Risk Assessment for ML-Based Applications in Satellite Systems

Simon Shigol[*], Roy Peled[*], Avishag Shapira, Yuval Elovici, Asaf Shabtai

Department of Software and Information Systems Engineering, Ben-Gurion University of the Negev, Israel
[*]These authors contributed equally.

***Abstract*—Machine learning (ML) is increasingly embedded in satellite systems, supporting both operational tasks and payload services. While ML provides greater efficiency and autonomy, it also exposes satellite systems to a new class of vulnerabilities known as adversarial ML (AML). Although AML threats have been studied extensively in other domains, their impact on satellite systems, which operate with limited power and computing resources and under latency-critical conditions, remains unexplored. This paper presents a structured risk assessment of AML threats to satellite ML applications. We review common types of cyber threats and AML techniques, providing clear definitions of AML categories and their relevance to satellite ML applications. We then map these threats to satellite operations and payloads, constructing a domain-specific framework that categorizes how adversarial attacks manifest under space conditions. Leveraging this framework, we apply a risk assessment methodology to evaluate the feasibility of attacks and their potential impact on missions. Our findings show that tasks such as anti-jamming control and telemetry-based fault detection are especially vulnerable, with integrity-focused attacks posing the most significant risk to the evaluated applications. In contrast, privacy-focused threats such as membership inference pose less risk in practice. We also suggest mitigation strategies tailored to space, including adversarial training, resilient data pipelines, and runtime monitoring. The results of our risk assessment highlight the need for further research aimed at strengthening ML security in aerospace environments and provide a foundation for the deployment of trustworthy ML in space missions.**

I. Introduction

Satellites play a fundamental role in modern infrastructure, supporting global communications, navigation, Earth observation, and defense [1]. As missions expand in scale and complexity, operators increasingly employ machine learning (ML) to enhance autonomy and efficiency in tasks such as beam hopping, interference mitigation, telemetry-based health monitoring, and mission planning [2], [3]. These capabilities improve operational resilience despite limited bandwidth, long communication delays, and dynamic orbital environments.

However, adversaries have repeatedly disrupted satellite services through jamming, spoofing, and cyberattacks [4]. Integrating ML into these systems introduces a new attack surface-adversarial machine learning (AML) [5]. While AML has been studied extensively on Earth, its implications for space missions, where computing power, energy, and reliability are tightly constrained, remain largely unexamined.

Satellites in low Earth orbit (LEO), medium Earth orbit (MEO), and geostationary orbit (GEO) support different mission profiles and ML applications, including communications, Earth observation, and space situational awareness. A detailed mapping of ML relevance across orbital regimes is provided in the Appendix. Yet, deploying ML models in orbit remains challenging due to strict size, weight, and power (SWaP) constraints, which limit onboard processing and increase reliance on ground-based analysis. Space agencies such as NASA and ESA are pursuing greater autonomy while expressing concern over the safety, certification, and explainability of ML components [6], [7]. Unlike deterministic flight software, ML models rely on data-driven logic with limited interpretability, complicating validation and assurance [8]–[10]. These characteristics create new risks in mission-critical environments where reliability and auditability are essential [11]–[13].

Recent studies demonstrate ML's potential for adaptive anti-jamming [14], object detection [15], and health forecasting [16]. Yet most evaluations focus on nominal performance rather than adversarial resilience.

Problem. The robustness of space ML systems under AML threats remains unexamined mainly, and no comprehensive framework exists for assessing the feasibility, access requirements, and mission impact of such attacks across diverse satellite use cases [8].

Our approach. We conduct a domain-grounded risk assessment by reviewing literature and operational reports to identify ML use cases across onboard and ground segments. Adversarial threat models are mapped to these applications using established space-cybersecurity frameworks and AML research. The mapping process evaluates the feasibility, access requirements, and potential mission impact of each threat, resulting in a relevance matrix for satellite ML risk assessment. Based on this mapping, we adapt the FRAME methodology [5] for space, incorporating constraints such as limited compute, intermittent connectivity, and certification barriers to enable realistic, domain-specific risk scoring.

Contributions.

- **Threat mapping for satellite ML applications:** A comprehensive mapping of AML threats across satellite

Workshop on Security of Space and Satellite Systems (SpaceSec) 2026
23 February 2026, San Diego, CA, USA
ISBN 978-1-970672-02-2
https://dx.doi.org/10.14722/spacesec.2026.23005
www.ndss-symposium.org

applications by deployment environment (onboard vs. ground) and functional role (operations vs. payload).
- **Domain-specific attack mapping:** Links between AML attack types and real satellite use cases (e.g., anti-jamming, telemetry prediction), specifying attack surfaces and access assumptions.
- **FRAME-based risk assessment:** An adaptation of FRAME for satellites, combining attacker modeling and comparative risk scoring by feasibility, access, and mission impact.
- **Operational mitigation:** Countermeasures tailored to space constraints, including model hardening, data validation, and secure ML lifecycle management.

By quantifying adversarial risks to ML in satellite systems, this study contributes to developing and validating secure and resilient AI-enabled space operations.

II. Cyber Threats to Satellite Systems

Cyber threats targeting satellites, including those that affect ML components, can be grouped into four main categories.

Communication-Based Attacks. These attacks exploit vulnerabilities in satellite communication channels and have been observed in real-world conflicts. *Jamming* uses radio frequency (RF) interference to disrupt uplinks or downlinks [17], [18]. *Spoofing* transmits falsified signals to mislead users or satellite systems by imitating legitimate channels such as GNSS, telemetry, or communications [19]. *Man-in-the-middle (MitM)* attacks intercept or alter data in transit between satellites and ground systems [20]. *Replay* attacks resend captured commands or telemetry to trigger unauthorized actions [21]. *Denial-of-service (DoS)* attacks overwhelm networks or ground equipment, as seen in the Viasat hack [22].

ML Model Manipulation Attacks. These attacks directly target ML models. *Adversarial examples* are crafted inputs that induce misclassification [23]. *Data poisoning* corrupts training datasets to bias model behavior [24]. *Backdoors (Trojans)* embed hidden triggers during training [25]. *Membership inference* attacks aim to determine whether specific records were part of a training set [26].

Hardware and Physical Attacks. These attacks target a satellite's physical components or hardware behavior, exploiting natural effects, deliberate faults, or manufacturing weaknesses. *Cosmic radiation* induces single-event upsets (SEUs), as observed in the Galaxy 15 and Hubble incidents [1]. *Fault injection* exploits lasers or voltage glitches to create controlled errors [27]. *Supply chain threats* arise when malicious modifications are inserted during the design or manufacturing of satellite components. *Side-channel attacks* extract information from power use, emissions, or timing [4].

Perception and Planning Manipulation Attacks. These attacks distort a satellite's awareness of its environment. *Sensor spoofing* introduces false sensor data, e.g., fake fire detections with MODIS [28]. Lasers can blind imaging satellites, while camouflaging deceives AI-based perception [29]. Speculative threats include the injection of fake debris to force unnecessary maneuvers [13].

III. Use of ML in Satellites

ML supports a growing range of satellite functions across subsystems and mission segments, as illustrated in Figure 1. These applications span communications, control, perception, and system health.

Anti-Jamming. ML models have been used to classify spectrum occupancy and adapt transmission strategies under interference. SVMs, CNNs, and GANs have been tested in ground-based studies [30]–[32], while DRL agents have been simulated onboard to evade adaptive jammers [14], [33]. These efforts remain experimental and mainly relevant to the TT&C subsystem, which handles both uplink and downlink communication.

Beam Hopping. ML has been proposed to replace rule-based beam scheduling with adaptive control. DRL and hybrid optimization methods have demonstrated higher throughput in simulations [34], [35], though no in-orbit deployments have yet occurred. Adaptive payloads such as ESA's JoeySat illustrate potential future integration [7]. This domain primarily involves the Payload subsystem, where configurable antennas and signal processors could host ML-based beam management.

Command and Control Optimization. RL and LSTM-based controllers have been explored for attitude stabilization and autonomous maneuver planning [16], [36]. These functions depend on the ADCS and TT&C subsystems, which govern satellite orientation and command execution.

Cybersecurity and Interference Detection. AEs, GANs, and other deep models have been applied to detect anomalies in telemetry or network traffic [37]–[39]. Despite promising accuracy, deployment is limited by the scarcity of labeled data and high false-alarm costs. Cyber threats span TT&C (command/telemetry integrity), C&DH (internal data flows), Payload (mission data), and the ground segment, which is often the first attack surface [6].

Mission Planning and Space Traffic Management. RL and meta-learning techniques have been used for collision avoidance, observation scheduling, and dynamic task allocation [40]–[42]. Learning-based conjunction prediction has also been demonstrated [43]. These applications mainly operate in the ground segment, interfacing with TT&C for command uplink and ADCS for maneuver execution.

Network Traffic Management. Within non-terrestrial networks (NTN) standardized by 3GPP [44], ML methods such as RL, DL, and LSTMs have been explored for resource allocation, routing, and anomaly detection [39], [45], [46]. Most work remains at the simulation level, with limited deployment in LEO or GEO systems. The use case mainly concerns the ground and user segments.

Object Detection and Vision. CNNs, LSTMs, and transformers process optical and hyperspectral imagery for debris detection, environmental monitoring, etc. [47]–[49]. Onboard inference, demonstrated by ESA's Φ-Sat-1 [50], reduces downlink needs. However, models remain vulnerable to adversarial perturbations [51], [52] and constrained by limited onboard computing. This domain mainly involves the Payload, supported by ADCS for precise pointing and C&DH for data handling.

Telemetry Analysis. CNNs, LSTMs, and AEs have been applied to detect anomalies in satellite telemetry [37], [53], [54]. Performance is often overestimated, and ML currently acts as an advisory tool for operators rather than an autonomous diagnostic system. Key subsystems include the EPS and TT&C, where anomalies in power or communication integrity signal early failures.

Overall, ML now supports communications, control, perception, and system monitoring across both onboard and ground segments, though operational maturity remains uneven and most applications are still validated through simulation or testbed studies.

IV. Related Work

A. Risk Analysis in the ML Domain

Several efforts have been made to formalize how risks in ML systems should be identified and prioritized. The NIST AI Risk Management Framework (RMF) provides governance-level guidance for trustworthy AI, standardizing terminology and emphasizing processes for risk identification and mitigation; however, it is intentionally domain-agnostic [55]. MITRE's ATLAS knowledge base complements this by documenting adversarial Tactics, Techniques, and Procedures (TTPs) used against ML systems, offering an ATT&CK-style taxonomy tailored to AML [56].

Building on this, NIST recently published a taxonomy dedicated to AML that systematizes attacks and mitigations across the ML lifecycle [57]. It distinguishes between predictive and generative AI systems, classifies attacks by attacker goals-such as compromising availability, integrity, or privacy-along with attacker capabilities and knowledge, and provides a common terminology for evaluating mitigations. In parallel, recent surveys emphasized the importance of applying such structured risk views in critical industries. Pelekis et al. [58] reviewed AML across the automotive, healthcare, energy, and large language model (LLM)-driven NLP domains, finding that while taxonomies and defenses exist, practical robustness and privacy assessments in high-stake domains remain fragmented and often outdated.

On the methodology side, FRAME is a general-purpose and automated framework for AML risk assessment [5]. FRAME quantifies adversarial risks by integrating attack feasibility, system context, and empirical success rates reported in the literature, producing prioritized risk scores across use cases.

While these frameworks and reviews provide a foundation for understanding ML risks, they have primarily been applied in terrestrial or general-purpose contexts. None directly addresses the challenges of ML when deployed in space missions, where constraints such as limited power, radiation, and communication windows influence both the feasibility and impact of attacks. This gap motivates the need for a framework that considers satellite-specific threat models and operational realities.

B. Risk Analysis in the Satellite Domain

Risk analysis in the satellite domain has largely focused on cyber and system vulnerabilities. Prior work has mapped attack paths across ground, space, and RF segments for CubeSats and small satellites [59], exposed exploitable flaws in in-orbit firmware and telecommand interfaces [60], and demonstrated hosted-payload compromise scenarios such as OPS-SAT [61], [62]. Cyber-threat frameworks including MITRE ATT&CK and SPARTA have been adapted to satellite architectures, enabling analysis of real attacks on ground infrastructure [63] and the development of LEO-specific taxonomy extensions [4], as well as informing security controls in SDN-based networks and live hacking exercises [64], [65]. However, existing studies emphasize traditional cyber risks and generally overlook AML threats, space-specific constraints (e.g., SWaP, latency, radiation), and distinct attack surfaces of onboard vs. ground ML pipelines.

C. Existing Gap

Despite the progress made by both research communities, a unified perspective on AML risks for satellites is still lacking. Specifically, there is a need for a framework that bridges these domains: (1) a framework that tailors adversarial tactics and defenses to satellite operational realities (SWaP limits, radiation effects, high latency, intermittent links, and certification constraints), (2) distinguishes between onboard and ground-based ML pipelines and their distinct attack surfaces, and links attacks to their mission-level impact on availability, integrity, and privacy. Our work addresses this gap by extending an established AML risk methodology (FRAME) through the integration of space-specific literature, datasets, and threat models. The result is a satellite-aware AML risk framework that adapts existing adversarial classifications to space environments and provides prioritized risk scores grounded in real satellite ML applications.

V. Threat Models and Risk Scenarios

This section outlines the threat landscape: first identifying the actors and their motives (*who* and *why*), then describing the attack methods (*how*) that compromise ML systems across the satellite lifecycle. Together, these perspectives form the basis for a structured threat mapping, summarized in Table I, which links each attack category to its relevance across representative satellite ML use cases.

A. Threat Actors: Who and Why

Adversarial threats to ML-enabled satellites arise from actors with differing motives, capabilities, and resources [4], [55], [66], [67]. They can be broadly categorized into six tiers:

Tier 1 – Individuals and Activists: Motivated by curiosity, ideology, or notoriety, these actors conduct low-skill attacks such as public red-teaming or surface-level ML manipulation to gain attention or test boundaries.

Tier 2 – Commercial Competitors: Driven by economic or strategic gain, competitors may steal proprietary models or

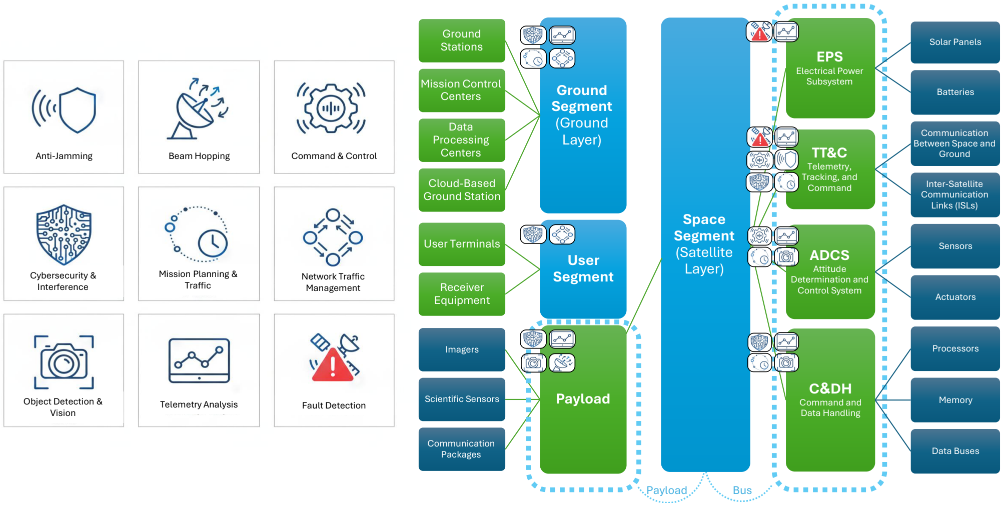

Anti-Jamming
Beam Hopping
Command & Control
Cybersecurity & Interference
Mission Planning & Traffic
Network Traffic Management
Object Detection & Vision
Telemetry Analysis
Fault Detection
Ground Stations
Mission Control Centers
Data Processing Centers
Cloud-Based Ground Station
Ground Segment (Ground Layer)
User Terminals
Receiver Equipment
User Segment
Imagers
Scientific Sensors
Communication Packages
Payload
Space Segment (Satellite Layer)
EPS Electrical Power Subsystem
TT&C Telemetry, Tracking, and Command
ADCS Attitude Determination and Control System
C&DH Command and Data Handling
Solar Panels
Batteries
Communication Between Space and Ground
Inter-Satellite Communication Links (ISLs)
Sensors
Actuators
Processors
Memory
Data Buses
Payload
Bus

Fig. 1: Mapping of ML use cases to satellite subsystems.

poison datasets to degrade rival performance, delay launches, or gain a technological advantage.

Tier 3 – Terrorist groups: Seeking disruption and public distrust, they may jam communications, spoof sensors, or inject adversarial noise into ML pipelines to amplify operational chaos.

Tier 4 – Insiders: Operators, contractors, or supply-chain staff with legitimate access who exploit trust for personal gain, coercion, or ideological purposes-by embedding backdoors, leaking data, or tampering with retraining processes.

Tier 5 – Organized crime: Financially motivated groups executing ransomware, data theft, or disruption-for-hire schemes, often via compromised supply chains or collaboration with insiders.

Tier 6 – State-Sponsored Actors: Pursuing geopolitical or military objectives, these actors employ advanced tactics such as model inversion, stealthy poisoning, or hardware compromise to degrade adversary capabilities covertly.

B. Adversarial Attack Methods: How

The *how* refers to the main attack vectors that exploit ML throughout the satellite lifecycle [56], [66], [67].

Communication-based attacks: Jamming and spoofing corrupt or block telemetry and sensor streams, resulting in control or anomaly detection models failing to function correctly. MitM, replay, or DoS attacks inject false or stale data into inference or training pipelines, undermining reliability.

Model manipulation: Adversarial examples induce misclassification, while data poisoning or backdoors corrupt models during training. Membership inference poses a threat to confidentiality, but it rarely directly impacts operations.

Hardware and physical attacks: Radiation or fault injection can alter model weights; supply-chain tampering can embed persistent vulnerabilities; side-channel attacks may expose parameters or sensitive data.

Perception and planning manipulation: False sensor inputs mislead ML models used in navigation and tasking, triggering unnecessary maneuvers or resource misallocation.

VI. Research Methodology

This study proposes a methodology for assessing the risks associated with the use of ML in satellite systems. The methodology combines system identification, threat modeling, and a structured evaluation stage based on the FRAME risk analysis framework [5], a comprehensive and automated approach for assessing risks posed by AML threats across diverse ML-based systems.

The proposed methodology, building on FRAME, encompasses three primary dimensions: the system's deployment environment, the characteristics of adversarial ML techniques, and empirical data from prior research. To adapt the assessment to each system's operational context, FRAME's system-profiling process utilizes a structured questionnaire guided by an LLM model, which provides automated, context-aware profiling. To further specialize this profiling for the satellite domain, we introduce additional parameters. These parameters help guide the LLM's questioning to effectively capture space-specific constraints and mission architectures.

FRAME then performs attack feasibility impact mapping, linking AML attacks to specific feasibility conditions and security impacts, informed by expert knowledge and a literature review. A comprehensive empirical dataset on AML attacks is used to estimate the realistic success rates of various attack techniques in the system's context. These inputs are integrated through a modeling component that quantifies the attack risks, which are then ranked and presented for actionable decision-making. FRAME balances ease of use for technical system owners without AML expertise with detailed, data-driven risk prioritization, supporting both effective mitigation strategies and the secure deployment of ML technologies.

To tailor FRAME to the satellite domain, we introduced two main changes. First, we expanded the dataset of documented

TABLE I: Threat model relevance to satellite ML use cases (●= high relevance ◐= medium relevance, ○= low or no relevance).

Use Case	Communication-Based Attacks					Model Manipulation Attacks					Hardware & Physical Attacks			Perception & Planning Manip.	
	Jamming	Spoofing	MitM	DoS	Replay Attacks	Adv. Attack	Backdoor	Model Invers.	Model Extract.	Mem. Infer.	Fault Inject.	Cosmic Rad.	Side Chan.	Sensor Spoof.	Debris Manip.
Anti-Jamming	●	●	●	◐	○	○	○	○	○	○	○	○	○	○	○
Beam Hopping	●	●	○	○	○	○	○	○	○	○	○	○	○	○	○
Command & Control	○	○	●	○	●	○	○	●	○	○	◐	○	○	○	○
CyberSec & Intrusion det.	○	○	○	◐	○	●	●	◐	●	◐	○	○	○	○	○
Mission Planning	○	○	○	○	●	○	○	○	○	○	●	○	○	○	●
Net. Traffic Mgmt	○	○	○	●	○	●	○	◐	●	○	○	○	○	○	○
Obj. Detection & Vision	○	○	○	○	○	○	○	○	○	○	●	○	○	●	○
Telemetry Analysis	○	○	○	○	○	○	○	○	○	●	○	●	●	○	○

attacks by incorporating AML research specific to satellites and space systems. This allowed us to refine the feasibility mappings and calibrate the attack success rates using domain-relevant evidence. Second, we extended the FRAME questionnaire with satellite-specific questions, capturing operational parameters such as the deployment segment, orbital regime, and organizational ownership. These adjustments ensure that FRAME produces risk scores that reflect the realities of ML deployments in both onboard and ground-based satellite missions. The following sections outline the approach taken to extend and apply FRAME for satellite-specific risk assessment.

A. Adjustment of FRAME

We extend FRAME with satellite-specific parameters to capture how deployment context and organizational setting affect attack feasibility. These adjustments ensure the framework reflects realistic conditions across different satellite environments.

a) Model Deployment Segment: The ML model's location in the satellite architecture (ground, bus, payload, or user segment) strongly affects attacker accessibility [66], [67]. User-level components present the highest feasibility (w=1.0) due to exposed interfaces and data endpoints, followed by ground systems (w=0.7), where terrestrial access and ground-station vulnerabilities are well known. Payload modules exhibit moderate feasibility (w=0.4), offering partial isolation while retaining mission-specific interfaces, whereas the bus remains the least accessible (w=0.1) due to its tightly integrated nature. These values are heuristic indicators for relative comparison rather than measured probabilities.

b) Orbital Regime: Attack feasibility also varies with orbit. LEO missions have the highest feasibility (w=1.0) as they rely on COTS components, dense constellations, and short lifecycles that limit per-satellite hardening. MEO missions are moderately exposed (w=0.8), balancing accessibility with higher operational costs and longer missions, while GEO platforms are least feasible (w=0.7) due to their strong control infrastructure and minimal physical access [6], [66].

c) Organizational Ownership: The type of operator further influences the security posture [55], [67]. Private and research operators face the highest feasibility (w=0.8), given their limited budgets and oversight. Commercial entities exhibit moderate feasibility (w=0.5), as protection levels vary by regulation and business scale. Governmental systems exhibit the lowest feasibility (w=0.2), due to strict certification, defense-grade controls, and active monitoring.

These heuristic parameters extend FRAME to capture satellite-specific architectural and organizational factors, enabling consistent, domain-relevant risk comparisons across deployment contexts.

To validate these heuristic weights, we performed a sensitivity analysis. We tested multiple logical scenarios, such as (1) increasing the risk variance between orbital regimes (e.g., w_{LEO}=1.0, w_{GEO}=0.4) and (2) assuming lower-variation risk (e.g., w_{LEO}=1.0, w_{GEO}=0.9). We observed that while absolute risk scores shifted, the relative prioritization of threats remained robust: Integrity-based attacks (e.g., Model Poisoning) consistently ranked as the dominant threat across all tested scenarios. This stability supports our chosen weights as a representative baseline for this domain.

d) Comparison with Terrestrial Domains: Unlike terrestrial ML deployments—for example, those in IoT or industrial control systems—satellite environments face constraints that strongly influence attack feasibility: intermittent connectivity, harsh physical conditions, radiation-induced faults, and reliance on SWaP-limited hardware. In its standard form, FRAME assumes continuous access and abundant computing resources, conditions rarely present in orbit. Our extension, therefore, ensures that risk scores reflect the economic, physical, and governance realities of space missions, enabling adversarial risk assessment in the context of satellite systems.

B. Identification of AI/ML Systems in Satellites

As an initial step, we surveyed contemporary literature, technical standards, and operational documentation to identify ML-based systems used in modern satellite infrastructure, including both onboard and ground-based deployments. The identified use cases were then classified across key operational areas, including beam hopping, anti-jamming, telemetry analysis, mission planning, and cybersecurity.

C. Threat Model Identification

After mapping the identified AI/ML components to their respective satellite subsystems and operational contexts, we compiled a comprehensive set of threat models relevant to satellite-based AI systems. The models encompass a wide range of attacks: Communication-based attacks (e.g., jamming, spoofing, MitM), Model manipulation attacks (e.g., adversarial examples, backdoors), Hardware and physical attacks (e.g., cosmic radiation, fault injection), and Planning and perception manipulation attacks (e.g., sensor spoofing, orbital debris simulation). These threats were adapted and refined based on existing standards such as CCSDS 350.1-G [68] and recent adversarial machine learning research.

D. Mapping Relevant Threat Models to Use Cases

Each threat model was systematically mapped to the AI/ML use cases identified in the previous subsection. The mapping accounted for both the deployment environment (onboard vs. ground) and the functional scope (mission-critical vs. payload-focused). The outcome is summarized in Table I, which presents a threat–use case relevance matrix highlighting the high-impact intersections—that is, where specific attack types are most feasible or consequential within a given operational context.

E. Application of the Extended FRAME Framework

We applied the extended FRAME framework [5] to each scenario, producing descriptive and score-based assessments of adversarial ML risks. The analysis encompassed various deployment environments and mission types, informed by prior studies on ML assurance in space systems [54], [69].

F. Analysis and Conclusions

The resulting FRAME outputs—risk scores and threat mappings—were analyzed to reveal recurring vulnerabilities across ML architectures and mission profiles (e.g., LEO telemetry vs. GEO beam hopping). High-risk intersections, such as onboard control models exposed to adversarial inference, were prioritized for further study. The results informed recommendations for risk mitigation, design hardening, and research to enhance the security and resilience of ML-based satellite systems.

VII. Evaluation

A. Evaluation Setup

The evaluation focuses on five representative ML use cases selected to capture the diversity of machine learning adoption in satellite missions. These cases span multiple orbital regimes (LEO–GEO), deployment layers (onboard and ground segments), and functional domains, each drawn from a distinct study. Together, they represent a broad spectrum of learning paradigms-from deep reinforcement learning for adaptive control to supervised and unsupervised models for anomaly detection. This diversity enables the extended FRAME framework to be evaluated across varied operational contexts and mission objectives. The five use cases are outlined below.

B. Representative Use Cases

To evaluate adversarial risks in realistic mission contexts, we analyze five representative ML applications across satellite domains and deployment environments. These use cases illustrate the breadth of ML adoption in both onboard and ground-based systems and form the basis for applying the adapted FRAME methodology.

1) Case A: Multi-Agent DRL-Based Anti-Jamming Spectrum Access (LEO): This case considers a multi-agent DRL framework for anti-jamming in LEO satellite networks. The study “A Multi-Agent Deep Reinforcement Learning Anti-Jamming Spectrum-Access Method in LEO Satellites” (Electronics, 2025) [70] proposes a VDN-based approach with centralized training and distributed execution. After offline ground training, the model is deployed onboard to enable real-time, decentralized spectrum-access decisions under jamming.

2) Case B: Beamforming Optimization (GEO, Beam Hopping Family): This case examines AI-driven dynamic beamforming for multibeam GEO satellites, considering both onboard and ground ML architectures. The study “Machine Learning for Radio Resource Management in Multibeam GEO Satellite Systems” (Electronics, 2022) [71] evaluates RL and supervised models for beam hopping to improve spectral efficiency. Our assessment focuses on the onboard configuration, which increases operational flexibility but also raises exposure to poisoning and evasion attacks due to limited retraining and security constraints.

3) Case C: Payload-Based Dynamic Frequency Allocation (Beam Hopping Family): This case addresses onboard AI control for beam hopping and frequency allocation. The work “Deep Reinforcement Learning-Based Beam Hopping Algorithm in Multibeam Satellite Systems” (IET Communications, 2019) [72] formulates illumination planning as a Markov decision process using a DQN framework. Simulation results demonstrate reduced latency and improved throughput, highlighting the transition from rule-based payload management to autonomous onboard optimization.

4) Case D: Traffic Anomaly Detection (Network Management): This case analyzes ML-based interference detection in the ground segment. The article “Machine Learning for Satellite Communications Operations” (IEEE Communications Magazine, 2021) [73] presents a CNN-based autoencoder deployed at the network operations center. By processing IQ signal samples from satellite transponders, the model identifies anomalies associated with interference or link degradation.

5) Case E: Satellite Health Anomaly Detection (Telemetry Analysis): This case examines telemetry-based health monitoring using ML techniques. The study “Artificial Intelligence for Satellite Communication: A Review” (Intelligent and Converged Networks, 2021) [16] describes a multivariate LSTM combined with probabilistic PCA for anomaly detection. The model analyzes temperature, voltage, current, and sensor data to detect deviations related to subsystem faults or degradation. Together, these cases cover the primary operational domains of satellite ML-communications, control, and health management-and serve as the foundation for the subsequent FRAME-based risk assessment.

C. Cross-Use-Case Insights and Observations

By analyzing the quantitative and qualitative outputs generated by the extended FRAME framework across five diverse satellite ML use cases, we identified several recurring security challenges and risk patterns. The following insights and observations are derived from this cross-use-case analysis, focusing on the **top-five attacks identified by FRAME** for each use case.

1) Common Attack Vectors: Table II summarizes the attack categories most frequently observed in the evaluated use cases, organized by their corresponding security objectives.

TABLE II: Summary of recurring attack vectors across the evaluated use cases.

Attack Category	**Description**	**Affected Use Cases**
Model Poisoning Attacks (Integrity)	Targeted manipulation of training or retraining pipelines to degrade model behavior	A, B, C, D, E
Evasion Attacks (Integrity)	Input manipulation to cause misclassification or bypass anomaly detection	A, B, C, D, E
Resource Latency Attacks (Availability)	Overloading model computation or inducing delays in real-time operations	A, B, D
Data Reconstruction / Model Extraction (Privacy)	Reverse-engineering model behavior or reconstructing sensitive data from outputs	B, D, E
Clean-Label Poisoning (Availability)	Introducing undetectable malicious samples to degrade performance over time	C, D, E

2) Risk Categorization by Objective and Attack Type: The heatmaps in Figure 2a visualize the relative risk scores aggregated across the five use cases, categorized by security objective (*integrity, availability, privacy*), while Figure 2b groups them by attack type (*poisoning, evasion, resource-latency, and data reconstruction*). Together, these visualizations highlight that integrity-related attacks consistently pose the highest risks, particularly through poisoning and evasion tactics.

3) Key Observations:

- **Integrity Risks Dominate:** *Model poisoning* and *evasion* attacks were identified in all use cases, and they obtained the highest average risk scores.
- **Model Poisoning is the Most Recurring Threat:** The *black-box interactive decision-based targeted model poisoning* attack was observed in four of the five use cases, with an average risk score of 8.42.
- **Evasion Attacks Are a Secondary Concern:** The *black-box interactive decision-based evasion or misclassification* attack was seen in two use cases, with an average risk score of 7.99.
- **Availability Threats Are Context-Specific:** Real-time systems (e.g., interference management, beam hopping) are more exposed to latency and *resource exhaustion* attacks than batch-processing use cases like telemetry analysis.
- **Privacy Risks Are Secondary but Present:** *Data reconstruction* and *model extraction* attacks were noted in onboard and telemetry-related use cases.
- **Importance of Pipeline Security:** All of the use cases require secure retraining and data ingestion pipelines to defend against *poisoning* attacks.

VIII. Discussion

A. Interpretation and Implications of Findings

The FRAME-based analysis reveals a clear pattern across satellite ML applications: integrity risks dominate due to the unique data and operational dynamics of space systems. These results reinforce the paper's central argument that ML introduces mission-level vulnerabilities not captured by traditional satellite cybersecurity frameworks.

Specifically, the high recurrence of poisoning and evasion attacks across both onboard and ground segments demonstrates that adversarial manipulation is not a theoretical threat but a feasible and cross-domain risk. Such threats correspond to different actor motives: poisoning risks align with commercial competitors seeking to degrade rival performance, while evasion attacks reflect state-backed attempts to disrupt communications or control functions. Lower-impact data extraction and reconstruction threats, more typical of activists or organized crime, highlight that motivations and capabilities vary widely across the actor spectrum. This variation underscores the importance of integrating adversarial resilience as a design requirement, rather than as a post-deployment safeguard.

Moreover, the context-specific exposure of availability attacks in real-time control and communication functions highlights that risk cannot be generalized across missions. Instead, risk prioritization must consider the mission profile, retraining frequency, and operational autonomy level.

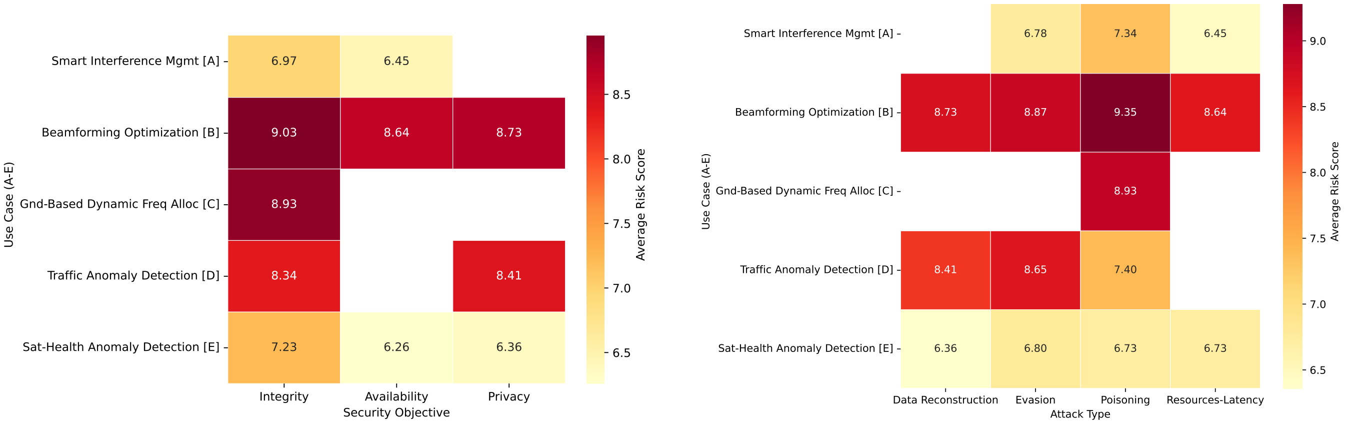
Taken together, these findings validate the need for a satellite-tailored risk framework, such as FRAME, adapted and scaled to address the unique conditions of the space domain, capable of quantifying not only attack feasibility but also mission impact. They provide the empirical basis for developing targeted mitigations that are relevant to the operational criticality of each ML component.

B. Mitigation Strategies

Building on the FRAME-based risk analysis and the identified threat landscape across multiple ML-enabled satellite use cases, this section outlines targeted mitigation strategies addressing recurring vulnerabilities observed in the evaluated systems.

1) Robust Model Training: Adversarial poisoning and evasion attacks were consistently ranked among the most severe threats. To mitigate these risks, satellite ML pipelines should incorporate adversarially robust training techniques, including the use of perturbations similar to known evasion examples and data sanitization procedures that limit the impact of poisoning during retraining or online learning. These measures align with prior recommendations [54], [69], which emphasize robustness in continuously learning and feedback-driven satellite systems.

2) Pipeline Security and Update Hardening: Continuous retraining on non-stationary data introduces new attack surfaces for model poisoning and drift. To reduce this exposure, data pipelines should enforce source verification, anomaly filtering, and authenticated ingestion. Automated integrity checks must validate incoming data before retraining is initiated,



(a) By security objective

(b) By attack type

Fig. 2: Expanded FRAME risk heatmaps for satellite ML use cases. (a) presents the average risk scores categorized by security objective (integrity, availability, privacy). (b) presents the average risk scores grouped by attack type (poisoning, evasion, resource depletion, latency, and data reconstruction). The risk levels are based on the top-5 adversarial attacks per use case.

and ML model updates should remain versioned, auditable, and reversible to ensure safe rollback in the event of post-deployment anomalies.

3) Runtime Monitoring and Anomaly Detection: Real-time adversarial detection is particularly critical for LEO satellites, which operate under rapidly changing conditions and limited contact windows. Lightweight onboard detectors, such as CNN-based autoencoders, can monitor inference consistency and alert operators to deviations. Cross-verification between onboard and ground-based inference results provides an additional layer of protection against signal spoofing or data tampering during transmission.

4) Redundancy and Fail-Safe Logic: Given the potential for radiation faults, denial-of-service conditions, or ML malfunction, critical satellite functions should not rely on a single inference path. ML models must be wrapped in fail-safe control logic that defaults to deterministic heuristics under abnormal conditions. Backup models or rule-based fallback policies should govern essential functions such as routing, beam allocation, or power management to ensure continued operation during degraded states.

5) Periodic Risk Audits and Threat Re-Evaluation: Because ML systems evolve and adversarial techniques advance, FRAME-based risk assessments should be periodically revisited. Annual re-evaluations using updated datasets and emerging adversarial findings help maintain system resilience. Furthermore, any major software, firmware, or retraining modification should trigger reapplication of the threat mapping and scoring process to ensure that mitigations remain valid and effective.

These combined strategies do not eliminate adversarial risk entirely but establish a layered, practical defense tailored to the operational, physical, and computational constraints of ML-based satellite systems.

IX. Conclusion

This study presented the first comprehensive risk assessment of AML threats in satellite systems, adapting the FRAME methodology to the space domain. By analyzing five representative use cases across onboard and ground segments, we identified recurring vulnerabilities-particularly model poisoning, evasion, and data manipulation. We demonstrated how risk exposure varies with deployment context and mission role.

Our adaptation of FRAME integrates space-specific literature and expert insights, enabling the quantification of realistic risks beyond generic AML evaluations. Findings indicate that integrity threats dominate most use cases, availability risks are critical for time-sensitive functions such as anti-jamming and beam hopping, and privacy issues emerge in telemetry and data services. These results offer a baseline for prioritizing ML robustness in space operations.

To bridge the gap between research and practice, we outlined mitigation measures tailored to satellite constraints, including adversarial training, hardened data pipelines, runtime anomaly detection, redundancy, and periodic risk audits. Together, these form a layered defense strategy that enhances ML resilience without compromising mission reliability.

Future work should focus on three directions: (1) hardware-in-the-loop validation of defenses under adversarial conditions, (2) development of satellite-specific AML datasets and benchmarks, and (3) integration with standards such as NIST AI RMF, MITRE ATLAS, and SPARTA to establish a unified framework for space AML risk assessment. As ML adoption in satellites expands, future work should calibrate the heuristic weights using expert feedback and structured surveys to achieve a more precise and empirically grounded characterization of satellite-specific risks.

By quantifying risks and proposing actionable defenses, this study offers a proactive step toward secure and trustworthy ML autonomy in orbit.

References

[1] J. Goodwill, C. Wilson, and J. MacKinnon, "Current ai technology in space," in *Precision Medicine for Long and Safe Permanence of Humans in Space*. Elsevier, 2025, pp. 239–250.

[2] A. Bhattacharyya, S. M. Nambiar, R. Ojha, A. Gyaneshwar, U. Chadha, and K. Srinivasan, "Machine learning and deep learning powered satellite communications: Enabling technologies, applications, open challenges, and future research directions," *International Journal of Satellite Communications and Networking*, vol. 41, no. 6, pp. 539–588, 2023.

[3] F. Ortiz, V. Monzon Baeza, L. M. Garces-Socarras, J. A. Vásquez-Peralvo, J. L. Gonzalez, G. Fontanesi, E. Lagunas, J. Querol, and S. Chatzinotas, "Onboard processing in satellite communications using ai accelerators," *Aerospace*, vol. 10, no. 2, p. 101, 2023.

[4] R. Peled, E. Aizikovich, E. Habler, Y. Elovici, and A. Shabtai, "Evaluating the security of satellite systems," *arXiv preprint arXiv:2312.01330*, 2023.

[5] A. Shapira, S. Shigol, and A. Shabtai, "Frame: Comprehensive risk assessment framework for adversarial machine learning threats," *arXiv preprint arXiv:2508.17405*, 2025.

[6] NASA Office of Inspector General, "Cybersecurity management and oversight at the jet propulsion laboratory," National Aeronautics and Space Administration, Tech. Rep. IG-19-022, June 2019, accessed: 2025-05-21. [Online]. Available: https://oig.nasa.gov/wp-content/uploads/2024/02/IG-19-022.pdf

[7] E. S. Agency, "Beam-hopping joeysat launched," European Space Agency News Release, 2023, (accessed Nov. 3, 2025).

[8] M. J. Veyette, K. Aylor, D. Stafford, M. Herrera, S. Jumani, C. Lineberry, C. Macklen, E. Maxwell, R. Stiles, and M. Jenkins, "Ai/ml for mission processing onboard satellites," in *AIAA SCITECH 2022 Forum*, 2022, p. 1472.

[9] P. Breda, R. Markova, A. F. Abdin, N. P. Mantı, A. Carlo, and D. Jha, "An extended review on cyber vulnerabilities of ai technologies in space applications: Technological challenges and international governance of ai," *Journal of Space Safety Engineering*, 2023.

[10] L. Jovanovic, N. Bacanin, V. Simic, J. Mani, M. Zivkovic, and M. Sarac, "Optimizing machine learning for space weather forecasting and event classification using modified metaheuristics," *Soft Computing*, vol. 28, no. 7, pp. 6383–6402, 2024.

[11] D. G. BOOK, "Wireless network communications overview for space mission operations," *CCSDS 880. 0-G, 2009*, 2009.

[12] A. Basrur, "Ai in space operations: Opportunities, challenges, and the path forward for india," Observer Research Foundation, 2025, available at https://www.orfonline.org/research/ai-in-space-operations/.

[13] P. Yue, J. An, J. Zhang, G. Pan, S. Wang, P. Xiao, and L. Hanzo, "On the security of leo satellite communication systems: Vulnerabilities, countermeasures, and future trends," *Authorea Preprints*, 2022.

[14] C. Han, L. Huo, X. Tong, H. Wang, and X. Liu, "Spatial anti-jamming scheme for internet of satellites based on the deep reinforcement learning and stackelberg game," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 5, pp. 5331–5342, 2020.

[15] E.-C. Chen, P.-Y. Chen, I. Chung, C.-R. Lee *et al.*, "Overload: Latency attacks on object detection for edge devices," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2024, pp. 24 716–24 725.

[16] F. Fourati and M.-S. Alouini, "Artificial intelligence for satellite communication: A review," *Intelligent and Converged Networks*, vol. 2, no. 3, pp. 213–243, 2021.

[17] J. Rainbow. (2022, October) Eutelsat says satellite jammers within iran are disrupting foreign channels. [Online]. Available: https://spacenews.com/eutelsat-says-satellite-jammers-within-iran-are-disrupting-foreign-channels/

[18] V. Insinna. (2022, April) Spacex beating russian jamming attack was 'eyewatering': Dod official. [Online]. Available: https://breakingdefense.com/2022/04/spacex-beating-russian-jamming-attack-was-eyewatering-dod-official/

[19] Center for Advanced Defense Studies, "Above us only stars: Exposing gps spoofing in russia and syria," Center for Advanced Defense Studies (C4ADS), Tech. Rep., March 2019, accessed: 2025-05-21. [Online]. Available: https://c4ads.org/wp-content/uploads/2022/05/AboveUsOnlyStars-Report.pdf

[20] A. A. Z. Hudaib, "Satellite network hacking & security analysis," *International Journal of Computer Science and Security (IJCSS)*, vol. 10, no. 1, p. 8, 2016.

[21] M. Liu, Z. Deng, and L. Jun, "Research of satellite receiver anti-replay attack techniques," in *China Satellite Navigation Conference (CSNC) 2015 Proceedings: Volume I*. Springer, 2015, pp. 503–516.

[22] A. Basrur, "Ai in space operations: Opportunities and challenges," Observer Research Foundation, ORF Issue Brief 791, April 2025, accessed: 2025-05-21. [Online]. Available: https://www.orfonline.org/research/ai-in-space-operations-opportunities-and-challenges

[23] A. Du, B. Chen, T.-J. Chin, Y. W. Law, M. Sasdelli, R. Rajasegaran, and D. Campbell, " Physical Adversarial Attacks on an Aerial Imagery Object Detector ," in *2022 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*. Los Alamitos, CA, USA: IEEE Computer Society, Jan. 2022, pp. 3798–3808. [Online]. Available: https://doi.ieeecomputersociety.org/10.1109/WACV51458.2022.00385

[24] Z. Wang, B. Wang, C. Zhang, Y. Liu, and J. Guo, "Defending against poisoning attacks in aerial image semantic segmentation with robust invariant feature enhancement," *Remote Sensing*, vol. 15, no. 12, p. 3157, 2023.

[25] T. Gu, B. Dolan-Gavitt, and S. Garg, "Badnets: Identifying vulnerabilities in the machine learning model supply chain," *arXiv preprint arXiv:1708.06733*, 2017.

[26] R. Shokri, M. Stronati, C. Song, and V. Shmatikov, "Membership inference attacks against machine learning models," in *2017 IEEE symposium on security and privacy (SP)*. IEEE, 2017, pp. 3–18.

[27] B. Cyr, Y. Long, T. Sugawara, and K. Fu, "Position paper: Space system threat models must account for satellite sensor spoofing." in *SpaceSec*, 2023.

[28] E. Salkield, S. Köhler, S. Birnbach, R. Baker, M. Strohmeier, and I. Martinovic, "Firefly: Spoofing earth observation satellite data through radio overshadowing," *SpaceSec 23*, 2023.

[29] K. Burke, *Where are the PLA's Other Laser Dazzling Facilities?* China Aerospace Studies Institute, 2023.

[30] R. Morales Ferre, A. De La Fuente, and E. S. Lohan, "Jammer classification in gnss bands via machine learning algorithms," *Sensors*, vol. 19, no. 22, p. 4841, 2019.

[31] W. Han, X. Song, Y. Huang, F. Yan, Q. Yin, and T. Zhang, "Deep learning based satellite communication anti-jamming system," in *2024 IEEE/ACIS 24th International Conference on Computer and Information Science (ICIS)*. IEEE, 2024, pp. 9–12.

[32] X. Ding, Y. Zhang, G. Li, X. Gao, N. Ye, D. Niyato, and K. Yang, "Few-shot recognition and classification framework for jamming signal: A cgan-based fusion cnn approach," *arXiv preprint arXiv:2311.05273*, 2023.

[33] Y. Li, Y. Xu, W. Li, G. Li, Z. Feng, S. Liu, J. Du, and X. Li, "Achieving hiding and smart anti-jamming communication: A parallel drl approach against moving reactive jammer," *arXiv preprint arXiv:2502.02385*, 2025.

[34] X. Xie, K. Fan, W. Deng, N. Pappas, and Q. Zhang, "Multi-satellite beam hopping and power allocation using deep reinforcement learning," *arXiv preprint arXiv:2501.02309*, 32025.

[35] S. Martínez Zamacola, F. Luna Valero, and R. Martínez Rogríguez-Osorio, "Hybrid moea with problem-specific operators for beam-hopping based resource allocation in multi-beam leo satellites," *Available at SSRN 5214437*, 2025.

[36] A. Harris and K. Naik, "Autonomous command and control for earth-observing satellites using deep reinforcement learning," in *2023 IEEE Aerospace Conference*. IEEE, 2023, pp. 1–12.

[37] L. Herrmann, M. Bieber, W. J. Verhagen, F. Cosson, and B. F. Santos, "Unmasking overestimation: a re-evaluation of deep anomaly detection in spacecraft telemetry," *CEAS Space Journal*, vol. 16, no. 2, pp. 225–237, 2024.

[38] M. Rath and S. Mishra, "Security approaches in machine learning for satellite communication," *Machine learning and data mining in aerospace technology*, pp. 189–204, 2020.

[39] N. Sitouah, F. Merazka, and A. Hedjazi, "Deep learning approach for interruption attacks detection in leo satellite networks," *arXiv preprint arXiv:2301.03998*, 2022.

[40] N. Bourriez, A. Loizeau, and A. F. Abdin, "Spacecraft autonomous decision-planning for collision avoidance: A reinforcement learning approach," *arXiv preprint arXiv:2310.18966*, 2023.

[41] P. Li, P. Cui, and H. Wang, "Mission sequence model and deep reinforcement learning-based replanning method for multi-satellite observation," *Sensors*, vol. 25, no. 6, p. 1707, 2025.

[42] W. Yao, X. Shen, G. Zhang, Z. Lu, J. Wang, and G. Gao, "Meta reinforcement learning method for dynamic mission scheduling of earth observation satellites," *Meta*, vol. 18, p. 5, 2025.

[43] F. Pinto, G. Acciarini, S. Metz, S. Boufelja, S. Kaczmarek, K. Merz, J. A. Martinez-Heras, F. Letizia, C. Bridges, and A. G. Baydin, "Towards automated satellite conjunction management with bayesian deep learning," *arXiv preprint arXiv:2012.12450*, 2020.

[44] 3rd Generation Partnership Project (3GPP), "Non-terrestrial networks (ntn) overview," https://www.3gpp.org/technologies/ntn-overview, 2024, accessed: 2025-09-10.

[45] Y. Feng, W. Cai, H. Yue, J. Xu, Y. Lin, J. Chen, and Z. Hu, "An improved x-means and isolation forest based methodology for network traffic anomaly detection," *Plos one*, vol. 17, no. 1, p. e0263423, 2022.

[46] J. Woo, S. Belvins, B. P. Michael, and T. Yates, "Detecting satellite laser ranging station data and operational anomalies with machine learning isolation forests at nasa's cddis," in *AGU Fall Meeting Abstracts, pp. IN12B–0269*, 2022.

[47] S. Gui, S. Song, R. Qin, and Y. Tang, "Remote sensing object detection in the deep learning era—a review," *Remote Sensing*, vol. 16, no. 2, p. 327, 2024.

[48] R. Rad, "Vision transformer for multispectral satellite imagery: Advancing landcover classification," in *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, 2024, pp. 8176–8183.

[49] R. G. Avilés, L. Scheibenreif, N. A. A. Braham, B. Blumenstiel, T. Brunschwiler, R. Guruprasad, D. Borth, C. Albrecht, P. Fraccaro, D. Lambhate *et al.*, "Hyperspectral vision transformers for greenhouse gas estimations from space," *arXiv preprint arXiv:2504.16851*, 2025.

[50] eoPortal, "Phisat-1 & -2 nanosatellite mission," 2023, accessed: 2025-05-21. [Online]. Available: https://www.eoportal.org/satellite-missions/phisat-1

[51] G. Tang, T. Jiang, W. Zhou, C. Li, W. Yao, and Y. Zhao, "Adversarial patch attacks against aerial imagery object detectors," *Neurocomputing*, vol. 537, pp. 128–140, 2023.

[52] W. Czaja, N. Fendley, M. Pekala, C. Ratto, and I.-J. Wang, "Adversarial examples in remote sensing," in *Proceedings of the 26th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*, 2018, pp. 408–411.

[53] K. Hundman, V. Constantinou, C. Laporte, I. Colwell, and T. Soderstrom, "Detecting spacecraft anomalies using lstms and nonparametric dynamic thresholding," in *Proceedings of the 24th ACM SIGKDD international conference on knowledge discovery & data mining*, 2018, pp. 387–395.

[54] K. Thangavel, R. Sabatini, A. Gardi, K. Ranasinghe, S. Hilton, P. Servidia, and D. Spiller, "Artificial intelligence for trusted autonomous satellite operations," *Progress in Aerospace Sciences*, vol. 144, p. 100960, 2024.

[55] E. Tabassi, "Artificial Intelligence Risk Management Framework (AI RMF 1.0)," Gaithersburg, MD, January 26 2023. [Online]. Available: https://doi.org/10.6028/NIST.AI.100-1

[56] MITRE, "MITRE ATLAS: Adversarial Threat Landscape for AI Systems," Fact Sheet; MITRE Corporation, 2024, a living knowledge base of adversarial tactics, techniques, and procedures (TTPs) for AI systems. [Online]. Available: https://atlas.mitre.org

[57] A. Vassilev, A. Oprea, A. Fordyce, and H. Andersen, "Adversarial machine learning: A taxonomy and terminology of attacks and mitigations," *NIST*, 2024.

[58] S. Pelekis, T. Koutroubas, A. Blika, A. Berdelis, E. Karakolis, C. Ntanos, E. Spiliotis, and D. Askounis, "Adversarial machine learning: a review of methods, tools, and critical industry sectors," *Artificial Intelligence Review*, vol. 58, no. 8, p. 226, 2025.

[59] G. Falco, A. Viswanathan, and A. Santangelo, "Cubesat security attack tree analysis," in *2021 IEEE 8th International Conference on Space Mission Challenges for Information Technology (SMC-IT)*. IEEE, 2021, pp. 68–76.

[60] J. Willbold, M. Schloegel, M. Vögele, M. Gerhardt, T. Holz, and A. Abbasi, "Space odyssey: An experimental software security analysis of satellites," in *2023 IEEE Symposium on Security and Privacy (SP)*. IEEE, 2023, pp. 1–19.

[61] M. Calabrese, G. Kavallieratos, and G. Falco, "A hosted payload cyber attack against satellites," in *AIAA SCITECH 2024 Forum*, 2024, p. 0270.

[62] B. Ruszczak, K. Kotowski, D. Evans, and J. Nalepa, "The ops-sat benchmark for detecting anomalies in satellite telemetry," *Scientific Data*, vol. 12, no. 1, p. 710, 2025.

[63] J. Hamill-Stewart and A. Rashid, "Threats against satellite ground infrastructure: A retrospective analysis of sophisticated attacks," in *Proceedings of the 2024 Workshop on Security of Space and Satellite Systems*, vol. 1, 2024.

[64] J. Slay, U. Uhongora, J. Plotnek, Y. W. Law, and R. Mulinde, "Applying the sparta matrix to develop intelligent security controls for space systems," in *ASCEND 2023*, 2023, p. 4770.

[65] J. Curbo and G. Falco, "Attack surface analysis for spacecraft flight software," in *2024 IEEE 10th International Conference on Space Mission Challenges for Information Technology (SMC-IT)*. IEEE, 2024, pp. 22–30.

[66] The Aerospace Corporation, "Sparta: Space attack research and tactic analysis," 2025, accessed: 2025-09-02. [Online]. Available: https://aerospace.org/sparta

[67] E. U. A. for Cybersecurity, "Space threat landscape," ENISA, Tech. Rep., 2025, (accessed Nov. 3, 2025).

[68] G. Book, "Security threats against space missions," *CCSDS Secretariat: Washington, DC, USA*, 2006.

[69] A. Weber and P. Franke, "Space-domain ai applications need rigorous security risk analysis," in *Proc. Workshop NDSS*, 2024, pp. 1–10.

[70] W. Cao, F. Chu, L. Jia, H. Zhou, and Y. Zhang, "A multi-agent deep reinforcement learning anti-jamming spectrum-access method in leo satellites," *Electronics*, vol. 14, no. 16, p. 3307, 2025.

[71] F. G. Ortiz-Gomez, L. Lei, E. Lagunas, R. Martinez, D. Tarchi, J. Querol, M. A. Salas-Natera, and S. Chatzinotas, "Machine learning for radio resource management in multibeam geo satellite systems," *Electronics*, vol. 11, no. 7, p. 992, 2022.

[72] X. Hu, S. Liu, Y. Wang, L. Xu, Y. Zhang, C. Wang, and W. Wang, "Deep reinforcement learning-based beam hopping algorithm in multibeam satellite systems," *IET Communications*, vol. 13, no. 16, pp. 2485–2491, 2019.

[73] M. Á. Vázquez, P. Henarejos, I. Pappalardo, E. Grechi, J. Fort, J. C. Gil, and R. M. Lancellotti, "Machine learning for satellite communications operations," *IEEE Communications Magazine*, vol. 59, no. 2, pp. 22–27, 2021.

Appendix

LEO, MEO, and GEO Relevance to ML Use Cases

Table III summarizes how LEO, MEO, and GEO systems relate to representative ML use cases. The mapping reflects the typical operational domain where each orbit is most applicable based on latency constraints, mission objectives, and system design trends. LEO missions commonly support dynamic or data-intensive tasks such as anomaly detection and anti-jamming, While GEO systems favor high-throughput functions, such as beam hopping and traffic management. MEO missions occupy an intermediate role, balancing coverage and resilience. Importantly, the relevance scores emphasize near-term applicability, which orbits are expected to host these ML functions within the coming generation of satellite systems, rather than long-term theoretical potential.

TABLE III: Satellite type (LEO, MEO, GEO) relevance to ML use cases (●= high relevance, ◐= medium relevance, ○= low or no relevance).

Use Case	**LEO**	**MEO**	**GEO**
Anti-Jamming	●	◐	◐
Beam Hopping Optimization	◐	○	●
Command and Control Optimization	●	○	◐
Cybersecurity and Intrusion Detection	●	◐	●
Mission Planning (Autonomy, SSA)	●	◐	○
Network Traffic Management	●	◐	●
Object Detection and Earth Observation	●	○	○
Telemetry Analysis (Anomaly, Health)	●	●	●