

Integration of Voting-Based Statistical Ensembles and Rule Mining for Anomaly Detection in Microsatellite Power System

Rizki Permala¹, Imas Sukaesih Sitanggang², Hendra Rahmawan³, Wahyudi Hasbi⁴

^{1,2,3}School of Data Science, Mathematics, and Informatics, IPB University, Bogor 16680, Indonesia

^{1,4}Research Center for Satellite Technology, BRIN, Bogor 16350, Indonesia

Article Info

Article history:

Received Dec 1, 2025

Revised Mar 25, 2026

Accepted Apr 16, 2026

Keywords:

Telemetry
Satellite
Anomaly
Rule Mining
Voting-Based Statistical
Ensembles

ABSTRACT

This study aims to detect anomalies, types of anomalies, and identify attributes that contribute to the cause of anomalies in the LAPAN-A2 microsatellite power system. Anomaly detection is cumulative-based, where the first year's dataset is added to the datasets of subsequent years for eight years. Anomaly detection uses a voting-based statistical ensemble (VBSE) approach and a combination of Apriori-Close to find associations and reduce rules. The VBSE obtained an average recall value of 86.49 and an average F1-score of 70.49. The F1-score value of VBSE increased by 4.2-fold compared with IForest (16.95), by 12.6-fold compared with ECOD (5.57), and by 5.6-fold compared with LOF (12.65). VBSE showed an increase in performance as the dataset complexity increased (DS1 → DS8), whereas IForest, LOF, and ECOD tended to decrease. The combination of metrics minSupp. 0.02%, minConf. 0.9, and lift is proven to be effective in capturing rare, reliable anomalies and significant association relationships. A strong correlation was observed between batteries (VBatt1–VBatt3) and a causal relationship between U_UMPB and VBatt. The types of anomalies detected included single, contextual, and correlation anomalies.

Copyright © 2026 Institute of Advanced Engineering and Science.
All rights reserved.

Corresponding Author:

Rizki Permala,
Research Center for Satellite Technology (PRTS),
National Research and Innovation Agency (BRIN),
Rancabungur, Bogor 16310, West Java, Indonesia.
Email: rizki.permala@brin.go.id

1. INTRODUCTION

Satellites play a vital role in many applications, such as Earth monitoring [1], navigation systems [2], demonstration of new technologies [3], and scientific research [4]. The reliability of a satellite depends heavily on its various components, particularly on the power system consisting of solar panels, batteries, and fuses. Solar panels are the primary energy source for satellites, converting sunlight into electricity that is needed to operate various satellite subsystems [5], [6]. When exposed to sunlight, solar panels continuously produce electricity that is directly used by the satellite system and also charges the batteries. In addition, solar panels play an important role in controlling the satellite's internal temperature by reflecting some of the heat received from sunlight. Batteries on satellites store the energy generated by solar panels so that this energy can be utilised when the satellite does not receive sufficient sunlight [5], [7], for example, when the satellite is in Earth's shadow (eclipse). Batteries also help balance fluctuations in electricity demand by providing additional power when the electricity demand exceeds the direct supply from the solar panels. Fuses are protective components that prevent excessive current in the satellite's power distribution lines, especially in the bus and payload subsystems.

Anomaly detection is crucial for ensuring the proper functioning of the power system, supporting continuous and uninterrupted satellite operations. Power system failure or performance degradation can result in data loss, communication disruptions, and impact the satellite's operational capabilities, even leading to the loss of control over the satellite itself [5]. Research conducted [8] on 1,584 Earth-orbiting spacecraft successfully

launched between January 1990 and October 2008 revealed that more than 25% experienced electrical power subsystem (EPS) failures.

Alves *et al.* [9] defines an anomaly as a pattern that deviates from the expected behaviour, occurring infrequently and being closely related to normal conditions. Research [10] suggests that anomalies are often caused by the combined action of several variables rather than by a single variable. Anomalies share common characteristics, including being rare, unexpected, and significant. This research focuses on three types of anomalies [11], namely, (1) point anomalies — single data points that differ significantly from other data points; (2) contextual anomalies — anomalies that are specific to situations where anomalies in one dataset are not necessarily anomalies in other datasets; and (3) collective anomalies — a group of data points that occur together and do not show a common pattern. Individual data points with the same behaviour may not be anomalies; however, when they occur collectively, they are considered anomalies.

This study detects anomalies using a voting-based statistical ensemble (VBSE) approach and searches for associations between attributes detected as anomalies using rule mining. This statistical approach was chosen because it is easy to understand and implement [12], ignores data distribution assumptions [13], and is computationally minimal [14]. The effectiveness of the VBSE approach (recall and F1-score) was compared with that of the isolation forest (IForest), local outlier factor (LOF), and empirical cumulative distribution-based outlier detection (ECOD).

1. IForest [15] detects outliers by measuring how quickly a value can be isolated from the rest of the data using random decision trees. Outliers are isolated more quickly because they have shorter path lengths and are fewer in number. [16] used IForest for anomaly detection in satellite power systems and [17] for satellite battery temperature.
2. LOF [18] works by measuring the local density of a data point with the local density of its neighbours. For univariate data, the distance is calculated based on the absolute difference between values, and the local neighbourhood consists of points with similar values. A set of points, whose density (based on their k -nearest neighbours) is significantly lower than that of their neighbours, is identified as an outlier. Research [19] used LOF for fault detection on GNSS/INS satellites, and [20] on GEO-SAR satellites.
3. ECOD [21] is a statistically based method designed to identify outliers, assuming that outliers are typically located at the tail end of a particular data distribution. The further a value is located at the tail end of the distribution (either left or right), the higher is the potential for it to be an outlier.

Research on anomaly detection for microsatellites and CubeSats is increasing with the rise in small satellite missions in low-Earth orbit (LEO). Jin *et al.* [22] detected anomalies in magnetometer data from the Tianping-2B satellite (a polar LEO microsatellite) using a cluster-based anomaly detection method and achieved an AUC > 0.9083. Cravidão *et al.* [23] proposed phase-folded kernel density estimation (PFKDE) to detect battery voltage anomalies on the BugSat-1 satellite, achieving a recall of 90.5% but a low F1-score (3.15). Cespedes *et al.* [24] applied LDA and One-Class SVM to the solar panel system of three BIRDS-4 satellites (polar LEO CubeSats) and achieved a precision of 0.97 and an F1-score of 0.85.

Ensemble approaches have proven more robust than single methods for time-series and imbalanced satellite telemetry data. Aleshko *et al.* [25] combined AAD, Isolation Forest Ensemble (IFE), and OCSVM with a hard voting mechanism to detect battery current and temperature anomalies on the SiriusSat-1 (CubeSat LEO) satellite, achieving a precision of 0.76, a recall of 1.0, and an F1-score of 0.86. Aslan and Çelebi [26] used an ANN-CNN-LSTM ensemble on the OPS-SAT (CubeSat LEO polar) satellite signal data with statistical feature extraction (mean, variance, skewness) and signal characterisation parameters, achieving an area under the curve (AUC) of >0.95.

Anderson *et al.* [27] developed a complex ensemble consisting of autoencoder (AE), GMM, LSTM, PCMCi, and Kalman filter with XGBoost as a classifier to detect anomalies in several subsystems (heater voltage, battery current, voltage, temperature) on AFRL's LEO satellite, resulting in an accuracy of > 99.79% and a true positive rate of 99.88%. Petkovic *et al.* [28] used an average-based ensemble of k -means, Isolation Forest, and LSTM-AE to detect thermal power consumption patterns on Mars Express. Jin *et al.* [29] compared various types of autoencoders in an ensemble framework for the power subsystem of a navigation satellite, with a precision of > 0.883 and a recall of 0.86. Xia and Meng [30] applied stacking ensemble learning to fault detection in a polar LEO satellite attitude control system (SACS), achieving a precision of 0.93, a recall of 0.91, an F1-score of 0.91, and an accuracy of 0.91.

Although the aforementioned methods have demonstrated high performance, most have been validated using satellites with differing subsystem configurations and the types of anomalies detected. Fair comparisons across studies should be made because of differences in satellites, parameter types, mission durations, and evaluation methods. Nevertheless, the proposed method (VBSE) demonstrated F1-score improvements of 4.2×, 12.6×, and 5.6× compared to iForest, ECOD, and LOF on the same dataset. Furthermore, unlike most comparable studies, this study identified the anomaly types (single, contextual, and correlation) and integrated explainability through rule mining to identify the causal attributes of anomalies using lift and leverage metrics.

For anomaly detection and explainability tasks, rule mining has several advantages, such as (1) it is effective in identifying complex patterns in high-dimensional data [31], [32], (2) the resulting association rules are easy to understand and explain logically [31], [33], and (3) it is flexible in parameter tuning [33]. According to [34], the association approach is effective for detecting anomalies in datasets because of its ability to detect anomalies (faults) in various ways, such as monitoring changes in the confidence level and support value of related rules (support), so that it can increase the reliability of anomaly detection. Table 1 summarises several anomaly studies on satellite power systems using rule mining.

Table 1. Research on anomalies in the electric power system using rule mining

Reference	Rule mining approach	Dataset
[5]	Combining FP-Growth with Trend Sign Aggregation Approximation (TSAX) to detect abnormal data and avoid unwanted model updates.	Satellite power system telemetry data (private).
[35]	FP-Growth algorithm to find frequent itemsets from telemetry data and association rules with support values of 0.1 and a confidence of 0.95.	Spacecraft telemetry data.
[36]	A combination of FP-Growth with random forest to improve the efficiency and accuracy of abnormal discharge detection in batteries. A support value of 0.4 and confidence of 0.9.	Satellite battery telemetry simulation data were used.
[34]	Time-series pattern clustering and Association Rule Mining (ARM). Support above 0.01.	Satellite telemetry data (private).

However, the studies presented in Table 1 do not clearly describe the types of anomalies detected. In this study, anomaly detection uses an accumulation-based scenario, where the first year's dataset is supplemented with datasets from subsequent years, such as 2016 (DS1); 2016–2017 (DS2); ...; 2016–2023 (DS8). This accumulation-based approach aims to obtain a comprehensive picture of the evolution of anomalies over time by integrating data from previous years. For example, a pattern that appears in DS1 may disappear in DS4 if it is no longer distinct. Conversely, new patterns may emerge after several years of collecting additional data. This study aims to detect anomalies, identify the types of anomalies detected, and identify the attributes that contribute to anomalies in satellite power systems, particularly in LEO.

However, there is a dearth of studies using datasets from equatorially orbiting microsatellites, such as LAPAN-A2. The LAPAN-A2 satellite is the only active microsatellite in an equatorial low-earth orbit (LEO) [37]. The orbit of the LAPAN-A2 satellite is around the equatorial line (inclination $\sim 6^\circ$) and moves from west to east. Most other microsatellites have polar orbits that move from north to south and vice versa. The LAPAN-A2 satellite operates under harsh space environmental conditions, with a sunlight–eclipse cycle of 14 times per day or approximately ~ 5110 times per year. LEO satellites also experience significantly more eclipse phases, resulting in more frequent battery charging and discharging cycles than geostationary Earth orbit (GEO) satellites [8]. Each of these cycles causes the power subsystem to experience an abrupt transition between sunlight and eclipse illumination phases. This cycle creates a non-stationary periodic pattern of battery status, namely, a charging and discharging cycle. During the eclipse phase, the solar array does not receive sunlight; however, the battery must distribute power to all subsystems. In addition, the speed of the LEO orbit of the LAPAN-A2 satellite is very high (~ 7.5 km/s), thereby limiting the contact window between the satellite and the ground station to only a few minutes per pass. Unlike GEO satellites that have continuous downlink telemetry, telemetry data from LEO satellites are received in small, irregular batches, creating time lags that affect the continuity of anomaly detection and require methods that are robust to uneven sampling.

Although research on anomaly detection in satellite telemetry data has been extensive, there are still gaps that can be improved: (1) Most research related to anomaly detection uses GEO-orbiting satellite datasets [38], [39] or LEO polar [40], [41], [42] datasets. Consequently, equatorial LEO satellites exhibiting distinct eclipse cycle patterns, rapid solar panel degradation, and short contact durations are underrepresented; (2) Existing approaches generally handle anomaly detection in isolation, without providing a diagnostic layer that can identify which combination of telemetry parameters contributes to a detected anomaly. This interpretability is operationally crucial for satellite operators to perform root cause analysis of anomalies; (3) Most proposed methods employ deep learning architectures, such as long short-term memory (LSTM), autoencoder, and transformer, which are black-box in nature, computationally complex [43], difficult to interpret [44], and highly dependent on the amount and quality of data [44].

Motivated by these gaps, this study proposes integrating a voting-based statistical ensembles (VBSE) approach for anomaly detection and a combination of Apriori-Close for association search and rule reduction. The framework integration was validated using eight years of telemetry data from the LAPAN-A2 satellite, an equatorial LEO microsatellite, and with an expert-annotated outlier-labeled dataset.

The main contribution of this research is the voting-based statistical ensembles (VBSE), which combine five robust statistical detectors. These detectors are adapted for non-normally distributed (skewed) data and use a

majority voting mechanism. VBSE achieved better anomaly detection performance (average F1 score of 70.49) compared to IForest, ECOD, and LOF. The evaluation was conducted over eight years (DS1–DS8) to demonstrate the scalability and consistency of VBSE performance despite increasingly complex datasets. This long-term validation is rare in the satellite anomaly detection literature. Furthermore, the diagnostic layer, AprioriClose, identifies combinations of telemetry parameters that occur together when anomalies occur. The AprioriClose results indicate that battery voltage (VBatt1–VBatt3) and main power bus voltage (U_UMPB) are key indicators of power system anomalies in LAPAN-A2. These findings help satellite operators analyse actionable root causes.

2. RESEARCH METHOD

2.1. Research Stages

This research has six main stages: data collection, data preprocessing, outlier detection, binning, rule mining, and rule interpretation. (1) Data is stored in the database immediately after the satellite has completely crossed the four ground stations (outside the coverage area of the easternmost ground station–Biak). Before being stored in the database, the telemetry data in text format are transformed into tabular form following the database structure. Separate software is used to process LAPAN-A4 satellite telemetry data based on Java. (2) Data preprocessing includes cleaning the database for duplicate data and noise. Data duplication arises from the repeated storage of the same telemetry data packet. Missing or corrupted data is considered noise and is removed from the dataset. The causes of data noise include incorrect data types (strings entered into numeric attribute columns) and missing values (data not recorded for certain attributes). Telemetry data with nadir and sunlight conditions were selected for further analysis, considering the apparent daily motion of the sun. The preprocessing stage serves to reduce false positives from nominal deviations, namely temporary outliers originating from satellite subsystem activations, such as turning on/off certain buses or payloads, battery cycling, current spikes, and/or transitions into and out of eclipse conditions. (3) Outlier detection using VBSE and comparison with expert labels. A hard voting mechanism requires the agreement of three out of five detectors before marking an anomaly. Only true positive outliers are processed further. (4) Binning involves converting numerical data into categorical (interval) data as input for the rule mining algorithm. (5) Rule mining, which implements the Apriori-Close algorithm to find associations and reduce rules. (6) Interpretation of the rule set and the types of anomalies detected. Figure 1 shows the research stages.

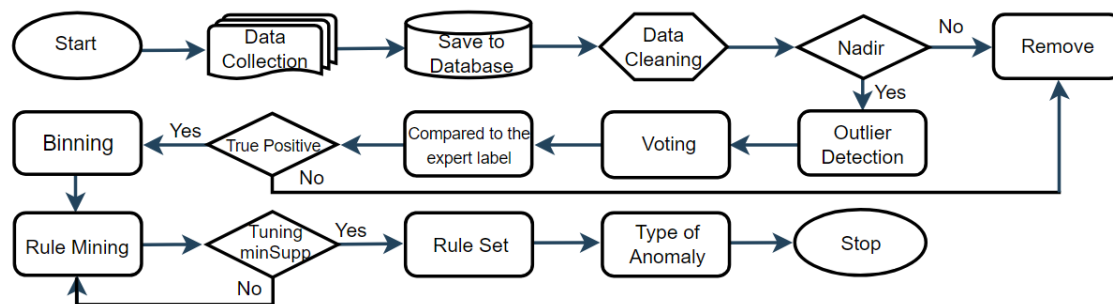


Figure 1. Research stages

2.2. Data Collection

The telemetry data used is not continuous because the LAPAN-A2 satellite passes over Indonesia every 103–105 minutes, with a maximum transit duration of 15 minutes, and passes over Indonesia 14 times per day. For technical and operational reasons, not all telemetry data are acquired during satellite transits. The ground stations used for LAPAN-A2 telemetry data acquisition include Kototabang (West Sumatra), Rancabungur/Bogor, (West Java), Parepare, (South Sulawesi), and Biak (Papua). Anomaly detection was conducted using a data-driven methodology utilising LAPAN-A2 satellite telemetry data from January 1, 2016, to December 31, 2023.

The dataset selected for further analysis included data with nadir and sunlight conditions, considering the apparent daily motion of the sun. The nadir is the satellite's position with the Z-axis facing the Earth and the X-axis pointing eastward from the direction of the satellite's approach. The nadir is determined based on the satellite's attitude, sunlight conditions, and the position of the sun as the satellite passes over the ground station. When the sun is north of the equator, the nadir is determined by the sun sensor current value on the Y-axis exceeding the sun sensor current value on the Y-axis. When the sun is south of the equator, the nadir conditions are determined based on the sun sensor current value on the plus Y (Y+) axis being greater than the sun sensor current value on the minus Y (Y-) axis. Sunlight conditions occur when the satellite is in an orbital area exposed to sunlight.

Off-nadir data were removed from the research data because they can cause normal data to be considered outliers. Satellite telemetry data during off-nadir manoeuvres, such as tumbling or attitude transitions, experience

rapid and natural changes in the power system and other sensor readings, which are operationally correct and not anomalies. Retaining such data without mode labelling increases false positives and systematically decreases detector precision [45]. For example, when a satellite is tumbling (an irregular rotational motion), the sun sensor current on the minus Z (Z-) axis, which should be facing the sky, fluctuates significantly and rapidly. This condition can be detected as an anomaly, which should be normal in tumbling conditions. Table 2 presents an example of the dataset attributes used.

Table 2. Example research dataset attributes

Attribute	Type Data	Unit	Description	Attribute Example
Time	Datetime	-	Data collection time	PCUTime
Current	Float	mili Ampere	Subsystem current value	I_Solar0, IBatt0, FusePCDH
Voltage	Float	Volt	Subsystem voltage value	V_Solar1, VBatt1, U_UMPB

2.3. Voting-Based Statistical Ensembles

VBSE is used to estimate whether an attribute value in a dataset is an outlier or not. A value was classified as an outlier when three of the five statistical approaches agreed. Table 3 presents the threshold formulas for the five statistical approaches used, and Figure 2 illustrates the statistical ensemble based on the voting.

Table 3. The upper and lower limits of the statistical approaches used

Method	Lower Limit	Upper Limit
Hampel Filter [46]	$\text{median} - (3 \times \text{MAD})$	$\text{median} + (3 \times \text{MAD})$
Modified Z-score [47]	-3.5	3.5
IQR [48]	$Q1 - 3 \times \text{IQR}$	$Q3 + 3 \times \text{IQR}$
Adjusted Boxplot for Skewed Distributions (ABSD) [49]	For $MC \geq 0$; $Q1 - 1.5e^{-4MC} \times \text{IQR}$ For $MC < 0$; $Q1 - 1.5e^{-3MC} \times \text{IQR}$	For $MC \geq 0$; $Q3 + 1.5e^{3MC} \times \text{IQR}$ For $MC < 0$; $Q3 + 1.5e^{4MC} \times \text{IQR}$
Z-score [50]	-2.5	2.5

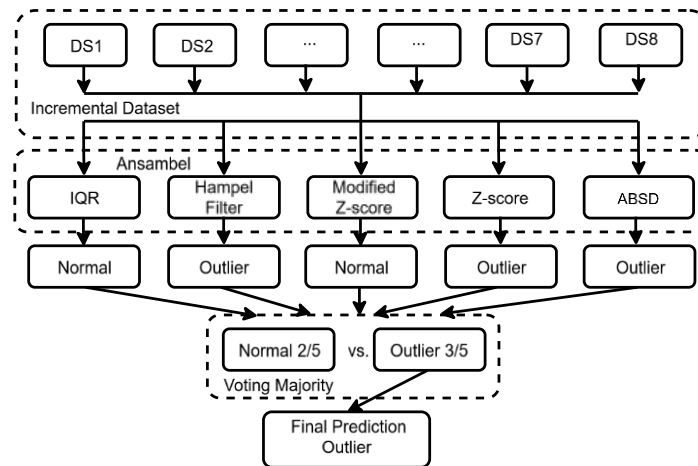


Figure 2. Illustration of a statistical ensemble based on voting

The Hampel filter uses the median absolute deviation (MAD) in a sliding window to identify outliers. The MAD is more reliable because it is less susceptible to extreme values [51]. Data points that deviate more than three times the MAD from the local median are classified as outliers [52]. The modified Z-score marks a value as an outlier if it exceeds a threshold of ± 3.5 MAD from the median [53]. The modified Z-score is insensitive to extreme values, thus helping to reduce false alarms and improve the reliability of error detection [54].

The interquartile range (IQR) categorises a value as an outlier if it falls outside the lower ($Q1 - 1.5 \times \text{IQR}$) or upper ($Q3 + 1.5 \times \text{IQR}$) limits [55]. This method is robust to extreme values [56] and data with high skewness [49]. The adjusted boxplot for skewed distributions (ABSD) is a modification of the standard boxplot that uses the robust medcouple estimator to adjust the IQR range for data with high skewness [49]. The Z-score identifies anomalies when the value exceeds a predetermined threshold, that is, less than -3 or greater than 3 [47], [54].

The results of anomaly detection using the VBSE approach were evaluated using the recall and F1-score. Recall measures the proportion of outliers correctly identified by the model [57], while the F1-score balances high outlier predictions (high recall) by minimizing false alarms (high precision) [57], [58]. The F1-score is an

important parameter for outlier detection in imbalanced datasets, where there are more normal data than outliers [59]. The formulas for the F1-score and recall are given in Equations 1 and 2, respectively [57].

$$\text{F1 Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (1)$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (2)$$

The statistical approach used only calculates the outlier score for each attribute (univariate) without considering its relationship with the other attributes. Attributes classified as true-positive outliers were selected for further analysis using the rule mining algorithm. Rule mining was used to examine the association relationships between attributes (multivariate) from the predicted outlier values calculated from the VBSE.

2.4 Binning

Telemetry data are numeric; therefore, they cannot be applied directly using the rule mining algorithm. The telemetry data were first converted into interval data. The process of grouping numeric data into several intervals is known as binning. Binning helps reduce data complexity by converting continuous attributes into discrete attributes, making it easier to identify association patterns. The binning method used in this study was the Freedman-Diaconis rule [60]. The Freedman-Diaconis rule uses two main parameters: the IQR and the total number of data points used (n). Equation 1 presents the calculation of the bin width [61], and Equation 2 presents the calculation of the number of bins [61].

$$\text{Bin width} = 2 \frac{\text{IQR}}{\sqrt[3]{n}} \quad (3)$$

$$\text{Number of bins} = \frac{\text{Range}}{\text{Bin width}} \quad (4)$$

If the IQR is 0 (zero) in equation 1, then the IQR is replaced with the range value (maximum value - minimum value). The Freedman-Diaconis rule was chosen because it has advantages, including (1) using IQR, which is more resistant to outliers [62] compared to other methods that use the full range such as the Sturges method [63], and (2) the bin width is adjusted to the data distribution [62], so it is more flexible for various types of data distributions, including non-normally distributed data.

2.5 Association Rule Mining

Association Rule Mining (ARM) is a data mining technique used to identify relationships or patterns that frequently occur between items in a dataset. The main goal of ARM is to find associative rules in the form of if-then statements, for example, "If A and B, then C". This association rule can be understood as a relationship between A and B, where A and B represent two sets of elements [64].

Support (Supp) is used to measure how often an itemset (combination of attributes) occurs in the entire data set. The support for the rule $A \Rightarrow B$ shows how often items A and B appear together in the dataset. Confidence (Conf) measures the strength of the rule, namely, how often item B appears in transactions that also contain A. The confidence for the rule $A \Rightarrow B$ shows the probability that a transaction containing A also contains B. Lift is the ratio of the probability of B occurring when A occurs to the overall probability of B occurring. Lift output defines the correlation between the antecedent and consequent [65]: lift > 1 (positive correlation), lift < 1 (negative correlation), and lift = 1 (independent correlation).

Leverage (Lev) measures the difference between the actual frequency of the combination of items A and B and the expectation if they were independent (i.e., uncorrelated) [66]. The more positive the leverage value, the stronger was the association between items. Conviction (Conv) measures the extent to which the presence of item A increases the probability of occurrence of item B. The formulas for support and confidence are given in Equations 5 and 6 [64], the lift formula is given in Equation 7 [67], and the leverage and conviction formulas are given in Equations 8 and 9 [66].

$$\text{Support}(A \Rightarrow B) = \frac{\text{Frequency}(A,B)}{\text{Total Transactions}} \quad (5)$$

$$\text{Confidence}(A \Rightarrow B) = \frac{\text{Frequency}(A,B)}{\text{Frequency}(A)} \quad (6)$$

$$\text{Lift}(A \Rightarrow B) = \frac{\text{Support}(A \cup B)}{\text{Support}(A) \times \text{Support}(B)} \quad (7)$$

$$\text{Leverage}(A \Rightarrow B) = \text{Support}(A \cup B) - [\text{Support}(A) \times \text{Support}(B)] \quad (8)$$

$$\text{Conviction}(A \Rightarrow B) = \frac{1 - \text{Support}(A)}{1 - \text{Confidence}(A,B)} \quad (9)$$

Association rules can reveal hidden patterns in data that might otherwise go undetected by identifying changes in the value of an attribute, whether influenced by other attributes or by the attribute that indicates an anomaly. The resulting set of rules can help identify various types of anomalies, such as single, correlational, and contextual anomalies. Each type of anomaly can indicate a different cause or source of the problem.

The minimum support value (minSupp) was selected using an elbow point curve by observing support values ranging from 0.1% (1 per 1,000 events) to 0.01% (1 per 10,000 events) with a range of 1,000 values. Based on previous research [68], a small support value was used to balance the rules. A small support value captures rare, but important, patterns. Rare patterns are important for understanding complex interactions that are not readily apparent using higher supports. A confidence value above 90% was chosen to obtain classification rules with a high level of confidence, maintain rule effectiveness, and avoid information loss that may occur at lower confidence levels [69]. According to [70], a confidence value above 90% was chosen to ensure that the rules used in the model were reliable and had a high probability of accurately representing interactions in the data, and to reduce false positives.

The AprioriClose algorithm was chosen because it simultaneously meets three operational requirements: (1) it does not require labelled data (unsupervised); (2) it generates association rules that can be directly interpreted by operators without specialised ML/DL expertise; and (3) the closed frequent itemset formula [71] is capable of eliminating redundant rules, that is, itemsets that do not have a superset (larger itemset) with the same support value, without losing important information. The resulting data-based rules can be cross-validated with expert-generated labels to determine their agreement. Rules that frequently appear as anomalies and match expert labels are more reliable than rules that appear on unlabelled data points. Several explainability techniques have been considered as alternatives to the proposed AprioriClose-based explanation layer, including SHAP, LIME, and rule-based classifiers such as decision trees.

However, each of these approaches presents fundamental limitations when applied to the operational context of this study. SHAP (SHapley Additive exPlanations) requires a pre-trained model to be explained [72]. Thus, SHAP cannot be directly used in unsupervised anomaly detection methods such as VBSE. SHAP models require additional modelling, which adds complexity to the system and can reduce transparency. SHAP outputs are numerical scores that are difficult to understand without a good knowledge of statistics. SHAP outputs make them less practical for operators who are not ML experts and need explanations that can be directly used in satellite missions [73]. Methods such as LIME also have limitations, such as explaining only one prediction at a time and not showing general patterns across the data [74]. To monitor the health of a satellite over eight years, operators need to know long-term patterns, such as which parameters frequently appear with anomalies over months or years. Alternatives such as rule-based classifiers (decision trees) generate rules that are easy for operators to understand; however, they require labelled anomaly data to create classification boundaries [75].

3. RESULTS AND DISCUSSION

3.1. Voting Implementation

Initially, telemetry data from 2016 to 2023 were divided into eight cumulative datasets. Each cumulative dataset was assigned an ID code from DS1 to DS8. Outlier scores were calculated for each attribute using the VBSE approach. Table 4 presents the VBSE approach.

Table 4. Illustration of VBSE methods

No.	Datetime	Fuse5_12	Hampel	IQR	Modified Z-score	ABSD	Z-score	Voting
1	2016-01-07 08:21:29	160.8	N/O	N/O	N/O	Outlier	N/O	N/O
2	2016-01-07 08:23:03	603.99	Outlier	Outlier	Outlier	N/O	N/O	Outlier
3	2016-01-07 08:23:18	1062.86	Outlier	Outlier	Outlier	N/O	Outlier	Outlier
4	2016-01-07 08:32:20	262.77	Outlier	N/O	N/O	N/O	N/O	N/O
5	2016-01-07 08:33:03	121.58	N/O	N/O	N/O	Outlier	N/O	N/O
6	2016-01-07 08:33:10	160.8	N/O	N/O	N/O	Outlier	N/O	N/O
7	2016-01-07 08:33:29	133.35	N/O	N/O	N/O	Outlier	N/O	N/O
8	2016-01-07 08:34:22	133.35	N/O	N/O	N/O	Outlier	N/O	N/O
9	2016-01-07 08:34:31	125.5	N/O	N/O	N/O	Outlier	N/O	N/O

*N/O = not outlier

Table 5 presents the dataset distribution. The "ID" attribute defines the dataset code. The "Dataset period" attribute defines the time span of the dataset used. The "Amount of data" attribute defines the number of observational data for each period, which increases over time. The "Ground truth" attribute defines the number of outlier data for each period based on a threshold value derived from expert experience and knowledge. The "% Outlier" attribute defines the percentage ratio between the total data and outliers based on the ground truth. The "Number of suspected outliers by VBSE" attribute defines the number of suspected outlier data based on VBSE.

Table 5. Distribution of datasets

ID	Dataset Period	Amount of data	Ground Truth	% Outlier	Number of suspected outliers by VBSE
DS1	2016	269,031	2,488	0.92	4,346
DS2	2016 to 2017	570,285	5,359	0.94	8,487
DS3	2016 to 2018	884,902	7,534	0.85	10,053
DS4	2016 to 2019	1,277,765	10,351	0.81	13,817
DS5	2016 to 2020	1,688,016	12,682	0.75	18,225
DS6	2016 to 2021	2,075,152	16,964	0.82	24,931
DS7	2016 to 2022	2,309,361	19,895	0.86	26,105
DS8	2016 to 2023	2,464,887	21,198	0.86	27,835

Eight years of accumulated telemetry data yielded 21,198 anomalies based on expert experience and knowledge, with an outlier rate ranging from 0.75% to 0.94%. The class imbalance ratio between normal and anomalous values is extreme (115:1), representing real-world satellite operations, in which anomalies are rare but significant events in a satellite's mission.

The recall and F1-score values of VBSE will be compared with those of other anomaly detection methods, namely IForest, LOF, and ECOD. Table 6 lists the comparison methods and the parameters used. The "Hyperparameter Tuning" attribute defines several parameter combinations tested to obtain the best recall and F1-score values. The "Selected Parameter" attribute defines the parameters selected based on the best hyperparameter tuning results.

Table 6. Comparative methods and parameters used

Method	Hyperparameter Tuning	Selected Parameter
IForest	contamination: [0.01, 0.05, 0.1, 0.15, 0.2],	contamination: 0.1,
	n_estimators: [25, 50, 100, 150],	n_estimators: 25,
	max_samples: [0.5, 0.8, auto],	max_samples: 0.5,
	max_features: [0.5, 0.8, 1.0]	max_features: 0.5
LOF	n_neighbors: [15, 20, 25, 30, 35, 40, 45, 50],	n_neighbors: 15,
	contamination: [0.1, 0.01, 0.001],	contamination: 0.1,
	metric: [euclidean, manhattan, minkowski],	metric: euclidean,
ECOD	p: [1, 2, 3]	p: 2
	Contamination: [0.1, 0.01, 0.001, 0.0001]	contamination: 0.001

Table 7 presents a comparison of the recall and F1-score values of VBSE with several other anomaly detection methods. Figure 3 shows the comparative performance of outlier detection methods on eight satellite telemetry datasets. VBSE demonstrated superior and consistent performance across all datasets, with an average F1-score of 70.48 and a recall of 86.49. This F1-score represents a 4.2-fold improvement over IForest (16.95), a 12.6-fold improvement over ECOD (5.57), and a 5.6-fold improvement over LOF (12.65). Although IForest achieved the highest average recall (91.92), it significantly sacrificed precision. IForest's average precision was approximately 9.34, indicating that ~90% of alerts were false positives. In contrast, VBSE maintained a high recall (86.49) while achieving significantly better precision (~ 59–65%) than IForest. The balance between recall and F1-score is crucial for real-world applications. Relying solely on precision or recall can flag too many normal events as anomalies, whereas the F1-score provides a single metric that balances both.

Table 7. Comparison between VBSE, IForest, ECOD, and LOF

ID/Method	VBSE		IForest		ECOD		LOF	
	Recall	F1-Score	Recall	F1-Score	Recall	F1-Score	Recall	F1-Score
DS1	68.51	50.04	91.36	18.14	2.81	5.08	31.55	11.24
DS2	72.42	55.98	92.42	18.83	3.10	5.67	21.31	12.57
DS3	88.86	74.93	95.77	17.76	4.95	8.86	17.20	13.24
DS4	90.28	76.45	95.22	16.63	4.22	7.57	14.24	14.20
DS5	93.64	75.57	97.03	16.09	3.60	6.38	12.35	14.47
DS6	90.96	71.81	88.96	15.74	2.30	4.14	9.46	12.62
DS7	93.51	79.50	88.04	16.36	1.97	3.54	8.10	11.53
DS8	93.76	79.60	86.58	16.01	1.84	3.34	7.78	11.30
Mean	86.49	70.48	91.92	16.95	3.10	5.57	15.25	12.65

VBSE's scalability demonstrates consistent performance improvement from DS1 (F1=50.04) to DS8 (F1=79.60), representing a ~62% improvement. This scalability demonstrates VBSE's ability to adapt to the increasing volume and complexity of data. In contrast, IForest exhibits performance degradation (18.14 → 16.01), ECOD exhibits fluctuating behaviour (5.08 → 3.34), and LOF exhibits stagnation (11.24 → 11.30). These trends indicate the poor generalisability of a single-algorithm approach when faced with operational scenarios of long-duration satellite missions.

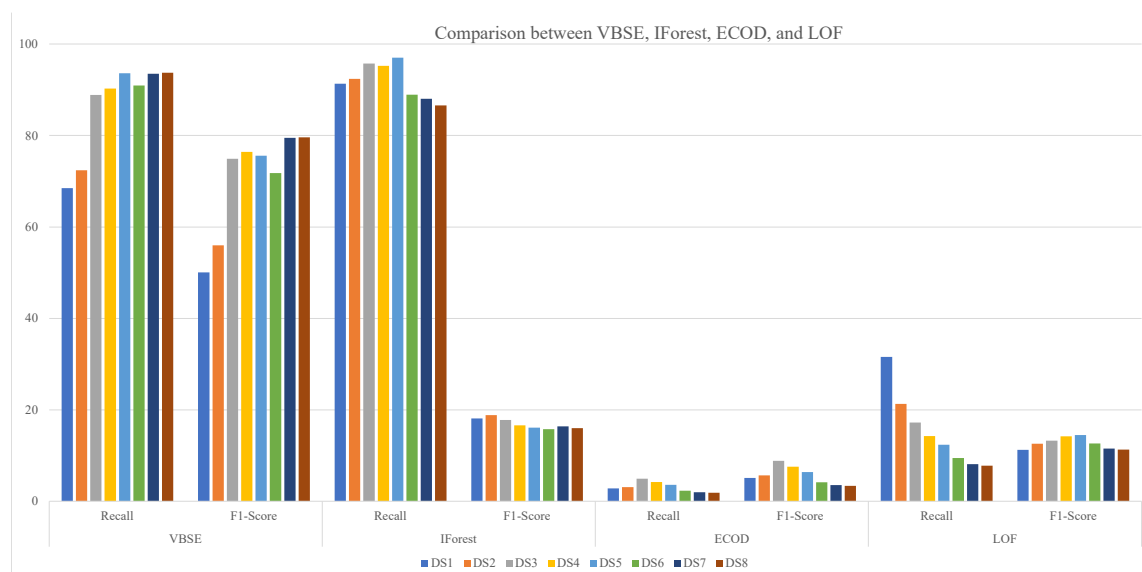


Figure 3. Comparative performance of outlier detection methods

Based on algorithmic characteristics, VBSE excels because it combines multiple univariate statistical detectors to balance sensitivity and specificity to skewed data distributions. Each attribute is evaluated independently from multiple perspectives. Furthermore, its direct voting mechanism has been shown to reduce false positives. In contrast, IForest is sensitive to parameter determination and variations in the density of skewed telemetry data. The random splits applied to IForest can isolate normal values at the tails of skewed data distributions.

ECOD is not robust to skewed data distributions, where the tails of the distribution do not represent true outliers because the data distribution is only skewed to one side. In skewed telemetry data, the tails of the distribution can contain many normal values (not outliers). Meanwhile, LOF, which is based on density estimation, is less robust to skewed data with varying densities. LOF is less effective when used in one dimension (univariate), because density differences between points become insignificant.

Table 8 presents a comparison of the 10 attributes with the most outlier values for DS1–DS4, and Table 9 presents a comparison of the 10 attributes with the most predicted outlier values for DS5–DS8. Each fuse manages several switches in satellite subsystems, such as the bus system (coil, gyro, and star sensor) and payload (AIS, APRS, and SpaceCam). Fuse0 controls the S-band switch, voice repeater, and S-band transmitter. Fuse1 controls the GPS, MFS, VAIS, and APRS switches. Fuse2 controls the STS-A, STS-B, stepper, and analogue camera switches. Fuse3 controls the gyro-1 and 2, ASR, gyro-3 and 4, and video recorder switches. Fuse4 controls the WDE-3 and WDE-4 switches. Fuse5 controls the WDE-1 and WDE-2 switches. Fuse6 controls the coil 1–3 switches. Fuse7 controls the SpaceCam and video camera switches. For example, "Fuse4_2" indicates that only switch 2 is active on Fuse4. While "Fuse1_124" indicates that switches 1, 2, and 4 are active on Fuse1.

The results presented in Tables 8 and 9 indicate that each dataset has certain attributes that dominate outliers. Table 8 presents the number of outliers from the first four years of the dataset (DS1–DS4), with the top three ranked by the Fuse attribute. Fuse4_2, which controls the wheel drive electronic (WDE-3), dominated the number of outliers from DS1 to DS3, but over time its position fell behind the other attributes. Attribute Fuse5_12, which controls the WDE-1 and WDE-2 switches, was in the top three in the first four datasets. The frequency of outliers in Fuse5 may indicate a continuous increase in load or mechanical activity on the WDE, which is affecting Fuse5.

Table 9 presents the number of outliers from the last four years of the dataset (DS5–DS8), where the top four rankings are occupied by attributes involving battery voltage. The subsequent rankings are dominated by attributes involving battery current. Furthermore, increased outlier scores for VBatt and IBatt may reflect decreased battery capacity or instability in charging cycles, which often occurs in Li-ion batteries after being exposed to extreme thermal cycling in orbit.

Table 8. Comparison of the number of outliers per dataset based on voting

Dataset	DS1		DS2		DS3		DS4	
No.	Attribute	Sum	Attribute	Sum	Attribute	Sum	Attribute	Sum
1	Fuse4_2	244	Fuse4_2	317	Fuse6_1	424	Fuse5_12	498
2	Fuse5_12	159	Fuse5_12	298	Fuse5_12	402	Fuse1_4	438
3	Fuse3_123	129	Fuse3_123	271	Fuse4_2	339	Fuse6_1	429
4	FusePCDH	70	Fuse1_24	167	VBatt3	290	VBatt3	416
5	Fuse3_2	64	VBatt3	134	Fuse3_123	285	VBatt2	402
6	Fuse7_1	57	Fuse6	124	VBatt2	272	VBatt1	401
7	Fuse1_24	56	I_Solar0	117	VBatt1	270	IBatt1	352
8	Fuse6	50	U_UMPB	113	Fuse1_4	261	Fuse4_2	348
9	Fuse2_2	29	FusePCDH	112	U_UMPB	241	IBatt3	341
10	Fuse1_124	28	VBatt1	104	IBatt3	225	IBatt2	311

Table 9. Comparison of the number of outliers per dataset based on voting

Dataset	DS5		DS6		DS7		DS8	
No.	Attribute	Sum	Attribute	Sum	Attribute	Sum	Attribute	Sum
1	U_UMPB	598	Fuse3_12	1217	VBatt3	837	VBatt2	918
2	VBatt1	544	VBatt2	673	VBatt2	833	VBatt3	911
3	VBatt3	544	VBatt3	673	VBatt1	759	VBatt1	838
4	VBatt2	542	VBatt1	660	IBatt1	673	IBatt3	767
5	Fuse5_12	535	U_UMPB	633	IBatt3	663	IBatt1	713
6	IBatt1	457	IBatt3	594	U_UMPB	633	U_UMPB	661
7	IBatt3	456	Fuse5_12	568	IBatt2	604	IBatt2	647
8	Fuse1_4	438	IBatt1	559	Fuse5_12	571	Fuse5_12	572
9	IBatt2	431	IBatt2	522	Fuse6_1	537	Fuse6_1	570
10	Fuse6_1	429	Fuse6_1	507	Fuse3_123	439	Fuse3_123	448

3.2 Rule Mining Implementation

Values detected as true positive outliers by VBSE are further processed to determine which bin they fall into. Table 10 presents an example of the binning results into item sets for mining on DS1. For example, item U_UMPB_14 indicates that the detected outlier value of U_UMPB falls into bin 14.

Table 10. Example itemsets in DS1

Datetime	Itemsets
2016-04-11 01:07	U_UMPB_14, VBatt1_11, VBatt2_11, VBatt3_11
2016-04-11 01:07	U_UMPB_14, VBatt1_12, VBatt2_12, VBatt3_12
2016-04-11 01:08	U_UMPB_9, VBatt1_9, VBatt2_10, VBatt3_10
2016-04-11 01:09	U_UMPB_4, VBatt1_6, VBatt2_6, VBatt3_6
2016-04-11 01:11	Fuse3_123_119, U_UMPB_0, VBatt1_0, VBatt2_0, VBatt3_0

The optimal support value was determined using an elbow point curve by testing several support values ranging from 0.1% (1 per 1,000 occurrences) to 0.01% (1 per 10,000 occurrences). The results of identifying a significant increase in the number of rules between support values indicated that the optimal support value was 0.02% (1 per 5,000 occurrences). Table 11 shows a comparison of the change in the support value with the number of generated rules. A MinSupp value of 0.02% was chosen because it best represented the point at which the increase in the number of rules began to slow significantly as the minSupp value decreased. The "Growth" attribute defines the average increase in the number of rules per 1,000 occurrences. Figure 4 shows a visualization of the elbow point graph.

Table 11. comparison of changes in support value to the number of rules

Support/ Dataset	0.1%	0.05%	0.033%	0.025%	0.02%	0.017%	0.014%	0.0125%	0.011%	0.01%
DS1	14	52	84	116	116	157	157	157	157	157
DS2	16	42	71	91	134	134	169	169	225	225
DS3	13	57	91	121	145	163	183	222	222	268
DS4	8	39	88	119	142	159	200	223	223	244
DS5	6	28	63	109	141	157	180	197	220	244
DS6	9	31	64	98	132	158	184	202	219	236
DS7	9	28	61	84	115	152	176	203	222	232
DS8	8	28	62	89	115	148	171	192	224	236
Mean	10.38	38.13	73.00	103.38	130.00	153.50	177.50	195.63	214.00	230.25
Growth	0	27.75	34.88	30.38	26.63	23.50	24.00	18.13	18.38	16.25

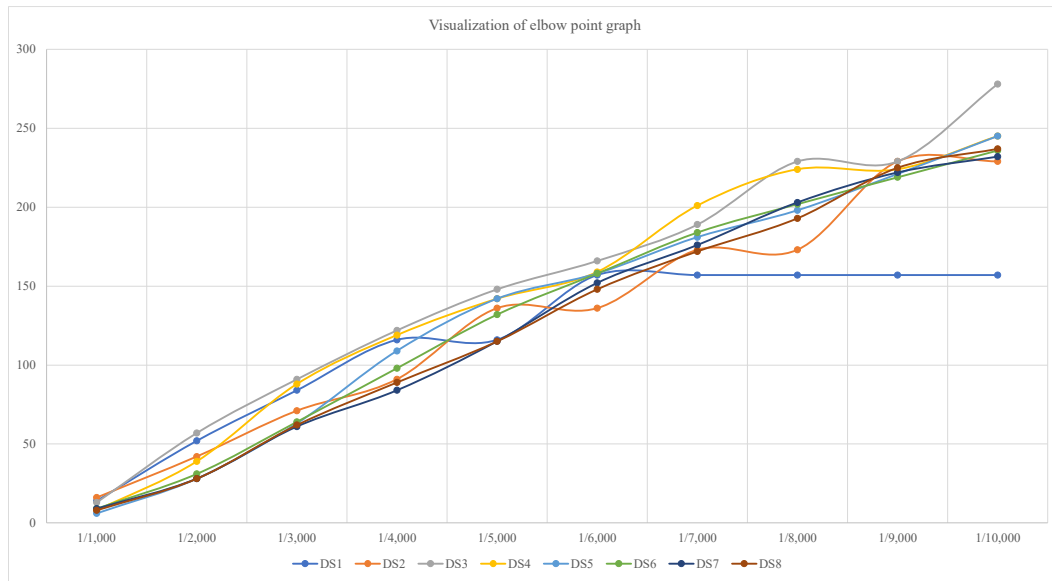


Figure 4. Visualization of elbow point graph

By applying a minSupp of 0.0002, each dataset had a different frequency value depending on the number of rows of data to be categorized as an outlier. Attributes with a frequency of occurrence below the frequency value were removed from the dataset and considered as noise. Table 12 shows the minimum frequency of occurrence of attributes for each dataset to be categorized as an outlier, and Figure 5 shows a comparison between the number of rows and the absolute support.

Table 12. Absolute support for each dataset

ID	Dataset Period	Number of Rows	Absolute support
DS1	2016	11,697	2
DS2	2016 to 2017	24,795	5
DS3	2016 to 2018	38,474	8
DS4	2016 to 2019	55,555	11
DS5	2016 to 2020	73,392	15
DS6	2016 to 2021	90,224	18
DS7	2016 to 2022	100,407	20
DS8	2016 to 2023	107,169	21

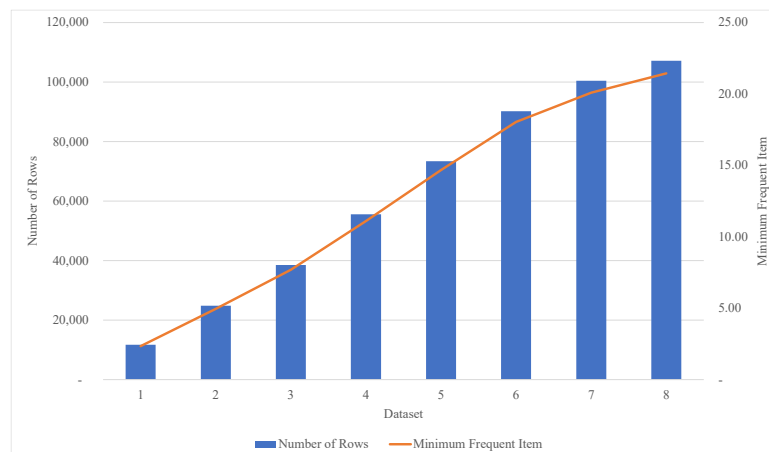


Figure 5. Comparison between the number of rows and absolute support

Table 13 shows examples of rules with the most frequent itemsets in each dataset, with a minSupp of 0.02%. The rules that emerged involved not only one attribute (L1) but also up to four attributes (L4). The more attributes involved, the more complex the attribute interactions in the system.

Table 13. The highest rule combinations in each dataset

Dataset	Length	Frequent	Supp.	Conf.	Lift	Lev.	Conv.	Antecedent	Consequent
DS1	L1	64	0.00547	1	0.0	0	0.0	Fuse3_2_0	-
	L1	50	0.00427	1	0.0	0	0.0	Fuse6_0	-
	L2	4	0.00034	1	2924.3	0.00034	99830.2	VBatt2_10	VBatt3_10
	L2	3	0.00026	1	182.8	0.00026	99318.0	Fuse1_134_45	Fuse3_2_0
	L3	3	0.00026	1	2924.3	0.00026	99830.2	U_UMPB_9, VBatt2_10	VBatt3_10
	L3	3	0.00026	1	2924.3	0.00026	99830.2	VBatt1_10, VBatt2_10	VBatt3_10
	L3	3	0.00026	1	2924.3	0.00026	99830.2	VBatt1_14, VBatt2_15	VBatt3_15
DS2	L1	124	0.005	1	0	0	0.0	Fuse6_0	-
	L1	81	0.00327	1	0	0	0.0	Fuse3_2_0	-
	L2	10	0.0004	1	1771.1	0.0004	99808.0	VBatt2_11	VBatt3_11
	L3	8	0.00032	1	1771.1	0.00032	99808.0	VBatt1_11, VBatt2_11	VBatt3_11
	L3	7	0.00028	1	1771.1	0.00028	99808.0	U_UMPB_8, VBatt2_11	VBatt3_11
	L2	5	0.0002	1	953.7	0.0002	99759.7	VBatt1_13	VBatt3_11
	L4	5	0.0002	1	2066.3	0.0002	99816.0	U_UMPB_6, VBatt1_8, VBatt2_8	VBatt3_13
DS3	L1	261	0.00678	1	0	0	0.0	Fuse1_4_0	-
	L1	134	0.00348	1	0	0	0.0	Fuse6_0	-
	L1	120	0.00312	1	0	0	0.0	U_UMPB_20	-
	L3	11	0.00029	0.917	1410.7	0.00029	12.0	U_UMPB_20, VBatt1_18	VBatt2_19
	L2	9	0.00023	0.9	2663.6	0.00023	10.0	VBatt1_11	VBatt2_12
DS4	L1	438	0.00788	1	0	0	0	Fuse1_4_0	-
	L1	135	0.00243	1	0	0	0	Fuse6_0	-
	L1	127	0.00229	1	0	0	0	U_UMPB_23	-
	L2	13	0.00023	1	677.5	0.00023	99717.0	Fuse1_14_219	Fuse3_2_0
DS5	L1	438	0.00597	1	0	0	0	Fuse1_4_0	-
	L1	286	0.0039	1	0	0	0	U_UMPB_33	-
	L1	136	0.00185	1	0	0	0	Fuse6_0	-
	L1	129	0.00176	1	0	0	0	U_UMPB_25	-
	L1	82	0.00112	1	0	0	0	Fuse3_2_0	-
DS6	L1	438	0.00485	1	0	0	0	Fuse1_4_0	-
	L1	312	0.00346	1	0	0	0	U_UMPB_35	-
	L1	250	0.00277	1	0	0	0	Fuse3_12_44	-
	L1	232	0.00257	1	0	0	0	Fuse3_12_14	-
	L1	182	0.00202	1	0	0	0	Fuse3_12_29	-
DS7	L1	438	0.00436	1	0	0	0	Fuse1_4_0	-
	L1	312	0.00311	1	0	0	0	U_UMPB_36	-
	L1	137	0.00136	1	0	0	0	Fuse6_0	-
	L1	133	0.00132	1	0	0	0	U_UMPB_27	-
	L1	128	0.00127	1	0	0	0	Fuse3_1_38	-
DS8	L1	438	0.00409	1	0	0	0	Fuse1_4_0	-
	L1	327	0.00305	1	0	0	0	U_UMPB_37	-
	L1	145	0.00135	1	0	0	0	U_UMPB_28	-
	L1	137	0.00128	1	0	0	0	Fuse3_1_41	-
	L1	137	0.00128	1	0	0	0	Fuse6_0	-

The priori-close results show that the majority of frequent itemsets with the highest support across all datasets (DS1–DS8) were dominated by anomalies in fuses and the main power bus. The single antecedent fuse (Fuse1_4_0, Fuse3_2_0, and Fuse6_0) that recurred in almost all datasets with perfect confidence (1) indicated a recurring anomaly in the satellite power protection path. The majority of detected fuse attributes had small bin sequence numbers (0-1), indicating small outlier values (undercurrent).

Furthermore, the frequent occurrence of U_UMPB_k in L1 indicates that anomalies tend to occur under certain bus conditions (e.g., a drop or spike at the kth level), indicating main power bus instability. For example, U_UMPB_8 on DS2 was detected as an anomaly because the voltage dropped to 14.53V from an average value of 15.87V. The repeated occurrence of fuses and the main power bus in rule L1 are good early indicators for anomaly detection.

Rules involving more than one attribute, confidence values approaching 1, and a very high lift, such as DS1:2924.3, DS2:2066.3, and DS3:2663.6, indicate a very strong and deterministic antecedent-consequence relationship. Rules with a high lift often involve the main power bus (U_UMPB) and battery (VBatt) variables, especially VBatt2 and VBatt3, as in DS1: U_UMPB_9, VBatt2_10 → VBatt3_10 with a lift value of 2924.3. This means that when U_UMPB_9 and VBatt2_10 (antecedent) occur, VBatt3_10 (consequent) is 2924.3 times more likely than under normal circumstances. This condition indicates that anomalies in the main power bus and one battery affect the other. Furthermore, this combination of rules can be used as an early indicator of anomalies. When U_UMPB reaches this value, the system can predictively monitor VBatt2 and VBatt3. If a complete pattern is formed, the onboard system can take protective actions, such as reducing the load, shutting down non-essential payloads, or transitioning to the safe mode.

A very high conviction value (e.g., 99830.2 in DS1) indicates that the rule is very strong, and if the antecedent occurs, the consequent is almost certain to occur. Furthermore, a very high conviction value reduces the number of false positives. The combination of high conviction with positive leverage reflects a rule that is nearly perfect but only valid in a specific context. In addition, the combination of positive leverage with high lift indicates a strong and non-random (not coincidental) relationship between items, even though the frequency of occurrence is low.

Several types of anomalies can be identified based on the rule set in Table 13. First, point anomalies are independent because their occurrence is not triggered or triggered by other components. Fuse1_4_0, Fuse6_0, Fuse3_2_0, or other L1 attributes can be categorized as point anomalies (global) because they are not associated with other parameters (lift = 0). These L1 attributes are not noisy because they appear consistently across multiple datasets and satisfy the minimum support threshold value. For example, Fuse3_2_0 is detected as an outlier in several datasets because when switch-2 is active, the ASR current consumption is below the nominal value (undercurrent) that should be (~91mA). Figure 6 Example of detected point anomalies.

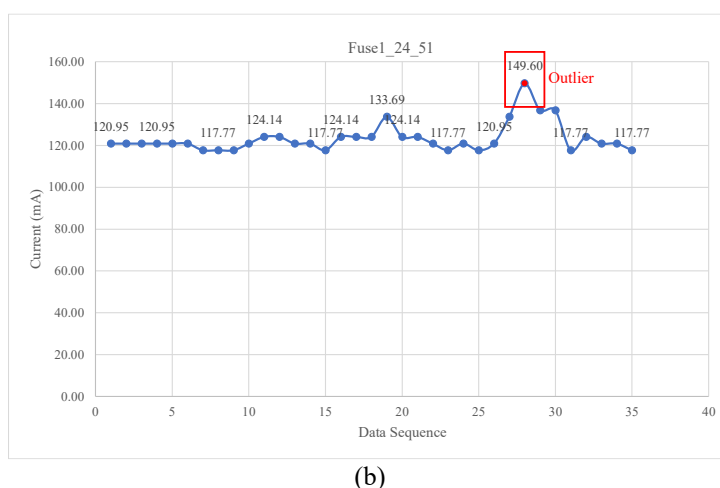
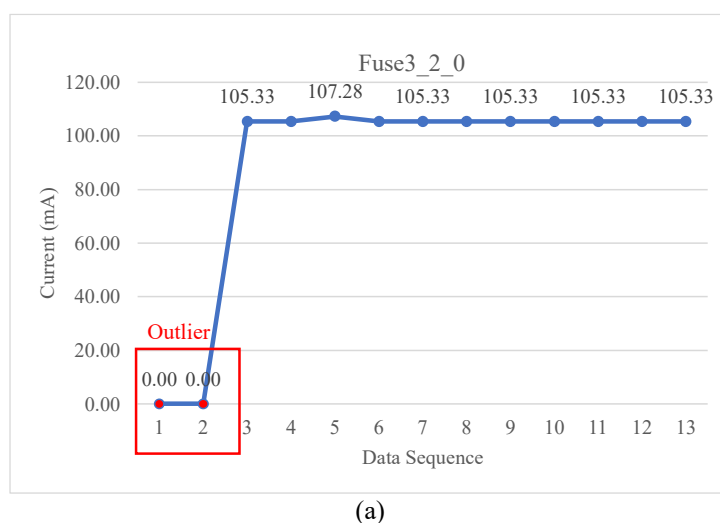


Figure 6. Example of detected point anomalies

Anomalies that appear regularly in DS1-DS4 with an attribute combination pattern involving U_UMPB_x, VBatt1_x, VBatt2_x, and VBatt3_x at unusual times are categorized as collective anomalies. This is characterized by a confidence level close to 1.0 and a high lift. For example, in DS3, the values U_UMPB_20, VBatt1_18, and VBatt2_19 are individual point anomalies, but when they occur together, they can be categorized as collective anomalies with a combined occurrence of 11 times. These three combinations appear together and are very consistent, indicated by a confidence level close to 1.0, and supported by a lift 1410.7 times higher than chance. Figure 7 shows an example of the collective anomalies of U_UMPB_20, VBatt1_18, and VBatt2_19.

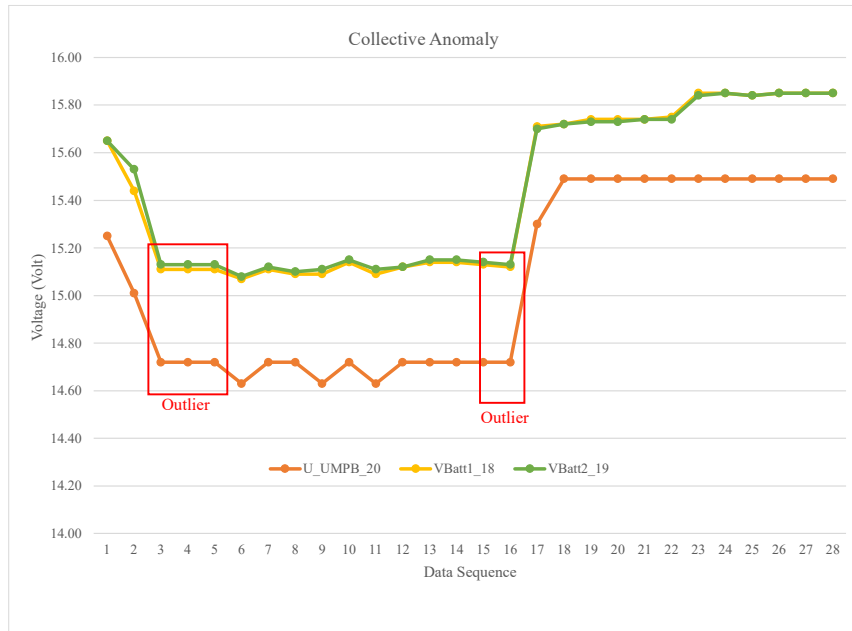


Figure 7. Example of detected collective anomalies

On DS4, Fuse1_14_219 and Fuse3_2_0 were detected as collective anomalies. By itself, Fuse1_14_219 is not an anomaly, whereas Fuse3_2_0 is a point anomaly. Fuse3_2_0 was detected as an outlier because when switch-2 was active, the ASR current consumption was below the nominal value (~91mA). The combination of these two fuses occurred 13 times, with a confidence level of 1.0, indicating that they always occurred together. This is supported by the lift, which indicates a 677.5 times greater probability of occurring together than simply by chance. Figure 8 shows an example of the collective anomalies of Fuse1_14_219 and Fuse3_2_0.

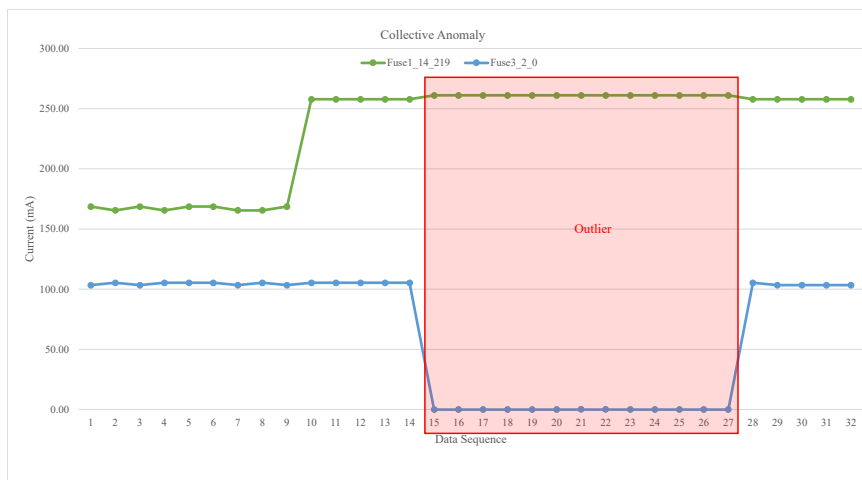


Figure 8. Example of detected collective anomalies

Rules that appear in DS1-DS4 and then disappear in DS5-DS8 can be classified as contextual anomalies, such as the combination VBatt2_10 → VBatt3_10 in DS1 and VBatt1_11, VBatt2_11 → VBatt3_11 in DS2. Although these combinations appear at the beginning of the dataset and are considered anomalies, when integrated with the subsequent datasets (DS4-DS8), they are no longer considered anomalies owing to the increase in the number of data without an increase in the frequency of their occurrence. In addition, distinct pattern changes were observed between the early and late phases. A clear transition boundary is seen in DS4 to DS5, namely, the change from complex multi-attributes to simpler single attributes.

The prolonged or repeated occurrence of Fuse1_4_0 is inconsistent with the operational schedule and may indicate a problem or failure in the satellite subsystem. However, we acknowledge that distinguishing true faults from nominal deviations requires cross-validation with satellite operational logs, which is a limitation of this study and may be addressed in future research.

4. CONCLUSION

VBSE demonstrated improved performance as dataset complexity increased (DS1 → DS8), whereas IForest, LOF, and ECOD tended to decline. VBSE's F1-score of 70.49 represents a 4.2–12.6 x improvement over the single-algorithm baseline while reducing false positives by 90% compared to IForest. Furthermore, a change in outlier patterns was detected, with the Fuse attribute dominating in the early datasets (DS1–DS4) but the VBatt attribute dominating in the later datasets (DS5–DS8).

An eight-year cumulative analysis revealed that a combination of battery voltage (VBatt) and main power bus voltage (U_UMPB) parameters is a key indicator of anomalies on the LAPAN-A2 satellite. The cumulative approach proved effective in maintaining detection sensitivity; however, computational refinements are needed for its long-term effectiveness.

Overall, rule mining yielded several important findings: (1) utilising closed frequent itemsets (AprioriClose) helps reduce rule redundancy without sacrificing critical information. (2) The combination of minSupp. 0.02%, minConf. 0.9, lift, leverage, and conviction metrics proved effective in capturing rare, reliable anomalies and significant association relationships, and provided a different perspective in assessing the significance of the rules, especially in the context of outlier detection. (3) AprioriClose successfully revealed hidden patterns between satellite power subsystems. (4) a strong correlation between batteries (VBatt1-VBatt3) and a causal relationship between U_UMPB and VBatt were found. (5) the use of accumulative scenarios allows for the identification of stable patterns over time, but also requires an algorithm that can handle data updates efficiently. (6) provides empirical evidence that association methods remain relevant and effective for anomaly detection on satellites and produce interpretable anomaly patterns.

Although formal expert validation of the association rules was not performed, the findings indicate that the rules demonstrate strong validity when interpreted against satellite power systems. An eight-year cumulative analysis revealed a causal relationship between battery voltage (VBatt) and main power bus voltage (U_UMPB), both key indicators of anomalies on the LAPAN-A2 satellite. However, we fully acknowledge that the lack of structured expert validation of the association rules is a limitation of this study. Formal validation, through structured expert interviews, where the extracted rules are presented and assessed for their operational relevance, is identified as a necessary direction and prioritised for future work.

This study also highlights the potential of a cumulative approach for long-term outlier detection but cautions against the potential for overfitting to past events. Further research is needed to explore weighting-based voting mechanisms and other optimal binning methods that consider data distribution.

ACKNOWLEDGMENTS

We would like to thank the LAPAN satellite operator at the Research Center for Satellite Technology (PRTS) for providing telemetry data for this research. In addition, the authors thank the experts who have been willing verified the dataset used in this study: Dr.-Ing. Mohammad Mukhayadi and Mr. Satriya Utama (attitude control), Mr. Abdul Karim (power systems), Mr. Arif Saifudin (sun and star sensor), Mrs. Rosza Madina (mission planning), Mr. Deddy El Amin (electronic and WDE), and Mr. Muhammad Taufik (electronic).

REFERENCES

- [1] D. Gadisa and H. Bang, "Small Satellite Electro-Optical System (EOS) Technological and Commercial Expansion," *Acta Astronaut.*, vol. 213, pp. 355–372, Dec. 2023, doi: 10.1016/j.actaastro.2023.09.010.
- [2] L. Miola, V. Muffato, E. Sella, C. Meneghetti, and F. Pazzaglia, "GPS use and navigation ability: A systematic review and meta-analysis," *J. Environ. Psychol.*, vol. 99, p. 102417, Nov. 2024, doi: 10.1016/j.jenvp.2024.102417.
- [3] J. W. Kim, V. K. Bhosale, K. S. Kim, S. Lee, and S. Kwon, "Room-temperature catalytically reactive ammonium dinitramide-H₂O₂ monopropellant for microsatellites," *Advances in Space Research*, vol. 69, no. 3, pp. 1631–1644, Feb. 2022, doi: 10.1016/j.asr.2021.11.022.
- [4] Y. Cai and Z. Wang, "GNSS receiver for Q-Sat and its analysis of precise orbit determination," *Acta Astronaut.*, vol. 200, pp. 357–370, Nov. 2022, doi: 10.1016/j.actaastro.2022.08.031.
- [5] C. Xing and Y. Hongmei, "Abnormal Discharge Detection Method of Satellite Battery System Based on FPGrowth-RF," in *International Conference on System Reliability and Safety Engineering, SRSE 2022*, Guangzhou, China: Institute of Electrical and Electronics Engineers Inc., Dec. 2022, pp. 284–288. doi: 10.1109/SRSE56746.2022.10067354.
- [6] I. P. Rodrigues, P. A. S. Oliveira, A. M. Ambrosio, and R. A. J. Chagas, "Modeling satellite battery aging for an operational satellite simulator," *Advances in Space Research*, vol. 67, no. 6, pp. 1981–1999, Mar. 2021, doi: 10.1016/j.asr.2020.12.031.
- [7] M. Cao, T. Zhang, Y. Liu, Y. Wang, and Z. Shi, "A Bayesian optimization hyperband-optimized incremental deep belief network for online battery behaviour modelling for a satellite simulator," *J. Energy Storage*, vol. 58, Feb. 2023, doi: 10.1016/j.est.2022.106348.

- [8] S. Y. Kim, J. F. Castet, and J. H. Saleh, "Spacecraft electrical power subsystem: Failure behavior, reliability, and multi-state failure analyses," *Reliab. Eng. Syst. Saf.*, vol. 98, no. 1, pp. 55–65, Feb. 2012, doi: 10.1016/j.res.2011.10.005.
- [9] P. V. A. Alves *et al.*, "Machine Learning Applied to Anomaly Detection on 5G O-RAN Architecture," in *Procedia Computer Science*, Elsevier B.V., 2023, pp. 104–113. doi: 10.1016/j.procs.2023.08.146.
- [10] M. Ma, L. Han, and C. Zhou, "BTAD: A binary transformer deep neural network model for anomaly detection in multivariate time series data," *Advanced Engineering Informatics*, vol. 56, Apr. 2023, doi: 10.1016/j.aei.2023.101949.
- [11] S. Cuéllar, M. Santos, F. Alonso, E. Fabregas, and G. Farias, "Explainable anomaly detection in spacecraft telemetry," *Eng. Appl. Artif. Intell.*, vol. 133, p. 108083, Jul. 2024, doi: 10.1016/j.engappai.2024.108083.
- [12] R. Yu, Y. Wang, and W. Wang, "AMAD: Active learning-based multivariate time series anomaly detection for large-scale IT systems," *Comput. Secur.*, vol. 137, Feb. 2024, doi: 10.1016/j.cose.2023.103603.
- [13] X. Gong, F. Zhang, T. Lu, and W. You, "Comparative Analysis of three outlier detection methods in univariate data sets," in *International Conference on Electronic Communication and Artificial Intelligence, IWECAI 2022*, Institute of Electrical and Electronics Engineers Inc., 2022, pp. 209–213. doi: 10.1109/IWECAI55315.2022.00048.
- [14] J. Murphy, J. E. Ward, and B. Mac Namee, "An Overview of Machine Learning Techniques for Onboard Anomaly Detection in Satellite Telemetry," in *European Data Handling & Data Processing Conference (EDHPC)*, IEEE, Oct. 2023, pp. 1–6. doi: 10.23919/EDHPC59100.2023.10396403.
- [15] F. T. Liu, K. M. Ting, and Z.-H. Zhou, "Isolation Forest," in *2008 Eighth IEEE International Conference on Data Mining*, IEEE, Dec. 2008, pp. 413–422. doi: 10.1109/ICDM.2008.17.
- [16] S. B. K. Preetha, J. V. M. Sai, V. S. Raj, M. J. Sekhar, and R. Lavanya, "Anomaly Detection in Satellite Power System using Deep Learning," in *Proceedings of CONECCT 2024 - 10th IEEE International Conference on Electronics, Computing and Communication Technologies*, Institute of Electrical and Electronics Engineers Inc., 2024. doi: 10.1109/CONECCT62155.2024.10677120.
- [17] E. H. Bensikaddour and B. Nasri, "Detection and Prediction of Satellite Telemetry Anomalies," in *Proceedings of EDiS 2024 - 2024 4th International Conference on Embedded and Distributed Systems*, Institute of Electrical and Electronics Engineers Inc., 2024, pp. 310–313. doi: 10.1109/EDiS63605.2024.10783288.
- [18] M. M. Breunig, H.-P. Kriegel, R. T. Ng, and J. Sander, "LOF: identifying density-based local outliers," in *Proceedings of the 2000 ACM SIGMOD international conference on Management of data - SIGMOD '00*, New York, New York, USA: ACM Press, 2000, pp. 93–104. doi: 10.1145/342009.335388.
- [19] W. Wang, W. Shangguan, J. Liu, and J. Chen, "Enhanced Fault Detection for GNSS/INS Integration Using Maximum Correntropy Filter and Local Outlier Factor," *IEEE Transactions on Intelligent Vehicles*, vol. 9, no. 1, pp. 2077–2093, Jan. 2024, doi: 10.1109/TIV.2023.3312654.
- [20] F. Chang, D. Li, and Z. Dong, "Block Optimization and Outlier Detection in GEO SAR Turbulent Tropospheric Error Compensation," *IEEE Geoscience and Remote Sensing Letters*, vol. 20, pp. 1–5, 2023, doi: 10.1109/LGRS.2023.3259399.
- [21] Z. Li, Y. Zhao, X. Hu, N. Botta, C. Ionescu, and G. H. Chen, "ECOD: Unsupervised Outlier Detection Using Empirical Cumulative Distribution Functions," *IEEE Trans. Knowl. Data Eng.*, vol. 35, no. 12, pp. 12181–12193, Dec. 2023, doi: 10.1109/TKDE.2022.3159580.
- [22] X. Jin, H. Q. Wang, and Z. H. Jin, "Anomaly detection of satellite telemetry data based on extended dominant sets clustering," *J. Phys. Conf. Ser.*, vol. 2489, no. 1, p. 012036, May 2023, doi: 10.1088/1742-6596/2489/1/012036.
- [23] F. Cravidão, J. P. Monteiro, P. Gil, and R. Ventura, "Phase-Folded KDE Anomaly Detection for Satellite Telemetry," 2026. doi: 10.2139/ssrn.6438828.
- [24] A. J. J. Cespedes, B. H. B. Pangestu, A. Hanazawa, and M. Cho, "Performance Evaluation of Machine Learning Methods for Anomaly Detection in CubeSat Solar Panels," *Applied Sciences (Switzerland)*, vol. 12, no. 17, Sep. 2022, doi: 10.3390/app12178634.
- [25] R. Aleshko, A. Sakhnik, A. Vasiliev, V. Berezovsky, and K. Shoshina, "Development of a methodology for identifying anomalies in telemetry data of small spacecraft using the ensemble method," *E3S Web of Conferences*, vol. 583, p. 04011, Oct. 2024, doi: 10.1051/e3sconf/202458304011.
- [26] A. Aslan and S. B. Çelebi, "ANOMALY DETECTION IN SATELLITE TELEMETRY USING DEEP LEARNING MODELS," May 2025.
- [27] M. Anderson, J. Muesing, K. Cahoy, and K. Center, "Ensemble Learning for Autonomous Onboard Satellite Fault Diagnosis with Validation Tool," in *38th Annual Small Satellite Conference*, Logan, UT, 2024. doi: https://doi.org/10.15142/dbas-0w70.
- [28] M. Petkovic, L. Lucas, T. Stepisnik, P. Panov, N. Simidjievski, and D. Kocev, "Discovering outliers in the Mars Express thermal power consumption patterns," in *2021 IEEE 8th International Conference on Space Mission Challenges for Information Technology (SMC-IT)*, IEEE, Jul. 2021, pp. 31–37. doi: 10.1109/SMC-IT51442.2021.00011.
- [29] W. Jin, B. Sun, Z. Li, L. Zhang, and S. Zhang, "Analytical Investigation of Anomaly Detection Methods based on Time-Domain Features and Autoencoders in Satellite Power Subsystem," in *Proceedings - 2021 8th International Conference on Dependable Systems and Their Applications, DSA 2021*, Yinchuan, China: Institute of Electrical and Electronics Engineers Inc., Aug. 2021, pp. 451–460. doi: 10.1109/DSA52907.2021.00068.
- [30] H. Xia and T. Meng, "A versatile feature selection learning method for satellite attitude control system fault diagnosis with limited data," *Neurocomputing*, vol. 619, Feb. 2025, doi: 10.1016/j.neucom.2024.129216.
- [31] W. Gan, L. Chen, S. Wan, J. Chen, and C. M. Chen, "Anomaly Rule Detection in Sequence Data," *IEEE Trans. Knowl. Data Eng.*, vol. 35, no. 12, pp. 12095–12108, Dec. 2023, doi: 10.1109/TKDE.2021.3139086.

- [32] D. Wang, J. Zhang, Y. Wang, D. Chang, W. Wang, and X. Cui, "Energy data anomaly detection based on association rule data mining," in *2022 7th International Conference on Intelligent Computing and Signal Processing, ICSP 2022*, Institute of Electrical and Electronics Engineers Inc., 2022, pp. 565–570. doi: 10.1109/ICSP54964.2022.9778824.
- [33] X. Duan, Z. Xiao, Y. Liu, Z. Li, Q. Zhu, and S. Lang, "Data relative currency repair and anomaly detection based on rules," *Expert Syst. Appl.*, vol. 243, Jun. 2024, doi: 10.1016/j.eswa.2023.122923.
- [34] T. Yairi, Y. Kato, and K. Hori, "Fault Detection by Mining Association Rules from House-keeping Data," in *Canadian Space Agency*, Quebec: St-Hubert, Jun. 2001. Accessed: Oct. 16, 2024. [Online]. Available: https://robotics.estec.esa.int/i-SAIRAS/isairas2001/papers/Paper_AS012.pdf
- [35] H. Wang, J. Yu, D. Tang, D. Han, L. Tian, and J. Dai, "Mining diagnostic knowledge from spacecraft data based on Spark cluster," in *Conference on Industrial Electronics and Applications, ICIEA 2020*, Institute of Electrical and Electronics Engineers Inc., Nov. 2020, pp. 1762–1767. doi: 10.1109/ICIEA48937.2020.9248423.
- [36] S. Kang, L. Yang, Y. Song, R. Zhou, and J. Pang, "Satellite Power System State Prediction Based on Online Learning With Parameter Association Rules," *IEEE Trans. Instrum. Meas.*, vol. 72, 2023, doi: 10.1109/TIM.2023.3291798.
- [37] CelesTrak, "Active Satellites," CelesTrak. Accessed: Oct. 10, 2025. [Online]. Available: <https://celestrak.org/NORAD/elements/table.php?GROUP=active&FORMAT=tle>
- [38] H. Zhao, S. Qiu, J. Yang, J. Guo, M. Liu, and X. Cao, "Satellite Early Anomaly Detection Using an Advanced Transformer Architecture for Non-Stationary Telemetry Data," *IEEE Transactions on Consumer Electronics*, vol. 70, no. 1, pp. 4213–4225, Feb. 2024, doi: 10.1109/TCE.2024.3357881.
- [39] Y. Du *et al.*, "A novel predictive algorithm for double difference observations of obstructed BeiDou geostationary earth orbit (GEO) satellites," *Advances in Space Research*, vol. 63, no. 5, pp. 1554–1565, Mar. 2019, doi: 10.1016/j.asr.2018.11.009.
- [40] N. Ferdiansyah, I. S. Sitanggang, Mushthofa, and M. Mukhayadi, "Hybrid Deep Learning Approach for Satellite's Battery Capacity Prediction," in *2025 IEEE International Conference on Aerospace Electronics and Remote Sensing Technology (ICARES)*, IEEE, Oct. 2025, pp. 1–7. doi: 10.1109/ICARES67579.2025.11371568.
- [41] V. Knap *et al.*, "Extending battery life in CubeSats by charging current control utilizing a long short-term memory network for solar power predictions," *J. Power Sources*, vol. 618, p. 235164, Oct. 2024, doi: 10.1016/j.jpowsour.2024.235164.
- [42] J. I. Minow, "Spacecraft Anomalies and Failures Workshop 2023: NASA Introductory Comments," Greenbelt, MD, Mar. 2023. Accessed: Oct. 17, 2024. [Online]. Available: <https://ntrs.nasa.gov/citations/20230003611>
- [43] A. Fejjari, A. Delavault, R. Camilleri, and G. Valentino, "A Review of Anomaly Detection in Spacecraft Telemetry Data," *Applied Sciences*, vol. 15, no. 10, p. 5653, May 2025, doi: 10.3390/app15105653.
- [44] A. Diro, S. Kaisar, A. V. Vasilakos, A. Anwar, A. Nasirian, and G. Olani, "Anomaly detection for space information networks: A survey of challenges, techniques, and future directions," *Comput. Secur.*, vol. 139, p. 103705, Apr. 2024, doi: 10.1016/j.cose.2024.103705.
- [45] S. Bhowmik, B. Jelfs, S. P. Arjunan, and D. K. Kumar, "Outlier removal in facial surface electromyography through Hampel filtering technique," in *Life Sciences Conference (LSC)*, Sydney: IEEE, Dec. 2017, pp. 258–261. doi: 10.1109/LSC.2017.8268192.
- [46] C. Leys, C. Ley, O. Klein, P. Bernard, and L. Licata, "Detecting outliers: Do not use standard deviation around the mean, use absolute deviation around the median," *J. Exp. Soc. Psychol.*, vol. 49, no. 4, pp. 764–766, Jul. 2013, doi: 10.1016/j.jesp.2013.03.013.
- [47] M. A. Ivanushkin, S. S. Volgin, I. V. Kaurov, and I. S. Tkachenko, "Analysis of statistical methods for outlier detection in telemetry data arrays, obtained from 'aIST' small satellites," in *Journal of Physics: Conference Series*, Institute of Physics Publishing, Oct. 2019. doi: 10.1088/1742-6596/1326/1/012029.
- [48] R. Dawson, "How Significant Is A Boxplot Outlier?," *Journal of Statistics Education*, vol. 19, no. 2, 2011, doi: <http://dx.doi.org/10.1080/10691898.2011.11889610>.
- [49] M. Hubert and E. Vandervieren, "An adjusted boxplot for skewed distributions," *Comput. Stat. Data Anal.*, vol. 52, no. 12, pp. 5186–5201, Aug. 2008, doi: 10.1016/J.CSDA.2007.11.008.
- [50] P. J. Rousseeuw and M. Hubert, "Anomaly detection by robust statistics," *Wiley Interdiscip. Rev. Data Min. Knowl. Discov.*, vol. 8, no. 2, Mar. 2018, doi: 10.1002/widm.1236.
- [51] S. Banerjee, A. Ghosh, A. Dey, and P. Pal, "Data Discovery in Images Using a Median-Absolute-Deviation Based Filter," *ITM Web of Conferences*, vol. 54, p. 01007, 2023, doi: 10.1051/itmconf/20235401007.
- [52] A. Koppula, A. Jain, S. Ganguly, L. V. Chaitanya Srinivas, and K. S. Sridharan, "Issues with the Use of the Hampel Filter for the Recovery of Event-Related Dynamics from Electrophysiological Signals," in *Asian Conference on Innovation in Technology, ASIANCON 2023*, Institute of Electrical and Electronics Engineers Inc., 2023. doi: 10.1109/ASIANCON58793.2023.10270545.
- [53] B. Iglewicz and D. C. Hoaglin, *How to Detect and Handle Outliers*, 1st ed., vol. 16. ASQC Quality Press, 1993. Accessed: May 22, 2024. [Online]. Available: https://books.google.co.id/books/about/How_to_Detect_and_Handle_Outliers.html?id=silnAQAAIAAJ&redir_esc=y
- [54] M. Wu *et al.*, "A fault detection method of electric vehicle battery through Hausdorff distance and modified Z-score for real-world data," *J. Energy Storage*, vol. 60, p. 106561, Apr. 2023, doi: 10.1016/J.EST.2022.106561.
- [55] J. W. Tukey, "Exploratory Data Analysis," Jan. 1977.
- [56] R. Swami, M. Dave, and V. Ranga, "IQR-based approach for DDoS detection and mitigation in SDN," *Defence Technology*, vol. 25, pp. 76–87, Jul. 2023, doi: 10.1016/J.DT.2022.10.006.

- [57] I. Souiden, M. N. Omri, and Z. Brahmi, "A survey of outlier detection in high dimensional data streams," May 01, 2022, *Elsevier Ireland Ltd.* doi: 10.1016/j.cosrev.2022.100463.
- [58] X. Sun, S. Guo, S. Liu, J. Guo, and B. Du, "Unsupervised anomaly detection of nuclear power plants under noise background based on convolutional adversarial autoencoder combining self-attention mechanism," *Nuclear Engineering and Design*, vol. 428, Nov. 2024, doi: 10.1016/j.nucengdes.2024.113493.
- [59] J. Mbuya, D. Pfoser, and A. Anastasopoulos, "Trajectory Anomaly Detection with Language Models," Sep. 2024, doi: 10.1145/3678717.3691257.
- [60] D. Freedman and P. Diaconis, "On the histogram as a density estimator:L2 theory," *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, vol. 57, no. 4, pp. 453–476, 1981, doi: 10.1007/BF01025868.
- [61] J. M. D. La Rubia, "Rice University Rule to Determine the Number of Bins," *Open J. Stat.*, vol. 14, no. 01, pp. 119–149, Feb. 2024, doi: 10.4236/ojs.2024.141006.
- [62] D. D. Saurette, A. Biswas, R. J. Heck, A. W. Gillespie, and A. A. Berg, "Determining minimum sample size for the conditioned Latin hypercube sampling algorithm," *Pedosphere*, vol. 34, no. 3, pp. 530–539, Jun. 2024, doi: 10.1016/j.pedsph.2022.09.001.
- [63] H. A. Sturges, "The Choice of a Class Interval," *J. Am. Stat. Assoc.*, vol. 21, no. 153, pp. 65–66, Mar. 1926, doi: 10.1080/01621459.1926.10502161.
- [64] R. Agrawal, T. Imieliński, and A. Swami, "Mining association rules between sets of items in large databases," *ACM SIGMOD Record*, vol. 22, no. 2, pp. 207–216, Jun. 1993, doi: 10.1145/170036.170072.
- [65] N. E. Oweis, M. M. Fouad, S. R. Oweis, S. S. Owais, and V. Snasel, "A Novel Mapreduce Lift Association Rule Mining Algorithm (Mrlar) for Big Data," *International Journal of Advanced Computer Science and Applications*, vol. 7, no. 3, 2016, doi: 10.14569/IJACSA.2016.070321.
- [66] P. J. Azevedo and A. M. Jorge, "Comparing Rule Measures for Predictive Association Rules," in *Machine Learning: ECML 2007*, Berlin, Heidelberg: Springer Berlin Heidelberg, 2007, pp. 510–517. doi: 10.1007/978-3-540-74958-5_47.
- [67] Z. Makani, S. Arora, and P. Kanikar, "A Parallel Approach to Combined Association Rule Mining," *Int. J. Comput. Appl.*, vol. 62, no. 15, pp. 7–13, Jan. 2013, doi: 10.5120/10154-5004.
- [68] H. Lan, X. Ma, L. Ma, and W. Qiao, "Pattern investigation of total loss maritime accidents based on association rule mining," *Reliab. Eng. Syst. Saf.*, vol. 229, Jan. 2023, doi: 10.1016/j.ress.2022.108893.
- [69] R. J. Bayardo, "Brute-Force Mining of High-Confidence Classification Rules," in *International Conference on Knowledge Discovery and Data Mining*, in KDD'97. Austin, Texas: AAAI Press, Aug. 1997, pp. 123–126. Accessed: Sep. 19, 2024. [Online]. Available: <https://www.bayardo.org/ps/kdd97.pdf>
- [70] I. K. Mwangi, L. Nderu, R. W. Mwangi, and D. G. Njagi, "Hybrid interpretable model using roughset theory and association rule mining to detect interaction terms in a generalized linear model," *Expert Syst. Appl.*, vol. 234, Dec. 2023, doi: 10.1016/j.eswa.2023.121092.
- [71] N. Pasquier, Y. Bastide, R. Taouil, and L. Lakhal, "Discovering Frequent Closed Itemsets for Association Rules," in *Lecture Notes in Computer Science*, vol. 1540, Springer, Berlin, Heidelberg, 1999, pp. 398–416. doi: 10.1007/3-540-49257-7_25.
- [72] S. M. Lundberg and S.-I. Lee, "A Unified Approach to Interpreting Model Predictions," in *Advances in Neural Information Processing Systems*, I. Guyon, U. Von Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett, Eds., Curran Associates, Inc., 2017. [Online]. Available: https://proceedings.neurips.cc/paper_files/paper/2017/file/8a20a8621978632d76c43dfd28b67767-Paper.pdf
- [73] A. Barredo Arrieta *et al.*, "Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI," *Information Fusion*, vol. 58, pp. 82–115, Jun. 2020, doi: 10.1016/j.inffus.2019.12.012.
- [74] N. Meenakshisundaram and G. Sajiv, "Interpreting Cervical Cancer Risk Predictions Using SHAP, LIME, and Counterfactual Explanations: A Human-in-the-Loop Explainable AI Framework," in *2025 International Conference on Communication, Computer, and Information Technology (IC3IT)*, IEEE, Oct. 2025, pp. 1–5. doi: 10.1109/IC3IT66137.2025.11341693.
- [75] A. Barbado, Ó. Corcho, and R. Benjamins, "Rule extraction in unsupervised anomaly detection for model explainability: Application to OneClass SVM," *Expert Syst. Appl.*, vol. 189, p. 116100, Mar. 2022, doi: 10.1016/j.eswa.2021.116100.

BIOGRAPHY OF AUTHORS



Rizki Permala    received the bachelor's degree in electrical engineering from Gadjah Mada University in 2006 and the master's degree in computer science from IPB University, Indonesia, in 2019. Currently, he is pursuing his doctoral studies (Ph.D.) in computer science at IPB University, Indonesia. He is currently a research engineer at Research Center for Satellite Technology, National Research & Innovation Agency (BRIN), Indonesia. His research interests include data mining, machine learning, databases, and software development. He can be contacted at email: rizki.permala@brin.go.id



Imas Sukaesih Sitanggang    is a Professor at the School of Data Science, Mathematics, and Informatics, IPB University, Bogor. She received a Ph.D. in computer science from the Faculty of Computer Science and Information Technology, Universiti Putra Malaysia, in 2013. She has been a Postgraduate Coordinator, Computer Science Department, IPB University since 2019. Her primary research interests are spatial data mining and smart agriculture. She can be contacted at email: imas.sitanggang@apps.ipb.ac.id



Hendra Rahmawan    received his B.Sc. from the IPB University, Indonesia, in 2004. He received his M.Eng. and Ph.D. degrees from the Bandung Institute of Technology, Indonesia, in 2009 and 2018, respectively. Currently, he is a lecturer in the Computer Science Department, School of Data Science, Mathematics, and Informatics, IPB University, Bogor. His current research interests include high-performance computing, computer architecture, networking, and parallel computation. He can be contacted at hrahmawan@apps.ipb.ac.id



Wahyudi Hasbi    received the bachelor's degree in physics from Hasanuddin University in 2000, the master's degree in computer science from IPB University, Indonesia, in 2011, and the Ph.D. degree from the Institut für Luft-und Raumfahrt, Technische Universität Berlin, in 2020. He is currently a Senior Researcher and a Head of Research Center for Satellite Technology, National Research & Innovation Agency (BRIN), Indonesia. He was involved as a System and a Chief Engineer in microsatellite projects in Indonesia, including LAPAN-TUBSAT, LAPAN-A2/LAPAN-ORARI, and LAPAN-A3/IPB. He is actively involved in research and has published articles related to satellite technology. He can be contacted at email: wahyudi.hasbi@brin.go.id