with few widely adopted public benchmarks, contributing to weak comparability and limited reproducibility.
- **Sim-to-real generalization is a persistent gap:** methods validated in controlled simulators often degrade under real disturbances (e.g., vibration, RF noise, weather variability), motivating more robust cross-platform evaluation.
- **Evaluation skews toward offline accuracy:** common reporting emphasizes accuracy/F1, while deployment-relevant factors (e.g., latency, energy, and real-time constraints) are less consistently assessed.

[1]The Straits Times, "4 men to be charged for offences related to operating unmanned aircraft in protected areas," Online, Jul. 2024, accessed: 2026-02-25. [Online]. Available: https://www.straitstimes.com/singapore/four-men-to-be-charged-for-offences-related-to-operating-unmanned-aircraft-in-protected-areas [7]

[2]M. Bryant, "Danish pm: airport drone incursion a 'serious attack' on critical infrastructure," The Guardian (Online), Sep. 2025, accessed: 2026-02-25. [Online]. Available: https://www.theguardian.com/world/2025/sep/23/drone-sightings-cause-disruption-delays-norway-denmark-airports [8]

[3]United States Department of Justice, U.S. Attorney's Office, Northern District of California, "Drone operator charged with violating temporary flight restrictions in area surrounding levi's stadium during nfl game," Press release (Online), Feb. 2026, accessed: 2026-02-25. [Online]. Available: https://www.justice.gov/usao-ndca/pr/drone-operator-charged-violating-temporary-flight-restrictions-area-surrounding-levis [9]

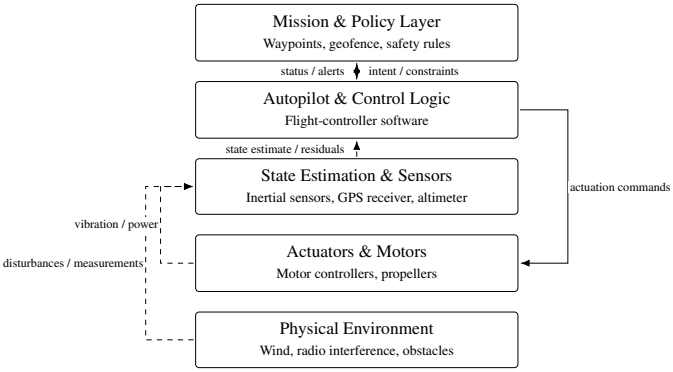


**FIGURE 1. Simplified UAV system stack, adapted by the authors from typical UAV platform and avionics architectures discussed in [3], [15]. The layers indicate where anomalies may originate and how their effects can propagate upward toward mission- and policy-level consequences.**

## II. BACKGROUND AND FOUNDATIONS

Modern UAVs are complex cyber-physical systems composed of tightly coupled sensing, control, communication, and mission software layers [12]. The interactions between these components shape how anomalies emerge, propagate, and ultimately affect flight behavior and policy compliance. Understanding anomaly management therefore requires grounding in the architectural structure of UAV systems and the operational contexts in which faults arise.

This section provides the technical background for the remainder of the survey. It first outlines key elements of UAV system architecture and flight-control stacks, then reviews how anomalies are defined and categorized in prior literature. These foundations support the lifecycle-oriented taxonomy and comparative analysis presented in the following sections.

### A. DRONE ARCHITECTURES AND FLIGHT STACKS

A typical UAV platform integrates onboard sensors, a flight controller, communication modules, and mission-specific payloads [15]. Core flight behavior is governed by autopilot stacks such as PX4, ArduPilot, and proprietary systems developed by manufacturers like DJI. These stacks implement stabilization, navigation, and mission execution while exposing telemetry and logging interfaces that enable monitoring, customization, and the development of anomaly detection or policy-enforcement mechanisms.

Simulation environments such as Gazebo [16], [17], AirSim [18], and JSBSim are commonly used to evaluate these mechanisms prior to deployment. The architectural configuration of a UAV (including its sensors, compute resources, and telemetry interfaces) directly influences how anomalies can be observed and detected.

Figure 1 illustrates a simplified UAV system stack used throughout this survey to highlight where anomalies originate, how they propagate across control layers, and where policy violations ultimately manifest.

### B. TYPES OF ANOMALIES IN UAVS

Figure 1 provides a stack-oriented view of how these anomaly types relate to the UAV system architecture. In particular, sensor-related anomalies arise primarily in the State Estimation & Sensors layer, actuator and propulsion anomalies in the Actuators & Motors layer, and software/control anomalies in the Autopilot / Control Logic layer. Environmental disturbances originate in the Physical Environment layer, while human and operational anomalies typically manifest at the Mission & Policy layer through unsafe decisions, mission deviations, or rule violations. Communication anomalies are cross-cutting: although not confined to a single layer, they affect the telemetry, coordination, and command pathways that connect the stack, and can therefore propagate upward into control degradation or policy-level failures.

Anomalies in drones can emerge from diverse sources and manifest in different ways:

- **Sensor Faults**: Degradation, noise, or drift in critical sensors such as IMUs, GPS receivers, or altimeters can lead to incorrect state estimation.
- **Software and Control Logic Anomalies**: Faults in autopilot software, navigation and state-estimation pipelines, or mission execution logic—such as incorrect parameter tuning, firmware bugs, timing errors, or unstable control loops—can result in intermittent or persistent deviations from expected flight behavior.
- **Actuator and Control Failures**: Motor degradation, propeller imbalance, or control signal corruption may result in unstable flight dynamics.
- **Communication Anomalies:** Packet loss, latency spikes, bandwidth degradation, or message corruption in command-and-control and telemetry links can disrupt coordination, delay safety responses, and trigger cascading mission failures, particularly in BVLOS and swarm operations [19].
- **Environmental Disturbances**: Sudden gusts of wind, magnetic interference, or GPS multipath effects can induce unexpected behavior.
- **Adversarial or Malicious Attacks**: GPS spoofing, sensor injection, or cyber intrusions deliberately trigger deviations from intended mission paths.

These anomalies can cause subtle performance degradation or catastrophic failures if not prevented, detected, or recovered promptly. Importantly, they often serve as precursors to policy violations. For example, a drifting GPS sensor may cause a drone to unknowingly cross into restricted airspace. Similarly, a faulty actuator would cause the failure of a mission, or even worse, a crash. Therefore, it is also important to generate such anomalies in a simulation environment and analyze them to devise methods to prevent, detect, and recover.

## III. RESEARCH OBJECTIVES AND QUESTIONS

This survey systematically reviews the literature on UAV anomaly management, with particular emphasis on how anomaly detection research relates to operational policies and safety constraints. Existing UAV surveys do not focus on examining how anomalies emerge, propagate, and are handled

across the broader anomaly management lifecycle. Moreover, the relationship between low-level technical anomalies and higher-level policy violations (e.g., geofence or altitude breaches) remains largely unexplored.

To address these gaps, this survey adopts a lifecycle-oriented perspective encompassing anomaly generation, prevention, detection, recovery, and post-incident analysis, while examining how current approaches relate to policy enforcement and operational safety. Specifically, we aim to: (i) identify the range of anomaly management techniques applied to UAV systems, (ii) analyze how these approaches relate to operational policies and regulatory constraints, (iii) document commonly used datasets, simulators, and evaluation practices, and (iv) synthesize key limitations and research opportunities for policy-aware and lifecycle-integrated UAV anomaly management frameworks.

We pose the following research questions in our study:

- **RQ1**: What types of anomalies occur in UAV systems, and how are they categorized?
- **RQ2**: What methods have been proposed for anomaly generation, prevention, detection, recovery, and analysis?
- **RQ3**: What datasets, simulators, and evaluation metrics are commonly used in UAV anomaly studies?
- **RQ4**: What are the main challenges, limitations, and open research issues reported?
- **RQ5**: What are the dominant research trends over time, and which gaps remain?

## IV. METHODOLOGY

To ensure methodological rigor, transparency, and reproducibility, our review procedure is informed by Kitchenham's established guidelines for conducting systematic literature reviews [20]. Although originally proposed in the context of software engineering, Kitchenham's methodology has since been widely adopted across adjacent domains, including cyber-physical systems, robotics, safety-critical systems, and applied artificial intelligence, due to its domain-agnostic emphasis on systematic search, explicit inclusion and exclusion criteria, structured data extraction, and traceable synthesis.

These characteristics are particularly well aligned with UAV anomaly detection research, which spans software, hardware, communication, and operational dimensions. Accordingly, Kitchenham's principles guided the formulation of research questions, search strategy, study selection, data extraction, and synthesis activities in this review, ensuring consistency and reproducibility across a heterogeneous and interdisciplinary body of literature.

### A. LITERATURE SEARCH STRATEGY

A structured search was conducted in December 2025 across major digital libraries: IEEE Xplore, ACM Digital Library and Google Scholar. The core search string used was: ("drone anomaly" OR "drone anomalies" OR "UAV anomaly" OR "UAV anomalies" OR ("anomaly detection" AND (drone OR UAV)) OR "flight anomaly" OR "flight anomalies" OR ("fault detection" AND UAV) OR "drone safety"). This string was adapted as needed to the syntax and search capabilities of IEEE Xplore, ACM Digital Library, and Google Scholar. To maintain focus on recent and relevant studies, the search was limited to publications between 2015 and 2025 written in English and involving either simulation or real-world UAV platforms. Citation chasing was also performed, where reference lists from key papers were also reviewed to capture additional studies not indexed by databases.

### B. SCREENING AND INCLUSION CRITERIA

An initial pool of papers was collected through database searches and backward/forward snowballing.

To handle duplicate records arising from overlapping digital libraries and citation chaining, all retrieved entries were first normalised using their titles, author lists, publication venues, and DOIs where available. Exact duplicates were automatically removed, while near-duplicates (e.g., extended journal versions of conference papers, tool companion papers for full-length papers) were manually inspected, retaining only the most complete or authoritative version. Following de-duplication, titles and abstracts were screened to exclude studies that were (i) not specific to UAVs (e.g., generic CPS or IoT anomaly detection without UAV instantiation), (ii) non-technical or conceptual in nature (e.g., opinion pieces, position papers), or (iii) not peer-reviewed (e.g., white papers, technical reports, theses).

Full-text reviews were then conducted to further exclude works that (iv) addressed perception or object-detection anomalies unrelated to system or operational behavior, or (v) lacked sufficient methodological detail to support reproducibility or comparison.

A study was included if it (i) explicitly analyzed anomaly detection or anomaly management in UAV systems, (ii) proposed, evaluated, or compared detection, prevention, recovery, or policy-enforcement mechanisms, or (iii) reported datasets, simulators, or evaluation metrics applicable to UAV anomaly research. This multi-stage screening process (see Figure 2) followed the PRISMA-style flow [21] and resulted in 111 studies, which constitute the core dataset analyzed in this review. The full set of papers can be explored at [1].

### C. DATA EXTRACTION AND CODING

For each included study, a structured data extraction process was conducted to collect attributes relevant to UAV anomaly management research. The extracted attributes captured several key dimensions of the literature, including the source of anomalies, the techniques used for detection, the type of data employed for evaluation, the application context, and the stage of the anomaly management lifecycle addressed by the study. These attributes were later used to construct the taxonomy described in Section V.

The analytical framework used for classification was based on a five-pillar anomaly management taxonomy. The initial structure of this taxonomy was informed by prior research on anomaly and fault management in cyber-physical systems, where system resilience is often conceptualized across lifecy-

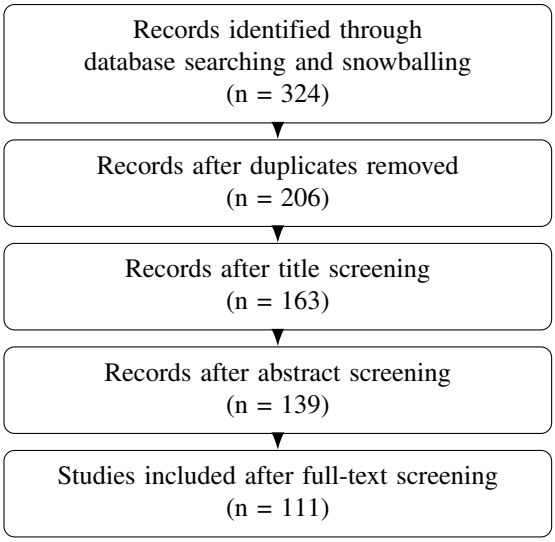


**FIGURE 2. PRISMA-style study selection flow.**

cle stages such as anomaly generation, detection, mitigation, and post-incident analysis [22]. The taxonomy was subsequently refined iteratively during the data extraction process to ensure that the categories reflect patterns observed from the reviewed literature. This iterative refinement followed a grounded categorization approach, where emerging themes from the studies were incorporated into the framework.

Data extraction and labeling were performed manually. The labels can be accessed on the website via the relevant filters [1]. To improve reliability and reduce classification bias, extracted attributes and assigned categories were reviewed and cross-checked for consistency across the dataset. Discrepancies in classification were resolved through re-examination of the original papers and refinement of the coding scheme where necessary. Because a single study may contribute to more than one lifecycle pillar or classification dimension, counts across categories are not necessarily mutually exclusive; percentages are reported with respect to the full corpus of 111 reviewed studies unless otherwise stated.

### *D. SYNTHESIS AND ANALYSIS*

Quantitative counts and qualitative synthesis were combined to reveal research trends, popular frameworks, and underexplored areas. Studies were grouped by taxonomy categories to analyze distribution patterns, simulator usage, dataset prevalence, and evaluation metrics. Descriptive statistics (frequency counts, annual trends) were complemented by thematic analysis to identify common challenges such as data scarcity, sim-to-real transfer limitations, and interpretability issues. This mixed-method approach supports a balanced understanding of how UAV anomaly-related research has evolved and where opportunities for policy-aware and adaptive frameworks remain.

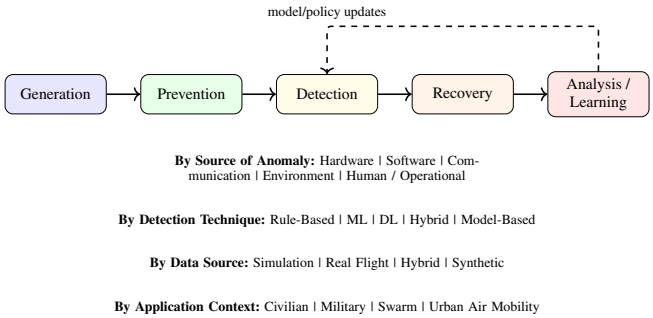


**FIGURE 3. Taxonomy of UAV anomaly management organized by lifecycle pillars and four complementary classification axes: anomaly source, detection technique, data source, and application context.**

## V. TAXONOMY

Based on the study selection and data extraction process (Section IV), we derived a taxonomy (Figure 3) that serves as the analytical framework for organising and interpreting the evaluation results presented in Section VI. Rather than being imposed a priori, the taxonomy emerged through iterative screening, data extraction, and cross-comparison of the selected studies. Across the literature, five recurring stages of anomaly handling consistently appeared: anomaly generation, prevention, detection, recovery/mitigation, and analysis/learning. Together, these stages capture the lifecycle through which UAV anomalies are managed, from pre-deployment testing to post-incident adaptation.

In addition to this lifecycle perspective, four cross-cutting dimensions were identified that contextualize prior work: anomaly source, detection technique, data source, and application context. These dimensions are orthogonal axes that reflect how studies define their problem settings, evaluation environments, and deployment assumptions.

As illustrated in Figure 3, UAV anomaly management can be understood as a lifecycle supported by complementary analytical dimensions. The horizontal sequence represents the progression from anomaly generation and prevention to detection, mitigation, and post-incident learning. The vertical axes capture how approaches vary according to anomaly types, detection methodology, available data, and operational contexts. This two-dimensional perspective highlights that effective UAV safety assurance extends beyond detection accuracy to include preventive design, resilient recovery strategies, and continuous learning from operational experience. The following subsections elaborate on each pillar.

### *A. TAXONOMY BY RESEARCH PILLARS*

The primary dimension of the taxonomy organizes UAV anomaly research according to the stage of the anomaly management lifecycle addressed by each study. Across the reviewed literature, five recurring pillars were identified aligned to the stages: anomaly generation, prevention, detection, recovery/mitigation, and analysis/learning. These pillars capture how UAV anomalies are handled across the system lifecycle, from pre-deployment testing and preventive design

to runtime monitoring and post-incident learning. Each pillar represents a distinct research focus, while the cross-cutting axes (anomaly source, detection technique, data source, and application context) describe how these approaches are implemented and evaluated. This section serves as an overview; detailed methodological comparisons across these pillars are presented later in the evaluation and discussion sections.

#### 1) Generation

The generation pillar covers methods that deliberately create or simulate anomalies in order to test, train, or benchmark UAV anomaly-management systems. Representative examples include simulator-based fault injection, fuzzing of autopilot or communication software, and synthetic augmentation of anomaly datasets.

#### 2) Prevention

The prevention pillar includes methods that aim to reduce the likelihood of anomalies before they occur. Representative examples include hardware redundancy, pre-flight safety and health checks, predictive maintenance, and policy- or security-oriented safeguards such as geofencing and anti-spoofing protections.

#### 3) Detection

The detection pillar focuses on identifying anomalies during flight or from post-flight data. Representative examples include rule-based threshold checks, machine-learning classifiers over telemetry, and deep learning methods such as autoencoders, recurrent networks, and transformer-based models.

#### 4) Recovery and Mitigation

The recovery/mitigation pillar concerns how UAV systems respond after an anomaly has been detected in order to maintain safety and reduce mission disruption. Representative examples include fail-safe actions such as return-to-home or emergency landing, switching to backup sensors or fault-tolerant control modes, and coordination-aware adaptation in multi-UAV systems.

#### 5) Analysis and Learning

The analysis/learning pillar focuses on post-incident investigation and mechanisms for improving future resilience. Representative examples include post-flight log analysis, anomaly replay, and interpretable or adaptive methods that support root-cause analysis and continuous improvement.

### B. BY SOURCE OF ANOMALY

Anomalies in UAVs can originate from a wide range of sources spanning hardware, software, communication systems, environmental interactions, and human or operational factors. Hardware-level anomalies include sensor degradation, IMU drift, GPS errors, motor or propeller failures, and power system faults such as sudden voltage drops or capacity loss. Examples of this are actuator-level anomalies, particularly those involving motors and propellers, which are commonly addressed through condition monitoring techniques that analyze vibration, acoustic, or electrical signals to detect early signs of mechanical degradation [23]. Software-level anomalies [24] arise from autopilot misconfigurations, unstable control logic, firmware bugs, or navigation errors in waypoint planning. At the network layer, communication anomalies result from packet loss, latency, or deliberate attacks such as jamming and GPS spoofing. Environmental disturbances, including wind gusts, rain, magnetic interference, or terrain occlusion, can trigger unexpected deviations in flight paths. Finally, human-induced anomalies encompass pilot mistakes, unsafe maneuvers, and policy breaches (e.g., entering restricted airspace or exceeding altitude limits). Understanding the root of each anomaly is essential for designing targeted detection mechanisms and appropriate mitigation responses.

### C. BY DETECTION TECHNIQUE

UAV anomaly detection methods vary widely in complexity, interpretability, and data requirements. Rule-based and statistical techniques offer simple, lightweight implementations using threshold checks or residual analysis, but often suffer from limited adaptability [25]–[27]. Classical machine learning (ML) methods, such as Support Vector Machines, Decision Trees, and K-Means clustering, capture complex relationships between telemetry variables while remaining interpretable. Deep learning (DL) approaches, including Autoencoders, CNNs, RNNs, and Transformers, learn hierarchical and temporal patterns from large-scale sensor or flight data, achieving high accuracy at the cost of computational demand. Model-based and physics-informed approaches leverage flight dynamics models or control-theoretic residuals to detect deviations from expected UAV behavior, providing physically grounded insights. Finally, hybrid and ensemble methods combine rule-based reasoning with learning-based models to balance robustness, explainability, and resource efficiency. The diversity of techniques highlights the ongoing trade-off between precision, interpretability, and real-time feasibility in UAV anomaly detection.

### D. BY DATA SOURCE

The effectiveness and generalizability of anomaly detection approaches depend strongly on the nature and quality of available data. Simulation-based data from environments such as Gazebo, AirSim, and UAV-oriented simulation frameworks allow controlled fault injection and safe experimentation with diverse anomaly scenarios [19], [28]–[30]. Such environments are especially useful for studying rare, hazardous, or difficult-to-reproduce faults, although they still face challenges related to the sim-to-real gap [14]. Real-world flight data, including PX4, ArduPilot and DJI telemetry logs offer high-fidelity insights but are expensive to collect and often lack sufficient examples of rare or critical anomalies. Hybrid datasets combine simulated and real data, using transfer learning or domain adaptation to bridge discrepancies between

virtual and physical domains. Meanwhile, synthetic or augmented data artificially expand limited datasets through noise injection, scaling, or adversarial perturbations to enhance training diversity. Selecting appropriate data sources thus remains a key determinant of anomaly detection performance, realism, and reproducibility.

### E. BY APPLICATION CONTEXT

UAV anomaly detection techniques are influenced by the operational domain in which they are applied. In civilian applications such as delivery [31], agriculture, or infrastructure inspection, anomalies often stem from sensor degradation, GPS inaccuracies, or mission-path deviations that threaten service reliability. In military and defense settings, research emphasizes resilience against adversarial interference, stealthy attacks, GPS spoofing, and communication-level threats, where detection systems must operate under uncertain or contested environments [6], [32], [33]. For swarm and multi-UAV systems, anomaly management extends to cooperative behaviors, including detection of inter-drone inconsistencies and reconfiguration of formations to sustain mission objectives. Finally, urban air mobility (UAM) introduces policy- and safety-critical constraints for passenger and cargo UAVs, requiring stringent compliance with airspace regulations and fail-safe recovery capabilities. This contextual classification highlights that anomaly detection strategies must be tailored not only to system design but also to mission-critical and regulatory environments.

## VI. EVALUATION

This section presents the comprehensive findings of our systematic review and analysis, answering the five research questions (RQ1–RQ5) outlined in Section III. The analysis draws upon more than one hundred publications from 2015–2025, coded along a few axes (pillar, source of anomaly, data source, detection approach, and application context). The synthesis combines quantitative statistics with qualitative interpretation of representative works.

### A. RQ1: WHAT TYPES OF ANOMALIES OCCUR IN UAV SYSTEMS AND HOW ARE THEY CATEGORIZED?

Although we briefly introduced anomaly types as a technical background classification at the start of the paper, the categories reported here reflect how the reviewed literature clusters around the abstraction layers in Figure 1. In particular, hardware anomalies map primarily to the sensing and actuation layers, software/control anomalies to the autopilot and control-logic layer, communication anomalies to the command, telemetry, and coordination interfaces spanning the stack, and environmental anomalies to the physical environment layer. Human and operational anomalies, while less frequently studied, typically manifest at the mission and policy layer, where low-level faults or disturbances propagate into unsafe decisions, mission failure, or rule violations.

Analysis of the selected studies reveals five dominant sources of UAV anomalies studied in the literature (see Table 1): **hardware** (≈27%), **software/control** (≈53%), **communication/network** (≈12%), **environmental** (≈2%), and **human/operational** (≈6%). These categories consistently emerge across the literature and reflect different layers of the UAV cyber-physical stack, from low-level sensing and actuation to high-level decision-making and mission execution.

Hardware-related anomalies primarily affect physical components and onboard sensors. Commonly reported cases include motor vibration, propeller imbalance, battery degradation, and power instability [14], [34]. Sensor-level faults such as IMU bias, accelerometer noise, gyroscope drift, and GPS signal degradation are particularly prevalent, as they directly impact state estimation and flight stability [25], [35]. These anomalies are often gradual and subtle, suitable for detection through telemetry monitoring and residual analysis.

Software and control anomalies arise from faults in autopilot logic, state estimation pipelines, or mission execution software. Reported issues include unstable control loops [36], [37], incorrect parameter tuning, firmware bugs, and navigation or waypoint-planning errors [38], [39]. Unlike hardware faults, software anomalies may manifest intermittently or only under specific operating conditions, complicating reproducibility and diagnosis.

Communication and network anomalies affect command-and-control (C2) and telemetry links between the UAV, ground control stations, and other agents. These include packet loss, latency spikes, bandwidth degradation, and deliberate attacks such as jamming, spoofing, or malicious message injection [6]. Such anomalies are especially critical in beyond-visual-line-of-sight (BVLOS) and swarm operations, where communication reliability directly influences coordination, safety, and regulatory compliance.

Environmental anomalies originate from external disturbances that interact with UAV dynamics. Frequently studied examples include wind gusts, turbulence, magnetic interference, illumination changes, and weather-related effects [40], [41]. These anomalies are commonly injected or approximated in simulation environments, yet their real-world manifestations are often non-stationary and context-dependent. As a result, models trained on simulated environmental disturbances may struggle to generalize to real flight conditions.

Human and operational anomalies represent the least explored category. These include pilot-induced errors, delayed or inappropriate manual interventions, mission misconfiguration, and violations of operational or regulatory policies such as geofencing and altitude constraints. The FAA's summary report on UAS in air carrier operations notes prior evidence that in a set of 56 reported UAS accidents, human error accounted for nearly 20% of accidents, with fatigue highlighted as an important underlying factor [42]. Only a small number of studies explicitly model such behaviors, typically through post-flight log analysis or rule-based policy monitoring [36], [37]. Consequently, most existing approaches focus on low-level telemetry deviations rather than higher-level intent, decision-making, or compliance-related anomalies. Operational anomalies are particularly pronounced in

**TABLE 1. RQ1: Distribution of anomaly sources across reviewed UAV studies (111 total papers, categorised by primary category).**

| Category | # Papers | % | Key Findings |
|---|---|---|---|
| Hardware | 30 | 27% | Dominated by sensor drift, motor degradation, and battery faults; strongly telemetry-centric. |
| Software / Control | 59 | 53% | Includes control instability and firmware-level faults; limited dataset availability. |
| Communicatior | 13 | 11.7% | GPS spoofing, jamming, and packet loss increasingly studied in security contexts. |
| Environmental | 2 | 1.8% | Wind and magnetic interference commonly simulated; fewer real-flight validations. |
| Operational / Policy | 7 | 6.3% | Human- or mission-level anomalies remain significantly underexplored. |

multi-UAV and swarm settings, where coordination failures, communication delays, or inconsistent shared state estimation can lead to collective deviations even when individual agents appear nominal [43]. Large-scale real-world experiments further demonstrate that anomalies can emerge at the collective level in multi-UAV systems, where deviations are not attributable to individual agents alone but arise from swarm interactions and environmental context, as observed in the pNEUMA drone swarm experiment [44].

Overall, the literature exhibits a strong bias toward anomalies that manifest as measurable deviations in sensor or control signals. Decision-level, intent-based, and policy-related anomalies, despite their significant safety and regulatory implications remain comparatively underrepresented. This imbalance highlights an important gap in operational-context modeling and motivates the need for anomaly detection frameworks that reason beyond raw telemetry to incorporate mission intent and policy awareness.

**RQ1 Takeaway — Types of Anomalies.** Existing UAV anomaly research is heavily skewed toward hardware and sensor-level faults, while human, operational, and policy-related anomalies remain underrepresented.

### *B. RQ2: WHAT METHODS HAVE BEEN PROPOSED FOR ANOMALY GENERATION, PREVENTION, DETECTION, RECOVERY, AND ANALYSIS?*

The SLR reveals uneven coverage among the pillars. **Generation** employs fuzzing and fault-injection frameworks, e.g., [28] and [29]. **Prevention** focuses on predictive maintenance and redundancy mechanisms, with examples such as vibration-based early warning [34] and multisensor fusion reliability modeling [14]. **Detection** dominates and diversifies into four methodological families:

- Rule-/Model-Based: Thresholding and analytical redundancy checks remain common in classical works [25], [26].
- Machine Learning: SVMs, Random Forests, and k-NN feature strongly in mid-period studies such as [26], [37].
- Deep Learning: Autoencoders (AE, VAE) [35], [38], RNN/LSTM [45], CNN and Transformer architectures [40], [41], and GNN extensions [46].
- Hybrid Approaches: Combinations of residual modeling and deep learning, such as CAE+SVDD [38] and distributed LSTM-AE [45].
- LLM-Assisted Approaches: Although still emerging in UAV anomaly management, LLM-based methods are beginning to support higher-level tasks such as scenario generation, policy reasoning, incident summarization, and human-in-the-loop decision support, rather than low-latency signal-level detection itself [47]–[50].

**Detection** is by far the most widely studied pillar in the reviewed literature, encompassing methods that identify anomalies during flight or from post-flight logs [24]. Existing approaches span a broad methodological spectrum. Classical methods include rule-based thresholds, residual monitoring, and statistical checks over sensor or control signals, while machine-learning approaches apply supervised and unsupervised techniques to classify abnormal behavior from telemetry features [27], [31].

Deep learning methods are especially prominent in recent work. Autoencoders, recurrent neural networks, and transformer-based models are commonly used to capture temporal dependencies and latent normality patterns in flight data [35], [51]. Some studies further extend these models to distributed or multi-agent settings, for example by combining LSTM autoencoders and standard autoencoders across multiple UAV agents without centralized data aggregation [45]. Hybrid approaches also appear frequently, combining model-based reasoning or residual modelling with learning-based methods to improve robustness.

Detection research is not limited to telemetry-centric settings. Security-oriented methods address GPS spoofing, jamming, and malicious packet injection [6], while video-based approaches leverage spatio-temporal visual cues to detect abnormal flight behavior or environmental interactions in surveillance scenarios [40], [52]. More recent studies also explore self-supervised and weakly supervised learning to reduce annotation cost in large-scale deployments where labeled anomaly data are scarce [53]. Overall, the breadth and maturity of detection methods help explain why this pillar dominates the literature, but they also reinforce the field's current imbalance relative to the much less explored prevention and recovery stages.

Detection is by far the most mature and methodologically diverse pillar in the reviewed literature, but this diversity also reveals clear trade-offs between accuracy, interpretability, and deployability [25]–[27], [35], [38], [45], [51], [54]. Rule-based and model-based approaches remain attractive because they are lightweight, transparent, and more readily aligned with certification and auditing requirements, although they may struggle with subtle or context-dependent anomalies [25]–[27], [46], [55]. For example, a GPS degradation event

may be relatively benign in open-space flight but become safety-critical near geofenced boundaries or during waypoint tracking, where even small localization errors can propagate into mission deviation or policy violation [6], [36], [37]. Classical machine-learning methods provide greater flexibility and can capture multivariate relationships in telemetry data while retaining some degree of interpretability, but their effectiveness often depends on careful feature engineering and stable data distributions [26], [27], [31], [37]. Deep learning approaches offer stronger representation capacity for temporal and multimodal flight behavior and frequently achieve superior offline performance, yet they are more data-intensive, less transparent, and less commonly validated under strict onboard resource constraints [35], [38], [40], [45], [51], [54]. Hybrid methods seek to combine the strengths of analytical reasoning and data-driven adaptation, and are therefore promising for safety-critical UAV settings, but they also increase architectural complexity and remain difficult to compare fairly because of fragmented datasets, inconsistent metrics, and differing evaluation assumptions [38], [45], [46], [54], [55]. Taken together, the literature suggests that no single detection family is uniformly superior; instead, method selection depends strongly on the intended balance between representational power, operational transparency, and real-time feasibility [25], [35], [46], [54], [55].

**Generation** methods are comparatively rare in the reviewed literature, but they play an important supporting role in enabling controlled testing and benchmarking. Existing studies mainly rely on simulation platforms such as Gazebo [16], [17] and AirSim [18] to reproduce flight dynamics, environmental disturbances, and fault scenarios in a safe and repeatable manner. Within these environments, fault injection is used to introduce sensor failures, actuator dropouts, or communication faults under controlled conditions. Beyond simulation-only approaches, fuzzing has also been applied to UAV autopilot software and communication protocols such as MAVLink to expose hidden faults and unexpected behaviours [30]. Some works further augment anomaly datasets by injecting synthetic faults or adversarial perturbations, helping address the scarcity of real-world anomaly data. Overall, however, generation remains a relatively underexplored pillar compared with detection, and is most often used as an enabling mechanism for downstream evaluation rather than as a primary research contribution in its own right.

**Prevention** methods are relatively sparse in the reviewed literature, but they address an important stage of anomaly management by seeking to reduce the likelihood of faults or unsafe conditions before they manifest in flight. Existing approaches include hardware redundancy, such as multiple IMUs or GPS units, pre-flight safety checks that validate system readiness, and predictive maintenance for batteries, motors, and other critical components [3], [23], [34]. On the software side, prevention also includes efforts to improve the correctness of control logic through static analysis, formal verification, and model checking [3]. Additional preventive measures target operational and security risks, for example through secure communication protocols that reduce vulnerability to spoofing or jamming, as well as policy-enforcement mechanisms such as geofencing and altitude constraints that limit unsafe behavior during flight [4], [6]. Reliability-oriented inspection methods, including vibration-based monitoring, further support prevention by identifying early signs of component degradation before they become in-flight anomalies [23], [34]. More broadly, adjacent reliability-engineering research suggests useful directions for strengthening the prevention pillar, particularly through uncertainty-aware maintenance planning that accounts for inspection error, degradation state, spare dynamics, and renewal effects. Although these studies are not UAV anomaly-management works in the narrow sense, they reinforce the view that prevention should extend beyond redundancy and pre-flight checks toward lifecycle reliability planning under uncertainty [55]–[57]. Overall, however, prevention remains much less developed than detection in the UAV anomaly literature, and is rarely treated as a primary research focus in its own right.

In contrast to the detection pillar, which is largely reactive and performance-driven, prevention-oriented approaches aim to reduce the likelihood of anomalous states before they manifest during operation [3], [23]. The reviewed studies show that preventive mechanisms are typically implemented through hardware redundancy, pre-flight safety checks, predictive maintenance, secure communication safeguards, or policy-enforcement constraints [4], [6]. Such approaches are attractive because they can reduce operational risk upstream and are often easier to justify from a safety-assurance perspective. However, their coverage is generally narrower than that of detection methods, as many preventive techniques address specific classes of faults rather than the broader range of interacting anomalies encountered in real deployments [14], [34]. Predictive-maintenance and inspection-based strategies offer the advantage of identifying gradual degradation before failure occurs, but they depend on long-term condition data that are rarely standardized or publicly available in UAV settings. Formal verification and model-checking techniques provide stronger guarantees for selected software and control properties, yet they remain difficult to scale to full-stack UAV behavior under open-environment uncertainty. As a result, prevention remains less standardized and less empirically developed than detection, even though it may offer greater operational value in safety-critical settings where post hoc anomaly response is inherently costly.

**Recovery and mitigation** methods address how UAVs respond once an anomaly has been detected, with the goal of maintaining safety while minimizing mission disruption. In the reviewed literature, the most common responses are fail-safe or fallback behaviors such as return-to-home, immediate landing, or controlled hovering, which are widely used in commercial UAV systems [3]. More advanced approaches involve switching to backup sensors or invoking fault-tolerance mechanisms that preserve stability under degraded sensing or actuation conditions [34], [39]. In multi-UAV settings, resilience may also be supported through coordination-aware

adaptation, where unaffected drones respond to anomalous or degraded peers to preserve mission continuity [32], [43].

Despite the practical importance of this pillar, recovery and mitigation remain weakly represented in the reviewed literature. Most reported mechanisms are rule-based and reactive, with relatively little evidence of adaptive, learning-based, or formally integrated recovery strategies in deployed UAV systems [44] (see Table 2).

Compared with detection, recovery and mitigation remain substantially less developed and are still dominated by conservative fail-safe mechanisms rather than adaptive resilience strategies. The most common responses in the reviewed literature include return-to-home, controlled landing, hovering, or switching to backup sensing and control pathways. These approaches are operationally useful because they are simple, auditable, and compatible with existing autopilot architectures, but they are also coarse-grained and frequently prioritize immediate safety over mission continuity. More advanced fault-tolerant and coordination-aware mitigation strategies, particularly in multi-UAV settings, offer the potential to preserve partial functionality under degraded conditions, yet such methods remain comparatively rare and are often evaluated only in simulation or tightly controlled scenarios. Learning-based recovery is conceptually appealing because it could support context-sensitive responses to different anomaly types and mission states, but current evidence remains limited, partly because recovery requires deeper coupling with controllers, safety envelopes, and actuation logic than detection alone. Consequently, the literature still treats recovery primarily as an emergency fallback layer rather than as a mature design space for intelligent, lifecycle-integrated resilience.

**Analysis and learning** methods focus on understanding anomalies after they occur and using that understanding to improve future resilience. Existing work includes post-flight log analysis, root-cause investigation, and forensic diagnosis, which help identify underlying anomaly sources and explain system behavior after incidents [24], [36], [46], [58], [59]. Replay and simulation environments are also used to reproduce UAV behavior under controlled conditions for anomaly replay, diagnosis, and stress testing [18], [30]. At a broader level, public datasets and benchmarks support knowledge sharing and more consistent comparison across studies [28], [29], [53]. Some approaches also incorporate online adaptation or continuous monitoring to improve anomaly handling over time [45], [54]. More recently, interpretability has become an important theme within this pillar, with explainable anomaly detectors designed to expose the reasoning behind alerts and improve trust and transparency [46], [58]. Despite these developments, analysis/learning remains much less represented than detection, and is often treated as a secondary follow-up activity rather than a primary design objective. The primary contribution lies not in immediate runtime response, but in strengthening future resilience through post-incident understanding, diagnosis, and adaptation. Existing approaches include forensic log analysis, root-cause investigation, anomaly replay, and interpretable learning mechanisms that seek to explain why abnormal behavior was detected. These methods are valuable because they can reveal causal structure, expose latent system weaknesses, and support refinement of future detection or mitigation strategies. However, much of the current work remains retrospective and weakly connected to operational feedback loops, meaning that insights derived from analysis do not always translate into updated controllers, preventive safeguards, or improved benchmark coverage. Replay and simulator-assisted analysis provide a controlled basis for debugging and stress testing, but their usefulness is still constrained by the realism and fidelity of the underlying environments. Similarly, explainability-focused methods improve transparency and trust, yet many remain proof-of-concept tools rather than standardized components of safety assurance. As a result, although analysis and learning have strong potential to connect isolated anomaly events into a broader resilience-improvement cycle, this pillar remains underdeveloped relative to detection and is still too often treated as a secondary follow-up activity rather than a central design objective.

The quantitative distribution in Table 2 highlights a structural fragmentation in the field: most UAV anomaly studies treat detection as the primary endpoint, rather than as one component of a broader anomaly management pipeline. This dominance is partly explained by the relative tractability of detection as a research problem. Detection methods can be developed using existing telemetry datasets and evaluated with familiar classification metrics [35], [38], [45], whereas prevention and recovery require deeper coupling with controllers, safety constraints, mission objectives, and operator workflows [3], [34], [39], [43]. However, the limited attention given to these latter stages reduces the practical value of many proposed methods for safety-critical deployment. In operational UAV systems, merely identifying an anomaly is insufficient unless the system can also anticipate likely failures, maintain safety under degraded conditions, and support timely mitigation or recovery. The current imbalance therefore risks producing high-performing detection models that remain weakly connected to real resilience requirements. Addressing this gap will require lifecycle-oriented research that integrates anomaly generation, prevention, detection, recovery, and post-flight analysis within unified evaluation settings [30], [36].

**RQ2 Takeaway — Methods Across the Lifecycle.** Table 2 highlights a strong structural imbalance in the literature: UAV anomaly research is overwhelmingly detection-centric, while anomaly generation, prevention, and especially recovery remain weakly explored and rarely integrated into lifecycle-oriented resilience.

**TABLE 2. RQ2: Distribution of studies across lifecycle stages (111 total papers). Note that papers can belong to more than one category, with each having a primary lifecycle stage and potentially one or more secondary lifecycle stages.**

| Lifecycle Stage | # Papers | Key Findings |
|---|---|---|
| Generation | 2 | Primarily fault injection and fuzzing; limited integration with downstream mitigation. |
| Prevention | 14 | Focused on redundancy and predictive maintenance; few policy-level prevention models. |
| Detection | 109 | Dominant research focus; majority employ deep or hybrid models with telemetry inputs. |
| Recovery / Mitigation | 30 | Mostly rule-based fail-safe mechanisms; adaptive or learning-based recovery remains rare. |
| Analysis / Learning | 9 | Growing interest in explainability and post-flight analysis; limited causal validation. |

### C. RQ3: WHAT DATASETS, SIMULATORS, AND EVALUATION METRICS ARE COMMONLY USED?

Table 3 summarizes the main patterns in datasets, simulator usage, evaluation metrics, and validation practices across the reviewed studies. Only a minority of the reviewed studies rely on publicly available datasets, with the majority instead using proprietary flight logs or hybrid data collected through custom simulation-real pipelines. Among public resources, UIT-ADrone [28] and ALFA [29] are the most frequently referenced. These datasets vary considerably in scope and realism. UIT-ADrone provides several hundred flight sequences collected in urban traffic-monitoring scenarios, with anomalies primarily annotated at the trajectory and behavioral level (e.g., abnormal motion patterns or traffic interactions). DroneSSL [53] focuses on self-supervised learning and contains large volumes of unlabeled or weakly labeled flight data, but offers limited explicit anomaly annotations. ALFA targets fault and anomaly detection more directly, incorporating multiple injected fault types such as sensor degradation and actuator-related anomalies, though it remains constrained in terms of platform diversity and environmental variability. Across these datasets, the number of distinct anomaly instances is typically limited, and coverage of coupled, cascading, or human-in-the-loop anomalies is sparse.

As a result of these limitations, roughly 21% of studies depend on proprietary or hybrid datasets, while approximately 36% rely exclusively on simulation. Simulation platforms play a central role in UAV anomaly research, with Gazebo [60], AirSim [18], UAVSim [19], and JSBSim [61] emerging as the most commonly used environments for controlled fault injection, validation, and benchmarking. Each simulator exhibits distinct strengths and weaknesses. Gazebo offers strong integration with robotic middleware and physics engines, making it suitable for control-loop testing and sensor fault injection, but its environmental realism and aerodynamic fidelity are limited. AirSim provides photorealistic environments and supports vision-based sensing, enabling perception- and trajectory-level anomaly studies; however, its physics abstractions and limited support for diverse airframe configurations constrain its use for detailed hardware fault modeling. UAVSim focuses on communication and network-level anomalies, particularly for cybersecurity analysis, but abstracts away low-level flight dynamics. JSBSim delivers high-fidelity flight dynamics modeling and is widely used in aerospace research, yet requires substantial manual configuration and lacks native support for complex environments or multi-UAV interaction.

No single simulator in the reviewed literature appears to comprehensively model all anomaly classes or operational conditions; instead, existing platforms typically emphasize selected aspects such as flight dynamics, perception realism, communication behavior, or controlled fault injection [14], [18], [19], [58], [59]. Most support only a subset of drone configurations, sensor suites, and environmental disturbances, and they typically inject faults in isolation rather than capturing coupled failures, hardware aging, pilot intervention, or regulatory context. Consequently, while simulators are effective for controlled experimentation and repeatability, they cannot fully represent the complexity of real-world UAV operations.

Evaluation practices remain largely conventional. Accuracy-based metrics dominate, with accuracy reported in 85% of studies, precision/recall/F1 in 78%, ROC-AUC in 41%, and confusion matrices in 35%. In contrast, deployment-relevant metrics such as detection latency, throughput, robustness under resource constraints, and energy consumption are reported in less than 20% of studies. Online anomaly detection methods emphasize low-latency processing and continuous monitoring of flight data streams, often trading off detection complexity for real-time responsiveness [54]. Recent benchmark-oriented works (e.g., [38], [62]) demonstrate strong offline detection performance (often F1 scores exceeding 0.9), but also report substantial degradation during real-flight validation (approximately 0.75), highlighting the persistent sim-to-real gap. Studies like [14] and [32] therefore argue for standardized evaluation protocols that jointly report anomaly types, environmental conditions, sensor modalities, and deployment constraints to improve comparability and reproducibility. Recent studies argue that accuracy-centric evaluation is insufficient for safety-critical UAV deployment, emphasizing the need for latency-aware and reliability-focused metrics under operational constraints [34].

**TABLE 3. RQ3: Datasets, simulators, and evaluation practices.**

| Aspect | Findings |
|---|---|
| Datasets | Approximately 36% of studies rely on simulation-only data, about 21% use proprietary or hybrid datasets, and fewer than 15% evaluate on public benchmarks, indicating limited dataset standardization and reproducibility. |
| Simulators | Gazebo, AirSim, and UAVSim are the most frequently used simulators for anomaly generation and evaluation. |
| Metrics | Accuracy is the most commonly reported metric (approximately 85% of studies), followed by Precision/Recall/F1 (approximately 78%) and ROC–AUC (approximately 41%), while latency, energy, and other deployment-oriented metrics are rarely included. |
| Validation | Real-flight validation appears in fewer than 25% of studies, indicating that most evaluations remain simulation-based, offline, or limited to controlled environments. |

**RQ3 Takeaway — Data, Simulators, and Evaluation.** UAV anomaly research is still supported by a relatively weak and fragmented evaluation base: many studies rely on simulated or proprietary data, use a narrow set of common simulators, and report accuracy-oriented metrics more often than deployment-relevant measures such as latency, energy, and real-flight robustness [14], [18], [19], [28], [29], [34], [62]. As a result, current evidence is stronger for offline detection performance than for real-world operational readiness.

### *D. RQ4: WHAT ARE THE MAIN CHALLENGES, LIMITATIONS, AND OPEN RESEARCH ISSUES?*

Across the literature, four macro-level challenges recur:

1) **Data Scarcity and Bias.** Public datasets remain limited; simulated faults rarely capture coupled failures or pilot response. Works such as [38], [62] stress domain adaptation as an unresolved bottleneck.
2) **Sim-to-Real Transfer.** Models trained in AirSim or UAVSim lose reliability under real disturbances (temperature, vibration, RF noise) [14], [28].
3) **Interpretability and Regulatory Assurance.** Deep and hybrid architectures, though accurate, are opaque. Explainable pipelines using attention [58] or causal graphs [46] begin to address this but remain research prototypes.
4) **Lifecycle Fragmentation.** Few studies integrate generation, prevention, and recovery into a continuous feedback loop. Frameworks like [30] argue for "closed-loop anomaly management" spanning simulation to policy enforcement.

The main challenges and limitations identified across the reviewed literature are summarized in Table 4. Emerging needs include (i) shared multimodal datasets that jointly capture telemetry, imagery, logs, and communication traces to improve robustness and cross-study comparability [14], [50], [53]; (ii) lightweight yet assurance-ready models suitable for resource-constrained onboard inference [3], [47], [62]; and (iii) closer alignment with operational safety frameworks and standards, including ASTM F3269 and the EASA U-space regulations [63], [64].

#### 1) Data Scarcity and Bias

Public datasets remain limited, and although high-fidelity simulators such as AirSim [18], Gazebo, and UAVSim [19] are widely used, simulated faults still fail to capture several classes of real-world anomalies. Public datasets such as UIT-ADrone and ALFA provide valuable benchmarks but remain limited in diversity, coverage, and real-world operational complexity [28], [29]. In particular, simulators struggle to reproduce coupled failures (e.g., simultaneous sensor drift, actuator degradation, and communication delay), human-in-the-loop effects (e.g., pilot intervention, delayed reactions, or misconfiguration), and context-dependent disturbances such as unmodeled wind turbulence, electromagnetic interference, hardware aging, and manufacturing variability. As a result, models trained on simulated data often overfit idealized conditions and exhibit degraded performance when deployed on real platforms. Works such as [38], [62] explicitly identify this sim-to-real mismatch and the resulting need for domain adaptation as a persistent and unresolved bottleneck.

#### 2) Sim-to-Real Transfer

Although high-fidelity simulators such as AirSim and UAVSim provide controlled environments for rapid experimentation and fault injection, models trained predominantly in simulation often lose reliability when deployed under real-world operating conditions. Thus, simulators inadequately capture the combined effects of temperature variation, airframe vibration, sensor aging, radio-frequency (RF) interference, timing jitter, and asynchronous sensor failures. These factors interact in non-linear and mission-dependent ways that are difficult to model explicitly.

Furthermore, real-world anomalies frequently manifest as coupled and cascading events. For example, motor degradation inducing vibration that affects IMU readings, which in turn destabilizes state estimation and control, whereas simulations typically inject faults in isolation. Human and operational factors, such as delayed pilot intervention, configuration errors, or recovery actions taken mid-flight, are also largely absent from simulated datasets. As a result, detectors trained on simulated data tend to overfit idealized dynamics and exhibit degraded performance when evaluated on real flight logs or live deployments. Empirical studies such as [28] and [14] explicitly report this performance gap and identify domain adaptation, cross-platform validation, and real-world benchmarking as critical open challenges for robust UAV anomaly detection.

#### 3) Interpretability and Regulatory Assurance

Deep and hybrid anomaly detection architectures, while achieving strong empirical performance, still remain largely

opaque, posing challenges for certification, regulatory approval, and operational trust in safety-critical UAV deployments. Aviation regulators and mission operators require not only anomaly alerts, but also transparent justifications that can be audited, reproduced, and traced back to sensor evidence or violated constraints. Black-box models complicate post-incident forensics, hinder root-cause analysis, and make it difficult to demonstrate compliance with safety cases or policy enforcement requirements.

Recent work has begun to address this gap by incorporating explainability mechanisms into anomaly detection pipelines. Attention-based models provide temporal or feature-level saliency that highlights which signals contributed most to a detected anomaly [58], while causality-enhanced graph neural networks attempt to expose dependencies between subsystems and distinguish primary faults from downstream effects [46]. However, these approaches remain largely research prototypes: explanations are often qualitative rather than verifiable, lack standardized evaluation metrics, and are rarely integrated into end-to-end safety assurance frameworks. Bridging the gap between explainable anomaly detection and certifiable UAV assurance remains an open research challenge, requiring tighter integration between interpretable learning, formal safety models, and regulatory processes.

#### 4) Lifecycle Fragmentation

Most existing UAV anomaly studies address isolated stages of the anomaly management lifecycle such as detection, fault injection, or post-flight analysis, without integrating these components into a continuous feedback loop. As a result, anomalies identified during deployment rarely inform upstream processes such as improved anomaly generation, prevention strategies, or controller refinement. This fragmentation limits cumulative learning across flights and hinders systematic improvement of safety and robustness over time.

Recent frameworks, such as [30], argue for *closed-loop anomaly management* in which simulated anomaly generation, runtime detection, policy-aware mitigation, and post-flight analysis are tightly coupled. In such frameworks, detected anomalies feed back into simulation environments and fuzzing engines to expand coverage, refine detection thresholds, and stress-test recovery policies under newly observed conditions. However, end-to-end implementations of this vision remain rare in practice, with most studies stopping at detection or offline evaluation. Bridging this gap requires unified system architectures, standardized interfaces between simulation and deployment, and lifecycle-aware benchmarks that evaluate not only detection accuracy but also adaptation, recovery effectiveness, and long-term safety improvement.

**TABLE 4. RQ4: Key challenges and limitations in UAV anomaly research.**

| Challenge | Description |
|---|---|
| Data Scarcity | Limited public datasets |
| Sim-to-Real Gap | 36% rely exclusively on simulation |
| Interpretability | >70% of studies use deep learning models |
| Lifecycle Fragmentation | Detection accounts for 98% of papers |

**RQ4 Takeaway — Challenges and Limitations.** Persistent challenges in UAV anomaly management include limited public and realistic datasets, degraded transfer from simulation to real deployment, the opacity of high-performing deep models, and poor integration between anomaly detection and downstream mitigation or recovery [14], [28]–[30], [46], [58]. Together, these limitations indicate that the field must move beyond accuracy-focused detection toward lifecycle-integrated and operationally grounded anomaly management.

### *E. RQ5: WHAT ARE THE DOMINANT RESEARCH TRENDS OVER TIME AND WHICH GAPS REMAIN?*

Table 5 summarizes the historical evolution of UAV anomaly research and the corresponding gaps that remain across each period.

**2015–2018 (Early Stage).** Early UAV anomaly research was dominated by rule-based methods, thresholding, and analytical redundancy frameworks, reflecting both the limited availability of large real-world UAV telemetry datasets and the need for interpretable safety monitoring in relatively simple deployment settings [25], [26]. During this phase, anomaly detection was typically formulated as fault identification in individual sensors or control variables, and evaluation relied heavily on synthetic or laboratory-generated data. As a result, the literature established initial baselines for detection, but offered little support for broader lifecycle concerns such as anomaly prevention, recovery, or post-incident analysis.

**2019–2021 (Expansion).** As UAV telemetry collection became more feasible and machine learning matured, the field expanded toward supervised, unsupervised, and autoencoder-based detectors [39], [45]. This period marked a shift from manually specified fault rules toward data-driven modelling of normal and anomalous flight behaviour. Problem formulations also broadened from single-variable fault monitoring to multivariate temporal pattern recognition, while evaluation increasingly emphasized classification metrics such as accuracy, precision, recall, and F1-score. However, benchmark diversity remained limited, with many studies still relying on private or platform-specific datasets, making cross-study comparison difficult.

**2022–Present (Deep, Hybrid, and Foundation-Model Era).** More recent work reflects the broader rise of deep learning, graph-based modelling, self-supervised learning, and hybrid architectures in cyber-physical anomaly detection [46], [53], [58]. At the same time, the growth of swarm UAV

applications, communication-aware threat models, and more complex multi-agent operating environments encouraged research on distributed, communication-centric, and security-oriented anomaly detection [6], [32], [33]. Compared with earlier phases, recent studies increasingly frame anomaly detection as representation learning over heterogeneous sensor, communication, or coordination signals, rather than simple fault thresholding. They also place greater emphasis on robustness, explainability, and lightweight deployment, particularly for edge execution on resource-constrained UAV platforms [62]. More recently, foundation models and large language models (LLMs) have begun to emerge as a complementary direction, especially for higher-level anomaly interpretation, incident summarization, operator support, and integration of heterogeneous textual and telemetry-derived evidence, rather than as replacements for signal-level detectors. Their broader role in UAV anomaly management is discussed in Section VII-H.

Recent studies increasingly emphasize lightweight and sensor-efficient anomaly detection models to meet onboard computational and energy constraints, particularly for edge deployment on resource-limited UAV platforms [62]. However, the evolution remains detection-heavy. Preventive modeling, recovery planning, and human-in-the-loop anomaly reasoning constitute under 25% of recent outputs. Notably, hybrid model complexity and energy cost are rising, with few efforts toward edge optimization or onboard deployment. Works such as [62] demonstrate memory-efficient architectures, but standard real-time benchmarks are still absent.

**Remaining Gaps.** The gaps reported in RQ5 are derived through cross-sectional synthesis of the reviewed studies rather than from a single dedicated coding field. Specifically, some gaps are directly supported by the quantitative distributions in RQ2 and RQ3 (e.g., detection-centricity, limited real-world validation, and benchmark fragmentation), while others emerged through thematic analysis of the limitations and future-work discussions reported across the reviewed papers. Despite recent progress, several important gaps remain in the UAV anomaly management literature. First, there is still limited unification of datasets, benchmarks, and evaluation metrics across the different pillars, which makes it difficult to compare methods fairly or understand how advances in one stage of anomaly management relate to another. Studies often evaluate generation, detection, prevention, or recovery using separate experimental setups, resulting in fragmented evidence rather than an integrated view of system performance. Second, many approaches are still insufficiently validated in real-world, noisy, and non-stationary operating environments. Models that perform well in simulation or controlled datasets may degrade when exposed to changing weather, sensor drift, communication instability, evolving mission conditions, or other sources of operational uncertainty. Third, there remains a lack of transparent and practical integration with flight controllers and onboard autonomy stacks. Many proposed methods are demonstrated as offline analytics or external monitoring tools, but fewer works show how anomaly-related decisions can be incorporated safely, explainably, and in real time within actual UAV control loops. Finally, the literature pays comparatively less attention to policy-aware anomaly governance, including how anomaly handling should align with safety rules, operational constraints, accountability requirements, and human oversight. As UAV deployment expands into safety-critical and regulated settings, these governance considerations are likely to become increasingly important alongside purely technical performance.

**TABLE 5. RQ5: Historical evolution of UAV anomaly research and remaining gaps (2015–2025).**

| Period | Main Trend | What Changed | Remaining Gap |
|---|---|---|---|
| 2015–2018 | Rule-based and analytical redundancy methods dominated. | Research focused on interpretable fault detection using synthetic or laboratory-generated data. | Limited attention to prevention, recovery, and post-incident analysis. |
| 2019–2021 | Machine-learning and autoencoder-based detectors expanded. | The field shifted toward data-driven modelling of multivariate flight behaviour and broader use of Accuracy, Precision, Recall, and F1-score. | Heavy reliance on private or platform-specific datasets limited comparability. |
| 2022–Present | Deep, hybrid, swarm-aware, and communication-centric models grew rapidly. | Attention, GNN, self-supervised, lightweight, and emerging LLM-assisted approaches increased emphasis on robustness, explainability, and edge deployment. | The literature remains detection-centric, with limited lifecycle integration, policy-aware frameworks, and standardized real-time benchmarks. |

**RQ5 Takeaway — Trends and Gaps.** Although UAV anomaly research has evolved from rule-based monitoring to machine learning, deep, hybrid, and increasingly lightweight or foundation-model-assisted approaches [19], [25], [45], [46], [58], [62], this evolution has not been matched by comparable progress in lifecycle integration, policy awareness, or real-world validation [6], [14], [30].

### F. LIMITATIONS

**Limitations.** Our findings are bounded by the search scope, indexing coverage of the selected digital libraries, and the terms used in the query strings; relevant studies may have been missed due to inconsistent naming or limited metadata. In addition, many included papers evaluate on proprietary logs or simulator-specific configurations, which constrains cross-study comparability and may bias reported per-

formance toward particular platforms. Finally, classification by lifecycle pillar is not always unambiguous. Many reviewed studies touch on multiple stages of anomaly management, even when one stage appears more central than the others. As a result, assigning each paper to a single primary pillar necessarily involves interpretive judgement and may be influenced by domain knowledge. Although we used consistent coding criteria and cross-checking to improve reliability, some studies could plausibly have been classified differently, especially where prevention, recovery, or analysis were discussed as secondary rather than primary contributions.

## VII. DISCUSSION: HOW DO DRONE-SPECIFIC CHALLENGES (REAL-TIME, SAFETY-CRITICAL, SIM-TO-REAL) INFLUENCE DESIGN OF ANOMALY MANAGEMENT SYSTEMS?

UAV anomaly research demonstrates strong methodological depth but insufficient lifecycle integration. The field has progressed from analytical redundancy toward deep and hybrid intelligence, yet suffers from fragmented datasets, weak real-world validation, and limited interpretability. Bridging these gaps will require standardized benchmarks, cross-domain data sharing, and safety-certifiable AI models that connect detection to prevention and recovery. UAV anomaly management is shaped by three intersecting constraints: real-time operation, safety-critical assurance, and sim-to-real transferability. This section synthesizes how these constraints drive design choices across the five pillars: generation, prevention, detection, recovery/mitigation, and analysis/learning, outlining concrete implications for future systems.

### A. REAL-TIME CONSTRAINTS AND ONBOARD COMPUTE BUDGETS

UAVs must detect and react to anomalies within stringent latency budgets while operating under tight computational and energy constraints [32], [62]. These demands discourage the use of heavy multi-stage inference pipelines or communication-bound cloud analytics. As a result, effective systems favor latency-aware architectures built around streaming or single-pass models [27], [54] such as lightweight RNNs, TCNs, or compact transformers. To maintain responsiveness, models are often pruned, quantized, or distilled so that they can execute entirely on the flight controller or an embedded companion computer. In scenarios where processing time exceeds real-time thresholds, watchdog mechanisms ensure the system reverts to simple rule-based thresholds instead of unsafe delays. Data fusion must also respect sensor-synchrony; causal filters are preferred to avoid look-ahead bias. Overall, UAV anomaly management must adopt an edge-first, link-optional philosophy where off-board computation can enhance situational awareness, but the drone's safety logic must remain fully autonomous even when communication links fail.

### B. SAFETY-CRITICAL ASSURANCE, CERTIFICATION, AND EXPLAINABILITY

In safety-critical flight domains, anomaly detectors must not only be accurate but also explainable, predictable, and certifiable [3]. Regulators and mission operators require models whose behavior can be audited and traced. Consequently, UAV research increasingly emphasizes hybrid assurance frameworks that combine formal runtime monitors based on temporal logic or invariants with data-driven detectors. The monitors act as safety supervisors that can block unsafe actuator commands. To support incident forensics and compliance, post-hoc interpretability tools such as feature-attribution visualizations or counterfactual traces are used to translate alerts into fault trees and causal explanations [46], [58]. Recovery controllers must also satisfy formal correctness properties, ensuring that fallback behaviors such as hover, loiter, or return-to-home respect safety envelopes even under uncertainty. Integrating policy constraints like geofences or altitude limits as hard boundaries further strengthens assurance, while multi-level mitigation strategies allow systems to degrade gracefully rather than fail catastrophically.

### C. THE SIM-TO-REAL GAP: DATA, FIDELITY, AND ROBUSTNESS

As real-world anomalies are rare and hazardous to reproduce, researchers often rely on simulated data or injected faults using platforms such as AirSim [18] and UAVSim [19]. Yet such data introduce a persistent sim-to-real gap, as simulation rarely captures the full stochasticity of wind gusts, vibration noise, sensor drift, or electromagnetic interference. To narrow this gap, anomaly management systems increasingly emphasize robust training strategies (e.g., realism-calibrated noise and latency models derived from flight logs) and validation on real flights, where empirical studies report measurable performance degradation when moving from offline evaluation to real-world deployments [14]. Beyond sensor-level realism, data augmentation through jitter, dropout, saturation, and event-level corruptions (e.g., GPS spoofing or ESC brownouts) helps increase robustness. Evaluation must extend beyond a single airframe or autopilot stack. Cross-platform validation on PX4, ArduPilot, and custom stacks is essential. Establishing open anomaly corpora and benchmarks such as ALFA [29], with standardized latency and energy metrics would greatly improve reproducibility and foster benchmarking across the community.

### D. LIFECYCLE INTEGRATION: FROM DETECTION TO MITIGATION AND LEARNING

Real-world resilience demands a feedback loop linking detection, mitigation, and post-flight learning [30]. An anomaly detector without an actionable recovery path offers little operational value. Therefore, modern frameworks connect alerts to risk-stratified playbooks that trigger corresponding behaviors such as hovering, reducing speed, or executing a controlled landing, based on the severity and time-to-react budget [65]. Controllers also exploit predictive uncertainty to

modulate aggressiveness, defaulting to conservative policies under low confidence [14]. After each flight, logged incidents can be replayed to refine detector thresholds and validate mitigation policies without endangering hardware [24]. Such forensic pipelines enable continual retraining while keeping certification traceability intact. To minimize operational risk, updated models can be evaluated in passive monitoring (shadow) modes before becoming active, supporting incremental reliability improvements over the UAV's lifecycle.

### E. PRACTICAL DEPLOYMENT CONSIDERATIONS

Effective anomaly management also depends on robust engineering practices. UAV operators need health-monitoring dashboards that consolidate estimator residuals, CPU utilization, and sensor integrity while avoiding operator overload through adaptive alert prioritization. Energy efficiency must be an explicit design objective; compute-intensive inference should scale with remaining battery, with clear power-per-inference metrics guiding runtime adaptation. Safety-critical updates require reproducible builds, signed binaries, and rollback mechanisms to ensure that firmware or model regressions can be reversed immediately after anomalous behavior [3]. Finally, incorporating human-in-the-loop supervision closes the operational trust gap: explainable alert summaries, manual override interfaces, and intuitive visualization of system confidence help pilots make informed corrective actions when autonomy falters [46], [58]. In addition to algorithmic accuracy, practical UAV anomaly management depends on whether proposed methods can be integrated safely and efficiently into real flight systems. In operational settings, anomaly-related decisions must interface with flight controllers, onboard state-estimation modules, supervisory safety logic, and mission-management components, rather than functioning as isolated analytics. This creates requirements that are less visible in offline evaluation, including low-latency inference, bounded computational and energy overhead, compatibility with embedded hardware, and predictable behavior under noisy, non-stationary conditions such as vibration, communication instability, weather variation, and sensor drift. Moreover, practical usefulness depends not only on detecting an anomaly, but also on enabling appropriate and auditable responses, such as degraded-mode flight, hovering, return-to-home, or controlled landing, without destabilizing the control loop. These considerations suggest that future UAV anomaly-management research should place greater emphasis on end-to-end deployment realism, including controller integration, resource-aware design, and validation under representative operational conditions.

### F. GOVERNANCE, PRIVACY, AND REGULATORY COMPLIANCE

Although UAV anomaly management is often framed primarily as a technical detection problem, practical deployment also raises important governance concerns [66]. Flight logs, telemetry streams, and mission traces may contain sensitive operational information, particularly in commercial, infrastructure, or security-related settings. These constraints can limit data sharing, reduce the availability of public benchmarks, and complicate reproducible evaluation. Accordingly, future UAV anomaly-management frameworks should be designed not only for detection effectiveness, but also for privacy, access control, and regulatory compliance, especially when deployed in safety-critical or policy-constrained environments.

### G. IMPLICATIONS AND RESEARCH DIRECTIONS

Beyond synthesizing existing approaches, this review highlights several concrete directions for future research, particularly for research on UAV anomaly management.

**First, research must move beyond detection-only formulations.** While anomaly detection dominates the literature, relatively little work addresses anomaly generation, prevention, or recovery in an integrated manner. Researchers should pursue end-to-end anomaly management pipelines in which detection outputs are explicitly linked to mitigation strategies and post-incident learning. This includes designing datasets and benchmarks that evaluate not only detection accuracy, but also recovery effectiveness, time-to-mitigation, and policy compliance.

**Second, there is a need for policy-aware and context-aware anomaly models.** Most existing detectors treat anomalies as abstract statistical deviations, without encoding operational rules such as geofences, altitude limits, or mission intent. Future work should investigate models that jointly reason over telemetry data and explicit policy constraints, enabling early detection of safety or regulatory violations before they manifest as physical failures.

**Third, sim-to-real robustness remains a fundamental open problem.** Researchers should prioritize cross-platform validation across different airframes, autopilot stacks, and environmental conditions. This includes developing physics-informed models, domain adaptation techniques, and standardized fault-injection protocols that better reflect real-world disturbances. Public datasets annotated with environmental, operational, and policy metadata would significantly improve reproducibility and comparability.

**Fourth, explainability and certification must be treated as first-class objectives.** Deep and hybrid models should be evaluated not only on predictive performance, but also on their interpretability, stability, and suitability for certification in safety-critical settings. Research into causal models, runtime monitors, and hybrid formal-learning frameworks is especially promising for bridging the gap between learning-based detection and regulatory assurance.

**Finally, lifecycle-oriented evaluation metrics are needed.** Current studies overwhelmingly report accuracy-based metrics yet overlook latency, energy consumption, and operational impact. Future research should adopt evaluation protocols that reflect real deployment constraints, including onboard inference budgets, battery-aware adaptation, and long-term system evolution through continual learning.

Taken together, these directions suggest a shift from algorithm-centric anomaly detection toward system-level, policy-aware, and lifecycle-integrated anomaly management for UAVs.

### H. FORWARD-LOOKING PERSPECTIVE ON EMERGING TECHNOLOGIES

Although the reviewed corpus remains dominated by conventional anomaly-detection pipelines, several emerging technologies may influence future UAV anomaly-management design. These include edge intelligence, advanced wireless connectivity, digital twins and high-fidelity simulation, federated and distributed learning, and foundation-model-based reasoning. Because these technologies are still only weakly represented in the reviewed UAV anomaly-management literature, this subsection is intended as a forward-looking perspective rather than a direct synthesis result. Nevertheless, recent developments in UAVs, robotics, cyber-physical systems, and log analysis suggest that such technologies may help address several persistent gaps identified in this review, including limited onboard deployment, weak lifecycle integration, fragmented evaluation, and the need for stronger post-incident interpretation and human oversight. In particular, recent advances in LLMs introduce new opportunities and challenges across the lifecycle of generation, prevention, detection, recovery, and analysis. While most existing UAV anomaly research focuses on numerical telemetry and sensor streams, LLMs offer complementary capabilities for reasoning over heterogeneous, symbolic, and unstructured information, including logs, policies, procedures, and human-machine interactions. As such, LLMs should be viewed not as replacements for traditional anomaly detectors, but as orchestration and reasoning components that augment existing systems [47], [48], [67].

Emerging technologies may support multiple stages of the anomaly-management lifecycle. In generation and testing, digital-twin-style simulation and foundation-model-based methods may enable richer scenario synthesis, adaptive fault injection, and more realistic virtual testing environments. LLM-guided approaches could assist in generating fault-injection scripts, fuzzing strategies, or edge-case mission scenarios from structured descriptions of UAV architectures, mission profiles, or regulatory constraints [68]. In prevention, edge intelligence and distributed monitoring may support earlier identification of degradation on resource-constrained platforms, while advanced connectivity and fleet-level learning may improve preventive maintenance and cross-platform knowledge sharing. LLMs may additionally support pre-flight analysis by identifying unsafe parameter choices, inconsistent mission constraints, or latent policy conflicts from configuration files, mission plans, and natural-language requirements [47], [49].

For detection, lightweight and hardware-aware models are likely to be especially important because many existing anomaly-detection methods remain too computationally demanding for strict onboard deployment. Advanced wireless connectivity may also support distributed anomaly reasoning in swarm or beyond-visual-line-of-sight settings. LLMs, however, are less suitable for low-latency signal-level anomaly detection from high-frequency telemetry because of timing and computational constraints. Their more natural role is at a higher level of abstraction, for example in identifying inconsistencies across logs, detecting abnormal command sequences, correlating alerts from multiple subsystems, or reasoning about coordination anomalies in multi-UAV settings [50]. Post-flight analysis may be even more promising: digital twins can support richer replay and diagnosis, federated or distributed learning may enable knowledge transfer across fleets, and LLMs may assist in summarizing incidents, extracting root-cause hypotheses, and linking low-level deviations to higher-level mission or policy violations [50].

Despite these opportunities, significant challenges remain. Edge models may trade representational power for resource efficiency; advanced connectivity can introduce new latency, reliability, and cybersecurity dependencies; digital twins require sufficient fidelity and synchronization with operational conditions; and federated learning must cope with heterogeneous, non-IID, and potentially unreliable participants. LLMs in particular lack deterministic guarantees, are sensitive to prompt formulation, and may hallucinate plausible but incorrect explanations [69], including outputs that conflict with safety-critical UAV requirements. Their use therefore demands strict architectural boundaries, with LLMs operating in advisory, offline, or supervisory roles rather than direct control loops. Overall, the most promising direction is likely to be hybrid UAV anomaly-management architectures that combine lightweight onboard intelligence, improved connectivity, richer simulation and replay, distributed learning, and language-based reasoning with runtime monitors, formal verification, and physics-based models, ensuring that new flexibility is balanced by the rigor required for safe and compliant UAV operations [70].

## VIII. RELATED WORK

### A. POSITIONING RELATIVE TO EXISTING SURVEYS

Several existing surveys provide useful background on UAVs, cyber-physical systems, and related anomaly or security issues, but their scope differs from the focus of this review. For example, surveys such as [10] and [12] provide broad overviews of deep learning and machine learning applications for UAVs, covering tasks such as perception, navigation, tracking, and mission planning, but they do not specifically focus on anomaly management as a lifecycle problem.

Other surveys are more closely related to security and fault diagnosis. For instance, [4] and [11] review UAV vulnerabilities, communication security, and privacy threats, while [13] surveys anomaly detection in cyber-physical systems more broadly. These surveys help contextualize UAV anomaly detection within larger safety and security concerns, but they typically do not analyse how anomaly generation, prevention, detection, recovery, and post-incident analysis connect across

the UAV lifecycle. Likewise, they give limited attention to how technical anomalies relate to operational policies, regulatory constraints, and deployment assumptions.

In contrast, this survey is explicitly lifecycle-oriented. Rather than treating anomaly detection as an isolated task, we organize the literature using a five-pillar taxonomy spanning anomaly generation, prevention, detection, recovery/mitigation, and analysis/learning. We further compare studies across anomaly source, technique family, data source, and application context, while also examining datasets, simulators, evaluation practices, and policy-aware considerations. This broader framing allows us to identify not only algorithmic progress, but also structural gaps in lifecycle integration, benchmark realism, and regulatory alignment that are less visible in existing detection- or security-focused surveys.

### B. ANOMALY DETECTION IN CYBER-PHYSICAL AND SWARM SYSTEMS

Beyond single-drone studies, several researchers explored anomaly detection in multi-agent or swarm UAV contexts. Works such as [33] and [58] extended deep models toward swarm coordination anomalies, while [45] proposed distributed LSTM-AE architectures that decentralize computation across agents. These align with broader CPS research emphasizing spatio-temporal modeling, self-supervised learning, and energy-aware anomaly detection. For example, [62] demonstrated that memory-augmented architectures with few-sensors can deliver accuracy comparable to full-sensor deep models, a critical insight for resource-constrained onboard computing. Collectively, such works advance scalability and autonomy but still rely heavily on simulation or curated data, underscoring the need for real-flight benchmark repositories that reflect operational noise, communication delays, and human-pilot variability.

### C. CROSS-DOMAIN INSIGHTS AND POLICY-AWARE PERSPECTIVES

Outside UAV-specific literature, methodologies from cyber-physical systems (CPS), IoT, and safety-critical robotics contribute valuable cross-domain insights. Structured sparse learning [25] and wavelet-SDAE methods [35] mirror advances in industrial CPS anomaly detection, emphasizing denoising, residual modeling, and adaptive thresholds. In parallel, policy-aware detection is gaining traction: studies such as [14] and [6] connect UAV anomalies to regulatory compliance and cybersecurity, showing how delayed or inaccurate detection can trigger operational violations. Yet, lifecycle-oriented taxonomies that integrate technical, operational, and regulatory perspectives remain scarce. This survey bridges technical detection methods with the broader ecosystem of anomaly generation, prevention, recovery, and policy enforcement, establishing a foundation for benchmarked, interpretable, and policy-aware UAV safety frameworks.

## IX. CONCLUSION

This systematic literature review analysed 111 studies on UAV anomaly management through a lifecycle-oriented taxonomy encompassing anomaly generation, prevention, detection, recovery/mitigation, and analysis/learning. By synthesising anomaly sources, detection techniques, datasets, and application contexts, the review provides a unified perspective that extends beyond detection-centric surveys and highlights how research assumptions influence real-world deployment feasibility and safety assurance.

Two structural imbalances emerge from the literature. First, research is overwhelmingly concentrated on anomaly detection, while generation, prevention, and particularly recovery remain comparatively underexplored. Second, most work targets telemetry-visible hardware and software faults, whereas operational and policy-related anomalies receive limited attention despite their clear safety and regulatory implications. Together, these patterns indicate that UAV anomaly research remains largely detection-centric and insufficiently integrated across the broader anomaly management lifecycle.

Addressing these gaps will require a shift toward lifecycle-integrated anomaly management. Key priorities include: (i) benchmark datasets and evaluation protocols reflecting realistic operating conditions and policy constraints; (ii) methods that improve sim-to-real robustness under environmental and communication disturbances; (iii) policy-aware anomaly reasoning linking telemetry deviations to mission-level violations; and (iv) adaptive recovery strategies that connect detection to mitigation and post-incident learning. By consolidating existing work and exposing these gaps, this survey provides a foundation for the development of safer, more transparent, and policy-aware UAV anomaly management systems. The proposed five-pillar taxonomy is useful not only for organizing prior literature, but also for guiding future UAV anomaly-management design and evaluation. By distinguishing generation, prevention, detection, recovery, and analysis, it helps reveal whether a system provides only anomaly recognition or supports broader lifecycle resilience. This makes the taxonomy a useful scaffold for identifying capability gaps and for encouraging more integrated, deployment-oriented UAV anomaly-management frameworks.

## REFERENCES

[1] I. Tan, "UAV anomaly papers," https://ivantanweihan.github.io/uav-anomaly-papers/, 2026.

[2] K. Kanistras, G. Martins, M. J. Rutherford, and K. P. Valavanis, "A survey of unmanned aerial vehicles (UAVs) for traffic monitoring," in *2013 international conference on unmanned aircraft systems (ICUAS)*. IEEE, 2013, pp. 221–234.

[3] D. Wanner, H. A. Hashim, S. Srivastava, and A. Steinhauer, "UAV avionics safety, certification, accidents, redundancy, integrity, and reliability: a comprehensive review and future trends," *Drone Systems and Applications*, vol. 12, pp. 1–23, 2024.

[4] A. Shafique, A. Mehmood, and M. Elhadef, "Survey of security protocols and vulnerabilities in unmanned aerial vehicles," *IEEE Access*, vol. 9, pp. 46 927–46 948, 2021.

[5] K. Dorling, J. Heinrichs, G. G. Messier, and S. Magierowski, "Vehicle routing problems for drone delivery," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 47, no. 1, pp. 70–85, 2016.

[6] M. Y. Alzahrani, "Enhancing drone security through multi-sensor anomaly detection and machine learning," *SN Computer Science*, vol. 5, no. 5, p. 651, 2024.

[7] The Straits Times, "4 men to be charged for offences related to operating unmanned aircraft in protected areas," Online, Jul. 2024, accessed: 2026-02-25. [Online]. Available: https://www.straitstimes.com/singapore/four-men-to-be-charged-for-offences-related-to-operating-unmanned-aircraft-in-protected-areas

[8] M. Bryant, "Danish pm: airport drone incursion a 'serious attack' on critical infrastructure," The Guardian (Online), Sep. 2025, accessed: 2026-02-25. [Online]. Available: https://www.theguardian.com/world/2025/sep/23/drone-sightings-cause-disruption-delays-norway-denmark-airports

[9] United States Department of Justice, U.S. Attorney's Office, Northern District of California, "Drone operator charged with violating temporary flight restrictions in area surrounding levi's stadium during nfl game," Press release (Online), Feb. 2026, accessed: 2026-02-25. [Online]. Available: https://www.justice.gov/usao-ndca/pr/drone-operator-charged-violating-temporary-flight-restrictions-area-surrounding-levis

[10] A. Carrio, C. Sampedro, A. Rodriguez-Ramos, and P. Campoy, "A review of deep learning methods and applications for unmanned aerial vehicles," *Journal of Sensors*, vol. 2017, no. 1, p. 3296874, 2017.

[11] Y. Zhi, Z. Fu, X. Sun, and J. Yu, "Security and privacy issues of UAV: A survey," *Mobile Networks and Applications*, vol. 25, no. 1, pp. 95–101, 2020.

[12] K. Teixeira, G. Miguel, H. S. Silva, and F. Madeiro, "A survey on applications of unmanned aerial vehicles using machine learning," *IEEE Access*, vol. 11, pp. 117 582–117 621, 2023.

[13] D. Abshari and M. Sridhar, "Cyber-physical systems security: A comprehensive review of anomaly detection techniques," *Internet Things*, vol. 37, p. 101938, 2026.

[14] P. Pathe, A.-L. Jousselme, B. Pannetier, and O. Bartheye, "Real-world validation of drone anomaly detection using evidential networks," in *2025 28th International Conference on Information Fusion (FUSION)*. IEEE, 2025, pp. 1–8.

[15] F. Ahmed and M. Jenihhin, "A survey on UAV computing platforms: A hardware reliability perspective," *Sensors*, vol. 22, no. 16, p. 6286, 2022.

[16] F. Furrer, M. Burri, M. Achtelik, and R. Siegwart, "RotorS—a modular gazebo MAV simulator framework," in *Robot Operating System (ROS) The Complete Reference (Volume 1)*. Springer, 2016, pp. 595–625.

[17] Open Robotics, "Gazebo: A Robot Simulation Framework," https://gazebosim.org, 2025.

[18] S. Shah, D. Dey, C. Lovett, and A. Kapoor, "AirSim: High-fidelity visual and physical simulation for autonomous vehicles," in *Field and service robotics: Results of the 11th international conference*. Springer, 2017, pp. 621–635.

[19] A. Y. Javaid, W. Sun, and M. Alam, "UAVSim: A simulation testbed for unmanned aerial vehicle network cyber security analysis," in *GLOBECOM Workshops*. IEEE, 2013, pp. 1432–1436.

[20] B. Kitchenham, S. Charters *et al.*, "Guidelines for performing systematic literature reviews in software engineering," *EBSE Technical Report EBSE-2007-01*, 2007.

[21] M. J. Page, J. E. McKenzie, P. M. Bossuyt, I. Boutron, T. C. Hoffmann, C. D. Mulrow, L. Shamseer, J. M. Tetzlaff, E. A. Akl, S. E. Brennan *et al.*, "The PRISMA 2020 statement: an updated guideline for reporting systematic reviews," *BMJ*, vol. 372, 2021.

[22] B. Dowdeswell, R. Sinha, and S. G. MacDonell, "Finding faults: A scoping study of fault diagnostics for industrial cyber-physical systems," *J. Syst. Softw.*, vol. 168, p. 110638, 2020.

[23] F. Pourpanah, B. Zhang, R. Ma, and Q. Hao, "Anomaly detection and condition monitoring of UAV motors and propellers," in *2018 IEEE SENSORS*. IEEE, 2018, pp. 1–4.

[24] L. K. Shar, W. Minn, T. N. B. Duong, J. Fan, L. Jiang, and D. L. W. Kiat, "DronLomaly: Runtime detection of anomalous drone behaviors via log analysis and deep learning," in *APSEC*. IEEE, 2022, pp. 119–128.

[25] Y. He, Y. Peng, S. Wang, D. Liu, and P. H. Leong, "A structured sparse subspace learning algorithm for anomaly detection in UAV flight data," *IEEE Transactions on Instrumentation and Measurement*, vol. 67, no. 1, pp. 90–100, 2018.

[26] C. Boucetta, B. Nour, S. E. Hammami, H. Moungla, and H. Afifi, "Adaptive range-based anomaly detection in drone-assisted cellular networks," in *2019 15th International Wireless Communications & Mobile Computing Conference (IWCMC)*. IEEE, 2019, pp. 1239–1244.

[27] Y. Liu and W. Ding, "A KNNS based anomaly detection method applied for UAV flight data stream," in *2015 Prognostics and System Health Management Conference (PHM)*. IEEE, 2015, pp. 1–8.

[28] T. M. Tran, T. N. Vu, T. V. Nguyen, and K. Nguyen, "UIT-ADrone: A novel drone dataset for traffic anomaly detection," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 16, pp. 5590–5601, 2023.

[29] A. Keipour, M. Mousaei, and S. A. Scherer, "ALFA: A dataset for UAV fault and anomaly detection," *Int. J. Robotics Res.*, vol. 40, no. 2-3, 2021.

[30] V. K. Malviya, W. Minn, L. K. Shar, and L. Jiang, "Fuzzing drones for anomaly detection: A systematic literature review," *Computers & Security*, vol. 148, p. 104157, 2025.

[31] V. Sindhwani, H. Sidahmed, K. Choromanski, and B. Jones, "Unsupervised anomaly detection for self-flying delivery drones," in *ICRA*. IEEE, 2020, pp. 186–192.

[32] J. Luo and A. Oracevic, "Adaptive machine learning for efficient anomaly detection in autonomous UAVs swarm operations," in *ISNCC*. IEEE, 2024, pp. 1–8.

[33] H. Ahn and S. Chung, "Deep learning-based anomaly detection for individual drone vehicles performing swarm missions," *Expert Systems with Applications*, vol. 244, p. 122869, 2024.

[34] M. H. M. Ghazali and W. Rahiman, "An investigation of the reliability of different types of sensors in the real-time vibration-based anomaly inspection in drone," *Sensors*, vol. 22, no. 16, p. 6015, 2022.

[35] R. Zhou, L. Chen, and X. Wang, "An anomaly detection method for UAV based on wavelet decomposition and stacked denoising autoencoder," *Measurement*, vol. 227, p. 114456, 2024.

[36] W. Minn, Y. N. Tun, L. K. Shar, and L. Jiang, "DronLomaly: Runtime log-based anomaly detector for DJI drones," in *ICSE Companion*. ACM, 2024, pp. 6–10.

[37] S. Shaw, K. Joshi, A. Pathak, A. K. Thyagarajan, G. Vidya, R. Hemal Shah, V. Ram Kishan, and J. S. R. Alex, "Anomaly detection in drones with machine learning algorithms," in *International Conference on Futuristic Communication and Network Technologies*. Springer, 2020, pp. 433–441.

[38] H. Chen, Y. Lyu, J. Shi, and W. Zhang, "UAV anomaly detection method based on convolutional autoencoder and support vector data description with 0/1 soft-margin loss," *Drones*, vol. 8, no. 10, p. 534, 2024.

[39] B. Wang, D. Liu, X. Peng, and Z. Wang, "Data-driven anomaly detection of UAV based on multimodal regression model," in *2019 IEEE International Instrumentation and Measurement Technology Conference (I2MTC)*. IEEE, 2019, pp. 1–6.

[40] A. Fakhry, J. Lee, and J. T. Lee, "Drone video anomaly detection by future segmentation prediction and spatio-temporal relational modeling," *IEEE Access*, vol. 13, pp. 22 395–22 406, 2025.

[41] P. Jin, L. Mou, G.-S. Xia, and X. X. Zhu, "Anomaly detection in aerial videos with transformers," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1–13, 2022.

[42] T. E. Nesthus, K. A. Fercho, J. D. Durham, T. C. Mofle, B. L. Nesmith, and P. Hu, "Summary final report for unmanned aircraft systems in air carrier operations: Uas operator fatigue," Federal Aviation Administration, Office of Aerospace Medicine, Washington, DC, Technical Report DOT/FAA/AM-21/16, May 2021.

[43] H. Ahn, H.-L. Choi, M. Kang, and S. Moon, "Learning-based anomaly detection and monitoring for swarm drone flights," *Applied Sciences*, vol. 9, no. 24, p. 5477, 2019.

[44] S. Kim, G. Anagnostopoulos, E. Barmpounakis, and N. Geroliminis, "Visual extensions and anomaly detection in the pNEUMA experiment with a swarm of drones," *Transportation Research Part C: Emerging Technologies*, vol. 147, p. 103966, 2023.

[45] G. Bae and I. Joe, "UAV anomaly detection with distributed artificial intelligence based on LSTM-AE and AE," in *International Conference on Multimedia and Ubiquitous Engineering*. Springer, 2019, pp. 305–310.

[46] C. Feng, J. Fan, Z. Liu, G. Jin, and S. Chen, "Unmanned aerial vehicle anomaly detection based on causality-enhanced graph neural networks," *Drones*, vol. 9, no. 6, p. 408, 2025.

[47] W. Xu, M. Liu, O. Sokolsky, I. Lee, and F. Kong, "LLM-enabled cyber-physical systems: Survey, research opportunities, and challenges," in *2024 IEEE International Workshop on Foundation Models for Cyber-Physical Systems & Internet of Things (FMSys)*. IEEE, 2024, pp. 50–55.

[48] S. Javaid, H. Fahim, B. He, and N. Saeed, "Large language models for UAVs: Current state and pathways to the future," *IEEE Open Journal of Vehicular Technology*, vol. 5, pp. 1166–1192, 2024.

[49] A. Sezgin, "Scenario-driven evaluation of autonomous agents: Integrating large language model for UAV mission reliability," *Drones*, vol. 9, no. 3, p. 213, 2025.
[50] S. Akhtar, S. Khan, and S. Parkinson, "LLM-based event log analysis techniques: A survey," *CoRR*, vol. abs/2502.00677, 2025.
[51] S.-H. Seo and H. Jung, "Anomaly detection method for drone navigation system based on deep neural network," *Journal of Positioning, Navigation, and Timing*, vol. 11, no. 2, pp. 109–117, 2022.
[52] J. Henrio and T. Nakashima, "Anomaly detection in videos recorded by drones in a surveillance context," in *2018 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*. IEEE, 2018, pp. 2503–2508.
[53] J. Akram, A. Anaissi, W. Othman, A. Alabdulatif, and A. Akram, "DroneSSL: Self-supervised multimodal anomaly detection in internet of drone things," *IEEE Transactions on Consumer Electronics*, vol. 70, no. 1, pp. 4287–4298, 2024.
[54] C. Titouna, F. Naït-Abdesselam, and H. Moungla, "An online anomaly detection approach for unmanned aerial vehicles," in *2020 International Wireless Communications and Mobile Computing (IWCMC)*. IEEE, 2020, pp. 469–474.
[55] J. Wang, X. Ma, and L. Yang, "Hybrid maintenance optimization for multi-state competing failure systems under inspection uncertainties and spare dynamics," *Reliability Engineering & System Safety*, p. 112524, 2026.
[56] J. Wang, Y. Zhao, X. Ma, H. Xiao, Y. Ma, R. Peng, and L. Yang, "A state-age-dependent maintenance-spare control strategy under inspection error compensation," *IEEE Transactions on Reliability*, 2025.
[57] D. Wu, K. Gao, R. Peng, J. Sun, L. Yang, and F. Wei, "Optimal preventive maintenance policy considering technical-innovation-empowered renewals," *Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability*, p. 1748006X251412931, 2025.
[58] J. Yang, D. Tang, J. Yu, J. Zhang, and H. Liu, "Explaining anomalous events in flight data of UAV with deep attention-based multi-instance learning," *IEEE Trans. Veh. Technol.*, vol. 73, no. 1, pp. 107–119, 2024.
[59] L. Watkins, D. Hamilton, K. Kornegay, and A. Rubin, "Triaging autonomous drone faults by simultaneously assuring autonomy and security," in *2021 55th Annual Conference on Information Sciences and Systems (CISS)*. IEEE, 2021, pp. 1–6.
[60] N. Koenig and A. Howard, "Design and use paradigms for gazebo, an open-source multi-robot simulator," in *Proceedings of the 2004 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Sendai, Japan, Sep. 2004, pp. 2149–2154.
[61] J. S. Berndt, "JSBSim: An open source flight dynamics model in C++," in *AIAA Modeling and Simulation Technologies Conference and Exhibit*, 2004.
[62] J. Do Yoo, G. M. Kim, M. G. Song, and H. K. Kim, "MeNU: Memorizing normality for UAV anomaly detection with a few sensor values," *Computers & Security*, vol. 150, p. 104248, 2025.
[63] ASTM International, "ASTM F3269-17: Standard practice for methods to safely bound flight behavior of unmanned aircraft systems containing complex functions," 2017, accessed: 2026-04-05.
[64] European Union Aviation Safety Agency, "U-space: Regulatory framework for the u-space," https://www.easa.europa.eu/en/regulations/u-space, 2024, commission Implementing Regulation (EU) 2021/664 and Easy Access Rules for U-space. Accessed: 2026-04-05.
[65] A. Mallesh, "Autonomous security response architecture for flight path anomaly detection in defense drone systems," *Journal of Computer Science and Technology Studies*, vol. 7, no. 8, pp. 652–662, 2025.
[66] T.-H. Tran and D.-D. Nguyen, "Management and regulation of drone operation in urban environment: A case study," *Social Sciences*, vol. 11, no. 10, p. 474, 2022.
[67] F. Zeng, W. Gan, Y. Wang, N. Liu, and P. S. Yu, "Large language models for robotics: A survey," *CoRR*, vol. abs/2311.07226, 2023.
[68] M. De la Cruz Cabello, T. P. Sales, and M. R. Machado, "AIOps for log anomaly detection in the era of LLMs: A systematic literature review," *Intelligent Systems with Applications*, p. 200608, 2025.
[69] Y. Yao, J. Duan, K. Xu, Y. Cai, Z. Sun, and Y. Zhang, "A survey on large language model (LLM) security and privacy: The good, the bad, and the ugly," *High Confid. Comput.*, vol. 4, no. 2, p. 100211, 2024.
[70] W. Xu, M. Liu, O. Sokolsky, I. Lee, and F. Kong, "LLM-enabled cyber-physical systems: Survey, research opportunities, and challenges." [Online]. Available: https://par.nsf.gov/biblio/10499418

**IVAN TAN** is currently pursuing a Ph.D. degree in Computer Science with Singapore Management University, Singapore. He received the Bachelor of Science (Information Systems) from Singapore Management University, Singapore, in 2023. His research interests include unmanned aerial vehicle anomaly detection, software testing, and cyber-physical systems.

**CHRISTOPHER M. POSKITT** is an Associate Professor of Computer Science (Education) at Singapore Management University (SMU), where he is part of the Centre for Research on Intelligent Software Engineering. Prior to SMU, he held postdoctoral research positions at ETH Zürich and SUTD, and obtained his PhD in Computer Science from the University of York (2014). His research broadly addresses the problem of engineering correct and secure software, especially in the context of cyber-physical systems (e.g. industrial control systems, autonomous vehicles). In addition to software engineering, his research interests span formal methods, cybersecurity, and computer science education.

**LINGXIAO JIANG** received his Ph.D. degree in Computer Science from the University of California, Davis, USA, in 2009, a Master's and a Bachelor's degree from the School of Mathematical Sciences at Peking University, China, in 2003. He is currently a Professor of Computer Science in the School of Computing and Information Systems (SCIS) at Singapore Management University (SMU) and a Co-Director of the Centre for Research on Intelligent Software Engineering (RISE) at SMU. He explores combinations of software engineering, systems, cybersecurity, program analysis, and deep learning techniques across languages at various abstraction levels, aiming to provide practical techniques and tools to increase developer productivity, enhance software reliability, and reduce development & maintenance cost.

**LWIN KHIN SHAR** obtained his PhD degree from Nanyang Technological University (NTU), Singapore. He was a postdoctoral research associate at SnT of the University of Luxembourg and then a research scientist at NTU. He is currently an Associate Professor in Computer Science (Practice), School of Computing & Information Systems, Singapore Management University. His research mainly focuses on security testing and analysis of web applications, mobile applications, and recently cyber-physical systems. Particularly, he focuses on analysing and detecting security vulnerabilities, privacy issues, malware, and anomalies. In his research, he often collaborates with industry partners spanning from healthcare to Government sectors. His recent awards include the runners-up of IEEE Best Paper Awards (2021) and the Best Paper Award nomination in ITiCSE (2022). He is also a member of IEEE and ACM.