CrossMark

Artificial intelligence for automatic prediction of required surface roughness by monitoring wear on face mill teeth

D. Yu. Pimenov[1] · A. Bustillo[2] · T. Mikolajczyk[3]

Abstract

Nowadays, face milling is one of the most widely used machining processes for the generation of flat surfaces. Following international standards, the quality of a machined surface is measured in terms of surface roughness, Ra, a parameter that will decrease with increased tool wear. So, cutting inserts of the milling tool have to be changed before a given surface quality threshold is exceeded. The use of artificial intelligence methods is suggested in this paper for real-time prediction of surface roughness deviations, depending on the main drive power, and taking tool wear, V_B into account. This method ensures comprehensive use of the potential of modern CNC machines that are able to monitor the main drive power, N, in real-time. It can likewise estimate the three parameters -maximum tool wear, machining time, and cutting power- that are required to generate a given surface roughness, thereby making the most efficient use of the cutting tool. A series of artificial intelligence methods are tested: random forest (RF), standard Multilayer perceptrons (MLP), Regression Trees, and radial-based functions. Random forest was shown to have the highest model accuracy, followed by regression trees, displaying higher accuracy than the standard MLP and the radial-basis function. Moreover, RF techniques are easily tuned and generate visual information for direct use by the process engineer, such as the linear relationships between process parameters and roughness, and thresholds for avoiding rapid tool wear. All of this information can be directly extracted from the tree structure or by drawing 3D charts plotting two process inputs and the predicted roughness depending on workshop requirements.

Keywords Face milling · Wear · Random forest · Surface roughness · Cutting power · Processing time

List of symbols

Ra	Surface roughness (μm)
V_B	Flank wear (mm)
N	Cutting power (kW)
t	Processing time (cutting pass) (min)
$L \times B \times H$	Length × width × height of parts (mm)

✉ D. Yu. Pimenov
   danil_u@rambler.ru

   A. Bustillo
   abustillo@ubu.es

   T. Mikolajczyk
   tami@utp.edu.pl

[1] Department of Automated Mechanical Engineering, South Ural State University, Lenin Prosp. 76, Chelyabinsk, Russia 454080

[2] Department of Civil Engineering, University of Burgos, Avda Cantabria s/n, 09006 Burgos, Spain

[3] Department of Production Engineering, UTP University of Science and Technology, Al. Prof. S. Kaliskiego 7, 85-796 Bydgoszcz, Poland

D	Cutter diameter (mm)
α	Back angle (°)
γ	Rake angle (°)
λ	Angle of the main cutting edge (°)
k_r	Major cutting edge angle (°)
k_{r1}	Minor cutting edge angle (°)
z	Number of teeth
d	Depth of cut (mm)
V	Cutting speed (m/min)
f_z	Feed per tooth (mm/tooth)
n	Spindle rotation speed (rpm)
T	Work cycle time (machining pass) (min)
N_{min}	Minimum drive power (kW)
N_{max}	Maximum drive power (kW)
N_{mean}	Average drive power (kW)

Introduction

Face milling is widely used today to obtain flat surfaces. One of the main precision criteria in face milling is sur-

face roughness, *Ra*. Surface roughness has in various papers (Fernández-Valdivielso et al. 2016; Pimenov 2013a, 2014) been shown to increase with increased flank wear of face mill teeth. Workpiece specifications will usually detail the maximum roughness value that has to be machined. As the worn surface of the cutting edge increases, the cutting edge of the replaceable indexable inserts (Gong et al. 2017; Mikoła-jczyk et al. 2017, 2018) must therefore be changed, before the maximum surface roughness is reached. Such an approach permits optimum use of the cutting edges of the replaceable indexable inserts (Grigoriev et al. 2015; Machado and Diniz 2017; Niaki and Mears 2017), while ensuring the specified surface roughness. Thus, the most relevant task is to guarantee a given surface roughness as the wear on the face mill surface increases (Pimenov 2013b). As we approach the sixth technological revolution, the analysis of large volumes of data with artificial intelligence (AI) and the integration of AI algorithms in computer-aided production are becoming increasingly relevant.

Many researchers have studied surface roughness in face milling (Baek et al. 1997; Adamczak et al. 2009; Miko and Nowakowski 2012; Arizmendi et al. 2009; Rosales et al. 2010; Muñoz-Escalona and Maropoulos 2015; Felho et al. 2015; Zhenyu et al. 2015; Moghaddam and Kolahan 2016; Popov and Schindelarz 2017; Jersák and Simon 2017). Baek et al. (1997) presented a mathematical model for surface roughness prediction taking into account the dynamic characteristics of the face-milling operation. Adamczak et al. (2009) and Miko and Nowakowski (2012) developed a generalized mathematical model of surface roughness formation for surfaces shaped with round-nose tools. Arizmendi et al. (2009) presented a model for the prediction of surface topography in peripheral milling operations processing signals that capture tool vibration during the cutting process. However, tool wear is not a parameter of that model. Rosales et al. (2010) investigated the dependence of rotation speed, feed rate, cutting depth, tool geometry, and runout errors in face milling. Muñoz-Escalona and Maropoulos (2015) designed a geometrical model for the prediction of surface roughness in face milling with square insert tools. Felho et al. (2015) presented the estimated relations between the calculated theoretical and measured real roughness values, allowing researchers to predict the machined surface roughness. Zhenyu et al. (2015) suggested an algorithm for predicting surface roughness that takes into account the influence of both static and dynamic factors on the roughness of the face milled surface. Moghaddam and Kolahan (2016) studied the influence of face milling parameters [cutting speed, (V), feed rate (f_z), and depth of cut, (a_p)] on surface roughness, *Ra,* in the face milling of AISI1045 steel. Popov and Schindelarz (2017) investigated the effect of hydraulic oil entering the cutting fluid on tool life and roughness when milling stainless steel. Jersák and Simon (2017) studied the influence of cooling lubricants

on surface roughness and the energy efficiency of cutting machine tools for the three following techniques: lubricant-free cutting, cutting with the use of a lubricant with the MQL technique, and only utilizing finish-turning and finish-face milling. However, no account of the impact of face mill wear on surface roughness as it forms has been found in these references (Baek et al. 1997; Adamczak et al. 2009; Miko and Nowakowski 2012; Arizmendi et al. 2009; Rosales et al. 2010; Muñoz-Escalona and Maropoulos 2015; Felho et al. 2015; Zhenyu et al. 2015; Moghaddam and Kolahan 2016; Popov and Schindelarz 2017; Jersák and Simon 2017).

Much of the research devoted to the study of surface roughness in face milling (Diniz and Filho 1999; Caldeirani Filho and Diniz 2002; De Souza et al. 2003, 2005; De Escalona and Maropoulos 2010; Prasad et al. 2011; Houchuan et al. 2015; Liu et al. 2016; Shi et al. 2016; Werda et al. 2017) takes tool wear into account. Diniz and Filho (1999) studied the influence of the relative positions of tool and workpiece on tool life, tool wear, and surface finish in the face milling process, taking tool wear into account. Caldeirani Filho and Diniz (2002) investigated the influence of cutting speed and feed per tooth on tool life and the surface roughness of machined surfaces during mechanical milling, taking into account tool wear. De Souza et al. (2003) defined the parameters of surface roughness, *Ra*, tool life, and burr formation and used them when comparing the performances of two face-milling cutter systems with PCBN (polycrystalline citric boron nitrite) tools. De Souza et al. (2005) defined the parameters of surface roughness, waviness, tool life (based on flank wear), and burr formation and used them to compare the performance of the A system with $24Si_3N_4$ ceramic inserts and the B system in face milling. De Escalona and Maropoulos (2010) studied the influence of feed per tooth and cutting speed on tool wear and surface roughness in the face milling of martensitic stainless steel. Prasad et al. (2011) studied the impact of offsets due to vibration during face milling and the correlation of surface roughness with tool wear under different cutting conditions while monitoring the tool state. Houchuan et al. (2015) investigated surface roughness, machining defects, microhardness, and microstructure variations at different cutting speeds and tool average flank wear values in the face milling of a titanium alloy, Ti–10V–2Fe–3Al (Ti-1023). Liu et al. (2016) investigated the process of tool damage and its effect on machined surface roughness in high-speed face milling of 17-4PH stainless steel. Shi et al. (2016) investigated tool wear behaviors and their effect on machinability in dry high-speed milling of magnesium alloy at high cutting speeds. Werda et al. (2017) compared tool life and surface roughness in the milling of X100CrMoV5 mold steel under different lubrication conditions: dry machining and minimum quantity lubrication (MQL). Even though many authors (Diniz and Filho 1999; Caldeirani Filho and Diniz 2002; De

Souza et al. 2003, 2005; De Escalona and Maropoulos 2010; Prasad et al. 2011; Houchuan et al. 2015; Liu et al. 2016; Shi et al. 2016; Werda et al. 2017) have presented the results of surface roughness studies that take into account tool wear, no solutions to the problem of how to establish a comprehensive correlation of surface roughness, tool wear, and processing time have been advanced from the perspective of managing these settings in automated production. This task is however made possible with AI.

Many researchers have developed surface roughness prediction models in face milling using AI (Srinivasa Pai et al. 2002; Benardos and Vosniakos 2002, 2003; Saglam and Unuvar 2003; Bruni et al. 2008; El-Sonbaty et al. 2008; Lela et al. 2009; Muñoz-Escalona and Maropoulos 2010; Razfar et al. 2011; Bharathi Raja and Baskar 2012; Grzenda et al. 2012; Bajić et al. 2012; Kovac et al. 2013; Simunovic et al. 2013; Grzenda and Bustillo 2013; Elhami et al. 2013; Saric et al. 2013; Rodríguez et al. 2017; Simunovic et al. 2016; Selaimia et al. 2017; Svalina et al. 2017). Srinivasa Pai et al. (2002) presented an estimation of flank wear in face milling based on the radial basis function (RBF) of neural networks using acoustic emission signals, surface roughness, and cutting conditions (cutting speed and feed). Benardos and Vosniakos (2002) introduced the neural network modeling approach to predict surface roughness, *Ra,* in Computer Numerical Control (CNC) machine face milling. The paper takes into account depth of cut, the feed rate per tooth, the cutting speed, the use of cutting fluid, and cutting tool wear. Saglam and Unuvar (2003) suggested the use of a multilayer neural network for status monitoring and the estimation of flank wear and surface roughness in face milling. Benardos and Vosniakos (2003) aimed to present the various methodologies and practices used in surface roughness prediction. Bruni et al. (2008) offered analytic and Artificial Neural Network (ANN) models for surface roughness prediction at various cutting speeds and under various cooling conditions in face milling finishing processes on AISI 420 B stainless steel. El-Sonbaty et al. (2008) used ANNs to develop models for predicting the correlation between the cutting conditions and the corresponding fractal parameters of machined surfaces in face-milling operations, including surface roughness, without taking into account the cutting tool wear. Lela et al. (2009) studied the influence of cutting speed, feed, and depth of cut on surface roughness in face milling. Three different modeling methodologies, namely regression analysis (RA), support vector machine (SVM), and Bayesian neural network (BNN), have been applied to data determined in the process of experimental design. Muñoz-Escalona and Maropoulos (2010) developed various ANNs for predicting the surface roughness (*Ra*) of 7075-T7351 aluminum alloys in face milling. Razfar et al. (2011) presented an approach that defines the optimum cutting parameters that provide minimum surface roughness in face milling of X20Cr13 steel

by combining an ANN and the harmony search algorithm. Bharathi Raja and Baskar (2012) conducted experimental studies of the influence of machining parameters such as the cutting speed, feed rate, and depth of cut on the surface roughness of aluminum and the provision of design surface roughness in face milling using Particle Swarm Optimization (PSO). Grzenda et al. (2012) presented a new strategy for improving the AI models for predicting surface roughness using small datasets tested in high-torque milling operations. Bajić et al. (2012) examined the impact of various cutting speeds, cutting depth, and tool wear on surface roughness and the cutting force components. Kovac et al. (2013) studied the influence of machining parameters on surface roughness in face milling, comparing prediction models based on fuzzy logic and RA. Simunovic et al. (2013) presented a study of surface roughness in the face machining of an aluminum alloy at low cutting speed using neural networks; however no changes in the degree of tool wear were monitored. Grzenda and Bustillo (2013) focused on the initial data transformation and its effect on the prediction of surface roughness in high-torque face milling operations. Elhami et al. (2013) suggested a genetically optimized neural network system for the prediction of constrained optimal cutting conditions in the face milling of a high-silicon austenitic stainless steel, in order to minimize surface roughness. Saric et al. (2013) presented a study of the prediction of machined surface roughness in the face milling of steel at various numbers of revolutions, cutting speeds, feeds, and depths. Rodríguez et al. (2017) presented an AI-based decision-making tool for selection of the right cutting tools for face milling, one of the criteria for which is the roughness of the machined surface. Simunovic et al. (2016) presented a machined surface roughness investigation based on the features of a digital image such as spindle speed, feed per tooth, and cutting depth, but without considering tool wear; the digital image was produced following a milling operation of an aluminum alloy Al6060. Selaimia et al. (2017) modeled the output responses, namely: surface roughness (*Ra*), cutting force (*FC*), cutting power (*PC*), specific cutting force (*KS*) and metal removal rate (*MRR*) during the face milling of the austenitic stainless steel X2CrNi18-9 with coated carbide inserts (GC4040). ANOVA was used for evaluating the influence of the cutting parameters: cutting speed (*VC*), feed per tooth and depth of cut (*aP*) on the output responses. Svalina et al. (2017) presented an evolutionary neuro-fuzzy system for recommending optimal cutting parameters and controlling surface roughness during mechanical milling without taking account of tool wear. A further development is a solution to the problem of managing surface roughness in a complex correlation of tool wear, time of machining, and cutting force.

Besides these well-established AI techniques, ensemble methods (Kuncheva 2014) simultaneously use several AI models, where each model provides its own prediction and all

the predictions are combined. The high accuracy of ensemble predictions has been demonstrated in many milling processes: Bustillo et al. (2011a) proposed the use of ensembles to predict surface roughness in ball-end milling operations. Maudes et al. (2017) used random forest (RF) ensembles for the prediction of dimensional parameters in laser micro-manufacturing of stents, and Ferreiro and Sierra (2012) used different kinds of ensembles for burr detection in a dry-drilling process on aluminum Al 7075-T6.

Finally, although the application of decision or regression trees to roughness prediction in face-milling operations is not very common, it is necessary to stress that decision trees are one of the most common machine-learning techniques, due to their ability to generate clear rules that the final user of the model can use immediately. Some complex manufacturing problems like laser polishing (Bustillo et al. 2011b), maintenance planning of five-axis milling (Freiburg et al. 2014), laser micromachining of cavities (Teixidor et al. 2015), and deep drilling (Bustillo et al. 2016) have been successfully modeled using regression trees. In all these cases, the decision trees create models that are as accurate as other standard machine-learning techniques, such as ANNs, and are suitable for both continuous and discrete outputs.

As the worn surface of the face mill tooth increases, the cutting force also increases. López De Lacalle et al. (2006) proposed an online correlation, taking account of the geometry of the machined surface and the cutting forces generated during surface milling, by means of a monitoring system capable of detecting cutting forces and cutter tool position of the. Rivero et al. (2008) evaluated the suitability of a tool wear monitoring system based on machine tool internal cutting force signals in dry high-speed milling of aluminium alloys. (Guzeev and Pimenov 2011) presented a mathematical model of cutting forces using various cutting conditions, material properties, and the flank worn surface of face mills as variables. Compeán et al. (2012) described a three-stage multivariable tool used for high-performance milling and the modelling of its dynamic behaviour to predict stability against chatter. However, tool wear was not considered in the paper. (Dugin and Popov 2013) investigated ways of increasing the accuracy of the effect of processing materials and cutting tool wear on the plowing force values. (Pimenov and Guzeev 2017) developed a mathematical model of plowing force to account for flank wear. Artetxe et al. (2017) described the implementation of a cutting force prediction model for milling that introduces radial engagement reduction caused by tool runout and workpiece flexibility, although tool wear is not considered.

An increase in the worn surface of face mill teeth also leads to an increase in cutting power. Shao et al. (2004) described a cutting power model for face milling using a cutting power threshold-updating strategy for tool-wear monitoring. Bhattacharyya et al. (2008) suggested a method for continuous

online evaluation of tool wear in face milling based on inexpensive measurements of current and voltage of the spindle motor. Bhattacharyya and Sengupta (2009) used a combination of signal processing techniques to obtain improved and robust estimates of tool wear. da Silva et al. (2016) demonstrated the use of a probabilistic neural network in monitoring tool wear in the end-milling operation via acoustic emission and cutting power signals. Pimenov (2015) suggested a mathematical model of main drive power that can be used for controlling face milling conditions in the process of cutting tool wear. Niaki et al. (2015) used two methods of stochastic filtering, Kalman and particle filter, to predict flank tool wear when machining difficult-to-machine materials through spindle power consumption measurements, but without considering the the correlation of these parameters with the roughness of the machined surface. Urbikain et al. (2017) proposed a methodology to reduce machining problems by means of a simulation utility, which uses the main variables of the system and process as input data, and generates results that help in the proper decision-making and machining plan. Direct benefits can be found in (a) the fixture/clamping optimal design; (b) the machine tool configuration; (c) the definition of chatter-free optimum cutting conditions and (d) the right programming of cutting toolpaths at the Computer Aided Manufacturing (CAM) stage. Many modern machines have such systems, which means that they are low-cost and can be used for online monitoring (Niaki and Mears 2017; Niaki et al. 2015).

It is possible to manage the process of changing surface roughness directly by controlling the amount of tool wear through dimensional wear. To do so, the CNC machine needs to have a contact control sensor, for instance a Renishaw. In this case, wear is monitored between tool replacements. Whenever the tool wear reaches a point where the design surface roughness cannot be attained or the tool wear approaches maximum values, forced cutting tool replacement is necessary. In this case, the direct method of controlling tool wear will be used (Niaki and Mears 2017).

However, not all machines are equipped with such sensors. Many CNC machines, though, are equipped with feedback sensors that, for instance, control the drive power and in particular the main drive power. As soon as the correlation between the tool wear and the main drive power has been established, the specified surface roughness can be attained by monitoring the current power value during the machining process.

The first two variants are impossible to implement with universal manual machine tools. However, by establishing a correlation between the worn surface of the tool and the processing time, the specified surface roughness can be obtained by monitoring the current processing time, for example, through the number of processed machine parts. Cutting tool replacement is forced whenever the current processing time

reaches the point where the design surface roughness can no longer be attained or the time value approaches the maximum. In this case, an indirect method of controlling tool wear will be used employing the current machining time.

The disadvantage of the first and third methods is that they can not be implemented in the pause during tool replacement. If, for instance, the tool is damaged during machining, real-time tool replacement is not possible with these methods; a drawback that is not shared by the second method. It may be avoided by combining the first method with the third method, provided that the machine has such a capability.

However, to implement the above methods, it is necessary to have the relevant models of surface roughness and main drive power depending on tool wear.

Thus, the purpose of this work is to create models for predicting the roughness of machined surfaces in a complex correlation between tool wear, machining time, and cutting power using AI with the aim of integrating AI algorithms in online monitoring of automated manufacturing. Furthermore, the analysis of how these models can provide useful and immediate information for the process engineer is considered: in some cases, such as regression or decision trees, this information can be directly extracted from the model structure, but in other cases, such as ANN models, interpretation of the black-box structure is not easy and the development of 3D charts becomes necessary (Bustillo et al. 2016).

Materials and method

Experimental face milling was performed, in order to establish a complex correlation between the machined surface roughness Ra, size of the flank surface V_B, input cutting power N, and machining time, t. To do so, a machine part manufactured from 45 steel with the dimensions $L = 200\,\text{mm} \times B = 75\,\text{mm} \times H = 100\,\text{mm}$ was machined on a Mori Seiki NMV 5000 CNC machining center for drilling, milling, and boring. The experiments were conducted in the Engineering Scientific and Educational Center of South Ural State University.

The composition of carbon-quality structural steel 45 (similar to ANSI 1045) according to GOST 1050-88 is listed in Table 1.

The hardness of the workpiece was HB190 when tested with a Brinell indenter TB 5004-03. The machining process included cooling and was carried out using a Pramet 8230 insert (the manufacturer recommends using the hard 8230 alloy in milling toughened corrosion-resistant alloy steels and dedicated alloys). The main cutting parameters of the tool are specified in Table 2.

All cutting parameters employed for the face milling operations are specified in Table 3. The geometry is acceptable for

Table 1 Chemical composition and hardness of 45 steels

Workpiece material	Chemical composition (%)										Brinell hardness, HB
	Carbon, C	Silicon, Si	Magnesium, Mn	Nickel, Ni	Sulfur, S	Phosphorus, P	Chrome, Cr	Cuprum, Cu	Arsenic, As	Iron, Fe	
Steel 45 (ANSI 1045)	0.42–0.5	0.17–0.37	0.5–0.8	to 0.25	to 0.04	to 0.035	to 0.25	to 0.25	to 0.08	≈97	190

Table 2 Tool parameters

Tool cutter material	Cutter diameter, D, mm	Rake angle, γ, °	Back angle, α, °	Angle of the main cutting edge, λ, °	Major cutting edge angle, k_r, °	Minor cutting edge angle, k_{r1}, °	Number of teeth, z
Coated with carbide material 8230 from the Pramet Company	100	20	8	−5	45	45	1

mild and mid-carbon steels. These cutting conditions were selected following the cutting tool provider specifications and the machine-tool capabilities. Although cutting conditions are regularly changed in experiments that focus on roughness quality prediction, in order to generate extensive datasets (Bustillo et al. 2011a), the number of different cutting conditions is strongly reduced (Prasad et al. 2011) or just limited to one cutting condition, as in this research (Rivero et al. 2008), in the case of including tool wear.

The main drive power was established in accordance with the online readings made using a Mori Seiki NMV 5000 during machining. The minimum and maximum readings were recorded at the beginning, in the middle, and at the end of the mill operating cycle as percentages of the machine power equal to 22 kW. There are therefore six cycles in each experiment ($k = 6$).

After each cycle, a macrograph of the flank surface of the face mill tooth was taken. The images were processed on a personal computer and the flank wear land of the mill tooth was measured by comparing the image with a dimensional ruler (see Fig. 1).

Surface roughness measurements, Ra, were taken from Abris-PM7.0 profilometer readings. The measurements were recorded for the basic length $L = 0.4$ mm at the beginning (point 1), in the middle (point 2), and at the end (point 3) of the working stroke of the mill. Points 1 and 3 are located at a distance of 90 mm from the center of the part (see Fig. 2). The processing time (cutting pass), t, values in each working stroke are established as follows: $t = L/fzn$ (see Table 3).

Thus, the experimental points were the main drive power, N, size of flank wear-land of the mill tooth for various worn surfaces, V_B, processing time, t, and roughness of the machined surface, Ra, and the experimental results for each point are shown in Table 4.

Graphs (a–d) of the average experimental values (see Table 4) are shown in Fig. 3: a—drive power primary motion, N, from Processing time, t; b—surface roughness, Ra, from Processing time, t; c—drive power primary motion N from Flank wear V_B; d—surface roughness Ra from Flank wear V_B.

Modeling

The dataset was generated from the experimental results listed in Table 4. Each dataset included five inputs: the first two were the processing time (t) and flank wear (V_B), and the last three were minimum (N_{min}) and maximum (N_{max}) drive power and average values (N_{mean}). The output is the measured surface roughness, Ra. All the inputs and the output are continuous numeric variables. As described in "Modeling" section, for each of the 35 experiments, the roughness and the drive power were evaluated in three positions or ranges: at

Table 3 Specifications of cutting conditions

Depth of cut, d, mm	Cutting speed, V, m/min	Feed per tooth, f_z, mm/tooth	Spindle rotation speed, n, rpm	Length of the workpiece, L, mm	Processing time (cutting pass), t, min	Work cycle time (machining pass), T, min
1	188.5	0.2	600	200	1.67	2.58
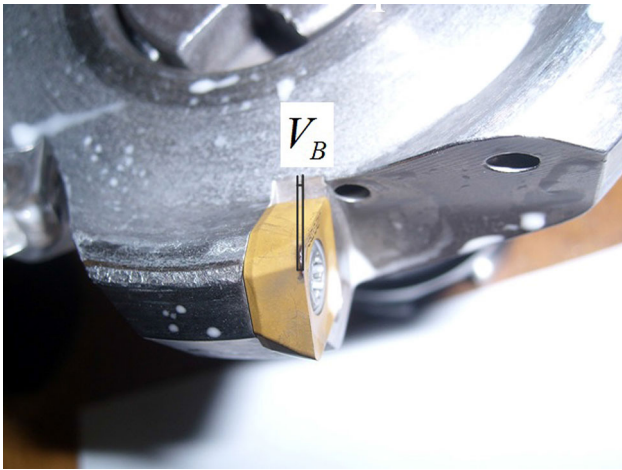
Fig. 1 Wear of changeable indexable insert (CII) on the plowing surface: flank wear V_B
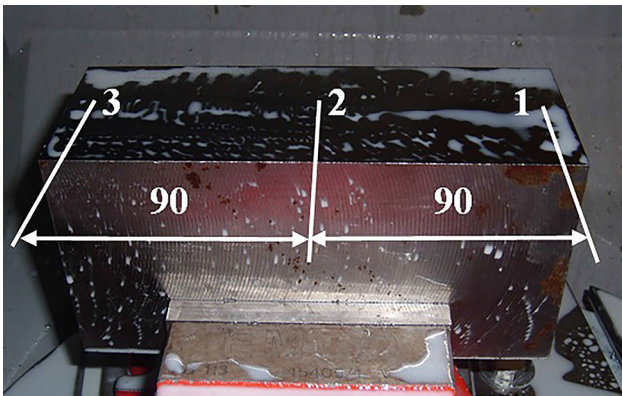
Fig. 2 Scheme of control points of surface roughness Ra

the beginning, in the middle, and at the end. Therefore each experimental test generated three instances of the dataset for each of these positions; the processing time and the flank wear were extrapolated considering each experiment and the previous experiment. Therefore the datasets include 105 instances. Table 5 summarizes the inputs and outputs, their units, and the range of values presented in the dataset; the output variable is shown in bold.

Surface roughness prediction is a regression problem from the AI point of view and the machine-learning techniques are known as regressors. The dataset is randomly divided into two sets: the first one with the instances used to train the model is called the training set, while the second one, used to measure accuracy, is called the validation set. Therefore the validation set includes only instances that have not been used during the training stage, limiting the overfitting tendency of the prediction models; i.e. their natural tendency to fit the training dataset perfectly and their loss of accurate predictive capability under new conditions. In this way it is possible to evaluate the generalization capabilities of the prediction model to deal with new instances.

When talking about roughness prediction (Benardos and Vosniakos 2003), the quality of the predictions is usually measured with the Root Mean Square Error (RMSE), calculated with Eq. 1. This indicator is the squared root of the sum of squares of the prediction errors for each instance divided by the number of instances. Although the RMSE does not provide a physical measure of the error variable such as, for example, the mean error, it has the advantage of penalizing the models with very wrong predictions for some instances, as it considers the squares of the errors and not the errors themselves. Therefore the RMSE is more suitable than the mean error for this study, because a roughness prediction that is far from the real value will mean the production of a workpiece that does not meet the end-user's requirements.

$$RMSE = \sqrt{\frac{\sum_{t=1}^{n} (\hat{y}_t - y)^2}{n}} \tag{1}$$

The following regressors have been tested using Weka machine learning software (Hall et al. 2009): multilayer perceptrons (MLP), radial basis functions (RBFs), regression trees and random forest (RF). Besides, two baseline methods, ZeroR and linear regression, were tested as baseline methods.

MLPs and RBFs, the most common ANN typologies, are considered as standards in the prediction of surface roughness in machining processes (Benardos and Vosniakos 2003). MLPs use a back-propagation algorithm to calculate the weights for each connection between neurons (Bishop 1995), while RBFs use a radial basis function as the activation function in the neurons of the hidden layer (Leonard and Kramer 1991). While MLPs can have more than one hidden layer of neurons, there is only one layer in an RBF that is able to perform non-linear calculations directly.

In contrast, ANNs are considered black-box models, because the end-user cannot extract direct knowledge from them. Regression trees provide visual models of special

Table 4 Experimental results of the drive power primary motion, N; processing time, t; flank wear, V_B; and, surface roughness, Ra

| No. | Drive power primary motion, N, kW | | | | | | | | | | Processing time, t, min | Flank wear, V_B, mm | Surface roughness, Ra, μm | | | |
| | Beginning (1) | | | Center (2) | | | End (3) | | | Average | | | | | | |
	Min	Max	Average	Min	Max	Average	Min	Max	Average				1	2	3	Average
0	0.88	1.54	1.21	1.1	1.54	1.32	0.88	1.54	1.21	1.25	0	0	0.37	0.36	0.37	0.37
1	1.1	1.76	1.43	1.1	1.54	1.32	1.1	1.76	1.43	1.39	1.67	0.08	0.39	0.37	0.39	0.38
2	1.32	1.54	1.43	1.1	1.54	1.32	1.32	1.76	1.54	1.43	3.34	0.16	0.38	0.35	0.47	0.4
3	1.32	1.54	1.43	1.32	1.54	1.43	1.1	1.54	1.32	1.39	5.01	0.33	0.41	0.41	0.49	0.44
4	1.1	1.54	1.32	1.32	1.76	1.54	1.32	1.98	1.65	1.5	6.68	0.39	0.47	0.45	0.56	0.49
5	1.32	1.76	1.54	1.32	1.54	1.43	1.32	1.54	1.43	1.47	8.35	0.39	0.51	0.53	0.59	0.54
6	1.32	1.54	1.43	1.1	1.54	1.32	1.1	1.32	1.21	1.32	10.02	0.46	0.59	0.63	0.59	0.6
7	1.32	1.54	1.43	1.32	1.76	1.54	1.32	1.76	1.54	1.5	11.69	0.48	0.63	0.78	0.84	0.75
8	1.1	1.54	1.32	1.32	1.76	1.54	1.32	1.54	1.43	1.43	13.36	0.52	0.92	0.73	0.67	0.77
9	1.32	1.76	1.54	1.54	1.98	1.76	1.54	1.76	1.65	1.65	15.03	0.52	0.85	0.94	0.84	0.88
10	1.54	1.98	1.76	1.54	1.98	1.76	1.1	1.32	1.21	1.58	16.7	0.54	0.89	0.99	0.85	0.91
11	1.54	1.76	1.65	1.76	1.98	1.87	1.54	1.76	1.65	1.72	18.37	0.55	0.85	0.83	0.76	0.81
12	1.54	1.76	1.65	1.76	1.98	1.87	1.32	1.76	1.54	1.69	20.04	0.57	0.96	1.12	1.09	1.06
13	1.76	1.98	1.87	1.98	1.98	1.98	1.32	1.54	1.43	1.76	21.71	0.57	1.08	0.92	1.18	1.06
14	1.76	1.98	1.87	1.76	1.98	1.87	1.54	1.54	1.54	1.76	23.38	0.58	1.07	1.01	0.95	1.01
15	1.54	1.76	1.65	1.76	1.98	1.87	1.76	1.98	1.87	1.8	25.05	0.58	1.22	1	1.04	1.09
16	1.76	1.98	1.87	1.54	1.76	1.65	1.54	1.76	1.65	1.72	26.72	0.58	1.32	1.25	0.87	1.15
17	1.76	1.98	1.87	1.98	2.2	2.09	1.54	1.76	1.65	1.87	28.39	0.58	1.26	1.16	1.03	1.15
18	1.76	1.98	1.87	1.54	1.76	1.65	1.76	1.98	1.87	1.8	30.06	0.59	0.84	0.85	0.74	0.81
19	1.76	1.98	1.87	1.54	1.76	1.65	1.76	1.98	1.87	1.8	31.73	0.59	0.87	0.6	0.61	0.69
20	1.54	1.76	1.65	1.98	2.2	2.09	1.98	2.2	2.09	1.94	33.4	0.59	0.27	0.49	0.4	0.39
21	1.98	2.2	2.09	1.98	2.2	2.09	1.98	2.2	2.09	2.09	35.07	0.6	0.25	0.39	0.91	0.52
22	1.98	2.2	2.09	2.2	2.42	2.31	2.2	2.42	2.31	2.24	36.74	0.61	0.5	0.49	0.74	0.58
23	2.2	2.42	2.31	1.98	2.2	2.09	2.42	2.64	2.53	2.31	38.41	0.62	0.42	0.55	0.9	0.62
24	1.98	2.2	2.09	1.98	2.2	2.09	2.2	2.42	2.31	2.16	40.08	0.64	1.32	0.95	0.97	1.08
25	1.98	2.2	2.09	2.64	2.64	2.64	2.42	2.64	2.53	2.42	41.75	0.64	1.35	1.42	1.29	1.35
26	2.42	2.64	2.53	2.64	2.86	2.75	2.64	2.86	2.75	2.68	43.42	0.65	1.69	1.45	1.62	1.59
27	2.42	2.42	2.42	2.86	3.08	2.97	2.64	2.64	2.64	2.68	45.09	0.69	1.82	1.73	1.62	1.72
28	1.98	2.2	2.09	2.86	2.86	2.86	2.86	3.08	2.97	2.64	46.76	0.69	2.33	1.91	1.75	2
29	2.86	2.86	2.86	2.64	2.64	2.64	3.08	3.3	3.19	2.9	48.43	0.7	2.33	2.63	2.07	2.34
30	2.86	3.08	2.97	2.86	3.08	2.97	2.86	2.86	2.86	2.93	50.1	0.7	2.24	2.25	1.89	2.13
31	2.64	2.86	2.75	2.86	3.08	2.97	2.86	2.86	2.86	2.86	51.77	0.7	2.73	2.52	1.93	2.39

Table 4 continued

No.	Drive power primary motion, N, kW										Processing time, t, min	Flank wear, V_B, mm	Surface roughness, Ra, μm			
	Beginning (1)			Center (2)			End (3)			Average			1	2	3	Average
	Min	Max	Average	Min	Max	Average	Min	Max	Average							
32	2.86	2.86	2.86	3.08	3.08	3.08	2.86	2.86	2.86	2.93	53.44	0.71	3.08	3.11	2.46	2.88
33	2.64	2.64	2.64	3.08	3.52	3.3	3.52	3.52	3.52	3.15	55.11	0.73	3.07	3.19	3.15	3.14
34	3.08	3.3	3.19	3.74	3.74	3.74	2.86	2.86	2.86	3.26	56.78	0.76	4.25	4.28	3.55	4.03

interest for the process engineers who are in this case the end-users. The engineer should begin to read a regression tree from the upside (root node) and follow the branch that fits the process conditions defined by a certain combination of model inputs until reaching the downside of the tree: the final leaf that provides a linear regression model of the surface roughness. The attributes and numeric thresholds in each decision node are fixed by the M5P algorithm (Muñoz-Escalona and Maropoulos 2010) when it subdivides the training instances into pure subsets (i.e., subsets where almost all training instances are prone to fit the final linear model included in the leaf). The M5P criterion for obtaining pure subsets is to maximize the expected error reduction (Quinlan 1992). Figure 5 in "Results and discussion" section shows an example of the regression trees built in this research.

Regression ensembles (Mendes-Moreira et al. 2012) were also tested to improve the accuracy of regression trees. The ensembles are combinations of base regressors, referred to as regression trees in this study, that combine their individual predictions in a final result. In this research, the base regressors are Random Trees, a variant of regression trees where a random subset of the attributes is considered when selecting the decision for each node. RF (Breiman 1996) is an ensemble built with Random Trees that are trained with different sub-datasets of the training dataset (Breiman 2001) to make different predictions and the final result is the average of the predictions of each tree.

The last two regressors are often used as baseline methods for the purpose of comparison with the naïve approach and the regression model, as stated by previous works on surface roughness (Maudes et al. 2017; Teixidor et al. 2015). The naïve approach uses the mean value of the output as a prediction that is independent of the input values; in this study, the naïve approach will always predict a surface roughness of $1.20\ \mu$m (the mean roughness value considering the whole dataset). As a consequence, the error of the prediction model should be significantly lower than the error for the baseline method to assure the ability of the machine-learning model to predict new situations in terms of drive power and tool wear. In the Weka software tool, the naïve approach is called ZeroR. The second baseline, a linear regression model, was selected because it is usually simple enough to fit the main relationship between the process inputs and the output, at least in some input ranges, besides having extensive applications in real workshops.

The technique selected to train the models and for their validation was tenfold cross-validation repeated 10 times (Kohavi 1995). In this technique, the original dataset is, first of all, randomly divided into 10 equal-size datasets called folds; then, ninefolds are used to train the regressor and the last fold is used for its validation. As previously mentioned, regressor accuracy is thereby measured on instances that have
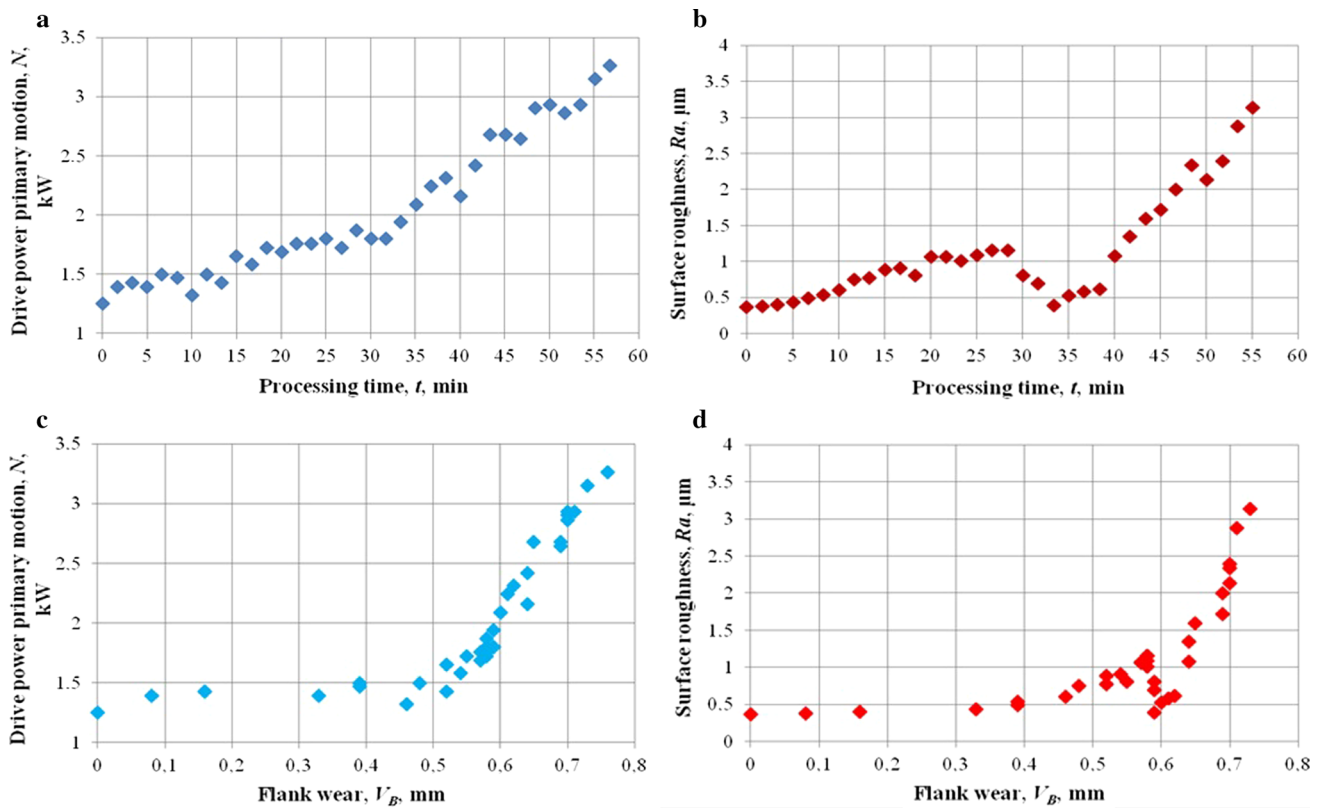
Fig. 3 Graphs **a**–**d** of the average experimental values: **a** drive power primary motion, N, from Processing time, t; **b** surface roughness, Ra, from Processing time, t; **c** drive power primary motion, N, from Flank wear, V_B; **d** surface roughness, Ra, from flank wear, V_B

Table 5 Dataset variables and their variation range

Variable	Abbreviation	Range	Units
Processing time	t	0–57.9	min
Flank wear	V_B	0–0.76	mm
Minimum drive power	N_{min}	0.88–3.74	KW
Maximum drive power	N_{max}	1.32–3.74	KW
Averaged drive power	N_{mean}	1.21–3.74	KW
Surface roughness	**Ra**	**0.25–4.28**	**μm**

not been used to train the model. This process is repeated 10 times, each time using a different fold for validation and the other nine for training, and the RMSE of the validation folds in the 10 repetitions is averaged. Hence, the variance of the regressor accuracy is reduced and its prediction results can be generalized (Kohavi 1995).

All the machine-learning algorithms have different parameters that should be tuned to find the optimal values for each dataset. A grid search was performed on the main parameters of each regressor, to provide a general overview of the effect of each tuning parameter; the change in accuracy of the prediction model was used to fix the number of steps and the increment in the search grid: if the parameter has an influence of more than 50%, the tuning range is divided into four steps; if the accuracy of the model changes due to a parameter with an extreme value in the tuning range that is higher than 50%, the parameter will be tuned in steps of the same size as the lowest parameter value, while if the change is lower than 50%, each step will double the value of the parameter until the upper extreme is achieved. In this grid search, all the values considered for any tunable parameter are tested against the other parameters, as in a full-factorial experiment; this extensive build-up of prediction models can only be done because the execution time for a small size dataset is not significant. Besides, this factor also allows the grid test to be performed on the tenfold cross-validation scheme.

Table 6 summarizes the tuned parameters, their tuning range, and the number of steps in which the parameters are varied. In the case of MLPs, three parameters are tuned: momentum, learning rate, and number of neurons in the hidden layer; for RBFs, regression trees, and RF, only one parameter was tuned: ridge, the minimum number of instances per leaf, or the number of iterations, respectively. Moreover, RF techniques are easily tuned and generate visual information for direct use by the process engineer, such as the linear relationships between process parameters and roughness, and thresholds for avoiding rapid tool wear. Therefore, the number of different configurations that were tested amounted to 336 for MLPs, 11 for RBFs, four for

Table 6 Parameters of each regressor that are tuned, variation range, and steps

Model	Parameter	Range	Steps
MLPs	Momentum	0.1–0.7	7
MLPs	Learning rate	0.025–0.3	12
MLPs	Number of neurons in the hidden layer	1, 2, 3, 4	4
RBFs	Ridge	1–0.00001	11
M5P tree	Minimum number of instances	2, 4, 8, 16	4
Random forest	Number of iterations	50, 100, 200, 400, 800	5

Table 7 Precision test results of the machine-learning methods

	With N_{min} and N_{max}		Without N_{min} and N_{max}	
	RMSE (µm)	Time (s)	RMSE (µm)	Time (s)
ZeroR	0.8883	0	0.8883	0
Linear regression	0.5308	0	0.5263	0
MLPs	0.3668	0.04	0.3762	0.03
RBFs	0.4342	0.01	0.358	0.03
M5P tree	0.3080	0.01	0.3077	0.01
Random forest	0.2758	0.08	0.2652	0.1

the regression trees, and five for RF. Because MLPs have three parameters to tune while the other models have only one, longer tuning times and an expert to perform the tuning process will be needed for MLPs compared with the other techniques.

Results and discussion

Table 7 shows the results of both RMSE and computational time for the best configuration of the tested machine-learning models and the two baselines methods for two datasets. One is presented in "Modeling" section and the other is a simplified version of this dataset without the minimum and maximum drive power. The best MLP configuration for the case of the complete dataset (columns under the heading "With N_{min} and N_{max}" in Table 7) is two hidden layers, a learning rate of 0.05, and a momentum of 0.3, while the most accurate models for RBFs, regression trees, and random forest (RF) are achieved for a ridge of 0.1, a minimum of eight instances, and 400 iterations, respectively. For the case of the incomplete dataset (columns under the heading "Without N_{min} and N_{max}" in Table 7), the best MLP configuration is one hidden layer (as expected due to the reduction in the number of attributes), a learning rate of 0.05, and a momentum of 0.2 while there are no changes in the parameters for the rest of the machine-learning methods except for the ridge value for the RBFs, which is 0.0001. The computational time has been calculated using an Intel Core i5 2300 2.8-GHz processor. The RMSE values listed in Table 7 show that all the machine-learning models are statistically more precise than ZeroR and linear regression, the two baseline methods con-

sidered; the statistically significant differences are calculated considering the corrected resampled t test (Nadeau and Bengio 2003) with a significance level of 5%. RBFs are more accurate than MLPs if N_{min} and N_{max} are not considered, while the opposite happens if both inputs are considered. Regression trees are at least 15% more precise than the best MLP or RBF configurations. RF has an even greater accuracy of between 33% for the complete dataset and 42% for the uncompleted dataset, with a statistically significant difference in both cases.

Considering the standard deviation of the roughness in the dataset, 0.88 µm, the RMSE can be considered a little high (30% for the best model), although still acceptable from the industrial point of view compared with similar works (Bustillo et al. 2011a, b; Maudes et al. 2017; Teixidor et al. 2015). Figure 4 shows the dataset prediction error for each instance (cross size) for the RF model to analyze this fact. It can be seen that the higher errors are all related to instances with high roughness values; this result can be expected due to the small account of instances with high roughness values: the machine-learning models are programmed to minimize the total error, so a limited number of instances with high roughness values will be disregarded by the model that will produce higher prediction errors for those instances.

With regard to the computational times listed in Table 7, the quickest machine-learning model was the simplest: the regression tree. Ensembles and ANNs required longer computational times. Although RF required twice as long as the MLPs, it has to be remembered that the tuning process of an MLP is more complex, with 336 configurations versus five configurations of RF, making the total time required to obtain
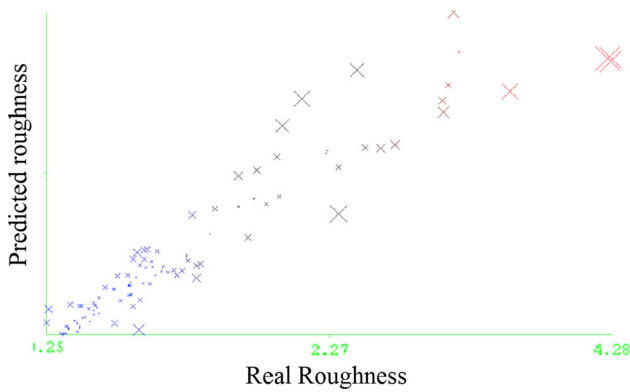
Fig. 4 Predicted roughness versus real roughness for the RF model

an accurate model significantly worse for the MLPs than the ensembles.

Besides this analysis, although regression trees provide a model that is 11.5% less accurate than RF ensembles, they provide visual and useful information on the cutting process to the process engineer. Figure 5 shows the M5P tree built with the whole of the completed dataset. All the linear models (LMs) built at the leaves of the tree have the same structure as described in Eq. 2, where C1 to C3 are different constants.

$$Ra = C_1 * N_{mean} + C_2 * t + C_3 \tag{2}$$

Besides, all the nodes evaluate the same input: the processing time, t. Therefore some immediate conclusions can be extracted from this tree: first, the relationship between drive power and roughness can be considered linear within small ranges of processing time. Second, V_B, N_{min}, and N_{max} are

not required by the model to predict Ra; this result does not mean that those variables play no role in the surface roughness of the face milling process. It does mean, however, that the regression tree is able to extract the information included in these inputs from the processing time and the average drive power. In this way, if the processing time with one cutting tool under fixed cutting conditions is known and if the RMSE of the regression tree is sufficient, the process engineer can predict surface roughness without stopping the machine to measure the flank wear. The third conclusion is related to the LMs: C_1 takes the same value (0.1933) for all LMs except LM6 where its value is more than twice as high (0.4279); therefore the process engineer knows that if the processing time is over 43 min, the average drive power should be carefully monitored because of its strong influence on roughness.

Finally, if the RF models are used to evaluate the different cutting conditions, further conclusions of industrial interest can be generated. First a 3D plot was built (Fig. 6), showing the average drive power, N_{mean}, on the X-axis, the processing time, t, on the Y-axis, and the predicted roughness, Ra, on the Z-axis. This 3D plot is built by erasing V_B, N_{min}, and N_{max} from the dataset and by retraining the model (with an RMSE in the same range, 0.2457 µm, as the completed dataset). As the accuracy of the model is within the same range, RF ensembles are also able to extract the information included in V_B, N_{min}, and N_{max} from the other inputs. The average drive power and the processing time are evaluated in the same experimental ranges as shown in Table 5, because no machine-learning model can securely predict that any process outside of its training range will achieve a reasonably accurate prediction.

Fig. 5 Regression tree for the prediction of surface roughness
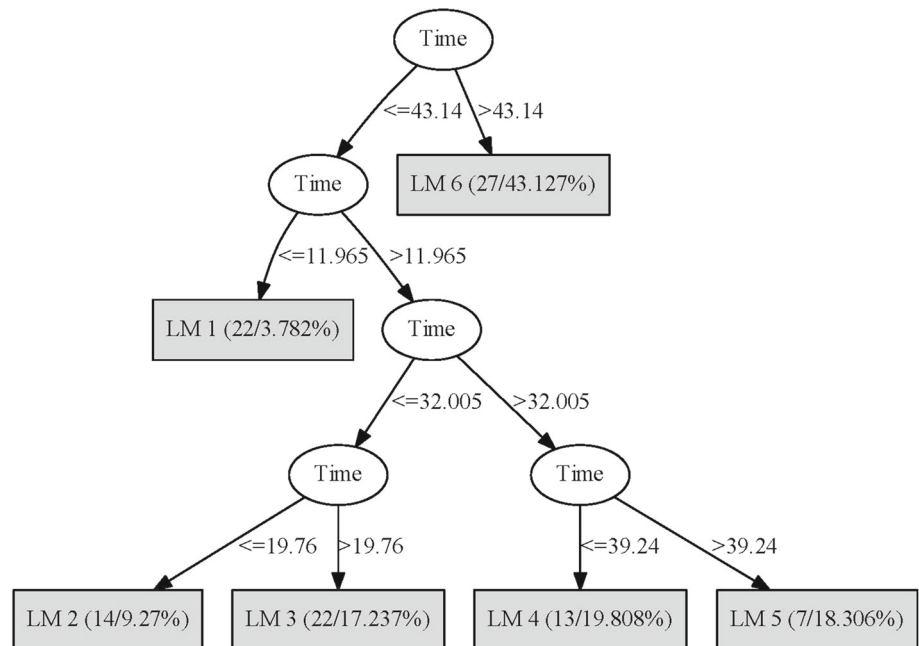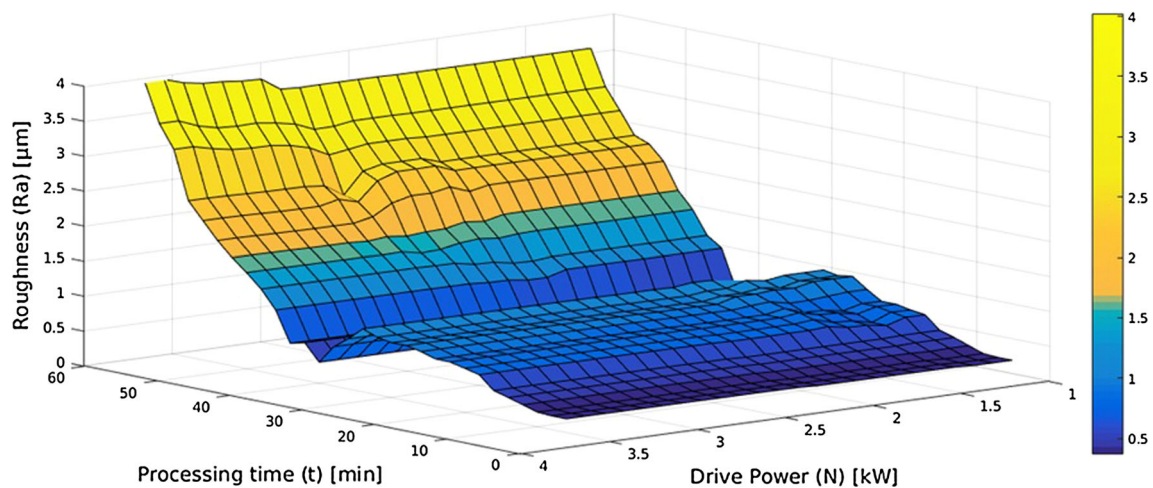
Fig. 6 Roughness predicted by RF ensemble as a function of time and drive power
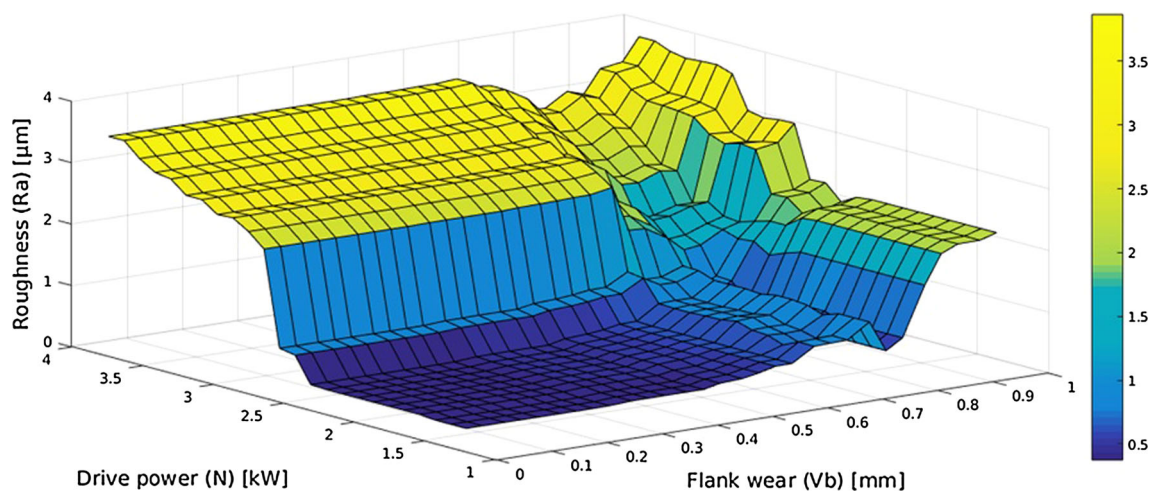
Fig. 7 Roughness predicted by the RF ensemble as a function of flank wear and drive power

Figure 6 shows that the processing time has the highest influence on workpiece roughness, especially if the processing time exceeds 40 min. The influence of drive power is smaller, but depending on the processing time, there is a border that will make the roughness increase by around 0.15–0.4 μm, a considerable step depending on the required surface quality of the workpiece.

A second useful 3D graph can be generated considering the average drive power on the X-axis, the flank wear on the Y-axis, and the predicted roughness on the Z-axis (Fig. 7). When generating this 3D plot, t, N_{min}, and N_{max} are erased from the dataset and the model is retrained. In this case, the accuracy of the model was lower: 0.4139 μm (RMSE), an expected value considering that the processing time was one of the main information sources for the machine-learning model. Despite the decreased accuracy of the model, the conclusions that the 3D plot can provide can be very useful for the process engineer. Again, the average drive power and

the flank wear were evaluated in the same range as for the experiments (Table 5).

This 3D plot shows the case where the processing time is unknown or cannot be evaluated. For example, the very common situation in production centers where different cutting conditions are applied during the machining with the same cutting tool. In this case, the process engineer can stop the machine before the cutting tool begins a new operation, evaluate the flank wear, and monitor the power drive to obtain a rough prediction of the surface roughness. So, if the drive power and flank wear are low, the roughness may remain under low values, but if either of these two parameters exceeds a certain threshold, namely 2.2 kW for drive power and 0.43 mm for flank wear, surface roughness will increase dramatically in a very short period of time (depending on the respective speed of the increase in both drive power and flank wear). Therefore the process engineer can set security thresholds in the cutting process that will not be exceeded

by monitoring the drive power, stopping the machine, and measuring the flank wear.

In summary, the proposed AI models can be used in two ways by the process engineer: first, regression trees can be built to provide immediate visual information such as the linear relationship between drive power and roughness in small ranges of processing time or up to a limit of 43 min of processing time; exceeding this limit implies that the drive power should be carefully monitored, because its influence on the roughness becomes very strong. Second, if high model accuracy is required, RF models can be used to build 3D charts considering two inputs. When drive power and processing time are considered, the same limit of around 40 min, defined for the regression trees, was identified. If the values of both drive power and flank wear are low, the roughness may remain at low values. However, if either of these two parameter values exceeds a certain threshold, namely 2.2 kW for drive power and 0.43 mm for flank wear, roughness levels will increase dramatically over a very short period of time (depending on the respective increase of either the drive power speed or the flank wear).

Conclusions

The productivity of milling operations is mainly limited by tool wear, which produces degradation in the surface quality of the machined workpiece. In real workshops, in some cases the cutting tool always works under the same cutting conditions and the total processing time of each tool is well-known, but in other cases the cutting tool is used under different cutting conditions and the only way to evaluate its present state is to measure the flank wear, V_B, directly. In both cases, the process engineer can consider the drive power as an indirect indicator of the state of the tool, because all CNC machine-tools provide this parameter in real time.

A practical approach based on AI models has been proposed in this research to provide useful information to the process engineer in real time on the expected workpiece surface roughness in both cases, based only on these parameters, and avoiding the inclusion of additional new sensors on the machine

First, the experimental data were collected to provide a broad number of wear conditions and processing times, while acquiring data on the power drive, for a fixed machining process: the face milling of carbon-quality structural steel 45 (similar to ANSI 1045). Second, different machine learning techniques were tested using a 10×10 cross-validation, from well-known ANNs such as MLPs and RBFs to more recent techniques such as ensembles of regression trees. The tuning processes of the main parameters of each technique were based on a grid search. The results have shown that:

- Regression trees are at least 15% more precise than the best MLP or RBF configuration,
- RF ensembles have even greater accuracy: between 33 and 44% depending on the inclusion in the dataset of the minimum and maximum drive power or only the average value for each cutting experiment.
- The tuning stage of ensembles and regression trees is shorter compared with ANNs, because they have only one parameter to be tuned, compared with the three parameters that have to be tuned in the case of MLPs.

Furthermore, the practical use of the most accurate models is presented following two possible strategies:

- Regression trees can provide immediate visual information, such as the linear relationship between inputs and outputs or thresholds in the behavior of the variables that should be carefully monitored, because exceeding these thresholds implies rapid degradation of the workpiece roughness if a tool change is not performed. Specially, V_B, N_{min}, and N_{max} are not required for roughness prediction by this model, because the regression tree is able to extract the information included in these inputs from the processing time and the average drive power.
- RF models can be used to build 3D charts considering different inputs: drive power and processing time can be used to identify limits in the behaviors of these inputs, as with regression trees, while drive power and flank wear can be used to identify ranges in the variables where the roughness may remain at low values. Both 3D charts are suitable for one of the previously presented real workshop situations: fixed cutting tool conditions and a well-known processing time, or changing cutting conditions and periodic measurements of flank wear.

Further research will focus on the application of specially designed machine-learning techniques for unbalanced datasets to this industrial problem, taking into account the unbalanced nature of the datasets presented in real workshops, due to the small number of instances that describe low-quality outputs, such as high roughness values, in this study case. This kind of technique may be expected to improve the overall accuracy of the models, especially in conditions outside of the main working range. Besides, broader conclusions could be drawn on the capability of RF to achieve highly accurate prediction models of tool wear, through the development of a more extensive dataset including information on wear evolution measured under different cutting conditions.

References

Adamczak, S., Miko, E., & Čuš, F. (2009). A model of surface roughness constitution in the metal cutting process applying tools with defined stereometry. *Strojniski Vestnik/Journal of Mechanical Engineering*, *55*(1), 45–54.

Arizmendi, M., Campa, F. J., Fernández, J., López de Lacalle, L. N., Gil, A., Bilbao, E., et al. (2009). Model for surface topography prediction in peripheral milling considering tool vibration. *CIRP Annals—Manufacturing Technology*, *58*(1), 93–96. https://doi.org/10.1016/j.cirp.2009.03.084.

Artetxe, E., Olvera, D., López de Lacalle, L. N., Campa, F. J., Olvera, Dn, & Lamikiz, A. (2017). Solid subtraction model for the surface topography prediction in flank milling of thin-walled integral blade rotors (IBRs). *International Journal of Advanced Manufacturing Technology*, *90*(1–4), 741–752. https://doi.org/10.1007/s00170-016-9435-1.

Baek, D. K., Ko, T. J., & Kim, H. S. (1997). A dynamic surface roughness model for face milling. *Precision Engineering*, *20*(3), 171–178.

Bajić, D., Celent, L., & Jozić, S. (2012). Modeling of the influence of cutting parameters on the surface roughness, tool wear and cutting force in face milling in off-line process control. *Strojniski Vestnik/Journal of Mechanical Engineering*, *58*(11), 673–682. https://doi.org/10.5545/sv-jme.2012.456.

Benardos, P. G., & Vosniakos, G. C. (2002). Prediction of surface roughness in CNC face milling using neural networks and Taguchi's design of experiments. *Robotics and Computer-Integrated Manufacturing*, *18*(5–6), 343–354. https://doi.org/10.1016/S0736-5845(02)00005-4.

Benardos, P. G., & Vosniakos, G.-C. (2003). Predicting surface roughness in machining: A review. *International Journal of Machine Tools and Manufacture*, *43*(8), 833–844. https://doi.org/10.1016/S0890-6955(03)00059-2.

Bharathi Raja, S., & Baskar, N. (2012). Application of particle swarm optimization technique for achieving desired milled surface roughness in minimum machining time. *Expert Systems with Applications*, *39*(5), 5982–5989. https://doi.org/10.1016/j.eswa.2011.11.110.

Bhattacharyya, P., & Sengupta, D. (2009). Estimation of tool wear based on adaptive sensor fusion of force and power in face milling. *International Journal of Production Research*, *47*(3), 817–833. https://doi.org/10.1080/00207540701403376.

Bhattacharyya, P., Sengupta, D., Mukhopadhyay, S., & Chattopadhyay, A. B. (2008). On-line tool condition monitoring in face milling using current and power signals. *International Journal of Production Research*, *46*(4), 1187–1201. https://doi.org/10.1080/00207540600940288.

Bishop, C. M. (1995). *Neural networks for pattern recognition*. New York, NY: Oxford University Press, Inc.

Breiman, L. (1996). Bagging predictors. *Machine Learning*, *24*(2), 123–140.

Breiman, L. (2001). Random forests. *Machine Learning*, *45*(1), 5–32. https://doi.org/10.1023/A:1010933404324.

Bruni, C., d'Apolito, L., Forcellese, A., Gabrielli, F., & Simoncini, M. (2008). Surface roughness modelling in finish face milling under MQL and dry cutting conditions. *International Journal of Material Forming*, *1*(SUPPL. 1), 503–506. https://doi.org/10.1007/s12289-008-0151-8.

Bustillo, A., Díez-Pastor, J.-F., Quintana, G., & García-Osorio, C. (2011a). Avoiding neural network fine tuning by using ensemble learning: Application to ball-end milling operations. *International Journal of Advanced Manufacturing Technology*, *57*(5–8), 521–532. https://doi.org/10.1007/s00170-011-3300-z.

Bustillo, A., Grzenda, M., & Macukow, B. (2016). Interpreting tree-based prediction models and their data in machining processes. *Integrated Computer-Aided Engineering*, *23*(4), 349–367. https://doi.org/10.3233/ICA-160513.

Bustillo, A., Ukar, E., Rodriguez, J. J., & Lamikiz, A. (2011b). Modelling of process parameters in laser polishing of steel components using ensembles of regression trees. *International Journal of Computer Integrated Manufacturing*, *24*(8), 735–747. https://doi.org/10.1080/0951192X.2011.574155.

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.

Compeán, F. I., Olvera, D., Campa, F. J., López De Lacalle, L. N., Elías-Zúñiga, A., & Rodríguez, C. A. (2012). Characterization and stability analysis of a multivariable milling tool by the enhanced multistage homotopy perturbation method. *International Journal of Machine Tools and Manufacture*, *57*, 27–33. https://doi.org/10.1016/j.ijmachtools.2012.01.010.

da Silva, R. H. L., da Silva, M. B., & Hassui, A. (2016). A probabilistic neural network applied in monitoring tool wear in the end milling operation via acoustic emission and cutting power signals. *Machining Science and Technology*, *20*(3), 386–405. https://doi.org/10.1080/10910344.2016.1191026.

De Escalona, P. M., & Maropoulos, P. G. (2010). Influence of cutting parameters and tool wear on martensitic stainless steel surface integrity after a face milling process. *American Society of Mechanical Engineers Pressure Vessels and Piping Division (Publication)*, *PVP 6*(PART B), 1707–1716. https://doi.org/10.1115/PVP2009-77024.

De Souza, A. M, Jr., Sales, W. F., Ezugwu, E. O., Bonney, J., & Machado, A. R. (2003). Burr formation in face milling of cast iron with different milling cutter systems. *Proceedings of the Institution of Mechanical Engineers Part B Journal of Engineering Manufacture*, *217*(11), 1589–1596. https://doi.org/10.1243/095440503771909962.

De Souza, A. M, Jr., Sales, W. F., Santos, S. C., & MacHado, A. R. (2005). Performance of single Si_3N_4 and mixed Si_3N_4+PCBN wiper cutting tools applied to high speed face milling of cast iron. *International Journal of Machine Tools and Manufacture*, *45*(3), 335–344. https://doi.org/10.1016/j.ijmachtools.2004.08.006.

Diniz, A. E., & Filho, J. C. (1999). Influence of the relative positions of tool and workpiece on tool life, tool wear and surface finish in the face milling process. *Wear*, *232*(1), 67–75. https://doi.org/10.1016/S0043-1648(99)00159-3.

Dugin, A., & Popov, A. (2013). Increasing the accuracy of the effect of processing materials and cutting tool wear on the ploughing force values. *Manufacturing Technology*, *13*(2), 169–173.

Elhami, S., Razfar, M. R., Farahnakian, M., & Rasti, A. (2013). Application of GONNS to predict constrained optimum surface roughness in face milling of high-silicon austenitic stainless steel. *International Journal of Advanced Manufacturing Technology*, *66*(5–8), 975–986. https://doi.org/10.1007/s00170-012-4382-y.

El-Sonbaty, I. A., Khashaba, U. A., Selmy, A. I., & Ali, A. I. (2008). Prediction of surface roughness profiles for milled surfaces using an artificial neural network and fractal geometry approach. *Journal of Materials Processing Technology*, *200*(1–3), 271–278. https://doi.org/10.1016/j.jmatprotec.2007.09.006.

Felho, C., Karpuschewski, B., & Kundrák, J. (2015). Surface roughness modelling in face milling. *Procedia CIRP*, *31*, 136–141. https://doi.org/10.1016/j.procir.2015.03.075.

Fernández-Valdivielso, A., López De Lacalle, L. N., Urbikain, G., & Rodriguez, A. (2016). Detecting the key geometrical features and grades of carbide inserts for the turning of nickel-based alloys concerning surface integrity. *Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science*, *230*(20), 3725–3742. https://doi.org/10.1177/0954406215616145.

Ferreiro, S., & Sierra, B. (2012). Comparison of machine learning algorithms for optimization and improvement of process quality in conventional metallic materials. *International Journal of Advanced Manufacturing Technology*, *60*(1–4), 237–249. https://doi.org/10.1007/s00170-011-3578-x.

Freiburg, D., Odendahl, S., Siebrecht, T., Steiner, M., Wagner, T., & Zabel, A. (2014). Simulation based process optimization for the milling of light weight components. *Procedia CIRP*, *18*, 132–137. https://doi.org/10.1016/j.procir.2014.06.120.

Gong, F., Zhao, J., Jiang, Y., Tao, H., Li, Z., & Zang, J. (2017). Fatigue failure of coated carbide tool and its influence on cutting performance in face milling SKD11 hardened steel. *International Journal of Refractory Metals and Hard Materials*, *64*, 27–34. https://doi.org/10.1016/j.ijrmhm.2017.01.001.

Grigoriev, S. N., Volosova, M. A., Gurin, V. L., & Seleznev, A. E. (2015). Wear of replaceable indexable inserts made of mixed cutting ceramics CC650 as a function of force parameters of steel ShKh15 face milling. *Journal of Friction and Wear*, *36*(6), 521–527. https://doi.org/10.3103/S1068366615060045.

Grzenda, M., & Bustillo, A. (2013). The evolutionary development of roughness prediction models. *Applied Soft Computing Journal*, *13*(5), 2913–2922. https://doi.org/10.1016/j.asoc.2012.03.070.

Grzenda, M., Bustillo, A., Quintana, G., & Ciurana, J. (2012). Improvement of surface roughness models for face milling operations through dimensionality reduction. *Integrated Computer-Aided Engineering*, *19*(2), 179–197. https://doi.org/10.3233/ICA-2012-0398.

Guzeev, V. I., & Pimenov, D. Y. (2011). Cutting force in face milling with tool wear. *Russian Engineering Research*, *31*(10), 989–993. https://doi.org/10.3103/S1068798X11090139.

Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P., & Witten, I. H. (2009). The WEKA data mining software: An update. *ACM SIGKDD explorations newsletter*, *11*(1), 10–18. https://doi.org/10.1145/1656274.1656278.

Houchuan, Y., Zhitong, C., & ZiTong, Z. (2015). Influence of cutting speed and tool wear on the surface integrity of the titanium alloy Ti-1023 during milling. *International Journal of Advanced Manufacturing Technology*, *78*(5–8), 1113–1126. https://doi.org/10.1007/s00170-014-6593-x.

Jersák, J., & Simon, S. (2017). Influence of cooling lubricants on the surface roughness and energy efficiency of the cutting machine tools. *International Journal of Applied Mechanics and Engineering*, *22*(3), 779–787. https://doi.org/10.1515/ijame-2017-0050.

Kohavi, R. (1995). A study of cross-validation and bootstrap for accuracy estimation and model selection. *Proceedings of the Fourteenth International Joint Conference on Artificial Intelligence (San Mateo, CA: Morgan Kaufmann)*, *2*(12), 1137–1143.

Kovac, P., Rodic, D., Pucovsky, V., Savkovic, B., & Gostimirovic, M. (2013). Application of fuzzy logic and regression analysis for modeling surface roughness in face milling. *Journal of Intelligent Manufacturing*, *24*(4), 755–762. https://doi.org/10.1007/s10845-012-0623-z.

Kuncheva, L. I. (2014). *Combining pattern classifiers: Methods and algorithms* (6th ed., pp. 1–357). ISBN: 978-111891456-4;978-111831523-1. https://doi.org/10.1002/9781118914564.

Lela, B., Bajić, D., & Jozić, S. (2009). Regression analysis, support vector machines, and Bayesian neural network approaches to modeling surface roughness in face milling. *International Journal of Advanced Manufacturing Technology*, *42*(11–12), 1082–1088. https://doi.org/10.1007/s00170-008-1678-z.

Leonard, J. A., & Kramer, M. A. (1991). Radial basis function networks for classifying process faults. *IEEE Control Systems*, *11*(3), 31–38. https://doi.org/10.1109/37.75576.

Liu, G., Zou, B., Huang, C., Wang, X., Wang, J., & Liu, Z. (2016). Tool damage and its effect on the machined surface roughness in high-speed face milling the 17-4PH stainless steel. *International Journal of Advanced Manufacturing Technology*, *83*(1–4), 257–264. https://doi.org/10.1007/s00170-015-7564-6.

López De Lacalle, L. N., Lamikiz, A., Sánchez, J. A., & Fernández De Bustos, I. (2006). Recording of real cutting forces along the milling of complex parts. *Mechatronics*, *16*(1), 21–32. https://doi.org/10.1016/j.mechatronics.2005.09.001.

Machado, Á. R., & Diniz, A. E. (2017). Tool wear analysis in the machining of hardened steels. *International Journal of Advanced Manufacturing Technology*, *92*(9–12), 4095–4109. https://doi.org/10.1007/s00170-017-0455-2.

Maudes, J., Bustillo, A., Guerra, A. J., & Ciurana, J. (2017). Random forest ensemble prediction of stent dimensions in microfabrication processes. *International Journal of Advanced Manufacturing Technology*, *91*(1–4), 879–893. https://doi.org/10.1007/s00170-016-9695-9.

Mendes-Moreira, J., Soares, C., Jorge, A. M., & De Sousa, J. F. (2012). Ensemble approaches for regression: A survey. *ACM Computing Surveys*, *45*(1), 10. https://doi.org/10.1145/2379776.2379786.

Miko, E., & Nowakowski, A. (2012). Analysis and verification of surface roughness constitution model after machining process. *Procedia Engineering*, *39*, 395–404. https://doi.org/10.1016/j.proeng.2012.07.043.

Mikołajczyk, T., Nowicki, K., Bustillo, A., & Pimenov, D. Y. (2018). Predicting tool life in turning operations using neural networks and image processing. *Mechanical Systems and Signal Processing*, *104*, 503–513. https://doi.org/10.1016/j.ymssp.2017.11.022.

Mikołajczyk, T., Nowicki, K., Kłodowski, A., & Pimenov, D. Y. (2017). Neural network approach for automatic image analysis of cutting edge wear. *Mechanical Systems and Signal Processing*, *88*, 100–110. https://doi.org/10.1016/j.ymssp.2016.11.026.

Moghaddam, M. A., & Kolahan, F. (2016). Application of orthogonal array technique and particle swarm optimization approach in surface roughness modification when face milling AISI1045 steel parts. *Journal of Industrial Engineering International*, *12*(2), 199–209. https://doi.org/10.1007/s40092-015-0137-3.

Muñoz-Escalona, P., & Maropoulos, P. G. (2010). Artificial neural networks for surface roughness prediction when face milling Al 7075–T7351. *Journal of Materials Engineering and Performance*, *19*(2), 185–193. https://doi.org/10.1007/s11665-009-9452-4.

Muñoz-Escalona, P., & Maropoulos, P. G. (2015). A geometrical model for surface roughness prediction when face milling Al 7075–T7351 with square insert tools. *Journal of Manufacturing Systems*, *36*(309), 216–223. https://doi.org/10.1016/j.jmsy.2014.06.011.

Nadeau, C., & Bengio, Y. (2003). Inference for the generalization error. *Machine Learning*, *52*(3), 239–281. https://doi.org/10.1023/A:1024068626366.

Niaki, F. A., & Mears, L. (2017). A probabilistic-based study on fused direct and indirect methods for tracking tool flank wear of Rene-108, nickel-based alloy. *Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture*. https://doi.org/10.1177/0954405416683432.

Niaki, F. A., Ulutan, D., & Mears, L. (2015). Stochastic tool wear assessment in milling difficult to machine alloys. *International Journal of Mechatronics and Manufacturing Systems*, *8*(3–4), 134–159. https://doi.org/10.1504/IJMMS.2015.073090.

Pimenov, D. Y. (2013a). Geometric model of height of microroughness on machined surface taking into account wear of face mill teeth. *Journal of Friction and Wear*, *34*(4), 290–293. https://doi.org/10.3103/S1068366613040089.

Pimenov, D. Y. (2013b). The effect of the rate flank wear teeth face mills on the processing. *Journal of Friction and Wear*, *34*(2), 156–159. https://doi.org/10.3103/S1068366613020104.

Pimenov, D. Y. (2014). Experimental research of face mill wear effect to flat surface roughness. *Journal of Friction and Wear*, *35*(3), 250–254. https://doi.org/10.3103/S1068366614030118.

Pimenov, D. Y. (2015). Mathematical modeling of power spent in face milling taking into consideration tool wear. *Journal of Friction and Wear*, *36*(1), 45–48. https://doi.org/10.3103/S1068366615010110.

Pimenov, D. Y., & Guzeev, V. I. (2017). Mathematical model of plowing forces to account for flank wear using FME modeling for orthogonal cutting scheme. *International Journal of Advanced Manufacturing Technology*, *89*(9–12), 3149–3159. https://doi.org/10.1007/s00170-016-9216-x.

Popov, A., & Schindelarz, R. (2017). Effect of hydraulic oil entering the cutting fluid on the tool life and roughness in milling of stainless steel. *Manufacturing Technology*, *17*(3), 364–369.

Prasad, B. S., Sarcar, M. M. M., & Ben, B. S. (2011). Real-time tool condition monitoring of face milling using acousto-optic emission—An experimental approach. *International Journal of Computer Applications in Technology*, *41*(3–4), 317–325. https://doi.org/10.1504/IJCAT.2011.042709.

Quinlan, J. R. (1992). Learning with continuous classes. In *Proceedings of the 5th Australian Joint Conference on Artificial Intelligence, Singapore* (Vol. 92, pp. 343–348).

Razfar, M. R., Farshbaf Zinati, R., & Haghshenas, M. (2011). Optimum surface roughness prediction in face milling by using neural network and harmony search algorithm. *International Journal of Advanced Manufacturing Technology* , *52*(5–8), 487–495. https://doi.org/10.1007/s00170-010-2757-5.

Rivero, A., López de Lacalle, L. N., & Penalva, M. L. (2008). Tool wear detection in dry high-speed milling based upon the analysis of machine internal signals. *Mechatronics*, *18*(10), 627–633. https://doi.org/10.1016/j.mechatronics.2008.06.008.

Rodríguez, J., Quintana, G., Bustillo, A., & Ciurana, J. (2017). A decision-making tool based on decision trees for roughness prediction in face milling. *International Journal of Computer Integrated Manufacturing*, *30*(9), 943–957. https://doi.org/10.1080/0951192X.2016.1247991.

Rosales, A., Vizán, A., Diez, E., & Alanís, A. (2010). Prediction of surface roughness by registering cutting forces in the face milling process. *European Journal of Scientific Research*, *41*(2), 228–237.

Saglam, H., & Unuvar, A. (2003). Tool condition monitoring in milling based on cutting forces by a neural network. *International Journal of Production Research*, *41*(7), 1519–1532. https://doi.org/10.1080/0020754031000073017.

Saric, T., Simunovic, G., & Simunovic, K. (2013). Use of neural networks in prediction and simulation of steel surface roughness.

International Journal of Simulation Modelling, *12*(4), 225–236. https://doi.org/10.2507/IJSIMM12(4)2.241.

Selaimia, A.-A., Yallese, M. A., Bensouilah, H., Meddour, I., Khattabi, R., & Mabrouki, T. (2017). Modeling and optimization in dry face milling of X2CrNi18-9 austenitic stainless steel using RMS and desirability approach. *Measurement: Journal of the International Measurement Confederation*, *107*, 53–67. https://doi.org/10.1016/j.measurement.2017.05.012.

Shao, H., Wang, H. L., & Zhao, X. M. (2004). A cutting power model for tool wear monitoring in milling. *International Journal of Machine Tools and Manufacture*, *44*(14), 1503–1509. https://doi.org/10.1016/j.ijmachtools.2004.05.003.

Shi, K., Ren, J., Zhang, D., Zhai, Z., & Huang, X. (2016). Tool wear behaviors and its effect on machinability in dry high-speed milling of magnesium alloy. *International Journal of Advanced Manufacturing Technology*, *90*(9–12), 3265–3273. https://doi.org/10.1007/s00170-016-9645-6.

Simunovic, G., Simunovic, K., & Saric, T. (2013). Modelling and simulation of surface roughness in face milling. *International Journal of Simulation Modelling*, *12*(3), 141–153. https://doi.org/10.2507/IJSIMM12(3)1.219.

Simunovic, G., Svalina, I., Simunovic, K., Saric, T., Havrlisan, S., & Vukelic, D. (2016). Surface roughness assessing based on digital image features. *Advances in Production Engineering and Management*, *11*(2), 93–104. https://doi.org/10.14743/apem2016.2.212.

Srinivasa Pai, P., Nagabhushana, T. N., & Ramakrishna Rao, P. K. (2002). Flank wear estimation in face milling based on radial basis function neural networks. *International Journal of Advanced Manufacturing Technology*, *20*(4), 241–247. https://doi.org/10.1007/s001700200148.

Svalina, I., Šimunović, G., Šarić, T., & Lujić, R. (2017). Evolutionary neuro-fuzzy system for surface roughness evaluation. *Applied Soft Computing Journal*, *52*, 593–604. https://doi.org/10.1016/j.asoc.2016.10.010.

Teixidor, D., Grzenda, M., Bustillo, A., & Ciurana, J. (2015). Modeling pulsed laser micromachining of micro geometries using machine-learning techniques. *Journal of Intelligent Manufacturing*, *26*(4), 801–814. https://doi.org/10.1007/s10845-013-0835-x.

Urbikain, G., Alvarez, A., López De Lacalle, L. N., Arsuaga, M., Alonso, M. A., & Veiga, F. (2017). A reliable turning process by the early use of a deep simulation model at several manufacturing stages. *Machines*, *5*(2), 15. https://doi.org/10.3390/machines5020015.

Werda, S., Duchosal, A., Le Quilliec, G., Morandeau, A., & Leroy, R. (2017). Minimum quantity lubrication advantages when applied to insert flank face in milling. *International Journal of Advanced Manufacturing Technology*, *92*(5–8), 2391–2399. https://doi.org/10.1007/s00170-017-0317-y.

Zhenyu, S., Luning, L., & Zhanqiang, L. (2015). Influence of dynamic effects on surface roughness for face milling process. *International Journal of Advanced Manufacturing Technology*, *80*(9–12), 1823–1831. https://doi.org/10.1007/s00170-015-7127-x.