Vibration Signal-based Tool Condition Monitoring Using Regularized Sensor Data Modelling and Model Frequency Analysis

Zepeng Liu, *Member, IEEE*, Zi-Qiang Lang, Yufei Gui, Yun-Peng Zhu, *Member, IEEE*, Hatim Laalej, David Curtis

***Abstract*—Tool condition monitoring (TCM) plays a vital role in maintaining product quality and improving productivity in advanced manufacturing. However, complex machining environments often limit the monitoring accuracy of conventional monitoring systems. In the present study, a new diagnostic framework is proposed for TCM during machining using a novel regularization-based sensor data modelling and model frequency analysis. For the first time, the physical information of the underlying machining process is incorporated into the modelling procedure for the design of the associated regularization parameter. This ensures that significant underlying physics can be taken into account during the modelling so as to enhance the TCM performance. This idea is referred to as tool condition monitoring-oriented regularization (TCMoR). After a model has been identified from TCMoR-based sensor data modelling, the frequency domain properties of the model are extracted to reveal unique and physically meaningful features of the underlying machining process for the TCM purpose. The effectiveness of the proposed diagnostic framework is validated by extensive *in-situ* experimental studies under both variable and controlled tool-workpiece engagement conditions, demonstrating its advantages over conventional TCM methods and its potential applications in industry.**

***Index Terms*—Tool condition monitoring, Regularization, Data-driven modelling, Nonlinear AutoRegressive with eXogenous input (NARX) model, Nonlinear output frequency response functions (NOFRFs)**

I. Introduction

In advanced manufacturing, the focus is naturally on producing high-performance components with tighter tolerances in the required time frame under conditions of maximized profit [1]. The failure of machine tools will result in the halt of the whole production, which in turn will result in economic losses [2]. Some scholars have estimated that the increase in unit machining costs due to the malfunction of machine tools is around 50% [3], while others put the figure closer to 100% [4]. Therefore, efficient and cost-effective tool condition monitoring (TCM) methods have gained extensive attention from both academia and industry.

According to different application scenarios, TCM methods can be categorized into direct and indirect methods [5]. Direct methods mainly utilize optical sensors, cameras, and other specialized equipment to directly observe and quantify changes in the material composition or geometric profile of the cutting tool [6]. You *et al* [7]. captured cutting tool images in an industrial environment and employed a lightweight network model based on the adaptive activation function to classify different tool wear levels. Miao *et al* [8]. applied a charge-coupled device (CCD) camera to capture the flank face of cutting tools from turning machining and used the U-Net-based network to detect and identify the tool wear levels. You *et al* [9]. proposed an evaluation metric for TCM based on wear distance dispersion in order to evaluate different flank wear forms throughout the tool life cycle. In addition to 2D imaging for TCM, there has been great interest in quantifying the 3D morphology of tool wear to evaluate cutting tool conditions. Zhu et al [10]. utilized microscopes to capture the raw 3D tool wear images and applied low-pass and high-pass filters to obtain reconstruction images for tool wear volume evaluations. Wang *et al* [11]. reconstructed a 3D crater profile to measure tool wear cracks relating to crater center, crater width, crater depth, and crater front distance. However, direct methods can usually only be used when the machine is stopped, and the relatively high cost of the associated measurement equipment prevents their widespread use in manufacturing. As a result, these direct methods have very limited applicability in real-world industrial manufacturing environments [12].

Alternatively, the indirect methods rely on the changes in sensor signals to monitor tool wear in machining processes [13]. The widely used sensor signals for TCM include cutting force, torque, temperature, acoustic emission and vibration [14], [15]. These approaches directly extract features from collected signals to reveal the physical states of cutting tools. The signal features can be the time-domain features including mean, kurtosis, etc., frequency domain features including frequency variance and time and frequency-domain features, such as wavelet packet energy [16]. Then, based on the extracted signal features, the classification models, such as support vector machines (SVM), artificial neural networks (ANN), etc., or regression models such as Gaussian process

This work was supported by the Engineering and Physical Sciences Research Council under Grant No. (EP/T024291/1). (Corresponding author: Zi-Qiang Lang)

Z.Liu is with the School of Engineering, Newcastle University, NE1 7RU (e-mail: zepeng.liu@newcastle.ac.uk).

ZQ.Lang and Y.Gui are with the Department of Automatic Control and Systems Engineering, University of Sheffield, Sheffield, S10 2TN, UK.

YP.Zhu is with the School of Engineering and Materials Science, Queen Mary University of London, E1 4NS, UK.

H.Laalej and D.Curtis are with the Advanced Manufacturing Research Centre, University of Sheffield, Sheffield, S10 2TN, UK.

regression are built to diagnose cutting tool health conditions or prognose remaining useful life (RUL) of the tools [8]. In [17], the signal features were extracted from force, vibration and acoustic emission (AE) signals, and the tool wear was subsequently monitored and predicted by feeding these extracted features into a deep learning regression model known as the sequence-to-sequence model with attention and monotonicity loss (SMAML). Wang *et al* [18]. utilized a t-distributed stochastic neighbor embedding (t-SNE) method to classify various tool wear levels, utilizing features extracted from both vibration and force signals. Liu *et al* [19]. incorporated the extracted signal features into the switching Hidden semi-Markov model (HSMM)), aiming to predict the tool wear level under time-varying cutting model conditions where machining parameters such as cutting speed, cutting depth and feed speed were set to be inconstant during machining. The aforementioned indirect methods are supervised learning methods requiring labeling of the extracted features. In situations where labeling becomes infeasible, resorting to an unsupervised learning algorithm becomes essential. For instance, Gui *et al* [20]. proposed an unsupervised tool breakage detection (TBD) framework based on both time and sensor domain data decomposition and analysis. Within this framework, the time series analysis is applied to generate the features for TBD, and sensor domain data is exploited to determine the threshold for the unsupervised learning model.

Although signal feature-based TCM has been extensively studied in the literature, only a very few implementations have been found in real-world applications [21]. This is because these techniques have two fundamental limitations and drawbacks. First, robustness is the most important concern. The techniques can often only be used on repeated and simple machining processes which would enable the signal features to be representative of cutting tool conditions. This implies that the techniques may not be suitable for real-world applications where the machining processes are often much more complicated. Secondly, different signal features must be designed for different milling processes, which makes the feature design a significantly demanding task to be completed every time [14].

In order to resolve the challenges faced by current signal feature-based methods, a model feature-based approach known as sensor data modelling and model frequency analysis was proposed [22], [23]. The main difference between the model feature-based approach and the signal feature-based approach is that the model feature-based approach does not directly extract features from collected sensor signals for TCM. Instead, the collected signals are firstly utilized to build a dynamic process model representing the relationship between the signals. After that, the frequency response characteristics of the built dynamic model are extracted and used as model features to conduct TCM. As has already been demonstrated by [22], the model feature-based method can resolve limitations and fundamentally address low robustness and complexity issues associated with signal feature-based methods.

The key to the model feature-based TCM introduced is that an effective model has to be built to appropriately reveal tool wear-induced dynamic changes in the machining process. However, our recent studies show that the conventional nonlinear system identification as applied in [22] sometimes cannot achieve this objective. This is because the identified model may not capture the key information of a machining process over tooth passing frequencies in many complicated scenarios even though overall the model may perform well in representing the time domain behaviors of the process.

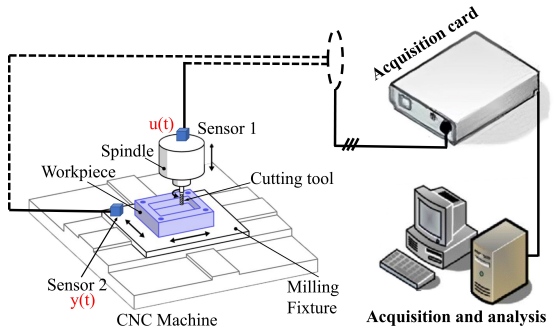


Fig. 1. The illustration of the machining and data acquisition process.

In order to address this problem, in the present study, a new diagnostic framework is proposed for TCM during machining using a novel regularization-based sensor data modelling and model frequency analysis. Under this framework, for the first time, the tooth passing frequency-related physical information of the underlying machining process is incorporated into the modelling procedure for the design of the associated regularization parameter. This ensures significant underlying physics can be taken into account during the modelling so as to enhance the model feature-based TCM performance. The idea is referred to as tool condition monitoring-oriented regularization (TCMoR). Compared with the conventional regularization (CR) method introduced in most studies [24], TCMoR can produce a model that can much more effectively represent the dynamics of the machining process induced by worn cutting tools, enabling model features to better reveal cutting tool status and be better exploited for TCM purposes. Extensive *in-situ* experimental studies under both variable and controlled tool-workpiece engagement conditions were carried out. These studies demonstrate that this novel TCMoR-based TCM framework is a significant development of the sensor data modelling and model frequency analysis-based TCM and has better potential to be applied in industry.

II. Sensor Data Modelling and Model Frequency Analysis for Tool Condition Monitoring

A. The Idea

Fig.1 shows a machining process of cutting tool system where $u(t)$ and $y(t)$ are the vibration signals measured on the spindle and on the milling fixture beneath the workpiece, respectively. To monitor the health status of cutting tools, a model feature-based method was proposed which involves two primary steps.

1) Sensor data modelling is applied to build a dynamic model to represent the dynamic relationship between $u(t)$ and $y(t)$ collected from the vibration sensors mounted on the

spindle and the milling fixture, respectively. According to [25], a general model representing the dynamics between $u(t)$ and $y(t)$ can be expressed as $y(t) = \mathbb{D}[u(t)]$ where the notation $\mathbb{D}$ indicates a dynamic relationship between the two measured signals.

2) Model frequency analysis is conducted to extract the frequency response characteristics of the built models, and the extracted frequency characteristics are used as the model features to evaluate the health conditions of the underlying machining process. In comparison with the conventional signal features, the proposed model features can reveal unique and physically meaningful characteristics of machining processes which can overcome the problems with conventional signal feature-based methods and be potentially applied to perform a more effective TCM.

For sensor data modelling, if the dynamic relationship $\mathbb{D}$ is linear, data-driven modelling methods will produce a linear model to represent the linear relationship, and the frequency response characteristics of the model can be obtained as $\frac{\mathscr{F}[y(t)]}{\mathscr{F}[u(t)]} = H(j\omega_l)$ where the notation $\mathscr{F}$ represents the Fourier transform and ω_l indicates the frequency variable. $H(j\omega_l)$ indicates the frequency response function (FRF) of the model, which can be used as the model features for TCM. However, if the dynamic relationship between $u(t)$ and $y(t)$ is nonlinear, nonlinear data-driven modelling and nonlinear model frequency analysis are proposed where Nonlinear Output Frequency Response Functions (NOFRFs) can be used to replace FRF for TCM.

Nonetheless, the conventional nonlinear data-driven modelling techniques sometimes have poor performance when applied to sensor data modelling and model frequency analysis-based TCM. This is because, in many complicated machining processes, the nonlinear models determined by conventional techniques often fail to capture critical information relating to the tooth passing frequency, i.e., the frequency at which the teeth of the tool cut the workpiece during machining. To overcome this issue, a novel TCMoR-based sensor data modelling is proposed which, for the first time, incorporates the physical information of the underlying machining process, i.e., tooth passing frequency, into the data-driven modelling procedure which can more effectively reveal the cutter status and be better exploited for the TCM purpose. More details of the sensor data modelling and model frequency analysis will be introduced in Section II-B to Section II-D, and the novel idea of the TCMoR will be introduced in detail in Section III.

B. Sensor Data Modelling

The relationship between $u(t)$ and $y(t)$ in Fig.1 can be represented by using the Nonlinear AutoRegressive with eXogenous input (NARX) model

$$y(t) = f^{\ell}[\mathbf{x}(t)] + e(t) \tag{1}$$

where $t = 1, ..., \Gamma$ with the sampling frequency f_s. $f^{\ell}[.]$ represents the polynomial function with maximum degree $\ell \in \mathbb{Z}^+$. $e(t)$ represents the noise and unmodeled dynamics.

$$\begin{aligned}\mathbf{x}(t) =&[y(t-1), \cdots, y(t-n_y), u(t-1), \cdots, u(t-n_u)] \\ =&[x_1(t), ..., x_\Delta(t)]\end{aligned} \tag{2}$$

where n_u and n_y are, respectively, the length of the input lag and output lag, and $\Delta = n_y + n_u$ [26]. Under rather general conditions [26], the polynomial function $f^{\ell}[.]$ can be represented by M weighted regressors

$$y(t) = \sum_{m=1}^{M} \theta_m \phi_m(t) + e(t), t = 1, ..., \Gamma \tag{3}$$

where $\phi_m(t)$, $m = 1, ..., M$ is a m-th order monomial term composed of $x_1(t), ..., x_\Delta(t)$, and M is the total number of candidate regressors or model terms. Define $\mathbf{y} = [y(1), .., y(\Gamma)]^{\mathrm{T}}$, $\mathbf{\Phi}_m = [\phi_m(1), ..., \phi_m(\Gamma)]^{\mathrm{T}}$, $\mathbf{\Phi} = [\mathbf{\Phi}_1, ..., \mathbf{\Phi}_M]$, $\boldsymbol{\theta} = [\theta_1, ..., \theta_m]^{\mathrm{T}}$ and $\mathbf{e} = [e(1), .., e(\Gamma)]^{\mathrm{T}}$, the regression model (3) can be written in the matrix form of $\mathbf{y} = \mathbf{\Phi}\boldsymbol{\theta} + \mathbf{e}$. In order to determine the structure of the model, the Forward Regression with Orthogonal Least Squares (FROLS) learning algorithm can be applied [26]. The final model structure is the linear combination of the M_0 ($M_0 \leq M$) significant model terms chosen from the M candidate regressors $\mathbf{\Phi}$, producing

$$\mathbf{y} = \mathbf{W}\mathbf{g} + \mathbf{e} \tag{4}$$

where $\mathbf{g} = [g_1, ..., g_{M_0}]^{\mathrm{T}}$ is the FROLS regression weight vector, and $\mathbf{W} = [\mathbf{\Phi}_1, ..., \mathbf{\Phi}_{M_0}]$ is the selected regressors. Therefore, the $\mathbf{g}$ can be evaluated by solving the following l_2-norm regularization problem.

$$\begin{aligned}\mathbf{g} =& \arg\min_{\mathbf{g}}\{\|\mathbf{W}\mathbf{g} - \mathbf{y}\|_2^2 + \lambda_g \|\mathbf{g}\|_2^2\} \\ =&\mathbf{W}^{\mathrm{T}}\mathbf{y}/(\mathbf{W}^{\mathrm{T}}\mathbf{W} + \lambda_g \mathbf{I})\end{aligned} \tag{5}$$

where λ_g is the regularization parameter, and $\mathbf{I}$ is an $M_0 \times M_0$ identity matrix.

C. Model Frequency Analysis

After $u(t)$ and $y(t)$ collected from the sensors fitted on the spindle and milling fixture, respectively, are used to build dynamic models, the frequency response characteristics of the model are extracted in order to conduct TCM. To extract nonlinear system frequency response characteristics, the Nonlinear Output Frequency Response Functions (NOFRFs) are applied here following the idea introduced in [27].

When the identified NARX model is stable at zero equilibrium, the output can be described by the Volterra series in the time domain as [26]

$$\begin{aligned}y(t) \approx& \sum_{n=1}^{N} y_n(t) \\ =& \sum_{n=1}^{N} \sum_{\tau_1=0}^{\infty} \cdots \sum_{\tau_n=0}^{\infty} h_n(\tau_1, \cdots \tau_n) \prod_{i=1}^{n} u(t-\tau_i) d\tau_i\end{aligned} \tag{6}$$

where N is the maximum order of the system nonlinearity. $y_n(t)$ denotes the nth order output, and $h_n(\tau_1, \cdots, \tau_n)$ is the nth order Volterra kernel of the system output. Based on [26], [27], the Fourier transform of $y(t)$ can be expressed by the NOFRFs-based representation as

$$\mathscr{F}(y(t)) = Y(j\omega) \approx \sum_{n=1}^{N} Y_n(j\omega) = \sum_{n=1}^{N} G_n(j\omega) U_n(j\omega) \tag{7}$$

where ω denotes frequency variables. $\mathscr{F}(\cdot)$ denotes the Fourier transform. $Y_n(j\omega) = \mathscr{F}(y_n(t))$ and $U_n(j\omega) = \mathscr{F}(u^n(t))$ are the nth-order output frequency spectrum and nth-order input frequency spectrum, respectively. $G_n(j\omega)$ is the NOFRFs and is defined by

$$G_n(j\omega) = Y_n(j\omega)/U_n(j\omega) \tag{8}$$

where $\omega \in \Omega_n$, $n = 1, ..., N$ with Ω_n indicating the frequency support of $U_n(j\omega)$ such that $U_n(j\omega) \neq 0$ with $\omega \in \Omega_n$ [28].

To determine NOFRFs, Generalized Associated Linear Equations (GALEs) are used which allow an evaluation of the NOFRFs up to any order from the solutions to a series of linear difference equations. More explanations of GALEs are presented in Appendix A. Suppose $u^*(t)$ is the input excitation signal for NOFRFs evaluation, the model output $y_n^*(t)$ can be evaluated using GALEs; and the NOFRFs, $G_n^*(j\omega)$, can be calculated based on the input signal $u^*(t)$ and the model output $y_n^*(t)$ as

$$G_n^*(j\omega) = \mathscr{F}[y_n^*(t)]/\mathscr{F}\{[u^*(t)]^n\} = Y_n^*(j\omega)/U_n^*(j\omega) \tag{9}$$

where $U_n^*(j\omega) \neq 0$. As a result, The NOFRFs-based model features are obtained as

$$\mathbf{F} = [\mathbf{f}_1, ..., \mathbf{f}_N] \tag{10}$$

with $\mathbf{f}_n = [|G_n^*(j\omega_n^1)|, ..., |G_n^*(j\omega_n^{L_n})|]$, $n = 1, ...N$ representing the n-th order NOFRFs-based model features. Here, L_n indicates the number of model features of the n-th order, and $\omega_n^1, ..., \omega_n^{L_n} \in \bar{\Omega}_n$, $\bar{\Omega}_n \in \Omega_n$.

D. Tool Condition Monitoring

The technique of sensor data modelling and model frequency analysis for TCM implements a strategy of off-line training and on-line monitoring, and a pictorial view of this framework is presented in Fig.2. As can be seen, the first stage is the off-line training which is summarized as follows.

Step 1: Sensor data modelling is conducted off-line in order to build a NARX model via the collected spindle vibration signal $u(t)$ and milling fixture vibration signal $y(t)$.

Step 2: Based on the built NARX model, model frequency analysis is applied to extract the NOFRFs-based model features $\mathbf{F}$.

Step 3: A classifier is trained off-line by using the labeled NOFRFs-based model features $\mathbf{F}$.

After the classifier is trained, the second stage is to conduct on-line TCM using the on-line extracted NOFRFs-based model features and the off-line trained classifier. The details are as follows:

Step 1: Sensor data modelling is used on-line to build a NARX model in order to find the dynamic relationship between the vibration signals measured on the spindle and milling fixture.

Step 2: Model frequency analysis is applied on-line to extract the NOFRFs-based model features $\mathbf{F}$.

Step 3: The on-line TCM is carried out by using the on-line extracted NOFRFs-based model features $\mathbf{F}$ and the off-line trained classifier.

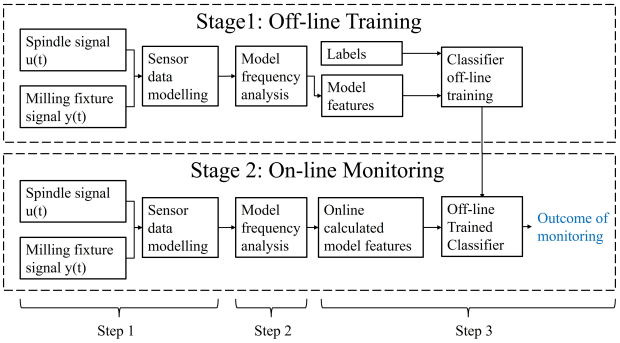


Fig. 2. The flowchart of the off-line training and on-line monitoring strategy.

E. The Need for Tool Condition Monitoring-oriented Regularization (TCMoR)

To obtain a satisfactory NARX model for TCM, one key factor is the regularization parameter λ_g that controls the bias-variance error associated with the identified NARX model. Different λ_g can produce different NARX models and, consequently, different NOFRFs. As a result, an appropriate λ_g is very important for the performance of the NOFRFs-based TCM.

In conventional nonlinear system identification, λ_g is often determined to find a mathematical model that can best fit the system input and output data in the time-domain. However, this can lead to over-fitted models which are either unstable or lack generalization. To overcome these difficulties, a specific procedure of CR was proposed which is to ensure that the identified NARX models can meet the requirements of stability, generalization, and accuracy at the same time [22]. But our recent studies show that for many more complicated machining processes this procedure can still not achieve a satisfactory TCM performance. This is because, in these complicated scenarios, CR often fails to facilitate the capture of key dynamics of machining processes over tooth passing frequencies, although overall the identified model may perform well in terms of stability, generalization and accuracy in the time domain.

It is to profoundly resolve this challenge that the novel regularization method of TCMoR is proposed in the present study. The objectives are not only to guarantee that the identified NARX model possesses stability, generalization and accuracy in the time domain, but also ensure that the NOFRFs extracted from the identified NARX model can well represent the dynamics of machining processes over tooth passing frequencies. The details of the novel TCMoR method will be introduced in the following section.

III. Tool Condition Monitoring-oriented Regularization (TCMoR)

A. Identified NARX Model and Model Predicted Output (MPO) in the Time and Frequency Domain

Assume that the original datasets $u(t)$ and $y(t)$ can be split into two datasets which are the training dataset used for model training and the testing dataset used for model evaluation (See Fig.3). The training dataset is expressed as $u_\alpha(t_1)$ and $y_\alpha(t_1)$

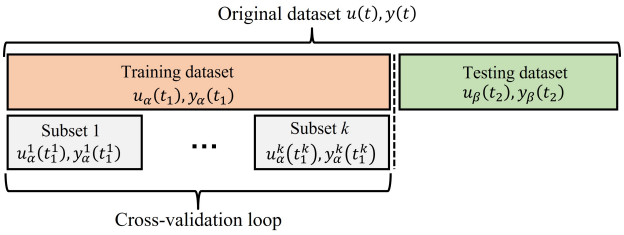


Fig. 3. Illustration of the data split.

with $t_1 = 1, ..., \Gamma_1$ and $\Gamma_1 < \Gamma$; and the testing dataset is represented as $u_\beta(t_2)$ and $y_\beta(t_2)$ with $t_2 = 1, ..., \Gamma_2$ and $\Gamma_2 < \Gamma$. Based on the training dataset, the identified NARX model using FROLS is given by

$$\begin{aligned} y_\alpha(t_1) =& F^\ell[y_\alpha(t_1-1), \cdots, y_\alpha(t_1-n_y), \\ & u_\alpha(t_1-1), \cdots, u_\alpha(t_1-n_u)], t_1 = 1, ..., \Gamma_1 \end{aligned} \tag{11}$$

where $F^\ell[.]$ represents the polynomial function of the variables $y_\alpha(t_1-1), \cdots, y_\alpha(t_1-n_y), u_\alpha(t_1-1), \cdots, u_\alpha(t_1-n_u)$. Then, the Model Predicted Output (MPO) [26] of the model can be evaluated by using the testing dataset as

$$\begin{aligned} \hat{y}_\beta(t_2) =& F^\ell[\hat{y}_\beta(t_2-1), .., \hat{y}_\beta(t_2-n_y), \\ & u_\beta(t_2-1), ..., u_\beta(t_2-n_u)], t_2 = 1, ..., \Gamma_2 \end{aligned} \tag{12}$$

Finally, to study the frequency domain properties of the MPO predictions of the testing dataset, the Fourier transform of MPO is obtained as

$$\hat{Y}_\beta(\omega) = \mathscr{F}(\hat{y}_\beta(t_2)) \tag{13}$$

B. Stability and Generalization Criteria

Stability is evaluated for the sensor data modelling in order to guarantee the built model is asymptotically stable at the zero equilibrium and can then be used for NOFRFs evaluation. More rigorously, the criterion of stability is as follow.

Criterion 1 (Stability)**.** Let $u_\beta(t_2) = 0, \forall t_2$ and $\hat{y}_\beta(t_2) \leq \mathbb{T}_s$, $\forall t_2 \leq 0$. The identified model (11) is stable if

$$e_s(t_2) = |\hat{y}_\beta(t_2)| - \epsilon_s \leq 0 \quad \text{with} \quad t_2 \to \infty, \epsilon_s \to 0 \tag{14}$$

where the notation $\mathbb{T}_s$ represents the initial condition threshold. ϵ_s denotes the stability threshold and e_s represents the stability error.

The generalization of the built model is also evaluated at the sensor data modelling stage in order to enable the built model to be less sensitive or even insensitive to an increasing amount of extreme values in the training dataset. The criteria of generalization is as follow.

Criterion 2 (Generalization)**.** k-fold cross-validation is used to assess the generalization performance of the model. The particulars are summarized in the following steps.

1) As can be seen in Fig.3, the training dataset $u_\alpha(t_1)$ and $y_\alpha(t_1)$ are partitioned into k subsets which are denoted as $u_\alpha^k(t_1^k)$ and $y_\alpha^k(t_1^k)$ with $k = 1, ..., K$ and $t_1^k = 1, ..., \Gamma_1^k$, $\Gamma_1^k < \Gamma_1$.
2) The modelling process is conducted K times by using the selected candidate λ_g. For each modelling process, the $k-1$ subsets are combined together to train the model and the remaining one subset is used to validate the model performance. The MPO of the remaining one subset is denoted as $\hat{y}_\alpha^k(t_1^k)$
3) An overall performance of the MPO across K validation sets is evaluated as

$$\mathrm{CV} = \sum_{k=1}^{K} \sum_{t_1^k=1}^{\Gamma_1^k} [\hat{y}_\alpha^k(t_1^k) - y_\alpha^k(t_1^k)]^2 \tag{15}$$

The identified model (11) has the required generalization performance if

$$e_r = \mathrm{CV} - \epsilon_r \leq 0 \tag{16}$$

where ϵ_r is a predefined generalization threshold and e_r denotes the generalization error.

C. Accuracy Criteria

Accuracy refers to how well the mathematical model that was found can fit the input-output data. In conventional system identification, time domain-based criterion is very commonly used as below

Criterion 3 (Accuracy in the time domain)**.** Let $\mathscr{E}(t)$ be the time-domain accuracy which can be defined as

$$\mathscr{E}(t_2) = \hat{y}_\beta(t_2) - y_\beta(t_2), t_2 = 1, ..., \Gamma_2 \tag{17}$$

The identified model (11) is said to be accurate if

$$e_t = \boldsymbol{\mathscr{E}}^{\mathrm{T}} \boldsymbol{\mathscr{E}} - \epsilon_t \leq 0 \quad \text{with} \quad \epsilon_t \to 0 \tag{18}$$

where $\boldsymbol{\mathscr{E}} = [\mathscr{E}(1), ..., \mathscr{E}(\Gamma_2)]^{\mathrm{T}}$. ϵ_t denotes a predefined accuracy threshold and e_t is the accuracy error.

However, **Criterion 3** often cannot guarantee the identified model captures the key information of a machining process over tooth passing frequencies in complicated scenarios. To solve this issue, a novel accuracy criterion is proposed to replace the conventional time-domain accuracy with frequency-domain accuracy that takes the tooth passing frequency related physical information into account. The new criterion is as follows.

Criterion 4 (Accuracy in the frequency domain)**.** Assume that the tooth passing frequency is located in the frequency range $[\omega_{\mathrm{low}}, \omega_{\mathrm{high}}]$, and the number of the discrete frequency components within this frequency range is N_t; therefore, the ςth harmonic of the tooth passing frequency is located in the frequency range $[\varsigma\omega_{\mathrm{low}}, \varsigma\omega_{\mathrm{high}}]$ where the number of the discrete frequency components is ςN_t. Based on the frequency ranges of the tooth passing frequency and its harmonics, the frequency-domain accuracy is defined as

$$\mathscr{E}_f(\omega^*) = \hat{Y}_\beta(\omega^*) - Y_\beta(\omega^*) \tag{19}$$

where $\omega^* \in [\omega_1^1, ..., \omega_1^{N_t}, ..., \omega_\varsigma^1, ..., \omega_\varsigma^{\varsigma N_t}]$ with $\omega_1^1 = \omega_{\mathrm{low}}$, $\omega_1^{N_t} = \omega_{\mathrm{high}}$, $\omega_\varsigma^1 = \varsigma\omega_{\mathrm{low}}$ and $\omega_\varsigma^{\varsigma N_t} = \varsigma\omega_{\mathrm{high}}$, and $\varsigma \in 1, ..., \Xi$ where the notation Ξ indicates the total number of the considered harmonics. Finally, the frequency-domain accuracy can be written in the matrix format in the frequency domain as

$$\boldsymbol{\mathscr{E}}_f = [\mathscr{E}_f(\omega_1^1), ..., \mathscr{E}_f(\omega_1^{N_t}), ..., \mathscr{E}_f(\omega_\varsigma^1), ..., \mathscr{E}_f(\omega_\varsigma^{\varsigma N_t})]^{\mathrm{T}} \quad (20)$$

where the length of the vector $\boldsymbol{\mathscr{E}}_f$ denoted as L_f can be derived as $L_f = N_t+, ..., +\varsigma N_t$. For this new criterion, the identified model (11) is said to be accurate if

$$e_f = \boldsymbol{\mathscr{E}}_f{}^{\mathrm{T}} \boldsymbol{\mathscr{E}}_f - \epsilon_f \leq 0 \quad \text{with} \quad \epsilon_f \to 0 \qquad (21)$$

where ϵ_f denotes a predefined frequency-domain accuracy threshold and e_f is the frequency-domain accuracy error.

Comparing **Criterion 3** and **Criterion 4**, **Criterion 4** integrates the physical information relating to the tooth passing frequency. It is the combination of **Criterion 1**, **Criterion 2** and **Criterion 4** that is referred to as TCMoR which can not only ensure that the identified NARX model has stability, generalization and accuracy in the time domain, but also ensure the NOFRFs extracted from the identified NARX model can well represent the dynamics of machining processes over tooth passing frequencies.

D. Implementation of Tool Condition Monitoring-oriented Regularization (TCMoR)

Based on TCMoR, the regularization parameter λ_g selection is determined by three criteria: 1) **Criterion 1** guarantees the model can be used for NOFRFs evaluation, 2) **Criterion 2** ensures the model is less sensitive to extreme data and 3) **Criterion 4** ensures the desired dynamics can be accurately modeled. Only when these three criteria are met at the same time, the established model can be used for TCM. To find an ideal λ_g that can meet these three criteria, the procedure of the regularization parameter selection is given as follows:

Step 1: To find the candidate regularization parameters λ_g which can satisfy (21) in **Criterion 4**, some algorithms based on exhaustive search, such as Randomized Search (RS), Evolutionary Algorithm (EA) or Genetic Algorithm (GA) [29] can be used to find a series of relatively optimal candidate regularization parameters denoted as $\boldsymbol{\Lambda} = [\lambda_1, ..., \lambda_Z]$ with $e_f(\lambda_z) \leq 0$ and $z = 1, ..., Z$.

Step 2: Rank the candidate regularization parameters in order from best to worst based on the performance of the accuracy error $e_f(\lambda_z)$. The ranked regularization parameters are denoted as $\boldsymbol{\Lambda^r} = [\lambda_1^r, ..., \lambda_Z^r]$ with $e_f(\lambda_1^r) \leq, ..., \leq e_f(\lambda_Z^r)$.

Step 3: Finally, these ranked parameters $\boldsymbol{\Lambda^r}$ are sequentially checked from λ_1^r to λ_Z^r by **Criterion 1** and **Criterion 2** until the one denoted as λ_ξ^r with $\xi \leq Z$ can meet both criteria. Therefore, the final selected regularization parameter is λ_ξ^r and the built model based on λ_ξ^r can be guaranteed to be stable and to provide a relatively optimal accuracy with a certain degree of generalization.

E. Summary of the Algorithm

In consequence, the final implementation of sensor data modelling with TCMoR can be summarized in Algorithm 1.

Algorithm 1 TCMoR-based sensor data modelling

Input: Input signals, **u**; output signals, **y**; output lag, n_y; input lag, n_u; polynomial degree ℓ.

Output: The identified NARX model.

1: Split the raw dataset into training and testing datasets.
2: Using FROLS to identify final model structure **W** by using the training dataset.
3: Define the accuracy threshold ϵ_f, stability threshold ϵ_s and generalization threshold ϵ_r.
4: Find a proper regularization parameter λ_g based on Step 1 to Step 3
5: Based on the selected λ_g, the regression weight vector **g** can be calculated according to (5), and the final identified NARX model can be obtained as (11).

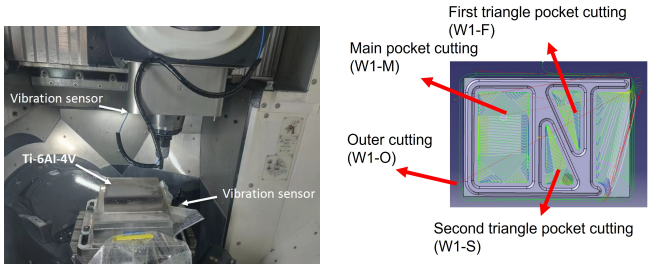
As can be seen, apart from the regularization parameter, there are several other critical parameters, which are input lag n_u, output lag n_y, polynomial degree ℓ, sampling frequency f_s, accuracy threshold ϵ_f, stability threshold ϵ_s and generalization threshold ϵ_r. In order to tune these parameters, the following methods are used.

1) n_u and n_y: Selecting appropriate values for n_u and n_y is crucial since a large value can provide more candidate model terms, including more useful information, but may also lead to the inclusion of redundant terms. The selection of n_u and n_y was made through trial and error in the present study. Based on Section III-D, a combination of n_u and n_y that can produce the smallest frequency-domain accuracy error is used for sensor data modelling.
2) ℓ*:* An ideal value of ℓ helps with determining appropriate nonlinear model terms. However, increasing ℓ can significantly increase the number of candidate model terms. For ease of use, ℓ is initially selected as 2, and it can then be incrementally adjusted until a satisfactory model is built.
3) f_s*:* An appropriate sampling frequency f_s is vital for generating an accurate NARX model that represents the dynamics of the milling process. As suggested by [30], f_s can be set to 10-20 times the tooth passing frequency.
4) ϵ_f*,* ϵ_s *and* ϵ_r*:* The accuracy threshold, stability threshold, and generalization threshold are crucial parameters for ensuring that the built model satisfies the accuracy, stability, and generalization criteria, respectively. It is advised that these thresholds are set to very small values for this purpose. In the context of our case studies, we suggest setting ϵ_f 1×10^{-1} for the accuracy threshold, ϵ_s 1×10^{-5} for the stability threshold, and ϵ_r 1×10^{-1} for the generalization threshold.

IV. Case Study 1: Variable Tool-Workpiece Engagement Conditions

A. Experiment Design

In this section, an *in-situ* experiment is designed at the Advanced Manufacturing Research Centre (AMRC), The University of Sheffield, Sheffield, U.K., where a DMU 40 evo universal machine tool, as shown in Fig.4(a), is used to conduct a run-to-failure test on cutting tools. The testing



(a) (b)

Fig. 4. (a) View of the experimental setup and (b) A CAD drawing of a complete toolpath.

cutting tools are OSG solid carbide end mills (UVX TI 6FL) which have six flutes, and the machining parameters used in the experiments are summarized in Table I. During the milling processes, a new cutting tool was used to create pocket features on the first titanium (Ti-6AI-4V) workpiece using the dynamic milling strategy; and this workpiece is denoted as W1. Fig.4(b) presents the computer-aided design (CAD) drawing which shows the complete toolpath when machining W1. As can be seen, the toolpath can be divided into four cutting processes which are outer cutting (denoted as W1-O), main pocket cutting (denoted as W1-M), first triangle pocket cutting (denoted as W1-F) and second triangle pocket cutting (denoted as W1-S). After the machining of the first workpiece was finished, a new cutting tool was used to cut the second workpiece (denoted as W2) with the same toolpath. The corresponding four cutting processes for the second workpiece are denoted as W2-O, W2-M, W2-F and W2-S which again represent outer cutting, main pocket cutting, first triangle pocket cutting and second triangle pocket cutting, respectively. Microscope images of each flute of the cutter were taken after completing each feature in order to measure the maximum flank wear of the profile of cutting tools. Table II summarizes the flank wear of each flute of the cutter. Based on the measured flank wear, the wear of cutting tools is categorized into 3 levels. If the flank wear of every flute is less than 0.300 mm, we define the level of tool wear as Level 1. If one of the cutter flute has wear between 0.300 to 0.400 mm, we define the level of tool wear as Level 2. If one of the cutter flute has a flank wear of more than 0.400 mm, we define the level of tool wear as Level 3. These levels of tool wear can represent different wear phases in a typical Tool Wear Curve (TWC) showing the relationship between the cutting distance and measured flank wear [31]. In a TWC, there are three well-defined wear phases which are i) initial wear phase, ii) steady state phase and iii) rapid wear phase. Level 1 corresponds to the initial wear and steady state stages. Level 2 relates to the rapid wear stage before tool failure. Level 3 refers to the rapid wear stage after tool failure. Table II also summarizes the three levels of cutting tool wear.

B. TCMoR-based Sensor Data Modelling

To conduct TCM, two PCB-type accelerometers (PCB 604B31) were mounted on the spindle and on the milling fixture beneath the workpiece, respectively. The sensitivity of the accelerometer is 100 mV/g and it has a constant frequency gain between 0.5 to 5 kHz. Vibration signals were collected by a data acquisition card (NI 9234) and the sampling rate was set to 51.2 kHz in order to cover the entire frequency range of the accelerometer. For each cutting process, 35 datasets were collected throughout the whole toolpath, and the time interval between each dataset was fixed. For the outer cutting, the time interval is 20 secs. For the main pocket cutting and triangle pocket cutting, the time interval is 10 secs and 5 secs, respectively. Furthermore, each dataset consists of two files that are one 1-second spindle vibration signal snapshot and one 1-second milling fixture vibration signal snapshot. Then, for each of the 35 datasets, sensor data modelling was applied to build a NARX model in order to find the dynamic relationship between the vibration signal collected from the spindle and the vibration signal collected from the milling fixture. The execution time for the proposed TCMoR-based sensor data modelling to identify a NARX model using one snapshot of sensor data is around 3 seconds [our tests were implemented using a desktop with a 12th Gen Intel(R) Core(TM) i9-12900k CPU and 64-GB memory]. As a result, for each machining process, 35 NARX models were produced from the 35 datasets in order to extract the NOFRFs features for TCM.

For TCMoR, the physical knowledge is the tooth passing frequency. Since the spindle rotation speed is 2586 RPM, the tooth passing frequency is 258.6 (= 2586/60 $\times$ 6) Hz. As a result, the frequency variables in (19) can be set to $\omega^* \in [250, 260] \cup [500, 520]$ Hz to include the tooth passing frequency and its harmonics for TCMoR. It is worth mentioning that, during the machining process, the tooth passing frequency is not strictly constant due to different machining strategies, cutting head deteriorations and fluctuations in spindle speed depending on the spindle motor control, and it is difficult to know the exact variation in advance. In Fig.5, the blue line is the first order NOFRF associated with the healthy tool where the dominant frequency component is 258.6 Hz which is the tooth passing frequency. In contrast, the red line is the first order NOFRF associated with the worn tool where the dominant frequency component has increased to 288.5 Hz which is a bit far away from the tooth passing frequency. The region demarcated by the dashed black rectangle represents a predetermined frequency range of [250, 260] Hz around the designed tooth passing frequency. From Fig.5, it can be observed that within the frequency range of [250, 260] Hz, the first order NOFRF associated with the healthy tool and the first order NOFRF associated with the worn tool are very different. This difference enables the differentiation between a health and worn tool.

To find an ideal λ_g that can meet **Criterion 1**, **Criterion 2** and **Criterion 4** introduced in Section III, the accuracy threshold ϵ_f, stability threshold ϵ_s and generalization threshold ϵ_r are set to 1×10^{-2}, 1×10^{-4} and 10, respectively; and the EA-based algorithm is used to search for a relatively optimal λ_g. More details of the EA-based algorithm have been introduced in [29].

Table I
Summary of the machining parameters.

Number of flutes	Axial depth of cut (mm)	Radial depth of cut (mm)	Feed rate (mm/min)	Spindle speed (rpm)	Chip thickness hex (mm)	Cutting Speed (m/min)	Coolant delivery
6	19.5	0.15	1738.32	2586	0.08	130	Flood

Table II
Flank wear of each cutter flute.

Workpiece 1	Flank wear (mm)						
	Flute 1	Flute 2	Flute 3	Flute 4	Flute 5	Flute 6	Wear level
W1-O	0.115	0.099	0.103	0.128	0.119	0.119	Level 1
W1-M	0.175	0.197	0.164	0.314	0.437	0.318	Level 2
W1-F	0.212	0.208	0.233	0.312	0.486	0.459	Level 3
W1-S	0.257	0.255	0.274	0.314	0.483	0.526	Level 3
Workpiece 2	Flank wear (mm)						
	Flute 1	Flute 2	Flute 3	Flute 4	Flute 5	Flute 6	Wear level
W2-O	0.128	0.126	0.106	0.122	0.158	0.192	Level 1
W2-M	0.133	0.126	0.136	0.125	0.194	0.301	Level 2
W2-F	0.121	0.169	0.134	0.138	0.189	0.302	Level 2
W2-S	0.155	0.170	0.135	0.142	0.260	0.302	Level 2

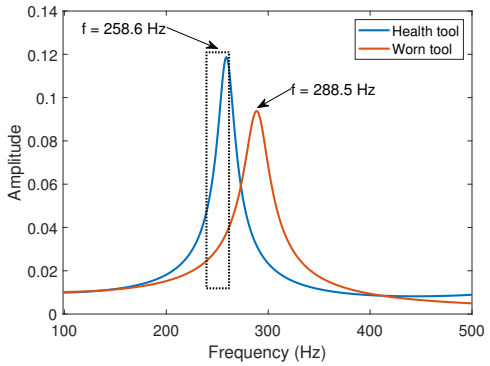


Fig. 5. First order NOFRFs of health and worn tools.

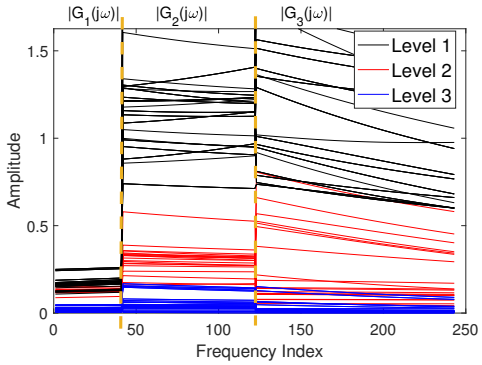


Fig. 6. NOFRFs-based model features of tool wear Level 1, Level 2 and Level 3.

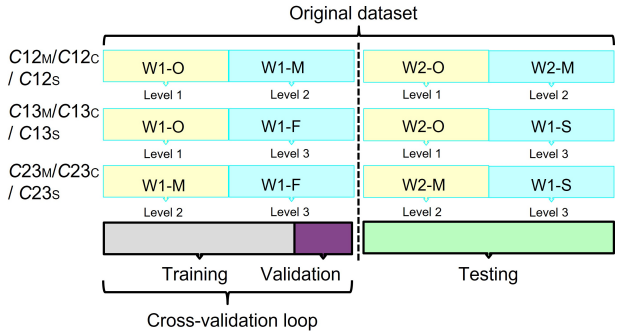


Fig. 7. Design of the diagnostic procedure.

C. Model Frequency Analysis

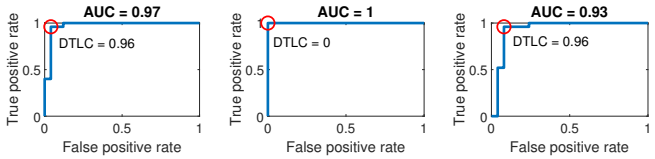
After TCMoR-based sensor data modelling, the next step is to conduct model frequency analysis via NOFRF evaluations. As the tooth passing frequency is 258.6 Hz, in order to observe the frequency domain properties at this frequency, we designed an input excitation with a frequency range from 250 to 260 Hz to evaluate the NOFRFs of each model. The input excitation signal $u^*(t)$ is designed as $u^*(t) = 3 \cdot [\sin(2 \cdot 260 \cdot \pi \cdot t) - \sin(2 \cdot 250 \cdot \pi \cdot t)]/2 \cdot \pi \cdot t$ where $-2 \leq t \leq 2$. Based on Eqs. (9) and (10), the NOFRFs of each identified model were determined. For this case study, the first, second and third-order NOFRFs were extracted as model features. The execution time for the evaluation of the NOFRFs of the identified NARX model takes around 1 second.

Fig.6 shows an experimental evidence of the NOFRFs-based model features under three different levels of tool wear (Levels 1, 2 and 3 introduced in Table II). The NOFRFs amplitudes under tool wear Level 1 (black lines) are mainly concentrated in the range from 0.15 to 2; under tool wear level 2 (red lines), the NOFRFs amplitudes are between 0.1 and 1; while under tool wear level 3 (blue lines), the NOFRFs amplitudes are limited to 0.18. These observations demonstrate that the NOFRFs-based frequency analysis can be effectively exploited to monitor the conditions of cutting tools.

D. Model Feature-based TCM

To conduct TCM, a classification approach was applied based on the extracted NOFRFs-based model features. Fig.7 displays the diagnostic procedure. As can be seen, three binary classification processes were designed in order to classify different levels of tool wear. For example, the classifier denoted as $C12_{\mathrm{M}}$ was designed to classify tool wear Level 1 and Level 2. The training and validation datasets were from W1-O and W1-M and the testing dataset was from W2-O and W2-M. In the same way, such nomenclature also applies to $C13_{\mathrm{M}}$ and $C23_{\mathrm{M}}$. It is worth mentioning that, in the present study, three binary classifiers were used instead of a single multiclass classifier. This is because the binary classifier can output Receiver Operating Characteristics (ROC) curves, within which the threshold and Area Under the Curve (AUC) can be analyzed in order compare the effectiveness of the developed NOFRFs-based model features with conventional features. On the contrary, a single multiclass classifier can only output a confusion matrix, making it difficult to evaluate the feature performance in detail.

During the training and validation stage shown in Fig.7, k-fold cross-validation was used, and we used a support



(a) (b) (c)

Fig. 8. The classification results on validation datasets by using the TCMoR-based method of the classifier (a) $C12_{\mathrm{M}}$, (b) $C13_{\mathrm{M}}$, (c) $C23_{\mathrm{M}}$.

vector machine (SVM) for classification training since it is a powerful classification approach in relevant applications. To quantify the performance of the classifier on the validation datasets, the ROC-AUC curve is utilized. If the ROC is close to the top-left corner with the coordinates $O(0,1)$, then the trained classifier has a good performance. Furthermore, as the trained SVM classifier has a particular threshold on the ROC curve, to quantify this trained classifier, an indicator termed as Distance to Top Left Corner (DTLC), is proposed [32], [33]. If DTLC is close to 0, then the classifier has an overall good classification performance. Finally, the combination of ROC-AUC and DTLC allows the performance of the classifier to be evaluated from several perspectives including quality and accuracy. Fig.8 shows the ROC-AUC curves of validation datasets from $C12_{\mathrm{M}}$, $C13_{\mathrm{M}}$ and $C23_{\mathrm{M}}$ where their AUCs are 0.97, 1 and 0.93, respectively. Furthermore, the red circle on the ROC curve corresponds to the trained SVM classifier. For example, the red circle on Fig.8(a) has the coordinates of $\mathbb{T}(0.04, 0.96)$, so the DTLC of the trained classifier is $0.96(\approx \sqrt{0.04^2 + 0.96^2})$.

In the training and validation stage, the classification training was repeated 5 times, so the mean AUC and mean DTLC can be recorded in the validation results; and Table III summarizes the validation results of $C12_{\mathrm{M}}$, $C13_{\mathrm{M}}$ and $C23_{\mathrm{M}}$.

Finally, After the SVM classifier was trained off-line by using the whole training and validation datasets, it was then utilized to predict the testing dataset. The prediction results including ROC-AUC and DTLC of the classifiers $C12_{\mathrm{M}}$, $C13_{\mathrm{M}}$ and $C23_{\mathrm{M}}$ on testing datasets are presented in Fig.9(a) to Fig.9(c), and summarized in Table III.

E. Comparative Study

1) *CR-based method:* To demonstrate the advantages of the proposed TCMoR method, some comparisons with the CR method were carried out where **Criterion 1**, **Criterion 2** and **Criterion 3** were taken into account during the sensor data modelling stage. After using the CR-based sensor data modelling, the corresponding NOFRFs-based model features were extracted for TCM. For the diagnostic procedure, as presented in Fig.7, the classifiers $C12_{\mathrm{C}}$, $C13_{\mathrm{C}}$ and $C23_{\mathrm{C}}$, were designed respectively to classify different tool wear levels. The results on validation datasets and testing datasets were summarized in Table III and presented in Fig.9(d) to Fig.9(f). As can be seen, the results on the training/validation datasets may perform well but the results on the testing dataset are unsatisfactory. This is because there are some variations between the training/validation datasets and the testing dataset due to different machining strategies, noise and some unknown reasons. Consequently, the extracted NOFRFs-based model features may fail to reveal the true cutting tool states. However, when applying the TCMoR-based method, the physical information of the machining process is incorporated into the modelling procedure. As a result, the TCM performance has been improved.

2) *Signal feature-based method:* An alternative yet an appealing method for TCM is the signal feature-based approach which extracts the candidate signal features from the collected datasets; and the significant signal features are selected from the candidate features in order to train the classifier for TCM [34]. To classify different tool wear levels, as presented in Fig.7, the classifiers $C12_{\mathrm{S}}$, $C13_{\mathrm{S}}$ and $C23_{\mathrm{S}}$, were designed respectively. The results of the signal feature-based approach were presented in Table III and demonstrated in Fig.9(g) to Fig.9(i). Compared with the model feature-based method using TCMoR shown in Fig.9(a) to Fig.9(c), the signal feature-based method often fails to identify the cutting tool states especially when the machining strategies are complicated. Furthermore, for different machining processes, signal features have to be redesigned increasing the difficulty of the feature extraction. One of the unique contributions of the model feature-based method is that the features are always the NOFRFs which can be fixed when applied to TCM of different machining processes.

3) *Linear model-based method:* The linear ARX model was used as a comparison to represent the dynamic relationship between $u(t)$ and $y(t)$ collected from the vibration sensors mounted on the spindle and milling fixture,

Table III
Case study 1: summary of the results under variable tool-workpiece engagement conditions by using different TCM methods.

Classifier	Cross-validation		Testing		Accuracy[1]	Specificity[2]	Recall[3]
$C12_{\mathrm{M}}$	AUC	0.94	AUC	0.74	84.00%	72.00%	96.00%
	DTLC	0.16	DTLC	0.28			
$C13_{\mathrm{M}}$	AUC	0.96	AUC	**1**	**100.00%**	100.00%	100.00%
	DTLC	0.14	DTLC	**0**			
$C23_{\mathrm{M}}$	AUC	0.92	AUC	**1**	**98.00%**	96.00%	100.00%
	DTLC	0.22	DTLC	**0.04**			
$C12_{\mathrm{C}}$	AUC	1	AUC	0.50	52.00%	4.00%	100.00%
	DTLC	0.12	DTLC	0.96			
$C13_{\mathrm{C}}$	AUC	1	AUC	1	52.00%	4.00%	100.00%
	DTLC	0	DTLC	0.96			
$C23_{\mathrm{C}}$	AUC	0.48	AUC	0.77	70.00%	56.00%	84.00%
	DTLC	0.67	DTLC	0.47			
$C12_{\mathrm{S}}$	AUC	0.98	AUC	**1**	**98.00%**	100.00%	96.00%
	DTLC	0.16	DTLC	**0.04**			
$C13_{\mathrm{S}}$	AUC	1	AUC	0.72	52.00%	100.00%	40.00%
	DTLC	0	DTLC	0.96			
$C23_{\mathrm{S}}$	AUC	0.94	AUC	0.85	80.00%	64.60%	86.00%
	DTLC	0.12	DTLC	0.36			
$C12_{\mathrm{L}}$	AUC	0.93	AUC	0.85	66.00%	36.00%	96.00%
	DTLC	0.12	DTLC	0.64			
$C13_{\mathrm{L}}$	AUC	0.97	AUC	1	86.00%	72.00%	100.00%
	DTLC	0.96	DTLC	0.28			
$C23_{\mathrm{L}}$	AUC	0.89	AUC	0.98	96.00%	96.00%	96.00%
	DTLC	0.89	DTLC	0.06			

[1] Accuracy = (TP+TN)/(TP+FP+TN+FN)
[2] Specificity = 1 - FPR = TN/(FP+TN)
[3] Recall = TPR = TP/(TP+FN)

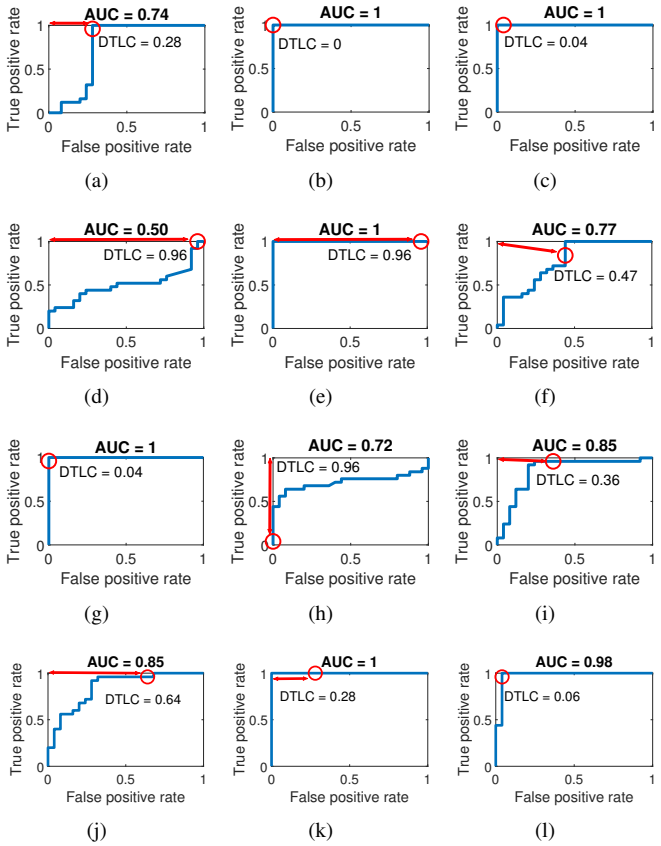


Fig. 9. The classification results on testing datasets by using the TCMoR-based method produced the classifier (a) $C12_{\mathrm{M}}$, (b) $C13_{\mathrm{M}}$, (c) $C23_{\mathrm{M}}$. The classification results on testing datasets by using the CR-based method produced classifier (d) $C12_{\mathrm{C}}$, (e) $C13_{\mathrm{C}}$ and (f) $C23_{\mathrm{C}}$. The classification results on testing datasets by using the signal feature-based method produced classifier (g) $C12_{\mathrm{S}}$, (h) $C13_{\mathrm{S}}$ and (i) $C23_{\mathrm{S}}$. The classification results on testing datasets by using the linear model-based method produced classifier (g) $C12_{\mathrm{L}}$, (h) $C13_{\mathrm{L}}$ and (i) $C23_{\mathrm{L}}$.

respectively; and the frequency response function (FRF) was extracted as model features to conduct TCM. The diagnostic procedure is shown in Fig.7, and the classifiers $C12_{\mathrm{L}}$, $C13_{\mathrm{L}}$ and $C23_{\mathrm{L}}$, were designed respectively to classify different tool wear levels. The diagnostic results are presented in Fig.9(j) to Fig.9(l), and summarized in Table.III. As can be seen, the linear model-based method has a lower accuracy rate. This is because the linear model cannot capture the nonlinear dynamics induced by the tool wear.

V. Case Study 2: Controlled Tool-Workpiece Engagement Conditions

The second case study is about controlled engagement milling, which was designed to cut the workpieces into different regular shapes. As can be seen in Fig.10(a), the workpiece was stratified into three discrete layers. The top layer was used to machine a round-shaped feature, the second layer was used to machine a diamond-shaped feature, and the third layer was used to machine a square-shaped feature. Lastly, the straight cutting milling was conducted where the cutting tool was used to cut straight lines.

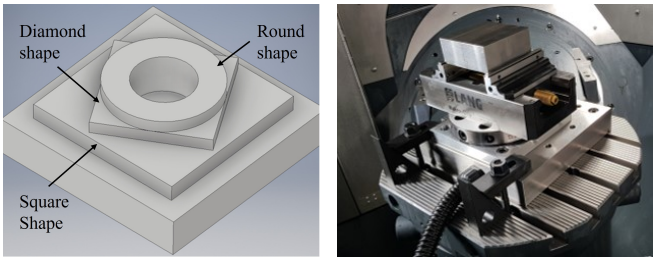


Fig. 10. (a) A CAD drawing of machining features and (b) view of the experimental setup.

For this case study, the machining operation was also performed using a DMU 40 evo machine, and, as presented in Fig.10(b), titanium (Ti-6Al-4V) workpieces with dimensions of 100 mm $\times$ 100 mm $\times$ 40 mm were used to conduct the experiment. Throughout the machining operation, the spindle operated at a speed of 300 rpm, and the cutter was equipped with 7 flutes. Consequently, the tooth passing frequency is computed to be 35 Hz, derived from the formula $= 300/60 \times 7$. Furthermore, the utilized accelerometers, data acquisition card, sampling rate are the same with the previous cast study.

During the experiments, the new cutting tools and worn cutting tools were used, respectively, to cut the workpieces of these regular shapes. The datasets collected from new cutting tools were denoted as Baseline and the datasets collected from worn cutting tools were denoted as Toolwear. For each cutting process of these regular-shaped features, 36 datasets were collected throughout the whole toolpath meaning that 36 NARX models were produced from the 36 datasets with the aim of extracting the NOFRFs features for TCM. For TCMoR, since the spindle speed is 35 Hz, the frequency variables in Eq.(19) were set to $\omega^* \in [30, 40] \cup [60, 80]$ Hz to include the tooth passing frequency and its harmonics for TCMoR.

The diagnostic procedure for TCM is similar to the previous case study where the original datasets from Baseline and Toolwear were divided into training, validation and testing datasets. Then, the TCMoR-based method, CR-based method, signal feature-based method and FRF-based method were used, respectively, in order to extensively compare the proposed method with other existing TCM techniques. The classification results by using different methods are summarized in Table IV. As can be seen, the proposed method achieves the highest accuracy rate for square-shaped, round-shaped and straight milling. Furthermore, the proposed method can produce the highest AUC and lowest DTLC for three of the four controlled tool-workpiece engagement machining processes. As a result, the proposed method performs better in terms of AUC, DTLC and accuracy, which demonstrates the superiority of the newly proposed TCMoR-based sensor data modelling and model frequency analysis over existing TCM methods.

VI. Discussion

In order to illustrate the strength and weakness of the proposed diagnostic approach, a discussion is presented below.

Table IV
Case study 2: summary of the results under controlled tool-workpiece engagement conditions by using different TCM methods.

Shapes	Methods			Accuracy	Specificity	Recall
Square	TCMoR	AUC	**0.79**	**76.0%**	68.6%	83.3%
		DTLC	**0.28**			
	CR	AUC	0.61	58.2%	78.1%	36.3%
		DTLC	0.65			
	Signal feature	AUC	0.75	50.0%	100.0%	0%
		DTLC	1			
	FRF	AUC	0.74	65.6%	42.5%	88.8%
		DTLC	0.59			
Diamond	TCMoR	AUC	**0.80**	72.9%	65.6%	80.2%
		DTLC	0.39			
	CR	AUC	0.79	**77.1%**	79.2%	75.0%
		DTLC	**0.33**			
	Signal feature	AUC	0.78	68.8%	70.8%	66.7%
		DTLC	0.44			
	FRF	AUC	0.69	58.3%	69.4%	47.2%
		DTLC	0.61			
Round	TCMoR	AUC	0.70	**72.9%**	62.5%	83.3%
		DTLC	**0.40**			
	CR	AUC	**0.79**	68.6%	70.8%	66.7%
		DTLC	0.44			
	Signal feature	AUC	0.6	50.0%	100%	0%
		DTLC	1			
	FRF	AUC	0.50	52.00%	4.00%	100%
		DTLC	0.96			
Straight	TCMoR	AUC	**0.99**	**99.0%**	99.0%	99.0%
		DTLC	**0.01**			
	CR	AUC	0.80	75.5%	74.8%	76.2%
		DTLC	0.35			
	Signal feature	AUC	0.98	84.8%	71.0%	98.6%
		DTLC	0.29			

1) *Time consumption:* The proposed TCMoR is a regularization method. In this respect, the performance of the TCMoR including execution time is the same as that of a CR-based method. The fundamental difference between the TCMoR and a CR is the way by which the regularization parameter λ_g is determined. When the TCMoR is applied, the regularization parameter λ_g is determined to make sure the identified NARX model can sufficiently capture the dynamics of the underlying machining process so that the NOFRFs of the identified NARX model can be used for TCM.
   The signal feature-based methods require a feature selection process in the offline training stage, which is time-consuming. On the contrary, for the proposed method, the features are always NOFRF features evaluated from the identified NARX models. Therefore, the proposed method is more efficient in terms of feature extraction in the training stage.
2) *Limitation:* The proposed framework has several tunable parameters, such as input lag n_u, output lag n_y, and maximum degree of nonlinearity ℓ, in the sensor data modelling stage.However, it is difficult to preset them in advance. Thus, a more effective algorithm is needed to automatically tune these parameters, which will be explored in future research.

VII. Conclusion

In the present study, a new diagnostic framework based on sensor data modelling and model frequency analysis is proposed for tool condition monitoring (TCM) during machining. Under this framework, for the first time, the physical knowledge of the underlying machining process is incorporated into the modelling procedure to design the associated regularization parameter. This is known as the tool condition monitoring-oriented regularization (TCMoR)-based sensor data modelling and model frequency analysis. The application of the TCMoR-based sensor data modelling and model frequency analysis to TCM has been investigated by comprehensive machining experimental studies. Compared with the existing sensor data modelling and model frequency analysis-based method and conventional signal feature-based approaches, the proposed method can better exploit the physics underlying monitored milling processes and has the advantage of not needing to design different features for different milling processes. The results of the experimental studies have demonstrated the effectiveness of the proposed method especially in the scenarios when conducting TCM for complicated milling processes including variable tool-workpiece engagement conditions.

The primary limitation associated with the proposed method relates to the computational speed of the NARX modelling. Additionally, there is a challenge in implementing the proposed method for real-time TCM in industrial-scale settings. As a result, future works will focus on the development of a real-time implementation approach to the TCMoR-based method. The aim is to apply the approach to perform real-time TCM in industrial-scale machining tests and evaluate the performance of the new diagnostic framework for TCM in industry application-oriented scenarios.

Appendix

Consider rewriting the built NARX model (6) into the following discrete-time equation,

$$\begin{aligned} y(t) = &\sum_{j=1}^{J}\sum_{p=0}^{j}\sum_{l_1,\ldots,l_{p+q}=1}^{L} c_{p,q}(l_1,\ldots,l_{p+q}) \\ &\times \prod_{i=1}^{p} y(t-l_i) \prod_{i=p+1}^{p+q} u(t-l_i) \end{aligned} \tag{22}$$

where J and L are integers, $p+q=j$, and $c_{p,q}(l_1,\ldots,l_{p+q})$ denotes the coefficients of the model [35]. The GALEs of the NARX model (22) can be described as [36]

$$\begin{aligned} y_n(t) = &\sum_{l_1=1}^{L} c_{1,0}(l_1) y_n(t-l_1) + \sum_{l_1,l_n=1}^{L} c_{0,n}(l_1,.,l_n) \\ &\times \prod_{i=1}^{n} u(t-l_i) + \sum_{q=1}^{n-1}\sum_{p=1}^{n-q}\sum_{t_1,t_{p+q}=1}^{L} c_{p,q}(l_1,\ldots,l_{p+q}) y^{\mathbf{L}}_{n-q,p}(t) \\ &\times \prod_{i=p+1}^{p+q} u(t-l_i) + \sum_{p=2}^{n}\sum_{l_1,l_p=1}^{L} c_{p,0,k}(l_1,\ldots,l_p) y^{\mathbf{L}}_{n,p,k}(t) \end{aligned} \tag{23}$$

where $n=1,\ldots,N$, $\mathbf{L}=[l_1,\ldots,l_n]$ and

$$\begin{cases} y^{\mathbf{L}}_{n,p}(t) = \sum\limits_{i=1}^{n-(p-1)} y_i\left(t-l_p\right) y^{\mathbf{L}_p}_{n-i,p-1}(t) \\ y^{\mathbf{L}}_{n,1}(t) = y_n\left(t-l_1\right) \end{cases} \tag{24}$$

REFERENCES

[1] T.-K. S. Schmitz, "Machining dynamics frequency response to improved productivity," 2009.

[2] Q. Zhu, B. Sun, Y. Zhou, W. Sun, and J. Xiang, "Sample augmentation for intelligent milling tool wear condition monitoring using numerical simulation and generative adversarial network," *IEEE Transactions on Instrumentation and Measurement*, vol. 70, pp. 1–10, 2021.

[3] C. Koulamas, B. Lambert, and M. Smith, "Optimal machining conditions and buffer space size for the two-stage case," *International Journal of Production Research*, vol. 25, no. 3, pp. 327–336, 1987.

[4] J. Taylor, "Carbide cutting tool variance and breakage: unknown factors in machining economics," in *Advances in Machine Tool Design and Research 1967*, pp. 487–504. Elsevier, 1968.

[5] Y. Li, X. Wang, Y. He, Y. Wang, Y. Wang, and S. Wang, "Deep spatial-temporal feature extraction and lightweight feature fusion for tool condition monitoring," *IEEE Transactions on Industrial Electronics*, vol. 69, no. 7, pp. 7349–7359, 2022.

[6] B. Lutz, L. Janisch, D. Kisskalt, D. Regulin, and J. Franke, "Interactive image segmentation using superpixels and deep metric learning for tool condition monitoring," *Procedia CIRP*, vol. 118, pp. 459–464, 2023, 16th CIRP Conference on Intelligent Computation in Manufacturing Engineering.

[7] Z. You, H. Gao, S. Li, L. Guo, Y. Liu, and J. Li, "Multiple activation functions and data augmentation-based lightweight network for in situ tool condition monitoring," *IEEE Transactions on Industrial Electronics*, vol. 69, no. 12, pp. 13 656–13 664, 2022.

[8] H. Miao, Z. Zhao, C. Sun, B. Li, and R. Yan, "A u-net-based approach for tool wear area detection and identification," *IEEE Transactions on Instrumentation and Measurement*, vol. 70, pp. 1–10, 2021.

[9] Z. You, S. Li, C. Li, H. Gao, L. Guo, and Y. Liu, "A novel evaluation metric based on dispersion of wear distance for in situ tool condition monitoring," *IEEE Transactions on Instrumentation and Measurement*, vol. 72, pp. 1–10, 2023.

[10] A. Zhu, D. He, J. Zhao, W. Luo, and W. Chen, "3d wear area reconstruction of grinding wheel by frequency-domain fusion," *The International Journal of Advanced Manufacturing Technology*, vol. 88, pp. 1111–1117, 2017.

[11] W. Wang, Y. Wong, and G. Hong, "3d measurement of crater wear by phase shifting method," *Wear*, vol. 261, no. 2, pp. 164–171, 2006.

[12] Y. Li, X. Wang, Y. He, Y. Wang, Y. Wang, and S. Wang, "Deep spatial-temporal feature extraction and lightweight feature fusion for tool condition monitoring," *IEEE Transactions on Industrial Electronics*, vol. 69, no. 7, pp. 7349–7359, 2022.

[13] H. Guo, X. Lin, and K. Zhu, "Pyramid lstm network for tool condition monitoring," *IEEE Transactions on Instrumentation and Measurement*, vol. 71, pp. 1–11, 2022.

[14] L. Huchel, T. C. Krause, T. Lugowski, S. B. Leeb, and J. Helsen, "Chasing the cut: A measurement approach for machine tool condition monitoring," *IEEE Transactions on Instrumentation and Measurement*, vol. 70, pp. 1–10, 2021.

[15] B. Chen, X. Chen, B. Li, Z. He, H. Cao, and G. Cai, "Reliability estimation for cutting tools based on logistic regression model using vibration signals," *Mechanical Systems and Signal Processing*, vol. 25, no. 7, pp. 2526–2537, 2011.

[16] D. Kong, Y. Chen, and N. Li, "Gaussian process regression for tool wear prediction," *Mechanical systems and signal processing*, vol. 104, pp. 556–574, 2018.

[17] G. Wang and F. Zhang, "A sequence-to-sequence model with attention and monotonicity loss for tool wear monitoring and prediction," *IEEE Transactions on Instrumentation and Measurement*, vol. 70, pp. 1–11, 2021.

[18] J. Wang, X. Cheng, Y. Gao, X. Wang, and J. Yang, "Cutting force embedded manifold learning for condition monitoring of vertical machining center," *IEEE Transactions on Instrumentation and Measurement*, vol. 71, pp. 1–12, 2022.

[19] T. Liu and K. Zhu, "A switching hidden semi-markov model for degradation process and its application to time-varying tool wear monitoring," *IEEE Transactions on Industrial Informatics*, vol. 17, no. 4, pp. 2621–2631, 2021.

[20] Y. Gui, Z.-Q. Lang, Z. Liu, Y. Zhu, H. Laalej, and D. Curtis, "Unsupervised detection of tool breakage: A novel approach based on time and sensor domain data analysis," *IEEE Transactions on Instrumentation and Measurement*, pp. 1–1, 2023.

[21] Renishawplc, "Nc4." [Online]. Available: https://www.renishaw.com/en/nc4-fixed-non-contact-tool-setting-probe--15147

[22] Y.-P. Zhu, Z. Q. Lang, and H. Laalej, "Data driven evaluation of nonlinear output frequency response functions with applications to structural system fault diagnosis," in *2019 1st International Conference on Industrial Artificial Intelligence (IAI)*, pp. 1–6, 2019.

[23] Z. Liu, Z. Q. Lang, Y. Zhu, Y. Gui, H. Laalej, and J. Stammers, "Sensor data modelling and model frequency analysis for detecting cutting tool anomalies in machining," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, pp. 1–1.

[24] G. Pillonetto, T. Chen, A. Chiuso, G. De Nicolao, and L. Ljung, *Regularized system identification: Learning dynamic models from data*. Springer Nature, 2022.

[25] K. Cheng, *Machining dynamics: fundamentals, applications and practices*. Springer Science & Business Media, 2008.

[26] S. A. Billings, *Nonlinear system identification: NARMAX methods in the time, frequency, and spatio-temporal domains*. John Wiley & Sons, 2013.

[27] Z. Lang and S. Billings, "Energy transfer properties of non-linear systems in the frequency domain," *International Journal of Control*, vol. 78, no. 5, pp. 345–362, 2005.

[28] Z. Peng, Z. Q. Lang, and S. A. Billings, "Non-linear output frequency response functions for multi-input non-linear volterra systems," *International Journal of Control*, vol. 80, no. 6, pp. 843–855, 2007.

[29] A. Kadochnikova, Y. Zhu, Z.-Q. Lang, and V. Kadirkamanathan, "Integrated identification of the nonlinear autoregressive models with exogenous inputs (narx) for engineering systems design," *IEEE Transactions on Control Systems Technology*, vol. 31, no. 1, pp. 394–401, 2023.

[30] S. Billings and L. A. Aguirre, "Effects of the sampling time on the dynamics and identification of nonlinear models," *International journal of Bifurcation and Chaos*, vol. 5, no. 06, pp. 1541–1556, 1995.

[31] L. Alhadeff, M. Marshall, D. Curtis, and T. Slatter, "Applying experimental micro-tool wear measurement techniques to industrial environments," *Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture*, vol. 235, no. 10, pp. 1588–1601, 2021.

[32] F. Habibzadeh, P. Habibzadeh, and M. Yadollahie, "On determining the most appropriate test cut-off value: the case of tests with continuous results," *Biochemia medica*, vol. 26, no. 3, pp. 297–307, 2016.

[33] I. Unal, "Defining an optimal cut-point value in roc analysis: an alternative approach," *Computational and mathematical methods in medicine*, vol. 2017, 2017.

[34] X. Qin, W. Huang, X. Wang, Z. Tang, and Z. Liu, "Real-time remaining useful life prediction of cutting tools using sparse augmented lagrangian analysis and gaussian process regression," *Sensors*, vol. 23, no. 1, p. 413, 2022.

[35] S. Billings and J. Peyton Jones, "Mapping non-linear integro-differential equations into the frequency domain," *International Journal of Control*, vol. 52, no. 4, pp. 863–879, 1990.

[36] Y.-P. Zhu, Z. Q. Lang, H. Mao, and H. Laalej, "Nonlinear output frequency response functions: A new evaluation approach and applications to railway and manufacturing systems' condition monitoring," *Mechanical Systems and Signal Processing*, vol. 163, p. 108179, 2022.

Zepeng Liu (Member, IEEE) received the B.Eng. degree in electrical engineering and electronics from the University of Liverpool, Liverpool, U.K., in 2015, the M.S. degree in power systems engineering from University College London, London, U.K., in 2016, and the Ph.D. degree in electrical and electronic engineering from the University of Manchester, Manchester, U.K, in 2021. From 2021 to 2023, he was a Research Associate with the Department of Automatic Control and Systems Engineering, University of Sheffield, Sheffield, U.K.

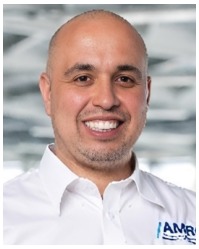
He is currently a Lecturer with the School of Engineering, Newcastle University, Newcastle upon Tyne, U.K. His research interests include complex system modelling, system identification, frequency analysis, time-frequency representation, sparse representation and nonlinear filtering, and data-driven analysis. The application areas include condition monitoring and fault diagnosis for wind turbine components and systems, digital manufacturing, and structural health monitoring.

Yufei Gui received the B.Eng. degree in Mechanical Design, Manufacturing and Automation from Tianjin University, Tianjin, China, in 2017, and the M.S. degree in Mechanical Engineering from Shanghai Jiao Tong University, Shanghai, China, in 2020. He is currently a Ph.D. student with the Department of Automatic Control and System Engineering, University of Sheffield, Sheffield, U.K.

His research interests include system identification, signal processing, data-driven condition monitoring and fault diagnosis for manufacturing processes.

Yun-Peng Zhu (Member, IEEE) received the B.S. degree in Mechanical Engineering and Automation in 2013, and the M.S. degree in Mechanical Manufacturing and Automation in 2015, all from the Northeastern University, China. He received the Ph.D. degree from the Department of Automatic Control and Systems Engineering, the University of Sheffield, U.K in 2020.

He was a Research Associate in the Department of Automatic Control and System Engineering, the University of Sheffield. He is currently an assistant professor in the School of Engineering and Material Science at Queen Mary University of London. His research interests include data driven dynamical systems, nonlinear dynamics, system identification, etc.



Hatim Laalej received an MEng degree in Computer Systems Engineering in 2007 and a Ph.D. in Control Systems in 2010 for his research on nonlinear vibration control systems of civil and mechanical structures at the department of Automatic Control and Systems Engineering, University of Sheffield.

He is currently a lead engineer in the process monitoring and control group at the University of Sheffield, Advanced Manufacturing Research Centre (AMRC). His research areas include condition monitoring and fault diagnosis for machine tools and cutting tools, digital manufacturing, vibration control systems and adaptive manufacturing. His research interests include sensor fusion systems, instrumentation, signal processing, data acquisition systems, control systems, vibration analysis, computational intelligence, and machine connectivity.

Zi-Qiang Lang received B.S. degree in automatic control in 1984 and M.Sc. degree in industrial automation in 1987 from Shenyang and Northeastern Universities in China, respectively. He received Ph.D. degree in systems and control engineering from University of Sheffield, U.K., in 1997.

He is currently a Chair Professor in Complex Systems Analysis and Design with the Department of Automatic Control and Systems Engineering, University of Sheffield. His research interests include theories and methods for complex systems modelling, analysis, design, signal processing, and the application of these to resolving various scientific and engineering problems. The application areas include smart structures and systems, civil and mechanical structure vibration control, structural health monitoring, condition monitoring and fault diagnosis for wind turbine components and systems, digital manufacturing, and medical diagnosis.

David Curtis has completed a MEng and PhD in Mechanical Engineering at the University of Birmingham, UK. His research concerned the application of point grinding technology for the production of aeroengine disc features.

He is currently a Professor of Subtractive Manufacturing, Research and Technology Manager at the University of Sheffield AMRC, and Co-director of the IDC in Machining Science. His research areas include Abrasive Machining, Emerging Machining Technology, Machining Science, Aerofoil and Gear Manufacturing Technology. He also leads activity within the AMRC on the machining of Ceramic Matrix Composite (CMC) materials, focusing on translational research which takes Machining Science fundamentals into an industrial context. He also plays an active role within the High Value Manufacturing Catapults (HVMC) technical community where he leads on Subtractive Manufacturing capability.