Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

# Physics-informed gradient boosting regression tree for tool wear prediction in milling

**Zhiwen Huang**

hzwedu@163.com

University of Shanghai for Science and Technology https://orcid.org/0000-0001-7325-3920

**Junjie Chen**

University of Shanghai for Science and Technology

**Jiajie Shao**

Tongji University

**Miaoxian Guo**

University of Shanghai for Science and Technology

**Xiaohui Jiang**

University of Shanghai for Science and Technology

**Jianmin Zhu**

University of Shanghai for Science and Technology

**Research Article**

**Keywords:** Tool wear monitoring, Physical model, Gradient boosting regression tree, Physical-data fusion, Milling

**Posted Date:** November 9th, 2026

**DOI:** https://doi.org/10.21203/rs.3.rs-10765267/v1

**License:** This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

# Physics-informed gradient boosting regression tree for tool wear prediction in milling

Zhiwen Huang [a, ✉], Junjie Chen[a], Jiajie Shao[b], Miaoxian Guo[a], Xiaohui Jiang[a], Jianmin Zhu[a]

[a] School of Mechanical Engineering, University of Shanghai for Science and Technology, Shanghai, 200093, China.

[b] School of Mechanical Engineering, Tongji University, Shanghai, 200092, China

✉ Corresponding author. E-mail address: hzwedu@163.com

## Abstract

Accurate prediction of tool flank wear is critical for maintaining machining quality, minimizing unplanned downtime, and enabling condition-based tool replacement. Data-driven models may violate physical wear laws under limited or varying data, whereas mechanistic models often lack the flexibility to capture complex, nonlinear wear evolution. To bridge this gap, this study proposes a physics-informed gradient boosting regression tree (PIGBRT) approach that integrates the generalized tool wear degradation law into the GBRT learning framework. In this method, the physical degradation law is formulated as a regularization term and integrated with the conventional data loss to construct a composite objective function, thereby jointly enforcing data fidelity and physical consistency during model training. The regression trees are then iteratively constructed along the negative gradient of the composite loss, enabling the physical prior to continuously constrain the learning trajectory rather than being imposed only during post-processing. Experimental validation on milling datasets demonstrates that the proposed PIGBRT achieves accurate tool wear prediction, with superior generalization and physical consistency compared with purely data-driven, mechanistic, and other advanced methods. These results highlight the effectiveness of embedding explicit physical knowledge into ensemble learning and establish PIGBRT as a promising framework for physics-consistent and generalizable tool monitoring in milling processes.

## Keywords

Tool wear monitoring, Physical model, Gradient boosting regression tree, Physical-data fusion, Milling

# 1 Introduction

Tool wear is a critical factor that directly influences machining quality, productivity, and manufacturing cost in milling operations. Accurate monitoring of tool flank wear enables timely tool replacement, reduces unexpected downtime, and ensures consistent part quality [1]-[3]. However, the process of tool wear is inherently complex, nonlinear, and time-varying, governed by multiple interrelated factors such as cutting parameters, work-piece material, etc [4]-[6]. These complexities render precise modeling of tool wear a substantial challenge.

Recently, numerous studies have aimed to improve the effectiveness and accuracy of tool wear monitoring, leading to a variety of prediction methods. These approaches generally fall into three categories: mechanistic models, data-driven models, and physics-data fused models [7][8]. Mechanistic models grounded in physical laws such as the Taylor tool life equation, offer direct insights but often rely on oversimplified assumptions, limiting their generalizability across varying operating conditions [9][10]. In contrast, data-driven methods achieve strong predictive performance by learning complex patterns from sensor data; however, they typically require large datasets and may produce unreal predictions, particularly in data-scarce scenarios [11][12]. Physics-data fused methods address these issues by embedding domain knowledge, such as the generalized tool wear degradation law, into data-driven learning frameworks, balancing learning flexibility with physical regularization [13][14].

Inspired by the physics-data integration philosophy, a physics-informed gradient boosting regression tree (PIGBRT) is proposed to improve the accuracy and generalization of tool wear prediction under multiple operating conditions and small-sample scenarios. GBRT excels at handling heterogeneous, non-stationary sensor data, capturing feature interactions, and tolerating outliers. However, as a purely data-driven model, it may produce physically inconsistent predictions, particularly when labeled data are scarce or imbalanced. Therefore, the proposed PIGBRT integrates the generalized tool wear degradation law as a regularization term into the GBRT loss alongside the conventional data prediction error. The composite objective, balanced by a weighting factor, is optimized to minimize both data and physical losses, yielding predictions that are accurate and physically consistent. This integration preserves the flexibility and strong generalization of GBRT while embedding physical constraints, effectively enhancing the prediction performance of too wear.

The remainder of this article is structured as follows. The related works of tool wear monitoring are reviewed in Section2. In Section 3, the overall framework and concrete implementation steps of the proposed methodology are described in detail. Section 4 introduces the experiment settings and data processing. The comparative results and model performances are analyzed and discussed in Section 5 comprehensively. The main conclusions and future work are drawn in Section 6.

## 2 Related works

Tool wear is an inevitable phenomenon in milling processes, directly affecting surface integrity, dimensional accuracy, and machining efficiency. Gradually intensifying over cutting time, this wear leads to excessive cutting forces, elevated temperatures, and eventual tool failure. Accurate prediction of tool wear enables proactive tool condition monitoring, reduces unplanned downtime, and optimizes tool replacement schedules, thereby lowering manufacturing costs. Over the years, extensive efforts have been dedicated to the development of robust tool wear prediction models, which can be broadly classified into mechanistic models, data-driven models, and physics-data fused models [15][16]. The characteristics of each tool wear prediction approach are summarized as **Fig. 1**.

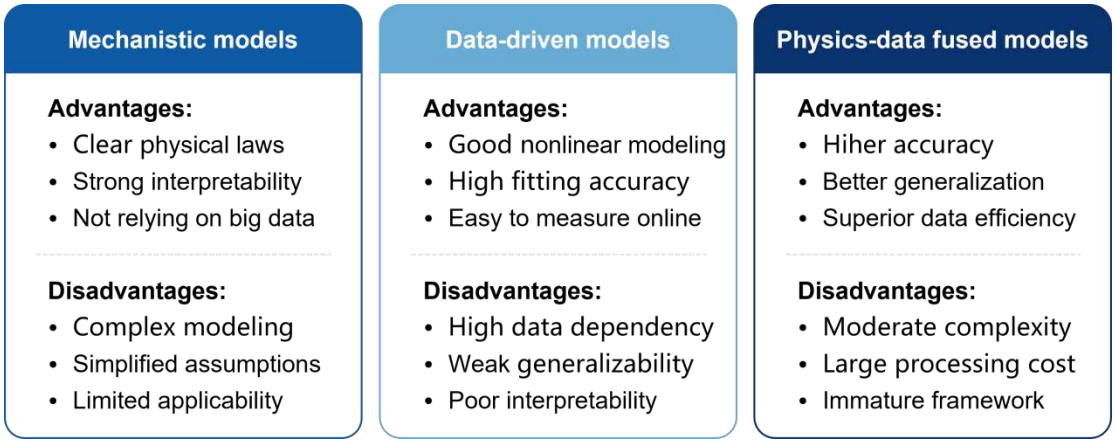



**Fig. 1** Comparison of characteristics of different tool wear prediction models

### 2.1 Mechanistic models for tool wear prediction

Mechanistic models leverage physical principles and mathematical formulations to accurately describe tool wear, enabling effective monitoring of the wear degradation process. Generally, these models can be classified into two categories: empirical models and analytical models. For example, as the most widely used empirical model, the Taylor tool life equation relates cutting speed to tool life under constant conditions. Building upon this formulation, the generalized Taylor model further integrates feed rate and depth of cut.

Analytical models based on physical wear mechanisms, such as abrasion and adhesion, offer more fundamental insights. For instance, Zhang et al. proposed an analytical model for tool wear of turning tungsten alloy, considering abrasive, adhesive, stress-driven, and thermal-driven wear [2]. Methon et al. presented a numerical method for flank wear prediction in milling 15-5 PH steel under abrasive and thermal-driven wear [3]. Jamshidi et al. built an analytical model for tool life prediction in grinding of Inconel 718, accounting for pullout, attrition, and fracture mechanisms [4]. Enrique et al. developed an analytical model for turning TiAl alloy, linking insert geometry, cutting speed, and contact stress to abrasive and thermal-driven wear [5]. Zhuang et al. focused on crater wear in ceramic insert turning of Ti6Al4V, proposing an

analytical model based on maximum temperature effects [6]. Bai et al. formulated a semi-analytical model for CFRP drilling, fusing two-body and three-body abrasive wear [7]. Ducroux et al. employed an exponential function to model abrasive, adhesive, and diffusion wear in milling of Inconel 718 [8]. Zhu and Zhang developed a mechanistic model with adjustable coefficients for flank wear prediction in ball-nose end milling of Inconel 718, accounting for running-in, adhesive, and abrasive wear [9]. A comprehensive summary of mechanistic models for tool wear prediction is presented in **Table 1**.

**Table 1** Relevant studies about mechanistic models for tool wear prediction

| Reference | Modeling object | Wear mechanism | Main contributions |
|---|---|---|---|
| Zhang et al. (2026) [2] | Turning tool for tungsten alloy | Abrasive, adhesive, stress, and thermal driven wear | A mechanistic model of tool flank wear considering force, temperature, and wear modes. |
| Methon et al. (2025) [3] | Milling cutter with inserts for 15-5 PH steel | Abrasive wear and thermal driven wear | 3D numerical method for predicting tool flank wear and handling milling tool paths and complex 3D geometries |
| Jamshidi at al. (2024) [4] | Grinding tool for Inconel 718 | Pullout, attrition, and fracture | Analytical model for tool life prediction in super abrasive machining |
| Enrique et al. (2023) [5] | Turning tool for TiAl alloy | Abrasive wear and thermal driven wear | Analytical model considering insert geometry, cutting speed, and contact stress |
| Zhuang et al. (2022) [6] | Turning for Ti6Al4V with ceramic inserts | crater wear and thermal driven wear | Analytical model for crater wear depth and width considering the effect of the maximum temperature |
| Bai et al. (2021) [7] | Drilling cutter for CFRP | two-body and three-body abrasive wear | Semi-analytical model incorporating tool wear rate and variation of half axis length |
| Ducroux et al. (2020) [8] | Milling cutter for Inconel 718 | Abrasive, adhesive, and diffusion wear | An exponential function for modeling tool wear along the cutting edge |
| Zhu and Zhang (2019) [9] | Ball-nose end milling cutter for Inconel 718 | running-in wear, adhesive wear, and abrasive wear | Adjustable coefficients-based mechanistic model for tool flank wear during ball-nose end milling. |

Although these models offer strong interpretability and require only limited data, they typically demand a detailed understanding of material properties, friction coefficients, and thermal fields. Moreover, the idealized assumptions embedded in the modeling process, such as homogeneous material behavior and constant contact conditions, further constrain their generalization ability across diverse operational scenarios. Consequently, purely mechanistic approaches suffer from limited applicability, motivating researchers to explore data-driven alternatives that learn complex wear patterns directly from process data.

## 2.2 Data-driven models for tool wear prediction

With advances in sensing, signal processing, and machine learning, data-driven approaches have emerged as an important paradigm for tool wear prediction. These methods generally leverage cutting force, vibration, acoustic emission, and spindle current signals as informative inputs. Earlier studies primarily employed artificial neural networks, support vector

regression, hidden Markov models, and tree-based learning algorithms. While these models perform satisfactorily under stable machining conditions, their effectiveness is critically contingent upon the quality, quantity, and representativeness of labeled training data.

Recent studies have increasingly investigated hybrid deep-learning architectures for extracting discriminative tool-wear representations. Lu et al. integrated a CNN-Transformer architecture to jointly capture local features and long-range dependencies for tool-wear monitoring [10]. Mundra et al. employed a CNN-BiLSTM model to characterize vibration features and bidirectional temporal dependencies [11]. Huang et al. developed a hybrid adaptation network to learn domain-invariant representations [12]. Shao et al. proposed a dual-branch Transformer-BiLSTM architecture with error-compensation to enhance prediction robustness [13]. Chen et al. developed a temporal-frequency enhanced Transformer to extract wear-sensitive representations across complementary domains [14]. Hou et al. combined a ResNet-GRU dual-channel network with subdomain adaptation to fuse spatial-temporal features and mitigate cross-condition distribution discrepancies [15][16]. Collectively, these studies demonstrate the effectiveness of hybrid architectures, multi-domain feature extraction, and domain adaptation in capturing complex tool-wear. Related data-driven models for tool-wear prediction are summarized in **Table 3**.

**Table 2** Recent studies about data-driven models for tool wear prediction

| Reference | Input signals | Intelligent model | Main contributions |
|---|---|---|---|
| Lu et al. (2026) [10] | Cutting force | CNN-Transformer hybrid network | Deep model embedding into digital twin framework for online monitoring |
| Mundra et al. (2026) [11] | Cutting vibration | CNN-BiLSTM network | Extracting localized features via CNN layers and modeling sequential dependencies via BiLSTM layers |
| Huang et al. (2024) [12] | Cutting vibration | Residual attention hybrid adaptation network | Enabling cross-domain tool wear condition monitoring via learning domain-invariant features |
| Shao et al. (2026) [13] | Cutting force and vibration | Dual-branch transformer-BiLSTM network | Enhancing model robustness via error compensation within a dual-branch structure |
| Chen et al.(2026) [14] | Cutting force, vibration, and acoustic emission | Temporal-frequency enhanced Transformer (TFETformer) | Enhancing wear-sensitive features simultaneously in the temporal and frequency domains and improving robustness against noise and redundant information |
| Hou et al. (2025) [15] | Cutting force, vibration, and acoustic emission | ResNet-GRU dual-channel network with deep subdomain adaptation | Fusing spatial and temporal wear features and aligning subdomain distributions to improve prediction generalization across machining conditions |
| Adishesha et al. (2026) [16] | Spindle power | Random forest (RF) and BiLSTM | Embedding RF and BiLSTM models into an intelligent system for tool wear prognosis and diagnosis |
| Jones and Cao (2025) [17] | Cutting force, vibration, and acoustic emission | Three-level data fusion architecture with RF-ANN-XGBoost stacking ensemble | Integrating raw-data, feature-level, and decision-level fusion to improve the accuracy and robustness of tool flank wear prediction |

Alongside deep neural networks, ensemble-learning methods have been attracted attention for their ability to capture nonlinear relationships in heterogeneous machining data with relatively low architectural complexity. Adishesha et al. evaluated random forest and BiLSTM models for tool wear prediction [16]. Jones and Cao developed a fusion framework for tool wear prediction by integrating random forest, neural network, and XGBoost-based stacking ensembles. [17]. Among tree-based ensembles, gradient boosting regression trees (GBRTs) are particularly attractive because their sequential residual-correction mechanism effectively captures complex nonlinear interactions while offering flexible control of model complexity through tree depth, learning rate, subsampling, and boosting iterations. These characteristics make GBRTs well suited to small-sized machining datasets with heterogeneous and engineered features. However, conventional GBRT remains purely data-driven, relying solely on prediction error minimization. Consequently, its predictions may become physically inconsistent under limited training data or when extrapolating beyond the training distribution, motivating the integration of machining physics into the learning framework.

## 2.3 Physics-data fused models for tool wear prediction

Physics-data fusion modeling has emerged as a promising paradigm for addressing the complementary limitations of mechanistic and purely data-driven approaches. By integrating physical knowledge into data processing, feature extraction, model architectures, objective functions, probabilistic priors, or prediction correction, this framework can enhance both the physical consistency and generalization capability of data-driven models.

Recent studies have explored physics-data fusion for tool-wear prediction by embedding physical knowledge at different stages of the learning process. Zhang et al. developed a dual-branch physics-informed neural network that jointly learns multisensory-wear mappings and degradation dynamics under data-fitting, physics-residual, and monotonicity constraints [18]. Wang et al. integrated a time-accumulated wear model with multi-domain features in a physics-informed Gaussian process to enable wear prediction and uncertainty quantification under limited data [19]. Liu et al. embedded mechanism-based wear relationships into a recurrent unit coupled with a Transformer to capture both local degradation and long-range temporal dependencies [20]. Kim et al. combined latent-space alignment, pseudo-label regularization, and physics-guided inference correction for unsupervised cross-condition tool-wear prediction [21]. Guo et al. integrated adaptive denoising, multi-scale feature extraction, and uncertainty-aware physics-informed optimization to enhance robustness under noisy conditions [22]. Chen et al. employed a physics-informed dynamic gated graph network to model evolving sensor dependencies while enforcing physically consistent wear progression [23]. Pashmforoush et al. combined a milling-force model with machine learning to generate physics-based training data for flank-wear estimation [24], whereas Sun et al. embedded a physical wear model into the

Gaussian-process prior to capture global degradation trends and condition-dependent deviations [25]. Related fusion approaches for tool-wear prediction are summarized in **Table 3**.

**Table 3** Recent studies about physics-data fused models for tool wear prediction

| Reference | Physic models | Fusion strategy | Main contributions |
|---|---|---|---|
| Zhang et al.(2026) [18] | Implicit tool-wear-rate and wear monotonicity constraint | Physics-informed neural network integrating wear-rate, data-fitting loss, and monotonicity | Integrating tool wear monitoring and multi-step forecasting and improving prediction robustness under limited data |
| Wang et al.(2026) [19] | Time-accumulated tool wear evolution model | Embedding a time-accumulated wear model into Gaussian process regression | Improving physical consistency and small-sample prediction while providing confidence intervals for uncertainty quantification |
| Liu et al.(2026) [20] | Abstracted tool wear mechanism model | Embedding wear mechanisms into a mechanism-based recurrent unit combined with a Transformer | Capturing short- and long-term wear dependencies and quantitatively revealing the evolution of different wear mechanisms under limited-sample conditions |
| Kim et al.(2026) [21] | Analytical tool wear model under varying machining conditions | Physics-guided prediction adjustment combined with unsupervised domain adaptation | Enabling physically consistent tool wear prediction under unseen cutting conditions without labeled target-domain wear data |
| Guo et al. (2026) [22] | Nonlinear tool wear accumulation model | Physics-informed deep adaptive denoising and multi-scale feature fusion network | Improving noise robustness and prediction accuracy by integrating adaptive spectral denoising, multi-scale temporal feature fusion, and nonlinear wear accumulation constraints |
| Chen et al.(2025) [23] | Exponential empirical wear-time evolution model | Dynamic gated graph convolutional network constrained by an exponential wear-evolution model | Modeling dynamic spatial-temporal sensor relationships while enforcing wear monotonicity and wear-rate consistency |
| Pashmforoush et al.(2025) [24] | Wear-inclusive thermo-mechanical milling-force model | Combining a wear-inclusive thermo-mechanical force model with inverse machine learning | Generating physics-based training data and estimating flank wear from measured cutting forces with reduced dependence on extensive wear experiments |
| Sun et al.(2024) [25] | Explicit health-indicator-based tool wear evolution model | Embedding a physical wear model into the mean function of Gaussian process regression | Linking wear-sensitive health indicators with physical wear evolution and improving generalization across multiple cutting conditions |

Existing physics-data fusion studies have focused predominantly on neural networks and probabilistic models, while the incorporation of physical degradation knowledge into ensemble tree-based models remains comparatively underexplored. Gradient boosting regression trees (GBRTs) are particularly well suited to small and heterogeneous machining datasets because their sequential residual-correction mechanism provides strong nonlinear approximation with relatively modest model complexity [26][27]. However, their discrete feature partitioning and tree-based optimization are fundamentally different from the differentiable architectures of

neural networks, precluding the straightforward use of automatic differentiation and physics-residual backpropagation. This methodological distinction highlights the need for alternative strategies to incorporate physical degradation knowledge into tree-based learning frameworks.

## 2.4 Research gap and motivation

Despite substantial progress in mechanistic, data-driven, and physics-informed approaches, achieving accurate and physically consistent tool-wear prediction remains challenging. Mechanistic models offer physical interpretability but require numerous parameters and simplifying assumptions, limiting their adaptability to the dynamically varying conditions. Data-driven methods capture complex nonlinear signal-wear relationships but remain highly dependent on labeled-data quality and coverage, particularly under data-scarce, noisy, or extrapolative conditions. Although GBRT is well suited to small and heterogeneous machining datasets, conventional GBRT minimizes prediction error alone and lacks explicit constraints on physically plausible wear evolution.

These limitations reveal three key research gaps. First, purely data-driven GBRT models may yield physically inconsistent wear predictions under sparse training data or operating conditions beyond the training distribution. Second, physics-informed learning remains largely concentrated on differentiable neural and probabilistic models, with limited application to ensemble tree-based methods. Third, practical frameworks for incorporating generalized tool-wear degradation knowledge into gradient boosting methods remain lacking, particularly those that avoid reliance on complete mechanistic models or extensive labeled datasets.

To address these gaps, this study proposes a physics-informed gradient boosting regression tree (PIGBRT) framework for milling flank-wear prediction. Rather than embedding a complete mechanistic model into the learning process, the proposed approach introduces a generalized tool-wear degradation constraint as a physics-based regularization term alongside the conventional prediction loss. This formulation enables GBRT to retain its strong nonlinear approximation and data efficiency while penalizing physically implausible degradation trends, thereby establishing a practical balance between data fidelity and physical consistency for robust tool-wear prediction under realistic machining conditions.

# 3 Research methodology

## 3.1 Overall framework of proposed method

The proposed framework for predicting tool flank wear using a physics-Informed gradient boosting regression tree (PIGBRT) is depicted in **Fig. 2**. This proposed approach consists of four distinct phases: (i) signal acquisition and wear measurement, (ii) data preprocessing and feature selection, (iii) PIGBRT modeling for wear prediction, and (iv) performance evaluation and analysis.

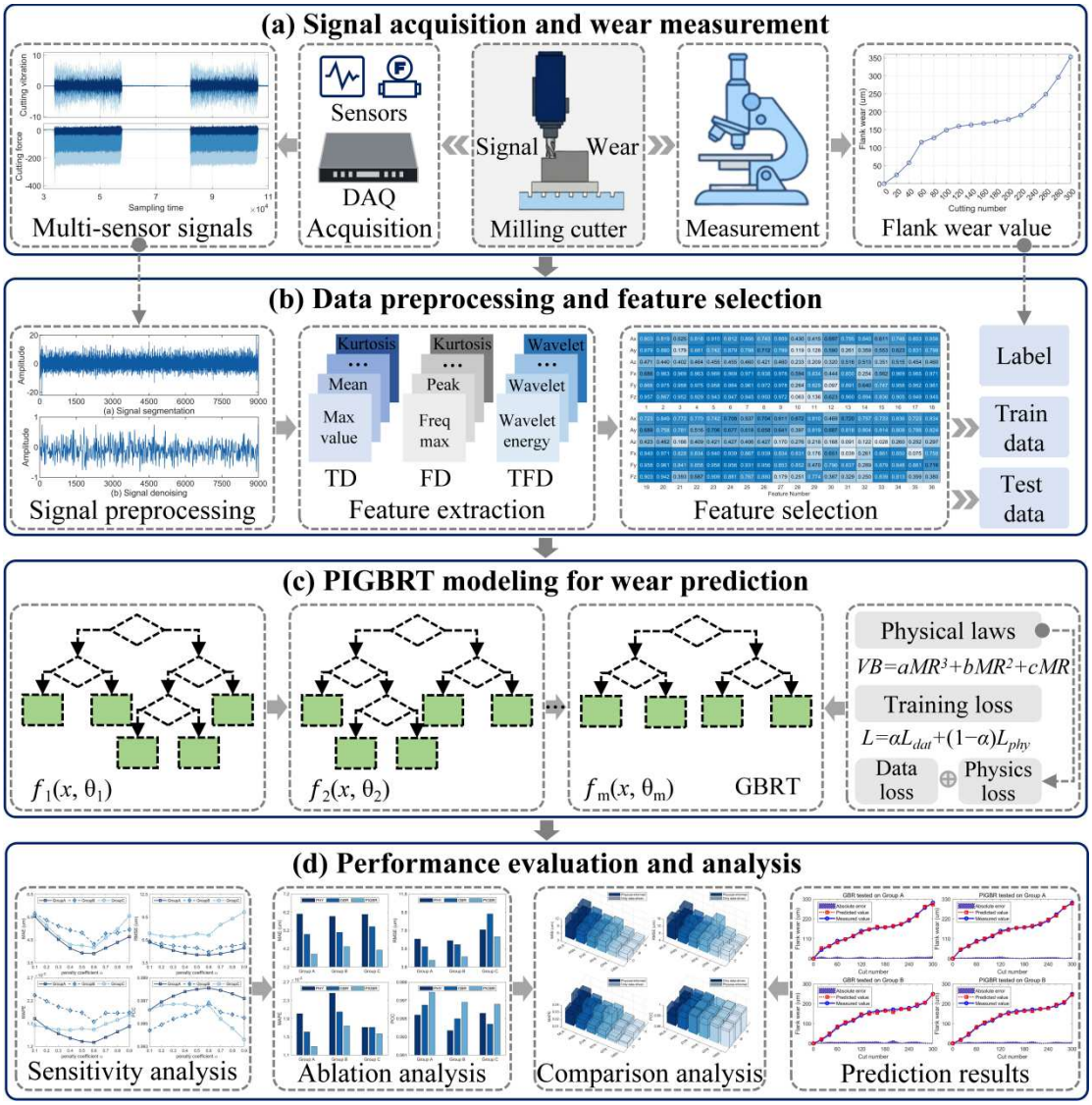



**Fig. 2** Overall framework of the proposed PIGBRT-based tool wear prediction approach

First, milling experiments were conducted on a CNC machining center under various cutting conditions to capture multi-sensory signals associated with tool flank wear. Cutting forces and vibrations were monitored in real time using piezoelectric dynamometer and accelerometer, while the flank wear width (VB) was measured offline via a high-resolution optical microscope. The resulting synchronized signal-wear dataset establishes a robust foundation for the predictive model.

Then, raw multi-sensory signals were segmented and denoised to eliminate noise and drift, followed by the extraction of time-domain (TD), frequency-domain (FD), and time-frequency domain (TFD) features. To reduce redundancy and mitigate the curse of dimensionality, a feature selection strategy using Pearson correlation analysis was employed. This approach retained only features strongly correlated with progressive flank wear, effectively reducing dimensionality while preserving predictive accuracy.

Next, the proposed PIGBRT model integrates conventional gradient boosting with a

physics-based regularization term from the Taylor tool wear equation, incorporating a penalty to constrain wear evolution to physically plausible trends. By embedding domain knowledge of wear mechanisms, it effectively mitigates over-fitting under sparse data and enhances extrapolation beyond the training parameter range. The model outputs continuous wear values per cutting pass, facilitating real-time tool flank wear monitoring.

Finally, to validate the effectiveness of PIGBRT, comparative experiments were conducted against both ablation models and benchmark methods under identical datasets and evaluation metrics. Through performance evaluation and analysis, PIGBRT exhibits superior robustness against noisy sensor inputs and maintains monotonic wear predictions. Overall, the proposed PIGBRT framework offers a reliable and interpretable solution for tool flank wear prediction in smart manufacturing environments.

## 3.2 GBRT modeling for tool wear prediction

### 3.2.1 GBRT-based data-driven prediction model

Boosting algorithm, as a powerful ensemble learning method, can significantly improve the predictive performance of a model by iteratively training a series of weak learners and weighting them together to form a strong learner. As a type of the Boosting algorithm, the fundamental principle of GBRT is to construct a new regression tree based on the negative gradient trend of the residuals from the current regression tree [28], and to iteratively update the model along this direction, thereby obtaining a stable and robust model. Consequently, GBRT was employed in this study to develop a data-driven model for the precise prediction of tool flank wear.

For the given selected feature $x$ related to tool flank wear, the *m*-th data-driven iteration model $\hat{y}_{\mathrm{d}}^{\mathrm{m}}(x)$ of GBRT is calculated as follows:

$$\hat{y}_{\mathrm{d}}^{\mathrm{m}}(x) = \hat{y}_{\mathrm{d}}^{\mathrm{m}-1}(x) + \omega_m h_m(x;\theta_m) \tag{1}$$

where $\hat{y}_{\mathrm{d}}^{\mathrm{m}-1}(x)$is the previous data-driven iteration model for tool flank wear prediction, $\omega_m$ denotes the weight of the regression tree, $h_m(x;\theta_m)$ represents the regression tree function, and $\theta_m$ denotes the estimated parameter of the regression tree.

In practice, the parameter $\theta_m$ is obtained by minimizing the loss function $L(y_i, \hat{y}_{\mathrm{d}}^{\mathrm{m}}(x_i))$ of the model $\hat{y}_{\mathrm{d}}^{\mathrm{m}}(x)$ during the regression process. It is expressed as follows:

$$\theta_m = argmin \sum_{i=1}^{N} L(y_i, \hat{y}_{\mathrm{d}}^{\mathrm{m}}(x_i)) \tag{2}$$

For regression problems, the loss function is defined as the squared error between the predicted and actual values, as follows:

$$L(y_i, \hat{y}_{\mathrm{d}}^{\mathrm{m}}(x_i)) = [y_i - \hat{y}_{\mathrm{d}}^{\mathrm{m}}(x_i)]^2 = [r_{mi} - \omega_m h_m(x_i;\theta_m)]^2 \tag{3}$$

where $r_{mi} = y_i - \hat{y}_{\mathrm{d}}^{\mathrm{m}-1}(x_i)$ is the residual error of the fitting model, corresponding to the minimum value of the previous loss function.

The residual error $r_{mi}$ is estimated using the negative gradient value of the loss function,

and its expression as follows:

$$r_{mi} = -\left[\frac{\partial L(y_i, \hat{y}_d(x_i))}{\partial\, \hat{y}_d(x_i)}\right]_{\hat{y}_d(x)=\hat{y}_{\mathrm{d}}^{\mathrm{m-1}}(x)} \tag{4}$$

Finally, the tool flank wear prediction model is obtained by continuously minimizing error iterations, and the calculation process of the GBRT-based prediction model is illustrated in **Fig.3**.

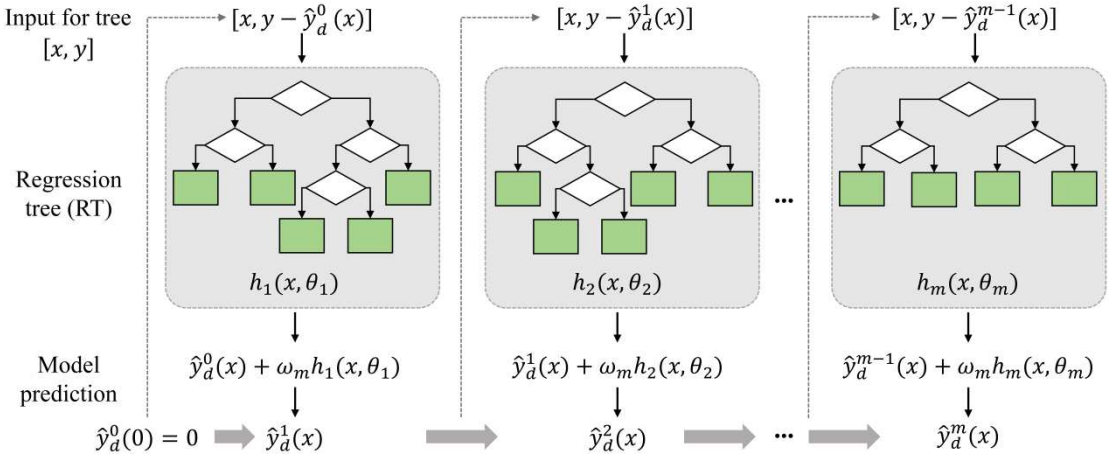



**Fig. 3** The calculation process of the GBRT-based prediction model for tool wear prediction

## 3.2.2 Physics-informed GBRT for tool wear prediction

By integrating physical laws of tool flank wear, a physics-informed GBRT approach effectively synthesizes the complementary strengths of both mechanistic and data-driven paradigms. This integration significantly enhances the model's generalization capability and robustness.

In general, the Taylor formula represents the most classical empirical model for tool wear; however, its applicability is constrained by overly simplified assumptions. To overcome these limitations, the generalized Taylor formula has been developed to characterize the relationship more comprehensively between tool wear, machining time, and key process parameters **错误!未找到引用源。**. The generalized Taylor formula for tool wear process is as follows:

$$VB(t) = C \cdot t^{E_t} V_c^{E_v} f_z^{E_f} a_p^{E_p} a_e^{E_e} \tag{5}$$

where $VB(t)$ is the flank wear width at the *t*-th moment, $V_c$, $f_z$, $a_p$, and $a_e$ denote cutting speed, feed rate, cutting depth, and cutting width, respectively. Furthermore, the remaining parameters ($C$, $E_t$, $E_v$, $E_f$, $E_p$, and $E_e$) represent empirical fitting coefficients, which are treated as constants within the model.

To model the degradation evolution of tool flank wear (VB) with respect to number of cutting passes, a third-order polynomial empirical function is employed to represent the physical degradation process **错误!未找到引用源。**. It is mathematically expressed as:

$$y_h \triangleq VB(t) = aMR^3 + bMR^2 + cMR \tag{6}$$

where $MR$ denotes the accumulated material removal at the *t*-th moment, while the remaining

parameters ($a$, $b$, and $c$) are empirical fitting coefficients.

In this study, the generalized tool wear degradation law is incorporated as a regularization term in the loss function of the data-driven model, serving as the physical loss component. Its expression is as follows:

$$L_{phy} = \sum_{i=1}^{N} \left\| y_i - \hat{y}_h^i \right\|_2^2 \tag{7}$$

where $L_{phy}$ denotes the physical loss, $y_i$ is the actual value of tool flank wear at the $i$-th moment, $\hat{y}_h^i$ is the prediction value for physical constraints at the $i$-th moment, $N$ represents the number of the input samples.

Additionally, to quantify the deviation from empirical observations, the prediction error of the tool wear is treated as the data loss component, as follows:

$$L_{dat} = \sum_{i=1}^{N} \left\| y_i - \hat{y}_d^i \right\|_2^2 \tag{8}$$

where $L_{dat}$ denotes the physical loss, $\hat{y}_d^i$ is the prediction value for pure data-driven model at the $i$-th moment.

The total objective function for training the GBRT model is formulated as a combination of the data loss and physical loss terms, as follows:

$$\min_{\theta^*, \alpha^*} L = \min_{\theta^*, \alpha^*} \left[ \alpha L_{dat} + (1 - \alpha) L_{phy} \right] \tag{9}$$

where $\alpha$ denotes the balancing factor that adjusts the relative contribution between the data loss and the physical loss, $\theta^*$ represent the parameters of the optimal GBRT model, and $\alpha^*$ corresponds to the optimal value of this balancing factor.

Ultimately, by effectively balancing the contributions of data-driven learning and physical constraints, the proposed PIGBRT-based method yields robust predictions of tool flank wear under varying operating conditions.

# 4 Experimental verifications

## 4.1 Experimental setup

To evaluate the performance of the proposed physics-informed tool wear prediction method, a series of high-speed milling experiments were conducted on die steel using four-flute solid carbide cutters on a computer numerical control (CNC) machining center. During the experiments, cutting forces in three orthogonal directions were measured using a Kistler 9347C triaxial dynamometer installed between the machine table and the workpiece. Simultaneously, vibration signals generated during machining were captured by a Kistler 8763B triaxial accelerometer magnetically attached to the lateral surface of the workpiece. The raw charge signals from both sensors are conditioned using a charge amplifier and digitized at a sampling frequency of 2500 Hz via a National Instruments data acquisition (NI-DAQ) system. These synchronized multi-sensor signals were subsequently used for the indirect monitoring and assessment of tool wear progression. After completing the milling trials for each cutter, flank

wear on all four cutting edges was measured using an optical microscope of Hirox KH-7700. In this study, the average flank wear among the edges was taken as the representative actual wear of the milling cutter. A schematic illustration of the data acquisition procedure within the experimental rig is presented in **Fig. 4**.

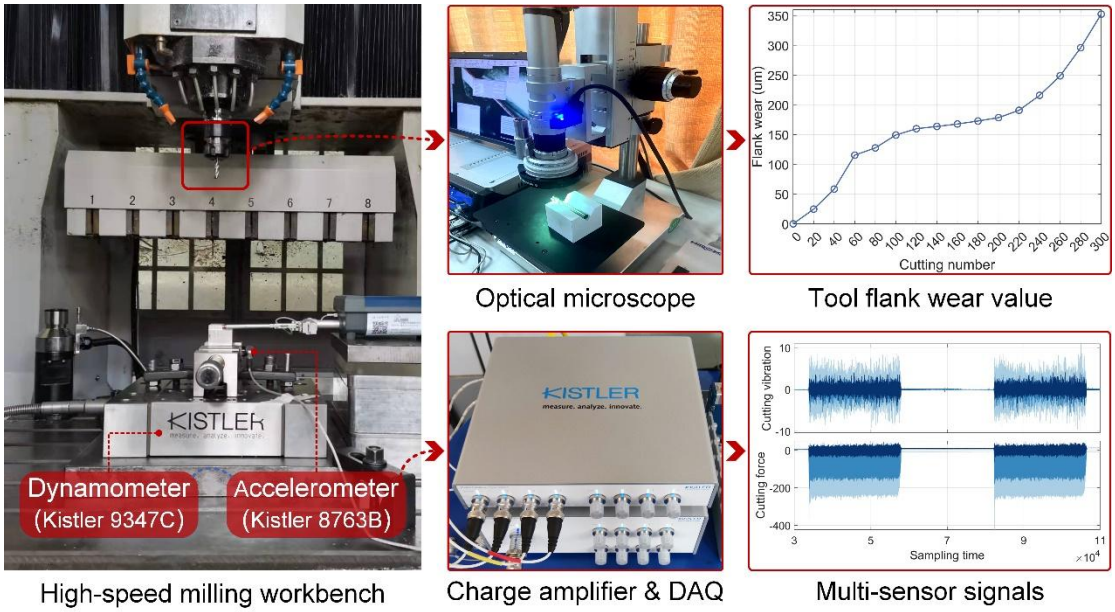



**Fig. 4** A schematic illustration of the data acquisition procedure within the experimental rig.

To further validate the generalization capability of the proposed method, three sets of milling experiments were conducted. Specifically, across the three experimental groups, the radial depth of cut and material removal rate were held constant, while the spindle speed, feed rate, feed engagement, and axial depth of cut were varied. The detailed machining parameters for each group are summarized in **Table 1**. Furthermore, recognizing that in real-world industrial production, it is generally infeasible to conduct offline flank wear measurements immediately after each cutting pass without disrupting the ongoing machining operation. To better simulate such practical conditions, a series of milling experiments was designed using a set of geometrically and materially identical cutting tools. Each experimental group consisted of fifteen individual milling cutters, with each cutter subjected to an increasing number of feeds passes in increments of twenty. Specifically, the first cutter completed 20 feed passes, the second cutter performed 40 feeds, and so forth, culminating in the fifteenth cutter performing 300 feed passes. This progressive experiment effectively simulates cumulative tool wear under uninterrupted, while enabling wear assessment without operational interference.

**Table 1** The machining parameters for milling operations in this study

| Group | Spindle speed (r/min) | Feed speed (mm/min) | Feed engagement (mm/r) | Axial cutting depth (mm) | Radial cutting depth (mm) | Material removal rate |
|---|---|---|---|---|---|---|
| A | 2000 | 500 | 0.0625 | 6 | 0.15 | 7.5 |
| B | 1500 | 750 | 0.125 | 6 | 0.10 | 7.5 |
| C | 1500 | 375 | 0.0625 | 6 | 0.20 | 7.5 |

## 4.2 Data preprocessing

In the milling experiments, multi-sensor signals are acquired continuously throughout the wear progression of each milling cutter. First, the cutting segments are separated from the raw collected signals to eliminate interference from non-cutting phases. Next, to mitigate the effects of occasional disturbances in the machining environment, outliers within the cutting segments are identified and replaced with the average of the maximum values from their adjacent left and right neighborhoods. Finally, a four-layer wavelet packet denoising algorithm with a soft threshold is applied to the processed signals to further enhance signal quality. To reduce signal redundancy and minimize the influence of dynamic characteristics during milling processes, multi-domain features are extracted from the preprocessed signals. Inspired by Ref. 错误!未找到引用源。, the descriptions and corresponding formulas for time-domain and frequency-domain features are provided in **Table 2**, where $\boldsymbol{x}$ represents the preprocessed signal, $N$ is the number of sampling points, $\boldsymbol{v}$ denotes the frequency calculated by fast Fourier transform, $K$ is the number of frequency points, $\boldsymbol{s}$ represents the power spectrum density using Welch method, $M$ is the number of power spectrum density points.

As shown in **Table 2**, the features $f_1 \sim f_{15}$ reflect the time-domain characteristics of the preprocessed signals. The features $f_1 \sim f_7$ capture the amplitude characteristics of the signals, while the features $f_8 \sim f_{15}$ characterize the distribution properties of the signals over time. The features $f_{16} \sim f_{28}$ reflect the frequency-domain characteristics of the preprocessed signals. The features $f_{16} \sim f_{22}$ both the amplitude and distribution characteristics of the frequency spectrum, whereas the features $f_{23} \sim f_{28}$ characterize the amplitude and distribution properties of the power spectrum density. In addition, wavelet packet energy is introduced as a time-frequency feature to overcome the limitation of reduced frequency resolution. A three-layer wavelet packet decomposition using the db8 basis is applied to the preprocessed signals, yielding eight energy values that constitute the features $f_{29} \sim f_{36}$.

**Table 2** Expressions of the extracted time-domain and frequency-domain features

| Description | Expression | Description | Expression |
|---|---|---|---|
| Maximum value | $f_1 = max(\boldsymbol{x})$ | Minimum value | $f_2 = min(\boldsymbol{x})$ |
| Mean value | $f_3 = \frac{\sum_{i=1}^{N} \boldsymbol{x}_i}{N}$ | Peak value | $f_4 = max(\lvert\boldsymbol{x}\rvert)$ |
| Absolute mean | $f_5 = \frac{\sum_{i=1}^{N} \lvert\boldsymbol{x}_i\rvert}{N}$ | Peak to peak value | $f_6 = max(\boldsymbol{x}) - min(\boldsymbol{x})$ |
| Root mean square | $f_7 = \sqrt{\sum_{i=1}^{N} \frac{\boldsymbol{x}_i{}^2}{N}}$ | Variance | $f_8 = \frac{\sum_{i=1}^{N} (\boldsymbol{x}_i - \bar{\boldsymbol{x}}_i)^2}{N}$ |
| Standard deviation | $f_9 = \sqrt{\frac{\sum_{i=1}^{N} (\boldsymbol{x}_i - \bar{\boldsymbol{x}}_i)^2}{N}}$ | Impulse factor | $f_{10} = \frac{max(\lvert\boldsymbol{x}\rvert)}{\left(\frac{\sum_{i=1}^{N} \boldsymbol{x}_i}{N}\right)}$ |

| | | | |
|---|---|---|---|
| Shape factor | $f_{11}=\frac{\sqrt{\sum_{i=1}^{N}\frac{\boldsymbol{x}_i^2}{N}}}{\left(\frac{\sum_{i=1}^{N}\boldsymbol{x}_i}{N}\right)}$ | Crest factor | $f_{12}=\frac{max(\lvert\boldsymbol{x}\rvert)}{\left(\sqrt{\sum_{i=1}^{N}\frac{\boldsymbol{x}_i^2}{N}}\right)}$ |
| Clearance factor | $f_{13}=\frac{max(\lvert\boldsymbol{x}\rvert)}{\frac{\sum_{i=1}^{N}\lvert\boldsymbol{x}_i\rvert}{N}}$ | Skewness | $f_{14}=\sum_{i=1}^{N}\left[\frac{(\boldsymbol{x}_i-\bar{\boldsymbol{x}})}{\sigma_x}\right]^3$ |
| Kurtosis | $f_{15}=\sum_{i=1}^{N}\left[\frac{(\boldsymbol{x}_i-\bar{\boldsymbol{x}})}{\sigma_s}\right]^4$ | Frequency maximum | $f_{16}=max(\boldsymbol{v})$ |
| Frequency Mean | $f_{17}=\frac{\sum_{i=1}^{K}\boldsymbol{v}_i}{K}$ | Frequency peak | $f_{18}=max(\lvert\boldsymbol{v}\rvert)$ |
| Frequency RMS | $f_{19}=\sqrt{\sum_{i=1}^{K}\frac{\boldsymbol{v}_i^2}{K}}$ | Frequency variance | $f_{20}=\frac{\sum_{i=1}^{K}(\boldsymbol{v}_i-\bar{\boldsymbol{v}}_i)^2}{K}$ |
| Frequency standard deviation | $f_{21}=\sqrt{\frac{\sum_{i=1}^{K}(\boldsymbol{v}_i-\bar{\boldsymbol{v}}_i)^2}{K}}$ | Frequency skewness | $f_{22}=\sum_{i=1}^{K}\left[\frac{(\bar{\boldsymbol{v}}_i-\bar{\boldsymbol{v}})}{\sigma_v}\right]^3$ |
| Frequency kurtosis | $f_{23}=\sum_{i=1}^{K}\left[\frac{\bar{\boldsymbol{v}}_i-\bar{\boldsymbol{v}}}{\sigma_v}\right]^4$ | Spectral mean | $f_{24}=\frac{\sum_{i=1}^{M}\boldsymbol{s}_i}{M}$ |
| Spectral variance | $f_{25}=\frac{\sum_{i=1}^{M}(\boldsymbol{s}_i-\bar{\boldsymbol{s}}_i)^2}{M}$ | Spectral standard deviation | $f_{26}=\sqrt{\frac{\sum_{i=1}^{M}(\boldsymbol{s}_i-\bar{\boldsymbol{s}}_i)^2}{M}}$ |
| Spectral skewness | $f_{27}=\sum_{i=1}^{M}\left[\frac{(\boldsymbol{s}_i-\bar{\boldsymbol{s}})}{\sigma_s}\right]^3$ | Spectral kurtosis | $f_{28}=\sum_{i=1}^{M}\left[\frac{(\boldsymbol{s}_i-\bar{\boldsymbol{s}})}{\sigma_s}\right]^4$ |

To eliminate the influence of differing scales among features, all extracted multi-domain features are normalized in this study. Subsequently, a correlation analysis conducted between the normalized features and the progression of tool wear during the milling process. Taking Group B as an example, the results of this correlation analysis are illustrated in **Fig. 5**, where Ax, Ay, and Az denote the vibration signals, and Fx, Fy, and Fz represent the force signals, in the x, y, and z directions, respectively. It is evident that certain features, such as time-domain and frequency domain features extracted from the cutting force signals, exhibit strong correlation with the flank wear progression of the milling cutter. In contrast, a subset of the time-domain features extracted from cutting vibration signals show limited sensitivity to flank wear. Given this variability in relevance and responsiveness, feature selection and fusion are essential to identify and integrate the most informative and wear-sensitive components from the multi-sensor signals for accurate tool wear monitoring.

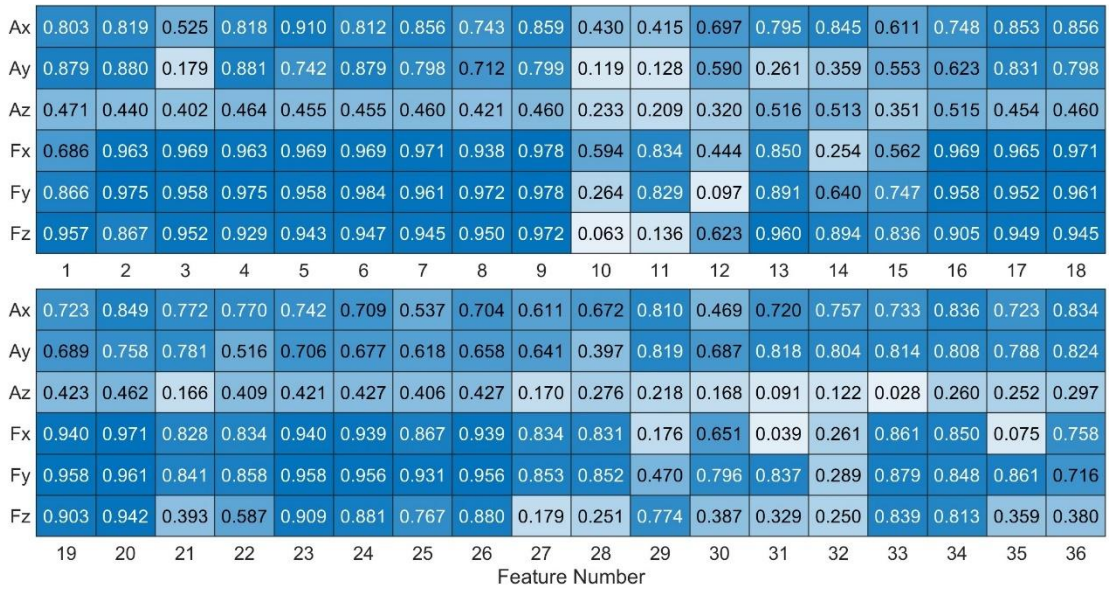



**Fig. 5** Correlation analysis of multi-domain features during tool wear process from Group B

In addition, to mitigate feature redundancy and suppress the influence of non-sensitive attributes, only features exhibiting pairwise correlation coefficients greater than 0.8 are retained as candidate indicators. Subsequently, eigenvalue decomposition is applied to the resulting feature matrix. The number of principal components is determined according to the cumulative explained variance ratio of the eigenvalues. Specifically, principal components are retained until their cumulative variance contribution exceeds 95% of the total variance, thereby ensuring that the reduced feature set captures the dominant modes of variation associated with tool wear. This refined subset comprising of the most sensitive and informative features for characterizing tool wear in milling operations is ultimately used as input for the modeling stage.

## 4.3 Evaluation metric

For a comprehensive quantitative evaluation of the proposed physical-informed tool wear prediction approach, four performance metrics are utilized in this study. These metrics comprise the mean absolute error (MAE), the root mean squared error (RMSE), the mean absolute percentage error (MAPE) and the Pearson correlation coefficient (PCC). Let $\boldsymbol{y}_i$ and $\hat{\boldsymbol{y}}_i$ denote the actual and predicted values of flank tool wear, the mathematical formulations for these evaluation metrics are given as follows:

$$\begin{cases} MAE = \frac{1}{N}\sum_{i=1}^{N}|\hat{\boldsymbol{y}}_i - \boldsymbol{y}_i| \\ RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(\hat{\boldsymbol{y}}_i - \boldsymbol{y}_i)^2} \\ MAPE = \frac{1}{N}\sum_{i=1}^{N}\frac{|\hat{\boldsymbol{y}}_i - \boldsymbol{y}_i|}{\boldsymbol{y}_i} \\ PCC = \frac{\sum_{i=1}^{N}(\boldsymbol{y}_i - \bar{\boldsymbol{y}})(\hat{\boldsymbol{y}}_i - \bar{\hat{\boldsymbol{y}}})}{\sqrt{\sum_{i=1}^{N}(\boldsymbol{y}_i - \bar{\boldsymbol{y}})^2 \sum_{i=1}^{N}(\hat{\boldsymbol{y}}_i - \bar{\hat{\boldsymbol{y}}})^2}} \end{cases} \tag{10}$$

where $N$ is the total number of test samples.

Specifically, MAE indicates the average magnitude of prediction residuals, the RMSE

accentuates larger deviations through quadratic weighting, MAPE provides a scale-independent relative error assessment, and PCC captures the linear dependency predicted and actual wear values. Additionally, MAE, RMSE, and MAPE are error-based metrics wherein smaller values indicate higher predictive accuracy. In contrast, the PCC increases with better performance, establishing it as a positively oriented metric.

# 5 Results and discussion

## 5.1 Hyperparametric sensitivity analysis

As shown in Eq. (10), the penalty coefficient $\alpha$ serves as a critical hyperparameter in the training process of the proposed PIGBRT model, governing the balance between physical-informed loss and data-driven loss. To determine its optimal value, a systematic sensitivity analysis is conducted by evaluating tool wear prediction performance across a range of penalty values. The evaluation results from three trials are visualized in **Fig. 6**, and the corresponding quantitative metrics are summarized in **Table 3**.

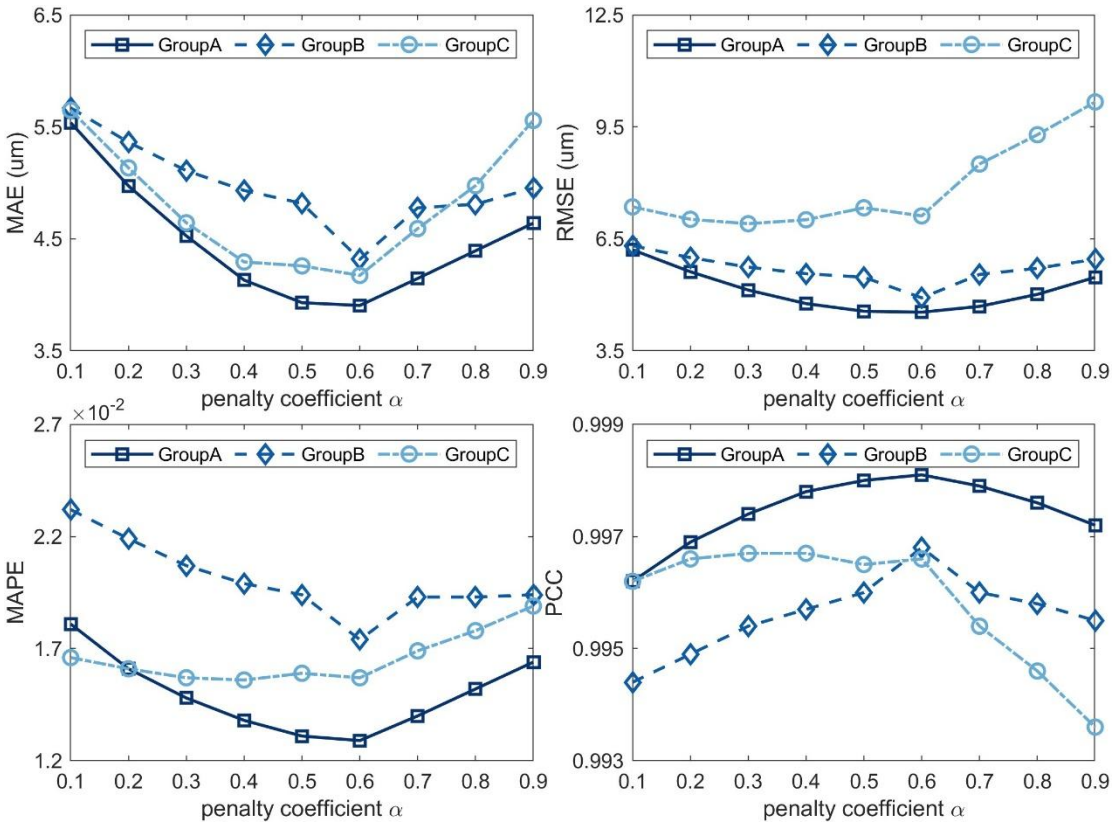



**Fig. 6** Evaluation results of tool wear prediction under different penalty coefficients

As shown in **Fig. 6**, the relationship between the penalty coefficient and the performance of the tool wear prediction model is non-monotonic. Specifically, when the penalty coefficient is below 0.6, its further reduction results in an increase in prediction error (i.e., MAE, RMSE, and MAPE). This occurs because the physical-informed model places excessive emphasis on satisfying physical constraints at the expense of fully leveraging the information content embedded in the sensor data. Conversely, when the penalty coefficient exceeds 0.6, any further

increase leads to a degradation in prediction accuracy. This is primarily due to the over-enforcement of physical constraints, which excessively limits the model's capacity to adapt to the empirical patterns captured in the sensor data. Therefore, as summarized in **Table 4**, a penalty coefficient of 0.6 is selected as the optimal value in this study, as it yields the lowest prediction error by effectively balancing the enforcement of physical constraints with the flexibility of data-driven modeling.

**Table 3** Overall evaluation of tool wear prediction under different penalty coefficients

| Alpha | MAE (um) | RMSE (um) | MAPE | PCC |
|---|---|---|---|---|
| 0.1 | 5.6188±0.0580 | 6.6248±0.5186 | 0.0193±0.0028 | 0.9956±0.0009 |
| 0.2 | 5.1559±0.1600 | 6.2116±0.5942 | 0.0180±0.0027 | 0.9961±0.0009 |
| 0.3 | 4.7593±0.2517 | 5.9243±0.7371 | 0.0171±0.0026 | 0.9965±0.0008 |
| 0.4 | 4.4520±0.3456 | 5.7795±0.9295 | 0.0165±0.0025 | 0.9968±0.0008 |
| 0.5 | 4.3347±0.3663 | 5.7866±1.1534 | 0.0161±0.0026 | 0.9968±0.0009 |
| **0.6** | **4.1313±0.1708** | **5.5214±1.1368** | **0.0153±0.0018** | **0.9972±0.0006** |
| 0.7 | 4.5042±0.2645 | 6.2459±1.6339 | 0.0167±0.0021 | 0.9964±0.0011 |
| 0.8 | 4.7252±0.2445 | 6.6699±1.8734 | 0.0174±0.0017 | 0.9960±0.0013 |
| 0.9 | 5.0499±0.3808 | 7.1961±2.1100 | 0.0182±0.0013 | 0.9954±0.0015 |

## 5.2 Results analysis of ablation study

To rigorously evaluate the contribution of each modeling component and substantiate the design rationale of the proposed framework for tool wear prediction, a comprehensive ablation study is conducted in this section. The study compares three modeling paradigms: (i) a physics-based model derived from prior knowledge (i.e., PHY); (ii) a data-driven model trained exclusively on sensor data without incorporating physical priors (i.e., GBRT); and (iii) the proposed PIGBRT model, which integrates physical prior knowledge with data-driven learning capabilities. Specifically, the predictive performance of tool wear is illustrated in **Fig. 7**, while the corresponding quantitative evaluation metrics are summarized in **Table 4**.

As depicted in **Fig. 7**, although the physics-based model (i.e., PHY) offers high interpretability, its predictive accuracy of tool wear is limited by the idealized assumptions underpinning its formulation. In contrast, the data-driven GBRT model demonstrates superior predictive performance of tool wear within the training distribution by leveraging information extracted from multi-sensor data. But it suffers from poor extrapolation when confronted with unseen operational conditions, which in turn restricts the model's generalization ability. Critically, the proposed physics-data fusion model (i.e., PIGBRT) outperforms both baseline approaches (i.e., PHY and GBRT) across all evaluation scenarios.

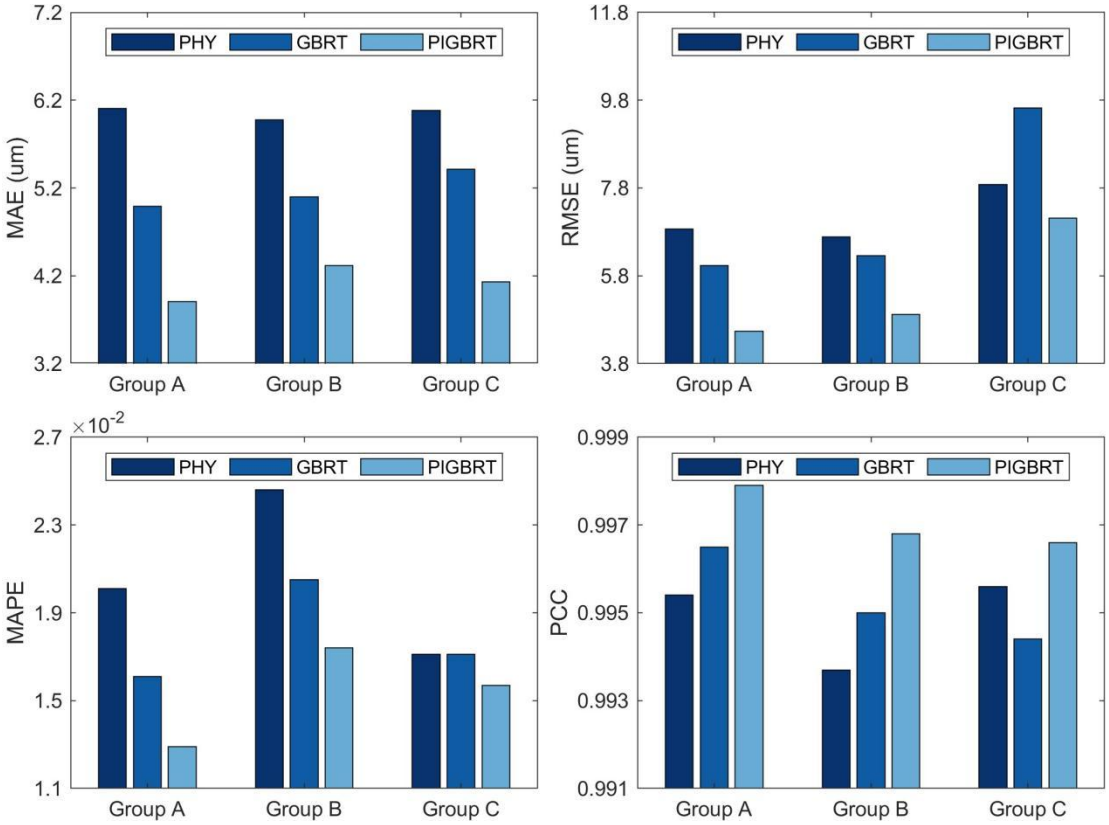



**Fig. 7** Evaluation results of tool wear prediction under different ablation models

**Table 4** Overall results of tool wear prediction under different ablation models

| Method | MAE (um) | RMSE (um) | MAPE | PCC |
|---|---|---|---|---|
| PHY | 6.0815±0.0778 | 7.1435±0.5216 | 0.0206±0.0030 | 0.9949±0.0009 |
| GBRT | 5.4124±0.5230 | 7.3046±1.6389 | 0.0179±0.0019 | 0.9946±0.0017 |
| **PIGBRT** | **4.1313±0.1708** | **5.5214±1.1368** | **0.0153±0.0018** | **0.9972±0.0006** |

As summarized in **Table 4**, the PIGBRT model achieves a reduction in MAE of 23.67% compared to the purely data-driven GBRT model and 32.07% relative to the physics-based PHY model, demonstrating substantial gains in predictive accuracy. More importantly, the PIGBRT model preserves physical consistency while effectively adapting to empirical patterns in multi-sensor data. This dual fidelity to both physical laws and observational evidence not only enhances prediction accuracy but also improves robustness under distributional shifts and increases model interpretability. Consequently, the PIGBRT model achieves a principled balance between mechanistic rigor and data-driven adaptability.

## 5.3 Comparison with other advanced models

To further verify the advancement of the proposed PIGBRT model, a comprehensive and systematic evaluation was conducted across two modeling paradigms. First, within the purely data-driven paradigm, the baseline GBRT model of the proposed method was evaluated against other advanced models, including multiple linear regression (MLR), artificial neural network

(ANN), support vector regression (SVR), random forest regression (RFR), and Gaussian process regression (GPR). This comparison is designed to evaluate the relative efficacy of GBRT as a high-performance data-driven foundation prior to the introduction of physical constraints, thereby establishing a baseline for assessing the added value of physics integration.

Additionally, to assess the contribution of the proposed physics-data fusion strategy, each of the above comparison models was extended into its corresponding physics-informed counterparts by incorporating the same governing physical constraints used in the proposed PIGBRT model. The resulting models, denoted as PIMLR, PIANN, PISVR, PIRFR, and PIGPR, were then directly compared with the PIGBRT model under identical experimental conditions. This two-tiered comparative design not only isolates the contribution of the base learning architecture but also rigorously evaluates the effectiveness of the specific fusion mechanism in harmonizing empirical learning with mechanistic principles. Correspondingly, the predictive performance of tool wear is illustrated in **Fig. 8** and **Fig. 9**, where light and dark colors represent the prediction errors of purely data-driven models and physics-data fusion models, respectively. The overall evaluation metrics are quantitatively summarized in **Table 5**.

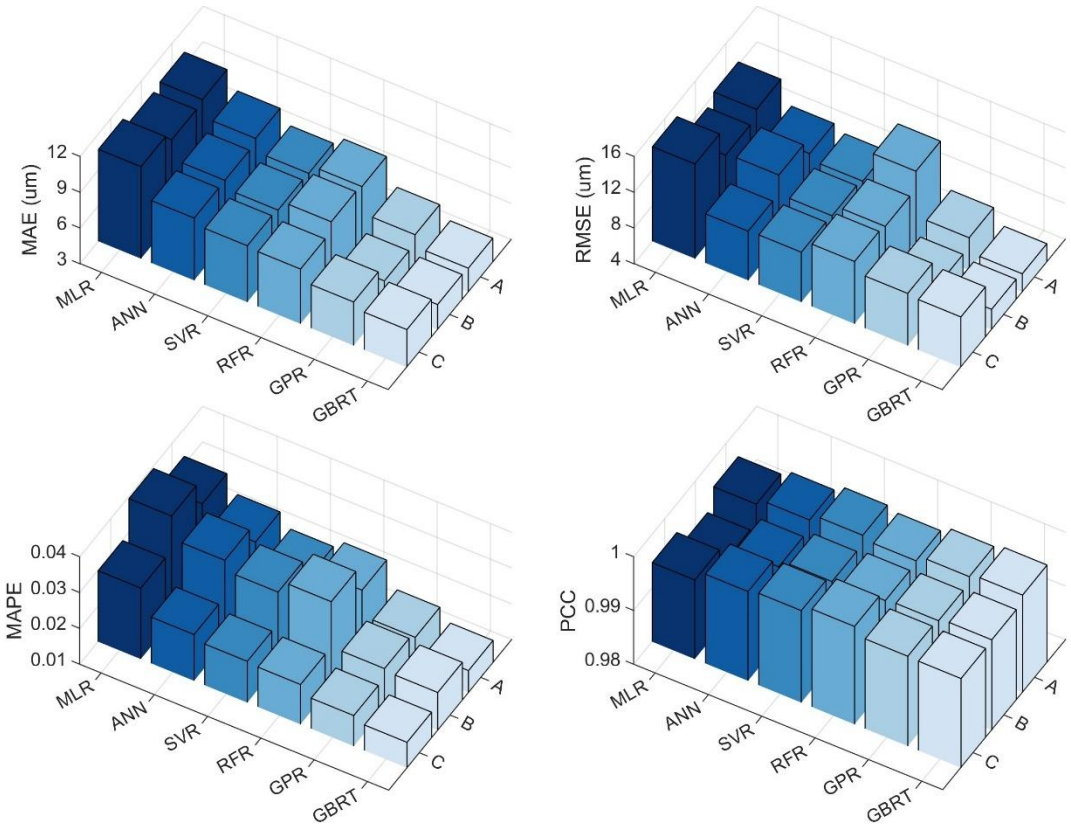



**Fig. 8** Evaluation results of tool wear prediction under different data-driven models

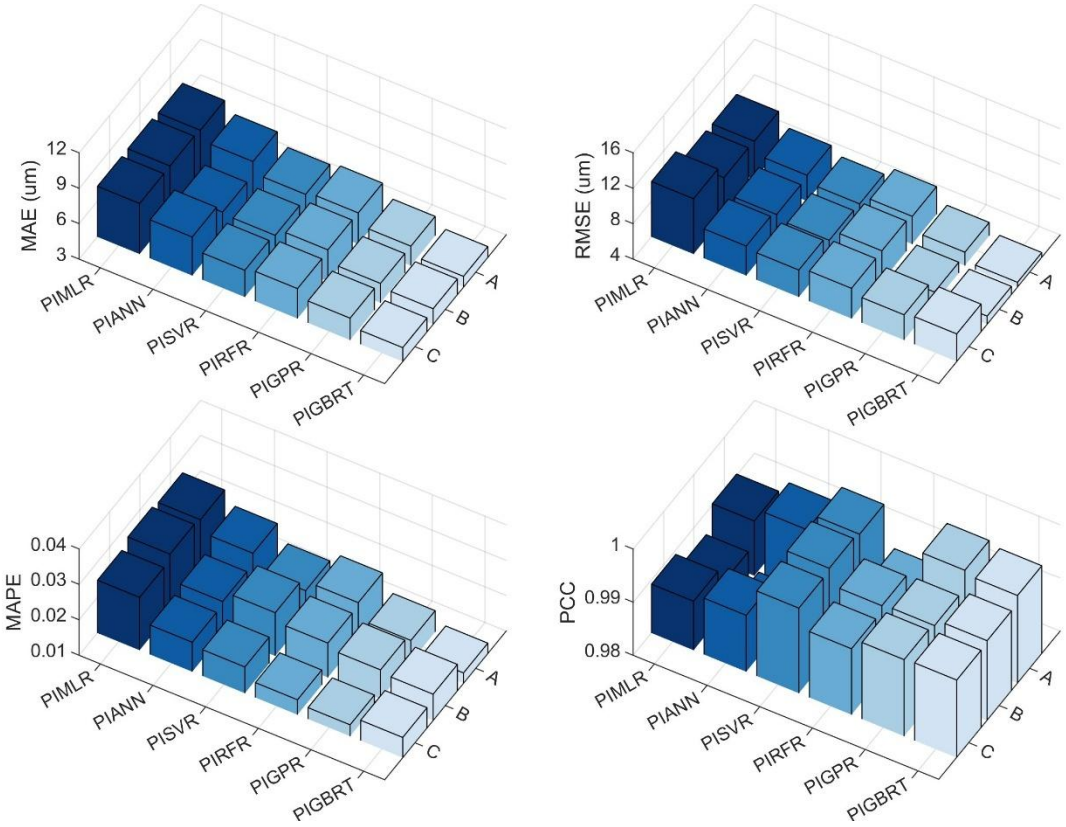



**Fig. 9** Evaluation results of tool wear prediction under different physical-informed models

As shown in **Fig. 8**, **Fig. 9**, and **Table 5**, among the purely data-driven models, GBRT consistently achieves the lowest prediction error of tool wear across all test scenarios. Specifically, GBRT reduces MAE by 46.83%, 13.07%, 17.90%, 34.99%, and 17.94% compared to MLR, ANN, SVR, RFR, and GPR, respectively. Correspondingly, it achieves reductions in RMSE of 42.64%, 12.42%, 10.30%, 43.54%, and 23.78% relative to the same set of baseline models. These results demonstrate the strong capability of GBRT to capture complex, nonlinear patterns in multi-sensor data of tool wear, even in the absence of explicit physical constraints.

**Table 5** Overall evaluation of tool wear prediction under different advanced models

| Category | Method | MAE (um) | RMSE (um) | MAPE | PCC |
|---|---|---|---|---|---|
| Data-driven | MLR | 10.178±0.3825 | 12.7351±1.3474 | 0.0338±0.0041 | 0.9892±0.0024 |
| | ANN | 8.3445±0.1720 | 10.2708±0.9220 | 0.0279±0.0042 | 0.9906±0.0056 |
| | SVR | 7.5844±0.0987 | 9.3694±0.1977 | 0.0253±0.0037 | 0.9960±0.0003 |
| | RFR | 8.0598±0.3448 | 11.4257±0.8655 | 0.0271±0.0052 | 0.9924±0.0029 |
| | GPR | 5.9606±0.5661 | 8.0217±1.3377 | 0.0196±0.0011 | 0.9943±0.0014 |
| | **GBRT** | **5.4124±0.5230** | **7.3046±1.6389** | **0.0179±0.0019** | **0.9946±0.0017** |
| Physical-informed | PIMLR | 7.2542±0.0558 | 8.9956±0.8821 | 0.0257±0.0007 | 0.9937±0.0016 |
| | PIANN | 5.9409±0.5124 | 6.9739±0.2638 | 0.0200±0.0017 | 0.9956±0.0010 |
| | PISVR | 5.2399±0.1311 | 6.4508±0.4231 | 0.0192±0.0022 | 0.9966±0.0008 |
| | PIRFR | 5.6105±0.0570 | 7.3247±0.1399 | 0.0182±0.0028 | 0.9969±0.0010 |
| | PIGPR | 4.7103±0.0990 | 5.8532±0.7097 | 0.0158±0.0020 | 0.9968±0.0005 |
| | **PIGBRT** | **4.1313±0.1708** | **5.5214±1.1368** | **0.0153±0.0018** | **0.9972±0.0006** |

As shown in **Fig. 9**, with the incorporation of physics-informed constraints, all fusion models exhibit improved prediction accuracy and enhanced generalization, particularly under extrapolative conditions involving unseen operational parameters. Specifically, compared to their purely data-driven counterparts, PIMLR, PIANN, PISVR, PIRFR, PIGPR, and PIGBRT achieve MAE reductions of 28.73%, 28.81%, 30.91%, 30.39%, 20.98%, and 23.67%, respectively. Corresponding RMSE reductions are 29.36%, 32.10%, 31.15%, 35.89%, 27.03%, and 24.41%, respectively. These results further demonstrate that embedding physical knowledge into data-driven architectures effectively regularizes model behavior beyond the training distribution.

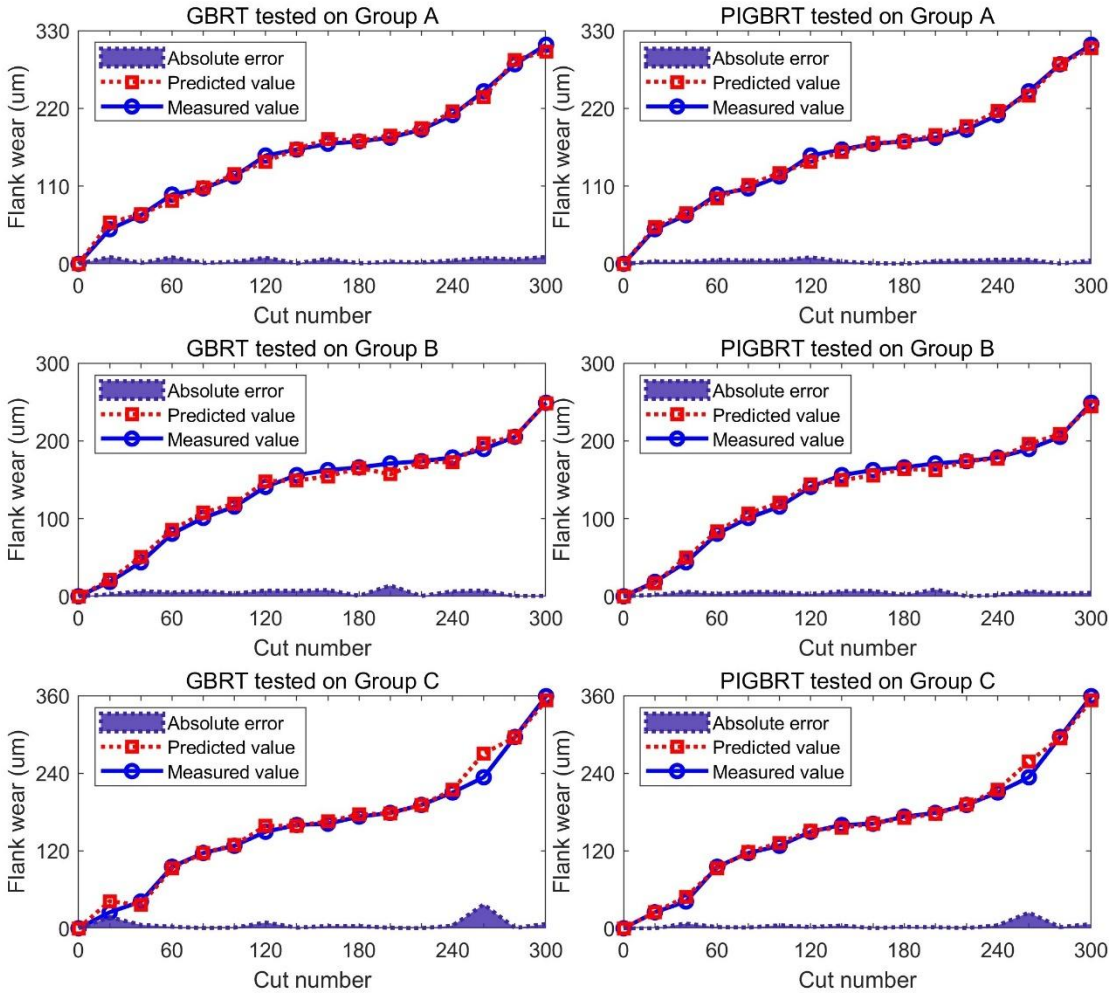



**Fig. 10** Prediction results of tool flank wear under different methods

While all hybrid models benefit from physics integration, the extent of improvement varies depending on the base learner's flexibility and inductive bias. Among the compared models, PIGBRT exhibits the most substantial performance gains, reducing MAE by 43.05%, 30.46%, 21.16, 26.36%, and 12.29% relative to PIMLR, PIANN, PISVR, PIRFR, PIGPR, respectively. These improvements are primarily attributable to the synergistic integration of gradient boosting with a tailored physics-informed loss function. This integration mitigates reliance on spurious statistical correlations and promotes solutions that are consistent with underlying

physical wear mechanisms. This variation underscores the critical interplay between model architecture and physical priors in physics-informed hybrid modeling.

Finally, the prediction results of tool wear are displayed in **Fig. 10**. It can be found that the tool wear predicted by the proposed PIGBRT model exhibits significantly closer agreement with the experimentally measured data than those generated by the benchmark models. This enhanced alignment not only visually demonstrates the superior predictive accuracy of the PIGBRT model but also indicates its capability to effectively capture the underlying tool wear dynamics as reflected in the experimental measurements. The consistency between prediction and measurement further validates the efficacy of integrating gradient boosting with physics-informed regularization for high-fidelity tool wear modeling.

# 6 Conclusion

This study proposed a physics-informed gradient boosting regression tree (PIGBRT) framework for milling flank-wear prediction. By incorporating a generalized tool-wear degradation law as a physics-based regularization term into the conventional GBRT objective, the proposed approach jointly accounts for data fidelity and physical consistency. The resulting objective function combines prediction loss with a physics-based penalty weighted by a tunable coefficient, enabling the model to retain the nonlinear learning capability of GBRT while constraining physically implausible wear evolution.

Experimental results demonstrate that PI-GBRT consistently outperforms purely data-driven, mechanistic, and hybrid models, particularly under limited-data conditions. Incorporating physical knowledge enhances predictive robustness, cross-condition generalization, and physical consistency while preserving GBRT's capability to capture nonlinear relationships and feature interactions. Moreover, the physics-based regularization mitigates overfitting without requiring a complete mechanistic wear model.

Despite these advantages, the current framework relies on a generalized degradation law and determines the physics-loss weighting factor through cross-validation. Future research will focus on incorporating multiple and process-specific physical constraints, developing adaptive weighting and online-learning strategies for real-time tool-wear monitoring, and validating the framework across diverse tool–workpiece combinations and machining conditions. Integrating uncertainty quantification will also be explored to provide reliable confidence bounds for physics-informed wear prediction and support risk-aware industrial decision-making.

# Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Data availability

Data will be made available on reasonable request.

## Acknowledge

This work was financially supported in part by the National Natural Science Foundation of China (Grant No. 52575520).

## References

[1] Wang R, Song Q, Peng Y, Qin J, Liu Z, Liu Z (2024) Toward digital twins for high-performance manufacturing: tool wear monitoring in high-speed milling of thin-walled parts using domain knowledge. Robot Comput Integr Manuf 88:102723. https://doi.org/10.1016/j.rcim.2024.102723

[2] Huang Z, Shao J, Zhu J, Zhang W, Li X (2024) Tool wear condition monitoring across machining processes based on feature transfer by deep adversarial domain confusion network. J Intell Manuf 35:1079-1105. https://doi.org/10.1007/s10845-023-02088-2

[3] Shah M, Vakharia V, Chaudhari R, Vora J, Pimenov DY, Giasin K (2022) Tool wear prediction in face milling of stainless steel using singular generative adversarial network and LSTM deep learning models. Int J Adv Manuf Technol 121:723-736. https://doi.org/10.1007/s00170-022-09356-0

[4] Dutta S, Pal SK, Mukhopadhyay S, Sen R (2013) Application of digital image processing in tool condition monitoring: a review. CIRP J Manuf Sci Technol 6:212-232. https://doi.org/10.1016/j.cirpj.2013.02.005

[5] Astakhov VP (2004) The assessment of cutting tool wear. Int J Mach Tools Manuf 44:637–647. https://doi.org/10.1016/j.ijmachtools.2003.11.006

[6] Liang X, Liu Z, Wang B (2019) State-of-the-art of surface integrity induced by tool wear effects in machining process of titanium and nickel alloys: a review. Measurement 132:150-181. https://doi.org/10.1016/j.measurement.2018.09.045

[7] Feng C, Ye H, Li W, Huang Z, Li M, Chen J, Xu X (2026) A hybrid model for tool wear monitoring via physics-based and data-driven utilizing unscented Kalman filter. Adv Eng Inform 72:104448. https://doi.org/10.1016/j.aei.2026.104448

[8] Han Z, Luo W, Liang X, Yang C, Gui W, Wang X, Zhang M (2026) Online dynamic parameter calibration of a physics-driven hybrid model for tool wear prediction. J Manuf Syst 86:676-693. https://doi.org/10.1016/j.jmsy.2026.04.012

[9] Li B (2012) A review of tool wear estimation using theoretical analysis and numerical simulation technologies. Int J Refract Met Hard Mater 35:143-151. https://doi.org/10.1016/j.ijrmhm.2012.05.006

[10] Ma Z, Zhao M, Dai X, Chen Y (2023) A hybrid-driven probabilistic state space model for

tool wear monitoring. Mech Syst Signal Process 200:110599. https://doi.org/10.1016/j.ymssp.2023.110599

[11] Li Y, Wang J, Huang Z, Gao RX (2022) Physics-informed meta learning for machining tool wear prediction. J Manuf Syst 62:17-27. https://doi.org/10.1016/j.jmsy.2021.10.013

[12] Wang J, Li Y, Zhao R, Gao RX (2020) Physics guided neural network for machining tool wear prediction. J Manuf Syst 57:298-310. https://doi.org/10.1016/j.jmsy.2020.09.005

[13] Fang X, Song Q, Wang X, Li Z, Ma H, Liu Z (2025) An intelligent tool wear monitoring model based on knowledge-data-driven physical-informed neural network for digital twin milling. Mech Syst Signal Process 232:112736. https://doi.org/10.1016/j.ymssp.2025.112736

[14] Yang Q, Pattipati KR, Awasthi U, Bollas GM (2022) Hybrid data-driven and model-informed online tool wear detection in milling machines. J Manuf Syst 63:329-343. https://doi.org/10.1016/j.jmsy.2022.04.001

[15] Zhang H, Jiang S, Gao D, Sun Y, Bai W (2024) A review of physics-based, data-driven, and hybrid models for tool wear monitoring. Machines 12:833. https://doi.org/10.3390/machines12120833

[16] Mohamed A, Hassan M, M'Saoubi R, Attia H (2022) Tool condition monitoring for high-performance machining systems–a review. Sensors 22:2206. https://doi.org/10.3390/s22062206

[17] Zhang S, Dong Z, Liu W, Zhang Y, Kang R, Bao Y (2026) Mechanistic modeling of machined surface topography evolution considering tool wear during cutting of tungsten heavy alloy. Tribol Int 221:111997. https://doi.org/10.1016/j.triboint.2026.111997

[18] Methon G, Courbon C, Auzenat F, M'Saoubi R, Girinon M, Rech J (2025) Efficient predictive simulation of tool wear in face milling using an advanced 3D numerical approach. Wear 570:205946. https://doi.org/10.1016/j.wear.2025.205946

[19] Jamshidi H, Bagherzadeh A, Budak E, Ghadbeigi H (2024) A predictive wear model in grinding using single-layer electroplated cBN tools. J Manuf Process 127:251-258. https://doi.org/10.1016/j.jmapro.2024.07.121

[20] García-Martínez E, Miguel V, Martínez-Martínez A, Ayllón J (2023) A new model to predict the tool life in turning of titanium aluminides. Int J Adv Manuf Technol 125:5757-5770. https://doi.org/10.1007/s00170-023-11090-0

[21] Zhuang K, Li M, Lin F, Hu C, Weng J, Gao Z, Wu C (2022) Crater wear prediction in turning Ti6Al4V considering cutting temperature effect. Int J Adv Manuf Technol 121:6763-6781. https://doi.org/10.1007/s00170-022-09773-1

[22] Bai Y, Wang F, Fu R, Hao J, Si L, Zhang B (2021) A semi-analytical model for predicting tool wear progression in drilling CFRP. Wear 486-487:204119. https://doi.org/10.1016/j.wear.2021.204119

[23] Ducroux E, Fromentin G, Viprey F, Prat D, D'Acunto A (2021) New mechanistic cutting force model for milling additive manufactured Inconel 718 considering effects of tool wear evolution and actual tool geometry. J Manuf Process 64:67-80. https://doi.org/10.1016/j.jmapro.2020.12.042

[24] Zhu K, Zhang Y (2019) A generic tool wear model and its application to force modeling and wear monitoring in high speed milling. Mech Syst Signal Process 115:147-161. https://doi.org/10.1016/j.ymssp.2018.05.045

[25] Lu Y, Yue C, Liu X, Liang SY, Dong S, Dou X, Yue Z (2026) Research on an online monitoring system for tool wear status based on digital twin. Mech Syst Signal Process 252:114249. https://doi.org/10.1016/j.ymssp.2026.114249

[26] Mundra S, Yadav M, Jadhav S (2026) Remaining useful life prediction of face milling cutter via vibration signatures and CNN-BiLSTM architecture. J Vib Eng Technol 14:150. https://doi.org/10.1007/s42417-026-02409-8

[27] Huang Z, Li W, Zhu J, Wang L (2024) Cross-domain tool wear condition monitoring via residual attention hybrid adaptation network. J Manuf Syst 72:406-423. https://doi.org/10.1016/j.jmsy.2023.12.003

[28] Shao J, Fang D, Huang Z, Zheng Z, Guo W, Zhu J (2026) A dual-branch transformer-BiLSTM network with error compensation for tool wear prediction. J Intell Manuf. https://doi.org/10.1007/s10845-026-02914-3

[29] Chen Z, Zhang J, Qin F, Tao G, Li D, Cao H (2025) Spatial-temporal dual-stream fusion network with embedded knowledge: a process-generalized framework for tool wear monitoring. J Manuf Syst 81:189-207. https://doi.org/10.1016/j.jmsy.2025.05.014

[30] Hou W, Wang J, Wang L, Zhang S (2025) Novel tool wear prediction method based on multimodal information fusion and deep subdomain adaptation. Mech Syst Signal Process 224:112128. https://doi.org/10.1016/j.ymssp.2024.112128

[31] Adishesha P, Mathew J, Lawrence KD (2026) Intelligent tool wear prognosis and diagnosis: a cloud analytics approach based on measurement of machining power data. Proc Inst Mech Eng B J Eng Manuf. https://doi.org/10.1177/09544054261427875

[32] Jones T, Cao Y (2025) Tool wear prediction based on multisensor data fusion and machine learning. Int J Adv Manuf Technol 137:5213–5225. https://doi.org/10.1007/s00170-025-15472-4

[33] Zhang P, Zhao H, Sun C, Li A, Zhang H, Sugita N, Shu L (2026) Tool wear monitoring and multi-step forecasting in deep hole boring based on physics-informed neural network and ARIMA model. Mech Syst Signal Process 249:114078. https://doi.org/10.1016/j.ymssp.2026.114078

[34] Wang Z, Zhou G, Zhang C, Xiao Q, Zio E (2026) Physics-informed Gaussian process regression for tool wear monitoring with uncertainty quantification. Reliab Eng Syst Saf

268:111984. https://doi.org/10.1016/j.ress.2025.111984

[35] Liu D, Liu Z, Zhao J, Wang B, Song Q, Wang P (2026) Physics-informed data-driven framework for tool wear monitoring with limited samples: mechanism-embedded learning. J Manuf Syst 86:277-299.

[36] Kim G, Yang SM, Jeon S, Park S, Choi JG, Park HW, Lim S (2026) Robust tool wear prediction under novel operating conditions via physics-guided unsupervised domain adaptation. Adv Eng Inform 69:103883. https://doi.org/10.1016/j.aei.2025.103883

[37] Guo M, Gong D, Huang Z, Guo W, Jiang X, Lv D (2026) Physics-informed deep adaptive denoising and multi-scale feature fusion method for tool wear prediction. J Intell Manuf. https://doi.org/10.1007/s10845-026-02837-z

[38] Chen Z, Zhang J, Qin F, Tao G, Li D, Cao H (2025) PIDGGCN: a novel physics-informed deep learning framework for tool wear monitoring. Adv Eng Inform 68:103790. https://doi.org/10.1016/j.aei.2025.103790

[39] Pashmforoush F, Ebrahimi Araghizad A, Budak E (2025) Tool wear prediction in milling process using physics-informed machine learning and thermo-mechanical force model with monitoring applications. J Manuf Syst 82:1192-1212. https://doi.org/10.1016/j.jmsy.2025.08.014

[40] Sun M, Wang X, Guo K, Huang X, Sun J, Li D, Huang T (2024) Tool wear monitoring based on physics-informed Gaussian process regression. J Manuf Syst 77:40-61. https://doi.org/10.1016/j.jmsy.2024.09.001

[41] Bagga PJ, Patel KM, Makhesana MA, Şirin Ş, Khanna N, Krolczyk GM, Pala AD, Chauhan KC (2023) Machine vision-based gradient-boosted tree and support vector regression for tool life prediction in turning. Int J Adv Manuf Technol 126:471-485. https://doi.org/10.1007/s00170-023-11137-2

[42] Domínguez-Monferrer C, Ramajo-Ballester A, Armingol JM, Cantero JL (2024) Spot-checking machine learning algorithms for tool wear monitoring in automatic drilling operations in CFRP/Ti6Al4V/Al stacks in the aircraft industry. J Manuf Syst 77:96-111. https://doi.org/10.1016/j.jmsy.2024.08.023

[43] Xu N, Wang Z, Dai Y, Li Q, Zhu W, Wang R, Finkelman RB (2023) Prediction of higher heating value of coal based on gradient boosting regression tree model. Int J Coal Geol 274:104293. https://doi.org/10.1016/j.coal.2023.104293