Remaining useful lifetime prediction for milling blades using a fused data prediction model (FDPM)

Teemu Mäkiaho[1] · Jouko Laitinen[1] · Mikael Nuutila[2] · Kari T. Koskinen[1]

Abstract

In various industry sectors, predicting the real-life availability of milling applications poses a significant challenge. This challenge arises from the need to prevent inefficient blade resource utilization and the risk of machine breakdowns due to natural wear. To ensure timely and accurate adjustments to milling processes based on the machine's cutting blade condition without disrupting ongoing production, we introduce the Fused Data Prediction Model (FDPM), a novel temporal hybrid prediction model. The FDPM combines the static and dynamic features of the machines to generate simulated outputs, including average cutting force, material removal rate, and peripheral milling machine torque. These outputs are correlated with real blade wear measurements, creating a simulation model that provides insights into predicting the wear progression in the machine when associated with real machine operational parameters. The FDPM also considers data preprocessing, reducing the dimensional space to an advanced recurrent neural network prediction algorithm for forecasting blade wear levels in milling. The validation of the physics-based simulation model indicates the highest fidelity in replicating wear progression with the average cutting force variable, demonstrating an average relative error of 2.38% when compared to the measured mean of rake wear during the milling cycle. These findings illustrate the effectiveness of the FDPM approach, showcasing an impressive prediction accuracy exceeding 93% when the model is trained with only 50% of the available data. These results highlight the potential of the FDPM model as a robust and versatile method for assessing wear levels in milling operations precisely, without disrupting ongoing production.

List of symbols

0-P	Zero-to-peak vibration [mm/s]
A_e	Radial depth of cut [mm]
A_p	Axial depth of cut [mm]
α	Contact angle [degrees]
BiLSTM	Bidirectional long short-term memory
CNN	Convolutional neural network
dF_a	Tangential force [N]
dF_r	Radial force [N]
dF_t	Feed force [N]
EF	External factor
ETS	Exponential Triple Smoothing
F_c	Average cutting force [N/mm^2]
F_t	ETS forecast
f_1	Rake surface
f_2	Flank surface
FDPM	Fused data prediction model
FRIEND	Feature selection approach on inconsistent data
gPC	Generalized polynomial chaos
LSTM	Long short-term memory
MACE	Mean acceptable error [%]
MAE	Mean absolute error [%]
mc	Chip thickness compensation factor
M_c	Simulated torque [Nm]
n	Spindle rotational speed [rpm]
P_c	Net power [kW]

✉ Teemu Mäkiaho
teemu.makiaho@tuni.fi

1 Automation Technology and Mechanical Engineering, Tampere University, Korkeakoulunkatu 7, 33720 Tampere, Finland

2 Pemamek Oy, Lamminkatu 47, FI-32200 Loimaa, Finland

PCC	Pearson correlation coefficient algorithm
PLC	Programmable logic controller
P-P	Peak-to-peak vibration [mm/s]
p-value	Probability value
PPX	Pay-per-x
Q	Material removal rate [mm^3/min]
ReLu	Rectified linear unit activation function
RMS	Root mean square [mm/s]
RMSE	Root mean square error [%]
RNN	Recurrent neural network
RUL	Remaining useful lifetime [%/min]
S_t	ETS seasonal forecast
t	Time
Tanh	The hyperbolic tangent activation function
T_{end}	Final running time [min]
T_i	Current running time [min]
T_t	ETS trends forecast
Vb	Flank wear [mm]
Vbf_1	Flank wear on the rake surface
Vbf_2	Flank wear on the flank surface
VbF_c	Cumulative flank wear value associated with average cutting force calculations [mm]
Vb_{max}	Maximum measured flank wear on individual blade [mm]
Vb_{mean}	Calculated mean of the blade-specific Vb_{max} values [mm]
VbM_c	Cumulative flank wear value associated with torque calculations [mm]
VbQ	Cumulative flank wear value associated with material removal rate calculations [mm]
VbRNN	Physics and data-informed recurrent neural network
V_c	Cutting speed [m/min]
V_f	Table feed speed [mm/min]
X_1	VbRNN input layer variable 'Cumulative VbF_c'
X_2	VbRNN input layer variable 'Milled meters'
X_3	VbRNN input layer variable 'Radial depth of cut'
X_4	VbRNN input layer variable 'Spindle motor torque'
X_5	VbRNN input layer variable 'Profile length'
X_6	VbRNN input layer variable 'Table feed'
$\widehat{x}_1$	Predicted variable 'Cumulative VbF_c'
Y_t	ETS forecast
z	Number of teeth in the spindle
*	Vb_{mean} value location in Fig. 9

Introduction

Milling machines play a crucial role in various industry sectors, enabling the cutting or shaping of raw materials to meet
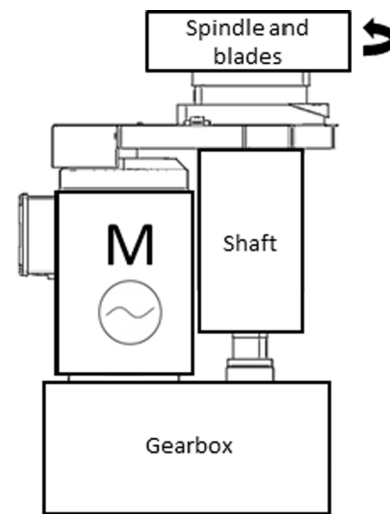

Fig. 1 The case peripheral milling machine's main components

specific requirements. This research focuses on a peripheral milling machine used in the marine industry, where continuous production is the norm and conventional methods for investigating milling blade wear progress are not often feasible. Therefore, the primary objective of this study is to increase the understanding of the wear phenomena in milling machine spindle cutting blades by combining different methods to accurately evaluate the wear progress. To achieve this, the study involves a thorough analysis of in-use data and physical characteristics associated with the milling process and combining measured wear data from used milling blades. The main physical components of the peripheral milling machine are illustrated in Fig. 1, being an alternating electric motor, gearbox, shaft, and spindle with cutting blades. The main cutting components of the test case peripheral milling machine comprise a spindle equipped with cutting blades arranged in four rows, with 18 blades per row around its circumference. The milling is operating in a down-feed direction.

Presently, the ongoing research in this field places a growing emphasis on the integration of diverse methodologies for prediction-making to overcome their respective limitations. One of the most critical aspects is integrating data that is relevant to the observed phenomena because inadequate data may also contain information that impacts the expected output negatively. Dimensionality reduction methods, which are based on machine learning, play a crucial role in unveiling hidden patterns and relationships within complex machining processes. The attractiveness of dimensionality reduction methods lies in their non-parametric nature, their efficiency in terms of computational requirements, and their straightforward implementation (Sarmadi et al., 2022). Wu et al. researched a physics-informed machine learning model to

demonstrate that the physics-informed model incorporation with the long short-term memory (LSTM) prediction model can achieve high-accuracy and reliable prediction performance in real-life milling operations surface roughness prognostics (Wu et al., 2022). However, the overall concern on model operation only on the limited datasets as well as potential shortcomings in the time-varying black box feature extraction process remains unanswered. In addition to the uncertainties created by the black box process, the direct usage of the input signals with LSTM may increase the possibility of bad-quality data. Processing raw time series data directly with the LSTM network might lack robustness due to the presence of noise in the sensor data (An et al., 2020). To address this issue, an integrated and transparent feature selection is needed to perform local feature extraction from the original signal sequence data.

Other components frequently incorporated into hybrid models include physics- and data-based models. Physics-based models provide essential data for prediction models that cannot be obtained through machine or sensor data, including information on structural integrity, material behavior, and machining dynamics (Elsayed, 2012). In general, the physics-based methodologies have the capability to assess the health status of a specific system by utilizing a set of equations that are derived from foundational principles in physics and engineering (Sikorska et al., 2011). Their drawback is that they often become excessively complex and require a deep understanding of the physical dynamics within the system of interest (Wu et al., 2017). Therefore, despite the progress in academia aimed at finding ways to optimize complex systems with multiple conflicting objectives, such as the data-driven sequential learning framework proposed by Khosravi et al. (Khosravi et al., 2024), it may still hinder the widespread implementation of the created model in other applications. In contrast, data-driven models are constrained by the extent of their training datasets (Arias Chao et al., 2022). These algorithms rely on historical data and big data rather than a comprehensive understanding of the system's physics (Heng et al., 2009). When evaluating the predictive uncertainties linked to the observed data, model parameters, and structures (Tian et al., 2023), data-driven models face limitations due to their training data.

Hybrid techniques may offer more in-depth information on the asset behaviour in contrast to physics-based modelling or data-based model used alone, as both models often suffer from their comprehensive applicability to complex real-world domains (Arias Chao et al., 2022). As such, hybrid approaches are continuously explored to leverage the strengths of both methods across research fields. Sahoo et al. proposed a hybrid model that merges the cutting force coefficient derived from finite element method (FEM) simulations with a revised undeformed chip thickness (UCT) algorithm

to predict cutting forces in micro end milling. The comparison between the forecasted and actual results showed a significant correlation, with the average peak force error ranging between 8.1 and 10.21% in the x- and y-directions, respectively. Despite the encouraging outcomes of their study in predicting cutting forces using hybrid models, it did not explore the relationship between cutting force and Vb predictions (Sahoo et al., 2019). Yang et al. (2022) proposed a novel hybrid method that merges data-driven strategies with insights from models for real-time wear detection in face milling machines, using power or force measurements. This model was put to the test with synthetic data created from simulations of a physics-based model, considering a variety of operational conditions, levels of measurement noise, and tool wear degrees. The model achieved an accuracy rate of 92% in data where 1% noise was artificially introduced. Importantly, this hybrid model significantly reduced the number of false alarms compared to using either data-driven or physics-based models on their own, demonstrating its effectiveness in accurately detecting tool wear and anomalies in real-time. Zhang et al. (2021) combined the digital representation of data with the physical inputs through a digital twin. They proposed a digital twin-enhanced dynamic scheduling methodology, which is based on the physical machine and virtual machine inputs. Their model outputs are used to enhance machine availability prediction, disturbance detection and performance evaluation. The highlighted limitation of the study emphasizes the time-consuming and costly work of the digital twins, which are required for the efficient implementation of the model. To overcome this challenge (Zhang et al., 2021), the authors are proposing the usage of a partial digital twin, comprising solely the relevant objects and essential model types (e.g. geometry models, physics models, or behaviour models) based on specific requirements.

In various research, the developed methods are implemented in controlled circumstances, yet, the real-world domain predictions require memory effects due to environmental noise and other natural disturbances in production. Li et al. (2022) developed a hybrid method to predict the Remaining Useful Life (RUL) of cutting tools by considering their wear state. They used support vector regression to map the relationship between sensor signals and tool wear in a controlled test setup. The findings indicated that this approach achieved enhanced prediction accuracy when contrasted with the utilization of exclusively physics-based or data-driven methods. However, the original support vector regression is known for its broad applicability (Santos et al., 2021) but is not widely acknowledged to accommodate large datasets (Rivas-Perea et al., 2013) or to effectively handle long-term dependencies in data (Bathla, 2020).

To capture memory effects more effectively from past occurrences, a version of the Recurrent Neural Network

(RNN) known as LSTM has demonstrated its potential in predicting RUL. Zhou et al. (2019) proposed a method involving the creation of a unified representation of working conditions and the extraction of wear characteristics from the processing signal. These extracted wear features, along with the corresponding working conditions, are combined into an input matrix for predicting tool wear. They utilize an LSTM model to capture the complex spatio-temporal relationships under variable working conditions and establish a model for predicting the remaining useful life of the tool. In another study, Nie et al. adopted an alternative approach by integrating a convolutional neural network (CNN), bidirectional long short-term memory (BiLSTM), and an attention mechanism to predict RUL of milling cutters. The CNN in their approach is responsible for handling sensor-monitored data, extracting crucial local feature information. Simultaneously, the BiLSTM neural network adaptively extracts temporal features, while the attention mechanism processes critical degradation features and extracts information related to the tool wear status. Their study demonstrated promising results compared to traditional approaches in terms of predictive accuracy (Nie et al., 2022).

In the context of milling blade wear prediction, there is a noticeable paucity of attainable and implementation-easy models concerning hybrid approaches suitable for deployment in ongoing production assets. Furthermore, prior studies have overlooked the utilization of prediction models that incorporate historical knowledge of past occurrences into physics-based simulation model usage and transparent feature extraction processes. In summary, this research introduces a novel Fused Data Prediction Model (FDPM) approach to fill this gap by combining advanced simulation model physics, rake wear results, and transparent feature extraction process with the recurrent neural network for RUL prediction.

The main contributions of this research are:

1. A novel data simulation model is established to emulate machine behaviour based on cumulative trend behaviour in terms of average cutting force (F_c), torque (M_c), and material removal rate (Q).
2. A recurrent neural network called VbRNN is established to predict rake surface wear in the milling machine context.
3. A novel FDPM model is developed, which combines simulated wear trend behaviour, real sensor data, a transparent feature extraction process, VbRNN and the Exponential Triple Smoothing (ETS) to extrapolate offset in the Vb predictions.

Methodology

A programmable logic controller (PLC) collects operational data from the peripheral milling machine, presented in Fig. 2a. The collected dynamic inputs consist of the following variables: table feed speed (V_f), cutting speed (V_c), radial depth of cut (a_e), and axial depth of cut (a_p). The online data collection process is described in Mäkiaho et al., (2023). The static inputs required for constructing the physics-based simulation model construction are summarized in Table 3. The machine static parameters Fig. 2b and dynamic variables Fig. 2d are used as the physics-based simulation model inputs, which creates simulated cumulative trend behaviour in terms of F_c, M_c, and Q. The blade wear laboratory measurements are performed for four (4) milling cycles (T1C4, T1C5, T2C4, and T2C5) presented in Fig. 2c. These results are connected to the simulation model, which further creates a time-integrated wear progress trend signal imitating temporal rake wear in the spindle blades. Due to the ongoing machine production, these measurements were taken only once, at the end of pre-determined milling meter targets. The results of the simulation model are evaluated by comparing them to the measured wear levels of individual cutting blades on the rake surface. In the experiments, the best-performing simulation model variable was found to be the cumulative rake wear, which is associated with average cutting force calculations (VbF_c). This dynamic trend is then combined with real machine data to create a hybrid dataset for the feature reduction phase.

The hybrid dataset is formed in Fig. 2e, where the dynamic variable data obtained from the PLC and the selected simulation model output VbF_c are conjugated. This merged dataset is utilized as input for the Pearson Correlation Coefficient (PCC) algorithm to ensure that only meaningful features are selected for the neural network training. The PCC is a statistical measure that calculates the linear correlation between two variables (Zhang et al., 2016), providing insight into the strength and direction of their relationship. The PCC is used to select optimal degradation features and linear correlations for the wear model (Cheng et al., 2019; Jiang et al., 2021). Therefore, the PCC is used to identify features that have a positive correlation with the VbF_c, yet, ensuring more informative and robust features for the RNN to learn and to improve its accuracy.

The subsampled features from the PCC are used as inputs $(X_1, \ldots, X_6)$ to train the LSTM neural network, named VbRNN due to its aim to predict accumulated rake wear. A generalized principle of a neural network is illustrated in Fig. 2, point (f). The output layer of the neural network supplies a probability value of a selected parameter to forecast the RUL of the system (Li et al., 2018; Zhang et al., 2016), which is detailed in Chapter 5.2. The Exponential Triple Smoothing (ETS) method is used in Fig. 2g

Fig. 2 Overall description of the data inputs and the FDPM model
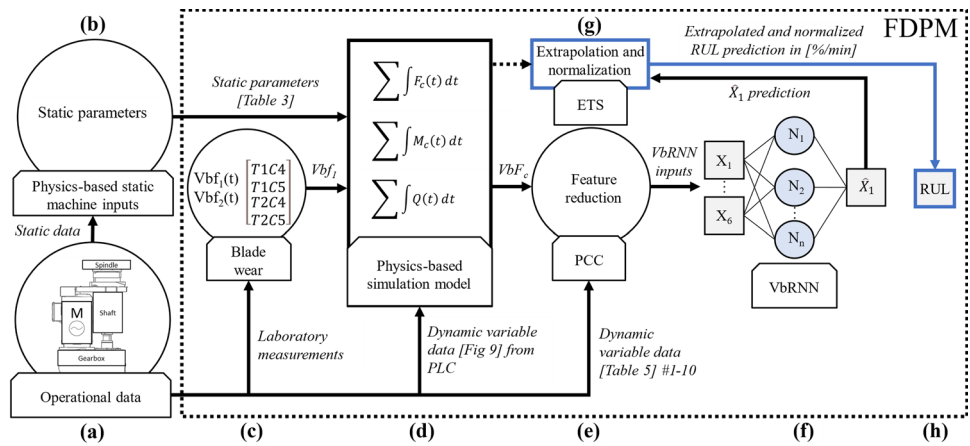
to visualize the compensation needed due to inaccuracies in the VbRNN predictions. The prediction model developed under this research is including the process steps depicted in Fig. 2c–h utilizing aggregated simulation model data, blade wear measurements, and real-life observations. The model is herein referred to as the Fused Data Prediction Model (FDPM).

Physics-based simulation model construction

To construct a physics-based simulation model, understanding the forces affecting the milling blade is essential. Cutting forces are the forces generated when a cutting tool is in contact with the milled material. These forces are generated due to the interaction between the tool's cutting edge and the workpiece, as illustrated in Fig. 3. The cutting forces can be divided into three components: feed force, radial force, and tangential force (Wayal et al., 2015). The feed force (dF_t) is the force that pushes the tool into the workpiece (Z-direction), while the radial force (dF_r) is the force that acts perpendicular to the cutting direction (Y-direction) (Moufki et al., 2015). The tangential force (dF_a) is the force that acts in the direction of the cutting edge (X-direction) (Moufki et al., 2015). These forces can vary depending on the machined material, V_c, chip depth of cut dimensions (a_e, a_p), and V_f, as annotated in Fig. 3.

The forces affecting the blade will wear the surfaces in contact with the milled material. The maximum surface wear (Vb_{max}) describes the maximum width of the surface wear land (Uhlmann et al., 2014) on each side of the contact flanks. These flanks are annotated as a rake surface (f_1) and flank surface (f_2) (Mia et al., 2017; Xie et al., 2012). The following subchapters present the physics-based equations, cutting blade wear measurements, and simulation model creation and validation.

Physics-based equations

The physics-based equations used in the simulation model are created with the input parameters V_f, spindle rotational speed (n) derived from the V_c, a_e, and a_p that can create as realistic variating output as possible to imitate physical phenomena occurring whilst the milling machine is in operation. The input parameters were collected from the peripheral milling machine data provided by the embedded PLC. The physical phenomena constructed in the simulation model are tested and compared with F_c, M_c, and Q calculations. The F_c formula is presented in Aaltonen et al. (1997):

$$F_c = h_m \times k_c \times b \times \left(\frac{360}{z \times \alpha} \right) \tag{1}$$

where average chip thickness (h_m), specific cutting force (k_c), and length of cut (b) are resolved by multiplication. This product is further multiplied with the result of 360 (degrees) over the product of the calculated contact angle (α), and the total number of cutting elements (z) in the spindle. The M_c is constructed according to Sandvik (2017):

$$M_c = \frac{P_c \times 30 \times 10^3}{\pi \times n} \tag{2}$$

Where the net power (P_c) calculations and n are dynamic parameters. Lastly, the Q is calculated directly from the simulation model inputs, as presented in Ersvik and Khalid (2015), Nee (2015):

$$Q = \frac{ap \times ae \times Vf}{1000} \tag{3}$$

Detailed employment of the dynamic aspects in Eqs. (1) and (2) can be observed in research paper (Mäkiaho et al., 2022), where the simulation model architecture was preliminarily prepared and used for vibration excitation and

Fig. 3 Blade contact parameters and affected forces
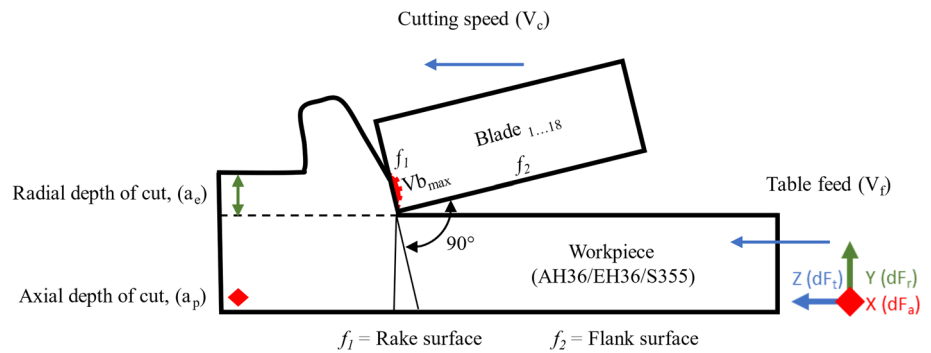
Table 1 Recorded operational cycles with cumulative milled meter values and recorded data points

Operational cycle	T1C4	T1C5	T2C4	T2C5
Milled meters [m]	2009	2505	2010	2518
Recorded data [points]	15,001	18,653	15,212	17,748

torque imitation to obtain additional information for pay-per-x (PPX) business model (Schroderus et al., 2022; Uuskoski et al., 2020) lifecycle calculations.

Cutting blade wear measurements

The blade surface wear measurements were performed from four (4) different operational cycles at the customer facility. The operational cycles generate the research datasets that are used in different phases of the FDPM. Table 1 presents the operational cycles along with their respective cumulative meters milled and the number of recorded data points for each cycle. The quantity of data points recorded during the milling process is contingent upon two variables: the duration of the milling operation and the specific profile type being manufactured. This relationship exists because material feed speeds, which are controlled automatically, vary depending on the type of profile being processed. The operational data is collected with an online data procurement system connected to the machine's PLC system. Once in operation, the data points were recorded with a frequency of 5 Hz, resulting in 200 ms between individual data points.

The naming convention of T1 and T2 in the operational cycles indicates the unique numbering of two separate spindles. This is done to prevent unnecessary production stoppages caused by the time required for blade changes. Therefore, T1 is the spindle set number 1, and T2 is the spindle number 2. During the blade wear measurement tests, the spindles are changed to the peripheral milling machine once the predetermined stage of milled meters is reached. The predetermined stages were 2000 m and 2500 m which were met in close approximation. In designations, C4 and C5, the letter 'C' refers to blade rows on a spindle, which are positioned at a 90-degree cutting angle perpendicular to the material being milled. During the normal milling process, only one blade
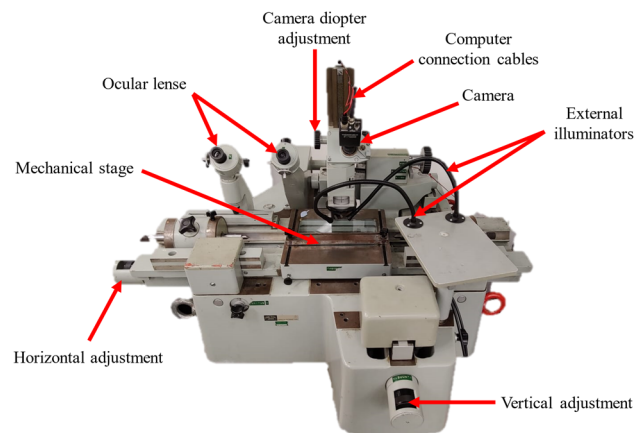

Fig. 4 Carl Zeiss Jena optical microscope setup with the digital connection to the NI Labview engineering software-based digital photo measurement tool

row is in contact with the material. The numbers '4' and '5' indicate the specific blade edge used in a given operational cycle.

Laboratory measurement setup and wear limit criteria

Optical microscopic measurements are accomplished with a 1200–2400% zooming range from the original pixel frame of 2560 times 2048 pixels. The Vb_{max} measurements were performed with calibrated Carl Zeiss Jena optics connected with internally created digital photo measurement software based on NI Labview engineering software. Microscopic measurement setup excluding the computer interface with the measurement software is illustrated in Fig. 4. Due to the confidential nature of the manufacturing equipment, a more

Fig. 5 Numbered individual cutting blade #46 flank surface 1–4 in (a), rake surface edges 4 (up left) and 8 (down right) in (**b**)
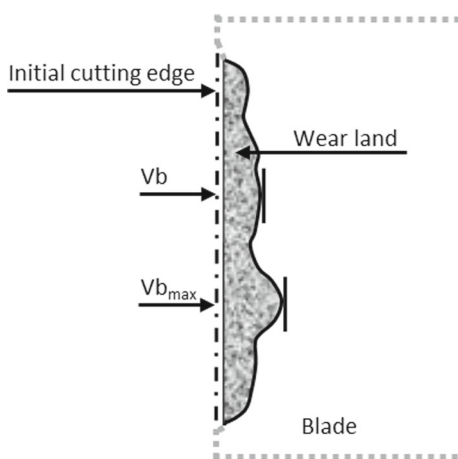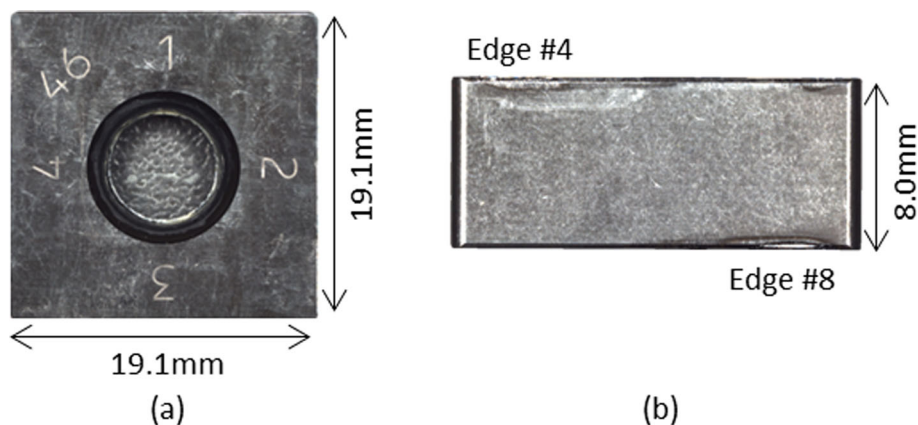
Fig. 6 Principle of the wear measurement

detailed illustration of the milling machine setup is not provided.

The sides of the cutting blades were marked with carved numbering to reduce the risk of confusion after the milled meters were received and the blade edge was turned or changed in the spindle as illustrated in Fig. 5. In addition to each cutting blade numbering, each of the edges was carved for similar reasoning, as presented with numbers 1–4 in Fig. 5a. Similar identification for edges 5–8 exists on the reverse side of the f_2. The f_1 surface side of the blade is presented in Fig. 5b.

In this study, the blade-specific wear measurements are presented by the maximum values of flank wear (Vb_{max}) for individual flank surfaces, where the maximum peak land width is measured (Siddhpura & Paurobally, 2013; The American Society of Mechanical Engineers, 1985), as annotated in Fig. 6.

The Vb_{max} tool life criterion can be considered as a wear criterion when the wear pattern in the measurement area results is relatively uneven (Siddhpura & Paurobally, 2013), which meets the criterion in the results of this research. The Vb_{max} is also referred to as the critical flank wear in which the tool can be observed to reach its end of life and requires replacement (Traini et al., 2019). In previous studies, different Vb_{max} values were used: 0.24 mm in Panda et al. (2008), 0.3 mm in Lin et al. (2020), and 0.7 mm in Caldeirani Filho and Diniz (2002). In this study, a blade change threshold of 0.4 mm was established for the prediction phase. The determination of this threshold was based on expert opinions regarding blade condition, which involved visually inspecting and comparing the blades after their operational cycle. To achieve our research objective, a thorough evaluation of how the blade condition affects milled material quality in the specific machine construction was carried out in collaboration with end users and machine manufacturer experts. This evaluation centred around a critical threshold of 0.4 mm, measured from the rake surface, which played a central role as a defining parameter in our investigation. Therefore, the normal operational rake wear and milling blade change limit used in this study can be summarized by the presented criterion as follows:

Normal operation: 0 mm $\geq Vb_{max} \leq$ 0.4 mm.

The individual cutting blade physical dimensions are given in millimetres to X, Y, and Z-directions according to Fig. 7a, being 19.1 mm, 8.0 mm, and 19.1 mm, respectively. Consequently, the selected Vb_{max} value corresponds to approximately 5% of the flank's physical dimensions in the f_1 direction and 2.1% in the f_2 direction.

The f_1 surface side of an individual cutting blade is presented in Fig. 7a where the observed cutting edge is pointed with a red colour rectangle. Each measured cutting edge surface was analyzed by the software-based measuring tool to accurately observe the maximum peak land width in Fig. 7b.

Fig. 7 Microscopic view of milling blade Vb$_{max}$ indicating observed blade #44 flank #4 in (**a**), close view in (b), and Laplacian transformed view in (**c**)
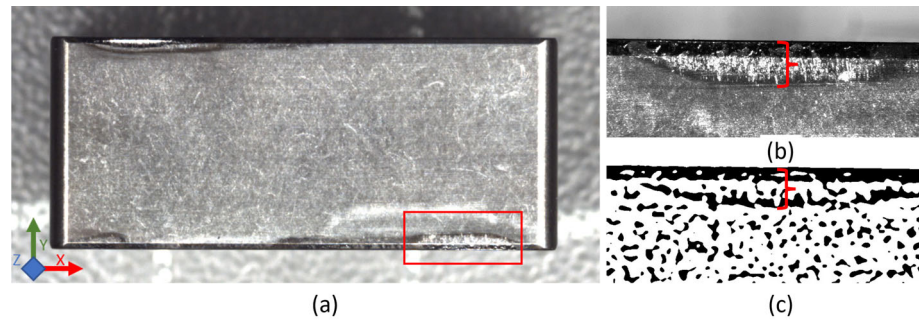
Fig. 8 Wear pattern comparison between rake surface (f$_1$) and flank surface (f$_2$) wear named as Vbf$_1$ and Vbf$_2$, respectively
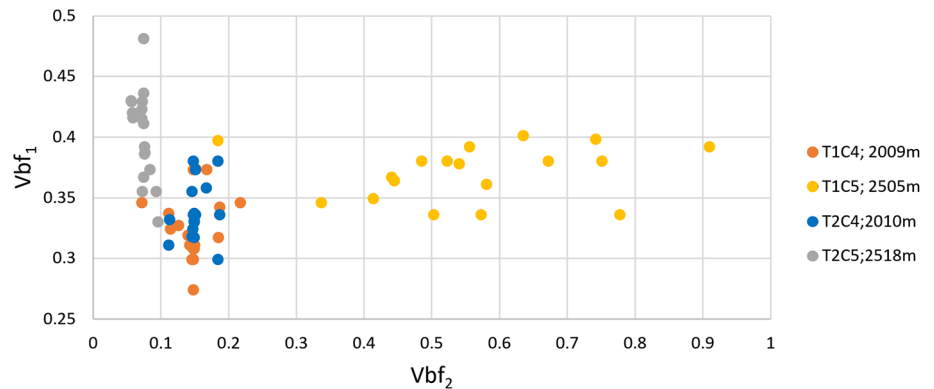
A Laplacian of Gaussian filter was occasionally employed in the Vb$_{max}$ pattern recognition, as the filter helps to estimate the scales, shapes, and orientations of an object (Siddhpura & Paurobally, 2013). An example of the Vb$_{max}$ examination with the help of the Laplacian of Gaussian filter is presented in Fig. 7c to determine the f_1 surface wear maximum value.

Analysis of the wear measurements

The wear measurement results were analyzed to quantify the amount of wear experienced by the milling blades during their designated milling cycle. As previously discussed, these forces act in three dimensions, with the primary effect being the wear on the two blade surfaces that come into direct contact with the milled material. Therefore, the milling blades' f_1 surface and f_2 surface measurements were performed. The f_1 surface-related inconsistency on test run T1C5 can be observed in Fig. 8 where the majority of the measured data points are scattered on the f_2 surface side of the blade, which can be interpreted as anomalous behaviour.

Consequently, the T1C5 flank surface wear results are left out from the wear prediction part of the research due to its examined inconsistency on the wear patterns in comparison to any researched time or cycle constraints. Despite this, the

rake surface measurements on the T1C5 remain consistent in comparison to the number of milled meters with the other test runs, which fortifies the f_1 surface measurement's usage in the simulation model creation. The scatter plot illustration in Fig. 8 also presents the natural variation in rake surface dimension between the individual blades. Naturally, some degree of variation in the f_1 surface dimension is anticipated in the measured results. After analyzing the measurements, the recorded variation is deemed to fall within reasonable limits.

Due to the natural variation in the results, the measured mean value (Vb$_{mean}$) of the blade wear is calculated for the simulation model validation purposes. Minimum and maximum values are also recorded to obtain the scale in which the Vb$_{max}$ values are present in the circumference of the blades in a specific row. The calculated mean values used in the simulation model construction (T2C4) and validation phase (T1C4, and T2C5) are found in Table 2 below. A complete list of the blade-specific f_1 surface and f_2 surface Vb$_{max}$ values of each test run are recorded in Appendix 1, yet only the f_1 surface results are used in the simulation model construction.

Table 2 The measured mean, minimum, and maximum f_1 wear values in millimetres

Dataset	Vb measured mean	Vb scale min	Vb scale max
T2C4	0.338	0.299	0.38
T1C4	0.327	0.274	0.374
T2C5	0.402	0.330	0.481

Table 3 Simulation model static input parameters

Static input parameters	Value	Unit
Spindle diameter, (D)	496	mm
Spindle number of effective teeth, (z)	18	pcs
Electric motor nominal power, (P)	45	kW
Gear unit ratio: (HSS/LSS ratio)	7.88	ratio
Low-speed shaft nominal torque, (Tg)	2290	Nm
Lead angle, (κ)	90	degrees
Material-specific cutting force, (Kc1.1) (Sandvik P1.2.Z.HT)	1820	N/mm^2
Chip thickness compensation factor, (mc)	0.25	–

Physics-based simulation model creation and validation

The objective of the simulation model is to reconstruct the continuous-time wear pattern to as close an approximation to the known measured Vb_{mean} -value as possible. All the simulations are performed in MATLAB-Simulink software, which is a commonly known design and programming platform for dynamic and embedded systems. Simulation inputs are divided into static parameters and dynamic input variables. Static input parameters used in the simulation model construction are listed in Table 3, containing physical dimensions of the milling machine components, milling lead angle, nominal component values, and feed material-related static properties. All the static data presented in Table 3, including the material-specific hardness factor *Kc1.1,* is received from the case machine's original equipment manufacturer. The chip thickness compensation factor *mc* is obtained (Sandvik, 2017).

The equations (Eqs. 1, 2, and 3) represent various calculation formulas that incorporate both the static parameters and the dynamic input variables for the simulation model in Fig. 9. Base calculations for the simulated formulas are presented in Mäkiaho et al. (2022). Due to diversity in the equation output units, a unique external factor (EF) is needed to accommodate the iterated simulation results as close as possible to the measured Vb_{mean} target value measured from the T2C4 dataset. The purpose of the EF is to correct any offset or discrepancy in the signal behaviour that may arise due to differences in the units used in the equation's output. By applying the EF, the simulation results can be calibrated to better match the measured target value, ensuring greater accuracy in the model's predictions. Also, the discrete values provided by the equations require continuous-time integration (1/t) to obtain the cumulative trend behaviour of the parameter's progress in the time domain. Consequently, the simulation model provides cumulative trends of the individual equation outputs to be validated, named VbF_c, VbM_c, and VbQ, respectively.

The VbF_c, VbQ, and VbM_c values presented in Table 4 are the final results received from the physics-based simulation model. The validation of the results is performed by comparing the received simulation results to the measured dataset-specific mean values already presented in Table 2.

The T1C4 dataset was primarily used in the simulation model validation due to its relatively similar milled meters values (2009 m) in comparison to the T2C4 (2010 m) used for the model construction. All the simulated results are cumulative values to present wear accumulation to the blades, which are encountering stress behaviour when in physical contact with the milled material. The simulated results illustrate relative error % in comparison to the measured mean values as VbM_c and VbQ indicating negative, and VbF_c positive error values with the T1C4. The scale of the variations in the T1C4 results is reasonable in accuracy but the variation in the results propagated in further testing of the constructed simulation model with another validation test run, T2C5. The model was calibrated using data for a shorter milling cycle (T2C4, 2010 m); however, its extrapolation capability for producing accurate rake wear simulation results for a longer time domain data cycle (T2C5, 2518 m) was exceeding expectations. The relative error % values obtained when compared to the measured Vb_{mean} at the final data cycle stayed within a relatively small range, indicating much better accuracy than the T1C4 results. The average error percentage of the dataset relative errors also indicates that the VbF_c method overcomes the VbM_c and VbQ in accuracy.

In conclusion, the VbF_c produces the most accurate values to meet the measured Vb_{mean} target value. The simulated VbF_c average error value of 2.38% is therefore bolded in Table 4 to highlight the most applicable simulated signal for the prediction algorithm. The accumulated VbF_c trend behaviour is visualized in Fig. 10 with the measured minimum and maximum values for the test run. All the simulated results are located inside the MinMax boundaries of each dataset. The red asterisk '*' symbols in the figure are presenting the measured Vb_{mean} value location of each test run at the end of the milling cycle.

To conclude, the results from the presented two validation rounds give good confidence in using the VbF_c in the data

Fig. 9 Structure of the simulation model
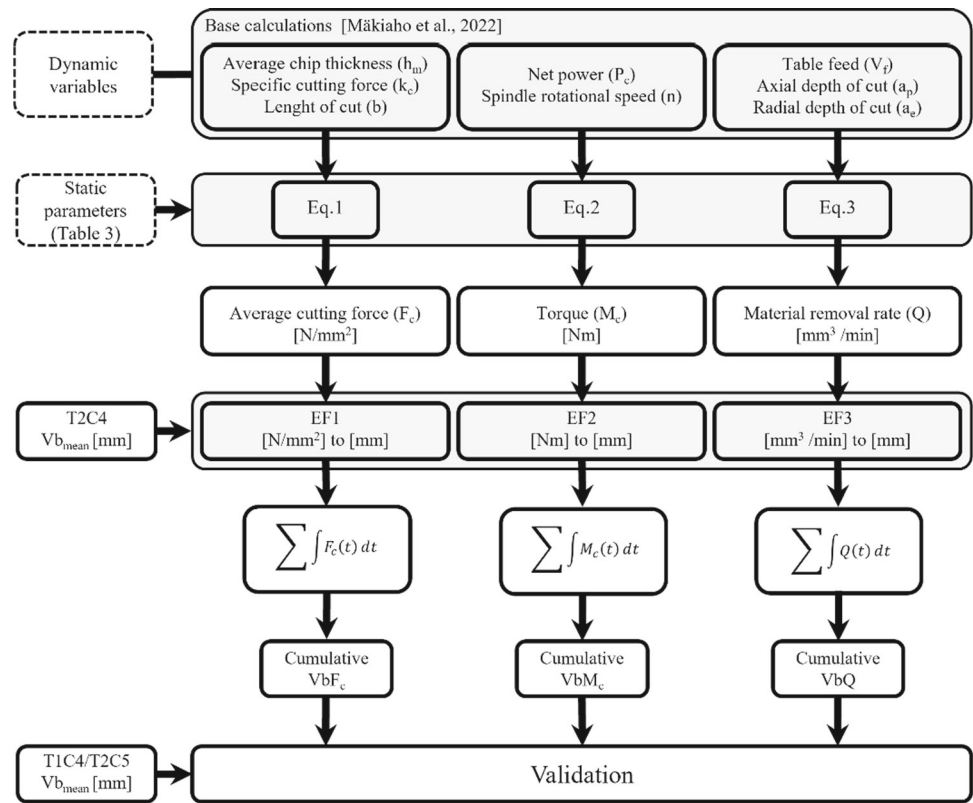
Table 4 Simulated Vb results in comparison to measured mean values in Table 2

| Dataset	VbF$_c$ [mm]	Relative error [%]	VbM$_c$ [mm]	Relative error [%]	VbQ [mm]	Relative error [%]
T2C4	0.338	0.00 %	0.340	0.59 %	0.340	0.59 %
T1C4	0.355	7.89 %	0.300	-9.00 %	0.290	-12.76 %
T2C5	0.399	-0.75 %	0.390	-3.08 %	0.390	-3.08 %
Average error [%]		**2.38 %**		-3.83 %		-5.08 %

merging phase together with the operational data obtained from the PLC.

Feature reduction process with Pearson correlation

This stage of the FDPM model integrates the operational data and physics-based simulation model data into one dataset. The data collection was designed to collect data from PLC as well as external vibration measurements during the blade wear test duration (54 calendar days). The collected vibration data contained vibration raw signal data from three (3) directions (axial, horizontal, and radial) as well as calculated Root Mean Square (RMS), zero-to-peak (0-P), and peak-to-peak

(P-P) amplitude values. However, a malfunction in the vibration collection setup at the start of the wear measurements hindered the ability to use the vibration excitation data in conjunction with the other process-related data collected. As a result, it was decided to only utilize the PLC data to obtain usable parameters for the RNN, thereby creating a consistent stream of data for the algorithm's training. Although, some methods are recognized in the literature for selecting features on inconsistent data like the feature selection approach on inconsistent data (FRIEND) in Qi et al. (2020) or mean acceptable error (MACE) in Kim et al. (2017), using such additional methods are excluded from this research due to the existence of other operational data adequate for the purpose. The collected data, excluding the vibration data, consisted of operational data from eleven (11) variables, listed in Table 5.

Fig. 10 Simulated VbF_c trends with a different test run T2C4 'blue', T1C4 'orange', and T2C5 'grey' with measured MinMax boundaries for the test run, as well as Vb_{mean} rake surface value with '*'
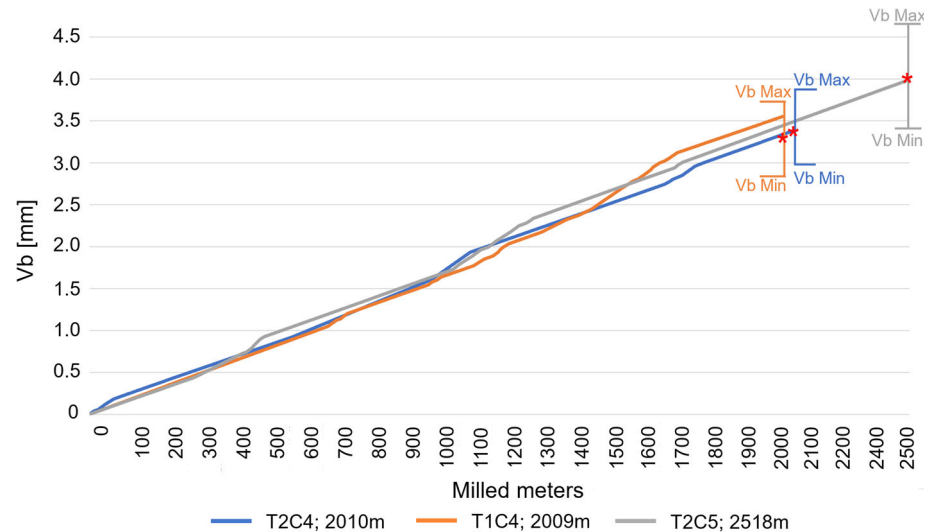
Table 5 Input variables to Pearson correlation coefficient algorithm

Variables	Unit	Variable #
Axial depth of cut, (a_p)	mm	1
Cutting speed (V_c)	m/min	2
Frame position	mm	3
Milled meters	m	4
Profile height	mm	5
Profile length	mm	6
ProfileType	1…2	7
Radial depth of cut, (a_e)	mm	8
Spindle motor torque (M_c)	Nm	9
Table feed, (V_f)	mm/min	10
VbF_c	mm	11

Variables 1–10 are obtained from PLC and variable 11 from the simulation model.

Two well-known correlation methods, Pearson and Spearman's (Myers & Sirois, 2006), were initially tested for the T2C4 dataset. Upon observing a significant similarity in the variable correlations between the two methods when comparing VbF_c to other variables, the selection of Pearson over Spearman was motivated by its renowned capability in detecting linear relationships between variables measured on continuous scales (Obilor & Amadi, 2018). Consequently, the Pearson correlation coefficient was deemed more appropriate for the analysis.

The selection of input variables for the VbRNN is based on two criteria: an overall positive average score and a positive correlation to the simulated VbF_c in at least two out of three data cycles. Applying these criteria resulted in the selection of six variables, denoted as X_1 to X_6, as presented in Table 6. After exposing the input variables to the PCC algorithm, the results are indicating the highest average correlation to VbF_c from 'Milled meters','Radial depth of cut','Spindle motor torque','Profile length', and 'Table feed', with the correlation of 0.834, 0.176, 0.166, 0.076 and 0.035, respectively. All the selected variables were found to be statistically significant with both the tested methods at a significance level of $p = < 0.01$, as correlations are deemed significant when p-values are below 0.05 (Obilor & Amadi, 2018).

Remaining useful lifetime prediction with a recurrent neural network

The RNNs are commonly used in real-life applications due to their proven applicability to detect dependencies in the data as well as solve several types of time series forecasting problems in high-dimensional data structures (Sagheer & Kotb, 2019). A recurrent neural network learns not only from the current time series input but can accommodate relevant information from past states of a neuron in the network (De Beaulieu et al., 2022). A widespread version of the RNN neural network is called LSTM, which was introduced by Hochreiter and Schmidhuber in 1997 (De Beaulieu et al., 2022; Samek et al., 2019).

The cell architecture in the LSTM algorithm comprises this specialized capability to learn and remember long-term dependencies with the help of dedicated backward flow. Therefore, the LSTM algorithm is selected as part of the network in a supervised manner to perform the forecasting for the FDPM model dataset containing multidimensional data

Table 6 PCC scores for the datasets, the highest denoted as X_1 to X_6 inputs to the VbRNN algorithm based on their average correlation to the VbF_c variable

Variable	Unit	PCC [T2C4]	PCC [T1C4]	PCC [T2C5]	Average	VbRNN input
VbF_c	mm	1.000	1.000	1.000	1.000	X_1
Milled meters	m	0.860	0.804	0.839	0.834	X_2
Radial depth of cut, (a_e)	mm	0.183	0.244	0.101	0.176	X_3
Spindle motor torque	Nm	0.146	0.185	0.166	0.166	X_4
Profile length	mm	0.630	− 0.690	0.288	0.076	X_5
Table feed, (V_f)	mm/min	0.101	− 0.138	0.143	0.035	X_6
Axial depth of cut, (a_p)	mm	− 0.049	0.284	− 0.168	0.022	
Cutting speed (V_c)	m/min	− 0.003	− 0.057	0.107	0.016	
ProfileType	1…2	0.114	− 0.204	0.049	− 0.014	
Profile height	mm	0.047	− 0.120	− 0.265	− 0.113	
Frame position	mm	0.044	− 0.120	− 0.273	− 0.116	

in time series. The predictions are performed with different learning ratios of 50%, 70%, and 90% to obtain information and knowledge on the model's prediction accuracy towards the end of the spindle blade set life cycle. The operational data from the test run 'T2C5' is selected for testing the RNN algorithm due to its length in data points as well as its absence in the physics-based simulation model construction. The computational evaluation was performed using an Intel(R) Core(TM) i5-8365U CPU processor with 1.60 GHz, 1896 MHz, and 4 Core(s), along with 16 GB of physical memory. The Python algorithm was executed using the Jupyter Notebook computing platform.

Architecture and training

Generally, the deep neural network developed consists of the input layer, LSTM layer, dense layer, and output layer, as presented in Fig. 11. To simplify, the complete deep neural network architecture developed in this research is further referred to as VbRNN, to contain its function to predict blade surface wear with the help of the recurrent neural network.

The input layer consists of six variables based on the PCC results, being 'VbF_c', 'Milled meters', 'Radial depth of cut', 'Spindle motor torque', 'Profile length', and 'Table feed', the first being the forecasted variable. The input neurons are depicted as $X_{1…6}$, respectively. To enhance network performance, sample values are normalized to fall within the range of [0, 1] in the model training phase. For predicting the next time step, a sliding window look-back technique with a value of 10 is employed to select the number of previous time steps used as input features. To prevent overfitting and conserve computational resources, early stopping is implemented to

monitor the validation loss with patience of 3. This means that if no improvement is observed in the validation loss for three consecutive epochs, the training process automatically stops. The combined use of the sliding window and early stopping ensures the model's performance on the validation set is optimized and prevents unnecessary training iterations.

After the input layer, the network has a single LSTM layer with 64 hidden units and hyperbolic tangent (Tanh) as an activation function, due to its robustness and non-linear insertion capability for neural networks (Sartin & da Silva, 2013). The Tanh function compresses the input between negative 1 and positive 1, being: $X \in (−1, 1)$ (Herawan et al., 2016; Zeng & Long, 2022). Rectified linear unit activation function (ReLu) and logistics sigmoid (sigmoid) activation functions were also tested for the dataset. The ReLu is commonly known for allowing to find of complex nonlinear relationships from the data. The ReLu retains the positive numbering and restrains negative numbering to zero (Nanni et al., 2022). Sigmoid is known for its capability to manage the output data of the network layer between 0 and 1 (Xu et al., 2021): $X \in (0, 1)$ (Zeng & Long, 2022). However, the comparison of the Keras activation functions with 50% of training data indicates the most sufficient prediction capability for the dataset is addressed with the Tanh activation function. Tanh indicates dominancy to comply with the T2C5 test run Vb value, where the probability to converge the actual Vb value is stated as 0.402(0.378|tanh).

The comparison of the activation function influence on the prediction accuracy, related Mean Absolute Error (MAE) and associated computational effect in terms of training time are found in Table 7. The training time for Tanh and Sigmoid activation functions is similar, while the ReLU activation

Fig. 11 The VbRNN architecture used for RUL prediction contains an input, an LSTM layer, a dense layer, and an output layer
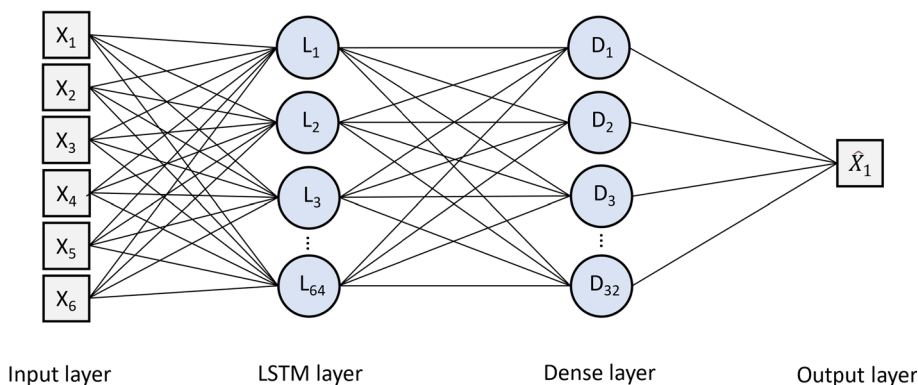
Table 7 Activation function Tanh, ReLu, and Sigmoid impact on the VbRNN prediction capability and associated training times

	T2C5 Vb	Tanh	ReLu	Sigmoid
Vb	0.402	0.378	0.344	0.205
MAE	–	0.011	0.030	0.104
Training time (s)	–	74.53	109.39	73.37

function requires significantly more time for training. The MAE indicates the model learning capability by estimating the performance degradation by comparing the estimated trend with actual performance (Pecht & Kang, 2018), where all individual differences have equal weight (Sagheer & Kotb, 2019). MAE calculation form is depicted in Chicco et al. (2021), Peng et al. (2010), Tong et al. (2022):

$$MAE(x, \widehat{x}) = \frac{1}{N} \sum_{i=1}^{N} |x_i - \widehat{x}_i| \tag{4}$$

Considering that each of the nodes is connected to the following layer, the architecture is called a fully connected network containing dependency between each active layer. Another terminology for a fully connected layer is a dense layer (Zeng & Long, 2022). The used dense layer consists of 32 nodes to reduce feature dimensionality from the LSTM layer. The dense layer is further connected to a single predicted parameter $\widehat{X}_1$(VbF$_c$) in the output layer.

The LSTM and dense layers are added with a dropout function with a rate of 0.2 to improve model generalization in the training (Cheng et al., 2017). Bayesian optimization function with log-uniform was tested with the following limits: low = 0.00001, high = 0.001 to uniformly sample in the logarithmic space between *low* and *high*. Adam optimization with a learning rate of 1.585E-5 was selected for the model based on the result from the Bayesian optimization function. The batch size was manually iterated receiving the best scores when the batch size = 30. Following, the MAE
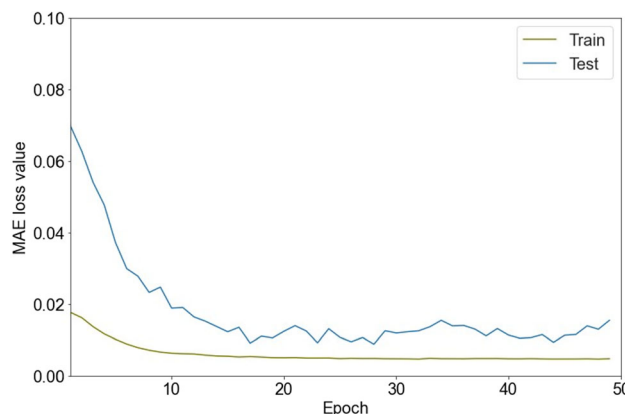
Fig. 12 The mean absolute error with the data set training size of 50%

loss function is used to evaluate the model performance with different tested training % shares. The MAE loss behaviour is illustrated in Fig. 12, where both the train and test trends are presented with a 50% training share by using the Tanh activation function and aforementioned optimization characteristics.

The learning iteration capability of a model is controlled with epochs. The results declare that the model learning capability is saturated approximately after 15 epochs, however, the network was further trained to achieve good accuracy, learning rate, and loss minimization as proposed in Poornima and Pushpalatha (2019). Another widely used model

Table 8 RMSE and MAE scoring with 50%, 70%, and 90% train set with the T2C5

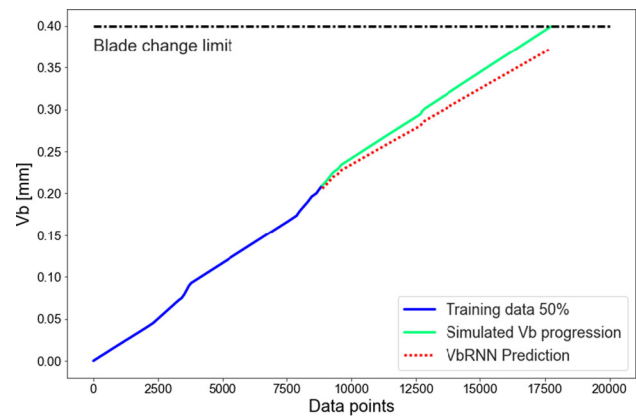| Data (trained data points)	Test RMSE [%]	Test MAE [%]
Full cycle (17,748)	–	–
Prediction with 90% train (15,972)	0.005	0.004
Prediction with 70% train (12,422)	0.012	0.011
Prediction with 50% train (8873)	0.016	0.014

Fig. 13 VbRNN prediction capability visualized with training data share of 50%

prediction error was tested by applying root mean square error (RMSE) for the train data. The RMSE measures the standard deviation of the errors that the RNN architecture yields in its predictions (Géron, 2017; Lughofer & Sayed-Mouchaweh, 2019) by squaring the errors before averaging, therefore, making it more sensitive to larger errors (Wang & Lu, 2018). RMSE calculation form is given as in Chicco et al. (2021), Peng et al. (2010), Tong et al. (2022):

$$RMSE(x, \widehat{x}) = \sqrt{\frac{1}{N} \sum_{i=1}^{N} |x_i - \widehat{x_i}|^2} \tag{5}$$

The lowest scores for RMSE and MAE are presented in Table 8. As seen, both the RMSE and MAE scores decrease as the training data amount increases. Considering the low RMSE and MAE scores the model performance is proven, as generally acknowledged that the lower the evaluation indexes of RMSE and MAE are, the better the model performance is considered (He et al., 2021; Wang et al., 2017).

Remaining Useful Lifetime (RUL) prediction

The training set size and data quality conclusively affect the prediction accuracy of the model due to its capability to retrieve dependencies from previously learned data. As described earlier, the LSTM layer architecture creates long-term dependencies with the help of dedicated backward flow, which ultimately results in higher accuracy on the prognosis as more data is fed to the model. The numeric results of the prediction accuracy are illustrated in Table 9, which declares the progressive learning impact on the Vb prediction accuracy for the simulated T2C5 dataset. The prediction result of 93,6% with the given 50% training set size incrementally increases being 94,7% with 70% train set size, and 97,6% with 90% train set size when $\widehat{x_1}/Vb_{mean}$. Overall, the results indicate that the rake wear phenomenon aggregation with F_c

calculations is relatively attainable by the VbRNN algorithm to learn as indicated by the prediction accuracy results.

The trend behaviour between the simulated Vb progression and VbRNN prediction is visualized in Fig. 13. The prediction is distinct to receive higher accuracy together with the increase in the training data set. The 'blue' trend indicates the training share, 'green' is the simulated Vb progression, and 'red' is the VbRNN model's capability to predict future trend behaviour. The horizontal line 'Blade change limit' is set to the selected rake surface blade change interval target of 0.4 mm to indicate the appointed blade change threshold. The VbRNNs prediction capability is visualized with a training data share of 50%.

The trend of remaining useful lifetime is further illustrated in Fig. 14, which employs a learning rate of 50%. The figure presents the RUL in minutes, as well as the normalized wear value of Vb for T2C5 (0.4019 mm), expressed as a percentage ranging from 100 to 0%. The normalized RUL for the system degradation is expressed as:

$$RUL = \frac{(T_{end} - T_i)}{T_{end}} \tag{6}$$

where, T_{end} corresponds to the time at the selected Vb value of 0.4019 mm at RUL [0%/0 min], and T_i is the current time in operation (Feng et al., 2023). Since the prediction value with a 50% training dataset resulted in the predicted Vb of 0.3761 mm at the last data point (17,748), therefore, extrapolation to reach the targeted Vb of 0.4019 mm was performed to visualize the variation of the model inaccuracy in terms of RUL minutes and %-values. The ETS is a generally acknowledged method used for time-series forecasting. The ETS uses three parameters with different weights: level, trend, and seasonal. The mathematical basis of the ETS is observable in Airlangga et al. (2019), Chen and Ho (2020),

Table 9 VbRNN prognostic capability for the simulated T2C5 dataset, evaluated by absolute error, related % -value, and the model prediction accuracy %

Data (trained data points)	Vb_{mean}, $(\hat{x}_1)$ [mm]	Absolute error	Absolute error [%]	Prediction accuracy [%]
Full cycle (17,748)	0.4019	–	–	–
Prediction with 90% train (15,972)	(0.3923)	0.0096	0.96%	97.6%
Prediction with 70% train (12,422)	(0.3806)	0.0214	2.14%	94.7%
Prediction with 50% train (8873)	(0.3761)	0.0259	2.59%	93.6%

Fig. 14 Remaining useful lifetime in minutes and % with VbRNN prediction and ETS forecast trends
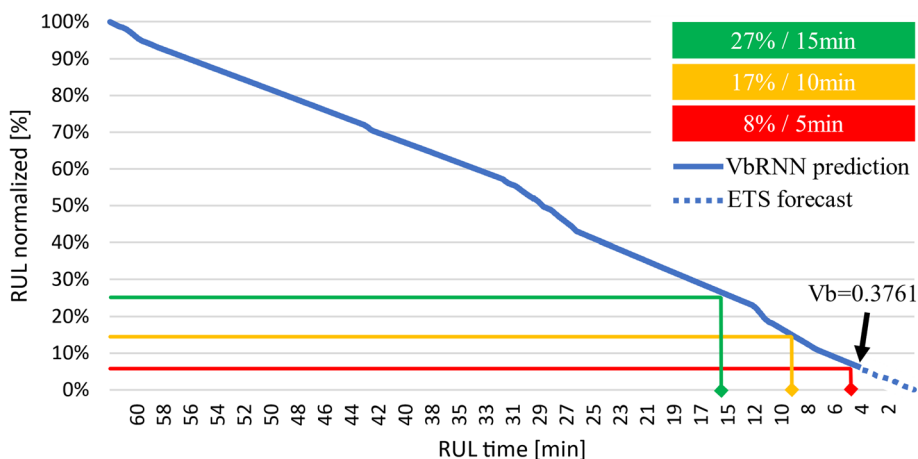
where the ETS forecast (Y_t) is described in the following equation by summing up level (F_t), trends (T_t), and seasonal (S_t) over time (t) forecasts.

$$Y_t = F_{t-1} + T_{t-1} + S_t \tag{7}$$

The ETS method was employed to integrate 2000 historical events (data points 15,748–17,748) of the predictions made by the VbRNN. The primary utilization in this instance was to display the number of timesteps by which the VbRNN algorithm predictions fell short of the simulated wear Vb = 0.4019 mm. Thus, the RUL RUL = 0%/0 min is illustrated with the extrapolated forecast targeted on the final Vb value. The extrapolated trend based on the ETS calculation performed in the simulation model is annotated with a dashed line in Fig. 14.

The prediction results indicate that the trained dataset with a 50% training deviates by approximately 4 min of machine operation, having 6% in RUL normalized % -value left in comparison to the original T2C5 data set simulated values.

Additionally, Fig. 14 contains proposed threshold limits, which are illustrated in traffic light colours. Such limits are feasible to indicate forthcoming milling blade change operations with predetermined warning thresholds. These thresholds are set and visualized, with the first notification at 27%/15 min, and the second at 17%/10 min of the remaining lifetime. The last threshold indication is undertaken at 5 min and 8% before the predicted end of life, giving the operator sufficient time to react and prepare for the spindle change operation.

Conclusions

The assessment and measurement of blade wear in real-life applications present a significant challenge, primarily due to the reluctance of real-life manufacturing processes to undergo operational stoppages unrelated to planned or forced downtime. This limitation hampers the ability to comprehensively investigate the wear progression of milling blades, a

process that often requires repetitive measurements at different stages of the milling cycle and throughout the predicted lifetime of the blades. In this research, a novel FDPM model is proposed enabling accurate blade wear predictions yet minimizing the need for additional downtimes in production often required for empirical studies. While previous studies consider physical dynamics with the inclusion of machining data, they often lack attainability and efficient implementation for in-production industrial assets. The FDPM's open architecture and demonstrated capability to predict availability with limited training data address these gaps left by previous research.

The validation of the physics-based simulation model demonstrates its excellent predictive capability for the VbF_c, yielding a low average relative error of 2.38% when compared to the measured mean of the milling cycle rake wear. This finding underscores the model's accuracy and reliability in simulating real-world scenarios. However, the model's foundation heavily relies on the physics-based simulation model, which incorporates the physical behaviour occurring during milling machine operations. As a consequence, the model's physics-based simulation model construction process requires only a limited amount of operational data. Whereas the model relies highly on the physical features, the reliance on specific physical effects in the simulation model's architecture limits its direct generalizability for use in other applications.

The PCC stage of the FDPM process is incorporated to retain the variables having a positive correlation to Vb, which enables faster training times and improves the model performance of the VbRNN prediction algorithm by reducing the dimensional space. The VbRNN prediction method is considered suitable for the peripheral milling machine setup due to its capability to retain information on past occurrences, as well as the capability to adapt changes in the production setup while offering transparent implementation of the selected features.

The prediction accuracy of 93.6% with the limited 50% training data gives a profound ground towards further developing the FDPM model's RUL prediction for onsite use. In the results, the prediction result of 93,6% with the given 50% training set size incrementally increases being 94,7% with 70% train set size, and 97,6% with 90% train set size. These results highlight the excellence of the FDPM model and lay a solid groundwork for the potential development of semi-autonomous or fully autonomous variable selection, prediction offset analyzing, and RUL estimation systems, thereby enhancing productivity and enabling on-site utilization of FDPM in milling machines.

The constrained availability of comprehensive machine-related datasets hindered the opportunity to conduct a more extensive and thorough investigation of the FDPM model over a prolonged time frame. In future research, it is recommended to give priority to incorporating the FDPM model into diverse applications. This will allow for an assessment of the model's generalizability across various configurations, considering the physical differences present in different applications. By testing the FDPM model in different contexts, researchers can gain valuable insights into its adaptability and robustness, ultimately enhancing its practical utility and widening its scope of applicability. Particular attention needs to be given to prediction accuracy made at the beginning of the milling cycle with low training share, where data may be scarce, resulting in higher uncertainty in predictions. Furthermore, future research should address any anomalous behaviour that could affect the prediction accuracy of the FDPM model and how effectively the ETS can be used to extrapolate the offset in predictions while affected by anomalous instances. It is essential to note that the FDPM model in this research is limited to detecting blade wear exclusively during normal operation, and any potential anomalous occurrences should be thoroughly investigated to understand their impact on the model's performance in such circumstances. However, this study has appointed significant avenue for improving availability in high-energy milling applications, especially where other wear measurement methods pose implementation challenges.

Appendix 1 illustrates the measured Vb values for each blade. The blades are used in four milling cycles, named T1C4, T1C5, T2C4, and T2C5, which generated the research datasets. The wear on blade surfaces f_1 and f_2 was recorded separately, named Vbf_1 and Vbf_2. The datasets are named using the following format: spindle (T), spindle #, row (C), blade edge #; milled meters per milling cycle.

Operational cycle	T1C4; 2009 m		T1C5; 2505 m			T2C4;2010 m		T2C5;2518 m	
Blade #	Vbf_1	Vbf_2	Vbf_1	Vbf_2	Blade #	Vbf_1	Vbf_2	Vbf_1	Vbf_2
41	0.346	0.072	0.397	0.184	141	0.318	0.146	0.33	0.096
42	0.373	0.168	0.378	0.541	142	0.337	0.149	0.367	0.075
43	0.327	0.126	0.346	0.337	143	0.324	0.147	0.355	0.093
44	0.311	0.143	0.364	0.445	144	0.38	0.184	0.386	0.076
45	0.346	0.217	0.38	0.485	145	0.331	0.149	0.355	0.073
46	0.324	0.114	0.336	0.573	146	0.336	0.187	0.392	0.076
47	0.299	0.148	0.367	0.441	147	0.311	0.112	0.373	0.084
48	0.337	0.112	0.336	0.503	148	0.33	0.149	0.387	0.077
49	0.311	0.15	0.361	0.581	149	0.373	0.151	0.415	0.072
50	0.373	0.148	0.38	0.751	150	0.355	0.146	0.42	0.058
51	0.317	0.185	0.392	0.91	151	0.336	0.148	0.429	0.057
52	0.274	0.148	0.336	0.778	152	0.38	0.148	0.411	0.075
53	0.299	0.146	0.398	0.742	153	0.336	0.151	0.423	0.072
54	0.342	0.187	0.401	0.635	154	0.358	0.167	0.481	0.075
55	0.319	0.14	0.38	0.672	155	0.33	0.149	0.436	0.075
56	0.308	0.149	0.392	0.556	156	0.299	0.184	0.416	0.059
57	0.374	0.151	0.38	0.523	157	0.332	0.113	0.429	0.073
58	0.308	0.148	0.349	0.414	158	0.317	0.149	0.43	0.056

References

Aaltonen, K., Andersson, P., & Kauppinen, V. (1997). *Koneistustekniikat*. WSOY Porvoo Finland.

Airlangga, G., Rachmat, A., & Lapihu, D. (2019). Comparison of exponential smoothing and neural network method to forecast rice production in Indonesia. *TELKOMNIKA (Telecommunication Computing Electronics and Control), 17*(3), 1367. https://doi.org/10.12928/telkomnika.v17i3.11768

An, Q., Tao, Z., Xu, X., El Mansori, M., & Chen, M. (2020). A data-driven model for milling tool remaining useful life prediction with convolutional and stacked LSTM network. *Measurement*, *154*, 107461. https://doi.org/10.1016/j.measurement.2019.107461

Arias Chao, M., Kulkarni, C., Goebel, K., & Fink, O. (2022). Fusing physics-based and deep learning models for prognostics. *Reliability Engineering and System Safety, 217*, 18. https://doi.org/10.1016/j.ress.2021.107961

Bathla, G. (2020). Stock Price prediction using LSTM and SVR. In: *2020 Sixth International Conference on Parallel, Distributed and Grid Computing (PDGC)*, 211–214. https://doi.org/10.1109/PDGC50313.2020.9315800

Caldeirani Filho, J., & Diniz, A. E. (2002). Influence of cutting conditions on tool life, tool wear and surface finish in the face milling process. *Revista Brasileira De Ciencias Mecanicas/journal of the Brazilian Society of Mechanical Sciences, 24*(1), 10–14. https://doi.org/10.1590/S0100-73862002000100002

Chen, S.-H., & Ho, Y.-L. (2020). *Application of Exponential Smoothing to Machining Precision of Nickel-based Superalloy Waspaloy* . https://doi.org/10.18494/SAM.2020.2596

Cheng, G., Peddinti, V., Povey, D., Manohar, V., Khudanpur, S., & Yan, Y. (2017). An Exploration of Dropout with LSTMs. In: *Interspeech 2017*, 1586–1590. https://doi.org/10.21437/Interspeech.2017-129

Cheng, Y., Zhu, H., Hu, K., Wu, J., Shao, X., & Wang, Y. (2019). Multisensory data-driven health degradation monitoring of machining tools by generalized multiclass support vector machine. *IEEE Access, 7*, 47102–47113. https://doi.org/10.1109/ACCESS.2019.2908852

Chicco, D., Warrens, M. J., & Jurman, G. (2021). The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation. *PeerJ Computer Science, 7*, 1–24. https://doi.org/10.7717/PEERJ-CS.623

da Santos, C. E. S., Sampaio, R. C., dos Coelho, L. S., Bestarsd, G. A., & Llanos, C. H. (2021). Multi-objective adaptive differential evolution for SVM/SVR hyperparameters selection. *Pattern Recognition, 110*, 107649. https://doi.org/10.1016/j.patcog.2020.107649

De Beaulieu, M. H., Jha, M. S., Garnier, H., & Cerbah, F. (2022). Unsupervised remaining useful life prediction through long range health index estimation based on encoders-decoders. *IFAC-PapersOnLine, 55*(6), 718–723. https://doi.org/10.1016/j.ifacol.2022.07.212

Elsayed, E. A. (2012). *Reliability engineering* (2nd ed.). Wiley.

Ersvik, E., & Khalid, R. (2015). *Milling in hardened steel—A study of tool wear in conventional- and dynamic milling* (Issue June) [Uppsala Universitet]. https://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-255646. Accessed 31 May 2021.

Feng, K., Ji, J. C., Ni, Q., Li, Y., Mao, W., & Liu, L. (2023). A novel vibration-based prognostic scheme for gear health management in surface wear progression of the intelligent manufacturing system. *Wear*. https://doi.org/10.1016/j.wear.2023.204697

Géron, A. (2017). *Hands-on Machine Learning with Scikit-Learn & TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems* (1st ed.). O'Reilly Media Inc. https://doi.org/10.5555/3153997

He, Z., Shi, T., Xuan, J., & Li, T. (2021). Research on tool wear prediction based on temperature signals and deep learning. *Wear*. https://doi.org/10.1016/j.wear.2021.203902

Heng, A., Zhang, S., Tan, A. C. C., & Mathew, J. (2009). Rotating machinery prognostics: State of the art, challenges and opportunities. *Mechanical Systems and Signal Processing, 23*(3), 724–739. https://doi.org/10.1016/j.ymssp.2008.06.009

Herawan, T., Ghazali, R., Nawi, N. M., & Deris, M. M. (Eds.). (2016). *Recent Advances on Soft Computing and Data Mining: The Second International Conference on Soft Computing and Data Mining (SCDM-2016), Bandung, Indonesia, August 18–20, 2016 Proceedings* (Vol. 549). Springer International Publishing. https://doi.org/10.1007/978-3-319-51281-5

Jiang, J. R., Kao, J. B., & Li, Y. L. (2021). Semi-supervised time series anomaly detection based on statistics and deep learning. *Applied Sciences (switzerland)*. https://doi.org/10.3390/app11156698

Khosravi, H., Olajire, T., Raihan, A. S., & Ahmed, I. (2024). A data driven sequential learning framework to accelerate and optimize multi-objective manufacturing decisions. *Journal of Intelligent Manufacturing*. https://doi.org/10.1007/s10845-024-02337-y

Kim, K.-Y., Park, J., & Sohmshetty, R. (2017). Prediction measurement with mean acceptable error for proper inconsistency in noisy weldability prediction data. *Robotics and Computer-Integrated Manufacturing, 43*, 18–29. https://doi.org/10.1016/j.rcim.2016.01.002

Li, X., Ding, Q., & Sun, J. Q. (2018). Remaining useful life estimation in prognostics using deep convolution neural networks. *Reliability Engineering and System Safety, 172*(October 2017), 1–11. https://doi.org/10.1016/j.ress.2017.11.021

Li, Y., Xiang, Y., Pan, B., & Shi, L. (2022). A hybrid remaining useful life prediction method for cutting tool considering the wear state. *The International Journal of Advanced Manufacturing Technology, 121*(5), 3583–3596. https://doi.org/10.1007/s00170-022-09417-4

Lin, Y., He, S., Lai, D., Wei, J., Ji, Q., Huang, J., & Pan, M. (2020). Wear mechanism and tool life prediction of high-strength vermicular graphite cast iron tools for high-efficiency cutting. *Wear*. https://doi.org/10.1016/j.wear.2020.203319

Lughofer, E., & Sayed-Mouchaweh, M. (2019). Prologue: Predictive maintenance in dynamic systems. In *Predictive Maintenance in Dynamic Systems: Advanced Methods, Decision Support Tools and Real-World Applications*. https://doi.org/10.1007/978-3-030-05645-2_1

Mäkiaho, T., Vainio, H., & Koskinen, K. (2022). Model-based wear prediction of milling machine blades. *Procedia Computer Science, 207*, 1113–1123. https://doi.org/10.1016/J.PROCS.2022.09.167

Mäkiaho, T., Vainio, H., & Koskinen, K. T. (2023). Wear Parameter Diagnostics of Industrial Milling Machine with Support Vector Regression. *MDPI Machines, 11*(3), 395. https://doi.org/10.3390/machines11030395

Mia, M., Khan, M. A., & Dhar, N. R. (2017). High-pressure coolant on flank and rake surfaces of tool in turning of Ti-6Al-4V: Investigations on surface roughness and tool wear. *The International Journal of Advanced Manufacturing Technology, 90*(5–8), 1825–1834. https://doi.org/10.1007/s00170-016-9512-5

Moufki, A., Dudzinski, D., & Le Coz, G. (2015). Prediction of cutting forces from an analytical model of oblique cutting , application to peripheral milling of Ti-6Al-4V alloy. 615–626. https://doi.org/10.1007/s00170-015-7018-1

Myers, L., & Sirois, M. J. (2006). Spearman Correlation Coefficients, Differences between. In *Encyclopedia of Statistical Sciences*. John Wiley & Sons, Ltd. https://doi.org/10.1002/0471667196.ess5050.pub2

Nanni, L., Brahnam, S., Paci, M., & Ghidoni, S. (2022). Comparison of different convolutional neural network activation functions and

methods for building ensembles for small to midsize medical data sets. *Sensors, 22*(16), 6129. https://doi.org/10.3390/s22166129

Nee, A. Y. C. (2015). *Handbook of Manufacturing Engineering and Technology* (1st ed.). Springer London. https://doi.org/10.1007/978-1-4471-4670-4

Nie, L., Zhang, L., Xu, S., Cai, W., & Yang, H. (2022). Remaining useful life prediction of milling cutters based on CNN-BiLSTM and attention mechanism. *Symmetry, 14*(11), 11. https://doi.org/10.3390/sym14112243

Obilor, E. I., & Amadi, E. C. (2018). Test for Significance of Pearson's Correlation Coefficient (r). *International Journal of Innovative Mathematic*, *Jan-Mar*. https://www.researchgate.net/publication/323522779_Test_for_Significance_of_Pearson%27s_Correlation_Coefficient. Accessed 18 July 2023.

Panda, S. S., Chakraborty, D., & Pal, S. K. (2008). Flank wear prediction in drilling using back propagation neural network and radial basis function network. *Applied Soft Computing, 8*(2), 858–871. https://doi.org/10.1016/j.asoc.2007.07.003

Pecht, M. G., & Kang, M. (2018). Prognostics and health management of electronics: Fundamentals, Machine Learning, and the Internet of Things. *John Wiley and Sons Ltd*. https://doi.org/10.1002/9781119515326.ch4

Peng, Y., Dong, M., & Zuo, M. J. (2010). Current status of machine prognostics in condition-based maintenance: A review. *International Journal of Advanced Manufacturing Technology, 50*(1–4), 297–313. https://doi.org/10.1007/s00170-009-2482-0

Poornima, S., & Pushpalatha, M. (2019). Prediction of rainfall using intensified LSTM based recurrent neural network with weighted linear units. *Atmosphere, 10*(11), 668. https://doi.org/10.3390/atmos10110668

Qi, Z., Wang, H., He, T., Li, J., & Gao, H. (2020). FRIEND: Feature selection on inconsistent data. *Neurocomputing, 391*, 52–64. https://doi.org/10.1016/j.neucom.2020.01.094

Rivas-Perea, P., Cota-Ruiz, J., Chaparro, D. G., Venzor, J. A. P., Carreón, A. Q., & Rosiles, J. G. (2013). Support vector machines for regression: A succinct review of large-scale and linear programming formulations. *International Journal of Intelligence Science, 03*(01), 5–14. https://doi.org/10.4236/ijis.2013.31002

Sagheer, A., & Kotb, M. (2019). Unsupervised pre-training of a deep LSTM-based stacked autoencoder for multivariate time series forecasting problems. *Scientific Reports*. https://doi.org/10.1038/s41598-019-55320-6

Sahoo, P., Pratap, T., & Patra, K. (2019). A hybrid modelling approach towards prediction of cutting forces in micro end milling of Ti-6Al-4V titanium alloy. *International Journal of Mechanical Sciences, 150*, 495–509. https://doi.org/10.1016/j.ijmecsci.2018.10.032

Samek, W., Montavon, G., Vedaldi, A., Hansen, L. K., & Müller, K.-R. (Eds.). (2019). *Explainable AI: Interpreting, Explaining and Visualizing Deep Learning* (Vol. 11700). Springer Nature Switzerland AG 2019. https://doi.org/10.1007/978-3-030-28954-6

Sandvik, C. (2017). *Training Handbook, Metal Cutting Technology*. AB Sandvik Coromant. https://www.sandvik.coromant.com/fi-fi/knowledge/materials/pages/workpiece-materials.aspx. Accessed 4 May 2022.

Sarmadi, H., Entezami, A., Behkamal, B., & De Michele, C. (2022). Partially online damage detection using long-term modal data under severe environmental effects by unsupervised feature selection and local metric learning. *Journal of Civil Structural Health Monitoring, 12*(5), 1043–1066. https://doi.org/10.1007/s13349-022-00596-y

Sartin, M. A., & da Silva, A. C. R. (2013). Approximation of hyperbolic tangent activation function using hybrid methods. In: *2013 8th International Workshop on Reconfigurable and Communication-Centric Systems-on-Chip (ReCoSoC)*, 1–6. https://doi.org/10.1109/ReCoSoC.2013.6581545

Schroderus, J., Lasrado, L. A., Menon, K., & Kärkkäinen, H. (2022). Towards a Pay-Per-X Maturity Model for Equipment Manufacturing Companies. *Elsevier, 196*, 226–234. https://doi.org/10.1016/j.procs.2021.12.009

Siddhpura, A., & Paurobally, R. (2013). A review of flank wear prediction methods for tool condition monitoring in a turning process. *International Journal of Advanced Manufacturing Technology, 65*(1–4), 371–393. https://doi.org/10.1007/s00170-012-4177-1

Sikorska, J. Z., Hodkiewicz, M., & Ma, L. (2011). Prognostic modelling options for remaining useful life estimation by industry. *Mechanical Systems and Signal Processing, 25*(5), 1803–1836. https://doi.org/10.1016/j.ymssp.2010.11.018

The American Society of Mechanical Engineers. (1985). *Tool life testing with single-point turning tools.* ANSI/ASME B94.55M (1985). https://www.asme.org/codes-standards/find-codes-standards/b94-55m-tool-life-testing-single-point-turning-tools/1985/drm-enabled-pdf. Accessed 1 Feb 2023.

Tian, M., Fan, H., Xiong, Z., & Li, L. (2023). Data-driven ensemble model for probabilistic prediction of debris-flow volume using Bayesian model averaging. *Bulletin of Engineering Geology and the Environment, 82*(1), 34. https://doi.org/10.1007/s10064-022-03050-x

Tong, X., Wang, J., Zhang, C., Wu, T., Wang, H., & Wang, Y. (2022). LS-LSTM-AE: Power load forecasting via Long-Short series features and LSTM-Autoencoder. *Energy Reports, 8*, 596–603. https://doi.org/10.1016/j.egyr.2021.11.172

Traini, E., Bruno, G., D'Antonio, G., & Lombardi, F. (2019). Machine learning framework for predictive maintenance in milling. *IFAC-PapersOnLine, 52*(13), 177–182. https://doi.org/10.1016/j.ifacol.2019.11.172

Uhlmann, E., Oberschmidt, D., Kuche, Y., & Löwenstein, A. (2014). Cutting edge preparation of micro milling tools. *Procedia CIRP, 14*, 349–354. https://doi.org/10.1016/j.procir.2014.03.083

Uuskoski, M., Kärkkäinen, H., & Menon, K. (2020). Rapid sales growth mechanisms and profitability for investment product manufacturing SMEs through pay-per-X business models. *Product Lifecycle Management Enabling Smart X*. https://doi.org/10.1007/978-3-030-62807-9_32

Wang, J. J., Zheng, Y. H., Zhang, L. B., Duan, L. X., & Zhao, R. (2017). Virtual sensing for gearbox condition monitoring based on kernel factor analysis. *Petroleum Science, 14*(3), 539–548. https://doi.org/10.1007/s12182-017-0163-4

Wang, W., & Lu, Y. (2018). Analysis of the Mean Absolute Error (MAE) and the Root Mean Square Error (RMSE) in assessing rounding model. *IOP Conference Series: Materials Science and Engineering, 324*, 012049. https://doi.org/10.1088/1757-899X/324/1/012049

Wayal, V., Ambhore, N., Chinchanikar, S., & Bhokse, V. (2015). Investigation on cutting force and vibration signals in turning: mathematical modeling using response surface methodology. *Journal of Mechanical Engineering and Automation, 5*(March), 64–68. https://doi.org/10.5923/c.jmea.201502.13

Wu, D., Jennings, C., Terpenny, J., Gao, R. X., & Kumara, S. (2017). A comparative study on machine learning algorithms for smart manufacturing: Tool wear prediction using random forests. *Journal of Manufacturing Science and Engineering, Transactions of the ASME, 139*(7), 1–9. https://doi.org/10.1115/1.4036350

Wu, P., Dai, H., Li, Y., He, Y., Zhong, R., & He, J. (2022). A physics-informed machine learning model for surface roughness prediction in milling operations. *The International Journal of Advanced Manufacturing Technology, 123*, 4065–4076. https://doi.org/10.1007/s00170-022-10470-2

Xie, J., Luo, M.-J., He, J.-L., Liu, X.-R., & Tan, T.-W. (2012). Micro-grinding of micro-groove array on tool rake surface for dry cutting of titanium alloy. *International Journal of Precision Engineering*

and Manufacturing, 13(10), 1845–1852. https://doi.org/10.1007/s12541-012-0242-9

Xu, S., Li, X., Xie, C., Chen, H., Chen, C., & Song, Z. (2021). A high-precision implementation of the sigmoid activation function for computing-in-memory Architecture. *Micromachines, 12*(10), 1183. https://doi.org/10.3390/mi12101183

Yang, Q., Pattipati, K. R., Awasthi, U., & Bollas, G. M. (2022). Hybrid data-driven and model-informed online tool wear detection in milling machines. *Journal of Manufacturing Systems, 63*, 329–343. https://doi.org/10.1016/j.jmsy.2022.04.001

Zeng, X., & Long, L. (2022). Neural networks. In L. Long & X. Zeng (Eds.), *Beginning deep learning with TensorFlow* (pp. 191–234). Apress. https://doi.org/10.1007/978-1-4842-7915-1_6

Zhang, C., Yao, X., Zhang, J., & Jin, H. (2016). Tool condition monitoring and remaining useful life prognostic based on a wireless sensor in dry milling operations. *Sensors (switzerland).* https://doi.org/10.3390/s16060795

Zhang, M., Tao, F., & Nee, A. Y. C. (2021). Digital Twin Enhanced Dynamic Job-Shop Scheduling. *Journal of Manufacturing Systems, 58*, 146–156. https://doi.org/10.1016/j.jmsy.2020.04.008

Zhou, J.-T., Zhao, X., & Gao, J. (2019). Tool remaining useful life prediction method based on LSTM under variable working conditions. *The International Journal of Advanced Manufacturing Technology, 104*(9), 4715–4726. https://doi.org/10.1007/s00170-019-04349-y