Physics-Informed Neural Networks With Weighted Losses by Uncertainty Evaluation for Accurate and Stable Prediction of Manufacturing Systems

Jiaqi Hua, Yingguang Li, *Member, IEEE*, Changqing Liu, Peng Wan, and Xu Liu

***Abstract*— The state prediction of key components in manufacturing systems tends to be risk-sensitive tasks, where prediction accuracy and stability are the two key indicators. The physics-informed neural networks (PINNs), which integrate the advantages of both data-driven models and physics models, are deemed as an effective approach and research trends for stable prediction; however, the potential advantages of PINN are limited for the situations with inaccurate physics models or noisy data, where the balancing of the weights of the data-driven model and physics model is very important for improving the performance of PINN, and it is also a challenge urgently to be addressed. This article proposed a kind of PINN with weighted losses (PNNN-WLs) by uncertainty evaluation for accurate and stable prediction of manufacturing systems, where a novel weight allocation strategy based on uncertainty evaluation by quantifying the variance of prediction errors is proposed, and an improved PINN framework is established for accurate and stable prediction. The proposed approach is verified with open datasets on tool wear prediction, and experimental results show that the prediction accuracy and stability could be obviously improved over existing methods.**

***Index Terms*— Intelligent manufacturing, physics-informed neural network (PINN), stable prediction, uncertainty evaluation.**

I. INTRODUCTION

THE extreme performance and scale demand of modern products is pursuing ultimate technologies in materials, part complexity, and manufacturing precision, which has imposed severe challenges to the existing manufacturing theory and technology. With the transformation and upgrading of the manufacturing industry, intelligent manufacturing processes have been put on the agenda. The monitoring and prediction of manufacturing systems is the key to intelligent manufacturing, especially in aerospace manufacturing [1].

Manuscript received 26 September 2021; revised 12 July 2022 and 20 September 2022; accepted 23 November 2022. This work was supported in part by the National Science Fund of China for Distinguished Young Scholars under Grant 51925505 and in part by the National Natural Science Foundation of China under Grant 52175467. *(Corresponding author: Yingguang Li.)*

Jiaqi Hua, Yingguang Li, Changqing Liu, and Peng Wan are with the College of Mechanical and Electrical Engineering and the National Key Laboratory of Science and Technology on Helicopter Transmission, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China (e-mail: liyingguang@nuaa.edu.cn).

Xu Liu is with the School of Mechanical and Power Engineering, Nanjing Tech University, Nanjing 210000, China.

Color versions of one or more figures in this article are available at https://doi.org/10.1109/TNNLS.2023.3247163.

Digital Object Identifier 10.1109/TNNLS.2023.3247163

Accurate prediction of key components in manufacturing systems can effectively improve manufacturing quality and efficiency and reduce the manufacturing cost [2], [3]. In addition, the predictions in manufacturing systems are always risk-sensitive tasks, especially for machining high-value-added parts in aerospace. Not only the prediction accuracy, i.e., the mean value of prediction errors, is important, but also the stability of prediction error is a key indicator for the prediction model. Unstable prediction errors will increase the risks of manufacturing problems mentioned above. Tool wear prediction in computer numerical control (CNC) machining is a typical prediction task in manufacturing systems [4], [5] where the prediction accuracy and stability are equally crucial. Take an example of tool wear prediction; a well-trained data-driven prediction model could achieve the prediction accuracy of 0.04 mm (the mean value of prediction errors of test samples), which satisfies the industrial requirement nominally. When we look into the prediction results, we, however, find that the prediction errors are unstable, where 33% of prediction errors of all the test samples are more than 0.05 mm, and 8.5% of samples are even more than 0.1 mm, resulting in that the prediction model could not be applied in the real industry. Unstable prediction errors will increase the risks in the machining process or even damage the machining quality. In conclusion, it is extremely important to realize accurate and stable predictions in manufacturing systems, as very important decisions are always made based on the prediction results in real time.

Traditional physics prediction models, which reflect the common state patterns, could be effective under certain assumptions or simplifications for simple situations, while they are not able to achieve a high prediction accuracy because of the complex manufacturing processes. Some advanced data-driven models could learn from data for accurate prediction but could not ensure the stability of prediction error due to the changes in data distribution under varying working conditions. The models that integrate data and physics information, i.e., physics-informed neural networks (PINNs), have the advantages of both data-driven method and physics modeling method and have the potential to avoid their limitations [6]; however, the potential advantages of PINN are limited for the situations with inaccurate physics models or noisy data, where the balancing of the weights of the data-driven model and physics model is very important for improving the performance of PINN, and it is also a challenge urgently need to be addressed. Therefore, it calls for a new integrating method for accurate and stable prediction, and we propose a kind of PINN with weighted losses (PINN-WLs) by uncertainty evaluation for accurate and stable prediction of manufacturing systems, which is verified in the case of tool wear prediction. Compared with other commonly used prediction methods, the developed model has several distinctive merits.

1) The information provided by the data-driven model and the physics model could be effectively used to improve the prediction accuracy and stability by using the weight allocation strategy based on uncertainty evaluation.
2) This work makes a contribution to the model integrating researches by raising an improved PINN framework for accurate and stable prediction of manufacturing systems.
3) The basic idea of this research work has the potential to be applied to tool life monitoring, part quality assurance, production automation, and other industrial applications.

II. Literature Review

Existing prediction methods for manufacturing systems can be generally classified into three groups: physics model-driven methods, data-driven methods, and the data-driven physics-integrated methods. This article will review the literatures from the above three aspects.

A. Physics Model-Driven Methods

Researchers have observed the physical and chemical changes of various components in the manufacturing process through experiments, or derived mathematical formulas to established physics models reflecting the nature of the manufacturing process in fault diagnosis or state prediction. Researches have been carried out for state prediction of manufacturing systems, such as machining chatter prediction [7] and the parameter prediction of the tool-tip [8]. Especially in tool wear prediction, some researchers have proposed a variety of empirical formulas describing tool wear by counting the changes in tool wear over time [9], [10], [11].

With the physics model-driven method, a prediction model is established based on experiences or some manufacturing mechanism; however, manufacturing is a complex physical and chemical process, while physics model can only consider some specific processes such as friction and deformation. During the considered specific processes, the manufacturing process is influenced by many factors, and the physics model can only consider some specific factors, such as cutting temperature, cutting force, and so on; even in this situation, the process can only be modeled by significant simplification and assumptions, while the solution searching procedure is still complex, and only approximation solution can be obtained. Accordingly, physics model-driven methods are difficult to realize accurate and stable predictions for complex conditions such as varying working conditions due to the limitations aforementioned above.

B. Data-Driven Methods

With the wide application of sensors and the greatly improved computer performance in recent years, the data-driven method has gradually become an effective way for the prediction of manufacturing systems [12].

By employing machine learning or other types of data processing techniques, a data-driven method predicts the states of manufacturing systems by establishing a complex mapping relationship between the monitored signals (such as the cutting force, the vibration, the spindle current, etc.) and the expected output. Data-driven methods have also been widely used in fault diagnosis of mechanical devices, such as gearboxes [13], bearings [14], [15], and motors [16], or in other state prediction of manufacturing systems, such as chatter [17], [18] and cutting force [19], [20]. In tool wear prediction, Fu et al. [21] proposed an adaptive feature extraction method for spindle vibration during the cutting process based on deep belief networks (DBNs), and accurate classification of tool wear stages is realized. Huang et al. [22] proposed an adaptive feature extraction approach based on a convolution neural network (CNN) for multisensor signals by reshaping the time series of raw monitoring signals, and accurate prediction of tool wear is realized. Li et al. [23] proposed a statistical analysis method through the entropy weight-gray correlation analysis to extract and select features in the time domain, frequency domain, and time-frequency domain of monitoring signals. Furthermore, the meta-learning algorithm is used to establish the relationship between features and tool wear, and an accurate prediction of tool wear is realized.

Although data-driven methods are effective means to achieve higher prediction accuracy, existing methods still suffer from two major problems: 1) only working conditions with small variations can be considered, i.e., the data are under the hypothesis of independent identical distribution, thus unsuitable under cross working conditions where data do not satisfy identical distribution and 2) a large number of labeled data are needed, which are difficult and expensive to obtain during manufacturing processes. Therefore, it is still difficult for data-driven methods to realize accurate and stable prediction under real manufacturing circumstances.

C. Data-Physics Integrated Methods

Data-physics integrated method, which integrates data and physics information, is deemed as an effective approach and research trends for accurate and stable prediction in recent years. It has the advantages of both the data-driven method and physics modeling method and the potential to avoid their limitations. Lutter et al. [24] predefined the robot dynamics equation in general form based on Lagrangian mechanics in the network structure and directly learned the physical parameters of the model in an end-to-end form. For the problem of identifying Hamiltonian systems, Jin et al. [25] maintained the symplectic structure of the basic Hamiltonian system by designing the network structure.

In the machine learning area, there exists a general paradigm of data-physics integration, where physics knowledge is used as regularization items to constrain the solution space of the machine learning model. Raissi et al. [6] proposed a deep learning framework, called PINNs, to solve the forward and inverse problems involving nonlinear partial differential equations. In this framework, the solution of the partial differential equation is expressed as a neural network, and the partial differential equation and its initial conditions and boundary conditions are constructed as parts of the loss function of the neural network, thereby constraining the search space of neural network parameters.

The data-physics integrated methods have also been applied in various areas [26], including manufacturing systems. Postel et al. [27] used the simulation data generated by the cutting dynamics models to pretrain the machine learning model and then used the experimental data to update the trained model to predict the stability lobe diagram. In order to solve the problem of inaccurate parameters in the chatter stability analysis model, Chen et al. [28] introduced Bayesian inference into the chatter stability analysis process, leveraging experimental data to infer the distribution of model parameters for computing a reliable probabilistic stability lobe diagram. A novel physics-informed machine learning algorithm that combined a multiscale mechanistic model with random forests was proposed to predict material removal rate in chemical mechanical planarization [29]. Huang et al. [30] propose a hybrid approach to predict product completion time by combining machine learning techniques and an analytical system model. In tool wear prediction, Wang et al. [31] proposed a regression model to obtain the final tool wear prediction value by concatenating the prediction value of the physics model and the prediction value of the data-driven model. Hanachi et al. [32] proposed a tool wear prediction method by combining the adaptive neuro-fuzzy inference system (ANFIS) with empirical formulas.

Existing data-physics integrated methods used in state prediction of manufacturing systems, however, only focus on improving the prediction accuracy instead of the stability; they are still suffering from the issues of unstable prediction, i.e., the potential advantages of PINN have not been fully performed, especially with inaccurate physics models or noisy data. The guiding effect of the inaccurate physics before the data-driven model is uncertain. How to effectively take advantages of the information provided by the two models is an urgent problem to be solved, and model weight balancing could be a potentially effective way.

D. Model Weight Balancing Methods

The model weight balancing methods are widely used in ensemble learning, such as boosting [33] and mixture of experts [34]. Some researchers tried to weigh the two models in data-driven physics-integrated method. In [6], the weights are designed by experience, and according to the specific problems or the validation results, it is still difficult to obtain a better solution. Other weighting approaches tried to train the weights and the model parameters simultaneously and finally converge to an optimal solution [35], [36]; the problem of this method is that the interpretability of weights is lacking, and more importantly, the objective of weighting is not to guarantee the stability of prediction error, but the prediction accuracy and the convergency.

III. Proposed PINNs-WLs by Uncertainty Evaluation

In order to bridge the gap of the existing PINN, a weight allocation approach based on the uncertainty evaluation of prediction error for model integration is introduced in this article, and a PINN-WLs is further proposed for accurate and stable prediction of manufacturing systems.

A. Problem Definition

Given the training data D_{train} and test data D_{test}, the prediction model f is trained on D_{train} and tested on D_{test}. In regression problems, which are quite common in manufacturing, the test error ε_i of the given test sample i in D_{test} is the distance between the model output and the true label, which could be measured by the absolute error (AE)

$$\varepsilon_i = |f_i(\boldsymbol{x}) - y_i| \tag{1}$$

where $f_i(\boldsymbol{x})$ and y_i are the output of the prediction model f and the true label of the test sample i, respectively. The test errors of all the test samples in D_{test} could be listed as $\boldsymbol{\varepsilon} = [\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_{\text{d}}]$, where d is the number of test samples.

The purpose of this article is to simultaneously improve the prediction accuracy and stability of the prediction model. The prediction accuracy of the prediction model is represented as the mean AE (MAE) on all test samples

$$\overline{\boldsymbol{\varepsilon}} = \mathbb{E}[|f(\boldsymbol{x}) - y|] = \frac{1}{\text{d}} \sum_{i=1}^{\text{d}} \varepsilon_i. \tag{2}$$

And the stability of the prediction model is represented as the variance of the prediction errors on all test samples

$$\mathbb{E}\big[(\varepsilon - \overline{\boldsymbol{\varepsilon}})^2\big] = \frac{1}{\text{d}} \sum_{i=1}^{\text{d}} (\varepsilon_i - \overline{\boldsymbol{\varepsilon}})^2. \tag{3}$$

It can be found from the above two formulas that the two indicators have a kind of difference but also have correlation, i.e., when $\overline{\boldsymbol{\varepsilon}}$ is approximating to 0, $\mathbb{E}[(\varepsilon - \overline{\boldsymbol{\varepsilon}})^2]$ is also approximating to 0. While in many of other cases, for a given $\overline{\boldsymbol{\varepsilon}}$, the uncertainty $\mathbb{E}[(\varepsilon - \overline{\boldsymbol{\varepsilon}})^2]$ of is also quite great, just as the example presented in the introduction of this article.

The problem defined here has not been paid much attention for the reasons that: 1) the variance of the prediction errors mainly exists for regression problems, which is quite common in the manufacturing area, while for classification problems, which is common in the popular machine learning areas of the Internet, machine vision, etc., the indicator of prediction results is category (i.e., 0 or 1) instead of the error magnitude and 2) researches in regression pay more attention to MAE instead of the variance of the prediction errors, as many issues are not risk-sensitive problems. While the distinction between MAE and the variance of errors are quite different, which has been illustrated above in formulas (2) and (3) so it is worth to be focused on risk-sensitive issues.

It is worth mentioning that there exists a series of methods for improving the robustness of data-driven models. According to [37], the expectation of the generalization error of the learning algorithm could be decomposed to the sum of bias $\mathrm{bias}^2(\boldsymbol{x})$, variance $\mathrm{var}(\boldsymbol{x})$, and noise e

$$\begin{aligned}\mathbb{E}\big[(f(\boldsymbol{x})-y)^2\big] &= \mathrm{bias}^2(\boldsymbol{x})+\mathrm{var}(\boldsymbol{x})+e\\ \mathrm{var}(\boldsymbol{x}) &= \mathbb{E}\big[(f(\boldsymbol{x})-\bar{f}(\boldsymbol{x}))^2\big]\end{aligned} \tag{4}$$

where y represents the true value, the $\mathrm{bias}(\boldsymbol{x})$ represents the distance of the predicted value and the true value. $f(\boldsymbol{x})$ represents the model output, and $\overline{f}(\boldsymbol{x})$ represents the expectation of the model output. The improvement of the expectation of the generalization error could be viewed as the improvement of the variance, which could improve the model robustness. The variance in formula (4), however, represents the anti-interference ability of the algorithm for data changes instead of the stability of prediction error mentioned in this article.

B. Strategy of Weight Allocation to the Physics Model and Data-Driven Model

Inspired by the measurement area [38], [39], where several measurement results with different measurement accuracy are allocated by different weights in order to improve the integral measurement accuracy, the weighting idea is introduced into the training process of PINN in this article. Instead of setting constant model weights with the traditional machine learning, the weights allocated in this article are correlated with the stability of the prediction error.

In order to appropriately use the information provided by the two components, respectively, the uncertainty of the physics model and the data-driven model is quantified by the variance of the prediction errors of the two components. The reciprocal of the variance of the prediction error is set as the weight, allocating to the loss function of the physics model component and the data-driven model component

$$\mathcal{L}(W,\sigma_1,\sigma_2,X)=\frac{1}{\sigma_1^2}\mathcal{L}_1(W,X)+\frac{1}{\sigma_2^2}\mathcal{L}_2(W,X) \tag{5}$$

where $\mathcal{L}_1(W,X) = \|\hat{y}-y_{\mathrm{label}}\|_2^2$ and $\mathcal{L}_2(W,X) = \|\hat{y}-y_{\mathrm{physic}}\|_2^2$ are the mean-square-error (MSE) loss function of the data-driven model component and the physics model component, σ_1^2 and σ_2^2 are the variance of the prediction error of the two models separately, $\hat{y}$ is the current prediction of the data-driven model component, y_{label} is the label of the dataset, and y_{physic} is the prediction value of the physics model component.

Essentially, the unstable prediction error is caused by the following reasons: 1) the inaccurate physics model, 2) the noisy data, and 3) the changes of data distribution caused by varying working conditions in the data-driven model. The weights allocated to each loss function aim to inhibit the impacts of the data volatility or the physics inaccuracy on each model component. The larger the error uncertainty, the smaller the weight allocated, and the more effective the inhibition of the model. It is an effective way to prevent over-fitting, as the unstable data points are restrained to the uttermost extent. As a result, the stability of the integrated model could be improved.

C. Proposed PINN-WLs

The weight allocation strategy proposed in this article is embedded into the framework of PINN for accurate and stable prediction of manufacturing systems, as shown in Fig. 1. PINN is a kind of neural network used to solve supervised learning tasks, following not only the distribution of training data samples but also the mechanism described by partial differential equations. Compared with pure data-driven neural network learning, PINN imposes physical information constraints during the training process, so it can acquire models with more generalization ability with fewer data samples.

From the perspective of mathematical function approximation theory, the neural network can be regarded as a general nonlinear function approximation, and the modeling process of the partial differential equation is also to find nonlinear functions that meet the constraints. According to the automatic differential technique widely used in deep neural networks (DNNs), the differential form constraints of partial differential equations could be combined into the loss function design of neural networks so as to realize the deep integration of the data-driven model and mechanism model. The neural network trained in this way cannot only approximate the observed data but also automatically satisfy the symmetry, invariability, conservation, and other physical properties of the partial differential equation.

Consider the general form of the differential equation parameterized by $\boldsymbol{\lambda}$ for the solution $u(\boldsymbol{x})$ with $\boldsymbol{x}=(x_1,\ldots,x_{\mathrm{d}})$ defined on a domain $\Omega\subset\mathbb{R}^{\mathrm{d}}$

$$f\left(\boldsymbol{x};\frac{\partial u}{\partial x_1},\ldots,\frac{\partial u}{\partial x_{\mathrm{d}}};\frac{\partial^2 u}{\partial x_1\partial x_1},\ldots,\frac{\partial^2 u}{\partial x_1\partial x_{\mathrm{d}}};\ldots;\boldsymbol{\lambda}\right)=0,\quad \boldsymbol{x}\in\Omega. \tag{6}$$

For time-dependent problems, time t is considered as a special component of $\boldsymbol{x}$, and Ω contains the temporal domain.

For the state prediction of the manufacturing process, a neural network $\hat{u}(\boldsymbol{x},\boldsymbol{W})$ is constructed to establish the mapping relationship between the input $\boldsymbol{x}$ (the extracted signal features of the cutting force, the vibration, the spindle current, etc., and time t) and the expected output u (for tool wear prediction, the output is tool wear $\mathrm{VB}_{\max}$). $\boldsymbol{W}$ is the set of all weight matrices in the network $\hat{u}$. Then, the loss function of the data-driven model could be set as the MSE

$$\mathrm{Loss}_D=\|\hat{u}-u\|_2^2 \tag{7}$$

where $\hat{u}$ is the expected prediction output of the neural network with the input $\boldsymbol{x}$, and u is the label of the dataset. While minimizing the Loss_D, the neural network is to approximate the observed data.

In the next step, $\hat{u}$ needs to be restricted to satisfy the differential equation of manufacturing mechanism. Taking the tool wear prediction as an example, the empirical formula provided in [32] is used here as the physics model

$$\frac{\mathrm{d}u}{\mathrm{d}t}-c_1\big(1+bt+2c_2t^2\big)\exp\big(a+bt+c_2t^2\big)=0 \tag{8}$$

where $\boldsymbol{\lambda}=(a,b,c_1,c_2)$ are the parameters of the differential equation. In order to constrain the solution space of the neural

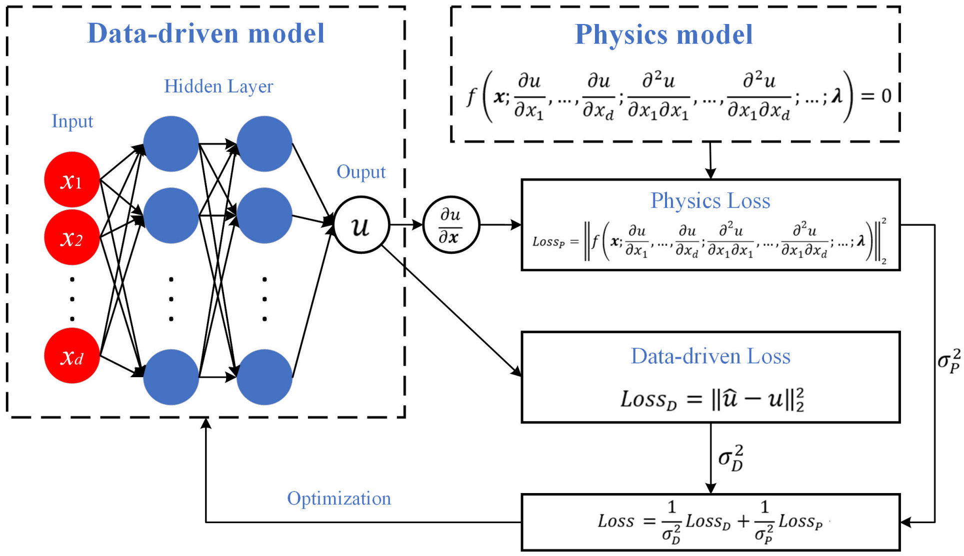


Fig. 1. Framework and general idea of the proposed PINN-WL.

network, the tool wear mechanism is integrated into the loss function of the neural network model as the regularization term

$$\begin{aligned}\text{Loss}_p &= \left\| f\left(\boldsymbol{x}; \frac{\partial u}{\partial x_1}, \ldots, \frac{\partial u}{\partial x_{\text{d}}}; \frac{\partial^2 u}{\partial x_1 \partial x_1}, \ldots, \frac{\partial^2 u}{\partial x_1 \partial x_{\text{d}}}, \ldots, \lambda\right)\right\|_2 \\ &= \left\| \frac{\partial u}{\partial t} - c_1\left(1 + bt + 2c_2t^2\right)\exp\left(a + bt + c_2t^2\right)\right\|_2^2. \quad (9)\end{aligned}$$

For each batch of training data, the variance of the prediction error of the neural network and the tool wear mechanism could be calculated as

$$\sigma^2 = \mathbb{E}\left[(\varepsilon - \bar{\varepsilon})^2\right] \quad (10)$$

where ε is the prediction error and $\bar{\varepsilon}$ is the expected error of the models. Finally, according to the weight allocation strategy, the loss function of the PINN-WL is written as

$$\text{Loss}(\boldsymbol{W}; \boldsymbol{\lambda}) = \frac{1}{\sigma_D^2}\text{Loss}_D + \frac{1}{\sigma_P^2}\text{Loss}_P. \quad (11)$$

The training procedure is to search for a good $\boldsymbol{W}$ by minimizing the loss $\text{Loss}(\boldsymbol{W}; \boldsymbol{\lambda})$ using gradient-based optimizers. For tool wear prediction, the parameters $\boldsymbol{\lambda}$ of physics model is unknown. PINN-WL optimizes the $\boldsymbol{W}$ and $\boldsymbol{\lambda}$ together: $\boldsymbol{W^*}, \boldsymbol{\lambda^*} = \text{argmin}_{\boldsymbol{W},\boldsymbol{\lambda}}\text{Loss}(\boldsymbol{W}; \boldsymbol{\lambda})$.

D. Training Strategy of PINN-WL

The PINN-WL is a general framework for the integration of the data-driven model and the physics model; any neural network or physics that satisfied the differential equation form could be integrated into the framework. The training of the PINN-WL is the procedure of searching for a good $\boldsymbol{W}$ of the neural network model by minimizing the loss $\text{Loss}(\boldsymbol{W}; \boldsymbol{\lambda})$ constrained by the data and the physics. Considering the fact that the loss is highly nonlinear and nonconvex with respect to $\boldsymbol{W}$, the loss function is usually minimized by gradient-based optimizers, such as gradient descent, Adam [40], and limited-memory BFGS (L-BFGS) [41]. The required number of iterations highly depends on the problem (e.g., the smoothness of the solution). The loss function or the residual of the mechanism model could be monitored using callback functions to check whether the network converges or not. Finally, acceleration of training can be achieved by using an adaptive activation function that may remove bad local minima [42], [43].

For further illustration of the training process of the proposed PINN-WL, the data-driven model—meta-learning model, and the physics model—the tool wear empirical formula, are integrated into the PINN-WL framework, as an example. According to the application scene of few-shot learning tasks, meta learning is used to induce the PINN-WL change law from multiple tasks, which facilitates fast model learning when faced with a new working condition with only very few samples [44], [45]. Model-agnostic meta learning (MAML) [46] is chosen in this article among several meta-learning models for the reason that it does not introduce other parameters for the learner, and the training strategy is a known optimization process (e.g., the gradient descent method) instead of building one from scratch, which makes the algorithm more efficient.

The meta learner is a PINN-WL model parameterized by Ø. The base learner is an PINN-WL model B_{φ_i} parameterized by φ_i, the parameter of the neural network $\hat{u}$. The parameter φ_i is updated along the descent gradient

$$\varphi_i = \emptyset - \alpha\nabla_{\varphi_i}L_{T_i}\left(B_{\varphi_i}\right) \quad (12)$$

where the base learning rate α is a fixed hyperparameter, and $L_{T_i}(B_{\varphi_i}) = (1/\sigma_D^2)\text{Loss}_D + (1/\sigma_P^2)\text{Loss}_P$ is the base loss function, which corresponds to $\text{Loss}(\boldsymbol{W}; \boldsymbol{\lambda})$ in Section III-C. Ø is the initial base parameter and also the metaparameter, which itself is updated by the descent gradient based on the base parameters

$$\emptyset = \emptyset - \beta\nabla_{\emptyset}\sum_{T_i \sim p(T)} L_{T_i}\left(B_{\varphi_i}\right)\left(B_{\varphi_i}\right) \quad (13)$$

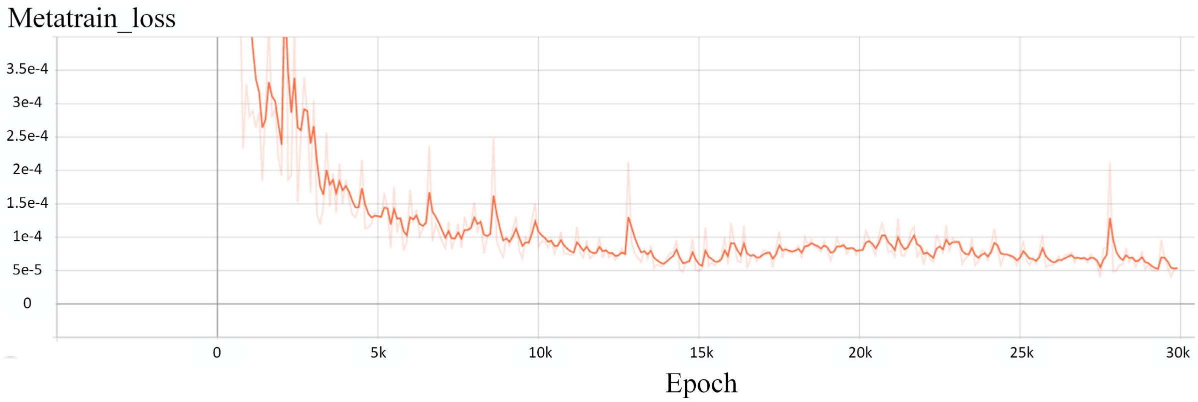


Fig. 2. Convergence curve of the training process of PINN-WL.

where the meta-learning rate β is a fixed hyper-parameter. The meta objective function is

$$\min_{\emptyset} \sum_{T_i \sim p(T)} L_{T_i}\left(B_{\varphi_i}\right) = \sum_{T_i \sim p(T)} L_{T_i}\left(B_{\emptyset - \alpha \nabla_{\varphi_i} L_{T_i}\left(B_{\varphi_i}\right)}\right). \quad (14)$$

The meta-learner optimization is to find the best initial base parameter. When faced with a new task T_{new}, the PINN-WL of T_{new} can be obtained with only a few stochastic gradient descent (SGD) steps to fine-tune $\emptyset$ to $\emptyset_{\text{new}}$, so that the new base learner can continue to perform accurate and stable tool wear prediction

$$\emptyset_{\text{new}} = \emptyset - \alpha \nabla_{\emptyset} L_{T_{\text{new}}}(u_{\emptyset}). \quad (15)$$

The base models are constructed by a fully connected neural network, which contains three hidden layers, and the number of neurons in each hidden layer is 100, 40, and 40, respectively. The model adopts the MAML meta-learning algorithm for training, in which the gradient descent optimizer is used for the base model, while the Adam optimizer is used for the meta-model. The initial learning rate is 0.001 for both, and the model iterations are set to 30 000 times. This article adopts the mini-batch gradient descent method to train the neural networks, and each batch contains 32 training samples. The initial weights of the integrating model are set to 1 equally, and after a certain number of iterations (1000 times) of model training, the variance of the prediction errors is calculated as the updated loss weights of the PINN-WL model to ensure the convergence. The convergence curve of the training process of the proposed PINN-WL is shown in Fig. 2.

E. Generalization Error Bound Estimation

The generalization error bound of the proposed method is estimated in order to guarantee stability and generalization performance. Let U be a bounded domain in $\mathbb{R}^{\mathrm{d}}$. Let $\boldsymbol{x} = (x_1, \ldots, x_{\mathrm{d}})$ be a point in $\mathbb{R}^{\mathrm{d}}$. For the function $u(x)$ fit by PINN-WL, it satisfies that for $\forall x, y \in \mathcal{R}$, the quantity below is finite [47]

$$[u]_{\alpha;U} = \sup_{x,y \in U, x \neq y} \frac{\|u(x) - u(y)\|}{\|x - y\|^{\alpha}} < \infty, \quad 0 < \alpha \leq 1 \quad (16)$$

where α is the Hölder index and $0 \leq \alpha \leq 1$. Therefore, $u(x)$ is uniformly Hölder continuous with exponent α in U. According to the existing research [47], the upper bound of the generalization error of the proposed PINN-WL can be controlled by the training samples and the training loss

$$\text{Loss}_{\text{test}}\,(u_m) \leq C\left[\text{Loss}_{\text{train}}\,(u_m) + m_r^{\frac{\alpha}{\mathrm{d}}} + m_b^{-\frac{\alpha}{\mathrm{d}-1}}\right] \quad (17)$$

where $\text{Loss}_{\text{test}}(u_m)$ is the test loss of PINN-WL, $\text{Loss}_{\text{train}}(u_m)$ is the training loss of PINN-WL, and u_m is the training data. m_r and m_b denote the number of training data in the region and on the boundary of the region, respectively. d is the dimension of the data.

According to the generalization error estimation formula (17), the solution of the PINN-WL can converge to the accurate mechanism solution as long as the number of training samples is large enough and the training error of the PINN-WL is small enough, under certain regularity assumptions.

IV. Experimental Results and Discussion

The proposed PINN-WL prediction model has been implemented and applied in tool wear prediction, a typical prediction problem in the manufacturing system. Its prediction accuracy and error stability are verified on two different tool wear datasets—the Ideahouse dataset [48] and the NASA dataset [49]; the indicator of prediction accuracy is the MAE and the indicators of prediction stability are the variance of the prediction error, the percentages of the test samples whose prediction errors are greater than 0.05 mm and greater than 0.1 mm. In order to demonstrate the improvement in prediction accuracy and stability, the above indicators of the proposed method are compared with those of the existing data-driven models, including a DNN model fine-tuned with meta learning and a common PINN model.

A. Validation in Ideahouse Dataset

The Ideahouse dataset is published by the authors' team, which is a typical dataset for varying working conditions. Milling experiments of titanium alloy part are carried out on a DMG 80P DuoBlock milling center using different cutting parameters, in which the working conditions are much

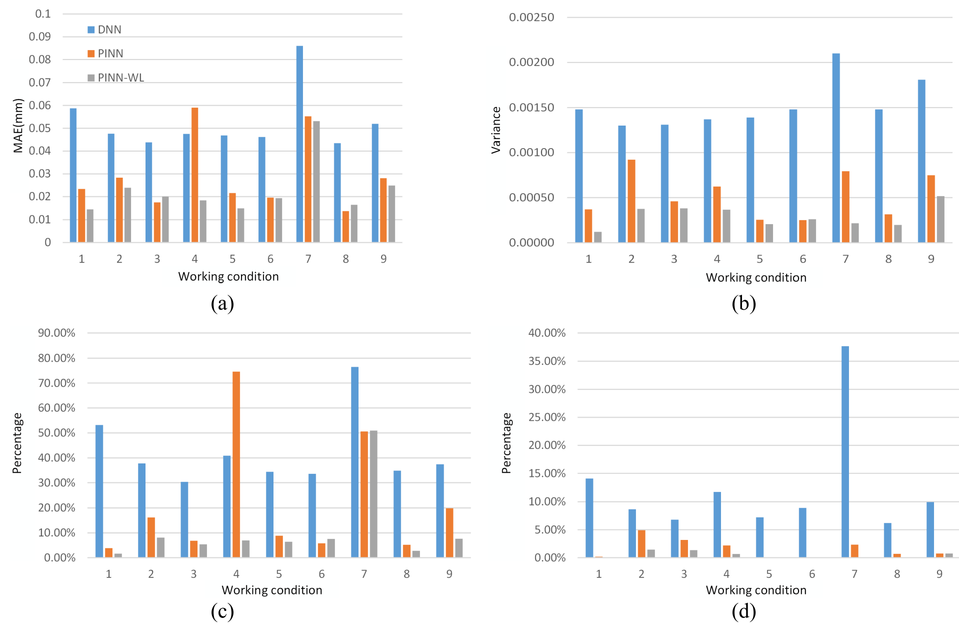


Fig. 3. Comparison results on the Ideahouse dataset (meta learning and the tool wear mechanism are integrated into the PINN-WL framework). (a) MAE. (b) Variance of prediction error. (c) Percentage of samples with an error greater than 0.05 mm. (d) Percentage of samples with an error greater than 0.1 mm.

TABLE I
WORKING CONDITIONS OF IDEAHOUSE DATASET

No.	*fz* (mm/r)	*n* (r/min)	*ap* (mm)	TD* (mm)	TM*	WM*
1	0.045	1750	2.5			
2	0.045	1800	3.0			
3	0.045	1850	3.5			
4	0.050	1750	3.0			
5	0.050	1800	3.5	12	EM_R*	TC4*
6	0.050	1850	2.5			
7	0.055	1750	3.5			
8	0.055	1800	2.5			
9	0.055	1850	3.0			

Note: TD*: Tool Diameter; TM: Tool Material; WM*: Workpiece Material; EM_R*: Carbide Endmill with Rounded Corner; TC4*: Titanium alloy

more varied. Both the straight line and the corner machining features are included in one experiment group, meaning that the working condition is complex and continuously changing.

In the first experiment, the data-driven model, meta learning, and the tool wear mechanism are integrated into the proposed framework to verify the performance of the proposed method in varying working conditions. Nine groups of machining tests are designed with different working conditions, as shown in Table I. The cross-validation test is designed, where eight groups are used for model training and the rest one group for the model test. The above process is performed with nine times; for each time, one different group is chosen for the test with the model trained by the other eight groups, and a total of nine tests are carried out. The meta learner is fine-tuned by one step with ten samples of the test data after each training process. The ranges of cutting parameters' variations: feed per tooth (fz: 0.045–0.055 mm/r), spindle speed (n: 1750–1850 r/min), and cutting depth (ap: 2.5–3 mm).

Under the above nine working conditions, the monitored signals, including vibrations, spindle current, and power, are collected, and feature extraction is performed to obtain the top 32 features which are most relevant to the tool wear. Note that the data labels in this article refer to the tool wear amount (i.e., VB_{max}, which is a certain width of the flank wear land, referring to ISO 8688-2, Part 2), which is measured using an XK-T600V[1] microscope (accuracy is calibrated as: 0.01 mm).

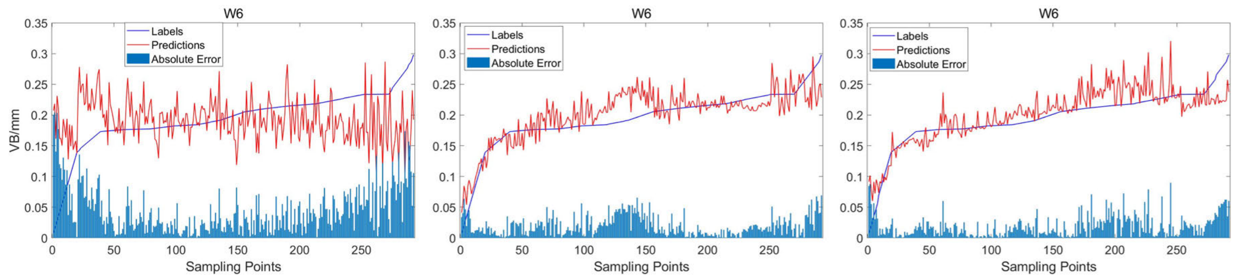
In this experiment, the prediction results of three models–data-driven model (DNN), PINN, and ours PINN-WL are compared on our data set. The MAE is used to verify the accuracy of prediction models, and three indicators—variance of error, percentage of samples with an error greater than 0.05 mm, and percentage of samples with an error greater than 0.1 mm, are used to verify the prediction stability. It is worth to be mentioned that the data-driven part of the traditional PINN is the same as the PINN-WL, and both of them are trained by a meta-learning strategy. The comparison results are shown in Fig. 3. The prediction details of each working condition are shown in Fig. 4, where the blue line represents the true labels of tool wear value, the red line represents the prediction value, and the blue bar below represents the prediction error, the left column represents the prediction details of DNN model, the middle column represents the prediction details of PINN model, and the right column represents the prediction details of the proposed model. It could be found that the proposed method can ensure prediction accuracy and error stability, where the variances of prediction error are mostly lower than the other two methods, the percentage of samples with an error greater than 0.05 mm can be controlled into 10%, and the samples with an error greater than 0.1 mm almost disappear.

[1]Trademarked.

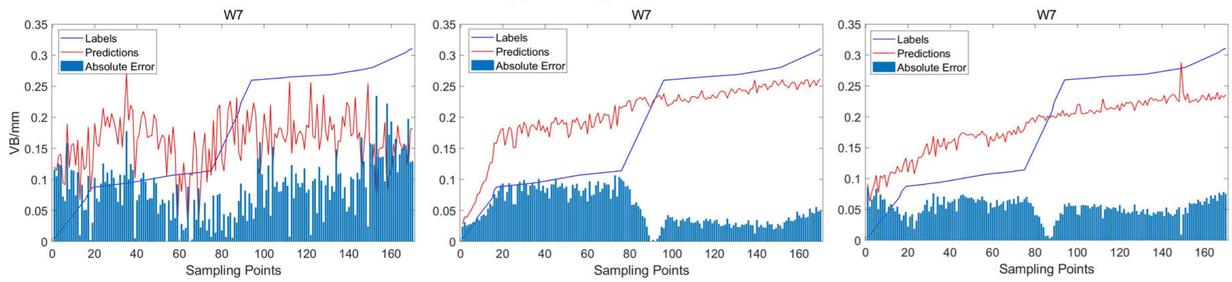
Fig. 4. Prediction details on the Ideahouse dataset. Left: DNN; middle: PINN; right: PINN-WL. (a) Working condition 1. (b) Working condition 2. (c) Working condition 3. (d) Working condition 4. (e) Working condition 5.

It means that the proposed method is reliable for varying working conditions in industrial applications.

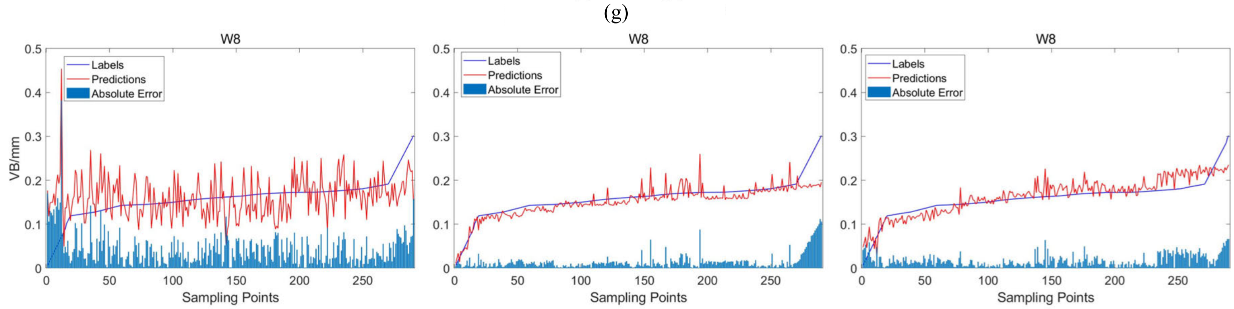
In the second experiment on the Ideahouse dataset, a different data-driven model, the traditional neural network, and the tool wear mechanism are integrated into the proposed framework to verify the performance of the proposed method in order to demonstrate that the framework could be applicated to different data-driven models. The data-driven model is constructed by a fully connected neural network, which contains three hidden layers, and the number of neurons in each hidden layer is 100, 40, and 40, respectively. The nine groups of machining tests are separately verified, where, in each of



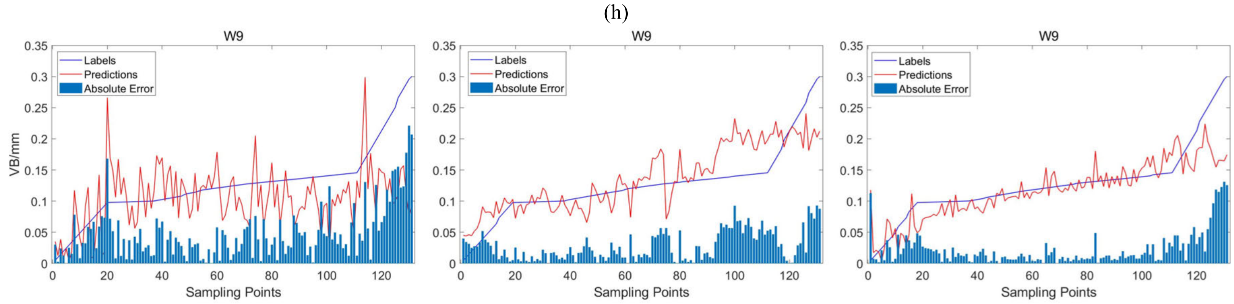
(f)



(g)



(h)



(i)

Fig. 4. *(Continued.)* Prediction details on the Ideahouse dataset. Left: DNN; middle: PINN; right: PINN-WL. (f) Working condition 6. (g) Working condition 7. (h) Working condition 8. (i) Working condition 9.

the groups, 80% of the samples are used for model training, and the rest 20% are used for the model test. The comparison results show that the PINN-WL could be applicated to different data-driven models and could still ensure prediction accuracy and improve the prediction stability, as shown in Fig. 5.

B. Validation in NASA Dataset

The NASA dataset is an open dataset in which eight groups of different working conditions are chosen for model tests. The cross-validation test is designed, as well, where seven groups are used for model training and the remaining one group for the model test; a total of eight tests are carried out. The ranges of cutting parameters' variations: feed per tooth (fz: 0.25–0.5 mm/r) and cutting depth (ap: 0.75–1.5 mm). Under the above eight working conditions, monitored signals, including ac spindle motor current, table and spindle vibration, acoustic emission at table and spindle, are collected, and 180 features are extracted. The data labels refer to the tool wear amount (VB_{max}), which is provided by the dataset.

In this experiment, the prediction results of the above three models are compared as well. The comparison results are shown in Fig. 6, and the prediction details of each working condition are shown in Fig. 7. The same conclusions could be made as that of the Ideahouse dataset.

C. Discussion

The experiment results on the Ideahouse dataset and the NASA dataset have shown the performance improvement of the proposed method compared with the data-driven method and the traditional PINN. Essentially, the unstable prediction

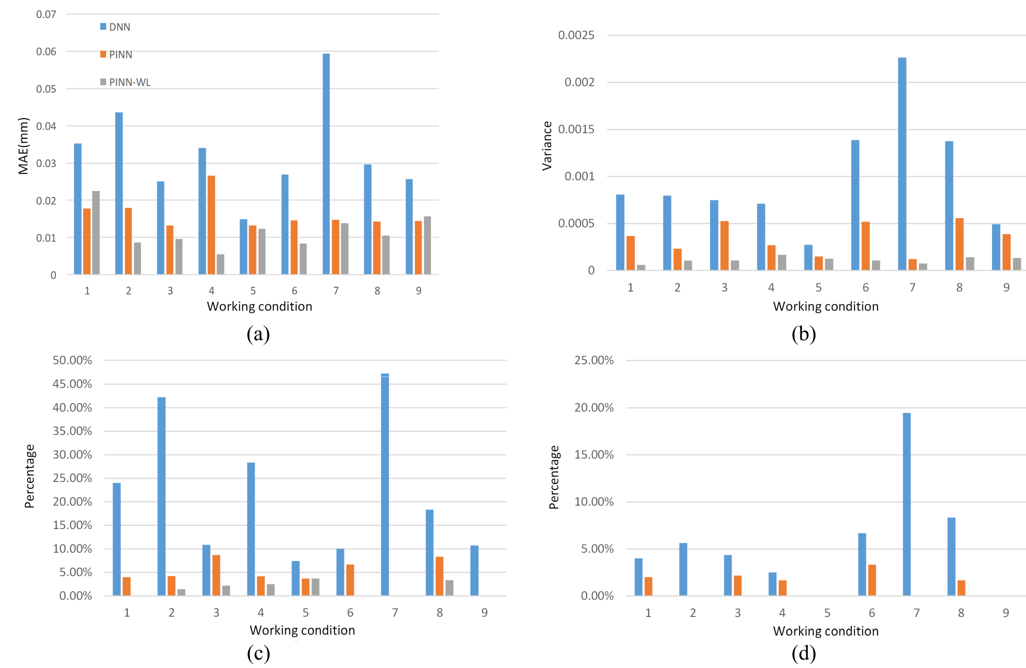


Fig. 5. Comparison results on the Ideahouse dataset (DNN and the tool wear mechanism are integrated into the PINN-WL framework). (a) MAE. (b) Variance of prediction error. (c) Percentage of samples with an error greater than 0.05 mm. (d) Percentage of samples with an error greater than 0.1 mm.

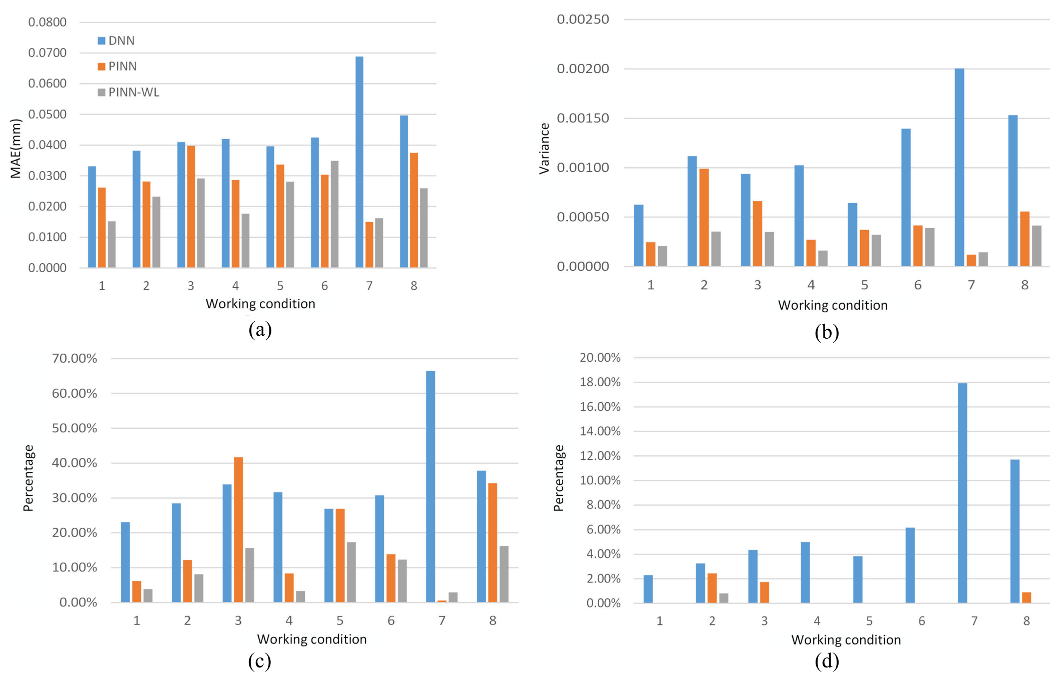


Fig. 6. Comparison results on NASA dataset. (a) MAE. (b) Variance of prediction error. (c) Percentage of samples with an error greater than 0.05 mm. (d) Percentage of samples with an error greater than 0.1 mm.

errors for existing methods in the above cases are caused by the following reasons.

1) The inaccurate physics model, i.e., the tool wear mechanism used, is inaccurate while the working condition changes.

2) The noisy data, i.e., the noise caused by the complex manufacturing environments, affects the signal data.

3) The changes in data distribution caused by varying working conditions in a data-driven model.

In the proposed framework, the weights allocated to each loss function aim to inhibit the impacts of the data volatility or the physics inaccuracy on each model component. The larger the error uncertainty, the smaller the weight allocated, and the more effective the inhibition of the model. It is an effective way to prevent over-fitting, as the unstable data points—the samples with large prediction errors in each group—are restrained to the uttermost extent. As a result, the stability of the integrated model could be improved.

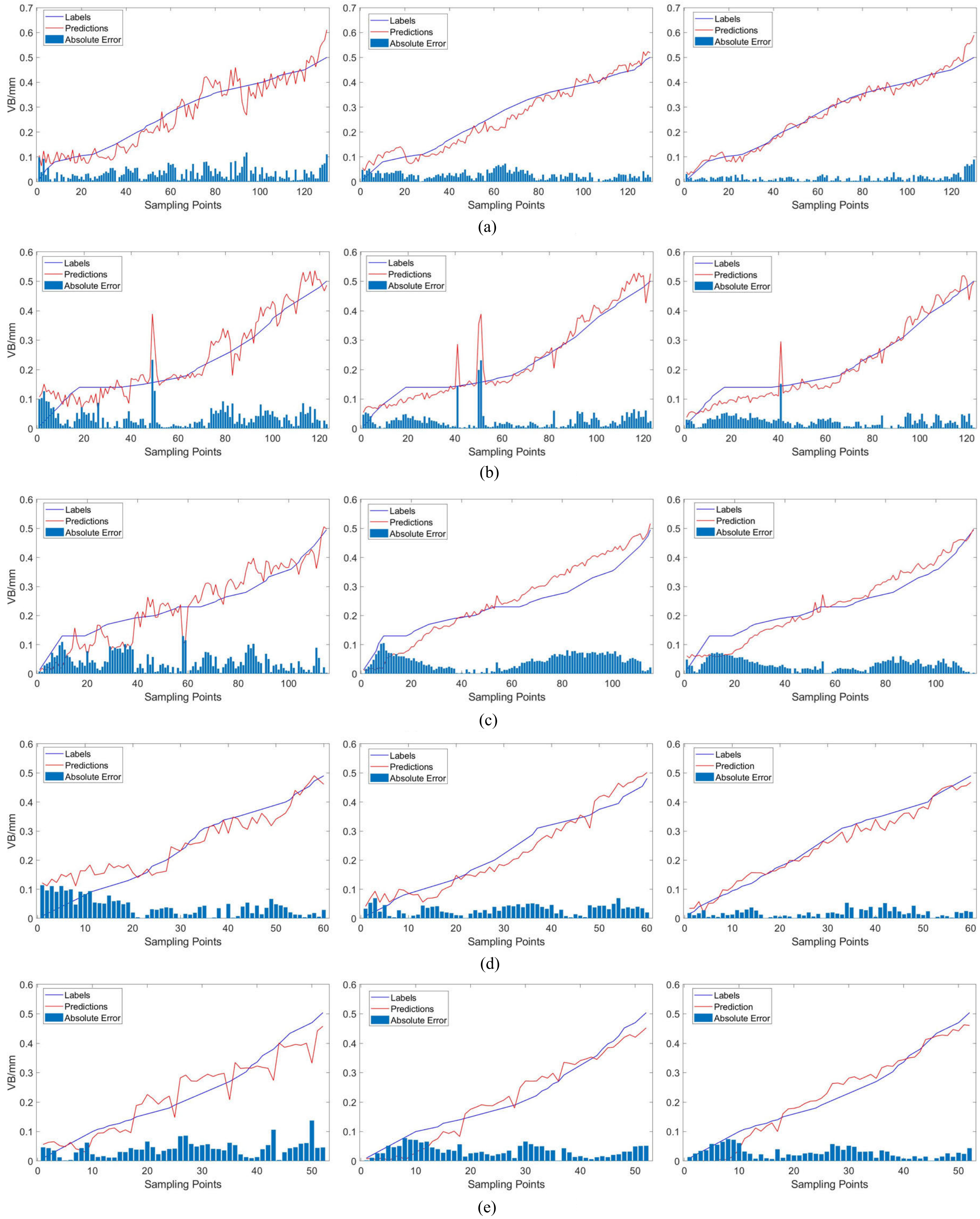
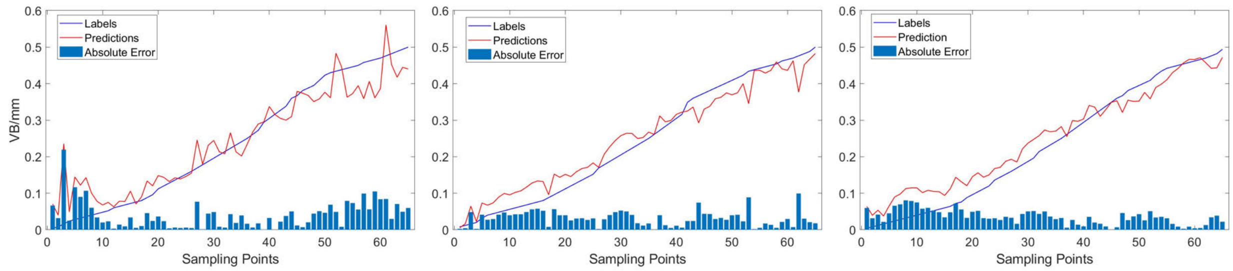


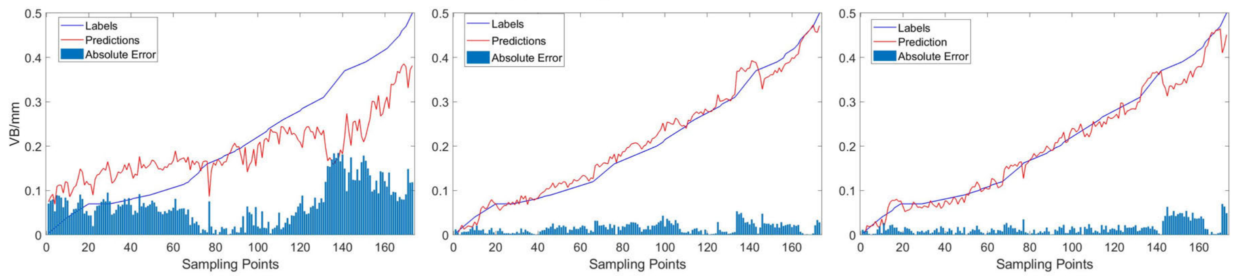
Fig. 7. Prediction details on NASA dataset. Left: DNN; middle: PINN; right: PINN-WL. (a) Working condition 1. (b) Working condition 2. (c) Working condition 3. (d) Working condition 4. (e) Working condition 5.

There are, however, still some unsatisfied results in the above experimental cases. The first problem that needs to be discussed is that in some specific cases (e.g., working condition 6 in the Ideahouse dataset and working condition 7 in the NASA dataset), the prediction accuracy or the stability of the proposed method is similar to the traditional PINN. The reason may be that the weights allocated to the data-driven model and the physics model in the proposed method are well similar to the traditional PINN, where the weights are set to 1 constantly.

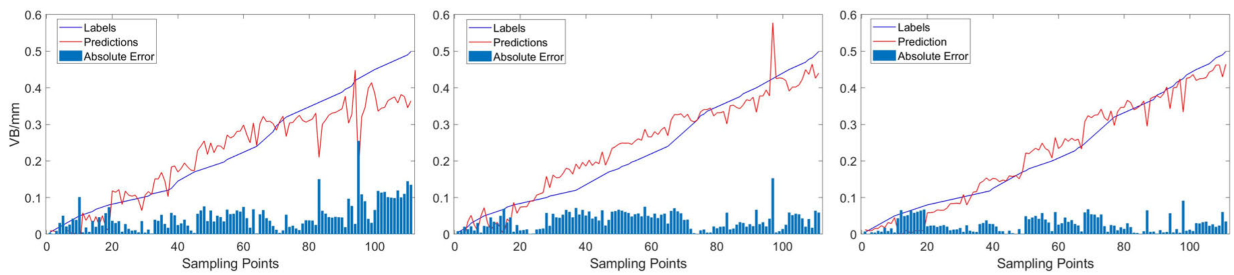
The second problem that needs to be discussed is that the prediction outputs of the three methods on the Ideahouse



(f)



(g)



(h)

Fig. 7. *(Continued.)* Prediction details on NASA dataset. Left: DNN; middle: PINN; right: PINN-WL. (f) Working condition 6. (g) Working condition 7. (h) Working condition 8.

dataset are not better than those on the NASA dataset. The reason may be that the working conditions in the Ideahouse dataset are more varied, where both the straight line and the corner machining features are included in one working condition. Instead, the cutting tool path is merely the straight line in the NASA dataset. Meanwhile, the material used in the Ideahouse dataset is titanium alloy, which is more difficult to cut than the cast iron used in the NASA dataset. The above effects may cause severe couple impacts on the monitoring signals, resulting in the prediction outputs.

V. Conclusion and Further Work

Accurate and stable prediction for manufacturing systems is very important in the area of aerospace machining. To address the challenge that the potential advantages of PINN are limited for situations with inaccurate physics models or noisy data, this article proposed a kind of PINN-WLs by uncertainty evaluation for accurate and stable prediction of manufacturing systems, which is verified in the case of tool wear prediction. Furthermore, the basic idea of this research work has the potential to be applied to tool life monitoring, part quality assurance, production automation, and other industrial applications. The contributions of this research are summarized as follows.

1) A novel weight allocation strategy based on uncertainty evaluation by quantifying the variance of prediction errors is proposed. As a result, the information provided by the data-driven model and the physics model could be effectively used to improve prediction accuracy and stability.
2) An improved PINN framework is established for accurate and stable prediction of manufacturing systems, which is a contribution to the model integrating researches, especially to that with inaccurate physics models, noisy data, and varying data distribution.
3) The experiments confirm the feasibility and effectiveness of the proposed method, as well as its advantages over existing methods, such as DNN and basic PINN, and the comparison is performed on our test data and NASA benchmark data.

Although promising results have been obtained, there are still some points that need to be further studied. As our future work, we will further verify the feasibility and effectiveness of the proposed method for different prediction cases of manufacturing systems.

References

[1] K. Xu et al., "Advanced data collection and analysis in data-driven manufacturing process," *Chin. J. Mech. Eng.*, vol. 33, no. 1, Dec. 2020, doi: 10.1186/s10033-020-00459-x.

[2] C. Liu, Y. Li, J. Li, and J. Hua, "A meta-invariant feature space method for accurate tool wear prediction under cross conditions," *IEEE Trans. Ind. Informat.*, vol. 18, no. 2, pp. 922–931, Feb. 2022, doi: 10.1109/TII.2021.3070109.

[3] Y. Gao, Y. Wen, and J. Wu, "A neural network-based joint prognostic model for data fusion and remaining useful life prediction," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 32, no. 1, pp. 117–127, Jan. 2021.

[4] J. Hua, Y. Li, W. Mou, and C. Liu, "An accurate cutting tool wear prediction method under different cutting conditions based on continual learning," *Proc. Inst. Mech. Eng., B, J. Eng. Manuf.*, vol. 236, nos. 1–2, pp. 123–131, Jan. 2022, doi: 10.1177/0954405421993694.

[5] J. Hua, Y. Li, C. Liu, and L. Wang, "A zero-shot prediction method based on causal inference under non-stationary manufacturing environments for complex manufacturing systems," *Robot. Comput.-Integr. Manuf.*, vol. 77, Oct. 2022, Art. no. 102356, doi: 10.1016/j.rcim.2022.102356.

[6] M. Raissi, P. Perdikaris, and G. E. Karniadakis, "Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations," *J. Comput. Phys.*, vol. 378, pp. 686–707, Mar. 2019.

[7] S. A. Tobias, "Machine tool vibration research," *Int. J. Mach. Tool Design Res.*, vol. 1, nos. 1–2, pp. 1–14, Sep. 1961.

[8] D. Hajdu, F. Borgioli, W. Michiels, T. Insperger, and G. Stepan, "Robust stability of milling operations based on pseudospectral approach," *Int. J. Mach. Tools Manuf.*, vol. 149, Feb. 2020, Art. no. 103516, doi: 10.1016/j.ijmachtools.2019.103516.

[9] Z. Pálmai, "Proposal for a new theoretical model of the cutting tool's flank wear," *Wear*, vol. 303, nos. 1–2, pp. 437–445, Jun. 2013.

[10] J. Rech, A. Giovenco, C. Courbon, and F. Cabanettes, "Toward a new tribological approach to predict cutting tool wear," *CIRP Ann.*, vol. 67, no. 1, pp. 65–68, 2018.

[11] Z. Liu, C. Yue, X. Li, X. Liu, and L. Wang, "Research on tool wear based on 3D FEM simulation for milling process," *J. Manuf. Mater. Process.*, vol. 4, no. 4, pp. 121–138, 2020.

[12] F. Tao, Q. Qi, A. Liu, and A. Kusiak, "Data-driven smart manufacturing," *J. Manuf. Syst.*, vol. 48, pp. 157–169, Jan. 2018.

[13] C. Li, R.-V. Sanchez, G. Zurita, M. Cerrada, D. Cabrera, and R. E. Vásquez, "Gearbox fault diagnosis based on deep random forest fusion of acoustic and vibratory signals," *Mech. Syst. Signal Process.*, vols. 76–77, pp. 283–293, Aug. 2016.

[14] X. Chen, B. Zhang, and D. Gao, "Bearing fault diagnosis base on multi-scale CNN and LSTM model," *J. Intell. Manuf.*, vol. 32, no. 4, pp. 971–987, Apr. 2021.

[15] J. Li, Y. Wang, Y. Zi, and Z. Zhang, "Whitening-Net: A generalized network to diagnose the faults among different machines and conditions," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 33, no. 10, pp. 5845–5858, Oct. 2022, doi: 10.1109/TNNLS.2021.3071564.

[16] P. Fu, J. Wang, X. Zhang, L. Zhang, and R. X. Gao, "Dynamic routing-based multimodal neural network for multi-sensory fault diagnosis of induction motor," *J. Manuf. Syst.*, vol. 55, pp. 264–272, Apr. 2020.

[17] B. Chen, J. Yang, J. Zhao, and J. Ren, "Milling chatter prediction based on the information entropy and support vector machine," in *Proc. Int. Ind. Informat. Comput. Eng. Conf.*, 2015, pp. 376–380, doi: 10.2991/iiicec-15.2015.86.

[18] J. Friedrich, C. Hinze, A. Renner, A. Verl, and A. Lechler, "Estimation of stability lobe diagrams in milling with continuous learning algorithms," *Robot. Comput.-Integr. Manuf.*, vol. 43, pp. 124–134, Feb. 2017.

[19] J. Wang, B. Zou, M. Liu, Y. Li, H. Ding, and K. Xue, "Milling force prediction model based on transfer learning and neural network," *J. Intell. Manuf.*, vol. 32, no. 4, pp. 947–956, Apr. 2021.

[20] S. Su, G. Zhao, W. Xiao, Y. Yang, and X. Cao, "An image-based approach to predict instantaneous cutting forces using convolutional neural networks in end milling operation," *Int. J. Adv. Manuf. Technol.*, vol. 115, pp. 1657–1669, May 2021.

[21] Y. Fu, Y. Zhang, H. Qiao, D. Li, H. Zhou, and J. Leopold, "Analysis of feature extracting ability for cutting state monitoring using deep belief networks," *Proc. CIRP*, vol. 31, pp. 29–34, Jan. 2015.

[22] Z. Huang, J. Zhu, J. Lei, X. Li, and F. Tian, "Tool wear predicting based on multisensory raw signals fusion by reshaped time series convolutional neural network in manufacturing," *IEEE Access*, vol. 7, pp. 178640–178651, 2019.

[23] Y. Li, C. Liu, J. Hua, J. Gao, and P. Maropoulos, "A novel method for accurately monitoring and predicting tool wear under varying cutting conditions based on meta-learning," *CIRP Ann.*, vol. 68, no. 1, pp. 487–490, 2019.

[24] M. Lutter, C. Ritter, and J. Peters, "Deep Lagrangian networks: Using physics as model prior for deep learning," 2019, *arXiv:1907.04490*.

[25] P. Jin, Z. Zhang, A. Zhu, Y. Tang, and G. E. Karniadakis, "SympNets: Intrinsic structure-preserving symplectic networks for identifying Hamiltonian systems," *Neural Netw.*, vol. 132, pp. 166–179, Dec. 2020.

[26] C. Oszkinat, S. E. Luczak, and I. G. Rosen, "Uncertainty quantification in estimating blood alcohol concentration from transdermal alcohol level with physics-informed neural networks," *IEEE Trans. Neural Netw. Learn. Syst.*, early access, Jan. 17, 2022, doi: 10.1109/TNNLS.2022.3140726.

[27] M. Postel, B. Bugdayci, and K. Wegener, "Ensemble transfer learning for refining stability predictions in milling using experimental stability states," *Int. J. Adv. Manuf. Technol.*, vol. 107, nos. 9–10, pp. 4123–4139, Apr. 2020.

[28] G. Chen, Y. Li, X. Liu, and B. Yang, "Physics-informed Bayesian inference for milling stability analysis," *Int. J. Mach. Tools Manuf.*, vol. 167, Aug. 2021, Art. no. 103767.

[29] T. Yu, Z. Li, and D. Wu, "Predictive modeling of material removal rate in chemical mechanical planarization with physics-informed machine learning," *Wear*, vols. 426–427, pp. 1430–1438, Apr. 2019.

[30] J. Huang, Q. Chang, and J. Arinez, "Product completion time prediction using a hybrid approach combining deep learning and system model," *J. Manuf. Syst.*, vol. 57, pp. 311–322, Oct. 2020.

[31] J. Wang, Y. Li, R. Zhao, and R. X. Gao, "Physics guided neural network for machining tool wear prediction," *J. Manuf. Syst.*, vol. 57, pp. 298–310, Oct. 2020.

[32] H. Hanachi, W. Yu, I. Y. Kim, J. Liu, and C. K. Mechefske, "Hybrid data-driven physics-based model fusion framework for tool wear prediction," *Int. J. Adv. Manuf. Technol.*, vol. 101, nos. 9–12, pp. 2861–2872, Apr. 2019.

[33] J. Shen, Y. Zhao, J. K. Liu, and Y. Wang, "HybridSNN: Combining bio-machine strengths by boosting adaptive spiking neural networks," *IEEE Trans. Neural Netw. Learn. Syst.*, early access, Dec. 10, 2021, doi: 10.1109/TNNLS.2021.3131356.

[34] S. E. Yuksel, J. N. Wilson, and P. D. Gader, "Twenty years of mixture of experts," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 23, no. 8, pp. 1177–1193, Aug. 2012, doi: 10.1109/TNNLS.2012.2200299.

[35] S. Arora, S. Du, W. Hu, Z. Li, R. Salakhutdinov, and R. Wang, "On exact computation with an infinitely wide neural net," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 32, 2019, pp. 8141–8150.

[36] Z. Fang, "A high-efficient hybrid physics-informed neural networks based on convolutional neural network," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 33, no. 10, pp. 5514–5526, Oct. 2022, doi: 10.1109/TNNLS.2021.3070878.

[37] Z. Zhou, *Machine Learning*. Beijing, China: Tsinghua Press, 2016.

[38] J. Wang, L. Pagani, R. K. Leach, W. Zeng, B. M. Colosimo, and L. Zhou, "Study of weighted fusion methods for the measurement of surface geometry," *Precis. Eng.*, vol. 47, pp. 111–121, Jan. 2017.

[39] B. Xiang, Y. Li, G. Chen, X. Liu, and W. Yang, "Multi-source integrated fusion for surface measurement," *Int. J. Adv. Manuf. Technol.*, vol. 109, nos. 7–8, pp. 1815–1823, Aug. 2020.

[40] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," 2014, *arXiv:1412.6980*.

[41] R. H. Byrd, P. Lu, J. Nocedal, and C. Zhu, "A limited memory algorithm for bound constrained optimization," *SIAM J. Sci. Comput.*, vol. 16, no. 5, pp. 1190–1208, 1995.

[42] A. D. Jagtap, K. Kawaguchi, and G. E. Karniadakis, "Locally adaptive activation functions with slope recovery for deep and physics-informed neural networks," *Proc. Roy. Soc. A, Math., Phys. Eng. Sci.*, vol. 476, no. 2239, Jul. 2020, Art. no. 20200334, doi: 10.1098/rspa.2020.0334.

[43] A. D. Jagtap, K. Kawaguchi, and G. E. Karniadakis, "Adaptive activation functions accelerate convergence in deep and physics-informed neural networks," *J. Comput. Phys.*, vol. 404, Mar. 2020, Art. no. 109136.

[44] Y. Li, J. Wang, Z. Huang, and R. X. Gao, "Physics-informed meta learning for machining tool wear prediction," *J. Manuf. Syst.*, vol. 62, pp. 17–27, Jan. 2022.

[45] X. Liu, X. Zhang, W. Peng, W. Zhou, and W. Yao, "A novel meta-learning initialization method for physics-informed neural networks," *Neural Comput. Appl.*, vol. 34, no. 17, pp. 14511–14534, 2022.

[46] C. Finn, P. Abbeel, and S. Levine, "Model-agnostic meta-learning for fast adaptation of deep networks," in *Proc. Int. Conf. Mach. Learn.*, 2017, pp. 1856–1868.

[47] Y. Shin, J. Darbon, and G. E. Karniadakis, "On the convergence of physics informed neural networks for linear second-order elliptic and parabolic type PDEs," 2020, *arXiv:2004.01806*.

[48] Y. Li, C. Liu, D. Li, J. Hua, and P. Wan, "Tool wear dataset of NUAA_IDEAHOUSE," IEEE Dataport, 2021, doi: 10.21227/3aa1-5e83.

[49] A. Agogino and K. Goebel, "Milling data set," NASA Ames Res. Center, Moffett Field, CA, USA, 2007. [Online]. Available: http://ti.arc.nasa.gov/project/prognostic-data-repository