18th International Conference on Space Operations, Montreal, Canada, 26 - 30 May 2025.

Detecting and Classifying Spacecraft Anomalies into Operational Categories Using Artificial Intelligence: From New to Recurrent Anomalies

Weicheng Qian, Joshua DeJong, Francis Senay, Bryan Cooke*

Calian Advanced Technologies, Saskatoon, Saskatchewan, Canada, winchell.qian@calian.com
* Corresponding Author

Abstract

Efficient spacecraft anomaly detection is essential for rapid fault detection, isolation and recovery (FDIR) and enhanced mission reliability. The increasing number of spacecraft in the coming years will challenge operational teams when screening high volumes of telemetry data, detecting anomalies by operational categories, and responding to anomalies according to standard operational protocols for isolation and recovery. Various deep learning models for time-series anomaly detection (TSAD) show promise in assisting operation teams to process high volumes of spacecraft telemetry data and detect anomalies among the growing number of spacecraft, each with its own advantages and disadvantages. On one hand, unsupervised TSAD models can detect the first instance of a new operational anomaly category (OAC) when labeled data for this signature is unavailable. On the other hand, unsupervised TSAD models may become less effective at detecting recurrent anomalies compared to supervised TSAD models, particularly when sufficient labeled data is available. When detecting an anomaly for the first time, no labeled data exists and the size of the accumulated labeled dataset grows at varying rates depending on the anomaly occurrence rate. Moreover, OACs can become obsolete or be mitigated due to special events such as permanent component failures or flight software upgrades. Integrating unsupervised and supervised machine learning models, considering the size of the labeled anomaly dataset, offers a promising approach to detect spacecraft anomalies and categorizing them from both the new and recurrent operational categories. We examined several strategies for integrating unsupervised and supervised machine learning models to detect and categorize anomalies throughout the lifecycle of operational categories. Our focus was particularly on the transition from unsupervised to supervised models, accounting for the variable growth of the labeled anomaly dataset, depending on the differing rates of occurrence and feature spaces of OACs. Our results suggest that integrating unsupervised and supervised models improves performance compared to using either model alone. The occurrence rates and feature space of OACs impact the optimal transition strategy, and ensemble learning can inform the development of this transition strategy. Our work has advanced the utility of machine learning models for practical failure detection applications, providing additional tools for operational teams to address real operational challenges.

Keywords: time-series, anomaly detection, operational anomaly category.

Acronyms/Abbreviations

AI&T	Assembly, Integration & Testing
ATOC	Advanced Technologies Operations Centre
CSA	Canadian Space Agency
FDIR	Fault Detection, Isolation and Recovery
ITM	Intelligent Telemetry Monitoring
LEOP	Launch and Early Operations Phase
ML	Machine Learning
OAC	Operational Anomaly Category
OvR	One vs. Rest
TSAD	Time-Series Anomaly Detection
UI	User Interface

1. Introduction

Reliable fault detection, isolation, and recovery (FDIR) mechanisms are foundational to spacecraft autonomy and mission success. As spacecraft fleets expand and telemetry data volumes increase, manual anomaly detection and classification become operational bottlenecks. The inherent complexity of telemetry data—coupled with the dynamic nature of spacecraft systems—necessitates scalable, data-driven anomaly detection solutions capable of generalizing across diverse operational regimes and failure modes [1–3].

Time-series anomaly detection (TSAD) methods based on deep learning offer powerful capabilities for analyzing multivariate telemetry streams [4, 5]. Unsupervised TSAD models—such as recurrent neural networks [6, 7], autoencoders [4], variational recurrent networks [8], and transformer-based architectures [1, 9, 10]—are particularly effective for identifying anomalies without labeled examples, making them well-suited for detecting novel or first-time anomalies. In contrast, supervised TSAD models—including deep classifiers [1, 3], sequence-to-sequence predictors [8], and transformer-based models trained on labeled sequences [1, 9]—typically achieve higher detection precision and interpretability when sufficient labeled data exists for known anomaly types.

Under the FDIR scenario, the goal of anomaly detection is to enable prompt and effective responses. For operational efficiency, anomalies are grouped into Operational Anomaly Categories (OACs) based on their root cause, operational impact (e.g., loss of functionality or degraded performance), severity, and the type of recovery actions required. In this context, the traditional anomaly taxonomy—comprising point, contextual, and collective anomalies—does not adequately capture the operational relevance of anomalies. Instead, TSAD models should aim to detect and classify anomalies in a manner that aligns with OACs, enabling more actionable, mission-relevant decisions and potentially triggering agentic workflows across multiple operational divisions. These workflows may involve coordination with payload operations, reconfiguration of mission objectives, or escalation to human-in-the-loop review, depending on the nature and impact of the anomaly.

Unsupervised TSAD models are well-suited for detecting novel or first-time anomalies without requiring labeled examples. They rely on learning effective representations of the hidden system state from observable multivariate time-series data, along with an appropriate discrimination criterion to distinguish between normal and anomalous system behaviors. Since these models are trained without labeled examples—where information of OAC classification would typically be embedded—it is less straightforward to expect them to be aware of or aligned with OACs. Nevertheless, enabling unsupervised models to detect anomalies associated with specific OACs could be highly valuable in practice. For example, such capability would allow operational teams to prioritize the detection of novel anomalies within critical OACs or to flag emerging variations within an existing OAC that may indicate a shift in underlying system behavior—due to non-uniform occurrence rates, differences in mission configurations, or system changes such as flight software upgrades and permanent component failures. These are often scenarios that traditional supervised models may fail to capture due to their dependence on historical labeled data.

Beyond the limitations of labeled data in capturing the evolving patterns of anomalies over time, operational practices introduce an additional challenge: dynamic labels. In reality, OACs are not static—particularly during the early stages of an anomaly's lifecycle, when only a few occurrences have been observed. Anomalies may be reclassified over time due to changes in operational responsibilities, evolving recovery procedures, or improved understanding of their root causes. This dynamic nature of OACs adds further complexity to the training and deployment of the detect-then-multi-class-classify learning paradigm, which typically assumes stable and well-defined labels. As OAC definitions shift—especially in the early phases—models trained under static labeling assumptions may misclassify anomalies, fail to detect important emerging patterns, or become costly to maintain. In practice, even minor changes to OAC definitions often necessitate retraining the entire multi-class classification model, making it difficult to ensure continued alignment with evolving operational knowledge.

As discussed above, the requirements and challenges of the TSAD problem in the context of FDIR applications can be summarized into three key aspects:

1. Coverage of the Anomaly Lifecycle—TSAD systems must handle the full lifecycle of anomalies—from first-time occurrences with no labels, to recurring anomalies with sufficient labeled data, and potentially obsolete or evolving anomaly patterns. This requires a combination of unsupervised detection for novelty and supervised classification for known patterns.
2. Classification Aligned with OACs—Effective anomaly detection must go beyond binary detection to support classification that aligns with OACs. Such alignment is essential for triggering operationally meaningful responses, including coordination across mission teams and subsystems.

3. Awareness of Dynamic and Evolving Labels—OACs are not static; they evolve as operational understanding improves, responsibilities shift, or recovery protocols change. TSAD models must be designed to handle dynamic labeling without requiring costly retraining or risking misalignment with current operational knowledge.

A straightforward approach to addressing these challenges is to adopt ensemble learning strategies that combine unsupervised and supervised models to cover the full anomaly lifecycle. One practical method is a one-vs-rest (OvR) architecture, in which a separate classifier is trained for each OAC. Compared to a single multi-class classification model, this approach may require more models, but its modular structure allows for easier updates, selective model deployment for specific anomaly types, and greater flexibility in dynamic operational environments. Moreover, the relative independence of each classifier enables embarrassingly parallel execution in distributed settings. With this OvR approach, supervised models can be retrained or updated at relatively lower cost, mitigating the challenges introduced by dynamic and evolving OAC labels.

Given this setup, the remaining challenges lie in (1) enabling unsupervised models to detect anomalies associated with specific OACs, and (2) designing an ensemble learning strategy that effectively integrates unsupervised and supervised models over time. We assume a lifecycle-driven dynamic, where unsupervised models initially outperform supervised models in detecting new or rare anomalies, but supervised models gain predictive strength as labeled data accumulates for a given OAC—until the next label redefinition or system change resets the cycle.

In this manuscript, we focus on developing and evaluating strategies for integrating unsupervised and supervised TSAD models to support OAC-aligned anomaly detection throughout the anomaly lifecycle. In addition to designing an ensemble learning framework, we investigate a method to guide unsupervised models toward detecting anomalies associated with selected OACs, thereby assisting supervised models during the early stages of the anomaly lifecycle when labeled data is limited or unavailable. Our work aims to inform the development of adaptive, operationally relevant TSAD systems that can respond effectively to the challenges of evolving anomaly patterns and dynamic constellation operations. Deployed within Calian Advanced Technologies' Operations Centre (ATOC) system, constellation operators are able to ensure rapid, automated and intelligent operations properly manage their fleet, ensuring constant availability to the users.

2. Methods

In this section, we present the methods developed to address the key challenges of TSAD for spacecraft operations, particularly the need to support the full anomaly lifecycle, align detection with Operational Anomaly Categories (OACs), and handle dynamic labels. Our approach consists of two main components: (1) guiding unsupervised models to detect anomalies aligned with selected OACs by refining the input time series streams, and (2) designing and evaluating ensemble learning strategies that integrate unsupervised and supervised models over time. For both components, our proposed approaches do not impose restrictions on the specific architectures of the unsupervised or supervised models used. Therefore, we expect our findings to be generalizable across a wide range of TSAD systems, beyond the specific examples demonstrated in this study.

2.1 Targeted Unsupervised Anomaly Detection via Time-Series Stream Refinement

Unsupervised learning approaches do not rely on information embedded in labeled examples to learn or align with OACs. Instead, their anomaly detection is based on learning to predict expected time-series behavior and identifying dissimilarities between incoming data and patterns learned from the training dataset. This can be formally expressed as the following contrastive meta-learning problem.

$$\underset{f}{\operatorname{argmin}}\ \mathbb{E}_{(A,p,n)\sim\mathcal{D}}\left[S\left(f(p\mid A),\ f(n\mid A)\right)\right] \tag{1}$$

where $A \sim D$ represents the training set of anchors sampled from the underlying distribution D, and $p, n \sim D$ denote positive and negative samples, respectively. Anchors serve as reference points that represents the baseline state of the system, which is typically assumed to be healthy or free of the specific type of anomalies we want to train the unsupervised model to detect, in contrast to the anomalous samples n. The positive samples p are expected to share characteristics with the anchors and reflect normal system behavior in the contrastive setup. $S(\cdot)$ is a general contrastive loss function that encourages the model to distinguish positives from negatives with respect to the anchors

[11, 12]. Such a loss function is designed to assign lower loss when the model $f(\cdot)$ correctly associates positives with the anchor context and separates negatives. For specific models, the contrastive loss $S(\cdot)$ can be either margin-based (e.g., triple loss), probabilistic/softmax-based (e.g., multiple negative ranking loss), or simply distance-based (including non-parametric dynamic thresholding error function employed by TelemAnom).

A common source of confusion in this contrastive setup is the assumption that the anchor set A must be completely free of anomalies. In practice, this is often unnecessary and overly restrictive. What matters is that A does not contain the specific types of anomalies the model is intended to detect. Since the learning objective encourages the model to distinguish between positives (which should resemble A) and negatives (which should differ from it), it is sufficient for A to reflect the normalcy baseline relevant to the detection task. For example, if the goal is to detect a specific failure mode, A need only exclude instances of that failure mode—other unrelated variations or mild deviations may still be present without harming model performance.

This perspective naturally motivates a broader strategy: input-space augmentation or projection as a means of shaping the anomaly categories an unsupervised model is attuned to. Specifically, by selecting subdimensions of the multivariate distribution D or by projecting data into a meaningful embedding manifold (e.g., via domain-informed encoders or representation learning), we can steer the model's focus toward particular types of deviations. This allows practitioners to maneuver the alignment between the model's contrastive learning objective and the semantic structure of the anomalies of interest. In other words, even when working with unlabeled data, the choice of representation—whether raw, selected, or learned—plays a critical role in determining which anomalies are detectable, and which are ignored as variation within the normal class. Formally, it can be represented as the following:

$$\underset{f;\,\phi}{\operatorname{argmin}}\ \mathbb{E}_{(A,p,n)\sim\phi(\mathcal{D})}\left[S\left(f(p \mid A),\ f(n \mid A)\right)\right] \tag{2}$$

where we denote ϕ as a hyperparameter that determines the data representation (e.g., selected subdimensions or projected embedding space). In practice, such data representation can take the form of selecting specific time-series variables over particular time spans, providing flexibility to enable the anchor set to define both the baseline behavior and the scope of detection. This enables the unsupervised learning model to focus on identifying specific types of anomalies, without requiring the anchor set to be entirely free of all anomalies. Instead, it is sufficient that the anchor set excludes the types of anomalies the model is intended to detect, such as new anomalies or first occurrences that deviate from the established baseline.

2.2 Evaluating Ensemble Learning Approaches to Combine Unsupervised and Supervised Model

There are numerous strategies for combining unsupervised and supervised models, ranging from simple rule-based switching to weighted voting, stacking, and meta-learning. In this work, we focus specifically on reinforcement learning (RL) as a means of learning an adaptive ensemble policy. This choice is motivated by the sequential and evolving nature of anomaly detection across the anomaly lifecycle, where model performance shifts over time and decisions must be made dynamically based on limited feedback. RL provides a principled framework for modeling this as a sequential decision-making problem, enabling the system to learn when to rely on unsupervised models, when to transition to supervised ones, and how to balance both during uncertain phases. In the following subsections, we first formalize the anomaly lifecycle as it manifests in spacecraft operations (Section 2.2.1) and then present our evaluation of typical RL approaches for dynamic TSAD model integration (Section 2.2.2).

2.2.1 Anomaly Lifecycles in Operation

Figure 1 illustrates the lifecycle of an Operational Anomaly Category (OAC), the corresponding availability of information, and our assumptions on the advantageous learning approaches at different stages given their characteristics. The lifecycle is divided into three main phases along a temporal axis:

- First Identification ("Day Zero"): At this initial stage, a new anomaly is observed for the first time. No labeled data is available, and the anomaly has not yet been classified into an existing OAC. Here, unsupervised learning is critical for detecting unexpected behavior without relying on the labeled examples.
- Establishing Protocol: As additional instances of the anomaly are observed, a small set of labeled data begins to accumulate. This phase involves assigning the anomaly to an Operational Anomaly Category (OAC), developing a recovery protocol, and refining the classification criteria. The amount of labeled data available depends on the anomaly's rate of occurrence. As more and more labeled examples are gathered, the performance of supervised models is expected to improve. However, the learning curve can vary significantly depending on the complexity and variability of each specific OAC. During this transitional

phase, ensemble learning—which combines the generalization capabilities of unsupervised models with the precision of early-stage supervised models—is especially valuable in bridging the gap between the two extremes.

- Recurrence of Known Anomalies: Once sufficient labeled data has been gathered, supervised learning becomes the preferred approach, due to its ability to deliver high classification accuracy and generally lower inference overhead in stable, well-labeled settings. At this stage, models can reliably classify recurring anomalies associated with established OACs—until the OAC is redefined or new anomaly patterns emerge, effectively restarting the cycle.

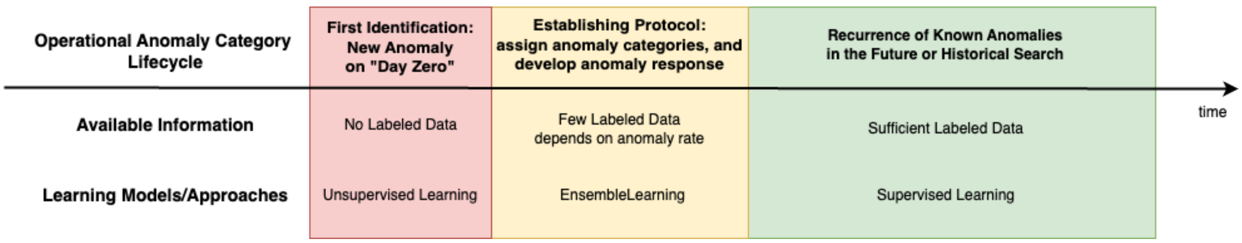

Fig. 1 Operational Anomaly Category Lifecycle

For dynamic labeling—where OAC definitions evolve over time—this process can be viewed as the obsolescence of an existing OAC and the initiation of a newly defined one. Such changes may arise from improved understanding of root causes, shifts in operational responsibilities, or updates to flight configurations, onboard software, or recovery strategies. In summary, the OAC lifecycle reinforces the need for an ensemble learning strategy that can flexibly combine model outputs as labeled data accumulates over time. In particular, reinforcement learning (RL) offers a natural framework for this setting, as it enables learning an optimal policy for dynamically selecting or weighting models in response to the evolving number of anomaly occurrences and system feedback—framing the challenge as a sequential decision-making problem with discrete choices.

2.2.2 Reinforcement Learning for Dynamic TSAD Model Integration

Ensemble learning with reinforcement learning (RL) offers a flexible and adaptive framework for combining multiple models to improve robustness and performance over time. In this work, we investigate RL due to its widespread use in sequential decision-making problems, which aligns well with the TSAD context—where, at each time step, the system must decide how to combine outputs from both models when processing new telemetry data. We focus on a simple yet representative scenario involving two models: one unsupervised and one supervised. This setting reflects the typical dynamics of anomaly lifecycles, where the utility of each model shifts over time. Although we demonstrate our approach in this two-model case, it is readily extensible to configurations involving multiple unsupervised and supervised models, allowing for broader applicability in more complex operational environments.

We further make the following assumptions to scope our discussion within a single anomaly lifecycle:

1. Unsupervised models outperform supervised models during the early stage, when no or few labeled examples are available.
2. The performance of supervised models is expected to improve as the number of labeled anomaly occurrences increases over time.
3. We restrict our analysis to a single anomaly lifecycle, treating any update to the OAC definition as the start of a new lifecycle. Therefore, we assume that no OAC redefinitions or label shifts occur within the lifecycle discussed in this subsection.

It is also straightforward to derive an oracle upper bound on the ensemble learning performance. This oracle serves as a theoretical benchmark, indicating the best possible performance achievable by an ideal ensemble strategy that always selects the better-performing model (unsupervised or supervised) at each point in time.

$$\hat{y}_t^{\text{oracle}} = \begin{cases} \hat{y}_t^{(0)} \vee \hat{y}_t^{(1)} & y_t = 1 \\ \hat{y}_t^{(0)} \wedge \hat{y}_t^{(1)} & y_t = 0 \end{cases}, \quad (3)$$

where $\widehat{y_t^{(0)}}$ and $\widehat{y_t^{(1)}}$ represent the binary predictions at time t from the unsupervised and supervised models, respectively, and y_t is the ground truth label indicating whether the input is anomalous ($y_t = 1$) or normal ($y_t = 0$). The oracle ensemble prediction $\widehat{y_t^{\text{oracle}}}$ is constructed such that if either model successfully detects an anomaly, the oracle will too, and if either model avoids a false alarm, the oracle prediction will also be correct. This ensures that the oracle achieves the best-case performance by capturing all true positives from both models while minimizing false positives, given perfect knowledge of the ground truth. As such, this oracle serves as a theoretical upper bound on the ensemble's performance—often unattainable in practice but useful as a benchmark to evaluate how closely practical strategies approach this ideal.

3. Experiment Setup

3.1 Unsupervised Learning with Topic-Driven Time-Series Stream Selection

We compared TelemAnom's performance on the NEOSSat dataset with and without applying the mapping function ϕ which refines the input representation to better align anomaly detection with the corresponding Operational Anomaly Categories (OACs). This mapping leverages domain knowledge of satellite operations and the structured nature of satellite telemetry data. For most OACs, we restrict the input time-series variables to those belonging to the telemetry packet associated with the OAC. In cases where a packet is associated with multiple OACs, we further narrow the input to a subset of variables based on operator insight, prioritizing those most relevant to the specific anomaly type. For temporal selection, we train the model using data from a randomly selected period known to be free of the target anomaly type, ensuring the unsupervised model is exposed only to nominal behavior during training.

3.2 Evaluating Typical RL Approaches to Combine Unsupervised and Supervised Model

We selected five representative reinforcement learning (RL) approaches to evaluate ensemble learning strategies that combine an unsupervised model and a supervised model. A summary of the configurations for each RL method is provided in Table 1. To ensure generality and repeatability, we developed a synthetic data generator that simulates the ground truth system status at each time step, indicating whether the system is healthy (label = 0) or anomalous (label = 1). The system status is sampled from a binomial distribution with a configurable anomaly rate, which we set to 0.01 in our experiments.

For simplicity, we assume each model's prediction accuracy at each time step is independent and identically distributed (i.i.d.). The unsupervised model is fixed to have a probability $p = 0.6$ of producing a correct prediction (i.e., true positive or true negative). In contrast, the supervised model starts with lower accuracy but improves over time, following a schedule defined as:

$$p(t) = 0.4 + 0.5 \frac{\text{current\_timestamp}}{\text{total\_timestep}} \tag{4}$$

At each time step, we simulate the arrival of new telemetry data and require the agent to make a prediction by selecting between the unsupervised and supervised model outputs, using the corresponding RL-based ensemble strategy. After the prediction, the agent receives feedback on whether the system was truly anomalous or not before the next time step, allowing the RL policy to update its decision-making process at each step. To ensure robust evaluation, we repeated the experiment 2,000 times, and in each run, we simulated a sequence of 200 time steps, modeling a single anomaly lifecycle.

We evaluate the performance of each RL approach using two metrics: the 10-step moving average of the F1 score, which reflects short-term responsiveness, and the cumulative F1 score, which captures overall performance across the full anomaly lifecycle.

Table 1. Selected Typical RL Approaches for Evaluation

Method	Reward Signal	Policy	Update Rule
Epsilon-Greedy [13]	Raw F1	$\arg\max_a \quad Q(a)\ (w.p.\ 1-\epsilon)$	$Q(a) \leftarrow Q(a) + \alpha\big(r - Q(a)\big)$
Thompson Sampling [14, 15]	Binary (F1 > 0.5)	$\arg\max_a \quad \text{Beta}(\alpha_a, \beta_a)$	α_a++ if TP/TN else β_a++
UCB [16]	Raw F1	$Q(a) + \sqrt{2\log t\,/N(a)}$	$Q(a) \leftarrow Q(a) + \alpha\big(r - Q(a)\big)$
REINFORCE [17]	Raw F1 (weighted)	$\text{softmax}(w_a)$	$w \leftarrow w + \alpha r \nabla \log \pi\,(a)$
Contextual Bandit [18]	Raw F1	$\arg\max_a \quad \theta_a^\top x_t$	$\theta_a \leftarrow \theta_a + \alpha r x_t$

4. Results

4.1 Unsupervised Learning with Topic-Driven Time-Series Stream Selection

Table 2 summarizes the performance of the topic-driven (ϕ-mapped) unsupervised model in detecting anomalies corresponding to each OAC, compared to the baseline unsupervised model. The results clearly demonstrate that the topic-driven approach is effective in improving the model's ability to detect anomalies aligned with specific OACs. To ensure a fair comparison, we did not apply additional fine-tuning to the topic-driven model beyond selecting a subset of time-series variables. These variables were chosen based on the known correspondence between OACs and telemetry packets, leveraging domain knowledge to isolate the most relevant features for each category.

Table 2. Comparison of Baseline and Topic-Driven Unsupervised Learning for OACs

Experiment ID	Model	Precision	Recall	F1 Score
OAC-1	Baseline	0.0093	0.3407	0.0181
	Topic-Driven	0.3361	0.8889	0.4878
OAC-2	Baseline	0.0041	0.2326	0.0081
	Topic-Driven	0.0066	0.8125	0.0131
OAC-3	Baseline	0.0014	0.3980	0.0028
	Topic-Driven	0.0009	0.7143	0.0018
OAC-4	Baseline	0.0060	0.3889	0.0119
	Topic-Driven	0.0236	0.6111	0.0455

4.2 Evaluating Typical RL Approaches to Combine Unsupervised and Supervised Model

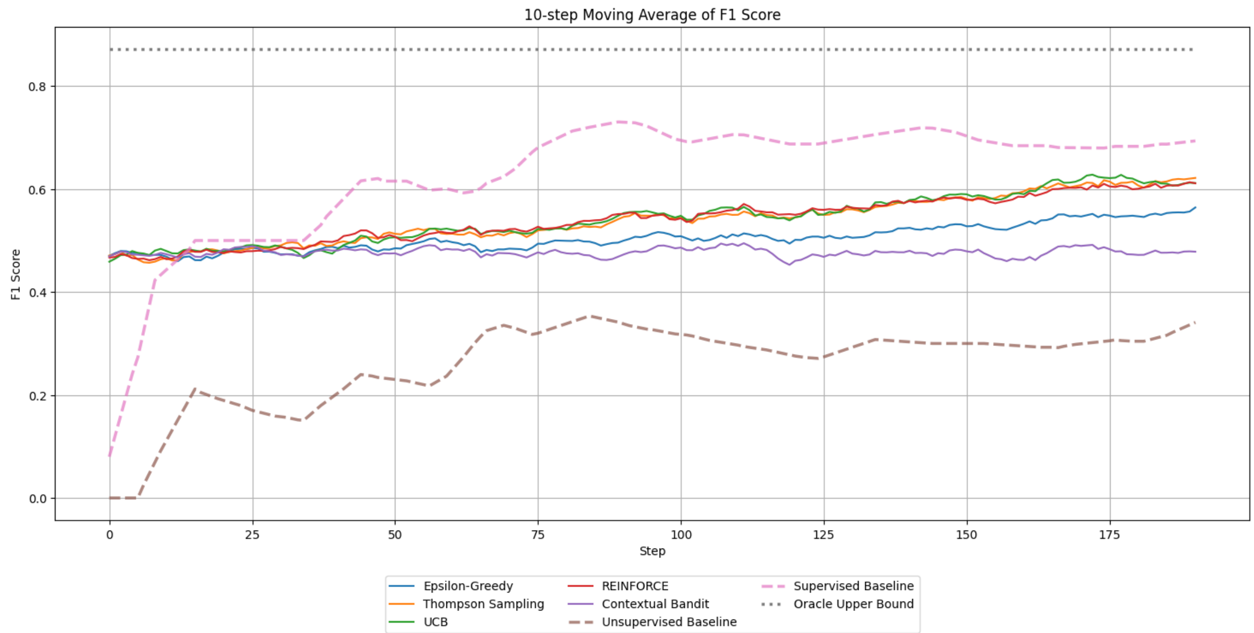

Fig. 2 10-Step Moving Average of F1 Score

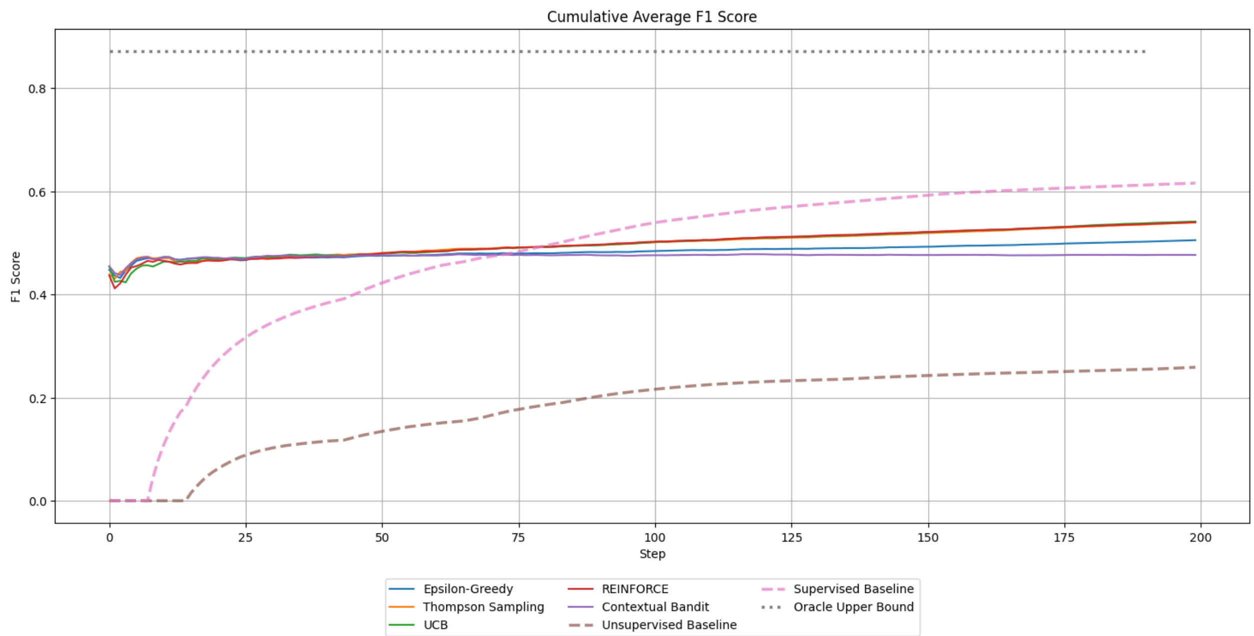

Fig. 3 Cumulative Average F1 Score

Figures 3 and 4 present the 10-step moving average and cumulative F1 scores, respectively, across 200 time steps for each reinforcement learning (RL) strategy, alongside unsupervised and supervised baselines, and the oracle upper bound. As expected, the oracle strategy consistently outperforms all approaches, serving as a theoretical upper limit. Among the RL-based ensemble methods, REINFORCE and Thompson Sampling achieve the highest performance, closely approaching the supervised baseline in the later stages while maintaining stronger early-stage performance compared to it. Notably, all RL strategies outperform the unsupervised baseline, which remains relatively flat over time due to its fixed accuracy and gradually catch up to or surpass the supervised baseline as more anomaly occurrences are observed and learned from.

The supervised baseline starts off weak but improves significantly over time, reflecting the assumed performance curve tied to increased labeled data. In contrast, the unsupervised model, while initially stronger, fails to adapt, leading to stagnant performance. The RL approaches demonstrate adaptability by dynamically balancing the early strength of the unsupervised model with the later-stage gains of the supervised model. The contextual bandit lags slightly behind other RL approaches, likely due to its limited memory of past outcomes, while epsilon-greedy and UCB show solid, consistent performance.

5. Discussion

While the RL-based ensemble strategies show strong performance across the anomaly lifecycle, one observation worth highlighting is that, after a sufficient number of time steps, the supervised model alone can outperform the ensemble strategy. This is expected, as the supervised model's accuracy improves with the accumulation of labeled data, eventually surpassing the benefit of dynamic switching. This suggests a potential optimization strategy: once the supervised model's performance consistently exceeds that of the ensemble, it may be more efficient to rely solely on the supervised model—unless there is evidence of label shifts, OAC redefinitions, or concept drift. Developing a mechanism to detect such changes and re-activate ensemble learning as needed could improve both performance and efficiency and is a promising direction for future work.

Another limitation of this study not yet explored are domain-specific customizations of RL algorithms. All RL approaches were applied using standard settings without tailoring to the unique structure of the TSAD task, such as the delayed but predictable improvement of the supervised model or the evolving class imbalance over time. Exploring reward shaping, temporal credit assignment, or hierarchical policies could further improve ensemble learning performance and better reflect operational priorities in real-world spacecraft anomaly detection systems.

6. Conclusion

In this work, we addressed the unique challenges of time-series anomaly detection (TSAD) in spacecraft operations, where anomalies follow dynamic lifecycles and are categorized into evolving Operational Anomaly Categories (OACs). We proposed a hybrid framework that integrates unsupervised and supervised models, and we explored reinforcement learning (RL) as a principled method for dynamically selecting between them over time. Our experiments demonstrated that RL-based ensemble strategies can effectively adapt to shifting model performance across the anomaly and constellation lifecycle, approaching oracle-level performance under idealized conditions. We also introduced a data refinement strategy to align unsupervised detection with OACs, improving the operational relevance of early-stage anomaly identification. Future work will explore extending our approach to multi-model ensembles, applying it to real telemetry datasets with partial or noisy labels, and improving robustness to dynamic label shifts and uncertainty in anomaly definitions.

Acknowledgements

We thank Hanwen Zheng, Francis Senay, Ayman Shahriar, and Jarin Tasnim, for their contribution to the project. We also appreciate Stephan Nappert's support with reviewing, as well as insightful advice on formatting and editing. Additionally, we thank Aimee Kraushar for help with formatting. This project was undertaken with the financial support of the Canadian Space Agency.

References

[1] Daniel Lakey and Tim Schlippe. "A Comparison of Deep Learning Architectures for Spacecraft Anomaly Detection". In: 2024 IEEE Aerospace Conference (2024), pp. 1–11.
[2] Muhamed Abdulhadi Obied et al. "Deep Clustering-Based Anomaly Detection and Health Monitoring for Satellite Telemetry". In: Big Data Cogn. Comput. 7 (2023), p. 39.
[3] Sara Cuéllar et al. "Explainable anomaly detection in spacecraft telemetry". In: Eng. Appl. Artif. Intell. 133 (2024), p. 108083.
[4] Pengfei Guo, Caisheng Wei, and Zeyang Yin. "Anomaly Detection of Spacecraft Reconstructed Signals Based on Attention Mechanism". In: 2023 6th International Symposium on Autonomous Systems (ISAS) (2023), pp. 1–5.
[5] L. Yang et al. "Improved deep learning based telemetry data anomaly detection to enhance spacecraft operation reliability". In: Microelectronics Reliability (2021).
[6] Kyle Hundman et al. "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.

[7] Jakub Nalepa et al. “ANTELOPE: TOWARDS ON-BOARD ANOMALY DETECTION IN TELEMETRY DATA USING DEEP LEARNING”. In: 2021.
[8] Bing Yu et al. “A novel end-to-end anomaly detection framework for spacecraft using MINE and LSTM-VAE with attention mechanism”. In: Measurement Science and Technology (2024).
[9] Ankit Srivastava, Neeraj Badal, and BS Manoj. “SADDLE: Spacecraft Anomaly Detection using Deep Learning”. In: 2024 IEEE Space, Aerospace and Defence Conference (SPACE) (2024), pp. 128–131.
[10] Jiehui Xu et al. “Anomaly transformer: Time series anomaly detection with association discrepancy”. In: arXiv preprint arXiv:2110.02642 (2021).
[11] Matthew Henderson et al. “Efficient Natural Language Response Suggestion for Smart Reply”. In: (May 2017). eprint: 1705.00652.
[12] Raia Hadsell, Sumit Chopra, and Yann LeCun. “Dimensionality reduction by learning an invariant mapping”. In: 2006 IEEE computer society conference on computer vision and pattern recognition (CVPR’06). Vol. 2. IEEE. 2006, pp. 1735–1742.
[13] Richard S. Sutton and Andrew G. Barto. Reinforcement Learning: An Introduction. MIT Press, 1998.
[14] William R. Thompson. “On the Likelihood that One Unknown Probability Exceeds Another in View of the Evidence of Two Samples”. In: Biometrika 25.3/4 (1933), pp. 285–294.
[15] Daniel Russo et al. “A Tutorial on Thompson Sampling”. In: Foundations and Trends in Machine Learning 11.1 (2018), pp. 1–96.
[16] Peter Auer, Nicol`o Cesa-Bianchi, and Paul Fischer. “Finite-time Analysis of the Multiarmed Bandit Problem”. In: Machine Learning 47.2-3 (2002), pp. 235–256.
[17] Ronald J. Williams. “Simple Statistical Gradient-Following Algorithms for Connectionist Reinforcement Learning”. In: Machine Learning 8.3 (1992), pp. 229–256.
[18] John Langford and Tong Zhang. “The Epoch-Greedy Algorithm for Multi-armed Bandits with Side Information”. In: Advances in Neural Information Processing Systems (NeurIPS). 2008.