Preemptive Failure Detection using Contamination-Based Stacking Ensemble in Missiles

Seong-Mok Kim[1], Ye-Eun Jeong[1], Yong Soo Kim[2], Youn-Ho Lee[3], and Seung Young Lee [4*]
[1] Department of Industrial and Systems Engineering, Kyonggi University
Suwon, 16227, South Korea
[e-mail: ksm5489@kyonggi.ac.kr (Seong-Mok Kim), dpdms925@kyonggi.ac.kr (Ye-Eun Jeong)]
[2] Major in Industrial and Systems Engineering, Kyonggi University
Suwon, 16227, South Korea
[e-mail: kimys@kgu.ac.kr]
[3] Integrated Product Support R&D, LIG Nex1
Seongnam, 13486, South Korea
[e-mail: younho.lee2@lignex1.com]
[4] C4ISTAR R&D Division 2, LIG Nex1
Seongnam, 13488, South Korea
[e-mail: seungyoung.lee@lignex1.com]
[*]Corresponding author: Seung Young Lee

Received October 29, 2023; revised December 28, 2023; revised March 20, 2024; accepted April 8, 2024; published May 31, 2024

Abstract

In modern warfare, missiles play a pivotal role but typically spend the majority of their lifecycle in long-term storage or standby mode, making it difficult to detect failures. Preemptive detection of missiles that will fail is crucial to preventing severe consequences, including safety hazards and mission failures. This study proposes a contamination-based stacking ensemble model, employing the local outlier factor (LOF), to detect such missiles. The proposed model creates multiple base LOF models with different contamination values and combines their anomaly scores to achieve a robust anomaly detection. A comparative performance analysis was conducted between the proposed model and the traditional single LOF model, using production-related inspection data from missiles deployed in the military. The experimental results showed that, with the contamination parameter set to 0.1, the proposed model exhibited an increase of approximately 22 percentage points in accuracy and 71 percentage points in F1-score compared to the single LOF model. This approach enables the preemptive identification of potential failures, undetectable through traditional statistical quality control methods. Consequently, it contributes to lower missile failure rates in real battlefield scenarios, leading to significant time and cost savings in the military industry.

Keywords: Weapon System, Missile, Quality Control, Anomaly Detection, Unsupervised Learning

This work was supported by Korea Research Institute for defense Technology planning and advancement (KIRT) – Grant funded by Defense Acquisition Program Administration (DAPA) (KRIT-CT-22-081, Weapon System CBM+ Research Center)

http://doi.org/10.3837/tiis.2024.05.009 **ISSN : 1976-7277**

1. Introduction

Missiles are precision weapons capable of rapidly and accurately attacking selected targets at a distance. This capability enables the effective execution of a wide range of military tasks, from disrupting enemy communications and supply routes to achieving air superiority and defense. Owing to this versatility, missiles play a pivotal role in modern warfare, being extensively employed for precision target engagements and strategic and tactical objectives on a global scale. With the increasing significance of missiles, the technology behind them, including various guidance systems like lasers, global positioning system (GPS), and infrared, has also evolved [1]. The integration and advancement of technology have made missile structures increasingly complex, with concentrated integration of mechanical and electronic components, raising the risk of various types of failures. However, missiles, which typically come with a 10-year warranty period, have long storage periods and are one-shot devices, making it difficult to detect failures [2]. Consequently, missile failures or malfunctions can lead to severe consequences such as safety hazards and mission failures, underscoring the need for high-reliability and safety assurance through failure prevention strategies.

Methods to prevent missile failures can be broadly categorized into macro-level and micro-level approaches. Macro-level approaches involve maintenance based on the condition of the missile, while micro-level approaches involve identification potential failures of missiles at the production phase. Studies of macro-level approaches include the development of condition monitoring systems to evaluate and enhance the reliability of missiles stored and exposed in the field [3], and reliability prediction using condition monitoring data [4]. Furthermore, with advancements in computational techniques, studies utilizing machine learning algorithms have been conducted, including missile lifespan prediction [5], failure characteristics and anomaly analysis of liquid rocket engine components [6-9], fault detection and diagnosis in solid rocket motors [10], and anomaly detection methods using flight data [11].

While study on maintenance based on the condition of missiles is actively pursued, there is a notable lack of studies identifying potential failures at the missile production phase. Particularly, over the past few years, almost no innovative models regarding quality, including potential failure identification at the production phase, have been proposed [12]. Identifying potential failures, specifically missile that will fail, at the production phase is crucial for equipment support capabilities and assessing combat mission performance. Moreover, identifying potential failures early can lead to design modifications or quality improvements, reducing additional costs that may arise and preventing system failures and safety incidents that may occur over time [13,14]. In response to these challenges, this study statistically examines quality inspection management (QIM) data acquired from the missile production phase and proposes a new methodology for detecting missiles that will fail based on the statistical analysis of this data. Missiles are classified according to solid and liquid fuels, and the proposed model's efficacy is evaluated using real data from the production phase of solid-fuel missiles. Although this study focuses on solid-fuel missiles, the proposed model is applicable to both solid and liquid fuel missiles.

The rest of the paper is organized as follows. In Section 2, an overview of missile components and functions is provided, and missile quality inspection management (QIM) is explained. This section also includes a literature review on cases of identifying potential failures within various industries. In Section 3, the distribution and deviation of the QIM data are statistically analyzed, and a novel methodology for detecting missiles that will fail is proposed. In Section 4, the proposed methodology's performance is evaluated experimentally. Finally, in Section 5, the conclusions and potential for future studies are presented.

2. Related Works

2.1 Missile Components and Functions

Missiles are self-propelled guided weapon systems that traverse through air or space, and are complex assemblies of interacting subsystems [15,16]. While missile configurations vary by type, they generally comprise components such as the guidance section, warhead section, sustainer/booster section, and control section, as illustrated by the cross-section diagram in **Fig. 1** [17]. These sections collectively enable the missile to accurately detect, track, and destroy its target.

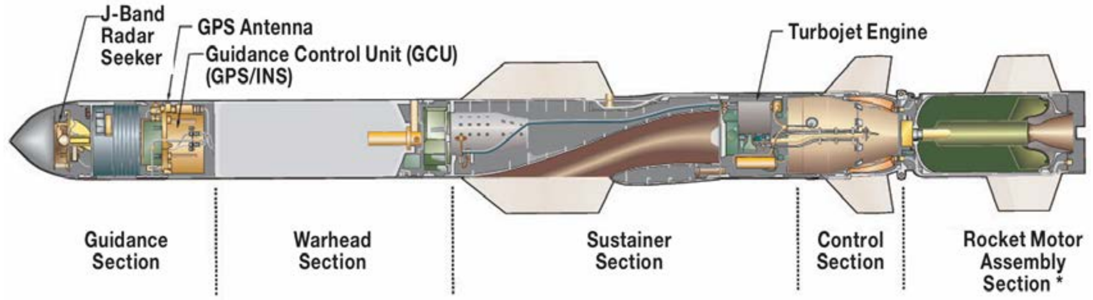


Fig. 1. The Missile Cross-Section Diagram.

The guidance section consists of complex electronic devices that enable the missile to accurately detect and track its target. It includes various sensors, computers, and algorithms to adjust the missile's flight path in real-time. The warhead section, which detonates upon reaching the target to achieve a destructive effect, can come in various forms such as high explosives, nuclear weapons, and chemical or biological warheads, with the choice depending on the missile's purpose and target. The sustainer/booster section contains rocket or jet engines that provide the necessary propulsion to carry the missile to the target, converting the chemical energy of the fuel into kinetic energy to accelerate the missile. The control section manages the missile's flight path and stability, including wings, rudders, and other control surfaces to adjust direction and altitude, ensuring the missile follows its designated path to accurately reach its target.

When classifying missile components based on their characteristics, they can be divided into mechanical items such as actuators and wings, electrical/electronic items like printed circuit board assemblies, and chemical items that include warheads and propulsion systems. A summary of the functions of missile components by category is presented in **Table 1**.

Table 1. Key Functions of the Missile by Characteristics

Category	Key Function
Electrical/Electronic	Signal processing/transmission, power conversion/distribution
Mechanical	Component protection, attitude/position control
Chemical	Power supply, detonation and explosion

2.2 Missile Quality Inspection Management

It is difficult to look at any single test result used to verify the quality of a product or system and know whether that number or result is accurate and reliable [18]. Therefore, there is a need for statistical quality control (SQC) to ascertain the statistical significance of test outcomes. SQC is a critical method for maintaining and enhancing quality in manufacturing and production processes. SQC is based on the assumption that the quality characteristics of interest in a product or process follow a normal distribution or a distribution similar to normal. It focuses on managing these characteristics. Methods to control the mean and variance within a specified range are employed to manage quality characteristics. QIM of missiles is an SQC activity conducting tests on produced missiles and managing deviations in products and processes based on test data, with the aim of minimizing these deviations. Ultimately, this QIM enhances missile quality.

2.3 Cases of Potential Failure Identification

In industries other than the missile field, studies are underway to identify potential failures at the production phase. Generally, traditional SQC is employed to detect potential failures and defects during production, but it often relies on univariate analysis, making it difficult to detect the interactions between factors affecting quality. Therefore, data-driven approaches are being actively studied because potential failures and defects that could not be detected through traditional SQC alone can be detected, and various studies are being conducted using machine learning and deep learning techniques.

Machine learning, a key component of data-driven approaches, has been increasingly adopted in recent study to identify potential failures at production phase. Susto et al. [19] compared various machine learning approaches for detecting anomalies or changes in the etch rate, a quality metric used during wafer etching in semiconductor manufacturing processes. The analysis results showed that local outlier factor (LOF) exhibited high performance in terms of precision, while angle-based outlier detection (ABOD) demonstrated excellent recall performance. Furthermore, the study of data-driven approaches for defect detection within semiconductor processes is actively being pursued [20-22]. Ko et al. [23] integrated manufacturing and inspection data with after-sales data, and then assessed product quality during the process using machine learning algorithms. To address the issue of class imbalance, a single-class classification algorithm was employed in conjunction with a machine learning-based anomaly detection model to calculate anomaly scores for each engine. Consequently, the approach enabled successful detection of defective engines with high probability before shipment. Kharitonov et al. [24] conducted an evaluation of ten machine learning models for the detection of anomalies in the manufacturing field, including unsupervised learning methods such as k-nearest neighbors (KNN), Autoencoder, LOF, and COPOD. These methods do not require prior knowledge of abnormal patterns but rely on understanding the characteristics and structure of the data. However, the limited applicability of these unsupervised learning approaches was also highlighted due to their heavy dependence on the shape or distribution of data. Alimohammadi et al. [25] employed classifier-based, angle-based, and proximity-based machine learning techniques for the detection of anomalies in gas production time-series data, with methods such as KNN, DBSCAN, and ABOD demonstrating excellent performance. Alelaumi et al. [26] utilized a sliding window to extract features for the detection of in-process anomalies in order to prevent print defects and reduce high rework costs in the manufacturing process, and patterns were detected using an AdaBoost model.

Deep learning, another pillar of data-driven methodologies, is also gaining traction in identifying potential failures at production phase. Liu et al. [27] developed a structured neural network by combining an event ordering relationship-based structuring technique and deep neural networks to detect anomalies in the manufacturing process using process data collected from 151 sensors. Through the event ordering relationship-based neural network structuring process, which involves determining important neuron connections and weight initialization before neural network training, the reduction in the complexity of the neural network led to an improvement in anomaly detection accuracy. The structured neural network demonstrated higher performance compared to widely used methods. Gao et al. [28] recognized that features extracted using AI methods may have limited discriminative power in anomaly detection due to slight variations in operational parameters. To address this issue and tackle the problem of imbalanced data, an integrated approach was developed that combines statistical feature extraction with unsupervised deep learning. Jakubowski et al. [29] aimed to identify the causes of anomalies and validate their findings using the SHAP method, which provides both local and global explanations. This method was applied to actual data collected during the hot rolling industrial process, utilizing a modified autoencoder architecture with long short-term memory layers.

In industries other than the missile field, data-driven approaches to identifying potential failures at the production phase are being actively studied, targeting improvements in yield, time savings, and cost reduction. Such studies encompass model performance evaluation, data preprocessing, and training methodologies. Given the outcomes of studies in these industries, it is crucial to conduct similar studies that can ensure a high level of reliability and safety by preventing failures within the missile field.

3. Proposed Methodology

The objective of this study is to detect missiles that will fail at the production phase in order to enhance missile reliability. In this section, the characteristics of the produced missiles are examined using statistical data analysis, and a model is proposed for detecting missiles that will fail based on this analysis.

3.1 Data Analysis

3.1.1 Dataset

Utilizing the QIM data of ○○ missiles, which were produced over a continuous period of approximately seven years and subsequently deployed in the military, the objective is to differentiate between normal missiles and those that will fail. The ○○ missile is a solid-fuel missile, and the QIM data comprises test results for various inspection items conducted on the missile immediately after production. This dataset encompasses 444 inspection items, including metrics such as connection resistance, detection performance, and squib-related tests, and is based on measurements from a total of 138 missiles. Due to security considerations, inspection item names are labeled as x1, x2, etc., and neither the measurement units nor the pass criteria are disclosed.

All missiles were found to be within the quality control limits based on measurements from post-production inspections and were thus successfully deployed in the military. After this deployment, annual sample inspections are conducted for each missile, enabling assessments of both operational status and failures across all inspection items to be made. From the sample inspection, failures were observed in 9 out of 138 missiles. All these failures were identified as simultaneous malfunctions of both the seeker and the inertial measurement unit (IMU), with no other failures detected. The seeker is responsible for target detection, the IMU is designed to sense the missile's movement. The malfunction in these components can cause the missile to deviate from its intended path, potentially leading to military operation failures or additional unintended damages.

To ascertain if any process issues were related to these failures, the normality of the inspection items in the QIM data was analyzed, and the histograms of both the normal missiles and those that will fail were compared. The results are visualized in **Fig. 2**. After the Anderson-Darling normality test was performed for all inspection items, all were found to satisfy normality at a significance level of 0.05.

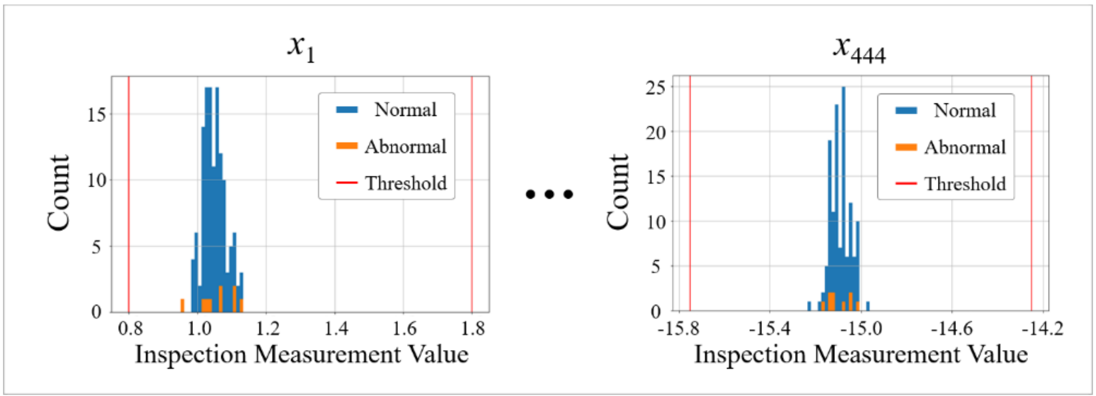


Fig. 2. Histogram Comparison of QIM Data between Normal Missiles and Those that Will Fail.

3.1.2 Failure Classification

When visualized as histograms using the QIM data, a broader deviation was exhibited by missiles that will fail compared to normal missiles. Upon conducting Bartlett's test, which is used to test if k samples have equal variances, for variances between normal missiles and those that will fail, significant differences in variances were observed in 312 out of 444 inspection items at a significance level of 0.05. Notably, for these 312 items, the missiles that will fail exhibited greater deviations than the normal missiles. This suggests that for inspection items, the measurement values for normal missiles tend to cluster closer to the mean, with density decreasing as values deviate from the mean. In **Fig. 3**, the approximate probability density functions of occurrences for both the normal missiles and those that will fail are illustrated, denoting the differences in their densities.

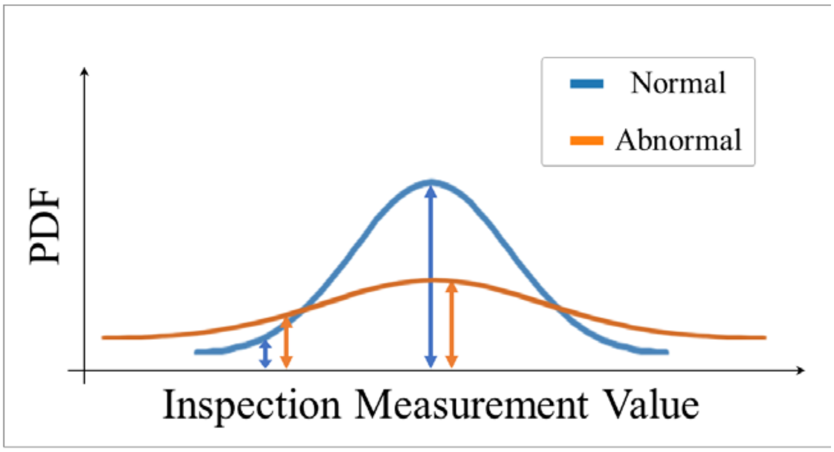


Fig. 3. Approximate Probability Density Function between Normal Missiles and Those that Will Fail

As the measurement values for the inspection items deviate further from the mean, the likelihood of missile failure is increased. It is inferred that missiles with a greater number of inspection items deviating significantly from the mean are more susceptible to failure. Therefore, missiles that will fail are identified based on the count of inspection items that fall outside a specific deviation range. Relying solely on accuracy to set the optimal classification criterion might result in imprecise outcomes due to the disparity in numbers between normal missiles and those that will fail. Given that the number of missiles that will fail is significantly lower than normal missiles, relying solely on accuracy to set the optimal classification criterion can result in most predictions being categorized as normal missiles, thereby achieving high accuracy but inadequately predicting missile that will fail. The F1-score, a metric representing the harmonic mean of Precision and Recall, mitigates this issue. Precision is the ratio of correctly predicted positive observations to all predicted positives, assessing the accuracy of positive predictions. Recall is the ratio of correctly predicted positive observations to all actual positives, evaluating the model's ability to identify all relevant cases. This metric is particularly effective in evaluating the predictive performance of failure-prone missiles in imbalanced datasets. Consequently, the F1-score is utilized to assess and establish the optimal classification criterion. When the criterion for deviation is defined as k and the criterion for the count of inspection items surpassing k as T, the optimal criterion (k^*, T^*) yielding the highest F1-score can be deduced through (1)-(5).

$$\bar{x}_m = \frac{1}{138}\sum_{i=1}^{138} x_{mi} \tag{1}$$

$$s_m = \sqrt{\frac{1}{137}\sum_{i=1}^{138}(x_{mi} - \bar{x}_m)^2} \tag{2}$$

$$C_{mj} = I\left(x_{mj} < \bar{x}_m - ks_m \ or \ x_{mj} > \bar{x}_m + ks_m\right) \tag{3}$$

$$C_j = \sum_{m=1}^{444} C_{mj} \tag{4}$$

$$(k^*, T^*) = \underset{k,T}{\text{argmax}} F1(k, T) \tag{5}$$

The term x_{mi} represents the measurement value of the m-th inspection item for the i-th missile. $\bar{x}_m$ and s_m denote the mean and standard deviation of the m-th inspection item measurement value, respectively. I is an indicator function, and C_{mj} is an indicator showing whether the measurement value of the m-th inspection item for the j-th missile deviates from the deviation criterion. C_j represents the number of inspection items that deviate from the deviation criterion for the j-th missile. If the number of inspection items that surpass the deviation criterion exceeds the criterion T, the respective missile is classified to be faulty. $F1(k, T)$ calculates the F1-score for a given combination of k and T. Through this process, the optimal criterion for ○○ missiles were derived with k^* being 1.02 and T^* being 76. The confusion matrix for this criterion is presented in **Table 2**, and the classification performance is summarized in **Table 3**.

Table 2. Confusion Matrix for the Optimal Criterion

		Predicted Class	
		Normal	Abnormal
Actual Class	Normal	125	4
	Abnormal	4	5

Table 3. Classification Performance for the Optimal Criterion

Accuracy	**F1-Score**	**Precision**	**Recall**
0.9420	0.5556	0.5556	0.5556

While the accuracy was notably high at approximately 94%, the F1-score displayed a lower performance at around 56%. This is attributed to the rare occurrences of failures from missiles and the problem of using the same deviation standard across all inspection items, disregarding their unique characteristics. To address this challenge, this study employs the local outlier factor (LOF) from unsupervised learning models. LOF considers the density of local sets and the degree to which objects are isolated from their neighbors.

3.2 Local Outlier Factor

3.2.1 Dataset and Preprocessing

Out of the 138 ○○ missiles, 9 have been confirmed to have failed. The QIM data for the ○○ missile is partitioned into training and testing datasets at an 8:2 ratio, yielding 110 data points for training and 28 for testing. The proposed model is an unsupervised anomaly detection model that trains exclusively on normal missile data. Therefore, the training dataset consists solely of normal missile data, while the test dataset includes all data related to missiles that will fail.

In the QIM data for the ○○ missile, various inspection items result in a diversity of measurement ranges and units. To address this inconsistency, min-max normalization should be applied. The normalization equation is given in (6).

$$x_{mi}^{*} = \frac{x_{mi} - \min(x_m)}{\max(x_m) - \min(x_m)} \tag{6}$$

3.2.2 Anomaly Detection

LOF addresses the challenge in traditional density-based methods where defining a density threshold is difficult due to clusters having varying densities. In the LOF framework, a pivotal metric, *k-distance*(*p*), is introduced. This metric represents the distance to the *k*-th nearest neighbor of a specific data point, *p*. It is possible for multiple data points to have the same *k-distance*(*p*). Hence, all points within this distance, excluding *p* itself, are designated as $N_k(p)$. For another data point, *o*, and its *k-distance*(*o*), the concept of reachability distance between *p* and *o* is introduced. This distance, termed *reachability-distance*(*p*, *o*), is defined in (7).

$$reachability - distance(p, o) = \max\big(k - distance(o), dist(p, o)\big) \tag{7}$$

Using the local reachability distance, *reachability-distance*(p, o), the local reachability density, *lrd*(p), is defined as per (8). Taking into account the local densities and the inverse relationship between distance and density, it is calculated as the inverse of the distance.

$$lrd(p) = \left[\frac{\sum_{o \in N_k(p)} reachability - distance_k(p, o)}{N_k(p)}\right]^{-1} \tag{8}$$

For the data point p, its local reachability density, *lrd*(p), and the ratio of points belonging to $N_k(p)$ are considered to evaluate the density. For this purpose, the *lof*(p) is defined, as in (9), by comparing the average distance between p and its neighbors to the average distance among the neighbors in $N_k(p)$.

$$lof(p) = \frac{\sum_{o \in N_k(p)} \frac{lrd(o)}{lrd(p)}}{N_k(p)} \tag{9}$$

In instances where surrounding points around a specific point do not exhibit significant differences, the *lof*(p) value is approximated to 1. Conversely, for points that deviate significantly from the average distance, the *lof*(p) value exceeding 1 is observed. Through this model, the detection of missiles that will fail based on QIM data is attempted. In this study, all parameters of the LOF model were adopted as initially proposed by Breunig et al. [30]. The results of the detection using this model are presented in **Table 4**.

Table 4. Anomaly Detection Performance Using the LOF Model

Accuracy	F1-Score	Precision	Recall
0.7314	0.1990	0.5359	0.1333

A decrease in overall performance was observed when relying solely on the LOF model compared to the deviation criterion-based classification method. This decline might be attributed to the high-dimensional nature of the QIM data, which consists of 444 inspection items. In high-dimensional data, the curse of dimensionality can occur. In such cases, data points are likely to be more widely dispersed, making density estimation difficult and leading to challenges for density-based anomaly detection models like LOF. To overcome these challenges and improve performance in this study, a contamination-based stacking ensemble model is proposed, which stacks various LOF models with different contamination values.

3.3 Contamination-Based Stacking Ensemble

This study proposes a contamination-based stacking ensemble model for detecting missiles that will fail. This model leverages the LOF, a renowned anomaly detection technique, and conducts its analysis using QIM data, which is normalized prior to training on individual sub-LOF models. The anomaly scores from these sub-models are subsequently combined through a stacking ensemble. The proposed model carries out unsupervised anomaly detection, utilizing inspection results obtained during the production of missiles that are deployed in military. The anomaly detection framework for this study is depicted in **Fig. 4**.

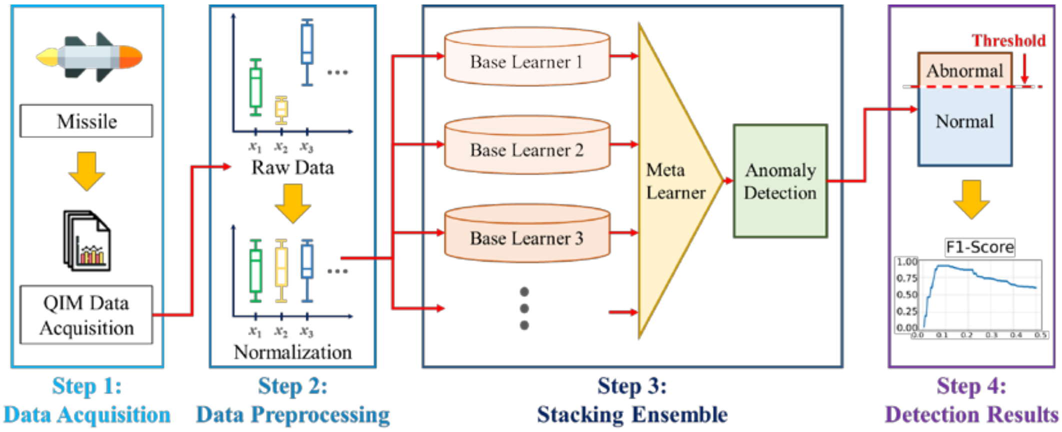


Fig. 4. The Overall Framework of the Anomaly Detection

The dataset and data preprocessing used for this model are consistent with those described in Section 2.2.1. The proposed model generates multiple sub-LOF models, each with a unique contamination value, from the single anomaly detection model LOF. In LOF models, when a point does not significantly differ from its neighboring points, the LOF value approximates to 1, while it exceeds 1 for points significantly different from the mean distance. Determining the threshold for considering how much greater than 1 the value should be to classify a point as an outlier is a challenging task, and various studies have been conducted on this matter [31]. In Python's Scikit-learn library, this threshold is set through the contamination parameter. A lower contamination value indicates fewer outliers, resulting in a higher threshold, while a higher contamination value implies a lower threshold. The anomaly score of the data also varies with different contamination values. This study proposes a contamination-based stacking ensemble model that accumulates anomaly scores produced differently according to this parameter and employs them for anomaly detection. Except for contamination, the other parameter settings follow the original configuration of the LOF model as proposed by Breunig et al. [30].

In this study, the training data involves individually training sub-LOF models generated based on different contamination values and stacking the estimated anomaly scores to create training stacking data. Training and estimation are conducted through k-fold cross-validation, preventing overfitting for each sub-model and generating a larger amount of stacking data. The test data is applied to the sub-models trained through previous k-fold cross-validation, resulting in k duplicate anomaly scores for each model. The estimated anomaly scores represent the average per missile, and by stacking the average anomaly scores, testing stacking data is generated. Utilizing the generated training stacking data and testing stacking data, anomaly detection is performed through a single LOF model. The architecture of the proposed model is illustrated in **Fig. 5**.

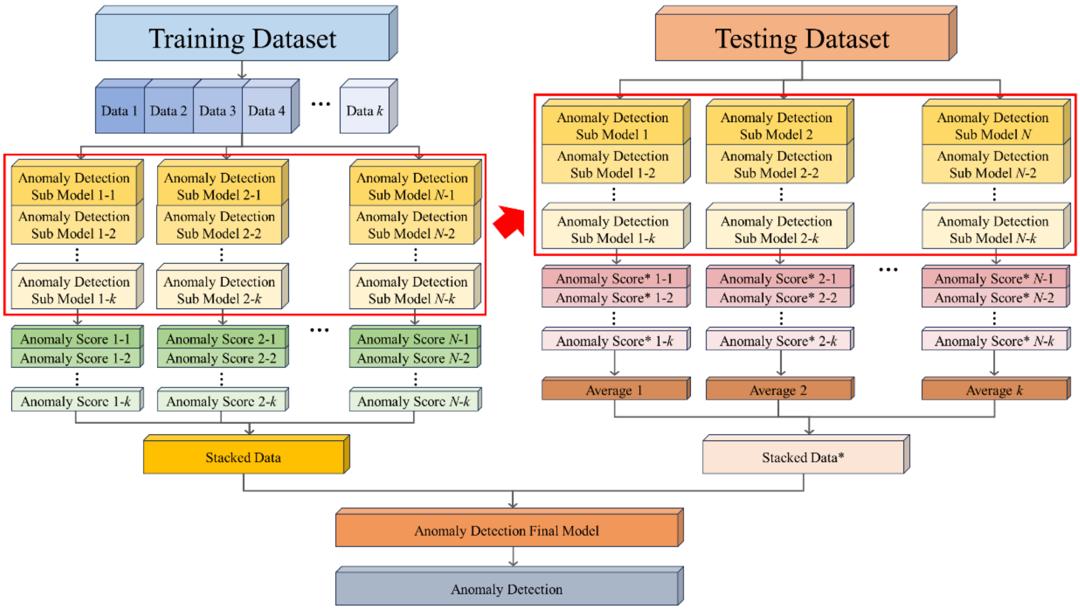


Fig. 5. The Architecture of the Contamination-Based Stacking Ensemble Model

In **Fig. 5**, *N* represents the number of unique contamination values, *k* denotes the value of *k* in *k*-fold cross-validation, sub-models refer to the base learners, and the final model represents the meta-learner. Here, the contamination values are calculated as shown in (10), and c_i signifies the contamination value of the *i*-th sub-model.

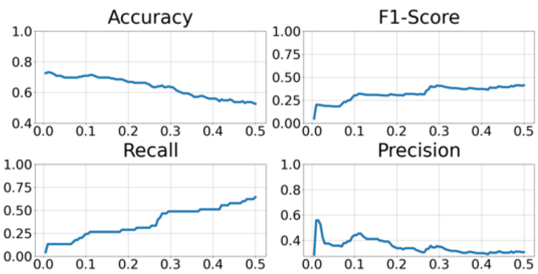
$$c_i = \frac{i}{2N} \quad (1 \leq i \leq N) \tag{10}$$

The model proposed in this study generates base learners with a variety of contamination values, stacks each base learner's anomaly scores, and the data thus stacked is used for anomaly detection through a single meta-learner. This model detects missiles that will fail during the production phase using an unsupervised learning-based approach. While the traditional LOF model may be limited when applied to high-dimensional data due to its susceptibility to data dimensionality, the proposed model integrates models with various thresholds, ranging from lenient to strict, to overcome issues arising in high-dimensional data. This integration results in a more stable and robust performance compared to the traditional LOF model. As a result, this model significantly enhances the performance of detecting missiles that will fail compared to applying the LOF model independently.

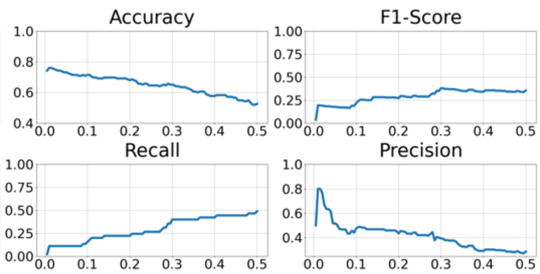
4. Experimental Evaluation

The performance and training time of the contamination-based stacking ensemble model for anomaly detection are evaluated in this section. Given the imbalance between the number of normal missiles and those that will fail, the F1-score has been selected as the evaluation metric to account for the recall and sensitivity. The proposed model's performance is assessed based on both accuracy and the F1-score, to ensure a comprehensive evaluation.

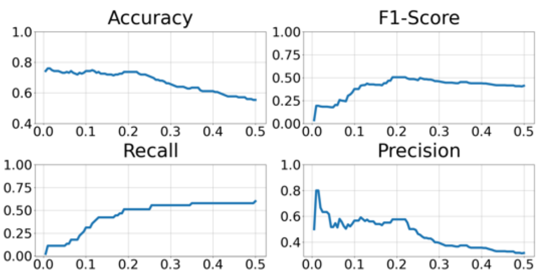
Fig. 6 displays the average accuracy, F1-score, recall, and sensitivity obtained through a *k*-fold cross validation with *k*=5, comparing the traditional LOF model to the one with varying numbers of base learners, n. In subfigure (a), the x-axis represents the contamination value, and the y-axis illustrates the respective performance metrics. For subfigures (b) through (g), the x-axis shows the contamination values of the meta learner, while the y-axis illustrates the respective performance metrics.



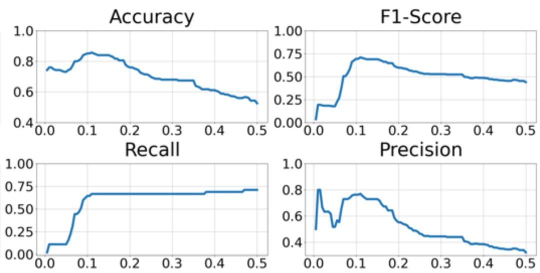
(a) LOF (Original)



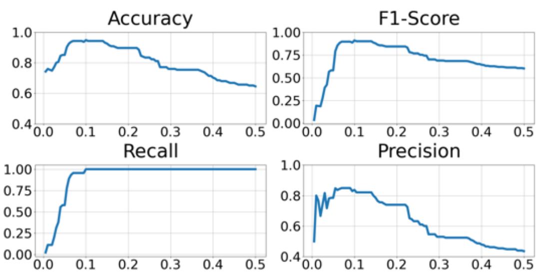
(b) Proposed Model (n=1)



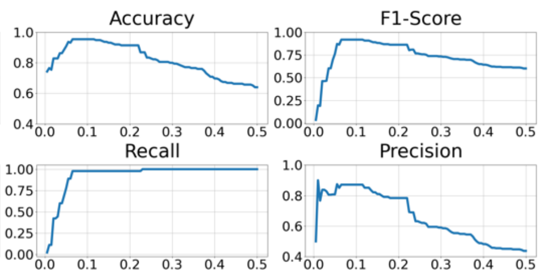
(c) Proposed Model (n=10)



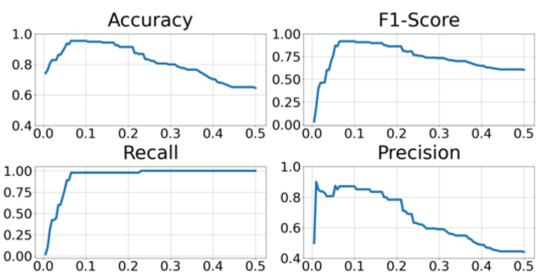
(d) Proposed Model (n=50)



(e) Proposed Model (n=100)



(f) Proposed Model (n=500)



(g) Proposed Model (n=1,000)

Fig. 6. Performance of the LOF Model and the Proposed Model with Varying Base Learner Counts

The proposed model shows performance improvement up to n=100, but beyond that point, there's only a minimal gain in performance while the learning time increases significantly. The highest average performance of the proposed model was observed at contamination parameter of 0.1, and the average accuracy, F1-score, and average learning time at contamination parameter of 0.1 are summarized in **Table 5**.

Table 5. Performance Evaluation of Models with a Contamination Parameter of 0.1

Index		Accuracy (%)	F1-Score (%)	Train Time (sec)
LOF Only		73.14	19.31	0.4
Contamination-Based Stacking Ensemble Model (Proposed)	n=1	71.43	21.56	0.7
	n=10	74.29	37.58	4.4
	n=50	85.14	69.37	21.3
	n=100	94.86	91.04	41.7
	n=500	95.43	91.87	206.3
	n=1000	95.43	91.87	411.9

At the contamination parameter of 0.1, the proposed model showed an approximate increase of 22 percentage points in accuracy and 71 percentage points in the F1-score compared to the traditional LOF model. This indicates the capability of the proposed model to detect missiles that will fail with a high performance.

5. Conclusion

Missiles are considered crucial weaponry in modern warfare, but due to their disposable and long-term storage characteristics, identifying faults can be challenging. To address these challenges, unlike previous studies, which focused solely on maintenance based on the condition of missiles, this study preemptively detected missiles that will fail. Specifically, the study focused on analyzing QIM data of missiles and, through the analysis results, identified distribution and density differences between normal missiles and those that will fail. Based on these findings, a new model for detecting missiles that will fail during the production phase was proposed.

This study utilized QIM data from missiles deployed in military and proposed a contamination-based stacking ensemble model based on the LOF model. This model generates multiple sub-LOF models with different contamination values. By considering various thresholds, a comprehensive approach is enabled by the model, resulting in performance that is more stable and robust than that of a single LOF model. Its effectiveness was demonstrated through a comparison with a single LOF model. The proposed model holds promise for applications beyond missile failure prediction, extending to various industries for anomaly detection. Furthermore, it is expected that this study will significantly contribute to reducing missile failure rates in real battlefield scenarios.

This study confirmed the capability to identify the specific missile failure preemptively. However, it relies on the quality and completeness of QIM data. Future studies will aim to address these limitations by exploring more adaptive and resilient modeling approaches. Additionally, future work will integrate these identification results with sensor data and operational missile information to obtain more accurate and diverse state information.

References

[1] J. R. Hoehn and S. Ryder, *Precision-Guided Munitions: Background and Issues for Congress*, Congressional Research Service, 2021.
[2] M. Zhao and M. Xie, "A model of storage reliability with possible initial failures," *Rel. Eng. & Syst. Saf.*, vol. 43, no. 3, pp. 269-273, 1994. Article (CrossRef Link)
[3] A. J. Kudiya and S. A. Marotta, "Tactical Missile Health Management," in *Syst. Health Manage: With Aerosp. Appl.*, New jersey, NJ, USA: Wiley, 2021, pp. 555-564. Article (CrossRef Link)
[4] S. A. Marotta, A. Kudiya, T. K. Ooi, H. A. Toutanji, and J. A. Gilbert, "Predictive reliability of tactical missiles using health monitoring data and probabilistic engineering analysis," in *Proc. of 1st Int. Forum on Integr. Syst. Health Eng. and Manage. in Aerosp.*, 2005.
[5] A. Lewis, P. Nalliah, C. Lomax, and C. Hawkins, "A methodology using health and usage monitoring system data for payload life prediction," in *Proc. of ISMA 2018 Int. Conf. on Noise and Vib. Eng.*, Leuven, Belgium, Sep. 2018.
[6] W. Zhang, G. Tian, Z. Xu, and Z. Yang, "Fault diagnosis of LRE based on Ann," in *Failure Characteristics Anal. and Fault Diagnosis for Liquid Rocket Engines, Berlin, BER, DEU: Springer*, 2016, pp. 121-164. Article (CrossRef Link)
[7] T. Wang, L. Ding, and H. Yu, "Research and Development of Fault Diagnosis Methods for Liquid Rocket Engines," *Aerosp*., vol. 9, no. 9, Aug. 2022. Article (CrossRef Link)
[8] S. Y. Park and J. Ahn, "Deep neural network approach for fault detection and diagnosis during startup transient of liquid-propellant rocket engine," *Acta Astronaut*., vol. 177, pp. 714-730, Dec. 2020. Article (CrossRef Link)
[9] Y. Feng, J. Chen, Z. Liu, H. Lv, and J. Wang, "Full Graph Autoencoder for One-Class Group Anomaly Detection of IIoT System," *IEEE Internet Things J*., vol. 9, no. 21, pp. 21886-21898, Jun. 2022. Article (CrossRef Link)
[10] D. Liu, L. Sun, and T. C. Miller, "Defect Diagnosis in Solid Rocket Motors Using Sensors and Deep Learning Networks," *AIAA J*., vol. 59, no. 1, pp. 276-281, Oct. 2021. Article (CrossRef Link)
[11] L. Li, S. Das, J. Hansman, R. Palacios, and A. N. Srivastava, A. N, "Analysis of flight data using clustering techniques for detecting abnormal operations," *J. Aerosp. Inf. Syst.*, vol. 12, no. 9, pp. 587-598, 2015. Article (CrossRef Link)
[12] A. Zonnenshain and R. S. Kenett, "Quality 4.0—the challenging future of quality engineering," *Qual. Eng.*, vol. 32, no. 4, pp. 614-626, Feb. 2020. Article (CrossRef Link)
[13] L. Irshad, S. Ahmed, H. O. Demirel, and I. Y. Tumer, "Computational Functional Failure Analysis to Identify Human Errors During Early Design Stages," *J. Comput. Inf. Sci. Eng.*, vol. 19, no. 3, p. 031005, Sep. 2019. Article (CrossRef Link)
[14] S. J. Ghoushchi, M. N. A. Rahman, D. Raeisi, E. Osgooei, and M. J. Ghoushji, "Integrated Decision-Making Approach Based on SWARA and GRA Methods for the Prioritization of Failures in Solar Panel Systems under Z-Information," *Symmetry*, vol. 12, no. 2, p. 310, Feb. 2020. Article (CrossRef Link)
[15] P. Sethunathan, A. Anupriya, K. H. Prabha, and T. Vaishnavi, "Aerodynamic Characterisitics of a Missile Components," *Int. J. Res. Eng. Technol.*, vol. 4, no. 3, pp. 100-104, Mar. 2015.
[16] E. G. C. Burt, "Theoretical principles of guided missile systems," *The Aeronautical J.*, vol. 63, no. 584, pp. 455-464, Aug. 1959. Article (CrossRef Link)
[17] Boeing, "Harpoon Product Card," *Boeing*. [Online]. Available: https://www.boeing.com/content/dam/boeing/boeingdotcom/defense/weapons-weapons/images/harpoon_product_card.pdf, Accessed on: Mar. 10, 2024.
[18] D. C. Montgomery, *Introduction to Statistical Quality Control*, Hoboken, NJ, USA:Wiley, 2019.
[19] G. A. Susto, M. Terzi, and A. Beghi, "Anomaly Detection Approaches for Semiconductor Manufacturing," *Procedia Manuf*., vol. 11, pp. 2018-2024, 2017. Article (CrossRef Link)
[20] S. K. Fan, C. Y. Hsu, D. M. Tsai, F. He, and C. C. Cheng, "Data-driven approach for fault detection and diagnostic in semiconductor manufacturing," *IEEE Trans. Autom. Sci. Eng*., vol. 17, no. 4, pp. 1925-1936, Apr. 2020. Article (CrossRef Link)

[21] K. Kerdprasop and N. Kerdprasop, "A Data Mining Approach to Automate Fault Detection Model Development in the semiconductor manufacturing process," *Int. J. Mech.*, vol. 5, no. 4, pp. 336-344, 2011.
[22] S. H. Kim, C. Y. Kim, D. H. Seol, J. E. Choi, and S. J. Hong, "Machine Learning-Based Process-Level Fault Detection and Part-Level Fault Classification in Semiconductor Etch Equipment," *IEEE Trans. Semicond. Manuf.*, vol. 35, no. 2, pp. 174-185, 2022. Article (CrossRef Link)
[23] T. H. Ko, J. H. Lee, H. C. Cho, S. z. Cho, W. J. Lee, and M. J. Lee, "Machine learning-based anomaly detection via integration of manufacturing, inspection and after-sales service data," *Ind. Eng. Manag. Syst.*, vol. 117, no. 5, pp. 927-945, Mar. 2017.
[24] A. Kharitonov, A. Nahhas, M. Pohl, and K. Turowski, "Comparative analysis of machine learning models for anomaly detection in manufacturing," *Procedia Comput. Sci.*, vol. 200, pp. 1288-1297, 2022. Article (CrossRef Link)
[25] H. Alimohammadi and S. N. Chen, "Performance evaluation of outlier detection techniques in production timeseries: A systematic review and meta-analysis," *Expert Syst. Appl.*, vol. 191, Apr. 2022. Article (CrossRef Link)
[26] S. Alelaumi, H. Wang, H. Lu, and S. W. Yoon, "A Predictive Abnormality Detection Model Using Ensemble Learning in Stencil Printing Process," *IEEE Trans. Compon. Packaging. Manuf. Technol.*, vol. 10, no. 9, pp. 1560-1568, Jul. 2020. Article (CrossRef Link)
[27] J. Liu, J. Guo, P. Orlik, M. Shibata, D. Nakahara, S. Mii, and M. Takáč, "Anomaly Detection in Manufacturing Systems Using Structured Neural Networks," in *Proc. of 2018 13th World Congr. on Intell. Control and Automat. (WCICA)*, Changsha, China, pp. 175-180, Jan. 2018. Article (CrossRef Link)
[28] Y. Gao, X. Yin, Z. He, and X. Wang, "A deep learning process anomaly detection approach with representative latent features for low discriminative and insufficient abnormal data," *Comput. Ind. Eng.*, vol. 176, Feb. 2023. Article (CrossRef Link)
[29] J. Jakubowski, P. Stanisz, S. Bobek, and G. J. Nalepa, "Explainable anomaly detection for Hot-rolling industrial process," in *Proc. of 2021 IEEE 8th Int. Conf. on Data Sci. and Adv. Analytics (DSAA)*, Porto, Portugal, pp. 1-10, Oct. 2021. Article (CrossRef Link)
[30] M. M. Breunig, H.-P. Kriegel, R. T. Ng and J. Sander, "LOF: Identifying density-based local outliers," in *Proc. of ACM SIGMOD*, pp. 93-104, May. 2000. Article (CrossRef Link)
[31] V. P. Kumaravel, M. Buiatti, E. Parise, and E. Farella, "Adaptable and Robust EEG Bad Channel Detection Using Local Outlier Factor (LOF)," *Sensors*, vol. 22, no. 19, pp. 7314, Sep. 2022. Article (CrossRef Link)

Seong-Mok Kim received the B.S. degree from the Department of Industrial and Management Engineering, Kyonggi University, South Korea. He is currently pursuing the M.S. degree with the Department of Industrial and System Engineering. His research interests include reliability data analysis, reliability test planning, and degradation modeling.

Ye-Eun Jeong received the B.S. degree from the Department of Industrial and Management Engineering, Kyonggi University, South Korea. She is currently pursuing the M.S. degree with the Department of Industrial and System Engineering. Her research interests include data mining, reliability, and statistics.

Yong Soo Kim is a Professor at the Department of Industrial and Management Engineering, Kyonggi University, Korea. He received B.S., M.S., and Ph.D. degrees from KAIST, respectively. His research interests include data mining and reliability.

Youn-Ho Lee received the M.S. degree in Industrial Engineering from Sungkyunkwan University, Suwon, Republic of Korea, in 2010. He is currently working for LIG Nex1. His research interests relate to life cycle management and CBM+ for ammunition and missile.

Seung Young Lee is a CTO at the LIG Nex1, South Korea. He received B.S. (Electronics). and M.S. (Computer Science) degrees from Inha University, I specialize in the development and planning of cutting-edge technologies in the realms of C4I, cyber security, electromagnetic warfare, artificial intelligence, unmanned systems including satellites and ground stations, and avionics electronic systems.