ENHANCING SPACE OPERATIONS WITH UNSUPERVISED ANOMALY DETECTION:

THE PITIA SYSTEM

Index

Introduction

Introduction: The problem.

Why anomaly detection matters in spacecraft missions?

Spacecraft missions rely on many onboard subsystems — from thermal control to power management — that continuously generate **massive volumes of telemetry data**.

Ensuring the correct operation of these systems is critical for mission success, but even a small undetected anomaly can lead to severe consequences.

Traditionally, anomaly detection relies on simple approaches such as predefined nominal ranges or manual thresholds. These methods quickly become insufficient as missions grow in complexity and the number of sensors increases — telemetry consists of **hundreds of interconnected time series**, and anomalies are often subtle or entirely new.

This growing complexity calls for more advanced, **data-driven solutions**. Artificial intelligence can automate anomaly detection, enabling faster, more accurate, and cost-efficient operations.

The real challenge: Scaling anomaly detection to mission-level data

Anomaly detection in spacecraft telemetry is not just about finding unexpected values — it’s about doing so reliably, automatically, and at scale.

Every month, a single mission can generate:

- ~900,000 telemetry records
- ~20,000 parameters across multiple subsystems
- ~72 GB of raw data to process and monitor

Traditional rule-based approaches cannot handle this level of complexity.

Manual inspection is unfeasible, anomalies may be subtle or unseen before, and static thresholds fail to capture evolving system behaviors.

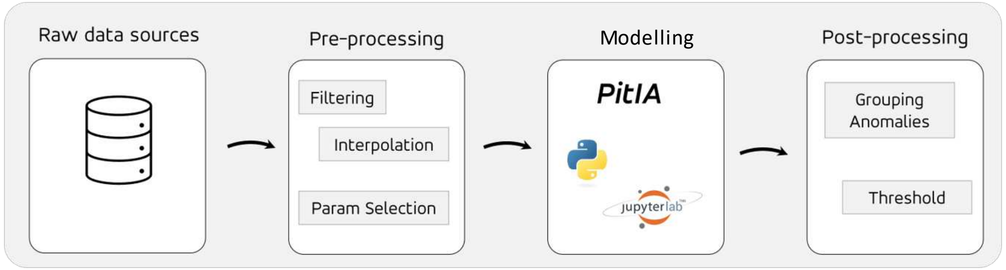
This is the challenge:
How can we efficiently detect anomalies in such massive, multivariate data streams — and do it continuously and autonomously to support mission-critical operations?

PitIA: Our Solution

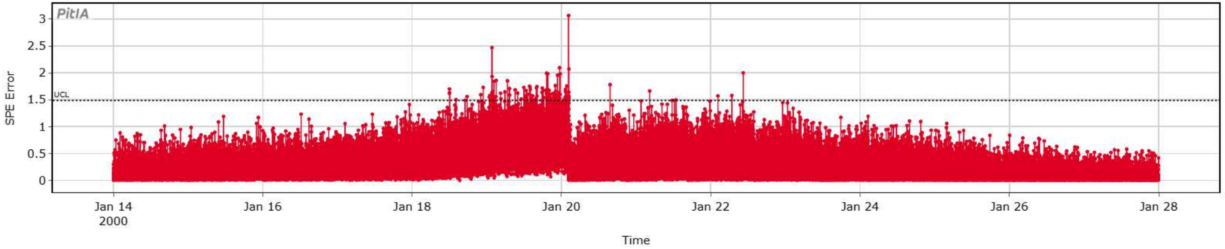
PitIA is an unsupervised anomaly detection system designed specifically for spacecraft telemetry.

Purpose:
Detect anomalous periods in multivariate time series and identify the most contributing channels — enabling faster diagnosis and decision-making.

Approach:
Based on Multivariate Statistical Process Control (MSPC) and Principal Component Analysis (PCA), adapted from industrial process monitoring to space operations.



PitIA turns raw telemetry into actionable anomaly insights — autonomously and in real time.



ESA dataset description

ESA Benchmark Dataset: A Foundation for Anomaly Detection

A new public benchmark
To advance anomaly detection in satellite telemetry, the European Space Agency released in 2024 the first open benchmark dataset with real, annotated telemetry data.

Real data from multiple missions
The dataset contains multivariate time series from **three ESA missions**, annotated by **domain experts** and cross-verified with state-of-the-art algorithms.

Designed for real operations
It simulates real mission conditions — with **rare, subtle, and multi-sensor anomalies** — and includes a **hierarchical evaluation pipeline** plus new metrics tailored to spacecraft operations.

Why it matters
This benchmark enables consistent evaluation and fair comparison of algorithms, providing the realistic data needed to train and assess PitIA.

	Mission1	Mission2	Both missions
Channels	**76**	**100**	**176**
Target / Non-target	58 / 18	47 / 53	105 / 71
Channel groups	18	29	47
Subsystems	4	5	6*
Telecommands	**698**	**123**	**821**
Priority 0/1/2/3	345 / 323 / 19 / 11	0 / 0 / 119 / 4	345 / 323 / 138 / 15
Total executions	1,594,722	1,918,002	3,512,724
Data points	**774,856,895**	**776,734,364**	**1,551,591,259**
Duration (anonymised)	14 years	3.5 years	17.5 years
Compressed size [GB]	3.51	3.81	7.32
Annotated points [%]	1.80	0.58	1.19
Annotated events	**200**	**644**	**844**
Anomalies	118	31	148
Rare nominal events	78	613	690
Communication gaps	4	0	4
Univariate / Multivariate	32 / 164	9 / 635	41 / 799
Global / Local	113 / 83	585 / 59	698 / 142
Point / Subsequence	12 / 184	0 / 644	12 / 828
Distinct event classes	**22**	**32****	**54**

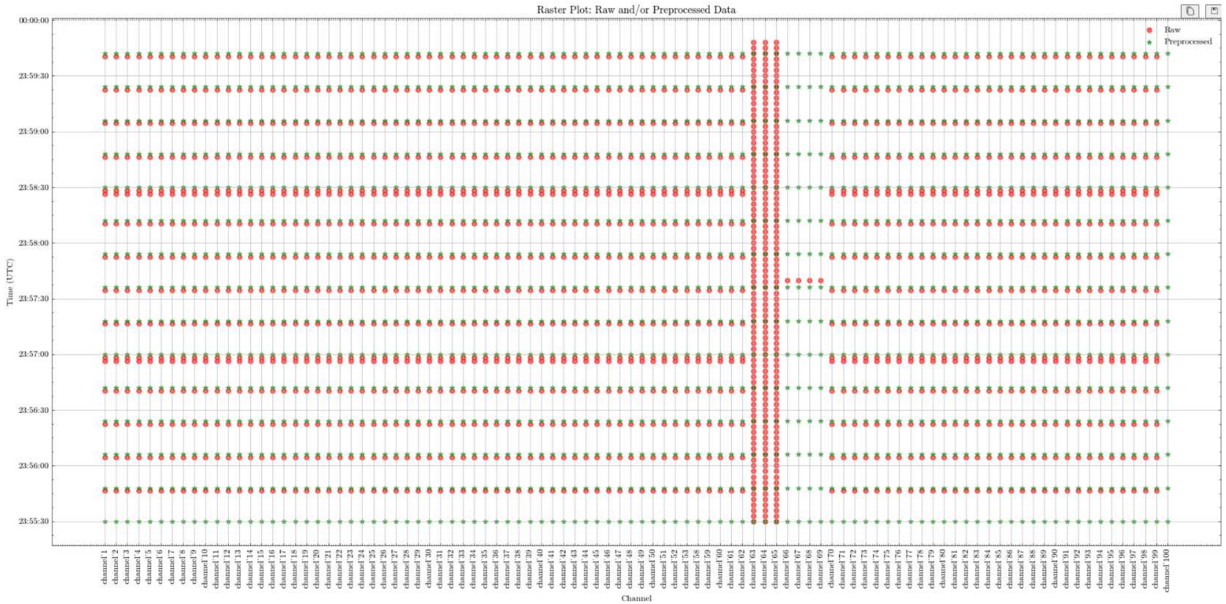
Data Preprocessing

Resampling and Synchronization

Before training, telemetry from different subsystems must be aligned and standardized.

- **Resampling:** unify data into a single time resolution to handle different sampling rates (e.g., 30s vs 18s).
- **Zero-order hold interpolation:** fill missing values by propagating the last known sample without using future information.

This ensures a clean, uniform dataset ready for anomaly detection.



Impact and Optimization

Implementing the zero-order hold preprocessing significantly reduces dataset size and memory needs:

- From ~15 million rows to ~1 million
- Over **50% memory reduction**, enabling longer time spans and faster processing

It also preserves anomalies during interpolation to ensure they are not lost in resampling.

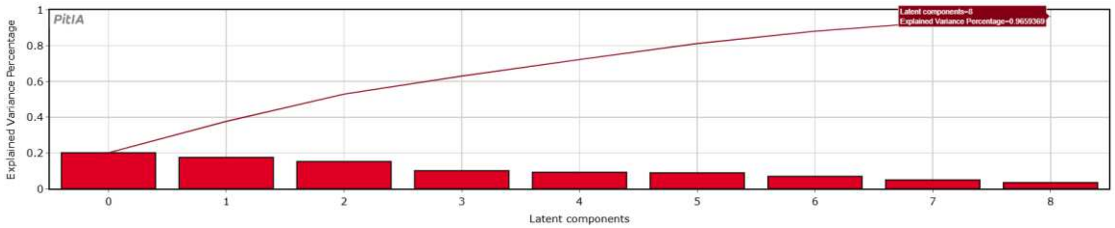
TIME RANGE	BEFORE	ZERO-ORDER
2000 – All channels	15 230 553	1 054 081
2000 – Target columns	14 743 973	

Training PitIA Model

Training PitIA: Unsupervised Anomaly Detection

PitIA learns what *normal* spacecraft telemetry looks like and detects deviations from this behavior — all **without labeled anomalies**.

- Uses **Multivariate Statistical Process Control (MSPC)** techniques such as **Principal Component Analysis (PCA)**.
- Builds a model of normal system behavior by capturing correlations among thousands of telemetry parameters.
- The number of principal components is **automatically selected** to capture **90% of the cumulative variance** in the data.
- New data points that lie far from this learned space are flagged as potential anomalies.

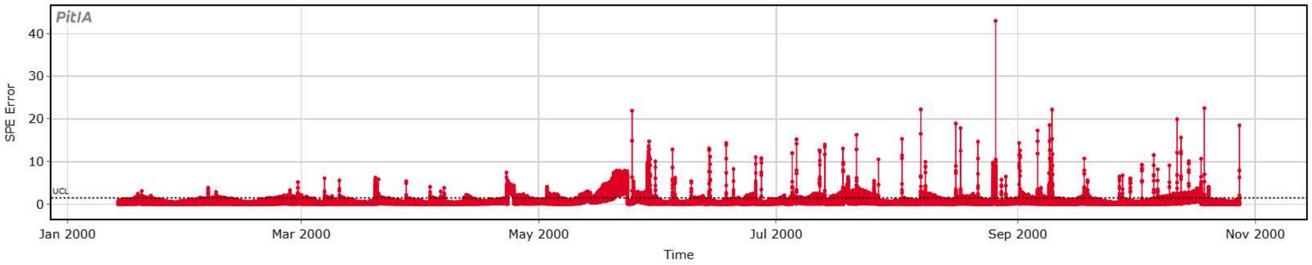


Statistical Indicators: SPE & Hotelling's T^2

PitIA uses two main statistical indicators to evaluate new observations:

- **Hotelling's T^2** – measures how far a new observation is from the center of the learned PCA space.
- **Squared Prediction Error (SPE)** – measures how much of the observation cannot be explained by the model.

To determine whether an observation is anomalous, a **statistical threshold (UCL)** is computed based on a **chi-square distribution** (typically at 95% confidence).
Observations with **SPE > UCL** or **T^2 > threshold** are flagged as **anomalies**.



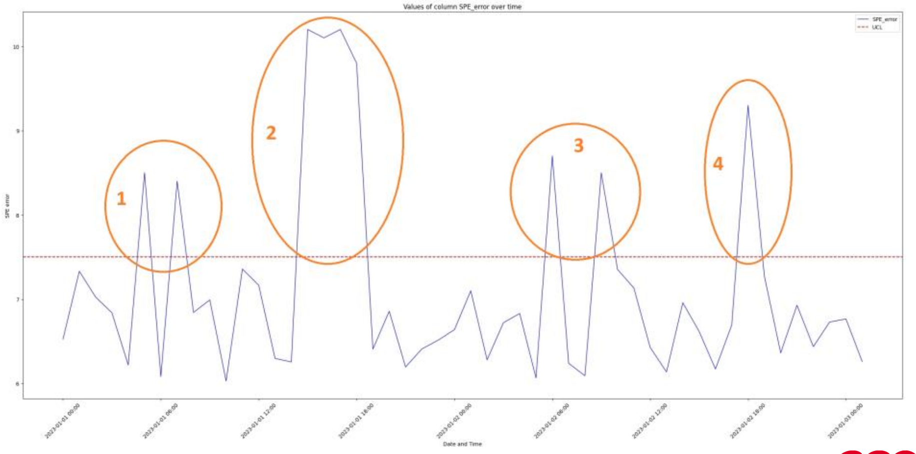
Postprocessing

Refining and Grouping Anomalies

After detection, PitIA applies a dedicated **postprocessing stage** to turn raw detections into meaningful alerts:

1. **Refinement** – we compute the gradient of the prediction error $\|\dot{SPE}\|$ to highlight subtle but significant changes over time.
2. **Grouping** – detections less than **6 hours apart** are merged into a single event, reducing redundant alarms and simplifying analysis.

Together, these postprocessing steps significantly improve detection reliability, reduce noise, and ensure anomaly alerts are **clear, meaningful, and actionable** for mission operations.



Evaluation

Evaluation. ESA Anomaly Detection Metrics

PitIA was evaluated using the official ESA metrics, which measure not only detection accuracy but also **operational relevance** — precision, timing, and usefulness of anomaly alerts.

Metric name	What it measures	Why it matters
Event Wise (F0.5)	Precision vs. recall at event level	Fewer false alarms and missed detections
Anomaly Detection Timing Quality Curve (ADTQC)	How close detection is to anomaly start	Early and actionable alerts
Context awareness	Correct subsystem/channel and duration	Operationally useful diagnostics

These metrics ensure PitIA is evaluated based on how well it supports **mission operations**, not just statistical performance.

Results

ESA Benchmark

PitIA was evaluated on the official ESA anomaly detection benchmark, demonstrating competitive performance across precision, timing, and contextual accuracy.

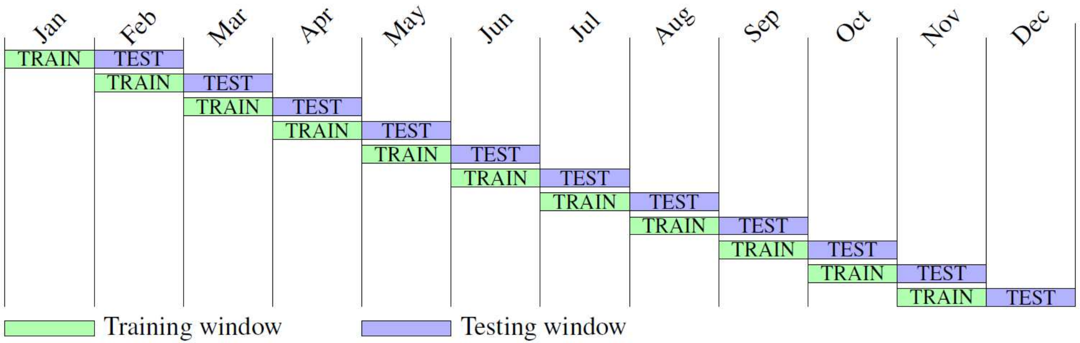
```
Mission 1

Channels:
•   Full set of channels (All)
•   Subset of channels (41 – 46)

Train: 01/01/2000 – 01/07/2007 (84 months)
Test:  01/10/2007 - 01/07/2009 (21 months)
```

```
Mission 2

Channels:
•   Full set of channels (All)
•   Subset of channels (18 – 28)

Train: 01/01/2000 – 01/07/2001 (18 months)
Test:  01/10/2001 - 01/07/2003 (21 months)
```

Configuration	PitIA	Best ESA Model	Second Best ESA Model
Mission 1 – Full channels	0.424	0.061 (Teleman-ESA P.)	0.008 (Teleman-ESA)
Mission 1 – Subset (ch. 41–46)	0.323	0.786 (Teleman-ESA P.)	0.253 (Global STD5)
Mission 2 – Full channels	0.760	0.241 (STD5)	0.100 (PCC)
Mission 2 – Subset (ch. 18–28)	0.794	0.949 (Window iForest)	0.842 (Teleman-ESA P.)

From Benchmark to Operations: Continuous Model Updating

PitIA transitions from static benchmark setups — which rely on years of historical data — to a **lightweight, continuous, and operational deployment**.
The model is trained using **only one month of telemetry** and is **updated monthly** with a *sliding window* strategy, ensuring it remains aligned with the most recent spacecraft conditions.



Configuration	PitIA
Mission 1 – Full channels	0.424
Mission 1 – Subset (ch. 41–46)	0.332
Mission 2 – Full channels	0.882
Mission 2 – Subset (ch. 18–28)	0.880

Conclusions

Conclusions

- Robust and scalable approach:
  - Valid in real scenarios with large volumes of data / TM.
  - Generic: valid for different missions, signal types, etc.
  - Unassisted: no action required from the operator
  - Low computational requirements
  - Adaptability to changes in the system thanks to automatic unassisted training and threshold calculation
  - Balanced results FP and FN
- Outstanding results for mission 2. Especially when training by subsystems.
- Promising and state-of-the-art results for mission 1. This mission seems to be more stable and refined, and its anomalies are more challenging.

Contact:

- GMV:
  - LinkedIn
- Miguel Tejedor:
  - E-mail: mtejedor@gmv.com
  - LinkedIn