

A Hybrid Model and Data Driven Approach for Anomaly Detection in Space Power Systems

M. A. Carbone and M. J. Muscatello and J. A. McCormick and J. T. Csank
NASA Glenn Research Center, Cleveland, OH, 44135, U.S.A.

C. A. Teubert and J. D. Frank
NASA Ames Research Center, Moffett Field, CA, 94035, U.S.A.

The future of crewed deep space exploration requires advanced system awareness and monitoring techniques to ensure the reliability of critical subsystems such as the electric power system (EPS). Currently, expert systems are used to detect and isolate large abrupt faults such as short circuits are sufficient to safe the power system. More nuanced and subtle fault behavior is analyzed by ground control, where corrective and troubleshooting actions can be made remotely or by the onboard crew. To ensure the safety of the crew, onboard computers are tasked with the detection and diagnosis of faults and disturbances onboard the vehicle. Spacecraft experience additional failure modes on a regular basis after launch that were unplanned, unexpected, or ambiguous, and NASA has identified improved software as a means of lowering the number of failures onboard spacecraft. This paper proposes a novel combination of model-based and data-driven techniques to monitor and identify known normal, faulty, and anomalous behavior. The method is demonstrated against a high-fidelity simulation of the spacecraft EPS and validated against real hardware.

I. Introduction

NASA is developing new technology to ensure that the crew on a mission to Mars have the tools and information necessary to respond to on-board problems without the rapid response from a large number of ground support resources [1], [2]. Communication delays for Mars missions are expected to take up to 44 minutes round trip, making traditional Earth-based response impractical [3]. To reduce the workload of necessary functions performed by crew, new technologies in fault detection and diagnosis (FDD) are needed to assist and guide fault recovery and restorative actions [4].

Previous efforts in onboard spacecraft power system fault detection has focused on rules-based expert systems and failure modes and effects analysis (FMEA) on vehicles such as the International Space Station (ISS) [5]. When suspicious behavior occurs in the power system and the FMEA is unable to identify the cause or location of the failure, Earth-based ground control performs a diagnostic study to understand and resolve the issue [6]. Spacecraft experience failure modes on a regular basis after launch that were unplanned, unexpected, or ambiguous, and NASA has identified improved software as a means of lowering the number of failures onboard spacecraft [7]. More modern developments include the Advanced Caution and Warning System (ACAWS) [8], [9], based on the Testability Engineering and Maintenance System (TEAMS) [10], that rely on real-time expert system and knowledge-based FDD [11]. ACAWS was tested on telemetered flight data from Orion test flight EFT-1 [12] and is being matured for use on Gateway [4].

In anticipation for future crewed missions to Mars, missions beyond low earth orbit will have greater communication delays, higher complexity, and lower risk tolerance and thus will require more advanced FDD and controls capability. In this paper we focus on the application of the electrical power system (EPS) onboard a spacecraft; however, the techniques presented can apply to many other systems. The number of research papers applied to onboard spacecraft subsystems is limited, but many research efforts in terrestrial systems have been conducted. In recent years, machine learning (ML) for fault classification have been used to assist or enhance more traditional model-based approaches found in many power system applications. Learning methods assist in reducing the uncertainty in solely model-based approaches. In [13] Random Forests and eXtreme Gradient Boosting (XGB) are used to detect and diagnose faults in wind turbines. Chen et. al. used an XGB-based algorithm to analyze fault transients to predict post-fault stability in a 39-bus AC power system [14]. In [15] Fuzzy C-means (FCM) clustering is combined with neural networks and support vector machines to provide classification of large data sets. Power transmission line faults are diagnosed using decision tree and Random Forest (RF) classifier. Combinations of model-based and ML approaches have been explored as well,

in particular, using the Kalman filter (KF) and its variants to perform power system dynamic state estimation [16]. High impedance faults (HIF) in AC power distribution feeders are detected using the extended Kalman filter (EKF) and neural networks in [17] and EKF with support vector machines (SVM) in [18]. In [19] an adaptive unscented Kalman filter (UKF) is used to generate features for a wavelet neural network for the application of soft fault diagnosis in analog circuits. Sensor anomalies in autonomous vehicles are detected using EKF and SVM in [20].

While these more recent efforts have been successful in diagnosing subtle faults, they are often limited to a narrow number of fault types, offering little "coverage" or number of fault modes capable of being detected and isolated. The problem of identifying unanticipated faults (anomalies) is an ongoing challenge in systems engineering in particular, due to the fact that unanticipated faults can cause misdiagnosis and result in erroneous corrective actions. Air vehicle anomalies were detected using particle filters in [21]. Anomalies in autonomous underwater vehicles are detected using online topic models in [22]. Accommodation of unanticipated faults using the Kalman filter and a feed-back learning system is shown in [23]. A bank of EKFs and the generalized likelihood ratio test is used to detect unanticipated faults in a grinding mill circuit [24]. For a review on fault tolerant control systems, see [25].

This work builds on the fault detection and diagnosis (FDD) design proposed in [26] that automatically generated banks of Kalman filters for fault detection and diagnosis in Direct Current (DS) space power systems. In this paper we propose an extension to the previous design that uses a hybrid model-based and ML-based FDD tool designed for spacecraft EPS. This approach provides better fault detection through adaptive threshold monitoring and more accurate fault classification. The unique contributions of this paper include (1) a high-fidelity spacecraft power system simulation capable of generating large quantities of nominal and faulty data, (2) an improved FDD method for the application of spacecraft subsystems using a combination of model-based and data-driven techniques, and (3) the ability to distinguish between known and unknown behaviors and identify unanticipated conditions (anomalies). Although the case study use in this paper is a spacecraft EPS distribution system the methodology presented can be applied to other industries requiring high levels of autonomous FDD.

The remainder of this paper is organized as follows: Section II formulates and outlines the proposed FDD method. Section III demonstrates the performance of the method applied to a DC electric power system. Lastly, Section IV provides conclusions and remarks about future areas of research.

II. Methodology

The detection and diagnosis of faults can be achieved using a model-based architecture shown in Fig. 1. Here, several different measured variables can be used to extract information of possible faults. These changes can be detected using a process model of the system. The extraction of different features (e.g. residuals, state variables, parameters, etc.) can be compared with their known values under various conditions and lead to fault diagnosis. In this paper we propose a data-driven layer using clustering and classification machine learning tools to achieve improved change detection and classification using known information about the behavior of the system to enhance the existing process model-based approach.

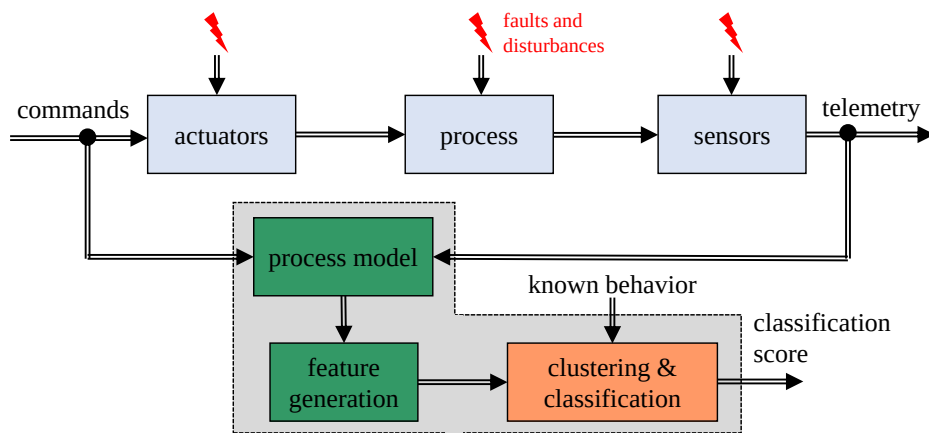


Fig. 1 Block diagram of fault detection using process models

A. Process Model: Kalman Filter

Process models are a valuable resource to track the dynamic states of a system and provide analytical redundancy for FDD. The innovations-based framework proposed by [27], [28], and [29], is adopted in this paper, relying on the innovations of the Kalman filter. This paper will investigate linear time invariant systems; however, extensions to non-linear systems exist (see EKF and UKF [30, 31]).

To construct the model of the system, consider the discrete-time, linear, multi-input multi-output (MIMO) process

$$x_{(t+1)} = \Phi x_{(t)} + G u_{(t)} + \Gamma w_{(t)} \quad (1)$$

$$z_{(t)} = H x_{(t)} + v_{(t)} \quad (2)$$

where $x_{(t)}$ is an $n \times 1$ vector of state variables (subscript t indicates the time instant), $u_{(t)}$ is a $p \times 1$ vector of control variables, $w_{(t)}$ is a $q \times 1$ vector of random variables, $z_{(t)}$ is an $m \times 1$ vector of output measurements, Φ is an $n \times n$ state transition matrix, G is an $n \times p$ input distribution matrix, and H is an $m \times n$ output matrix. Γ is an $n \times q$ noise distribution matrix, which modifies the $q \times 1$ vector of random variables $w_{(t)}$ and adds to the state vector. Then, $v_{(t)}$ is an $m \times 1$ vector of random measurement errors which adds directly to the output measurements $z_{(t)}$. The random vectors $w_{(t)}$ and $v_{(t)}$ are assumed to be independent Gaussian white noise sequences with known mean and covariance matrices.

The Kalman filter is the best linear mean-square estimator of the state vector $x_{(t)}$ given the input variables $u_{(t)}$ and the measured output variables $z_{(t)}$. If it is assumed that all system parameters and statistics are known exactly, the innovation sequence of the Kalman filter can be generated in the following form

$$\bar{x}_{(t+1|t)} = \Phi[\bar{x}_{(t|t-1)} + K_{(t)}v_{(t)}] + G_{(t)}u_{(t)} + \Gamma_{(t)}\bar{w}_{(t)} \quad (3)$$

$$\hat{x}_{(0|-1)} = x_{(0)} \quad (4)$$

$$v_{(t)} = z_{(t)} - H_{(t)}\hat{x}_{(t|t-1)} - \bar{v}_{(t)} \quad (5)$$

$$K_{(t)} = P_{(t|t-1)}H_{(t)}^T(H_{(t)}P_{(t|t-1)}H_{(t)}^T + R_{(t)})^{-1} \quad (6)$$

$$P_{(t+1|t)} = \Phi_{(t)}P_{(t)}\Phi_{(t)}^T + \Gamma_{(t)}Q_{(t)}\Gamma_{(t)}^T \quad (7)$$

$$P_{(t|t)} = (I - K_{(t)}H_{(t)})P_{(t|t-1)} \quad (8)$$

where $P_{(t|j)}$ is the error covariance matrix of $\hat{x}_{(t|j)}$ such that $P_{(t|j)} = E\{(x_{(t)} - \hat{x}_{(t|j)})(x_{(t)} - \hat{x}_{(t|j)})^T\}$, and $K_{(t)}$ is the $n \times m$ Kalman gain matrix [32]. The innovation sequence $v_{(t)}$ is a vector representation of the difference between the system output measurements and model output estimates. The innovation sequence, or residuals, make for a convenient measure for fault detection and diagnosis as $v_{(t)}$ should be near zero when no faults in the system are present. Fault detection and diagnosis can take place when changes to the innovation sequence occur, corresponding to changes in the system behavior, parameters, and measured or unmeasured states. Known faults may be classified by the unique magnitude and direction of the innovation vector.

1. Feature Generation: Innovations

A convenient feature of the Kalman filter residuals is that when the process and model parameter matrices match, then the statistics of the innovation sequence are well known. That is, the innovation sequence $v_{(t)}$ is a zero mean Gaussian white noise sequence with covariance equal to

$$\text{cov}(v_{(t)}) = (H_{(t)}P_{(t|t-1)}H_{(t)}^T + R_{(t)}). \quad (9)$$

For fault isolation, it is more useful to consider the Standardized Innovation Sequence

$$\eta_{(t)} = (H_{(t)}P_{(t|t-1)}H_{(t)}^T + R_{(t)})^{-1/2}v_{(t)} \quad (10)$$

where $(\cdot)^{-1/2}$ represents the square root of the inverse of a matrix so that

$$E\{\eta_{(t)}\eta_{(j)}^T\} = I\delta_{(t,j)} \quad (11)$$

where I denotes the identity matrix and $\delta_{(t,j)}$ denotes the Kronecker delta.

The known statistics of the standardized innovation sequence simplifies the fault hypothesis testing for detecting faults such as sensor bias, noisy measurements, changes in the level of noise, changes in the system parameters, and

more. These faults can be identified using the tests for whiteness via the autocorrelation function, c_k , of a stationary process for lag $k = 1, 2, \dots$ is,

$$c_k = \frac{1}{N} \sum_{i=k}^N (\eta_i - \hat{\eta})(\eta_{i-k} - \hat{\eta})^T \quad (12)$$

where $\hat{\eta}$ denotes the mean of η_i , and $\hat{\eta}$ represents the sample mean. Under the null hypothesis, $\hat{c}_k$ is asymptotically independent and normal with zero mean and covariance equal to I/N . The tests of mean simply determine whether the innovation sequence is zero mean or not. The mean of the innovation sequence is estimated by

$$\hat{\eta} = \frac{1}{N} \sum_{i=1}^N \eta_i. \quad (13)$$

where N is the sample size and $\bar{\eta}$ is the true mean. The null hypothesis states that $\hat{\eta}$ has a Gaussian distribution with zero mean and covariance

$$E\{\hat{\eta}\hat{\eta}^T\} = I/N. \quad (14)$$

Next, a test of covariance is performed to check if the measurements are independent. The covariance of the innovation sequence is estimated by

$$\hat{c}_0 = \frac{1}{N} \sum_{i=1}^N (\eta_i - \hat{\eta})(\eta_i - \hat{\eta})^T. \quad (15)$$

Under the null hypothesis, $\hat{c}_0$ has a Wishart Distribution and the trace of $\hat{c}_0$ has a Chi-Squared distribution with $(N-1)r$ degrees of freedom, where r is the number of observed measurements. Therefore, $\hat{c}_0$ can be tested for its null hypothesis covariance equal to the identity matrix, I .

B. Clustering & Classification

The model-based approach described above paired with static fault detection thresholds may not work reliably as EPS interactions become more complicated or as sensors age and degrade. Therefore, a machine learning classifier should provide a more robust diagnostic method. A learned classifier should adapt to minor changes that can cause variations in the response to the signal output during abnormal conditions. This work proposes an ensemble learning process to achieve highly efficient and robust FDD that is adaptive to minor changes in the feature vector. A combination of FCM and XGB classifiers are used to learn information about the power system behavior. The implementation combines a practical way to cluster features paired with a supervised classifier for improved accuracy. Below we outline the methods used in our approach.

1. Fuzzy C-means (FCM)

Dunn introduced the first fuzzy generalization of the hard c-means algorithm in [33] with Bezdek providing important improvements [34]. Today, tens of thousands of research papers have been published related to fuzzy cluster analysis and applications [35]. The objective of any clustering algorithm is to partition a set of N data points based on some variable or feature of the data [36]. Let $X = \{X_i | i = 1 \dots c\}$ be the set of samples or feature vectors. The FCM aims to group these points into a set of C clusters, where C is an integer $2 \leq C \leq n$. A fuzzy partition of X can be modeled by matrix $U = [u_{ij}]$ be a $C \times n$ matrix with the following conditions:

- 1) No fuzzy subset is 'empty', at least one sample has partial membership in each cluster

$$\sum_{j=1}^n u_{ij} > 0 \quad \forall i \quad (16)$$

- 2) No fuzzy subset is all of X

$$\sum_{j=1}^n u_{ij} < n \quad \forall i \quad (17)$$

- 3) The sum of the membership values for each X_j is one

$$\sum_{i=1}^C u_{ij} = 1; \quad \forall j \quad (18)$$

The FCM algorithm performs an iterative optimization to determine the minimum of an objective function which determines the membership of the data set. The objective function, $J(U, \rho)$, describes the value, or quality, of the fuzzy partitioning U of the feature vectors into C clusters where ρ is the tuple $(\rho_1, \rho_2, \dots, \rho_c)$.

$$J(U, \rho) = \sum_{i=1}^c \sum_{j=1}^n (u_{ij})^m (d_{ij})^2 \quad (19)$$

ρ_i is the cluster center of class i , $1 \leq i \leq C$, $m \in [1, \infty)$ is an exponential weighting factor called the *fuzzifier*, and $(d_{ij})^2$ is the distance from a feature point X_j ,

$$d_{ij}^2 = \|X_j - \rho_i\|^2. \quad (20)$$

Then for a fixed number of clusters, $2 \leq C \leq n$, where n equals the number of feature items and fixed m , $1 < m < \infty$ the cluster centers ρ_i can be calculated as

$$\rho_i = \frac{\sum_{j=1}^n (u_{ij})^m x_k}{\sum_{j=1}^n (u_{ij})^m}, \quad (21)$$

and the new membership values for feature item j are

$$u_{ij} = \frac{1}{\sum_{k=1}^C (d_{ij}/d_{ik})^{2/(m-1)}}. \quad (22)$$

The FCM algorithm performs an iterative optimization of J which generates a fuzzy partition of the data set X . To run the algorithm, select a fixed number of clusters C . There is no theoretical method for selecting the exponential weighting factor, m , but $1.1 \leq m \leq 5$ is recorded as the most useful range. The algorithm as outlined above iterates until convergence is reached when $\|U^t - U^{(t+1)}\| < \epsilon$ for a specified tolerance level ϵ . Then a sufficient optimal set of clusters C have been defined each with membership value u_{ij} .

2. eXtreme Gradient Boosting (XGB)

Extreme Gradient Boosting, or XGB, proposed by Chen and Guestrin in 2016, is a highly scalable supervised machine learning algorithm that uses parallel tree boosting to find patterns in datasets based on labeled data or features. The algorithm has shown state-of-the-art results in a wide range of applications and standard benchmarks [37]. The fundamental concept behind boosting is to combine a set of weak classifiers with low accuracy to build a strong classifier with better performance. Gradient boosting is achieved by taking the gradient direction of the loss function to iteratively generate a new tree, thus correcting the mistakes made by the previous one. The traditional gradient boosting decision tree (GBDT) algorithm relies on first-order derivative information, however XGB uses a second-order Taylor expansion of the loss function to find the optimal solution to the model and avoid over-fitting [38]. The fundamental process of XGB is shown as follows.

Let a model with K decision trees and a classification result, $\hat{y}_i$, be defined as,

$$\hat{y}_i = \sum_{k=1}^K f_k(x_i), \quad f_k \in F, \quad (23)$$

where f_k is an independent regression tree. Regression trees differ from decision trees in that each tree contains a continuous score on each of the leaves, ω . To define the set of equations in this learning model, we minimize the loss (objective) function,

$$J = \sum_{i=1}^n l(y_i, \hat{y}_i) + \sum_{k=1}^K \Omega(f_k), \quad (24)$$

where l is the training loss function that measures the difference between predictions, $\hat{y}_i$ and the object y_i . Then,

$$\Omega(f) = \gamma T + \frac{1}{2} \lambda \|\omega\|^2 \quad (25)$$

is an over-fitting term that reduces the complexity of the trees (number of nodes and branches). Here, T and ω are the number and weight of the leaves, respectively. γ is a regularization parameter that controls the complexity complexity of

the tree, and λ penalizes the squared weight of the leaves. Let $\hat{y}_i^t$ be the prediction of the i^{th} instance at the t^{th} iteration. Each iteration t needs to add a new tree f_t to minimize the second-order Taylor expansion of the loss function,

$$J^{(t)} = \sum_{i=1}^n \left[l(y_i, \hat{y}_i^{(t-1)}) + g_i f_t(x_i) + \frac{1}{2} h_i f_t^2(x_i) \right] + \Omega(f_t) \quad (26)$$

where $g_i = \partial_{y_i^{(t-1)}} l(y_i, \hat{y}_i^{(t-1)})$ and $h_i = \partial_{y_i^{(t-1)}}^2 l(y_i, \hat{y}_i^{(t-1)})$ are the first and second order partial derivatives on the loss function l .

Substituting Equation (25) in (26) and removing the constant term in $l(y_i, \hat{y}_i^{(t-1)})$ yields,

$$J^{(t)} = \sum_{i=1}^n \left[g_i f_t(x_i) + \frac{1}{2} h_i f_t^2(x_i) \right] + \gamma T + \frac{1}{2} \lambda \|\omega\|^2 = \sum_{j=1}^T \left[\left(\sum_{i \in I_j} g_i \right) \omega_j + \frac{1}{2} \left(\sum_{i \in I_j} h_i + \lambda \right) \omega_j^2 \right] + \gamma T \quad (27)$$

where I_j represents all the data samples of leaf j . Finding the stationary points of Equation (27) gives, ω_j^* , the optimal weight of leaf j ,

$$\omega_j^* = - \frac{\sum_{i \in I_j} g_i}{\sum_{i \in I_j} h_i + \lambda}. \quad (28)$$

Inserting the optimal weight, ω_j^* , into the objective function gives

$$J^{(t)} = - \frac{1}{2} \sum_{j=1}^T \frac{\left(\sum_{i \in I_j} g_i \right)^2}{\sum_{i \in I_j} h_i + \lambda} + \gamma T. \quad (29)$$

This objective function provides high performance classification and prevents over-fitting. Equation (29) can be used as a scoring function to measure the quality of a tree structure. The model generates many simple trees which are used to score the leaf nodes during splitting. A greedy algorithm is used to add branches to the starting leaf to substitute for enumerating all of the tree structures,

$$J_{split} = \frac{1}{2} \sum_{j=1}^T \left[\frac{\left(\sum_{i \in I_L} g_i \right)^2}{\sum_{i \in I_L} h_i + \lambda} + \frac{\left(\sum_{i \in I_g} g_i \right)^2}{\sum_{i \in I_g} h_i + \lambda} - \frac{\left(\sum_{i \in I} g_i \right)^2}{\sum_{i \in I} h_i + \lambda} \right] - \gamma T. \quad (30)$$

3. KF-FCM-XGB Integration

The careful integration of the methods outlined above are critical to the success of the FDD. The model-based approach using the Kalman filter and statistical tests of the standardized innovations provides an analytical comparison of the measured system to the known behavior of the linearized model as well as noise filtering. The model based method alone is susceptible to modeling inaccuracies, natural disturbances, and lacks adaptive threshold monitoring. To account for these shortfalls, the data driven approach is appended to the output of the model-based system.

Using telemetry from the power system, a combination of clustering and classifier algorithms are used to train a model to detect nominal and faulty power system behaviors. Due to the unsupervised nature of FCM and the unpredictable nature of many faults, integration with a supervised method can improve the accuracy of the cluster identification. Furthermore, classification techniques such as XGB provide important information related to the probability of a given fault condition, $P(f)$. Training the XGB classifier using only the cluster centers of the FCM results in a large reduction of data points, thus improving computational efficiency and simplicity of the classification [15]. In order to train the classifier to detect unknown fault types, a sample of points outside of the existing clusters are generated and labeled during the training process of the XGB. These data points may exist outside of the trained clusters and thus indicate the region in the feature space that does not contain any known faults. The distance to the cluster centers may provide some useful information related to the nature of the fault which may assist with fault isolation and corrective actions. Fig. 2 displays a block diagram of the proposed system. The steps of the FDD method are outlined below.

- 1) Collect raw data and control inputs: $z(t), u(t)$
- 2) KF performs innovation sequence update : $v(t)$

- 3) Compute Standardized Innovation Sequence from KF output $v_{(t)}$: $\eta_{(t)}, \sigma^2$ (features for input to FCM)
- 4) FCM clustering of $\eta_{(t)}, \sigma^2$: extract fuzzy cluster centers ρ_i from final cluster $U = [u_{ij}]$ (input to XGB)
- 5) XGB learns classifier to predict $\hat{y}$ from ρ_i (probability of fault type $P(f)$ given $\eta_{(t)}, \sigma^2, \rho_i$, trained on labeled power system data)

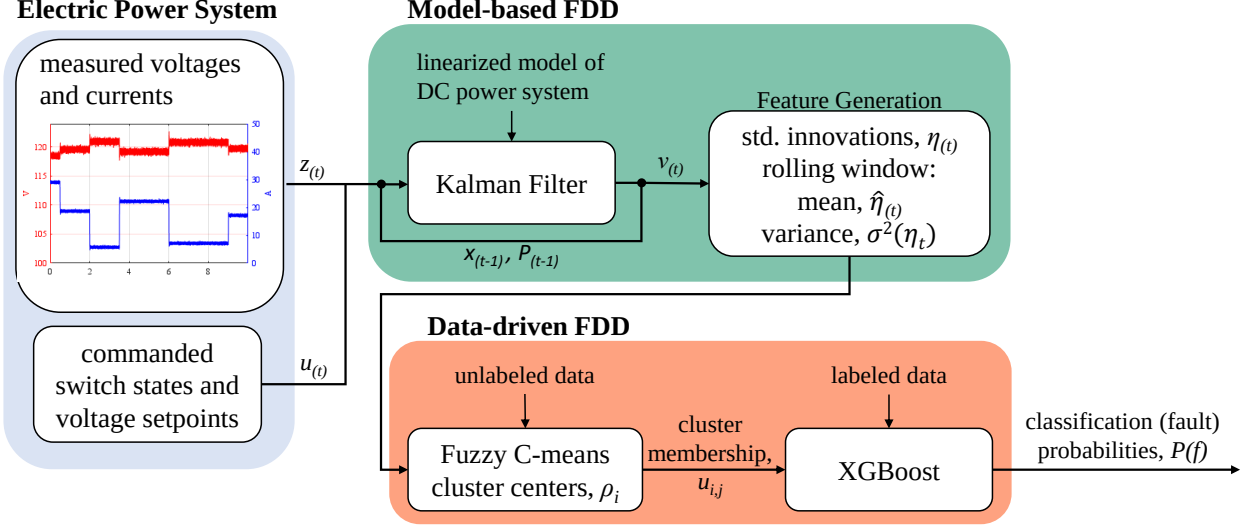


Fig. 2 Flow diagram of proposed methodology for fault detection and diagnosis

III. Case Study: DC Distribution Feeder

To demonstrate and evaluate the success of the method proposed in this paper a simple example is illustrated. Consider a single distribution feeder of a dc EPS for spacecraft as shown in Fig. 3. Voltage and current measurements are collected from Main Bus Switch Unit (MBSU) 1-3 Remote Bus Isolator (RBI) 2 and Power Distribution Unit (PDU) 1-2.

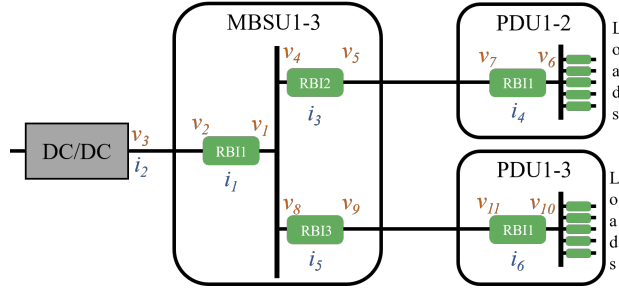


Fig. 3 One line diagram of the test system

This case study will demonstrate the ability of the proposed method to classify a number of known behaviors (faulted and normal) between MBSU1-3 and PDU1-2. Normal operations consist of standard 120 V DC distribution with variable constant power load. Pre-considered fault scenarios include sensor offset (sensor output offset by constant value), excessive sensor noise, and failed switch (switch commanded closed but is physically open). Training files with randomly generated load profiles and uniformly random fault magnitudes were simulated. To train the ML model, 400 files (consisting of 50,000 data points in each) were generated with randomized changes in the load and fault magnitude. To evaluate the performance of the method, an additional 800 test files were generated. Although this paper demonstrates a simple example across one distribution feeder, this architecture is designed to be expanded to larger systems such as the one shown in [26].

A. Criteria for Performance Evaluation

To measure the performance of the FDD method a confusion matrix captures the classification results of each timestep of the rolling window. Here the true class is evaluated against the predicted class of the KF-FCM-XGB. The following metrics are useful to evaluate a confusion matrix performance [39]: (1) precision, the ratio of true positive (TP) predictions to total number of positive predictions, (2) recall, the ratio of true positive predictions to the sum of TP and false negative (FN) instances, (3) F_1 score, the harmonic mean of precision and recall, (4) and accuracy, the ratio of correct instances to total instances. These parameters can be calculated as follows:

$$\text{Precision} = \frac{TP}{TP + FP} \quad (31)$$

$$\text{Recall} = \frac{TP}{TP + FN} \quad (32)$$

$$F_1 = \frac{2 \cdot \text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}} \quad (33)$$

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (34)$$

where TN and FP are the true negative and false positive instances respectively.

B. Implementation

In this work Matlab/Simulink 2023b is used to build a high-fidelity model of the spacecraft EPS. The simulation uses Thevenin equivalent models of the individual power system components to generate dynamics at 5 kHz, and can run faster than real time. Faults are able to be programmed and set into the model at random times with random fault magnitudes. More details about the model can be found in [40]. Synchronized data is sampled and sent to the central controller, then the standardized innovations and statistics are stored as features in time series CSV files. Training and testing of the ML models used python version 3.8.10, scikit-learn 1.3.2, numpy 1.24.3, scikit-fuzzy 0.4.2, and xgboost 2.0.3.

C. Process Model & Feature Generation

Using the DC nodal equations described in our previous work [26], applied to Fig. 3, the linear time-invariant model for the distribution line can be written as follows. Let the state vector $x = [x_1, x_2, x_3, x_4] = [v_1, v_3, v_6, v_{10}]$ represents the estimated voltage at each bus. Then measurement vector $z = [v_1, v_2, i_1, v_3, i_2, v_4, v_5, i_3, v_6, v_7, i_4, v_8, v_9, i_5, v_{10}, v_{11}, i_6]$ is the synchronized measured voltages and currents available to the controller.

Next let the measurement matrix equal

$$H = \begin{bmatrix} 1 & 1 & Y_{12} & 0 & -Y_{12} & 1 & 1 & Y_{13} & 0 & 0 & -Y_{13} & 1 & 1 & Y_{14} & 0 & 0 & -Y_{14} \\ 0 & 0 & -Y_{12} & 1 & Y_{12} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & -Y_{13} & 1 & 1 & Y_{13} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & -Y_{14} & 1 & 1 & Y_{14} \end{bmatrix}^T \quad (35)$$

where Y_{ij} represents the admittance between bus i and j . Then let $Q = I_{4 \times 4}$, $P_0 = I_{4 \times 4}$, and $A = I_{4 \times 4}$. Sensor noise is modeled as a Gaussian process with standard deviation, $\sigma_v = 0.01$ for voltage sensors, and $\sigma_i = 0.24$ for current sensor. Then σ_v^2 and σ_i^2 are used to populate the diagonal elements of the noise distribution matrix, R . Now the standardized innovations, mean, and covariance can be computed for the system as described in Section II. When calculating the statistics of the innovations η , a moving window of samples N , was set at 100.

To improve the efficiency and accuracy of the FDD method a heuristic feature selection method is employed. Here we select the residual sample mean and variance of current at MBSU1-3 RBI2 i_3 , $\hat{\eta}_8$ and $\sigma^2(\eta_8)$, respectively, and the residual sample mean of output voltage at MBSU1-3 RBI2 v_5 to be our selected features.

The FDD method is tested against 800 files of known behavior with randomly varying fault magnitude. First, Fig. 4 column one shows the signals under normal conditions and a failed switch fault condition. There are clear changes to the three selected features, making this fault type easy to detect and classify. Next, Fig. 4 column two shows a combination of signals with varying constant power load (0.48-3.84kW) load under normal and sensor bias fault on sensor i_3 . In this figure, the residual sample mean $\hat{\eta}_7$ and $\hat{\eta}_8$ and residual sample variance $\sigma^2(\eta_8)$ are shown. From these plots, it

appears that the sensor bias can be differentiated from the normal condition using the magnitude of $\hat{\eta}_8$. Similarly, Fig. 4 column three shows the same signals under normal conditions and a sensor noise fault on sensor i_3 . In this instance, the magnitude of $\sigma^2(\eta_8)$ can be used to distinguish normal from a sensor noise fault condition. In this scenario the distribution feeder switch MBSU1-2 RBI2 has failed ‘open’ while the measured state reads ‘closed’. Observations show that the magnitude of $\hat{\eta}_7$ is the main distinguishing factor during this condition.

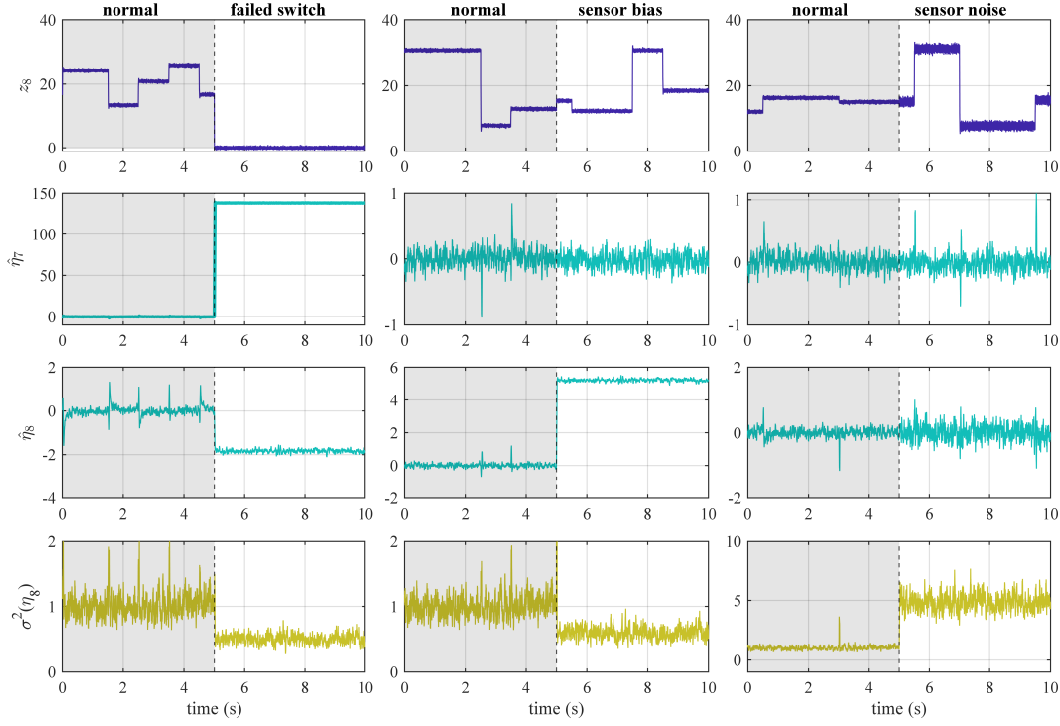


Fig. 4 Features of the power system during normal and faulted operation. Top row: raw measurements from the simulation sensors, second row: sample mean of the residual 7, third row: sample mean of the residual 8, fourth row: sample variance of the residual 8.

D. Semi-Supervised Machine Learning: FCM-XGB

While the statistics of the Kalman filter innovations are well known, variations in the system may occur over time, making threshold based FDD undesirable. The implementation of the FCM-XGB semi-supervised ML approach provides adaptive fault detection thresholds that may adjust to changes in the spacecraft over the mission life.

The FCM is trained on the three features identified to be most responsive to the fault conditions, $\hat{\eta}_8$, $\hat{\eta}_7$, and $\sigma^2(\eta_8)$. The unsupervised clustering method classifies the unlabeled data into $C = 4$ clusters (one for each of three faults, plus nominal), and exponential weighting factor (fuzzifier) is set to $m = 1.1$. The FCM clusters for the test data set are shown in Fig. 5a, and cluster centers are shown in Table 1. The unsupervised clustering method is able to group the test data into four regions resembling the four known conditions (one normal and three faults). However, the FCM fails to identify the threshold between the normal condition and low magnitude region of the sensor noise and sensor bias faults. As we will see in the performance evaluation section, this impacts the location of the cluster centers and reduces the performance of fault classification using FCM alone.

In order to improve the results of the change detection function, the method is modified into a semi-supervised ML method. Here the cluster centers are used to train a supervised XGB algorithm. We set the number of decision trees, $K = 149$, the number of leaves $T = 291$, and the max depths of the tree equal to 9. Tuning the XGB classifier is a complex task, however, these values were shown to provide good performance in this application. The training data set is used to generate the features, then the FCM groups the data into four clusters. Those cluster centers are then used to train the XGB decision tree with feedback from the labeled data. Fig. 5b displays the classification results of the 800 test files. Each test file contains the simulated current and voltage telemetry, Kalman filter standardized residual, as well as

Table 1 Cluster centroids (ρ_i) generated by the FCM

Cluster	Centroid $\hat{\eta}_8$	Centroid $\hat{\eta}_7$	Centroid $\sigma^2(\eta_8)$
C_1	-0.011	2.479	1.264
C_2	-0.009	0.001	13.92
C_3	-0.042	41.71	0.495
C_4	138.2	-1.839	0.499

the rolling window mean and variance of the standardized residuals. Each simulation is run for 10 seconds and sampled at 5 kHz. The high frequency was used to speed up the simulation time since the time-invariant nature of the simulations does not impact the dynamic behavior of the system. Each file either has no fault (normal) or contains a single fault and randomized fault magnitude. The task of diagnosing more than one fault is beyond the scope of this work. It is evident

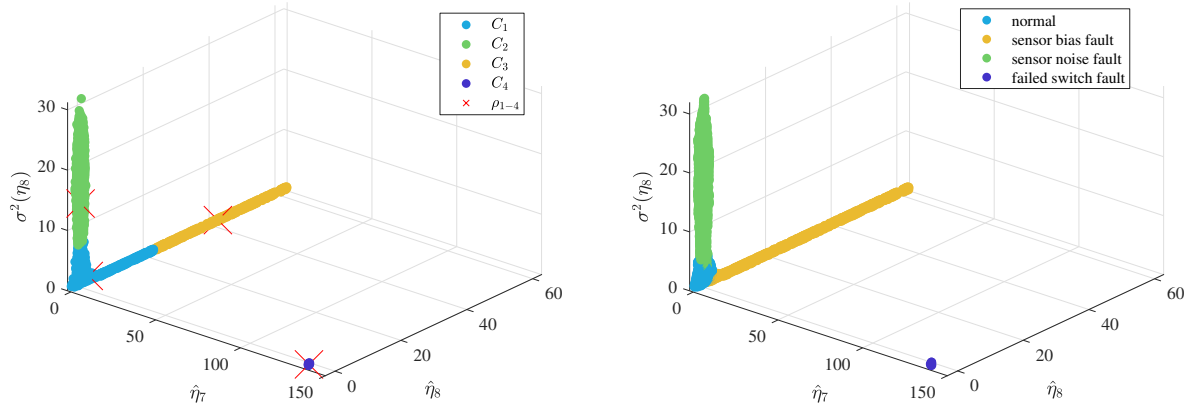


Fig. 5 (a) Fuzzy clusters identified by the KF-FCM algorithm, (b) Improved classification of the test files of the KF-FCM approach with supervised learning through the XGB method

that the XGB improves the classification accuracy, in particular, for low fault magnitudes such as the sensor bias.

E. Anomaly Detection: High Impedance Fault

In order to observe the behavior of the FDD method when an unanticipated fault is introduced, we use a HIF, which is not one of the faults used to generate training data. The increased resistance of the HIF results in a low fault current, often accompanied by random, unpredictable, and unstable dynamic fluctuations in the amplitude and harmonic levels of the fault current. As a result, HIF often go undetected by traditional protection methods and therefore new FDD methods are needed to address these faults. It is critical to detect and isolate short circuit faults in order to prevent further damage to the power system components such as generators, converters, loads, and energy storage. Failure to properly isolate these faults can lead to blackouts in the power system.

Some research in terrestrial power systems has been performed. In [41] discrete wavelet transform and K-nearest neighbor-based classifiers are used on local current measurements to detect HIFs on DC naval shipboards. Bayati et. al. used a support vector machine (SVM) to detect HIFs in DC microgrids [42]. In [43], a superconducting fault current limiter is used to detect and isolate both low and high impedance faults in DC microgrids. Although research in terrestrial grids has been conducted to detect these faults, we will utilize this special case of low magnitude fault as a case study in anomaly or unanticipated fault detection.

A HIF is used to test the proposed FDD method in response to an unanticipated fault. To test this, a series of HIFs of varying magnitude are passed through the KF-FCM-XGB method without pre-training the models on this type of

fault. The condition for determining an unanticipated fault is a maximum classification probability $\max(P(f)) > P_k$, where P_k is the probability threshold for classification of a known condition.

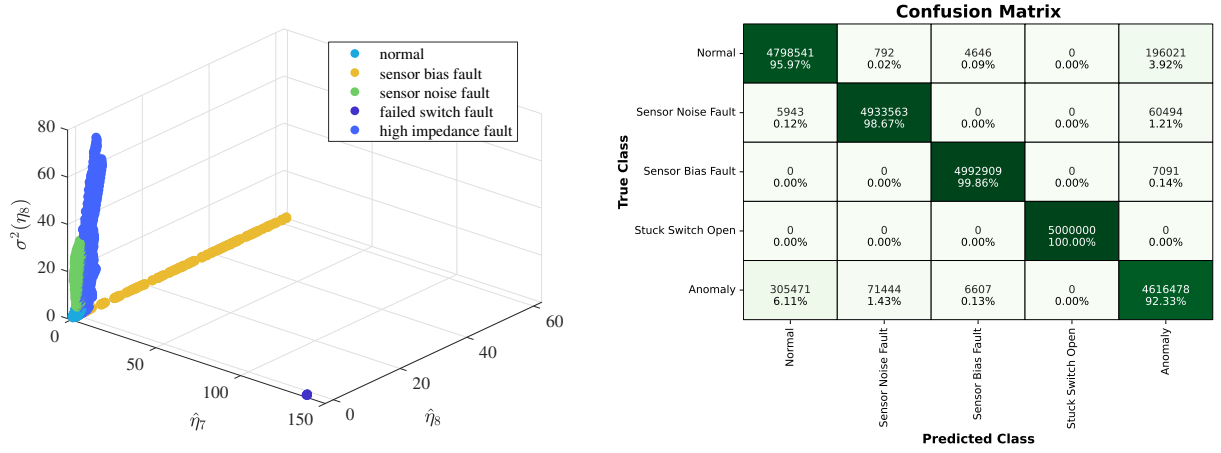


Fig. 6 (a) Classifier results of the KF-FCM-XGB with known faults and anomalies of varying magnitude, (b) Confusion matrix classification results of the test data

F. Performance Evaluation

The features of the HIF data set is shown in Fig. 6a. This set of unlabeled anomaly data is processed through the KF-FC-XGB algorithm for classification. The results of the test data through the classifier is captured in Fig. 6b. Here each data sample is classified and recorded in the confusion matrix. The model was most successful in classifying failed switch open faults, due to the large, abrupt change in feature. Sensor noise and bias faults were also highly successful, however low magnitude faults were shown to be more difficult to detect. The overall accuracy of the classifier is 97.4%, with an F-1 score = 0.974, the harmonic mean of the precision and recall of the model. The combination of methods proposed in this paper improved the classification accuracy compared to a limit threshold implementation of the Kalman Filter approach with accuracy and F-1 score = 0.921. Also, the KF-FCM approach without supervised feedback from the XGB was shown to have accuracy and F-1 score = 0.899. These results demonstrate the value of the supervised learning approach as well as the importance of careful feature selection.

IV. Conclusion

In this paper a novel method for fault detection and diagnosis of faults in dc electric power systems is developed. The technique uses features extracted from the innovations of the Kalman filter and its statistical properties. The change detection functionality is achieved using a ensemble combination of FCM clustering and XGB classification. Using a high-fidelity model of a spacecraft electrical power system to generate both nominal and faulty conditions, the fault detection and diagnosis strategy was evaluated. The sample mean and variance of the voltage and current measurements at MBSU1-2 RBI2 were selected as the features the machine learning model would be trained on. The FCM computed the cluster centers and membership values that the XGB classifier would use to identify the behavior of the power system. The model was trained on 400 test scenarios with 50,000 data samples of normal and faulted events. The scenarios are generated with random fault magnitudes and load profiles. The success of the method was evaluated on 800 test files where the results are captured in the confusion matrix. The success of the method was also tested in the application of anomaly detection, where an unknown/untrained fault was introduced into the system. Here a high impedance fault was simulated and processed through the FDD tool to produce an decision about the current state of the power system. The combination of model-based and machine learning-based methods provides a strengthened approach towards FDD and AR, making it a good candidate for other autonomous system applications.

To improve this method, a full spacecraft power system with many of fault types should be investigated to determine the scalability of the solution, as well as the testing on real spacecraft power system hardware. Other future work includes the implementation of validation and verification process of ML based algorithms for spaceflight.

Acknowledgments

The authors would like to thank NASA's Exploration Systems Development Mission Directorate, the Mars Campaign Office, and the Earth Independent Operations domain, including Andres Martinez, Alonso Vera, and Ian Maddox for supporting this work.

References

- [1] Maddox, I. D., Vera, A., and Martinez, A., *Development of Earth Independent Operations Technologies for NASA's Mars Campaign Office*, AIAA, 2024, Chap. 4869, pp. 1–12. <https://doi.org/10.2514/6.2024-4869>, URL <https://arc.aiaa.org/doi/abs/10.2514/6.2024-4869>.
- [2] Aaseng, G., Frank, J., Iatauro, M., Knight, C., Levinson, R., Ossenfort, J., Scott, M., Sweet, A., Csank, J., Soeder, J., et al., "Development and testing of a vehicle management system for autonomous spacecraft habitat operations," *AIAA Space Forum*, 2018, p. 5148. <https://doi.org/10.2514/6.2018-5148>.
- [3] Frank, J., Spirkovska, L., McCann, R., Wang, L., Pohlkamp, K., and Morin, L., "Autonomous mission operations," *2013 IEEE Aerospace Conference*, IEEE, 2013, pp. 1–20. <https://doi.org/10.1109/AERO.2013.6496927>.
- [4] Badger, J. M., Strawser, P., and Claunch, C., "A Distributed Hierarchical Framework for Autonomous Spacecraft Control," *2019 IEEE Aerospace Conference*, 2019, pp. 1–8. <https://doi.org/10.1109/AERO.2019.8742199>.
- [5] International Space Station Alpha Program, N. A., and Administration, S., "Failure Detection, Isolation, and Recovery (FDIR)," *Architecture Design Document (ADD) D684-10504*, 1994.
- [6] Crook, D. W., "DCSU 2A Power On Reset - channel failure," *International Space Station (ISS) Anomaly Log*, 2007, pp. 1–3.
- [7] Prokop, L., "Considerations for Software Fault Prevention and Tolerance," *NASA Engineering Safety Center Technical Bulletin*, Vol. 1, No. 23-06, 2023. <https://doi.org/NTRSDOCID20230013383>.
- [8] Aaseng, G., Barszcz, E., Valdez, H., and Moses, H., "Scaling Up Model-Based Diagnostic and Fault Effects Reasoning for Spacecraft," *Proceedings of the AIAA Conference on Space Operations*, 2015, pp. 1–21.
- [9] Aaseng, G., and Frank, J., "Transitioning Autonomous Systems Technology Research to a Flight Software Environment," *Proceedings of the AIAA Conference on Space Operations*, 2016, pp. 1–21.
- [10] Mathur, A., Deb, S., and Pattipati, K., "Modeling and Real-Time Diagnostics in TEAMS-RT," *Proceedings of the American Control Conference*, 1998, pp. 1–5.
- [11] McCann, R. S., Spirkovska, L., and Smith, I., "Putting integrated systems health management capabilities to work: Development of an advanced caution and warning system for next-generation crewed spacecraft missions," *AIAA Modeling and Simulation Technologies Conference*, 2013, pp. 1–15.
- [12] Aaseng, G. B., "Scaling up model-based diagnostic and fault effects reasoning for spacecraft," *AIAA Space 2015 Conference and Exposition*, 2015, p. 4465.
- [13] Zhang, D., Qian, L., Mao, B., Huang, C., Huang, B., and Si, Y., "A Data-Driven Design for Fault Detection of Wind Turbines Using Random Forests and XGboost," *IEEE Access*, Vol. 6, 2018, pp. 21020–21031. <https://doi.org/10.1109/ACCESS.2018.2818678>.
- [14] Chen, M., Liu, Q., Chen, S., Liu, Y., Zhang, C.-H., and Liu, R., "XGBoost-Based Algorithm Interpretation and Application on Post-Fault Transient Stability Status Prediction of Power System," *IEEE Access*, Vol. 7, 2019, pp. 13149–13158. <https://doi.org/10.1109/ACCESS.2019.2893448>.
- [15] Karlik, B., "The positive effects of fuzzy c-means clustering on supervised learning classifiers," *International Journal of Artificial Intelligence and Expert Systems (IJAE)*, Vol. 7, 2016, pp. 1–8.
- [16] Zhao, J., Gómez-Expósito, A., Netto, M., Mili, L., Abur, A., Terzija, V., Kamwa, I., Pal, B., Singh, A. K., Qi, J., et al., "Power system dynamic state estimation: Motivations, definitions, methodologies, and future work," *IEEE Transactions on Power Systems*, Vol. 34, No. 4, 2019, pp. 3188–3198. <https://doi.org/10.1109/TPWRS.2019.2894769>.
- [17] Samantaray, S., Dash, P., and Upadhyay, S., "Adaptive Kalman filter and neural network based high impedance fault detection in power distribution networks," *International Journal of Electrical Power & Energy Systems*, Vol. 31, No. 4, 2009, pp. 167–172. <https://doi.org/https://doi.org/10.1016/j.ijepes.2009.01.001>, URL <https://www.sciencedirect.com/science/article/pii/S0142061509000027>.

- [18] Samantaray, S., and Dash, P., "High impedance fault detection in distribution feeders using extended kalman filter and support vector machine," *European Transactions on Electrical Power*, Vol. 20, No. 3, 2010, pp. 382–393.
- [19] Xu-sheng, G., Wen-ming, G., Zhe, D., and Wei-dong, L., "Research on WNN soft fault diagnosis for analog circuit based on adaptive UKF algorithm," *Applied Soft Computing*, Vol. 50, 2017, pp. 252–259. <https://doi.org/https://doi.org/10.1016/j.asoc.2016.11.012>, URL <https://www.sciencedirect.com/science/article/pii/S1568494616305816>.
- [20] Wang, Y., Masoud, N., and Khojandi, A., "Real-Time Sensor Anomaly Detection and Recovery in Connected Automated Vehicle Sensors," *IEEE Transactions on Intelligent Transportation Systems*, Vol. 22, No. 3, 2021, pp. 1411–1421. <https://doi.org/10.1109/TITS.2020.2970295>.
- [21] Zhang, B., Sconyers, C., Byington, C., Patrick, R., Orchard, M., and Vachtsevanos, G., "Anomaly detection: A robust approach to detection of unanticipated faults," *2008 International Conference on Prognostics and Health Management*, 2008, pp. 1–8. <https://doi.org/10.1109/PHM.2008.4711445>.
- [22] Raanan, B.-Y., Bellingham, J., Zhang, Y., Kemp, M., Kieft, B., Singh, H., and Girdhar, Y., "Detection of unanticipated faults for autonomous underwater vehicles using online topic models," *Journal of Field Robotics*, Vol. 35, No. 5, 2018, pp. 705–716. <https://doi.org/https://doi.org/10.1002/rob.21771>, URL <https://onlinelibrary.wiley.com/doi/abs/10.1002/rob.21771>.
- [23] Farrell, J., Berger, T., and Appleby, B., "Using learning techniques to accommodate unanticipated faults," *IEEE Control Systems Magazine*, Vol. 13, No. 3, 1993, pp. 40–49. <https://doi.org/10.1109/37.214943>.
- [24] Olivier, L. E., and Craig, I. K., "Model-based fault-tolerant control with robustness to unanticipated faults**This work is based on the research supported in part by the National Research Foundation of South Africa (Grant Number 90533)." *IFAC-PapersOnLine*, Vol. 50, No. 1, 2017, pp. 1157–1162. <https://doi.org/https://doi.org/10.1016/j.ifacol.2017.08.401>, URL <https://www.sciencedirect.com/science/article/pii/S2405896317307486>, 20th IFAC World Congress.
- [25] Zhang, Y., and Jiang, J., "Bibliographical review on reconfigurable fault-tolerant control systems," *Annual reviews in control*, Vol. 32, No. 2, 2008, pp. 229–252.
- [26] Carbone, M. A., and Loparo, K. A., "Fault Detection and Diagnosis in Spacecraft Electrical Power Systems," *Journal of Aerospace Information Systems*, Vol. 20, No. 6, 2023, pp. 308–318. <https://doi.org/10.2514/1.I011136>, URL <https://doi.org/10.2514/1.I011136>.
- [27] Mehra, R. K., and Peschon, J., "An Innovations Approach to Fault Detection and Diagnosis in Dynamic Systems," *Automatica*, Vol. 4, 1971, pp. 637–640. [https://doi.org/https://doi.org/10.1016/0005-1098\(71\)90028-8](https://doi.org/https://doi.org/10.1016/0005-1098(71)90028-8).
- [28] Isermann, R., *Fault-diagnosis systems: an introduction from fault detection to fault tolerance*, Springer Science & Business Media, 2006.
- [29] Huang, S., Tan, K. K., Er, P. V., and Lee, T. H., *Intelligent Fault Diagnosis and Accommodation Control*, CRC Press, 2020. <https://doi.org/https://doi.org/10.1201/9780429263880>.
- [30] Julier, S. J., and Uhlmann, J. K., "New extension of the Kalman filter to nonlinear systems," *Signal processing, sensor fusion, and target recognition VI*, Vol. 3068, Spie, 1997, pp. 182–193.
- [31] Wan, E., and Van Der Merwe, R., "The unscented Kalman filter for nonlinear estimation," *Proceedings of the IEEE 2000 Adaptive Systems for Signal Processing, Communications, and Control Symposium (Cat. No.00EX373)*, 2000, pp. 153–158. <https://doi.org/10.1109/ASSPCC.2000.882463>.
- [32] Kalman, R. E., and Bucy, R. S., "New results in linear filtering and prediction theory," *Journal of basic engineering*, Vol. 83, No. 1, 1961, pp. 95–108. <https://doi.org/https://doi.org/10.1115/1.3658902>.
- [33] Dunn, J. C., "A fuzzy relative of the ISODATA process and its use in detecting compact well-separated clusters," *Journal of Cybernetics*, 1973, pp. 32–57.
- [34] Bezdek, J. C., *Pattern recognition with fuzzy objective function algorithms*, Springer Science & Business Media, 2013.
- [35] Ruspini, E. H., Bezdek, J. C., and Keller, J. M., "Fuzzy Clustering: A Historical Perspective," *IEEE Computational Intelligence Magazine*, Vol. 14, No. 1, 2019, pp. 45–55. <https://doi.org/10.1109/MCI.2018.2881643>.
- [36] Cannon, R. L., Dave, J. V., and Bezdek, J. C., "Efficient Implementation of the Fuzzy c-Means Clustering Algorithms," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. PAMI-8, No. 2, 1986, pp. 248–255. <https://doi.org/10.1109/TPAMI.1986.4767778>.

- [37] Chen, T., and Guestrin, C., “XGBoost: A scalable tree boosting system,” *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*, 2016, pp. 785–794.
- [38] Xie, J., Li, Z., Zhou, Z., and Liu, S., “A Novel Bearing Fault Classification Method Based on XGBoost: The Fusion of Deep Learning-Based Features and Empirical Features,” *IEEE Transactions on Instrumentation and Measurement*, Vol. 70, 2021, pp. 1–9. <https://doi.org/10.1109/TIM.2020.3042315>.
- [39] Fürst, K., Chen, P., and Gu, I. Y.-H., “Hierarchical LSTM-Based Classification of Household Heating Types Using Measurement Data,” *IEEE Transactions on Smart Grid*, Vol. 15, No. 2, 2024, pp. 2261–2270. <https://doi.org/10.1109/TSG.2023.3296020>.
- [40] Carbone, M., Muscatello, M., Collazo Fernandez, X., and Csank, J., “Adaptive Coordinated Protection for DC Microgrids with Lunar Surface Power Applications,” *IEEE Transactions on Power Delivery*, 2024, pp. 1–10.
- [41] Subramaniam, K., and Illindala, M. S., “High impedance fault detection and isolation in DC microgrids,” *2019 IEEE/IAS 55th Industrial and Commercial Power Systems Technical Conference (I&CPS)*, IEEE, 2019, pp. 1–8.
- [42] Bayati, N., Balouji, E., Baghaee, H. R., Hajizadeh, A., Soltani, M., Lin, Z., and Savaghebi, M., “Locating high-impedance faults in DC microgrid clusters using support vector machines,” *Applied Energy*, Vol. 308, 2022, p. 118338.
- [43] Xi, J., Pei, X., Song, W., Xiang, B., Liu, Z., and Zeng, X., “Experimental tests of DC SFCL under low impedance and high impedance fault conditions,” *IEEE Transactions on Applied Superconductivity*, Vol. 31, No. 5, 2021, pp. 1–5.