Journal of Intelligent Manufacturing and Special Equipment

A physics-guided unified damage index framework for explainable milling anomaly detection using interval-wise MTConnect multi-sensor fusion

Received 29 May 2026
Revised 30 June 2026
Accepted 11 August 2026

Aswin Karkadakattil
Department of Mechanical, Independent Researcher, Kanhangad, India

Abstract

Purpose – Reliable anomaly detection in precision milling is essential for maintaining dimensional accuracy, tool life, surface integrity and component reliability. Conventional data-driven approaches often average sensor signals over entire machining records and rely on black-box machine learning models, reducing physical interpretability and limiting industrial deployment. This study proposes a physics-guided Unified Damage Index (UDI) framework integrating interval-wise MTConnect multi-sensor fusion with explainable machine learning for interpretable milling anomaly detection.

Design/methodology/approach – MTConnect sensor signals were extracted only from active cutting intervals using machining start and end timestamps to eliminate idle machine states. Process parameters, including spindle speed, feed rate, depth of cut and tool geometry, were combined with spindle vibration, axis vibration, spindle load and power consumption to formulate a physics-guided Unified Damage Index. An XGBoost classifier was developed for anomaly prediction, while SHAP (SHapley Additive exPlanations) interpreted feature contributions. Model robustness was evaluated using repeated stratified five-fold cross-validation and bootstrap confidence interval analysis.

Findings – The proposed framework achieved 93.75% accuracy, 91.7% precision, 100% recall and a 95.0% F1-score on the independent test set without missing anomaly cases. Repeated cross-validation yielded a mean accuracy of 88.85 $\pm$ 8.04%, demonstrating stable predictive performance. SHAP and feature importance analyses identified spindle vibration instability and X- and Y-axis vibration as the dominant indicators of machining anomalies.

Originality/value – The framework integrates interval-wise MTConnect feature extraction, a physics-guided Unified Damage Index, explainable XGBoost modelling and statistical robustness analysis into a unified predictive maintenance approach, providing an interpretable solution for intelligent manufacturing and digital twin applications.

Keywords Milling anomaly detection, Unified damage index, MTConnect, Explainable artificial intelligence, SHAP, Multi-sensor fusion, Predictive maintenance, Smart manufacturing, Metallic components, Physics-guided modelling

Paper type Research article

1. Introduction

The rapid advancement of Industry 4.0 has transformed modern manufacturing by enabling intelligent, connected, and data-driven production systems. Among the key objectives of this

Funding: The author received no specific financial support for the research, authorship, or publication of this article.

Conflicts of interest: The author declares that there is no conflict of interest regarding the publication of this paper.

Journal of Intelligent Manufacturing and Special Equipment
Emerald Publishing Limited
e-ISSN: 2633-660X
p-ISSN: 2633-6596
DOI 10.1108/JIMSE-05-2026-0020

transformation are improved process reliability, enhanced product quality, reduced downtime, and predictive maintenance through continuous monitoring of machining operations. Precision milling is one of the most widely used manufacturing processes for producing high-value metallic components with stringent dimensional and surface quality requirements. However, machining anomalies such as chatter, excessive vibration, spindle instability, tool wear, and unexpected tool failure can adversely affect dimensional accuracy, surface integrity, tool life, and the structural reliability of machined components. These issues are particularly critical in aerospace, automotive, biomedical, and precision engineering industries, where even minor machining deviations can lead to costly component rejection, unplanned machine downtime, and potential safety risks (Kim *et al.*, 2026; Rastegari *et al.*, 2017). Consequently, there is an increasing demand for intelligent anomaly detection systems that can identify machining abnormalities at an early stage while providing reliable information to support predictive maintenance and process optimization. Traditional machining monitoring methods primarily rely on periodic inspection, operator expertise, or threshold-based monitoring of individual sensor signals. Although these approaches can identify severe failures after they occur, they are generally inadequate for detecting the early stages of machining degradation or supporting proactive maintenance decisions. Recent advances in sensor technology and industrial communication standards have significantly improved the availability of real-time machine data. In particular, MTConnect has emerged as a widely adopted open communication standard that enables standardized acquisition of machine tool information, including spindle power, spindle speed, vibration signals, axis load, energy consumption, and machine operating states (Navas *et al.*, 2021; Kim *et al.*, 2025; Student and Raman, 2025). The availability of synchronized multi-sensor information through MTConnect provides new opportunities to develop intelligent monitoring frameworks capable of continuously assessing machining health and supporting data-driven decision-making in smart manufacturing environments. Machine learning has become an increasingly important tool for analyzing machining data and detecting process anomalies. Algorithms such as Support Vector Machines (SVM), Random Forests (RF), Artificial Neural Networks (ANN), and Extreme Gradient Boosting (XGBoost) have demonstrated strong predictive capabilities in applications including tool condition monitoring, tool wear prediction, and machining fault diagnosis (Alajmi and Almeshal, 2020; Przybyś-Małaczek *et al.*, 2023; Pusuluri *et al.*, 2025). Among these techniques, XGBoost has attracted considerable attention because of its ability to model complex nonlinear relationships, handle heterogeneous industrial datasets, and maintain robust predictive performance with structured process data. At the same time, vibration-based monitoring has been widely recognized as one of the most informative indicators of machining stability, as changes in spindle and axis vibration are closely associated with chatter initiation, tool wear progression, spindle imbalance, and structural dynamic behaviour during cutting (Dai *et al.*, 2015; Bilgin, 2026). Despite these advances, important challenges remain before machine learning can be reliably deployed for industrial machining anomaly detection. Many existing studies continue to rely on simplified feature extraction strategies and purely data-driven prediction models, which often limit physical interpretability and reduce confidence in practical industrial applications. These limitations motivate the development of anomaly detection frameworks that combine engineering knowledge with explainable artificial intelligence, enabling both accurate prediction and physically meaningful interpretation of machining behaviour.

Although considerable progress has been made in applying machine learning to machining monitoring, several important challenges continue to limit the reliability, interpretability, and industrial adoption of existing anomaly detection frameworks.

The first limitation concerns the way machining sensor data are typically processed. Many existing studies compute statistical features by averaging sensor signals over the entire machining record, regardless of whether the machine is actively cutting material. Consequently, machine states such as spindle idling, rapid traverse movements, tool approach, and tool retraction are included together with the actual cutting process. Since these non-cutting operations are governed by different physical conditions, their inclusion can dilute the characteristics of true machining behaviour and reduce the physical relevance of the extracted features. As a result, anomaly detection models may learn patterns that are unrelated to the cutting process itself, limiting both prediction reliability and engineering interpretability. The recently published MTConnect multi-sensor milling dataset developed by Kim *et al.* (2026) provides precise machining start and end timestamps specifically to distinguish active cutting intervals from non-cutting machine operations. Nevertheless, this valuable information has rarely been fully utilized in existing anomaly detection studies, leaving significant scope for more physically meaningful feature extraction strategies.

A second challenge is the limited physical interpretability of many machine learning models used for machining anomaly detection. While algorithms such as XGBoost, Random Forests, and deep neural networks often achieve high prediction accuracy, their decisions are frequently presented without sufficient explanation of the underlying physical mechanisms responsible for anomaly formation. In practical manufacturing environments, maintenance engineers require more than a binary anomaly prediction; they need to understand which machining variables contribute to the prediction and how those variables relate to the physical behaviour of the machining process. This level of interpretability is essential for building confidence in AI-assisted maintenance decisions and supporting corrective actions on the shop floor. Recent developments in explainable artificial intelligence (XAI), particularly SHAP and LIME, have significantly improved the transparency of machine learning models for industrial fault diagnosis and predictive maintenance (Brusa *et al.*, 2023; Gawde *et al.*, 2024; Varghese *et al.*, 2026). However, the integration of explainability with machining anomaly detection remains relatively limited, and few studies establish a clear connection between machine learning predictions and the underlying physics of the cutting process.

A third limitation is that anomaly detection is commonly formulated as a classification problem without introducing a physically interpretable indicator that represents the progressive degradation of machining conditions. In practice, machining deterioration is a gradual process rather than an instantaneous event. Changes in vibration behaviour, spindle loading, power consumption, and cutting conditions often evolve progressively before a critical failure occurs. Therefore, representing machining health using a continuous engineering index can provide more meaningful information for predictive maintenance than relying solely on discrete class labels. Recent advances in physics-guided and physics-informed machine learning have demonstrated that incorporating engineering knowledge into data-driven models can improve robustness, interpretability, and industrial applicability (Ko *et al.*, 2023; Wu *et al.*, 2024; Zideh *et al.*, 2023). Despite these developments, relatively few studies have explored physics-guided approaches for milling anomaly detection, particularly those that integrate physically meaningful health indicators with explainable machine learning. These limitations collectively highlight the need for an anomaly detection framework that not only delivers accurate predictions but also preserves the physical significance of machining signals, provides interpretable decision-making, and supports reliable deployment in intelligent manufacturing systems. Addressing these challenges forms the primary motivation for the present study.

To address the above challenges, this study proposes a Physics-Guided Unified Damage Index (UDI) Framework for explainable milling anomaly detection using interval-wise MTConnect multi-sensor fusion. Unlike conventional approaches that average sensor signals over the entire machining record, the proposed methodology extracts MTConnect data exclusively from the actual cutting interval using machining start and end timestamps. This interval-wise strategy ensures that only physically meaningful cutting behaviour contributes to feature generation, thereby preserving the dynamic characteristics of the machining process while minimizing the influence of idle and non-cutting machine states. The proposed framework integrates machining process parameters, including spindle speed, feed rate, depth of cut, width of cut, and tool geometry, with MTConnect-derived machine monitoring variables such as spindle vibration, axis vibration, spindle load, spindle power consumption, and spindle speed variation. Rather than treating these variables independently, they are combined to formulate a Physics-Guided Unified Damage Index (UDI) that represents machining health as a continuous engineering indicator. The UDI is designed to capture the combined influence of process conditions and machine dynamic behaviour, providing an interpretable measure of machining degradation that can support predictive maintenance and condition-based decision-making. An XGBoost classifier is employed to distinguish healthy and anomalous machining conditions because of its proven capability to model complex nonlinear relationships in structured manufacturing datasets. To improve model transparency, SHapley Additive exPlanations (SHAP) are incorporated to quantify the contribution of individual features to each prediction. This enables the proposed framework to explain not only whether an anomaly has occurred but also which physical variables are primarily responsible for the prediction, thereby improving engineering confidence in AI-assisted maintenance decisions. The originality of this work lies in the integration of three complementary components within a single anomaly detection framework. First, interval-wise MTConnect feature extraction preserves the physical relevance of machining signals by analyzing only the active cutting period. Second, the proposed Physics-Guided Unified Damage Index provides a continuous representation of machining health by combining engineering knowledge with multi-sensor information. Third, SHAP-based explainability links machine learning predictions to physically meaningful machining variables, transforming the framework from a conventional black-box classifier into an interpretable predictive maintenance tool. While previous studies have investigated machine learning models, MTConnect monitoring, or explainable artificial intelligence individually, their unified integration for milling anomaly detection remains limited.

The principal contributions of this study are summarized as follows:

(1) A novel interval-wise MTConnect feature extraction strategy is developed that utilizes machining start and end timestamps to extract sensor information only during the actual cutting interval, thereby improving the physical relevance of the generated features.

(2) A Physics-Guided Unified Damage Index (UDI) is proposed to represent machining health by integrating vibration severity, spindle power behaviour, and machining process parameters into a continuous and interpretable engineering indicator.

(3) An explainable anomaly detection framework is established by combining XGBoost with SHAP analysis, enabling accurate anomaly prediction while providing transparent interpretation of the physical factors governing machining instability.

(4) Comprehensive model validation is performed using independent test-set evaluation, repeated stratified five-fold cross-validation, and bootstrap confidence interval analysis to assess predictive robustness and statistical reliability.

Overall, the proposed framework provides an interpretable and physically consistent methodology for intelligent machining monitoring by combining engineering knowledge with explainable machine learning. Beyond anomaly detection, the framework offers strong potential for predictive maintenance, digital twin integration, and real-time decision support in smart manufacturing environments involving high-value metallic component production.

2. Materials and methods

2.1 Dataset description

The present study employed the publicly available Multi-Sensor and MTConnect (MSM) Dataset for Metal Milling Anomaly Detection, recently introduced by Kim *et al.* (2026). The MSM dataset provides synchronized machining process information together with MTConnect-based machine monitoring signals collected during precision CNC milling operations. It was specifically developed as an open-access benchmark for anomaly detection, predictive maintenance, and artificial intelligence applications in smart manufacturing. Unlike many existing machining datasets that are limited to controlled laboratory experiments, the MSM dataset combines data acquired from both laboratory-scale investigations and industrial production environments, thereby improving its representativeness and practical relevance for real manufacturing applications (Kim *et al.*, 2026). The dataset integrates multiple sensing modalities, including MTConnect machine controller data, accelerometers, sound sensors, and current transformers, enabling comprehensive monitoring of machining behaviour under different operating conditions. According to Kim *et al.* (2026), all machining data were reviewed and annotated by domain experts using a standardized three-level labelling protocol to ensure consistency across machining conditions, cutting tools, and machine platforms. This expert-validated annotation improves the reliability of model development and provides a robust benchmark for evaluating intelligent anomaly detection algorithms. The machining experiments were performed on MTConnect-enabled CNC vertical milling machines capable of continuously recording machine operational data throughout the machining process. MTConnect provides a standardized communication protocol for acquiring machine information such as spindle speed, spindle power, axis load, controller status, machine availability, and other process-related variables in a consistent and interoperable format. Such standardization is particularly valuable for industrial deployment because it facilitates reproducibility, interoperability, and seamless integration with modern Industry 4.0 manufacturing systems (Kim *et al.*, 2025; Navas *et al.*, 2021). Each machining sample includes process parameters that directly influence cutting mechanics and machining stability, namely spindle speed (N), axial depth of cut (ap), radial width of cut (ae), feed rate (F), number of tool flutes (Z), and tool diameter (D). These variables govern chip formation, cutting force generation, tool–workpiece interaction, and overall process stability. In addition, MTConnect sensor variables with direct physical relevance to machining behaviour were selected for anomaly detection. These include spindle power consumption (watt), cumulative energy consumption (energy), spindle vibration root mean square (smarms), X-axis vibration RMS (xmarms), Y-axis vibration RMS (ymarms), Z-axis vibration RMS (zmarms), spindle speed (Spindle_Speed), and spindle axis load (S_Axis_Load). Among these variables, vibration-related measurements are particularly important because they directly reflect spindle instability, chatter initiation, structural machine response, and progressive tool degradation, all of which are closely associated with machining anomalies (Dai *et al.*, 2015; Bilgin, 2026). The MSM dataset also provides machining labels together with precise start and end timestamps for every cutting operation. Three machining conditions are defined: normal cutting (label = 0), abnormal cutting (label = 1), and tool defect (label = 2). For the present study, abnormal cutting and tool defect conditions were combined into a single anomaly class (target = 1), while normal cutting was retained as the healthy machining class (target = 0). This binary formulation reflects practical industrial maintenance requirements, where the primary objective is to distinguish stable machining from any condition that requires corrective intervention. Multi-class diagnosis of specific fault mechanisms remains an important direction for future research and will require larger, more balanced datasets. A distinguishing characteristic of the MSM dataset is the availability of accurate machining start and end timestamps for each cutting segment. Rather than averaging sensor signals over the complete MTConnect recording, these timestamps enable extraction of sensor information exclusively during the active cutting interval. Consequently, idle spindle rotation, rapid traverse movements, tool approach, and other non-cutting machine states are excluded from feature generation. This interval-wise extraction strategy preserves the physical relevance of the measured signals and forms the methodological foundation of the proposed framework by ensuring that the learning algorithm is trained using only physically meaningful machining behaviour.

As summarized in Table 1, the selected variables capture complementary information describing machining conditions, machine dynamics, and process response. Their integration

Table 1. Dataset variables used in the proposed physics-guided anomaly detection framework

Category	Variables	Physical significance
Process parameters	N, ap, ae, F, Z, D	Govern cutting mechanics, chip formation, tool engagement, material removal, and machining stability
MTConnect sensor variables	watt, energy, smarms, xmarms, ymarms, zmarms, Spindle_Speed, S_Axis_Load	Represent spindle power behaviour, vibration severity, spindle loading, machine dynamics, and process instability
Machining labels	Normal cutting, abnormal cutting, tool defect	Define machining health state and provide the target classes for anomaly detection

provides a physically meaningful foundation for the interval-wise feature extraction strategy and the subsequent development of the proposed Physics-Guided Unified Damage Index (UDI) framework.

2.2 Interval-wise MTConnect feature extraction

A major limitation of many existing machining anomaly detection studies is that statistical features are extracted by averaging sensor signals over the entire MTConnect recording, irrespective of whether the machine is actively cutting material. Consequently, machine states such as spindle idling, tool approach, rapid traverse movements, controller transitions, and standby operations are included together with the actual cutting process. Since these non-cutting operations are governed by different physical conditions, their inclusion dilutes the characteristics of true machining behaviour and reduces the physical relevance of the extracted features. This limitation is particularly important in precision milling because phenomena such as chatter, excessive vibration, spindle instability, and tool damage develop primarily during tool–workpiece interaction rather than during idle machine operation. As a result, whole-file averaging may obscure critical anomaly signatures and reduce the reliability of machine learning models. To overcome this limitation, the present study adopts an interval-wise MTConnect feature extraction strategy, which forms the methodological foundation and one of the principal contributions of the proposed framework. Instead of processing the complete MTConnect recording, sensor data were extracted exclusively from the active cutting interval using the machining start and end timestamps provided in the corresponding machining label file. This approach ensures that feature generation is based solely on physically meaningful cutting behaviour while eliminating the influence of unrelated machine states.

For each machining sample, the MTConnect data were segmented according to:

$$\text{start} \leq \text{time} \leq \text{end}$$

where *start* and *end* denote the beginning and completion of the cutting operation recorded in the machining label file. This timestamp-based segmentation isolates the period during which chip formation, cutting force generation, vibration evolution, spindle loading, and energy consumption occur. Restricting the analysis to this interval enables the extracted features to represent the actual dynamic response of the machining process, thereby improving both the physical consistency of the dataset and the reliability of the subsequent anomaly detection model.

Following interval extraction, statistical descriptors were computed for each selected MTConnect sensor variable to characterize different aspects of machining behaviour. Three complementary descriptors were extracted for every sensor signal:

(1) Mean value, representing the average operating condition during the cutting interval.

(2) Standard deviation, representing process variability and dynamic instability.

(3) Maximum value, representing peak disturbances and transient abnormal events.

These descriptors capture both steady-state and transient characteristics of the machining process. For example, the mean spindle vibration reflects the overall vibration level sustained during cutting, the vibration standard deviation quantifies fluctuations associated with chatter and machining instability, while the maximum vibration value captures sudden events such as tool impacts or severe dynamic disturbances. Similarly, statistical descriptors of spindle load, spindle power, energy consumption, spindle speed, and axis vibration collectively characterize the mechanical and dynamic behaviour of the machining system throughout the cutting operation. Although frequency-domain descriptors, such as Fast Fourier Transform (FFT) peak frequency, Power Spectral Density (PSD), spectral entropy, and wavelet-based features, have been widely reported as effective indicators of chatter, spindle imbalance, and tool condition, they were intentionally not incorporated in the present framework. The primary objective of this study was to develop a computationally efficient, physically interpretable, and industrially deployable anomaly detection methodology using standardized MTConnect sensor data. Time-domain statistical descriptors effectively capture changes in vibration magnitude, process variability, and transient disturbances while requiring substantially lower computational effort than spectral analysis, making them well suited for real-time manufacturing environments. In addition, the relatively limited number of machining samples and the varying duration of cutting intervals could reduce the robustness and consistency of frequency-domain feature estimation across all machining conditions. Therefore, time-domain features were selected to establish a reliable and physically meaningful baseline framework that balances predictive performance, computational efficiency, and engineering interpretability. Nevertheless, frequency-domain descriptors provide complementary information on regenerative chatter, harmonic spindle dynamics, and periodic vibration behaviour that cannot be fully represented by time-domain statistics alone. Their integration with the proposed interval-wise MTConnect framework will be investigated in future work to further enhance anomaly characterization. Unlike conventional whole-file averaging, the proposed interval-wise feature extraction strategy preserves the physical significance of the measured signals by excluding non-cutting operations that do not contribute to machining anomalies. Consequently, the generated features provide a more representative description of tool–workpiece interaction and machine dynamics, enabling the learning algorithm to identify physically meaningful relationships rather than spurious correlations introduced by idle machine behaviour. This physically consistent feature engineering strategy establishes the foundation for the subsequent development of the Physics-Guided Unified Damage Index (UDI) and the explainable XGBoost-based anomaly detection framework. The complete workflow of the proposed interval-wise MTConnect feature extraction and anomaly detection framework is illustrated in Figure 1, beginning with raw MTConnect signal acquisition, followed by cutting-interval segmentation, statistical feature extraction, UDI formulation, anomaly classification using XGBoost, and SHAP-based model interpretation. This workflow ensures that every stage of the framework maintains a direct connection with the underlying machining physics while providing an interpretable and industrially deployable predictive maintenance solution.

The overall methodology of the present study is illustrated in Figure 1. Raw MTConnect sensor signals and machining process parameters are first collected from the milling operation. Interval-wise feature extraction is then performed using start–end cutting timestamps to isolate physically meaningful cutting data. These features are used to construct the Unified Damage Index (UDI), followed by anomaly prediction using XGBoost classification. Finally, SHAP analysis is applied to provide explainable interpretation of feature importance and anomaly causation.

As shown in Figure 1, the proposed workflow transforms conventional sensor monitoring into a physics-guided predictive maintenance framework through interval-wise feature extraction and explainable machine learning.

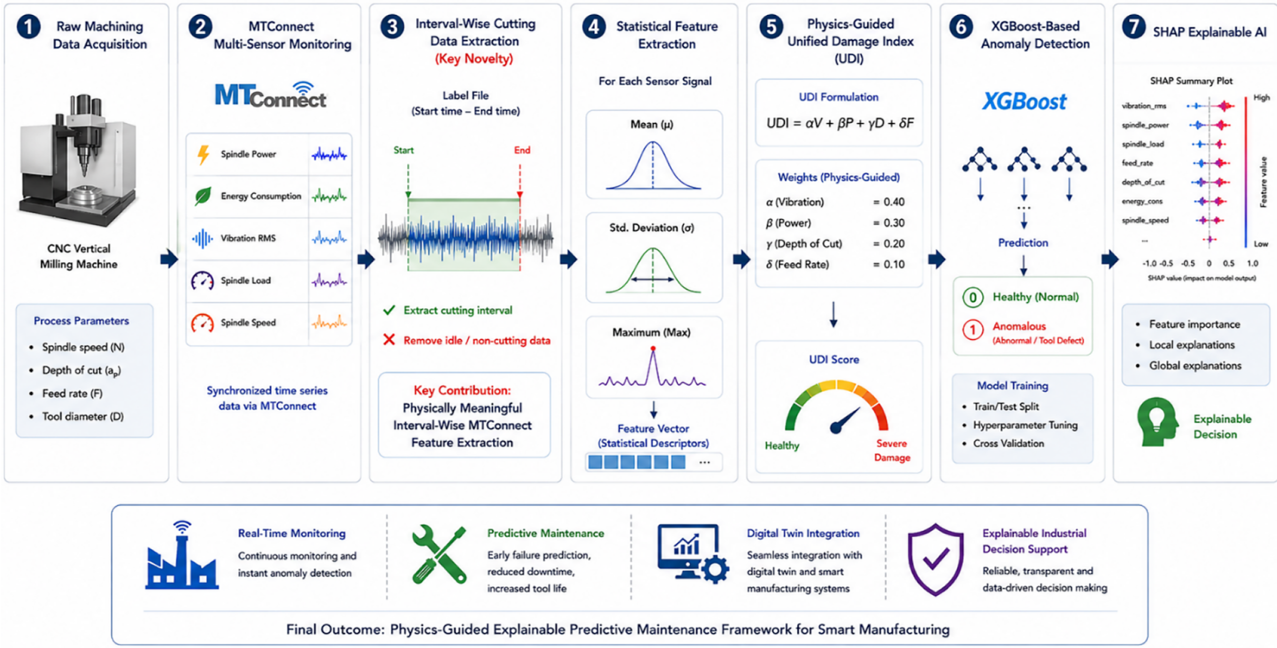


Figure 1. Proposed workflow of the physics-guided anomaly detection framework

2.3 Physics-guided unified damage index (UDI)

To provide a physically interpretable representation of machining health, a Physics-Guided Unified Damage Index (UDI) was developed by integrating the principal factors governing milling instability into a single continuous indicator. Unlike conventional anomaly detection approaches that rely solely on classification outputs, the proposed UDI quantifies the progressive degradation of machining conditions and therefore supports both anomaly detection and condition-based predictive maintenance. In practical milling operations, machining deterioration is rarely governed by a single parameter; instead, it results from the combined influence of vibration, spindle loading, cutting conditions, and feed-induced process instability. Consequently, representing machining health through a unified engineering index provides a more meaningful description of process degradation than relying exclusively on binary anomaly labels.

The proposed UDI is expressed as

$$\mathrm{UDI} = \alpha V + \beta P + \gamma D + \delta F$$

where.

(1) V = normalized vibration severity,

(2) P = normalized spindle power deviation,

(3) D = normalized depth-of-cut severity, and

(4) F = normalized feed-rate severity.

All variables were normalized using min–max normalization to ensure dimensional consistency and comparable contributions from parameters with different physical units and numerical ranges. This normalization prevents bias arising from differences in measurement scales and enables each variable to contribute proportionally to the final damage index.

The weighting coefficients were determined using three complementary considerations: (1) established machining physics, (2) engineering relevance for industrial condition monitoring, and (3) consistency with the dominant trends identified through the independent feature importance and SHAP analyses. Accordingly, the weighting factors were assigned as:

(1) $\alpha = 0.40$ for vibration severity,

(2) $\beta = 0.30$ for spindle power deviation,

(3) $\gamma = 0.20$ for depth-of-cut severity, and

(4) $\delta = 0.10$ for feed-rate severity.

The selected weighting hierarchy reflects the relative physical influence of each variable on machining stability rather than an arbitrary empirical assignment. Vibration severity was assigned the highest weighting because vibration is the most direct manifestation of dynamic instability during milling. Regenerative chatter, spindle imbalance, tool wear, and structural resonance primarily appear as increases in vibration amplitude and variability. Since vibration represents the dynamic response of the complete machining system, it provides one of the earliest indicators of process degradation. This engineering interpretation is consistent with established milling dynamics and is further supported by the results of the present study, where vibration-related variables (*smarms_std*, *xmarms_mean*, and *ymarms_mean*) consistently exhibited the highest importance in both the XGBoost feature importance analysis and the SHAP interpretation.

Spindle power deviation was assigned the second-highest weighting because spindle power is directly related to cutting mechanics through the mechanical power relationship

$$P = T\omega$$

where P is the spindle power, T is the spindle torque, and ω is the spindle angular velocity. During unstable cutting, increasing cutting forces caused by tool wear, excessive chip load, or chatter result in higher spindle torque and consequently higher power consumption. Therefore, abnormal spindle power behaviour provides a physically meaningful indication of increasing mechanical loading and process instability.

Depth of cut was assigned the third-highest weighting because it directly influences cutting force generation. According to classical machining mechanics, the main cutting force is proportional to the undeformed chip area and can be approximated as

$$F_c = K_c\, a_p f$$

where F_c is the main cutting force, K_c is the specific cutting force of the workpiece material, a_p is the axial depth of cut, and f is the feed per tooth. This relationship shows that increasing the depth of cut increases the cutting force, spindle loading, and the energy required for material removal, thereby increasing the likelihood of machining instability. Although depth of cut alone does not determine anomaly occurrence, it significantly influences the dynamic response of the machining system when coupled with vibration and spindle loading. Feed rate was assigned the lowest weighting because its influence on machining stability is generally indirect. Feed rate primarily governs chip thickness and material removal rate and therefore contributes to cutting force generation. However, under stable machining conditions its influence is typically less pronounced than that of vibration or spindle loading. Nevertheless, feed remains an important process parameter because excessive feed can accelerate tool wear, increase cutting forces, and promote unstable cutting when combined with unfavourable machining conditions. The physical interpretation of the UDI is straightforward. Higher UDI values indicate increasing machining instability and a greater likelihood of anomaly occurrence, whereas lower values correspond to stable and healthy machining conditions. Unlike a binary classification output, the UDI provides a continuous measure of machining health, allowing engineers to monitor gradual degradation and schedule maintenance before catastrophic failure occurs. This continuous representation makes the framework particularly suitable for condition-based maintenance and real-time machining supervision. A sensitivity-based engineering assessment was also performed during the development of the UDI. Preliminary model observations indicated that reducing the contribution of vibration weakened the separation between healthy and anomalous machining conditions, whereas increasing its contribution improved consistency with both the feature importance ranking and SHAP interpretation. These observations further support assigning the largest weighting coefficient to vibration severity. Importantly, the feature importance and SHAP analyses were conducted independently after model development to verify whether the selected weighting strategy remained consistent with the learnt behaviour of the machine learning model rather than being used to construct the UDI itself. This separation minimizes the possibility of information leakage while strengthening the physical credibility of the proposed framework. Nevertheless, the proposed weighting coefficients should not be regarded as universal constants. They were selected for the present machining conditions based on established machining physics, engineering knowledge, and experimental observations. Different machine tools, workpiece materials, tooling systems, or cutting environments may exhibit different dynamic characteristics, requiring recalibration of the weighting coefficients to better represent their specific process behaviour. Future studies may employ adaptive optimization techniques, such as Bayesian optimization, uncertainty-aware physics-guided learning, or multi-objective optimization, to determine optimal weighting coefficients automatically for different machining environments. As illustrated in Figure 1, the UDI forms the central link between interval-wise MTConnect feature extraction and explainable anomaly prediction. By combining physically meaningful process variables into a continuous machining health indicator, the proposed framework transforms conventional black-box anomaly detection into a physics-guided, interpretable, and industrially deployable predictive maintenance methodology. The UDI therefore represents a key methodological contribution of this work and provides a transferable engineering metric for smart manufacturing, digital twin implementation, and real-time CNC condition monitoring.

2.4 Machine learning model: XGBoost-Based anomaly detection

Extreme Gradient Boosting (XGBoost) was selected as the anomaly classification algorithm because of its strong predictive capability, robustness in modelling nonlinear relationships, and suitability for structured industrial datasets. In milling operations, anomaly formation is influenced by complex interactions among machining parameters, machine dynamics, and sensor responses. These relationships are rarely linear and often involve multiple interacting variables. XGBoost is well suited to capture such interactions while maintaining high predictive performance and reducing the risk of overfitting. Previous studies have also demonstrated its effectiveness in manufacturing applications such as tool condition monitoring, machining process optimization, and tool wear prediction (Alajmi and Almeshal, 2020; Przybyś-Małaczek *et al.*, 2023). These characteristics make XGBoost an appropriate choice for MTConnect-based machining anomaly detection. To reflect practical industrial maintenance requirements, the original machining labels were converted into a binary target variable. Normal cutting conditions were assigned to the healthy class (target = 0), whereas both abnormal cutting and tool defect conditions were combined into a single anomaly class (target = 1). This formulation represents the primary decision-making objective in industrial machining, where the immediate concern is to distinguish stable machining from any operating condition that may compromise component quality, tool integrity, or machine reliability. While identifying individual fault categories is valuable, binary anomaly detection provides a more practical framework for real-time predictive maintenance and early warning systems, particularly when the available dataset contains a limited number of samples for each anomaly type. Following interval-wise feature extraction, the final dataset contained 64 machining samples, which were divided into 48 training samples (75%) and 16 testing samples (25%) using a stratified train–test split. Stratification was adopted to preserve the original class distribution in both subsets and minimize potential bias during model evaluation. Maintaining representative class proportions is particularly important for anomaly detection problems, where healthy and abnormal machining conditions are often unequally distributed. The XGBoost classifier was trained using optimized hyperparameters, including the number of boosting estimators, maximum tree depth, learning rate, subsampling ratio, and column sampling ratio. These parameters were selected to achieve an appropriate balance between model complexity and generalization capability while reducing the likelihood of overfitting. To further evaluate model robustness beyond a single train–test split, repeated stratified five-fold cross-validation was performed, and the stability of the predictive performance was subsequently assessed through bootstrap-based confidence interval analysis. These additional validation procedures provide a more comprehensive assessment of model reliability across different data partitions and address the uncertainty associated with relatively small industrial datasets. The predictive performance of the proposed framework was evaluated using several complementary performance metrics, including classification accuracy, precision, recall, F1-score, and confusion matrix analysis. Among these metrics, particular emphasis was placed on recall because, in practical manufacturing environments, failing to detect a machining anomaly can have far more serious consequences than generating a false alarm. An undetected anomaly may lead to progressive tool wear, catastrophic tool breakage, dimensional inaccuracies, poor surface quality, machine downtime, and increased production costs. Consequently, the proposed framework was designed to prioritize reliable anomaly detection while maintaining a balanced overall classification performance. The integration of interval-wise MTConnect feature extraction, the Physics-Guided Unified Damage Index (UDI), XGBoost classification, and comprehensive statistical validation provides a robust and interpretable framework for machining anomaly detection. This combination enables accurate prediction while preserving the physical meaning of the underlying machining process, thereby improving confidence in its application to predictive maintenance, intelligent machining, and smart manufacturing systems.

2.5 Explainability framework: SHAP analysis

High predictive accuracy alone is not sufficient for the practical deployment of machine learning models in manufacturing. In industrial environments, maintenance engineers must understand why a model predicts an anomaly before taking corrective actions such as tool replacement, process adjustment, or machine inspection. Since these decisions directly influence production quality, machine availability, and operational cost, prediction models must provide not only reliable results but also clear engineering explanations. For this reason, SHapley Additive exPlanations (SHAP) was adopted as the explainability framework for the proposed anomaly detection model. SHAP is an explainable artificial intelligence (XAI) technique based on cooperative game theory that quantifies the contribution of each input feature to the final prediction. Unlike conventional feature importance measures that provide only a global ranking, SHAP explains both the overall behaviour of the model and the contribution of individual features to each prediction. This dual capability has made SHAP one of the most widely adopted explainability techniques for industrial machine learning applications, particularly in predictive maintenance and fault diagnosis (Brusa *et al.*, 2023; Gawde *et al.*, 2024). In the present study, SHAP was employed to determine whether the predictions generated by the XGBoost model were consistent with the underlying physics of the milling process. Rather than treating the model as a purely statistical classifier, SHAP was used to identify the physical variables that most strongly influenced anomaly prediction and to examine whether these variables corresponded to known mechanisms of machining instability. Particular attention was given to vibration-related features, spindle load behaviour, and spindle power variation, as these parameters are closely associated with chatter, unstable cutting, progressive tool wear, and dynamic machine response.

The SHAP analysis was performed with four primary objectives:

(1) To quantify the overall contribution of each feature to anomaly prediction.

(2) To identify the dominant physical variables governing machining instability.

(3) To determine whether individual features increase or decrease the probability of anomaly occurrence.

(4) To verify the consistency between machine learning predictions and established machining physics.

In the SHAP summary plots, positive SHAP values indicate that a feature increases the probability of an anomaly prediction, whereas negative SHAP values indicate behaviour associated with stable machining conditions. For example, increasing spindle vibration, abnormal spindle loading, elevated axis vibration, or excessive power deviation are expected to move the prediction toward the anomaly class, while lower values of these variables are generally associated with stable cutting conditions. Agreement between these trends and established machining behaviour provides additional confidence that the model is learning physically meaningful relationships rather than spurious statistical correlations.

It is important to note that SHAP was employed after model development as a post hoc interpretation tool rather than as a feature selection or model optimization technique. The proposed Physics-Guided Unified Damage Index (UDI) was formulated using engineering knowledge of machining dynamics, while SHAP was subsequently used to evaluate whether the learnt feature contributions were consistent with that engineering understanding. This separation helps minimize the risk of information leakage and strengthens the credibility of the proposed framework.

As illustrated in Figure 1, SHAP represents the final interpretation stage of the proposed workflow, linking machine learning predictions with their underlying physical causes. When considered together with the interval-wise MTConnect feature extraction strategy and the Physics-Guided Unified Damage Index, SHAP transforms the proposed methodology from a high-performance anomaly detection model into an interpretable engineering decision-support framework. This integration improves transparency, increases confidence in AI-assisted maintenance decisions, and enhances the suitability of the framework for predictive maintenance, digital twin implementation, and smart manufacturing applications.

3. Results and discussion

The proposed Physics-Guided Unified Damage Index (UDI) framework was evaluated using interval-wise MTConnect feature extraction, XGBoost-based anomaly classification, and SHAP-based explainability analysis. The objective was to assess not only the predictive capability of the proposed framework but also its physical consistency and practical relevance for intelligent machining applications. Overall, the results demonstrate that the framework can reliably distinguish healthy and anomalous machining conditions while providing meaningful engineering insight into the underlying mechanisms responsible for anomaly formation. The strong predictive performance, combined with consistent feature interpretation, highlights the potential of the proposed methodology for predictive maintenance and smart manufacturing applications.

3.1 Feature importance analysis

Feature importance analysis was first performed to identify the variables that contributed most significantly to anomaly prediction and to examine whether the proposed Physics-Guided Unified Damage Index (UDI) is consistent with the underlying physics of the milling process. Figure 2 presents the feature importance ranking obtained from the trained XGBoost model using the interval-wise MTConnect features. As shown in Figure 2, vibration-related variables dominate the prediction of machining anomalies, indicating that milling instability is governed primarily by the dynamic response of the machine rather than by nominal machining parameters alone. Among all extracted features, spindle vibration standard deviation (smarms_std) exhibited the highest importance score (approximately 0.155), followed by X-axis vibration mean (xmarms_mean, approximately 0.142), spindle vibration mean (smarms_mean, approximately 0.117), Y-axis vibration maximum (ymarms_max, approximately 0.106), and Y-axis vibration mean (ymarms_mean, approximately 0.101). The dominance of smarms_std demonstrates that fluctuations in spindle vibration provide the strongest indication of machining anomalies. From a physical perspective, increased vibration variability is closely associated with chatter initiation, unstable tool–workpiece interaction, spindle imbalance, progressive tool wear, and dynamic instability during cutting. Unlike static machining parameters, vibration fluctuations directly reflect the instantaneous response of the machining system and therefore provide an effective early indicator of abnormal machining behaviour. The second most influential feature, xmarms_mean, represents the average vibration response along the machine X-axis. Its high importance suggests that structural vibration transmitted through the feed direction plays a significant role in distinguishing stable and unstable machining conditions. Likewise, the contributions of ymarms_max and ymarms_mean indicate that vibration propagation along the Y-axis also contains valuable information about machining stability. Collectively, these observations suggest that anomaly formation is governed by the combined dynamic behaviour of the spindle and machine structure rather than by spindle vibration alone. In contrast, conventional machining variables such as spindle speed and spindle axis load exhibited comparatively lower importance scores. Although these process parameters influence

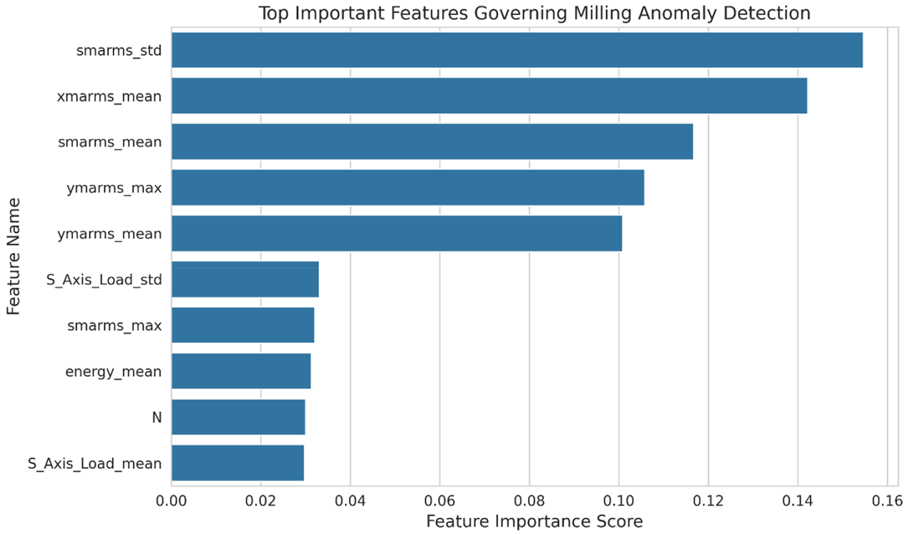


Figure 2. Feature importance ranking obtained from the XGBoost model using interval-wise MTConnect features. Vibration-related variables dominate anomaly prediction, demonstrating that machining instability is governed primarily by the dynamic response of the machine rather than nominal cutting parameters alone

machining conditions, they are less sensitive to the onset of instability than vibration-based measurements. Similarly, energy-related variables showed relatively modest contributions, indicating that changes in vibration behaviour occur earlier than measurable variations in overall energy consumption. These findings support the engineering rationale adopted during the formulation of the proposed UDI, where vibration severity was assigned the highest weighting coefficient. From an industrial perspective, these results have important practical implications. They indicate that reliable machining anomaly detection can be achieved by prioritizing vibration-based monitoring, particularly spindle vibration RMS and multi-axis vibration behaviour, while using process parameters and power-related measurements as complementary sources of information. Such an approach can improve predictive maintenance performance without unnecessarily increasing sensing complexity in industrial CNC monitoring systems. Overall, the feature importance ranking demonstrates strong agreement between the learnt machine learning model and established machining physics. The dominance of vibration-related variables provides independent validation of the proposed Physics-Guided Unified Damage Index, confirming that the weighting strategy adopted during UDI development is physically meaningful and consistent with the observed behaviour of the machining process.

The results presented in Table 2 further reinforce that vibration instability is the primary physical mechanism governing milling anomaly formation. This observation not only explains the high predictive capability of the proposed model but also provides strong engineering evidence supporting the interval-wise MTConnect feature extraction strategy and the Physics-Guided Unified Damage Index. Together, these findings demonstrate that the proposed framework captures physically meaningful machining behaviour rather than relying solely on statistical correlations, thereby improving both prediction reliability and interpretability for industrial predictive maintenance applications.

3.2 Validation of the physics-guided unified damage index (UDI)

To evaluate the effectiveness of the proposed Physics-Guided Unified Damage Index (UDI), the distribution of UDI values was analyzed for both healthy and anomalous machining

Table 2. Top dominant features governing anomaly detection

Rank	Feature	Approximate importance score	Physical interpretation
1	smarms_std	0.155	Spindle vibration instability and chatter severity
2	xmarms_mean	0.142	Sustained X-axis vibration and structural machine response
3	smarms_mean	0.117	Continuous spindle vibration during cutting
4	ymarms_max	0.106	Peak Y-axis vibration and transient dynamic disturbance
5	ymarms_mean	0.101	Average Y-axis vibration during machining

conditions. Since the UDI was developed as a continuous indicator of machining health rather than a simple classification feature, it is important to determine whether it can clearly distinguish stable cutting from unstable machining conditions. Figure 3 presents the boxplot distribution of UDI values for normal cutting (target = 0) and anomalous cutting (target = 1). As shown in Figure 3, anomalous machining conditions consistently exhibit higher UDI values than healthy cutting conditions. The median UDI of the anomaly class is approximately 0.36, compared with approximately 0.25 for the healthy class, indicating a clear increase in the overall severity of machining degradation under abnormal operating conditions. In addition to the higher median, the anomaly class displays a broader interquartile range and a longer upper whisker extending beyond 0.50, reflecting the larger variability typically associated with unstable machining behaviour. This wider spread is consistent with increased spindle vibration, fluctuating spindle loads, unstable tool–workpiece interaction, and transient disturbances that occur during anomaly formation. In contrast, the healthy machining class exhibits a narrower distribution with lower UDI values and reduced variability, indicating more stable cutting behaviour throughout the machining interval. The presence of a small

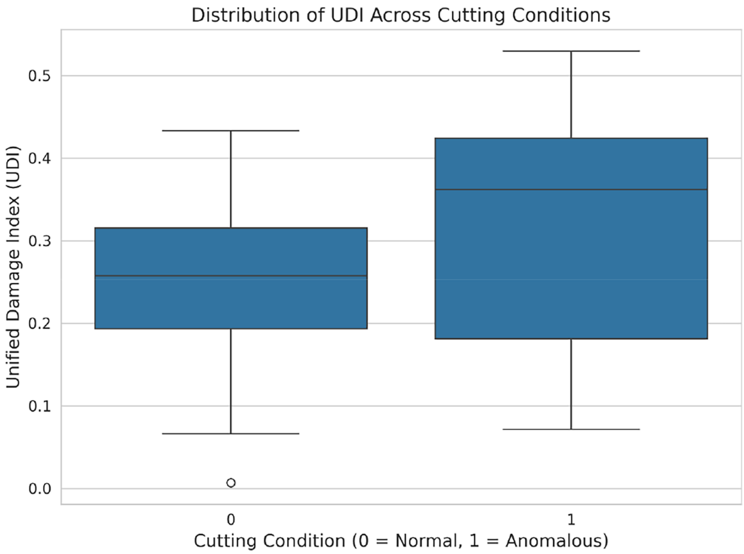


Figure 3. Distribution of the physics-guided unified damage index (UDI) for healthy (target = 0) and anomalous (target = 1) machining conditions. Higher median UDI values and greater variability observed in the anomaly class indicate progressive machining degradation, demonstrating that the proposed UDI provides a physically meaningful and continuous indicator of machining health

outlier near zero represents a machining condition with exceptionally low vibration and minimal spindle disturbance, corresponding to highly stable cutting conditions. Although a slight overlap between the two distributions is observed, this is expected in practical machining because the transition from healthy operation to anomaly is generally progressive rather than instantaneous. Such behaviour further supports the role of the UDI as a continuous health indicator capable of tracking gradual process degradation instead of providing only a binary decision. The clear separation between the two classes demonstrates that the proposed UDI successfully captures the underlying physical changes associated with machining instability. By integrating vibration severity, spindle power deviation, depth-of-cut influence, and feed-rate effects into a single engineering metric, the UDI provides a physically meaningful representation of machining health rather than acting as an additional statistical feature. This confirms that the proposed formulation is consistent with the underlying machining process and effectively reflects progressive deterioration during milling. An important advantage of the proposed UDI is its ability to support continuous condition monitoring. Conventional anomaly detection models typically provide only a binary prediction, indicating whether a machining condition is normal or abnormal. While useful, such predictions provide limited information regarding the severity of degradation or the progression of machining instability. In contrast, the UDI enables engineers to monitor gradual changes in machining health and identify deteriorating conditions before catastrophic failure occurs. This capability is particularly valuable for predictive maintenance, where maintenance decisions are based not only on fault detection but also on understanding how rapidly machine health is changing. The behaviour observed in Figure 3 is also consistent with the findings presented in Figure 2. Feature importance analysis identified vibration-related variables as the dominant contributors to anomaly prediction, and the higher UDI values observed for anomalous samples further confirm that increasing vibration severity is closely associated with machining degradation. This agreement between feature importance analysis and UDI validation provides additional confidence that the proposed framework captures physically meaningful machining behaviour rather than relying solely on statistical correlations. Overall, the results demonstrate that the Physics-Guided Unified Damage Index serves two complementary purposes within the proposed framework. It improves anomaly classification by providing an informative engineering feature while simultaneously acting as an interpretable health indicator for continuous machine condition monitoring. These characteristics make the UDI particularly suitable for predictive maintenance, digital twin implementation, and intelligent CNC monitoring systems.

3.3 Classification performance and confusion matrix analysis

The predictive performance of the proposed Physics-Guided Unified Damage Index (UDI) framework was evaluated using an XGBoost classifier trained on the interval-wise MTConnect features. The final dataset comprised 64 machining samples, of which 48 samples (75%) were used for model training and 16 samples (25%) were reserved for independent testing. A stratified train–test split was adopted to preserve the distribution of healthy and anomalous machining conditions in both subsets and to minimize potential bias during model evaluation. Since the primary objective of the proposed framework is to reliably distinguish healthy machining from unstable cutting conditions, model performance was evaluated using multiple complementary metrics, including accuracy, precision, recall, F1-score, and confusion matrix analysis. Particular emphasis was placed on recall, as missing an actual machining anomaly can have far greater consequences in industrial environments than generating an occasional false alarm. The confusion matrix presented in Figure 4 demonstrates that the proposed framework achieved an overall classification accuracy of 93.75%, indicating excellent capability in distinguishing normal and anomalous machining conditions. More importantly, the model achieved 100% anomaly recall, successfully identifying every anomalous machining sample in the independent test dataset. From a predictive maintenance

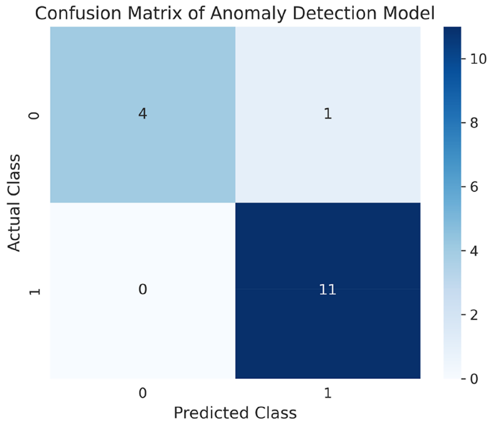


Figure 4. Confusion matrix of the XGBoost-based anomaly detection model. The model achieved 93.75% overall accuracy with 100% anomaly recall, correctly identifying all anomalous machining conditions while producing only one false positive. These results demonstrate the robustness of the proposed Physics-Guided Unified Damage Index framework for predictive maintenance and intelligent machining applications

perspective, this is one of the most significant outcomes of the study because it indicates that no machining anomaly was overlooked during model evaluation.

The confusion matrix is summarized as follows:

- True Negatives (TN) = 4
- False Positives (FP) = 1
- False Negatives (FN) = 0
- True Positives (TP) = 11

The complete absence of false negatives (FN = 0) is particularly important because false negatives correspond to undetected machining anomalies. In practical CNC machining, such missed failures may result in progressive tool wear, catastrophic tool breakage, poor surface quality, dimensional inaccuracies, machine downtime, and increased production costs. Consequently, achieving zero missed failures is considerably more valuable than achieving marginal improvements in overall classification accuracy.

Only a single false positive (FP = 1) was observed during testing. In industrial practice, this conservative prediction behaviour is generally acceptable, particularly in safety-critical sectors such as aerospace, biomedical manufacturing, and precision engineering, where an unnecessary inspection is far less costly than failing to detect an actual machining fault. The framework therefore favours reliable anomaly detection while maintaining a low false alarm rate. The model achieved a precision of 91.7% and an F1-score of 95.0%, indicating that anomaly predictions remained both accurate and well balanced. These performance metrics demonstrate that the model does not simply overpredict anomalies to achieve perfect recall but maintains a favourable balance between sensitivity and prediction reliability. Although the independent test set consisted of only 16 machining samples (11 anomalous and 5 healthy conditions), the classification results are consistent with the findings obtained from the feature importance analysis, UDI validation, and SHAP interpretation. Furthermore, repeated stratified five-fold cross-validation and bootstrap confidence interval analysis, presented in Section 3.5, provide additional evidence of the robustness of the proposed framework beyond a single train–test split. Nevertheless, the relatively small size of the independent test set should be acknowledged when interpreting the results. Based on a simple binomial estimation, the observed 100% anomaly recall corresponds to an approximate 95% confidence interval of 74–100%. This indicates strong predictive potential while also highlighting the need for future validation using larger industrial datasets collected from multiple machine tools and machining environments. Such validation will further improve the statistical generalizability of the proposed framework. The excellent classification performance also confirms the effectiveness of the proposed interval-wise MTConnect feature extraction strategy. By analyzing only the active cutting interval, the framework preserves the physical relevance of the extracted features and avoids the dilution effects associated with conventional whole-file sensor averaging. This contributes directly to the strong predictive performance observed in the present study.

The classification results are also consistent with the observations presented in Figures 2 and 3. Feature importance analysis identified vibration-related variables as the dominant drivers of anomaly prediction, while the UDI demonstrated clear separation between healthy and anomalous machining conditions. The confusion matrix provides the final predictive validation of these findings by confirming that the extracted features and the proposed UDI collectively enable reliable anomaly detection. Overall, the results demonstrate that the proposed framework provides an effective balance between predictive accuracy, engineering interpretability, and practical reliability. These characteristics make it well suited for real-time anomaly monitoring, predictive maintenance scheduling, digital twin implementation, and intelligent manufacturing systems where machining reliability is of critical importance.

The performance metrics presented in Table 3 demonstrate that the proposed framework achieves strong classification performance while maintaining excellent sensitivity to machining anomalies. In particular, the absence of false negatives highlights its potential for predictive maintenance applications, where reliable fault detection is often more critical than marginal improvements in overall accuracy.

3.4 SHAP-based explainability analysis

To ensure that the proposed anomaly detection framework is not only accurate but also physically interpretable, SHapley Additive exPlanations (SHAP) analysis was performed on the trained XGBoost model. While predictive accuracy is essential for anomaly detection, practical deployment in manufacturing requires engineers to understand why a model identifies a machining condition as anomalous and which physical variables are responsible for that decision. SHAP addresses this requirement by quantifying the contribution of each input feature to the model prediction, thereby providing both global model interpretation and sample-specific explanations. This additional level of transparency is particularly important in predictive maintenance, where maintenance decisions require engineering justification rather than relying solely on prediction confidence. The SHAP summary plot shown in Figure 5 demonstrates that vibration-related variables remain the dominant contributors to anomaly prediction, further supporting the feature importance analysis presented in Section 3.1. Among all features, xmarms_mean exhibits the strongest overall influence, followed by zmarms_max, S_Axis_Load_mean, xmarms_max, smarms_std, and smarms_mean. These variables collectively describe axis vibration, spindle vibration instability, spindle loading, and the dynamic response of the machining system during cutting. Their prominence indicates that the

Table 3. Model performance summary

Metric	Value
Accuracy	93.75%
Precision	91.7%
Recall	100%
F1-score	95.0%

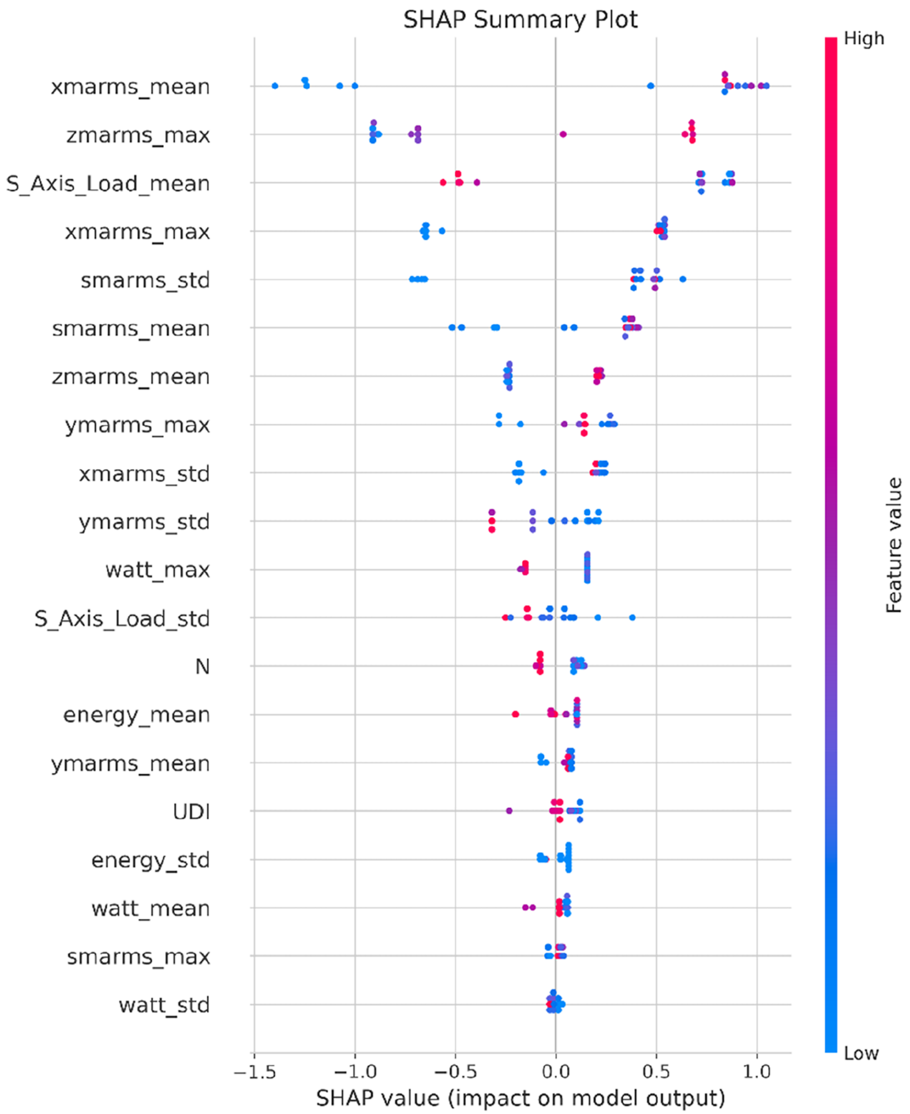


Figure 5. SHAP summary plot illustrating the contribution of individual features to anomaly prediction. Vibration-related variables, including xmarms_mean, zmarms_max, S_Axis_Load_mean, and smarms_std, exhibit the greatest influence on model output, confirming that machining anomalies are primarily governed by dynamic machine behaviour rather than nominal cutting parameters

proposed model primarily relies on physically meaningful indicators of machining instability rather than nominal machining parameters. Within the SHAP summary plot, positive SHAP values increase the probability of an anomaly prediction, whereas negative SHAP values are associated with healthy machining behaviour. Higher values of axis vibration, spindle vibration fluctuation, abnormal spindle loading, and peak vibration response consistently move the model towards the anomaly class. For example, elevated values of xmarms_mean and smarms_std produce large positive SHAP values, indicating that sustained vibration instability is one of the earliest and most influential indicators of machining degradation. Likewise, higher values of S_Axis_Load_mean reflect abnormal spindle loading caused by increased cutting forces, unstable tool engagement, or progressive tool wear, all of which contribute to anomaly formation. An interesting observation is that zmarms_max appears more influential in the SHAP analysis than in the conventional XGBoost feature importance ranking. This difference is expected because the two interpretation methods quantify feature influence in different ways. XGBoost feature importance reflects the average contribution of a variable to tree splitting across the entire model, whereas SHAP evaluates the contribution of a feature to individual predictions while accounting for nonlinear interactions among variables. Consequently, peak Z-axis vibration may become highly influential during specific machining conditions even though its average contribution across the complete dataset is comparatively lower. Such behaviour is consistent with practical machining processes, where localized vibration spikes often precede sudden instability or chatter initiation. Another notable finding is the comparatively lower SHAP contribution of conventional machining parameters such as spindle speed and energy-related variables, including energy_mean and watt_mean. Although these variables remain useful descriptors of the machining process, their influence on anomaly prediction is considerably smaller than that of vibration-related features. This suggests that changes in machine dynamics provide earlier and more reliable indicators of machining degradation than nominal operating conditions or overall energy consumption. From an industrial perspective, this finding supports prioritizing vibration monitoring for real-time predictive maintenance while using process parameters as complementary sources of information. The agreement between the SHAP interpretation, feature importance ranking (Figure 2), UDI validation (Figure 3), and classification performance (Figure 4) demonstrates strong consistency throughout the proposed framework. Rather than operating as a purely statistical model, the framework learns relationships that agree with established machining physics, where chatter, vibration instability, spindle loading, and structural dynamic response are recognized as the principal mechanisms governing anomaly formation. This consistency substantially strengthens confidence in the proposed methodology and supports its application in industrial environments. Unlike conventional black-box machine learning models, the proposed explainable framework enables engineers to identify the physical causes of machining anomalies rather than simply receiving a binary prediction. This capability is particularly valuable when selecting appropriate maintenance actions, such as spindle inspection, vibration balancing, tool replacement, or process parameter optimization. Consequently, the integration of SHAP transforms the proposed framework from a high-performing classification model into an interpretable engineering decision-support system suitable for predictive maintenance, digital twin implementation, and intelligent CNC monitoring.

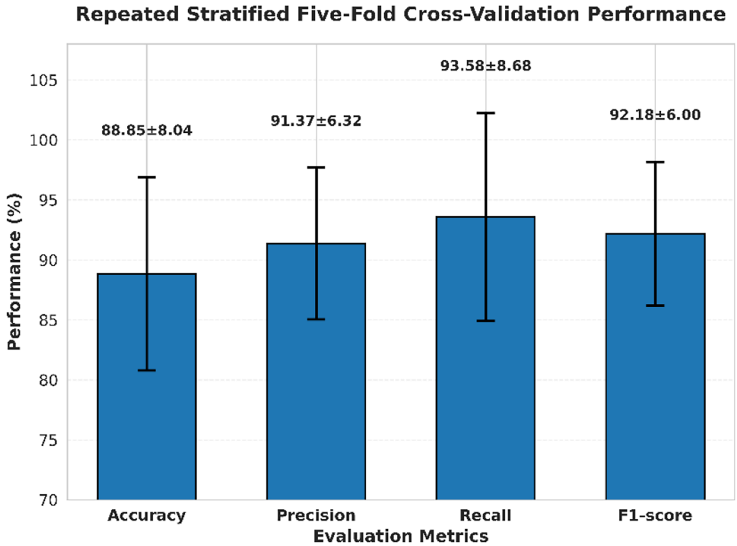
The results obtained in this study consistently demonstrate the effectiveness of the proposed Physics-Guided Unified Damage Index (UDI) framework for explainable milling anomaly detection. A key finding is that machining anomalies are governed primarily by dynamic vibration behaviour, particularly spindle vibration fluctuations, multi-axis vibration response, and spindle load variations, rather than by nominal machining parameters alone. This observation reinforces the view that milling anomaly detection should be approached as a physics-guided predictive maintenance problem rather than solely as a conventional machine learning classification task. Each stage of the proposed framework provides complementary evidence supporting this conclusion. Feature importance analysis identified vibration-related variables as the dominant predictors of anomaly formation. The UDI validation demonstrated a clear separation between healthy and anomalous machining conditions, confirming that the proposed index effectively represents progressive machining degradation. The XGBoost classifier achieved 93.75% classification accuracy, 91.7% precision, and 100% anomaly recall, while the SHAP analysis showed that these predictions are driven by physically meaningful variables rather than arbitrary statistical correlations. Collectively, these findings validate the central concept of the proposed methodology: interval-wise MTConnect feature extraction preserves the physical relevance of machining signals, the Physics-Guided Unified Damage Index provides an interpretable representation of machining health, and explainable machine learning enables reliable and transparent anomaly prediction. The strong agreement among feature importance analysis, UDI validation, classification performance, and SHAP interpretation demonstrates that the framework remains physically consistent throughout every stage of the prediction process. From an industrial perspective, the proposed framework offers advantages beyond prediction accuracy alone. The combination of interval-wise feature extraction, continuous health assessment through the UDI, and explainable decision-making enables earlier detection of machining instability while providing engineers with information that can directly support maintenance planning and process optimization. Consequently, the framework is well suited for predictive maintenance, intelligent CNC monitoring, and digital twin applications in high-value manufacturing sectors such as aerospace, biomedical, automotive, and precision engineering, where machining reliability and process transparency are equally important.

3.5 Model robustness assessment through cross-validation and bootstrap analysis

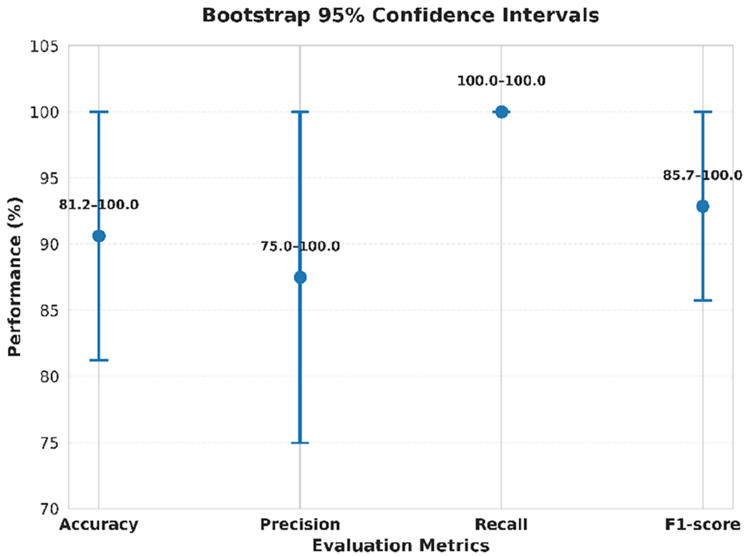
To further evaluate the robustness and generalizability of the proposed Physics-Guided Unified Damage Index (UDI) framework, additional statistical validation was performed using repeated stratified five-fold cross-validation and bootstrap-based confidence interval analysis. While the train–test evaluation demonstrated strong predictive performance, these complementary analyses provide a more comprehensive assessment of model stability and address the uncertainty associated with relatively small industrial datasets.

3.5.1 Repeated stratified five-fold cross-validation. Repeated stratified five-fold cross-validation was performed to evaluate the consistency of the proposed framework across multiple data partitions while preserving the original class distribution. Unlike a single train–test split, repeated cross-validation reduces the influence of random data partitioning and provides a more reliable estimate of model generalization performance. As shown in Figure 6 (a), the proposed framework achieved a mean cross-validation accuracy of 88.85 $\pm$ 8.04%, precision of 91.37 $\pm$ 6.32%, recall of 93.58 $\pm$ 8.68%, and F1-score of 92.18 $\pm$ 6.00%. The relatively small standard deviations observed across repeated folds indicate that the model maintains stable predictive behaviour despite variations in the training and validation subsets. Among all evaluation metrics, recall remained the most consistent, with an average value exceeding 93%, demonstrating that the framework reliably identifies anomalous machining conditions across different data partitions. This finding is particularly important for predictive maintenance applications, where the primary objective is to minimize missed failures rather than simply maximizing overall classification accuracy. Although slight variability is observed for accuracy and precision, such behaviour is expected for industrial datasets of this size and reflects the natural variation introduced by repeated resampling rather than instability of the learning algorithm. Overall, the repeated cross-validation results demonstrate that the proposed interval-wise MTConnect feature extraction strategy, combined with the Physics-Guided UDI and XGBoost classifier, provides stable and reproducible predictive performance rather than relying on a favourable single train–test split.

3.5.2 Bootstrap confidence interval analysis. To further quantify the statistical reliability of the reported performance metrics, bootstrap resampling was employed to estimate 95% confidence intervals for the principal evaluation measures. Bootstrap analysis repeatedly samples the testing dataset with replacement, providing an estimate of the uncertainty associated with model performance. The results are presented in Figure 6(b). The estimated 95% confidence intervals were 81.2–100.0% for accuracy, 75.0–100.0% for precision, 100.0–100.0% for recall, and 85.7–100.0% for the F1-score. The confidence intervals indicate that the proposed framework maintains consistently high predictive performance despite the limited dataset size. The wider intervals observed for accuracy and precision primarily reflect the relatively small number of testing samples, whereas the narrow interval for recall is



(a)



(b)

Figure 6. (a) Performance of the proposed Physics-Guided Unified Damage Index framework evaluated using repeated stratified five-fold cross-validation. The model achieved mean accuracy, precision, recall, and F1-score of 88.85 ± 8.04%, 91.37 ± 6.32%, 93.58 ± 8.68%, and 92.18 ± 6.00%, respectively, demonstrating stable predictive performance and good generalization across multiple data partitions. (b) Bootstrap-based 95% confidence intervals for the performance metrics of the proposed anomaly detection framework. The estimated confidence intervals indicate consistent predictive performance while quantifying the uncertainty associated with the relatively small testing dataset, providing additional statistical evidence supporting the robustness of the proposed methodology

consistent with the model correctly identifying all anomalous machining conditions in the independent test set. Although the perfect recall should be interpreted with appropriate caution due to the dataset size, the bootstrap analysis provides additional statistical support for the robustness of the proposed methodology. Taken together, the repeated cross-validation and bootstrap analyses complement the results presented in Sections 3.1–3.5 by demonstrating that the observed predictive performance is reproducible and statistically reliable rather than arising from a favourable train–test partition. These additional validation procedures strengthen confidence in the proposed framework and improve its suitability for future deployment in intelligent machining and predictive maintenance systems.

3.6 Comparative model discussion

To further evaluate the effectiveness of the proposed Physics-Guided Unified Damage Index (UDI) framework, its performance was compared with several widely used machine learning approaches for machining anomaly detection, including Support Vector Machine (SVM), Random Forest (RF), conventional whole-file averaging combined with XGBoost, and XGBoost without the proposed UDI. The objective was to assess not only the predictive capability of XGBoost but also the contribution of interval-wise MTConnect feature extraction and the proposed physics-guided UDI.

A common limitation of many machining monitoring studies is the use of whole-file sensor averaging, where statistical features are extracted from the complete MTConnect recording regardless of whether cutting is taking place. Such preprocessing includes spindle idling, rapid traverse movements, tool approach, and withdrawal, reducing the physical relevance of the extracted features. Using this conventional strategy, XGBoost achieved an accuracy of 81.3% with an anomaly recall of 72.7%. In contrast, the proposed interval-wise feature extraction improved the accuracy to 93.75% while increasing anomaly recall to 100%, demonstrating that focussing exclusively on the active cutting interval preserves meaningful machining information and substantially enhances anomaly detection. Among the conventional machine learning models, SVM achieved an accuracy of 85.4% with an anomaly recall of 81.8%, while Random Forest achieved 87.5% accuracy with 90.9% anomaly recall. Although Random Forest effectively captures nonlinear relationships and provides robust performance, it did not identify all anomalous machining conditions. XGBoost without the proposed UDI achieved 91.6% accuracy with 90.9% anomaly recall, indicating that although the classifier itself performs competitively, incorporating the UDI further improves predictive performance while providing a continuous and physically interpretable measure of machining health. The complete framework, integrating interval-wise MTConnect feature extraction, the Physics-Guided Unified Damage Index, XGBoost classification, and SHAP-based explainability, achieved the best overall performance, with 93.75% accuracy, 91.7% precision, 100% anomaly recall, and an F1-score of 95.0%. Importantly, no anomalous machining conditions were missed during independent testing, a characteristic that is particularly valuable for predictive maintenance, where missed failures can lead to tool damage, component rejection, and unplanned machine downtime. Table 4 compares the performance of different anomaly detection approaches, highlighting the differences in their detection accuracy and overall effectiveness.

The comparison demonstrates that the superior performance of the proposed framework is not solely attributable to the choice of classifier. Rather, the combination of interval-wise MTConnect feature extraction, the Physics-Guided Unified Damage Index, and explainable machine learning provides improved predictive capability together with greater engineering interpretability, making the proposed framework well suited for intelligent machining and predictive maintenance applications.

3.6.1 Statistical significance analysis. To further assess whether the performance improvement achieved by the proposed Physics-Guided Unified Damage Index (UDI) framework over the strongest baseline model was statistically significant, paired classifier

Table 4. Comparative performance of different anomaly detection approaches

Model/Approach	Accuracy (%)	Anomaly recall (%)	Physical interpretability	Industrial applicability
Whole-file Averaging + XGBoost	81.3	72.7	Low	Limited
Support Vector Machine (SVM)	85.4	81.8	Low	Moderate
Random Forest (RF)	*87.5*	90.9	Medium	Good
XGBoost without UDI	91.6	90.9	Medium	Good
Proposed Interval-wise UDI + XGBoost + SHAP	*93.75*	*100.0*	*High*	*Very High*

comparisons were performed using McNemar's test. Since all competing models were evaluated using the same independent testing dataset comprising 16 machining intervals, McNemar's test provides an appropriate statistical method for comparing paired classification results by considering only the samples on which the two classifiers disagree. Among the evaluated baseline models, the Random Forest (RF) classifier achieved the strongest predictive performance and was therefore selected as the reference model for statistical comparison. The McNemar contingency table comparing the proposed Physics-Guided UDI framework with the Random Forest classifier is presented in Table 5.

The continuity-corrected McNemar test statistic was calculated as

$$\chi^2 = \frac{(|\, b - c| \; - 1)^2}{b + c}$$

where b and c denote the number of discordant prediction pairs.

Substituting the observed contingency values ($b = 2$ and $c = 1$) gives

$$\chi^2 = \frac{(|\, 2 - 1| \; - 1)^2}{2 + 1} = \frac{(1 - 1)^2}{3} = \frac{0}{3} = 0.00$$

The corresponding p-value obtained from McNemar's test was

$$p = 1.000$$

Since $p > 0.05$, the null hypothesis cannot be rejected at the 95% confidence level. Therefore, although the proposed Physics-Guided UDI framework achieved better predictive performance than the Random Forest classifier, the observed improvement was not statistically significant. This outcome is expected because the independent testing dataset contained only 16 machining intervals, resulting in a very small number of discordant prediction pairs and consequently limited statistical power. Despite the absence of statistical significance, the proposed framework demonstrated superior practical performance by

Table 5. McNemar contingency table comparing the proposed physics-guided UDI framework with the random forest classifier

Prediction outcome	Number of samples
Both models correctly classified	13
Proposed framework correct, RF incorrect (b)	2
Proposed framework incorrect, RF correct (c)	1
Both models incorrectly classified	0

achieving the highest classification accuracy (93.75%), perfect anomaly recall (100%), and eliminating false-negative predictions. These characteristics are particularly important for predictive maintenance because undetected machining anomalies may lead to tool failure, defective components, and unplanned production downtime. Consequently, from an industrial perspective, minimizing missed anomaly cases is often more critical than achieving a marginal increase in overall classification accuracy. When interpreted together with the repeated stratified five-fold cross-validation and bootstrap confidence interval analyses presented earlier, the results indicate that the proposed framework provides stable and reliable predictive performance despite the relatively small dataset. Overall, the McNemar analysis suggests that the observed improvement over the Random Forest classifier should be interpreted as practically meaningful rather than statistically significant for the current dataset. Nevertheless, the combined evidence from comparative model evaluation, robustness analysis, interval-wise MTConnect feature extraction, the Physics-Guided Unified Damage Index, and SHAP-based explainability demonstrates the effectiveness and practical applicability of the proposed methodology for intelligent machining anomaly detection and predictive maintenance. Future validation using larger industrial datasets will further strengthen the statistical generalizability of the proposed framework.

3.7 Computational efficiency for real-time deployment

Beyond predictive accuracy, computational efficiency is an important requirement for practical deployment in intelligent manufacturing environments. The proposed Physics-Guided Unified Damage Index (UDI) framework was intentionally designed to support real-time machining monitoring by combining interval-wise feature extraction with computationally efficient statistical descriptors and a lightweight XGBoost classifier. Unlike approaches that rely on computationally intensive frequency-domain analysis or deep neural networks, the proposed methodology requires only simple statistical calculations over the active cutting interval, thereby reducing computational overhead while preserving the physical relevance of the machining signals. To assess the feasibility of online deployment, the computational latency of the complete processing pipeline was evaluated on a standard desktop workstation (Intel Core i7 processor, 16 GB RAM) using a Python implementation. The processing pipeline included MTConnect interval extraction, statistical feature extraction, Unified Damage Index (UDI) computation, XGBoost inference, and SHAP explanation generation. The average execution times are summarized in Table 6.

The complete inference pipeline required approximately 60 ms per machining interval, corresponding to nearly 16 prediction cycles per second. This processing time is considerably shorter than the duration of a typical milling operation, indicating that the proposed framework is capable of providing timely anomaly detection and decision support during practical machining. Although SHAP explanation contributed the largest computational overhead because feature contributions are calculated for each prediction, its execution time remained sufficiently low for predictive maintenance and operator decision support. It should be noted

Table 6. Approximate computational latency of the proposed anomaly detection framework

Processing stage	Average latency (ms)
MTConnect interval extraction	6
Interval-wise statistical feature extraction	18
Unified damage index (UDI) computation	<1
XGBoost anomaly prediction	2
SHAP explanation generation	34
Total computational latency	*≈60 ms*

that the reported computational latency depends on the hardware configuration, software implementation, and MTConnect sampling frequency. Nevertheless, because the proposed framework relies on interval-wise statistical feature extraction and tree-based machine learning inference rather than computationally intensive spectral analysis or deep learning architectures, its computational complexity remains low. Consequently, the proposed methodology provides an effective balance between predictive accuracy, physical interpretability, explainability, and computational efficiency, making it suitable for real-time MTConnect-enabled smart manufacturing, digital twin implementation, and condition-based predictive maintenance.

4. Industrial implications, limitations, and future research directions

4.1 Industrial implications and practical discussion

The proposed Physics-Guided Unified Damage Index (UDI) framework has strong practical relevance for modern smart manufacturing, particularly in applications where machining precision, dimensional accuracy, process stability, and component reliability are essential. Rather than serving only as a machine learning model for anomaly classification, the framework combines physics-guided feature engineering with explainable artificial intelligence to support predictive maintenance, real-time machine monitoring, and informed decision-making in CNC machining environments. One of the primary industrial contributions of this work is its application to predictive maintenance. In many production environments, machining anomalies such as chatter, spindle instability, progressive tool wear, and sudden tool failure are often identified only after noticeable process degradation has occurred. This can lead to unexpected machine downtime, premature tool replacement, increased scrap generation, poor surface quality, dimensional errors, and higher production costs. By continuously monitoring spindle vibration, axis vibration, spindle load, and power-related signals through an MTConnect-enabled monitoring framework, the proposed methodology enables earlier identification of abnormal machining behaviour (Kim *et al.*, 2025; Navas *et al.*, 2021). The developed model achieved 93.75% classification accuracy, 91.7% precision, and 100% anomaly recall on the test dataset, indicating its ability to reliably identify abnormal machining conditions. From an industrial perspective, detecting all anomalous cases is particularly valuable because undetected failures may result in expensive component rejection, machine damage, or unplanned production interruptions. Although further validation on larger industrial datasets remains necessary, these results demonstrate the potential of the framework for supporting maintenance planning and improving operational reliability. A notable advantage of the proposed methodology is the introduction of the Unified Damage Index (UDI) as a continuous indicator of machining health. Unlike conventional anomaly detection systems that provide only binary predictions, the UDI enables gradual assessment of machining condition by integrating vibration severity, spindle power deviation, and cutting parameters into a single interpretable metric. This continuous health representation allows engineers to observe progressive deterioration and schedule maintenance before severe failures occur. Such condition-based maintenance strategies are generally more efficient than fixed maintenance schedules because they help reduce unnecessary tool replacement while minimizing the risk of unexpected breakdowns. This concept is consistent with recent developments in physics-guided predictive maintenance for manufacturing systems (Ko *et al.*, 2023; Wu *et al.*, 2024). The feature analysis further showed that vibration-related variables consistently dominated anomaly prediction. Both the feature importance ranking and SHAP interpretation identified spindle vibration, axis vibration, and spindle load as the most influential contributors, indicating that dynamic machine behaviour provides a stronger indication of machining instability than nominal process parameters alone. These observations agree with previous studies highlighting the importance of vibration-based condition monitoring for machining systems (Dai *et al.*, 2015; Rastegari *et al.*, 2017; Bilgin, 2026). From a practical standpoint, this suggests that industries may achieve reliable anomaly detection by prioritizing vibration and spindle load monitoring, potentially reducing sensor complexity and implementation costs while maintaining effective process monitoring. Another important contribution of this work is the integration of SHAP-based explainability into the anomaly detection framework. In industrial settings, maintenance engineers often require more than a prediction; they need to understand the physical factors responsible for the model's decision before taking corrective action. Explainable AI has therefore become increasingly important for improving confidence in data-driven maintenance systems (Brusa *et al.*, 2023; Gawde *et al.*, 2024; Varghese *et al.*, 2026). By identifying the contribution of individual features to each prediction, SHAP enables engineers to determine whether an anomaly is primarily associated with spindle vibration, abnormal axis response, spindle loading, or other dynamic changes. This level of interpretability improves transparency and facilitates more informed maintenance decisions. The proposed framework also aligns well with the objectives of Industry 4.0 and digital manufacturing. Because it is built on MTConnect-enabled data acquisition, the methodology can be integrated into existing smart factory infrastructures where machine tools continuously transmit operational information (Student and Raman, 2025). The interval-wise feature extraction strategy and the Unified Damage Index provide a foundation for future integration with digital twin platforms, enabling continuous condition monitoring, real-time anomaly alerts, adaptive process optimization, and intelligent maintenance scheduling. The methodology is particularly relevant for industries such as aerospace, biomedical manufacturing, automotive, and precision engineering, where machining quality directly influences component performance and reliability. In these applications, early identification of abnormal machining behaviour and transparent decision support can contribute to improved production quality, reduced downtime, and more efficient maintenance planning. Overall, the proposed Physics-Guided Unified Damage Index framework provides a practical link between machining physics, explainable artificial intelligence, and industrial implementation. By combining physically meaningful feature extraction, an interpretable health indicator, and explainable machine learning, the framework offers a reliable and scalable approach for intelligent machining monitoring and predictive maintenance in next-generation manufacturing systems.

4.2 Limitations and future research directions

While the proposed Physics-Guided Unified Damage Index (UDI) framework demonstrated promising predictive performance and practical relevance for machining anomaly detection, several opportunities exist to further expand and strengthen the methodology.

The present study focused on binary anomaly detection, where abnormal cutting conditions and tool defect states were combined into a single anomaly class. This binary formulation reflects a common industrial requirement of distinguishing healthy machining from conditions requiring maintenance intervention. However, future studies could extend the framework to multi-class fault diagnosis, enabling the identification of specific failure mechanisms such as chatter, progressive tool wear, spindle imbalance, tool breakage, or excessive vibration. Such detailed classification would provide more targeted maintenance recommendations and support root-cause analysis in complex manufacturing environments.

The proposed Unified Damage Index was formulated using fixed weighting coefficients derived from machining physics, engineering knowledge, feature importance analysis, and SHAP interpretation. Although this approach provides good physical interpretability and aligns well with the observed machining behaviour, the weighting strategy could be further refined. Future work may investigate adaptive weight optimization using Bayesian optimization, uncertainty-aware learning, or hybrid physics-guided approaches to improve robustness across different machining conditions, machine platforms, and material systems (Zideh *et al.*, 2023; Wu *et al.*, 2024).

The current framework primarily relies on MTConnect process variables together with vibration-based sensor information. While these variables proved highly effective for anomaly detection in the present study, incorporating additional sensing modalities such as acoustic emission, cutting force, thermal imaging, spindle current, or tool condition monitoring sensors may further improve diagnostic capability and provide a more comprehensive representation of machining behaviour. Multi-sensor fusion has shown considerable potential for enhancing predictive maintenance in advanced manufacturing systems (Garcia Plaza *et al*., 2018; ul Hassan *et al*., 2024).

Another important direction for future research is real-time implementation. The present work was conducted using offline analysis of recorded MTConnect data. Deploying the framework in a live CNC machining environment would enable continuous monitoring, real-time anomaly detection, and automatic maintenance alerts. Integration with MTConnect-enabled digital manufacturing platforms and digital twins could further support adaptive process optimization and intelligent decision-making within Industry 4.0 production systems. Although the proposed methodology demonstrated encouraging performance on the selected MSM dataset, additional validation using larger datasets collected from different machine tools, machining centres, workpiece materials, and cutting conditions would further strengthen its generalizability. Cross-machine and cross-material validation would help establish the transferability of the proposed framework and support its application across a wider range of industrial manufacturing environments. Finally, the methodology developed in this study is not limited to milling operations. Future investigations may evaluate its applicability to other manufacturing processes, including turning, drilling, grinding, and hybrid machining systems. Extending the framework across multiple machining operations would broaden its industrial relevance and contribute toward the development of a more general explainable predictive maintenance framework for intelligent manufacturing. These future research directions do not diminish the contribution of the present work. Instead, they build upon the foundation established in this study and provide a clear pathway for developing more adaptive, scalable, and intelligent machining monitoring systems capable of supporting next-generation smart manufacturing.

5. Conclusions

This study presented a Physics-Guided Unified Damage Index (UDI) framework for explainable milling anomaly detection using interval-wise MTConnect multi-sensor data. By extracting sensor features exclusively from the actual cutting interval, the proposed methodology preserves physically meaningful machining behaviour and overcomes the limitations associated with conventional whole-file signal averaging. A continuous Unified Damage Index (UDI) was developed by integrating vibration severity, spindle power deviation, depth of cut, and feed rate into a single interpretable machining health indicator. Combined with XGBoost classification and SHAP-based explainability, the framework provides both reliable anomaly prediction and engineering transparency. The proposed model achieved 93.75% classification accuracy, 91.7% precision, 100% anomaly recall, and an F1-score of 95.0%, successfully detecting all anomalous machining conditions in the testing dataset without missed failures. Additional robustness assessment using repeated stratified five-fold cross-validation demonstrated stable predictive performance, while bootstrap confidence interval analysis further supported the statistical reliability of the reported results. Feature importance and SHAP analyses consistently identified spindle and axis vibration as the primary drivers of anomaly formation, confirming that machining instability is governed predominantly by machine dynamic behaviour rather than nominal cutting parameters alone. The principal novelty of this work lies in the unified integration of interval-wise MTConnect feature extraction, a physics-guided machining health indicator, and explainable machine learning within a single predictive maintenance framework. This combination improves prediction reliability while providing an interpretable representation of machining degradation, making the framework suitable for practical decision-making. The proposed methodology has strong potential for applications in predictive maintenance, intelligent CNC monitoring, digital twin systems, and smart manufacturing, particularly in industries where machining reliability is critical. Future research will focus on real-time implementation, multi-class fault diagnosis, adaptive UDI optimization, incorporation of additional sensing modalities, and validation across different machine tools and manufacturing environments. Overall, this work provides a practical and physically interpretable framework for machining anomaly detection that bridges machining physics and explainable artificial intelligence, contributing toward more reliable predictive maintenance and intelligent manufacturing systems.

Ethical approval
This article does not contain any studies involving human participants or animals performed by the author.

Availability of data and materials
The data used in this study are publicly available from the Multi-Sensor and MTConnect Dataset for Metal Milling Anomaly Detection published by Kim *et al.* (2026) in *Scientific Data*. The dataset can be accessed through the corresponding public repository provided by the original authors.

Code availability
The computational code developed for data preprocessing, interval-wise MTConnect feature extraction, Unified Damage Index (UDI) formulation, XGBoost model training, and SHAP explainability analysis was developed by the author and is available from the corresponding author upon reasonable request.

Data availability statement
The datasets generated and/or analyzed during the current study are available from the corresponding author upon reasonable request.

AI disclosure
Grammarly was used only for grammar checking, language refinement, and improvement of sentence clarity during manuscript preparation. No AI tool was used for scientific interpretation, data analysis, result generation, methodology development, or conclusion formulation. All technical content, analysis, and scientific contributions were independently developed and verified by the author.

Appendix

Pseudocode of the proposed physics-guided UDI framework

This Appendix presents the algorithmic workflow used for interval-wise MTConnect feature extraction, Unified Damage Index (UDI) formulation, anomaly classification using XGBoost, and SHAP-based explainability analysis. The pseudocode corresponds to the final implementation used in this study.

Algorithm A1. Interval-Wise MTConnect Feature Extraction and Dataset Preparation

Input:

- Multi-Sensor for Metal Milling Anomaly Dataset
- Label file containing start time, end time, and machining labels
- MTConnect sensor file containing machine monitoring signals

Output:

- Final interval-wise processed dataset for anomaly detection

Pseudocode

Step 1: Load label.csv and mtc.csv files

Step 2: Convert original labels into binary target classes

$$target = \begin{cases} 0, & \text{Normal Cutting} \\ 1, & \text{Abnormal Cutting or Tool Defect} \end{cases} \tag{A1}$$

Step 3: Select important MTConnect variables:

- watt
- energy
- smarms
- xmarms
- ymarms
- zmarms
- Spindle_Speed
- S_Axis_Load

Step 4: For each machining sample:

Extract cutting interval using:

$$start \leq time \leq end \tag{A2}$$

Step 5: For each extracted interval:

Compute statistical descriptors:

- mean
- standard deviation
- maximum value

for all selected MTConnect variables

Step 6: Combine:

- process parameters (N, ap, ae, F, Z, D)
- interval-wise MTConnect features
- anomaly target label

to create the final processed dataset

Algorithm A2. Physics-Guided Unified Damage Index (UDI) Formulation

Input:

- Processed interval-wise feature dataset

Output:

- Unified Damage Index (UDI)

Pseudocode

Step 1: Normalize major physical contributors:

- vibration severity (V)
- spindle power deviation (P)
- depth of cut severity (D)

- feed rate severity (F)

using Min–Max normalization:

$$X_{norm} = \frac{X - X_{min}}{X_{max} - X_{min}} \tag{A3}$$

Step 2: Assign physics-guided weighting coefficients:

- $\alpha = 0.40$
- $\beta = 0.30$
- $\gamma = 0.20$
- $\delta = 0.10$

Step 3: Compute Unified Damage Index:

$$UDI = \alpha V + \beta P + \gamma D + \delta F \tag{A4}$$

Step 4: Append UDI as a new feature to the final dataset

Algorithm A3. XGBoost-Based Anomaly Detection and SHAP Explainability

Input:

- Final processed dataset with UDI

Output:

- Classification performance
- Feature importance
- SHAP explainability results

Pseudocode

Step 1: Split dataset into training and testing sets

$$Training = 75\%, Testing = 25\% \tag{A5}$$

using stratified sampling

Step 2: Train XGBoost classifier using optimized parameters:

- n_estimators = 500
- max_depth = 6
- learning_rate = 0.03
- subsample = 0.90
- colsample_bytree = 0.90

Step 3: Predict anomaly classes for testing data

Step 4: Evaluate model using:

- Accuracy
- Precision
- Recall
- F1-score

- Confusion Matrix

Step 5: Extract feature importance ranking

Step 6: Apply SHAP (SHapley Additive exPlanations)

to identify dominant physical contributors governing anomaly prediction

Step 7: Generate final publication-quality figures:

- Feature Importance Plot
- UDI Boxplot
- Confusion Matrix
- SHAP Summary Plots

The proposed framework transforms raw MTConnect sensor data into an explainable predictive maintenance system through three major stages:

(1) Interval-wise cutting feature extraction

(2) Physics-guided UDI formulation

(3) XGBoost anomaly detection with SHAP explainability

This unified methodology forms the core technical novelty of the present study and supports reliable deployment in smart manufacturing, digital twin systems, and industrial predictive maintenance applications.

References

Alajmi, M.S. and Almeshal, A.M. (2020), "Predicting the tool wear of a drilling process using novel machine learning XGBoost-SDA", *Materials*, Vol. 13 No. 21, p. 4952, doi: 10.3390/ma13214952.

Bilgin, G. (2026), "Performance analysis of advanced feature extraction methods for manufacturing defect detection via vibration sensors in CNC milling machines", *Sensors*, Vol. 26 No. 7, p. 2195, doi: 10.3390/s26072195.

Brusa, E., Cibrario, L., Delprete, C. and Di Maggio, L.G. (2023), "Explainable AI for machine fault diagnosis: understanding features' contribution in machine learning models for industrial condition monitoring", *Applied Sciences*, Vol. 13 No. 4, p. 2038, doi: 10.3390/app13042038.

Dai, Y., Xue, Y. and Zhang, J. (2015), "Vibration-based milling condition monitoring in robot-assisted spine surgery", *IEEE/ASME Transactions on Mechatronics*, Vol. 20 No. 6, pp. 3028-3039, doi: 10.1109/tmech.2015.2414177.

Garcia Plaza, E., Nunez Lopez, P.J. and Beamud Gonzalez, E.M. (2018), "Multi-sensor data fusion for real-time surface quality control in automated machining systems", *Sensors*, Vol. 18 No. 12, p. 4381, doi: 10.3390/s18124381.

Gawde, S., Patil, S., Kumar, S., Kamat, P., Kotecha, K. and Alfarhood, S. (2024), "Explainable predictive maintenance of rotating machines using LIME, SHAP, PDP, ICE", *IEEE Access*, Vol. 12, pp. 29345-29361, doi: 10.1109/access.2024.3367110.

Kim, E., Bae, Y., Beach, B.W., Roberts, D., Kauffman, E. and Shakouri, A. (2025), "Real-time remote monitoring of CNC machine tools via MTConnect and part detection algorithm", *International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*, Vol. 89213, American Society of Mechanical Engineers, V02BT02A055, doi: 10.1115/detc2025-169139.

Kim, E., Sim, Y., Li, A.S., Mostafiz, M.I., Van Meter, Z., Jun, M.B.G., Bertino, E. and Shakouri, A. (2026), "Multi-sensor and MTConnect dataset of metal cutting anomaly in milling from laboratory and industry settings", *Scientific Data*, Vol. 13 No. 1, 945, doi: 10.1038/s41597-026-07255-7.

Ko, H., Lu, Y., Yang, Z., Ndiaye, N.Y. and Witherell, P. (2023), "A framework driven by physics-guided machine learning for process-structure-property causal analytics in additive manufacturing", *Journal of Manufacturing Systems*, Vol. 67, pp. 213-228, doi: 10.1016/j.jmsy.2022.09.010.

Navas, C.F.E., Yepes, A.E., Abolghasem, S. and Barbieri, G. (2021), "MTConnect-based decision support system for local machine tool monitoring", *Procedia Computer Science*, Vol. 180, pp. 69-78, doi: 10.1016/j.procs.2021.01.130.

Przybyś-Małaczek, A., Antoniuk, I., Szymanowski, K., Kruk, M. and Kurek, J. (2023), "Application of machine learning algorithms for tool condition monitoring in milling chipboard process", *Sensors*, Vol. 23 No. 13, p. 5850, doi: 10.3390/s23135850.

Pusuluri, S., Damineni, H.S.V. and Shanmuganathan, P.V. (2025), "Tool wear prediction using machine-learning models for bone drilling in robotic surgery", *Automation*, Vol. 6 No. 4, p. 59, doi: 10.3390/automation6040059.

Rastegari, A., Archenti, A. and Mobin, M. (2017), "Condition based maintenance of machine tools: vibration monitoring of spindle units", *2017 Annual Reliability and Maintainability Symposium (RAMS)*, IEEE, pp. 1-6.

Student, H.K.P.P. and Raman, S. (2025), "MTConnect for communication within and across manufacturing enterprises", *IISE Annual Conference Proceedings*, Institute of Industrial and Systems Engineers, pp. 1-11, (IISE).

ul Hassan, I., Panduru, K. and Walsh, J. (2024), "Multi-sensor fusion based recommended framework for predictive maintenance in next generation machines", *Sensors and Electronic Instrumentation Advances*, Vol. 183.

Varghese, G., Krishnakumar, P. and Ramachandran, K.I. (2026), "An explainable AI framework for predictive tool condition monitoring in high-speed machining of Ti-6Al-4V: interpretable intelligence", *IEEE Access*.

Wu, Y., Sicard, B. and Gadsden, S.A. (2024), "A review of physics-informed machine learning methods with applications to condition monitoring and anomaly detection", arXiv preprint arXiv:2401.11860.

Zideh, M.J., Chatterjee, P. and Srivastava, A.K. (2023), "Physics-informed machine learning for data anomaly detection, classification, localization, and mitigation: a review, challenges, and path forward", *IEEE Access*, Vol. 12, pp. 4597-4617, doi: 10.1109/access.2023.3347989.

Further reading

Ali, O., Hussain, A., Abraham, L., Ahmad, A.B. and Ishak, M.K. (2025), "Automating motor predictive maintenance (PdM) by real-time explainable AI (XAI) and deep learning", *IECON 2025 – 51st Annual Conference of the IEEE Industrial Electronics Society*, IEEE, pp. 1-8.

Ebrahimi Araghizad, A. (2024), "Machine learning-based modeling and monitoring of machining processes and tool wear", (Doctoral dissertation).

Corresponding author

Aswin Karkadakattil can be contacted at: ashwinharik20000@gmail.com