Eksploatacja i Niezawodnosc - Maintenance and Reliability
Volume 29 (2027), Issue 1
https://ein.org.pl
The Polish Maintenance Society

# Hyperparameter optimization of isolation forest using a variable neighborhood search approach for anomaly detection in electric motors

Carlos Mario Garcia-Pena[1], Thalita Monteiro Obal[2], Olga Usuga-Manco[1], Freddy Hernandez-Barajas[3], Fernando Guevara-Carazas[3], Carmen Elena Patino-Rodriguez[1,*]

[1] Universidad de Antioquia, Colombia
[2] Universidade Tecnológica Federal do Paraná, Brazil
[3] Universidad Nacional de Colombia, Colombia
* Corresponding Author: elena.patino@udea.edu.co

*Received: 19 February 2026*
*Revised: 2 March 2026*
*Accepted: 21 April 2026*
*Online: 8 June 2026*



**Abstract**

Anomaly detection in electric motors supports reliability-centered maintenance in electric mobility systems. Isolation Forest (iForest) is commonly used for unsupervised detection of abnormal operating conditions; however, its performance depends on hyperparameter configuration under non-stationary real-world conditions. This study proposes a systematic hyperparameter optimization framework based on Variable Neighborhood Search (VNS) to enhance iForest robustness in BLDC motor monitoring. The proposed VNS–IF–IQR framework integrates structured hyperparameter exploration with adaptive threshold calibration. The methodology is evaluated using 78,886 records from rear-hub BLDC motors operating under real urban conditions in Medellín, Colombia, and validated through a synthetic dataset with controlled anomaly injection. Results show that VNS-based optimization reduces anomaly score variability compared to default configurations. Sensitivity analysis across multiple scenarios indicates marginal hyperparameter influence below 6%, confirming convergence toward broad near-optimal regions and supporting the robustness of the proposed approach.

**Keywords:** *isolation forest, hyperparameter optimization, variable neighborhood search (VNS), anomaly detection, BLDC motors, two-wheeled EVs*

*Article citation:*
Garcia-Pena C M, Obal T M, Usuga-Manco O, Hernandez-Barajas F, Guevara-Carazas F, Patino-Rodriguez C E, Hyperparameter optimization of isolation forest using a variable neighborhood search approach for anomaly detection in electric motors, Eksploatacja i Niezawodnosc – Maintenance and Reliability 2027: 29(1) http://doi.org/10.17531/ein/220967

## Highlights

- A VNS-based framework is used to systematically optimize Isolation Forest hyperparameters under non-stationary operating conditions.
- The VNS–IF–IQR approach reduces anomaly score variability and improves detection stability in BLDC motors without labeled data.
- Validation using real urban data and synthetic anomaly injection confirms robustness and limited hyperparameter sensitivity.

## 1. Introduction

The increasing demand for sustainable mobility has accelerated the adoption of electric vehicles (EVs), driven by advances in energy storage, power electronics, and control systems. Two-wheeled EVs are particularly relevant for urban transport due to low emissions, reduced noise, and high operational efficiency [1–4]. Their performance and reliability depend on rear-hub BLDC motors and rechargeable batteries [5–8], which are exposed to variable loads, thermal stress, and transient events that accelerate degradation [7–13].

In real operating environments, failures do not occur under controlled or labeled conditions. Fault events are infrequent, uncertain in timing, and often undocumented. This makes supervised learning approaches difficult to deploy in reliability-oriented monitoring systems. For industrial assets operating under non-stationary conditions, anomaly detection methods capable of learning directly from unlabeled data are essential. Such methods allow deviations from normal behavior to be identified without prior knowledge of failure signatures, aligning more closely with real reliability assessment scenarios [14–16].

Isolation Forest (*iForest*) is widely used as an unsupervised method for high-dimensional operational data [17–19], yet its performance is highly sensitive to hyperparameter selection. Default or manually tuned configurations often yield

suboptimal detection, limiting reliability in variable-load systems such as e-bikes [10,20,21].

From a reliability perspective, improving hyperparameter stability and detection consistency in unlabeled, real-world data is critical. Methods that approximate actual operational variability, rather than laboratory conditions, provide more realistic foundations for predictive maintenance and decision-making. This justifies the integration of metaheuristic optimization, such as Variable Neighborhood Search (VNS), and estimators like the Interquartile Range (IQR), to maintain discrimination capacity under operational variability—a framework that is further developed in the following sections.

The main scientific contributions of this study are summarized as follows:

1. This work formalizes the problem of hyperparameter instability in *iForest* when applied to non-stationary, real-world operating conditions, highlighting the limitations of default and empirically tuned configurations for reliability-oriented anomaly detection.
2. VNS metaheuristic is systematically employed to optimize key *iForest* hyperparameters, enabling structured exploration of the search space and improving robustness against operating variability, in contrast to conventional single-neighborhood or random tuning approaches.
3. A unified VNS–IF–IQR framework is proposed, integrating metaheuristic optimization with adaptive, unsupervised threshold calibration, preserving the non-parametric nature of anomaly detection while enhancing detection stability.
4. The proposed approach is validated using a synthetic fault injection dataset and a real-world dataset comprising 78,886 records from rear-hub BLDC motors operating under urban conditions, providing empirical evidence of improved anomaly separation and hyperparameter stability.

The remainder of this paper is structured to progressively develop and validate the proposed framework. First, relevant literature on electric vehicle anomaly detection, *iForest*, and hyperparameter optimization is reviewed to position the research gap. The methodological section then describes the data acquisition process, the anomaly detection and thresholding procedures, and the integration of VNS into a unified optimization framework, including a synthetic fault modeling strategy for controlled validation. The results section presents both synthetic and real-world evaluations, followed by a robust analysis of hyperparameter sensitivity under non-stationary operating conditions. Finally, the discussion interprets the findings from a reliability perspective, and the conclusions summarize the practical implications and future research directions.

## 2. Related work

### 2.1. Vehicle electric anomalies state

EVs have gained strategic relevance for sustainable urban mobility due to decarbonization policies and efficiency targets [1,3,22]. Two-wheeled EVs have seen rapid adoption because of low emissions, reduced noise, and high operational efficiency in congested environments, with global sales surpassing 10 million units in 2024 and projections up to 60 million by 2035. China leads this market, followed by Europe and the US [1,3,23].

Two-wheeled EV propulsion systems face stringent requirements due to compact layouts, limited thermal dissipation, frequent start–stop cycles, and transient loads. BLDC motors are critical for performance, reliability, and lifecycle management in these vehicles [7–9,24,25]. Their high torque density, efficiency, and mechanical simplicity make them dominant in e-bikes and scooters. However, lightweight EV operating conditions exacerbate failure mechanisms such as stator insulation breakdown, rotor demagnetization, bearing wear, eccentricity, and inverter faults, producing distinct electrical, vibrational, and electromagnetic signatures [4,7–13,26,27].

Recent trends show a shift from threshold-based diagnostics to data-driven models capable of handling non-stationary conditions. Machine learning and deep learning methods, including CNNs, autoencoders, LSTM, and CNN–LSTM hybrids, have enhanced anomaly detection using current, back-EMF, torque ripple, vibration, and acoustic signals [24,25,28,29]. Data-driven diagnostics for BLDC motors are thus essential to anticipate degradation under real operating conditions [16].

### 2.2. Isolation forest algorithm

*iForest* algorithm, introduced by Liu et al. [19], constitutes a different paradigm within unsupervised anomaly detection. Rather than modeling normality through density estimation or distance metrics, *iForest* explicitly operationalizes the assumption that anomalies are sparse and structurally distinct observations that can be isolated with fewer random partitions. This isolation-based principle shifts the focus from estimating probability distributions to analyzing how data geometry interacts with stochastic partitioning processes [17,18,20,28].

The algorithm constructs an ensemble of randomly generated binary trees, where recursive splits are applied along randomly selected features and split values. The expected path length required to isolate at a point becomes a proxy for its structural rarity. Unlike density-based or distance-based techniques, the anomaly score in *iForest* does not derive from local density comparisons but from the interaction between random recursive partitioning and the intrinsic geometry of the feature space. Consequently, detection behavior is directly influenced by statistical properties such as dimensionality, correlation structure, heteroscedasticity, and distributional shifts [17,30,31].

This structural dependency is particularly relevant in non-stationary environments. Variations in load conditions, transient dynamics, or evolving correlations alter the spatial organization of data in the feature space, thereby modifying the distribution of path lengths used as anomaly scores. In such contexts, the algorithm's response is not merely stochastic but geometrically conditioned. Empirical analyses show that shifts in local density or correlation structures can significantly affect isolation depth distributions, leading to unstable score separation when default configurations are used [18,32].

Hyperparameters such as *n_estimators, max_samples*, and *contamination* assume a structural role rather than a purely operational one. While *n_estimators* controls ensemble stability, the *max_samples* parameter dictates the granularity at which local structures are captured, directly affecting the algorithm's sensitivity to "swamping" and "masking" effects. Furthermore, the *contamination* parameter acts as a static prior for the decision boundary. However, studies report that fixed settings degrade performance in heterogeneous datasets where noise levels and density distributions evolve [20,33]. This underscores a significant research gap: the need for systematic optimization of the partitioning mechanism and adaptive calibration of the decision threshold [34].

In reliability-oriented monitoring, such as BLDC motor diagnostics, these effects are amplified by evolving signal regimes and dynamic coupling among electrical variables [11,13,28]. Fixed configurations often fail to adapt to these changes, resulting in in inconsistent separation between nominal and abnormal regions. To address this, current literature suggests that robustness emerges not from the default *iForest* formulation, but from the alignment between stochastic partitioning and data geometry.

Overall, the literature indicates that the effectiveness of *iForest* in complex environments is not intrinsic to its default formulation but emerges from the alignment between stochastic partitioning and data geometry. Hyperparameters therefore act as structural regulators of this interaction, underscoring the necessity of systematic optimization strategies when applying *iForest* to real-world, non-stationary systems.

### 2.3. Hyperparameters optimization

Hyperparameter optimization is a crucial step in the *iForest* methodology, as it directly influences detection accuracy, robustness, and computational efficiency. Proper tuning of key hyperparameters such as *n_estimators, max_samples, and contamination* is essential to adapt *iForest* to the statistical structure of the input data, particularly in high-dimensional, correlated, and non-stationary signals [17,35].

Several studies have explored advanced optimization techniques for hyperparameter tuning across different algorithms and domains. Bayesian optimization (BO) and sequential model-based approaches have been applied to improve predictive accuracy and computational efficiency in unsupervised and supervised contexts, including random forests, CNNs, LSTM models, and ultra-high-performance concrete monitoring [36–39]. Metaheuristic strategies such as particle swarm optimization (PSO), genetic algorithms (GA), and ant colony optimization (ACO) have been widely employed to automate hyperparameter selection and improve model generalization, convergence, and classification performance [40–43].

More recently, metaheuristic-based frameworks have been

extended beyond deep neural networks to broader machine learning contexts [36–39,44]. Lankford and Grimes developed OpenNAS, an open-source system that integrated PSO and ACO for neural architecture search in image classification, showing higher accuracy on complex datasets [45], while Yuliana et al. applied hyperparameter optimization to random forest models for 5G coverage prediction, improving predictive performance and efficiency [42]. Although most studies focus on supervised or deep learning models, hyperparameter optimization has also been applied to unsupervised anomaly detection algorithms, particularly *iForest*. Strategies include grid search combined with Prediction-Area (P-A) plots for *iForest*, deep autoencoder, and one-class SVM models [33,46], Bayesian optimization across multiple algorithms including *iForest* [47], and evaluations in smart home environments using *iForest*, elliptic envelope, local outlier factor, and one-class SVM, showing that optimized configurations outperform default settings [48].

Despite these advances, metaheuristic-based tuning of *iForest* using the VNS algorithm has not been previously explored. To fill this gap, the present study introduces the VNS-IF-IQR framework, integrating VNS for systematic hyperparameter optimization with *iForest* for anomaly detection. The framework employs an IQR–based adaptive threshold criterion, enhancing the model's ability to distinguish normal from faulty operational states in BLDC motors operating under real conditions. This approach provides a scalable and reliable optimization strategy, aligning hyperparameter selection with the complexity of real-world electromechanical signals, and improving robustness and detection accuracy under variable load and transient conditions.

## 3. Methodology

### 3.1. Data acquisition and signal processing

The experimental platform consisted of two rear-hub BLDC motors installed on lightweight electric vehicles (an e-bike and an e-scooter), operating under real urban traffic conditions. The sensing architecture was designed to capture electrical, thermal, mechanical, and environmental variables commonly reported in electric mobility reliability studies [5–8]. The monitored variables included motor current and voltage, motor temperature, battery voltage (both total and series-level), battery current and charging current, battery temperature, linear acceleration, vibration, angular velocity, and ambient temperature, relative humidity, and barometric pressure. Electrical variables were acquired through shunt-based current sensing modules and instrumentation amplifiers integrated into an embedded data acquisition system based on an Arduino Mega platform. Vibration and acceleration were measured using MEMS accelerometers, while environmental variables were recorded through integrated pressure–temperature–humidity sensors. Thermal measurements were obtained via NTC thermistors and analog temperature sensors placed in direct contact with the motor casing, controller housing, and battery cells to ensure representative thermal readings. The selection of these variables follows established EV instrumentation practices, where motor and battery electrical quantities are considered primary indicators of drivetrain condition and degradation [12,13,15,16].

Sensors were physically distributed to minimize mechanical disturbance and environmental interference. Electrical and current sensing modules were mounted near the battery housing and controller assembly, while environmental sensors were positioned within protected compartments to prevent direct exposure to wind and moisture. Thermal sensors were installed with direct conductive contact to the monitored components to improve measurement fidelity under non-stationary operating conditions.

Data were collected in Medellín, Colombia (approximately 1460 m above sea level) on a near-flat urban circuit (~130 m per lap) without mechanical modification of the vehicles, preserving factory configuration. This ensures that detected anomalies arise from natural operational variability rather than artificially induced structural alterations. Across all experimental campaigns, a total of 78,886 multivariate electrical signal records were obtained from three configurations: Motor A with Battery A, Motor B with Battery A, and Motor B with Battery B, as well as a combined dataset integrating all configurations. Sampling frequency was defined according to the dynamic response of electrical and mechanical signals and the constraints of the embedded acquisition system. All sensor channels were temporally synchronized and, where required, resampled to ensure alignment prior to feature matrix construction.

Prior to anomaly detection, a standardized preprocessing pipeline was applied. Noise attenuation was performed through smoothing and filtering, particularly for current and vibration signals exhibiting high-frequency fluctuations. Incomplete or corrupted records caused by communication interruptions were removed. Sensor streams were temporally aligned, and consistency checks were conducted to detect physically implausible readings. Importantly, no manual filtering of extreme but physically admissible operational values was conducted, preserving the unsupervised nature of the study and avoiding bias toward predefined fault assumptions.

A key methodological clarification concerns the representation of input features. The proposed framework operates on synchronized multivariate operational state vectors constructed directly from instantaneous electrical, thermal, and environmental measurements. Each record corresponds to a structured snapshot of the system's operational condition. No handcrafted fault-specific indicators—such as spectral signatures, harmonic distortion metrics, or manually derived health indices—were introduced. This design choice preserves the non-parametric and assumption-light characteristics of the *iForest* algorithm [19,31,32] and avoids embedding prior knowledge of failure mechanisms into the feature space. Consequently, anomaly detection is performed on statistically structured but minimally engineered multivariate representations of real operating behavior, consistent with recent unsupervised approaches in electrical systems [16,24,49].

Given the heterogeneous physical units of the monitored variables (e.g., volts, amperes, degrees Celsius, acceleration), statistical normalization was applied prior to model training to prevent scale dominance effects. Feature scaling parameters were computed exclusively from the training portion of each configuration to avoid information leakage, and the same transformations were subsequently applied to evaluation data. Although *iForest* is less sensitive to scaling than distance-based methods, normalization improves numerical stability and consistency during hyperparameter optimization in high-dimensional, heterogeneous datasets.

This experimental setup and data processing pipeline ensure reproducibility, methodological transparency, and consistency with established EV monitoring practices, while preserving the unsupervised, non-parametric nature of the proposed VNS–IF–IQR anomaly detection framework.

### 3.2. iForest procedure

*iForest* procedure initiates by constructing a collection of isolation trees, each aimed at separating individual instances in the dataset. For each tree, a feature $q$ randomly selected from the feature set $Q$ of the dataset $X = \{x_{q1}, x_{q2}, x_{q3}, \ldots, x_{qk}\}$. where each subscript indicates the corresponding instance of the feature. A split point $p$ is then chosen uniformly at random within the observed range of the selected feature, $[min\,(q), max\,(q)]$. This split divides the dataset into two partitions, assigning instances with values less than $p$ to the left child node and those greater than or equal to $p$ to the right child node, creating the initial bifurcation of the isolation tree. This mechanism is like the partitioning performed in Binary Search Trees (BST) [19,31,35].

At each subsequent node, the procedure repeats independently: a new feature $q'$ is randomly selected from the subset of features $Q$ and a new split point $p'$ is determined based on the range of values for the instances contained in that node. This recursive partitioning continues until termination criteria are met, which typically includes reaching a single instance per node or attaining a maximum tree depth. Each path from the root to a terminal node represents the number of splits required to isolate a specific instance.

Once all trees in the forest are constructed, the algorithm computes the average path length for each instance across all trees. Instances that are isolated with fewer splits (shorter average path length) are assigned higher anomaly scores, indicating a greater likelihood of being outliers. Conversely, instances requiring longer paths are considered more representative of normal behavior. This mechanism enables a probabilistic ranking of anomalies without explicit assumptions about data distribution. The overall effectiveness of *iForest* is determined by its hyperparameters: *n_estimators* influences ensemble diversity and stability, *max_samples* affect partition granularity, and *contamination* adjusts the sensitivity threshold for anomaly detection.

### 3.3. Interquartile range procedure

Once the anomaly scores are obtained, a threshold is defined to separate anomalous data from non-anomalous data. In this work, we use a statistical thresholding method based IQR applied to

the score distribution. The IQR is determined in Equation (1).

$$IQR = Q3 - Q1 \quad (1)$$

where $Q1$ is first quartile and $Q3$ is the third quartile of the scores generated by the model. With this value, we define the failure threshold ($T$) as shown in Equation (2).

$$T = Q1 - 1.5 * IQR \quad (2)$$

It is important to note that the distribution of anomaly scores and the resulting threshold obtained using the IQR criterion are influenced by the configuration of the model's hyperparameters because theses affect the isolation process and modify both the range and variability of the scores.

### 3.4. Variable neighborhood search

The VNS is a metaheuristic proposed by Mladenović and Hansen that systematically explores different neighborhood structures during the search process [50].

In combinatorial optimization problems, a neighbor solution $x$' of a solution $x$ is obtained by applying a small modification to $x$, such as swapping, inserting, or removing elements in a sequence. Formally, the neighborhood of $x$ is defined in Equation (3).

$$N(x) = \{x' \in S |\ x' \text{ is obtained from } x \text{ by a predefined move}\} \quad (3)$$

where $S$ represents the set of all feasible solutions. Exploring $N(x)$. promotes local improvement and increases the likelihood of escaping local optima through structured neighborhood changes.

The choice of the VNS in this research is motivated by its efficiency in exploring the solution space and by the fact that it does not require the tuning of many parameters. The method relies mainly on a single control parameter, that is the maximum number of neighborhoods ($k_{max}$), to guide its operation.

In general terms, the VNS algorithm begins by generating an initial solution $x$ and defining the maximum number of neighborhood structures ($k_{max}$). The objective function $f(x)$ is evaluated, and the search starts with $k = 1$. For each iteration, a new solution $x$' is generated by applying the shaking procedure to $x$ within the $k$-th neighborhood. Then, a local search is performed starting from $x$' to find a local optimum $x$''. If $f(x'') < f(x)$, the solution $x$ is updated to $x$'' and the process returns to the first neighborhood ($k = 1$); otherwise, the algorithm increases $k$ and explores the next neighborhood. This process is repeated until a stopping criterion is met, such as a maximum number of iterations or convergence tolerance.

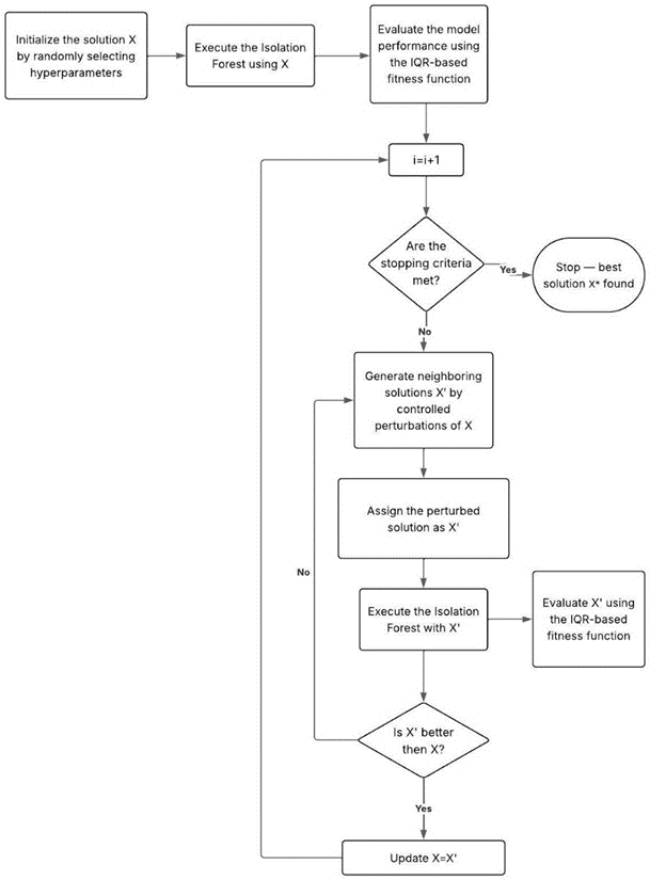


Figure 1. Flowchart of VNS-IF-IQR.

### 3.5. The proposed VNS-IF-IQR methodology

The VNS-IF-IQR methodology is proposed to optimize the hyperparameters of the *iForest* algorithm for enhanced evaluations of unsupervised structural anomaly discrimination under non-stationary operating conditions. The approach integrates the VNS metaheuristic for systematic parameter tuning with IQR-based criterion for adaptive anomaly thresholding. This framework is applied to electrical signals collected from sensors on an electric bicycle motor, aiming to improve structural discrimination. Figure 1 presents an overview of the methodology.

Three hyperparameters are considered in this study: *n_estimators*, *max_samples*, and *contamination*, respectively. These hyperparameters constitute the solution generated by the VNS algorithm, denoted as $X$ in Algorithm 1 and expressed in Equation (4):

$$X = [n.of\ estimators;\ max\ sample;\ contamination] \quad (4)$$

For each hyperparameter, a discrete set of possible values is defined as follows (equations (5), (6) and (7)), respectively for

the number of trees, the sample size per tree, and the expected proportion of anomalies.

$$X_{\square}^{E} = \{set\ of\ n.of\ possible\ estimators\ value\} \quad (5)$$

$$X_{\square}^{MS} = \{set\ of\ n.of\ possible\ max\ sample\ values\} \quad (6)$$

$$X_{\square}^{C} = \{set\ of\ n.of\ possible\ contamination\ values\} \quad (7)$$

$X_{\square}^{E}$, $X_{\square}^{MS}$ and $X_{\square}^{C}$ are respectively with dimension $e, ms$ and $c$.

To perform this optimization, a metaheuristic based on the VNS is employed. The procedure begins by randomly selecting values for each hyperparameter from their respective predefined sets: $X^{E}, X^{MS}$ and $X^{C}$; providing the initial solution $X$. For each proposed solution $X$, the *iForest* algorithm is executed, and the anomalies are labeled.

The real-world dataset does not include synchronized maintenance logs or expert-labeled fault intervals. Therefore, the proposed VNS–IF–IQR framework is formulated as an unsupervised anomaly-screening methodology. Detected anomalies correspond to statistically isolated observations within the learned feature space and are not assumed to represent confirmed mechanical failures. Diagnostic validation against verified fault events is outside the scope of the present study and is identified as future work.

### 3.6. Synthetic Fault Modeling Strategy

Synthetic fault injection is introduced to provide controlled ground-truth perturbations for methodological validation under known deviation conditions. Although the VNS–IF–IQR framework is intended for deployment under real non-stationary operating conditions, methodological assessment requires evaluation under controlled scenarios where anomaly intervals are explicitly defined. Since real urban datasets lack reliable fault annotations or maintenance logs, a synthetic fault modeling strategy was implemented to provide internal validation of the optimization criterion and adaptive thresholding mechanism prior to real-world experimentation.

The synthetic dataset was designed to emulate the fundamental oscillatory behavior observed in electrical signals of BLDC motors. In electromechanical systems, voltage and current waveforms are primarily governed by rotational dynamics and electromagnetic induction, producing signals characterized by a dominant fundamental frequency component [51,52]. Even under load variability, these signals retain a quasi-periodic structure, with anomalies typically manifesting as localized amplitude variations, transient distortions, or phase deviations [53,54]. The $t$-th value of the $i$-th time series $X$ is expressed by $X_i^t$ and can be obtained as shown Equations (8).

$$X_i^t = \begin{cases} \sin\big(2\pi f(t-t_0)\big) + \Lambda\varepsilon_i, & \text{with probability 0.5,} \\ \cos\big(2\pi f(t-t_0)\big) + \Lambda\varepsilon_i, & \text{with probability 0.5,} \end{cases} \quad (8)$$

where the frequency of the waves is $f \in \left\{\frac{1}{80\pi}, \frac{1}{100\pi}\right\}$, the time delay is $t_0$, the noise level is $\Lambda$ and $\varepsilon_i \sim N(0,1)$ and the contamination level is $\delta$.

The temporal location and magnitude of injected perturbations were explicitly defined, enabling the establishment of ground-truth anomaly segments. This controlled configuration allows systematic evaluation of whether the VNS-based hyperparameter optimization improves separation between baseline and perturbed signal distributions, and whether the IQR-based adaptive threshold correctly discriminates injected anomalies without supervised labels.

The synthetic modeling framework therefore establishes internal validity by isolating methodological effects from environmental variability. External validation under real urban operating conditions is addressed separately in Section 4.2.

The workflow in Figure 1 is implemented through two algorithms. Algorithm 1 constructs and evaluates *iForest* with candidate hyperparameters, while Algorithm 2 applies VNS optimization and IQR-based thresholding. Together, they operate the VNS-IF-IQR framework for systematic hyperparameter tuning and anomaly detection in motor signal data.

**Algorithm 1:** VNS_IF_IQR methodology.

**Algorithm 1** VNS-IF-IQR: Hyperparameter optimization for Isolation Forest

**Require:** $X$: Initial parameter vector $[n\_estimators, max\_samples, contamination]$
**Require:** $k_{max}$: Maximum number of neighborhoods
**Require:** $max\_iter$: Maximum iterations without improvement

$X^{*} \leftarrow X$ {Initialize best configuration}
$D_{max}^{*} \leftarrow$ EvaluateDmax($X^{*}$) {Initial $D_{max}$ evaluation}
$iter \leftarrow 0$
**while** $iter < max\_iter$ **do**
  $k \leftarrow 1$
  **while** $k \leq k_{max}$ **do**
    $Y \leftarrow$ Shake($X^{*}, k$)
    $X' \leftarrow$ LocalSearch($Y$)
    $D'_{max} \leftarrow$ EvaluateDmax($X'$)
    **if** $D'_{max} > D_{max}^{*}$ **then**
      $X^{*} \leftarrow X'$ {Update optimal configuration}
      $D_{max}^{*} \leftarrow D'_{max}$ {Update maximum distance}
      $k \leftarrow 1$
      $iter \leftarrow 0$
    **else**
      $k \leftarrow k + 1$
      $iter \leftarrow iter + 1$
    **end if**
  **end while**
**end while**
**return** $X^{*}$

The optimization strategy described in Algorithm 1 operates as the main search engine of the framework, exploring the hyperparameter space to identify the configuration $X^*$ that yields the highest separation. To evaluate each candidate solution, the VNS relies on the $D_{max}$, which serves as the measuring rule for separation, as detailed in Algorithm 2.

**Algorithm 2:** Evaluate distance.

**Algorithm 2** EvaluateDmax: Calculation of the average anomalous distance ($D_{max}$)

**Require:** $X = [n\_estimators, max\_samples, contamination]$

Train the Isolation Forest model using parameters $X$
Generate anomaly scores $s_i$ for all data instances
Calculate quartiles $Q1, Q3$ and $IQR \leftarrow Q3 - Q1$
$T \leftarrow Q1 - 1.5 \cdot IQR$ {Adaptive threshold}
Identify the set of anomalies $A = \{s_i \mid s_i < T\}$
Calculate distances $d_i = T - s_i$ for each $s_i \in A$
$D_{max} \leftarrow \frac{1}{|A|} \sum_{s_i \in A} d_i$

**return** $D_{max}$

To assess the quality of this configuration, the average distance between the anomaly scores and the threshold is calculated. This distance quantifies the degree of separation between normal and anomalous observations. The fitness function is defined in Equations (9 and 10), a higher fitness value indicates a configuration that enhances the separation between both groups.

$$D = \frac{1}{N}\sum_{i=1}^{N}|s_i - T| \tag{9}$$

$$f(X) = -\frac{1}{N}\sum_{i=1}^{N}|s_i - T| \tag{10}$$

where $s_i$ is the anomaly score generated for observation $i$, $T$ is the threshold obtained from the IQR criterion, $N$ is the number of processed observations, and $X$ is the candidate hyperparameter configuration.

The VNS-IF-IQR procedure evaluates candidate hyperparameter configurations by quantifying the separation between normal and anomalous samples, thereby enhancing the *iForest* unsupervised structural anomaly discrimination. At each iteration, new neighboring solutions are generated through controlled perturbations of the current configuration and refined via local search. Solutions are considered superior when failure case scores lie farther from the IQR-based threshold while normal observations remain consistent.

The algorithm explores a wide range of configurations: low contamination values and moderate numbers of estimators generally produce clearer separation, whereas high contamination or excessively large ensembles can reduce discriminatory power. Intensification phases focus on local neighborhoods, revealing incremental gains, while diversification allows the search to escape plateaus and identify configurations yielding substantial improvements. These performance jumps reflect the irregularity of the hyperparameter landscape and the critical role of specific parameter combinations in optimizing anomaly detection.

This structured search procedure establishes the foundation for the experimental evaluation presented in the following chapter, where the effectiveness of the VNS-IF-IQR framework is assessed using a validation using synthetic fault injection and real BLDC motor datasets.

The VNS–IF–IQR framework integrates metaheuristic hyperparameter optimization with adaptive threshold calibration, enabling unsupervised anomaly detection while preserving non-parametric properties. The overall procedure systematically explores the hyperparameter space, evaluates anomaly score distributions, and applies interquartile-range-based thresholding to separate anomalous and nominal samples.

To further assess robustness against hyperparameter instability under non-stationary operating conditions, a systematic hyperparameter sensitivity analysis was incorporated into experimental design. Multiple independent executions were performed across the explored search space, and for each configuration, anomaly score dispersion and separation metrics were computed. Statistical summaries (mean and standard deviation) were used to quantify variability and identify stable parameter regions.

In addition, a marginal sensitivity index was defined, in Equation (11), to quantify the relative influence of each hyperparameter $p$ on the optimization outcome.

$$S_p = \frac{\max(\bar{f}_p) - \min(\bar{f}_p)}{\bar{f}_{global}} \tag{11}$$

where $\bar{f}_p$ represents the mean distance quantifies the degree of separation between normal and anomalous observations obtained for each distinct value of hyperparameter $p$, and $\bar{f}_{global}$ denotes the overall mean Best Score across the entire search space. This normalized index provides a scale-independent measure of the dispersion induced by each hyperparameter, allowing identification of parameters contributing most to performance variability.

## 4. Results

### 4.1. Synthetic fault injection

Synthetic fault injection was introduced to provide controlled ground-truth conditions for evaluating the optimization criterion. The synthetic dataset is not intended to reproduce the full non-stationary complexity of real urban operations. Its purpose is methodological: to isolate the structural behavior of the proposed VNS–IF–IQR framework under known perturbations.

In electromechanical systems such as permanent magnet BLDC motors, voltage and current signals exhibit oscillatory behavior governed by electromagnetic induction and rotational dynamics [13,51]. Even under variable load, the dominant spectral component remains associated with the fundamental electrical frequency. Representing signals through sinusoidal functions therefore preserves the primary dynamical backbone observed in motor current and voltage measurements [6,12,52–54]. While synthetic signals do not capture harmonic distortion, stochastic load variation, or nonlinear coupling effects, they retain the oscillatory structure upon which localized deviations can be systematically introduced.

The proposed algorithm does not assume perfect periodicity or purely harmonic signals. It operates on structural features such as temporal coherence, local deviations from a dominant oscillatory component, and transient distortions. In practical systems, anomalies often appear as short-duration amplitude variations, phase shifts, or localized waveform distortions superimposed on regular oscillatory patterns [52–54]. By maintaining the fundamental spectral structure while introducing controlled perturbations, the synthetic dataset allows isolation of algorithmic behavior from environmental variability and enables evaluation under known ground-truth conditions.

Although the synthetic data simplify aspects such as harmonic distortion, stochastic noise, and nonlinear load effects, they preserve the underlying oscillatory backbone present in real signals. Accordingly, transferability is supported under a principle of partial structural equivalence: while statistical complexity differs between synthetic and real datasets, both share an oscillatory structure and anomaly mechanisms characterized by localized deviations from periodic behavior. Synthetic validation establishes methodological consistency, whereas evaluation on real-world data provides empirical confirmation of robustness.

A total of 30 random time series $X$ were generated, each consisting of $t$ =1.000 time points. The $t$-th value of the $i$-th time series $X$ is expressed by $X_i^t$ and is obtained as shown Equation (8), $t_0 \in \{50,100\}$, $\Lambda = 0.0$, and $\varepsilon_i \sim N(0,1)$. By altering consecutive values $t, t+1, \ldots, t+k$ of $X$, we induce artificial anomalies in the trigonometric series through the contamination level $\delta \in \{0.0, 0.1, 0.2, 0.3, 0.4, 0.5, 1.0, 2.0\}$. We randomly selected five of the 30 available variables and introduced anomalies $\delta$ into them. Each selected variable was contaminated five times, with each anomaly lasting three periods. This procedure yielded a new binary vector, denoted by $Y$, where 0 indicates the absence of artificial anomalies and 1 indicates that the observation contains at least one artificial anomaly in any variable. The vector $Y$ was used as the response variable in the classification machine learning models.

Under these controlled conditions, performance metrics (Recall, Precision, and F1-score) were computed alongside the proposed distance metric $D$. The metric $D$ is not defined as deviation from an arbitrary threshold. Instead, it quantifies separability between anomaly-score distributions associated with normal and perturbed observations. In this sense, $D$ measures structural discrimination rather than classification accuracy. Figures 2, 3 and 4 present the scatterplots between the distance $D$ and the performance metrics show a consistent monotonic, non-linear increasing relationship across contamination regimes. For $\delta = 0$, corresponding to absence of anomalies, values of $D$concentrate near zero and performance metrics remain minimal, as expected due to the absence of true positives. As $\delta$increases, point clouds shift toward larger values of $D$and higher Recall, Precision, and F1-score. Intermediate contamination levels exhibit rapid growth in metrics as separability increases, while higher contamination levels approach saturation regions with diminishing marginal gains.

When the analysis is stratified by the values of $\delta$ (represented by colors), a clear structural pattern emerges. As $\delta$ increases, indicating higher contamination levels, the corresponding point clouds shift toward larger values of $D$ and simultaneously toward higher values of each performance metric. This progression reflects a systematic increase in the separation between anomalous and non-anomalous observations. For intermediate values of $\delta$, the metrics increase

rapidly with $D$, revealing a non-linear growth regime in which improvements in separability translate into substantial gains in detection performance. For the largest values of $\delta$, the metrics approach high and relatively stable levels, indicating that once anomalies are sufficiently separated from normal observations, additional increases in distance produce smaller marginal improvements.

This behavior is consistent with the definitions of evaluation measures. Larger distances reduce overlaps between classes, which increases Recall by decreasing false negatives and improving Precision by reducing false positives. Since the F1-Score is the harmonic means of Precision and Recall, it increases as both quantities improve simultaneously. Overall, the color-coded structure of the point clouds indicates that higher contamination levels are associated with greater distances and reduced overlap between score distributions under controlled conditions, supporting the interpretation of the proposed distance as a structural separability index.

The results shown in Figure 2 indicate that the proposed metric $D$ closely tracks the traditional classification measures (Recall, F1-score, and Precision). Low values of Recall, F1-Score, and Precision are associated with low values of $D$, while high values of these metrics correspond to high values of $D$. In real anomaly detection problems, the true labels Y indicating the presence or absence of anomalies are typically unavailable. Therefore, $D$ cannot be interpreted as an absolute performance measure but rather as a comparative criterion for evaluating competing anomaly detection models. In this sense, $D$ acts as a proxy for Recall, F1-score, and Precision. For example, when comparing two anomaly detection methods, A and B, the method with the larger value of $D$ may be preferred, as higher values of $D$ are associated with better anomaly identification performance.

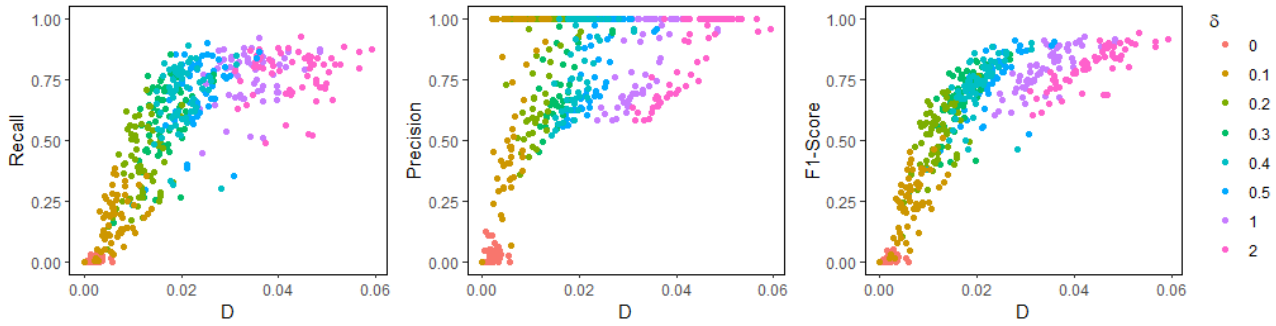


Figure 2. Scatterplot of Distance ($D$) vs performance measures.

Synthetic experiments validate the optimization criterion $D$ under controlled perturbations, demonstrating that increased score separation corresponds to improved structural discrimination. While these results establish internal validity, synthetic conditions do not encompass the full variability of urban environments. Consequently, real-world datasets are used to assess hyperparameter stability and robustness under non-stationary regimes. The monotonic behavior of $D$ across various contamination levels supports its application as a comparative separability index. Rather than a universal detection metric, $D$ serves as a specialized optimization objective for unsupervised contexts where labels are unavailable. Synthetic validation shows that maximizing the proposed distance is systematically associated with increased score separability under controlled perturbations, while real-world evaluation demonstrates stable anomaly-score dispersion across heterogeneous operating regimes. This dual-level approach validates the framework's internal logic while preserving its non-parametric, unsupervised nature.

### 4.2. Evaluation under real urban operating conditions

Real-world experiments evaluate hyperparameter stability, anomaly-score distribution behavior, and structural separation under heterogeneous, non-stationary operating regimes.

The proposed VNS-IF-IQR methodology was applied to 78,886 electrical signal records collected from two rear-hub BLDC motors installed in electric bicycles operating under real, non-stationary urban traffic conditions. Unlike controlled laboratory experiments conducted under steady-state regimes, the data acquisition process was carried out during actual route operation, where stop-and-go traffic, slope variations, rider-dependent load changes, and transient accelerations introduce

significant temporal variability and nonlinear dynamics in voltage and current signals. To partially compensate for the lack of labeled real faults, we performed a qualitative temporal coherence analysis. Detected anomalies were not uniformly distributed, but clustered in segments associated with transient load changes (e.g., acceleration, slope variation), suggesting that the algorithm is capturing structurally meaningful operational deviations rather than random statistical outliers. The monitored variables and sensing modalities used in the experimental platform are summarized in Table 1. All signals were acquired directly from embedded sensing units without feature engineering prior to anomaly detection.

Table 1. Monitored variables and sensing configuration in the experimental platform.

| Variable | Sensor Type / Source | Measurement Domain | Sampling Strategy | Processing Level |
|---|---|---|---|---|
| Phase Voltage | Hall-effect voltage sensor | Electrical | Synchronous acquisition | Raw time-series |
| Phase Current | Hall-effect current sensor | Electrical | Synchronous acquisition | Raw time-series |
| Motor Temperature | NTC thermistor | Thermal | Periodic sampling | Raw scalar |
| Ambient Temperature | Digital temperature sensor | Environmental | Periodic sampling | Raw scalar |
| Battery Voltage | Embedded BMS measurement | Electrical | Synchronous acquisition | Raw scalar |
| Battery Current | Embedded BMS measurement | Electrical | Synchronous acquisition | Raw scalar |
| Vehicle Speed | Hall-effect wheel sensor | Kinematic | Synchronous acquisition | Raw scalar |
| Timestamp | Internal controller clock | Temporal reference | Continuous | Reference index |

Table 2. Execution and results of the algorithm for Motor A, operating with Battery A.

| Execution | max_ samples | n_ estimators | Contami- nation | $D_{max}$ | Standard deviation | Fault Detection | Normal Instances | Threshold | Time [min] |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 6912 | 250 | 0.19 | 0.0503 | 0.04423 | 377 | 6670 | -0.0412 | 11.39 |
| 2 | 6656 | 150 | 0.16 | 0.0499 | 0.04314 | 363 | 6684 | -0.0385 | 8.95 |
| 3 | 6912 | 50 | 0.06 | 0.0513 | 0.04467 | 354 | 6693 | -0.0058 | 13.11 |
| 4 | 6912 | 250 | 0.19 | 0.0503 | 0.04423 | 377 | 6670 | -0.0412 | 37.60 |
| **5** | **6912** | **50** | **0.17** | **0.0513** | **0.04467** | **354** | **6693** | **-0.0401** | **10.69** |
| 6 | 6912 | 50 | 0.17 | 0.0513 | 0.04467 | 354 | 6693 | -0.0401 | 17.52 |
| 7 | 6912 | 350 | 0.33 | 0.0491 | 0.04382 | 385 | 6662 | -0.0569 | 103.39 |
| 8 | 6912 | 250 | 0.29 | 0.0503 | 0.04423 | 377 | 6670 | -0.0537 | 92.87 |
| 9 | 6912 | 50 | 0.27 | 0.0513 | 0.04467 | 354 | 6693 | -0.0537 | 11.67 |
| 10 | 6912 | 50 | 0.18 | 0.0513 | 0.04467 | 354 | 6693 | -0.0418 | 167.08 |

Data were collected from three experimental setups: Motor A operating with Battery A, Motor B with Battery A, and Motor B operating with Battery B, as well as a combined dataset. Each *iForest* hyperparameter (*max_samples*, *n_estimators*, *contamination*) was defined: $X^E = 36$, $X^{MS} = 36$ and $X^C = 36$, generating a total search space of 46,656 combinations. The VNS algorithm explored this space via neighborhood perturbations and local search to maximize the separation between anomalous and normal samples.

Optimization results, summarized across ten independent runs for each dataset, show consistent convergence to stable hyperparameter configurations. For Motor A with Battery A, the highest average anomaly distance (0.05133) was achieved in multiple runs, with the optimal configuration selected based on minimal computational time. Detected faults remained stable across runs, ranging from 354 to 385, while normal instances remained near 6,690, indicating robust detection without erratic classification. Computational time varied from 9 to 167 minutes, reflecting differences in local search intensity. These results demonstrate the VNS-IF-IQR framework's effectiveness in optimizing *iForest* hyperparameters for analyzing hyperparameter stability and anomaly-score robustness, rather than establishing supervised diagnostic accuracy, in BLDC motors operating under non-stationary conditions.

Following the analysis of the parameter configurations reported in Table 2, the behavior of the optimization process can be examined in detail through the evolution of the best objective values across global iterations. Figure 3 presents this

progression and provides direct evidence of how the VNS explores the search space, identifies improved configurations, and stabilizes once no additional gains can be achieved.

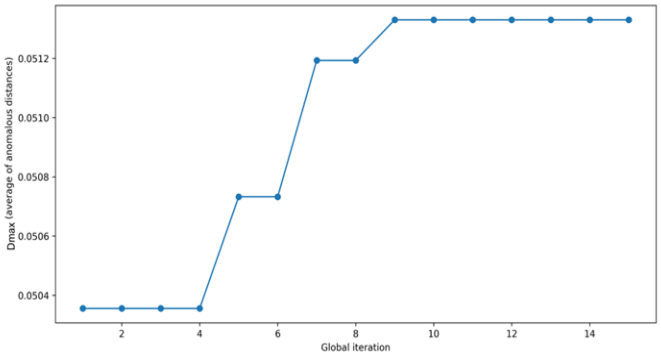


Figure 3. Evaluation $D_{max}$ Motor A, operating with Battery A.

The evolution $D_{max}$ for Motor A with Battery A shows progressive improvement during the initial iterations, reflecting the algorithm's exploration of different neighborhoods. Significant gains were observed between iterations 5 and 9, corresponding to configurations that enhanced the average anomalous distance. After iteration 9, the scores stabilized, indicating convergence to an optimal region of the hyperparameter space.

For Motor B with Battery A (Table 3), the $D_{max}$ achieved across runs were consistently lower than those of Motor A, reflecting differences in motor behavior and signal characteristics under the same battery conditions.

Table 3. Execution and results of the algorithm for Motor B, operating with Battery A.

| Execution | max_samples | n_estimators | contamination | $D_{max}$ | Standard deviation | Fault Detection | Normal Instances | Threshold | Time [min] |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 6400 | 250 | 0.34 | 0.0416 | 0.03668 | 1637 | 34348 | -0.0729 | 31.94 |
| 2 | 6912 | 50 | 0.01 | 0.0411 | 0.03708 | 1535 | 34450 | 0.0600 | 21.89 |
| 3 | 5632 | 150 | 0.28 | 0.0419 | 0.03697 | 1707 | 34278 | -0.0690 | 32.86 |
| 4 | 4096 | 150 | 0.16 | 0.0416 | 0.03707 | 1859 | 34126 | -0.0472 | 21.98 |
| **5** | **5632** | **150** | **0.25** | **0.0419** | **0.03697** | **1707** | **34278** | **-0.0655** | **14.44** |
| 6 | 5376 | 50 | 0.09 | 0.0418 | 0.03690 | 2005 | 33980 | -0.0215 | 17.08 |
| 7 | 4352 | 800 | 0.18 | 0.0408 | 0.03567 | 1530 | 34455 | -0.0603 | 19.11 |
| 8 | 4096 | 150 | 0.29 | 0.0416 | 0.03707 | 1859 | 34126 | -0.0643 | 23.50 |
| 9 | 5632 | 150 | auto | 0.0419 | 0.03697 | 1707 | 34278 | -0.0071 | 20.25 |
| 10 | 5632 | 200 | 0.14 | 0.0419 | 0.03617 | 1632 | 34353 | -0.0478 | 47.42 |

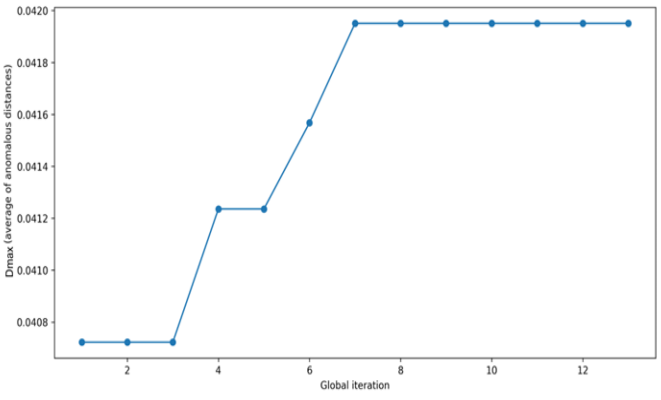


Figure 4. $D_{max}$ evaluation for Motor B operating with Battery A.

For Motor B with Battery A, anomalies are closer to the IQR-based decision threshold. Run 5 achieved the highest mean anomalous distance (0.0419506), after which scores stabilized, indicating convergence. Run 4 obtained a lower score, reflecting a less effective hyperparameter configuration. Execution times were generally longer than for Motor A, likely due to increased local search complexity or specific hyperparameter settings. The top-scoring run, however, had a moderate computation time. Overall, the optimization identified configurations suited to the characteristics of Motor B under these conditions. $D_{max}$ evolution, across runs is shown in Figure 4, highlighting the convergence behavior and the impact of hyperparameter tuning on anomaly separation.

Figure 4 shows the evolution of the $D_{max}$ for Motor B with Battery A. The score stabilizes from run 7, and although its absolute value is lower than Motor A, convergence indicates that the metaheuristic identifies suitable hyperparameter settings for this dataset. Comparison across the two motors with Battery A reveals that optimization adapts to different operational conditions, with Motor A allowing greater performance gains and Motor B constrained by its data structure.

The process demonstrates robustness by successfully stabilizing in both scenarios but also highlights the need for motor-specific calibration strategies because a single configuration cannot be assumed to generalize uniformly across heterogeneous operational profiles.

To extend the analysis to the second energy source, the following section introduces Table 4, which summarizes the execution parameters and results obtained for Motor B operating with Battery B.

Table 4. Execution and results of the algorithm for Motor B operating Battery B.

| Execution | max_samples | n_estimators | contamination | $D_{max}$ | Standard deviation | Fault Detection | Normal Instances | Threshold | Time [min] |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 6656 | 500 | 0.18 | 0.0443 | 0.03627 | 1455 | 34399 | -0.07011 | 26.64 |
| 2 | 5888 | 50 | 0.35 | 0.0450 | 0.03521 | 1459 | 34395 | -0.10521 | 59.07 |
| 3 | 6144 | 50 | 0.07 | 0.0466 | 0.03724 | 1721 | 34133 | -0.02277 | 54.42 |
| 4 | 5888 | 50 | 0.18 | 0.0450 | 0.03521 | 1459 | 34395 | -0.07527 | 34.26 |
| 5 | 6656 | 500 | 0.19 | 0.04439 | 0.03627 | 1455 | 34399 | -0.07193 | 56.95 |
| 6 | 6656 | 450 | 0.35 | 0.04422 | 0.03625 | 1450 | 34404 | -0.09667 | 36.50 |
| 7 | 5888 | 50 | 0.35 | 0.04503 | 0.03521 | 1459 | 34395 | -0.10521 | 25.39 |
| 8 | 5888 | 100 | 0.24 | 0.04447 | 0.03517 | 1491 | 34363 | -0.08575 | 23.46 |
| 9 | 6656 | 900 | 0.26 | 0.04422 | 0.03579 | 1460 | 34394 | -0.08409 | 26.45 |
| **10** | **6144** | **50** | **0.09** | **0.04662** | **0.03724** | **1721** | **34133** | **-0.03666** | **20.33** |

Figure 5 shows the evolution of the $D_{max}$ for Motor B with Battery B. The score stabilizes after the initial flat phase during the first five iterations, indicating consistent separation between failure observations and the IQR-based threshold. Runs 3 and 10 achieved the highest score of 0.0466 (Table 4), corresponding to hyperparameter configurations with reduced *n_estimators*, low *contamination*, and *max_samples* between 6,144 and 6,656, which generate clearer isolation boundaries. Configurations with higher contamination did not improve performance. Detected failures range from 1,450 to 1,721, reflecting the sensitivity of the threshold to score dispersion. Execution times varied from 20 to 50 minutes depending on the extent of the VNS exploration.

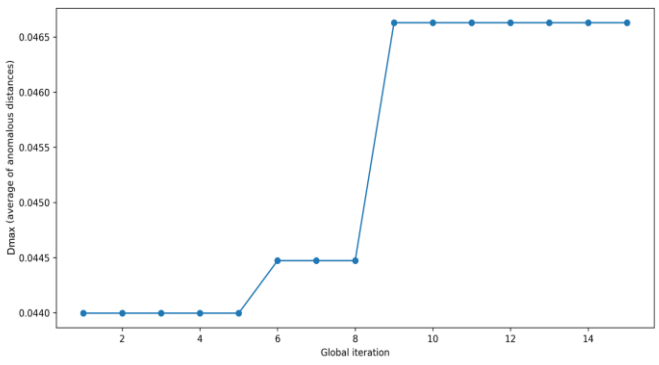


Figure 5. $D_{max}$ evaluation for Motor B operating with Battery B.

The results indicate that the algorithm effectively identified hyperparameter configurations suited to this dataset, with the highest-scoring runs achieving a balance between separation of anomalies and computational efficiency. The plateau observed after the initial iterations suggests convergence to a stable region of the search space, while variations in detected failures across runs reflect the influence of different hyperparameter choices on model sensitivity. Overall, the VNS-IF-IQR procedure demonstrates the ability to adapt to the specific characteristics of Motor B operating with Battery B, supporting robust anomaly detection in real-world signal data.

In summary, Figure 7 shows a convergence pattern characterized by long intervals of stagnation interrupted by one decisive improvement. This confirms that the optimization process eventually reaches a robust solution around iteration 9. No subsequent exploration yields additional benefits after this point.

The proposed methodology is applied to the full dataset. Table 5 consolidates the global results obtained, integrating observations from both motors and both batteries. Introducing this table marks the transition from the scenario-specific analyses toward an aggregated evaluation, where the model faces the highest heterogeneity and variability of the study.

Table 5. Execution and results of the algorithm for complete data.

| Execution | max_samples | n_estimators | contamination | $D_{max}$ | Standard deviation | Fault Detection | Normal Instances | Threshold | Time [min] |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 6912 | 50 | 0.08 | 0.04086 | 0.03307 | 3022 | 75864 | -0.03344 | 36.64 |
| 2 | 5376 | 150 | 0.19 | 0.04172 | 0.03354 | 3494 | 75392 | -0.05967 | 60.25 |
| 3 | 6656 | 600 | 0.35 | 0.04022 | 0.03225 | 3418 | 75468 | -0.08381 | 30.05 |
| 4 | 6912 | 50 | 0.12 | 0.04086 | 0.03307 | 3022 | 75864 | -0.04872 | 71.69 |
| 5 | 6912 | 50 | 0.2 | 0.04086 | 0.03307 | 3022 | 75864 | -0.06806 | 114.90 |
| 6 | 5376 | 150 | auto | 0.04172 | 0.03354 | 3494 | 75392 | -0.01344 | 93.85 |
| **7** | **5376** | **150** | **0.06** | **0.04172** | **0.03354** | **3494** | **75392** | **-0.01408** | **45.90** |
| 8 | 6656 | 50 | 0.13 | 0.04047 | 0.03232 | 3401 | 75485 | -0.04570 | 65.39 |
| 9 | 4864 | 300 | 0.28 | 0.04005 | 0.03131 | 3342 | 75544 | -0.08047 | 30.03 |
| 10 | 5376 | 600 | 0.26 | 0.04045 | 0.03162 | 3361 | 75525 | -0.07476 | 40.19 |

The $D_{max}$ values range from 0.04005405 to 0.04172204, indicating a narrow performance margin and suggesting that even small improvements rely on specific hyperparameter configurations. Executions 2, 6, and 7 achieved the highest scores, reflecting parameter combinations that effectively captured the joint dynamics of the system. Execution 9 obtained the lowest score, showing reduced discrimination capability under more complex interactions between motors and batteries.

The top-performing configurations share common values for *max_samples* (5,376) and *n_estimators* (150), while contamination levels vary, indicating convergence to a cluster of competitive solutions rather than a single optimum. The number of observations classified as failures ranges from 3,342 to 3,494, consistent with the scale of the global dataset. The IQR-based threshold shows notable variation, reflecting changes in the anomaly score distribution when combining subsystems. Execution times vary significantly, highlighting the trade-off between computational cost and model accuracy.

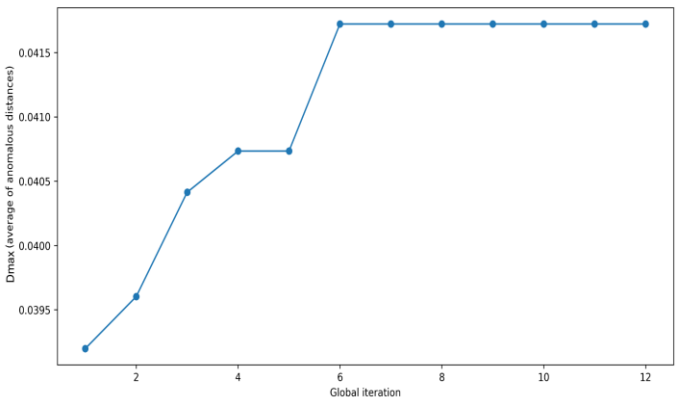


Figure 6. $D_{max}$ evaluation for complete data.

Figure 6 shows change $D_{max}$ during the optimization process for the global scenario, which combines data from both motors and batteries. The curve shows a gradual improvement in the $D_{max}$ during the first six iterations, followed by a clear stabilization phase beginning with iteration seven.

Overall, the results demonstrate that the VNS-IF-IQR procedure can identify robust hyperparameter configurations for the integrated dataset, unsupervised anomaly-screening consistency across multiple motors and batteries while adapting to increased data complexity. The analysis confirms the method's ability to generalize across heterogeneous operational conditions and provides a foundation for evaluating anomaly detection performance under realistic, multi-component scenarios.

The experimental evaluation of the VNS-IF-IQR framework across individual and combined BLDC motor datasets demonstrates its capability to identify hyperparameter configurations that enhance *iForest* anomaly detection. The method consistently stabilized detection performance, maintaining clear separation between normal and anomalous instances while adapting to the distinct operational and data-driven characteristics of each motor–battery combination. Variations in performance metrics, execution times, and IQR-based thresholds highlight the influence of dataset topology and system dynamics on model outcomes, emphasizing the need for scenario-specific calibration. These results provide a comprehensive to evaluate unsupervised structural anomaly discrimination under non-stationary operating conditions. Furthermore hyperparameter stability and anomaly-score

robustness rather than to establish supervised diagnostic accuracy, which is addressed in the following discussion.

### 4.3. Hyperparameter sensitivity

This subsection presents the results of the hyperparameter sensitivity analysis across different motor–battery configurations evaluated under real urban operating conditions. The VNS–IF–IQR framework was executed independently for Motor A with Battery A, Motor B with Battery A, Motor B with Battery B, and the combined dataset. For each configuration, used mean and standard deviation presented in Tables 2, 3, 4, and 5, the marginal sensitivity index $S_p$ was calculated to quantify the relative influence of each hyperparameter. The marginal sensitivity index assesses the relative influence of each hyperparameter on the optimization objective.

For the Motor A–Battery A configuration, the analysis indicates limited sensitivity to *max_samples* (1.62%), suggesting that variations in subsampling size within the explored range produce minor changes in the objective function. In contrast, *n_estimators* and *contamination* exhibit slightly higher sensitivities (4.34%), reflecting their influence on ensemble diversity and adaptive threshold definition, respectively. However, all sensitivity values remain below 5%, indicating the absence of sharply dominant parameters. This behavior suggests that the optimization process converges toward relatively broad near-optimal regions rather than relying on highly specific configurations.

In the Motor B–Battery A scenario, identical marginal sensitivity values (2.64%) were observed for *max_samples, n_estimators, and contamination*. This homogeneous profile indicates a comparatively flat response surface across the explored hyperparameter space. The comparable and limited influence of each parameter suggests reduced dependency on individual hyperparameter adjustments and reinforces the stability of the optimization process under these operating conditions.

For the Motor B–Battery B configuration, marginal sensitivity increases moderately for *contamination* (5.34%) and *max_samples* (5.17%), while *n_estimators* shows lower influence (3.18%). This pattern suggests that under comparatively more heterogeneous operating conditions, the optimization outcome becomes somewhat more dependent on subsampling strategy and adaptive threshold calibration than on ensemble size. Nevertheless, all sensitivity values remain below 6%, indicating that the optimization landscape does not exhibit extreme parameter sensitivity.

Across all evaluated configurations, marginal sensitivity indices remain within a narrow range and consistently below 6%. Combined with the observed dispersion levels (mean ± standard deviation), these results indicate that the VNS–IF–IQR framework operates over relatively flat performance regions and does not depend on finely tuned hyperparameter combinations. From a reliability perspective, this behavior supports the stability of the proposed optimization strategy under varying electromechanical and energy supply conditions. Importantly, the structured sensitivity assessment provides quantitative evidence that the reported improvements are associated with consistent parameter behavior rather than isolated favorable configurations, thereby strengthening the robustness claims of the proposed framework.

Overall, the sensitivity analysis indicates that the proposed VNS–IF–IQR framework maintains stable optimization behavior across different motor–battery configurations, with no evidence of extreme parameter dominance or sharp performance discontinuities within the explored search space. These findings suggest that the observed improvements in anomaly separation are associated with consistent hyperparameter regions rather than isolated configurations. The framework provides a consistent anomaly-screening baseline suitable for integration with future maintenance-aligned datasets.

## 5. Discussion

The absence of confirmed fault labels in the real-world dataset requires a conceptual distinction between diagnostic classification and structural anomaly discrimination. In urban electric mobility, maintenance records are frequently asynchronous with high-frequency sensor streams or statistically insufficient due to the low incidence of failure events. Consequently, the VNS–IF–IQR framework is evaluated as an unsupervised anomaly screening mechanism. The objective is to assess whether systematic hyperparameter optimization improves structural separation between nominal and statistically distinct observations under non-stationary

conditions.

Validation is conducted at two complementary levels. Synthetic fault injections provide a controlled environment to assess sensitivity and score consistency against known perturbations in oscillatory electrical signals. This stage isolates methodological effects from environmental uncertainty. The real-world evaluation subsequently examines hyperparameter stability and separation consistency across heterogeneous operating regimes. Together, these levels address internal methodological validity and external operational stability without relying on labeled fault intervals.

Results across the four motor–battery configurations indicate that anomaly-score separation is sensitive to dataset structure and operational variability. The optimization process identifies hyperparameter regions that improve structural discrimination relative to default settings. For instance, in Motor A, stable separation is achieved when ensemble sizes and contamination align with intrinsic variability. Conversely, Motor B exhibits increased operational noise, requiring more selective adjustments of subsampling size to maintain isolation depth. These results confirm that hyperparameters regulate how stochastic partitioning interacts with data geometry under non-stationarity.

Marginal sensitivity analysis further supports this interpretation. Low sensitivity values across scenarios indicate relatively flat performance landscapes, suggesting convergence toward broad near-optimal regions rather than unstable, over-fitted solutions. This behavior demonstrates that VNS-driven optimization improves robustness without introducing excessive dependence on precise parameter values. While execution times (14.44 to 20.33 minutes) reflect data complexity, they remain compatible with near-real-time monitoring in electric mobility. Ultimately, the study demonstrates that systematic tuning enhances the structural consistency of isolation-based detection, providing a statistically sound baseline for future integration with synchronized maintenance records.

## 6. Conclusion

This study addresses hyperparameter instability in *iForest* as a structural limitation for non-stationary signal analysis. By integrating VNS with adaptive IQR-based thresholding, the proposed framework achieves stable anomaly-score separation in fully unsupervised settings.

The dual-level validation confirms both internal methodological consistency—via synthetic injection—and external operational robustness across real-world motor–battery scenarios. The marginal sensitivity analysis is particularly significant, as it proves that the framework converges to stable near-optimal regions, reducing the risk of performance degradation under minor dataset shifts.

The findings establish that dataset-specific tuning is a prerequisite for reliable isolation-based detection in heterogeneous environments. While default configurations fail to capture evolving data geometries, structured optimization enables a consistent representation of anomalies. This work provides a scalable, unsupervised screening baseline for reliability-oriented monitoring.

Future research will focus on the integration of streaming temporal segmentation and incorporate synchronized maintenance logs to assess correspondence between anomaly scores and confirmed failure events.

## References


1. IEA. Global EV Outlook 2025. 2025.
2. Yan S, Kaundanya C, O'Connor N E, Little S, Liu M. Machine Learning in Micromobility: A Systematic Review of Datasets, Techniques, and Applications 2025. https://doi.org/10.48550/ARXIV.2508.16135.
3. IEA. Global EV Outlook 2024. 2024.
4. Jiang X, Wang Y, Li J, Ye L. Comprehensive importance analysis for repairable system components based on the GO method. *Eksploatacja i Niezawodnosc – Maintenance and Reliability* 2022; 24 (4): 785–794, http://doi.org/10.17531/ein.2022.4.18.
5. Fadil N D B M, Zaidi A F A, Leong J H, Azalan M S Z, Azmi S A, Wahab S P A. Assessing Torque-Ripple Mitigation Strategies for BLDC Motors in Electric Vehicles. *2025 9th International Conference on Man-Machine Systems (ICoMMS)*, Malacca, Malaysia: IEEE; 2025: 216–221. https://doi.org/10.1109/ICoMMS66553.2025.11200236.
6. Khaneghah M Z, Alzayed M, Chaoui H. Fault Detection and Diagnosis of the Electric Motor Drive and Battery System of Electric Vehicles.

*Machines* 2023; 11(7): 713-725. https://doi.org/10.3390/machines11070713.

7. Kudelina K, Asad B, Vaimann T, Rassolkin A, Kallaste A, Lukichev D V. Main Faults and Diagnostic Possibilities of BLDC Motors. *2020 27th International Workshop on Electric Drives: MPEI Department of Electric Drives 90th Anniversary (IWED)*, Moscow, Russia: IEEE; 2020: 1–6. https://doi.org/10.1109/IWED48848.2020.9069553.
8. Tashakori A, Ektesabi M. Fault diagnosis of in-wheel BLDC motor drive for electric vehicle application. *2013 IEEE Intelligent Vehicles Symposium (IV)*, Gold Coast City, Australia: IEEE; 2013: 925–930. https://doi.org/10.1109/IVS.2013.6629585.
9. Hur J-H, Lee T-G, Moon S-A, Lee S J, Yoo H, Moon S J, Lee J H. Thermal reliability analysis of a BLDC motor in a high-speed axial fan by the accelerated-life test and numerical methods. *Heat and Mass Transfer* 2008; 44(11): 1355–1369. https://doi.org/10.1007/s00231-008-0375-7.
10. Aslan E, Aoulmi Z. Temperature Prediction and Performance Comparison of Permanent Magnet Synchronous Motors Using Different Machine Learning Techniques for Early Failure Detection. *Eksploatacja i Niezawodność – Maintenance and Reliability* 2025; 27(1). https://doi.org/10.17531/ein/192164.
11. Karpenko M, Prentkovskis O, Skačkauskas P. Analysing the impact of electric kick-scooters on drivers: vibration and frequency transmission during the ride on different types of urban pavements. *Eksploatacja i Niezawodność – Maintenance and Reliability* 2025; 27(2). https://doi.org/10.17531/ein/199893.
12. Shifat T A, Hur J W. An Effective Stator Fault Diagnosis Framework of BLDC Motor Based on Vibration and Current Signals. *IEEE Access* 2020; 8: 106968–106981. https://doi.org/10.1109/ACCESS.2020.3000856.
13. Kumar M P, Velpula S, Saiprakash C, Sahoo B. Advancements in fault detection and diagnosis methods for electric vehicles: a review. *Discover Applied Sciences* 2025; 7(11): 1235. https://doi.org/10.1007/s42452-025-07758-9.
14. Moussa A, Aoulmi Z. Improving Electric Vehicle Maintenance by Advanced Prediction of Failure Modes Using Machine Learning Classifications. *Eksploatacja i Niezawodność – Maintenance and Reliability* 2025; 27(3). https://doi.org/10.17531/ein/201372.
15. Belgacem H, Chihi I. Toward Reliable and Intelligent Sensor Systems: A Comprehensive Study of Fault Diagnosis and Mitigation. *IEEE Sensors Reviews* 2025; 2:511–536. https://doi.org/10.1109/SR.2025.3601092.
16. Calabrese F, Regattieri A, Bortolini M, Galizia F G. Data-Driven Fault Detection and Diagnosis: Challenges and Opportunities in Real-World Scenarios. *Applied Sciences* 2022; 12(18): 9212. https://doi.org/10.3390/app12189212.
17. Al Farizi W S, Hidayah I, Rizal M N. Isolation Forest Based Anomaly Detection: A Systematic Literature Review. *2021 8th International Conference on Information Technology, Computer and Electrical Engineering (ICITACEE)*, Semarang, Indonesia: IEEE; 2021: 118–122. https://doi.org/10.1109/ICITACEE53184.2021.9617498.
18. Li C, Guo L, Gao H, et al. Similarity-Measured Isolation Forest: Anomaly Detection Method for Machine Monitoring Data. *IEEE Transactions on Instrumentation and Measurement* 2021; 70: 1–12. https://doi.org/10.1109/TIM.2021.3062684.
19. Liu F T, Ting K M, Zhou Z-H. Isolation Forest. *2008 Eighth IEEE International Conference on Data Mining*, Pisa, Italy: IEEE; 2008: 413–422. https://doi.org/10.1109/ICDM.2008.17.
20. Pastor D, Martínez J, Muñoz L, Li Y.. Multivariate Automatic Tuning of Isolation Forest for Anomaly Detection in Critical Infrastructures: A Solution for Intelligent Information Systems: *Proceedings of the 17th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management*, Marbella, Spain: SCITEPRESS - Science and Technology Publications; 2025: 393–400. https://doi.org/10.5220/0013710700004000.
21. Lis A, Dworakowski Z, Czubak P. An anomaly detection method for rotating machinery monitoring based on the most representative data. *Journal of Vibroengineering* 2021; 23(4): 861–876. https://doi.org/10.21595/jve.2021.21622.
22. Calvin K, Dasgupta D, Krinner G. IPCC, 2023: Climate Change 2023: Synthesis Report. Contribution of Working Groups I, II and III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change [Core Writing Team, H. Lee and J. Romero (eds.)]. *IPCC*, Geneva, Switzerland. First. Intergovernmental Panel on Climate Change (IPCC); 2023. https://doi.org/10.59327/IPCC/AR6-9789291691647.
23. Tuma J M, Gobry F M J, Osment S E. Thematic Research Summary: Intelligent Transport Systems. *Transport Research Knowledge Centre* 2010. https://doi.org/10.1002/9781118557495.ch6.
24. Rincón-Maya C, Acosta-González D, Guevara-Carazas F, Hernández-Barajas F, Patino-Rodríguez C E, Usuga-Manco O. Predictive

Modeling of Electric Bicycle Battery Performance: Integrating Real-Time Sensor Data and Machine Learning Techniques. *Sensors* 2025; 25(5): 1392. https://doi.org/10.3390/s25051392.

25. Rincón-Maya C, Guevara-Carazas F, Hernández-Barajas F, Patino-Rodríguez C E, Usuga-Manco O. Remaining Useful Life Prediction of Lithium-Ion Battery Using ICC-CNN-LSTM Methodology. *Energies* 2023; 16(20): 7081. https://doi.org/10.3390/en16207081.
26. Krishnan R. Permanent Magnet Synchronous and Brushless DC Motor Drives. 1st ed. CRC Press; 2017. https://doi.org/10.1201/9781420014235.
27. Mohanraj D, Aruldavid R, Verma R, Sathiyasekar K, Barnawi A B, Chokkalingam B, Mihet-Popa L. A Review of BLDC Motor: State of Art, Advanced Control Techniques, and Applications. *IEEE Access* 2022; 10: 54833–54869. https://doi.org/10.1109/ACCESS.2022.3175011.
28. Shu X, Zhang S, Li Y, Chen M. An anomaly detection method based on random convolutional kernel andisolation forest for equipment state monitoring. *Eksploatacja i Niezawodność – Maintenance and Reliability* 2022; 24(4): 758–770. https://doi.org/10.17531/ein.2022.4.16.
29. Zhang J, Zhang Q, Qin X, et al. A two-stage fault diagnosis methodology for rotating machinery combining optimized support vector data description and optimized support vector machine. *Measurement* 2022; 200: 111651. https://doi.org/10.1016/j.measurement.2022.111651.
30. Xu D, Wang Y, Meng Y, Zhang Z. An Improved Data Anomaly Detection Method Based on Isolation Forest. *2017 10th International Symposium on Computational Intelligence and Design (ISCID)*, Hangzhou: IEEE; 2017: 287–291. https://doi.org/10.1109/ISCID.2017.202.
31. Lesouple J, Baudoin C, Spigai M, Tourneret J Y. Generalized isolation forest for anomaly detection. *Pattern Recognition Letters* 2021; 149: 109–119. https://doi.org/10.1016/j.patrec.2021.05.022.
32. Hariri S, Kind M C, Brunner R J. Extended Isolation Forest. *IEEE Transactions on Knowledge and Data Engineering* 2021; 33(4): 1479–1489. https://doi.org/10.1109/TKDE.2019.2947676.
33. Kang H-S, Choi Y-S, Yu J-S, Jin S-W, Lee J-M, Kim Y-J. Hyperparameter Tuning of OC-SVM for Industrial Gas Turbine Anomaly Detection. *Energies* 2022; 15(22): 8757. https://doi.org/10.3390/en15228757.
34. Ali M Q, Al-Shaer E, Khan H, Khayam S A. Automated Anomaly Detector Adaptation using Adaptive Threshold Tuning. *ACM Transactions on Information and System Security* 2013; 15(4): 1–30. https://doi.org/10.1145/2445566.2445569.
35. Carletti M, Terzi M, Susto G A. Interpretable Anomaly Detection with DIFFI: Depth-based Isolation Forest Feature Importance. 2021. https://doi.org/10.48550/arXiv.2007.11117.
36. Hanifi S, Cammarono A, Zare-Behtash H. Advanced hyperparameter optimization of deep learning models for wind power prediction. *Renewable Energy* 2024; 221: 119700. https://doi.org/10.1016/j.renene.2023.119700.
37. He Y, Gao S, Li Y, Guan Y, Zhang J, Hu D. Adaptive machine learning framework: Predicting UHPC performance from data to modelling. *Results in Engineering* 2025; 27: 106724. https://doi.org/10.1016/j.rineng.2025.106724.
38. Suryadi M K, Herteno R, Saputro S W, Faisal M R, Nugroho R A. Comparative Study of Various Hyperparameter Tuning on Random Forest Classification With SMOTE and Feature Selection Using Genetic Algorithm in Software Defect Prediction. *Journal of Electronics, Electromedical Engineering, and Medical Informati*cs 2024; 6(2): 137–147. https://doi.org/10.35882/jeeemi.v6i2.375.
39. Aach M, Sarma R, Neukirchen H, Riedel M, Lintermann A. Resource-adaptive successive doubling for hyperparameter optimization with large datasets on high-performance computing systems. *Future Generation Computer Systems* 2026; 175: 108042. https://doi.org/10.1016/j.future.2025.108042.
40. Serizawa T, Fujita H. Optimization of Convolutional Neural Network Using the Linearly Decreasing Weight Particle Swarm Optimization; 2022. https://doi.org/10.11517/pjsai.JSAI2022.0_2S4IS2b03.
41. Singh P, Chaudhury S, Panigrahi B K. Hybrid MPSO-CNN: Multi-level Particle Swarm optimized hyperparameters of Convolutional Neural Network. *Swarm and Evolutionary Computation* 2021; 63: 100863. https://doi.org/10.1016/j.swevo.2021.100863.
42. Hajiar Yuliana. Hyperparameter Optimization of Random Forest for 5G Coverage Prediction. *Buletin Pos dan Telekomunikasi* 2024; 22(1). https://doi.org/10.17933/bpostel.v22i1.390.
43. El-Hassani F Z, Amri M, Joudar N-E, Haddouch K. A New Optimization Model for MLP Hyperparameter Tuning: Modeling and Resolution by Real-Coded Genetic Algorithm. *Neural Processing Letters* 2024; 56(2): 105. https://doi.org/10.1007/s11063-024-11578-0.

44. Raiaan M A K, Sakib S, Fahad N M, Mamun A A, Rahman M A, Shatabda S, Mukta M S H. A systematic review of hyperparameter optimization techniques in Convolutional Neural Networks. *Decision Analytics Journal* 2024; 11: 100470. https://doi.org/10.1016/j.dajour.2024.100470.
45. Lankford S, Grimes D. Neural Architecture Search using Particle Swarm and Ant Colony Optimization. 2024. https://doi.org/10.48550/ARXIV.2403.03781.
46. Mirzabozorg S A A S, Saremi M, DehghanNiri R, Abedi M, Yousefi M, Beiranvand Pour A, Hezarkhani A, Maghsoudi A. Hyperparameter optimization in unsupervised anomaly detection for mineral prospectivity mapping. *Ore Geology Reviews* 2025; 181: 106627. https://doi.org/10.1016/j.oregeorev.2025.106627.
47. Elshewey A M, Shams M Y, El-Rashidy N, Elhady A M, Shohieb S M, Tarek Z. Bayesian Optimization with Support Vector Machine Model for Parkinson Disease Classification. *Sensors* 2023; 23(4): 2085. https://doi.org/10.3390/s23042085.
48. Iturbe-Araya J I, Rifà-Pous H. Enhancing unsupervised anomaly-based cyberattacks detection in smart homes through hyperparameter optimization. *International Journal of Information Security* 2025; 24(1): 45. https://doi.org/10.1007/s10207-024-00961-6.
49. Jia Z, Wang Z, Sun Z, Liu P, Zhu X, Sun F. A Data-Driven Approach for Battery System Safety Risk Evaluation Based on Real-World Electric Vehicle Operating Data. *IEEE Transactions on Transportation Electrification* 2024; 10(3): 5660–5676. https://doi.org/10.1109/TTE.2023.3324450.
50. Mladenović N, Hansen P. Variable neighborhood search. *Computers & Operations Research* 1997; 24(11): 1097–1100. https://doi.org/10.1016/S0305-0548(97)00031-2.
51. Williamson S S, Emadi A, Rajashekara K. Comprehensive Efficiency Modeling of Electric Traction Motor Drives for Hybrid Electric Vehicle Propulsion Applications. *IEEE Transactions on Vehicular Technology* 2007; 56(4): 1561–1572. https://doi.org/10.1109/TVT.2007.896967.
52. Giantomassi A, Ferracuti F, Iarlori S, Ippoliti G, Longhi, S. Signal Based Fault Detection and Diagnosis for Rotating Electrical Machines: Issues and Solutions. In: Zhu Q, Azar AT, editors. *Complex System Modelling and Control Through Intelligent Soft Computations*, vol. 319, Cham: Springer International Publishing; 2015: 275–309. https://doi.org/10.1007/978-3-319-12883-2_10.
53. Niu G, Dong X, Chen Y. Motor Fault Diagnostics Based on Current Signatures: A Review. *IEEE Transactions on Instrumentation and Measurement* 2023; 72: 1–19. https://doi.org/10.1109/TIM.2023.3285999.
54. Faiz J, Bazrafshan M A, Tabarniarami Z. Demagnetisation fault analysis and diagnosis based on different methods in permanent magnet machines - An overview. *IET Electric Power Applications* 2024; 18(12): 1860–1893. https://doi.org/10.1049/elp2.12519.