ORIGINAL ARTICLE

Roundness prediction in centreless grinding using physics-enhanced machine learning techniques

Hossein Safarzadeh [1,2] · Marco Leonesio [3] · Giacomo Bianchi [3] · Michele Monno [1]

Abstract
This work proposes a model for suggesting optimal process configuration in plunge centreless grinding operations. Seven different approaches were implemented and compared: first principles model, neural network model with one hidden layer, support vector regression model with polynomial kernel function, Gaussian process regression model and hybrid versions of those three models. The first approach is based on an enhancement of the well-known numerical process simulation of geometrical instability. The model takes into account raw workpiece profile and possible wheel-workpiece loss of contact, which introduces an inherent limitation on the resulting profile waviness. Physical models, because of epistemic errors due to neglected or oversimplified functional relationships, can be too approximated for being considered in industrial applications. Moreover, in deterministic models, uncertainties affecting the various parameters are not explicitly considered. Complexity in centreless grinding models arises from phenomena like contact length dependency on local compliance, contact force and grinding wheel roughness, unpredicted material properties of the grinding wheel and workpiece, precision of the manual setup done by the operator, wheel wear and nature of wheel wear. In order to improve the overall model prediction accuracy and allow automated continuous learning, several machine learning techniques have been investigated: a Bayesian regularized neural network, an SVR model and a GPR model. To exploit the a priori knowledge embedded in physical models, hybrid models are proposed, where neural network, SVR and GPR models are fed by the nominal process parameters enriched with the roundness predicted by the first principle model. Those hybrid models result in an improved prediction capability.

1 Introduction

1.1 Centreless grinding

As Dhavlikar et al. [1] describe centreless grinding is a common manufacturing grinding process for round workpieces, thanks to its unique workpiece (WP) holding system. The WP is sustained along three contact lines, with the grinding wheel, the regulating wheel and the supporting blade (Fig. 1). This method removes the need to clamp the workpiece and create centring holes on the workpiece. WP loading and unloading is easier, which results in reduced cycle time and higher productivity. Nevertheless, Zhou et al. [2] mention, due to this setup, centreless grinding is exposed to roundness errors generated by two types of instabilities: dynamic regenerative chatter and geometric lobing.

Dynamic chatter, due to the interaction between the cutting process and the main resonances of the machine structure, is a very usual phenomenon in machining processes. Gallego [3] defines the geometric lobing as the product of the peculiar geometric setup of the WP, i.e. blade angle and WP height. As Klocke [4] shows in his study, it is one of the main constraints to WP roundness accuracy: WP centre can oscillate, provoking an irregular material removal which in turn increases WP waviness.

✉ Hossein Safarzadeh
hossein.safarzadeh@polimi.it

1   Department of Mechanical Engineering, Politecnico di Milano, Via Giuseppe La Masa, 1, 20156 Milan, Italy

2   Mechanical Engineering Department, Monzesi srl, Via Dalmazia, 18, 20834 Nova Milanese, Italy

3   CNR STIIMA, Institute of Intelligent Industrial Technologies and Systems for Advanced Manufacturing National Research Council of Italy, Via Alfonso Corti 12, 20133 Milan, Italy

Fig. 1 Centreless grinding geometry
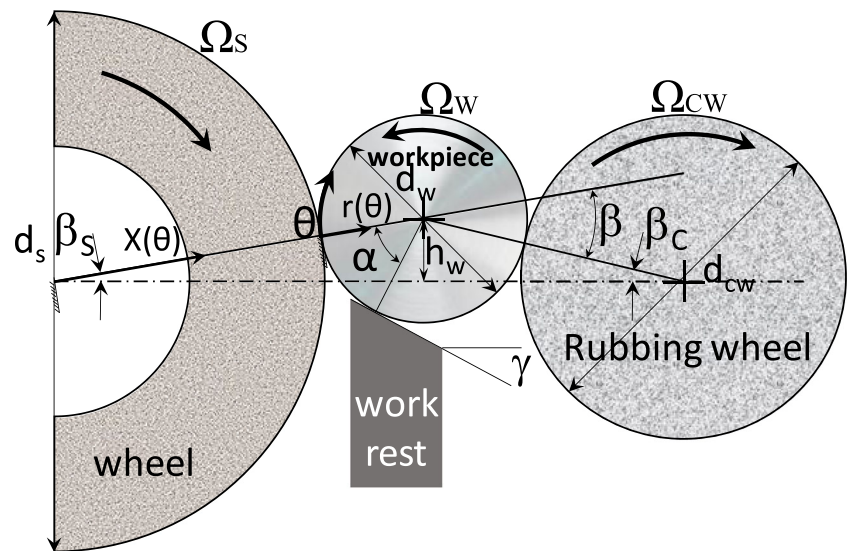
Dall [5] conducted the first significant research on the out-of-roundness WP problem, systematically relating the roundness error to the geometric configuration. Two main parameters were considered: the tangent angle and the top supporting blade angle. Then, Yonetsu [6] defined the interactions between pre- and post-grinding amplitudes of harmonics of the WP profile. And Rowe et al. [7] simulated the centreless grinding lobing problem on a digital computer taking into account all previous geometrical contemplations and presented the analytical model of the so-called *geometric rounding mechanism*. Later Marinescu et al. [8] used these results to explain the geometric roundness error regeneration process. After describing the basic geometrical connections of the process, it was possible to use different stability criteria, like Nyquist criterion, to show the theoretical instabilities produced by different configurations. Various researchers such as Bueno et al. [9] claimed that it is possible to create a stability map using the Nyquist criterion for each possible number of complete undulations that can be generated on the WP. Zhou et al. [2] presented the periodic characteristic roots distribution of the lobing loop and recommended a nominal stability diagram to suggest the range of the centre-height angle in order to lessen the lobing effect. Rowe et al. [10] introduced the geometric stability parameter derived from the Nyquist stability criterion, limited to integer lobes. Using these bases, Bianchi et al. [11] considered the nonlinearity due to the loss of contact under large waviness, investigating its effect on process stability. Also Lizarralde [12] has applied similar approaches to guide setup and optimization of centreless plunge grinding processes, in order to reduce setup time and avoid geometric instabilities as a function of WP height and blade angle, taking into account machine-WP dynamic interaction. These techniques lead to models that

quantitatively predict the evolution of profile error for each geometric configuration.

The main objective of the previous works was to identify the region of instability, to be avoided during process setup configuration. It has to be noted that, depending on WP initial profile and process nonlinearities, final waviness can be negligible despite process instability, resulting in an acceptable WP. Time domain simulations have been exploited in literature for considering nonlinear phenomena and trying to predict the final WP roundness for unstable processes. Nevertheless, the limits of simulative approaches are well explained by Marinescu et al. [8]: "It is seen that the general [simulated] shapes have a similar tendency to those predicted […], but it would appear that other errors are additionally present in the experimental results". As a matter of fact, limited quantitative comparisons are available in literature despite the significant research effort spent. As Zhou et al. [13] explain, different physical phenomena are involved in this phenomenon: while geometric variables roughly define the shape of the contact zone between WP and grinding wheel, measurements show that the actual contact length is much greater than the expected geometrical intersection. Furthermore Liu et al. [14] discuss that the main parameters influencing contact length are workpiece energy intensity, wear length and contact time of the abrasive grains, number of abrasive grains in contact, chip thickness and grain forces, wheel wear and nature of wheel wear, surface roughness and contact temperatures. In addition to the contact length, variability in material properties of the grinding wheel and workpiece, systematic precision/ error of the manual setup done by the operator as a parameter, should be tackled in overall calculations with a robust approach, in order to reach accurate results.

Considering the various combinations of process geometry and grinding wheels, a proper estimation of all required parameters is very time-consuming and barely robust. Many

researchers tried to develop algorithms to consider a reduced set of input data, such as workpiece diameter, wheel diameters and wheel properties, for suggesting an optimal grinding set-up. Hashimoto [15] developed a model, named Opt-Setup Master that can generate the optimum setup conditions to ensure safe operations, better roundness and chatter-free grinding. This model, referring to Fig. 1, finds the sets of blade angle (γ), centre-height angle (β) and workpiece rotational speed (Ω_w) satisfying all three stability criteria and then determines the optimum set by calculating the so-called PI (performance index) function, based on the process targets in terms of accuracy and productivity. Zakharov et al. [16] showed that the setup of centreless superfinishing machine tools entails the creation of geometric, kinematic and mechanical models of the shaping process and the use of formal optimization methods. Moreover Barrenetxea [17] developed an assistant tool for the setup and optimization of the centreless grinding process, optimizing productivity considering process data and machine characteristics (dynamics included).

All aforementioned approaches need a series of complex machine and process manual characterization in order to provide realistic output. Those characterizations are time-consuming and results uncertainty is usually not explicitly taken into account. Furthermore, traditional parametric identification does not allow to automatically alleviate epistemic errors by learning new functional relationships. These problems are amplified by the fact that the system is often inherently unstable and sensitive to initial states and parameters perturbation. A relevant known issue is concerned with nonlinear system identification, including structural dynamics, which is still an open problem in literature.

In addition to models based on first principles, as the one in the studies discussed above, an empirical approach is possible. We will focus on process characterization by artificial intelligence techniques in the following section.

1.2 Parameters selection based on AI

As Sjöberg et al. [18] suggest, a widespread approach to perform an input/output regression is based on artificial neural networks (NN). Many researchers have used those methodologies for prediction of chatter and surface roughness, both in grinding and other process domains. Rowe et al. [19] have suggested the application of AI technologies in grinding using modern computers and controllers as a way forward to produce higher quality components more efficiently. Junkar et al. [20] used machine learning technics for classification, through grinding signal detection, to assess performance classes, and discussed the possibility of upgrading this approach to a control algorithm by associating class assignments with appropriate control actions by a binary tree approach. Moreover Filipic et al. [21] used the same method to classify dielectric fluids in electrical discharge machining and for tool selection in an

industrial grinding process, showing that the approach is beneficial in preventing poor process performance and improving product quality. Cherukuri [22] applied an artificial neural network (NN) to model stability in turning operations using analytical stability study to generate a dataset that trains the NN. Additionally, Khasawneh [23] had combined supervised machine learning with topological data analysis to obtain a descriptor of the process which can detect chatter in turning. And Zhang et al. [24] used Gaussian process regression (GPR) for modelling and predicting surface roughness in end face milling with accuracy of 84%. Furthermore, Aguiar et al. [25] have developed a neural network using a multisensor method to predict the final roughness on the grinded workpiece with 70% success rate. Lela et al. [26] examined the influence of cutting speed, feed and depth of cut on surface roughness in face milling by three different modelling methodologies, namely, regression analysis (RA), support vector machines (SVM) and Bayesian neural network (BNN), and found out that, when the training dataset is small, both BNN and SVR modelling methodologies are comparable with RA methodology and, furthermore, they can even offer better results. In particular, the best results were achieved by BNN, with less error with respect to SVR.

Human learning builds on observations and empirical evidence from the surrounding world: this accumulated knowledge is synthetized, through scientific development, in first principles (FP) models. Similarly, data-driven AI approaches use machine learning (ML) to derive models based on collected data. A combined model allows physical models to be enhanced by AI solutions, to enhance performance and trust. In this paper, a hybrid model structure is proposed: a FP model is built, based on known physical characteristics of the system, and coupled with three data-driven models, to minimize residual errors. In particular, an SVR model with polynomial kernel function, a back-propagation NN with Bayesian regularization and a GPR model with ardexponential kernel function are used as novel methods to attune the FP output to be closer to experimental results. The improved prediction capability of the proposed hybrid models could be profitably used to drive setup of process parameters.

2 Process model

2.1 First principle model

2.1.1 Nonlinear process kinematic

As Rowe [27] describes, considering a plunge centreless grinding, the shape of the cylindrical WP is defined by the radial reduction $r(\theta)$ occurring at the grinding point (i.e. at the contact between wheel and WP), where θ defines WP rotation, starting from zero at the beginning of grinding. The reduction

$r(\theta)$ cumulates during grinding and it is equal to the reduction at the previous revolution $r(\theta - 2\pi)$ summed to the current wheel WP *depth of cut* $I(\theta)$ (also known as *infeed*). Namely:

$$r(\theta) = I(\theta) + r(\theta - 2\pi) \tag{1}$$

He [28] also mentions that, assuming a perfectly rigid system, the depth of cut is obtained from a pure kinematic model, computing the intersection between the grinding wheel and the WP, taking into account the feed movement $X(\theta)$, projected along the WP radius at grinding point, and WP centre displacement due to WP profile contacts at work rest and rubbing wheel.

Given its physical meaning, $I(\theta)$ cannot become negative, as the process either subtracts material, or leaves the surface unvaried when the wheel detaches: thus, the actual depth of cut must be "clipped" to zero if negative:

$$I(\theta) = y_{NL}(X(\theta) + K_1 r(\theta - \alpha) - K_2 r(\theta - (\pi - \beta)) - r(\theta - 2\pi)) \tag{2}$$

where

- K_1 and K_2 are the well-known coefficients relating WP displacement at the grinding contact to radius variation at work rest and rubbing wheel contacts, respectively: $K_1 \triangleq \sin\beta / \sin(\alpha + \beta)$ and $K_2 \triangleq \sin\alpha / \sin(\alpha + \beta)$. The contact angles depend on work-height h_w and work rest angle γ where $\alpha = \pi/2 - \gamma - \beta$ and $\beta = \beta_s + \beta_c$ (Fig. 1). Angles β_s and β_c are given by $\beta_s = \sin^{-1}(2 \cdot h_w/(d_s + d_w))$ and $\beta_c = \sin^{-1}(2 \cdot h_w/(d_{cw} + d_w))$. By introducing $\upsilon \triangleq \beta/\beta_s$, the grinding setup is completely defined by the 3-uple $\{\gamma, \beta, \upsilon\}$.
- $y_{NL}(\cdot)$ is a two-segment piecewise function expressing the "clipping" nonlinearity due to wheel WP detachment:

$$y_{NL}(u) \triangleq \begin{cases} u & \text{if } u \geq 0 \\ 0 & \text{otherwise} \end{cases} \tag{3}$$

Substituting Eq. (2) into Eq. (1), it yields:

$$r(\theta) = y_{NL}(X(\theta) + K_1 r(\theta - \alpha) - K_2 r(\theta - (\pi - \beta)) - r(\theta - 2\pi)) + r(\theta - 2\pi) \tag{4}$$

Knowing that $\theta = \Omega t$, the phase associated to a pulsation ω can be written as $\omega t = (\omega/\Omega)\theta = n\theta$, where n is the number of oscillation cycles in a WP revolution, i.e. the number of lobes. Then, in order to study system stability, Eq. (4) can be rewritten in Fourier domain with $s = jn$ (for sake of readability, the dependency on jn is omitted in the notation):

$$r = y_{NL}\left(X + K_1 r e^{-jn\alpha} - K_2 r e^{-jn(\pi - \beta)} - r e^{-2jn\pi}\right) + r e^{-2jn\pi} \tag{5}$$

2.1.2 Wave filtering

Due to a well-known geometrical interference phenomenon, once a critical limit is exceeded, the amplitude of the waves engraved in WP profile becomes smaller than the amplitude of the relative vibration. According to Hashimoto, contact filtering occurring at regulating wheel/WP contact (denoted with cr) and at wheel/WP contact (denoted with cs) is modelled as:

$$Z_{cs/cr}(n) = \frac{1}{2}\left(1 + \cos\left(\frac{l_{cs/cr}}{d_w}n\right)\right) \tag{6}$$

where $l_{cs/cr}$ is the contact length at the interface and d_w is the WP diameter.

The actual contact length depends both on geometrical factors and bodies compliance:

$$l_{cs} = \sqrt{l_g^2 + l_f^2} \tag{7}$$

where l_g is the geometric contact length ($l_g = \sqrt{I(d_w^{-1} + d_s^{-1})^{-1}}$) and l_f the deflection contact length, estimated as:
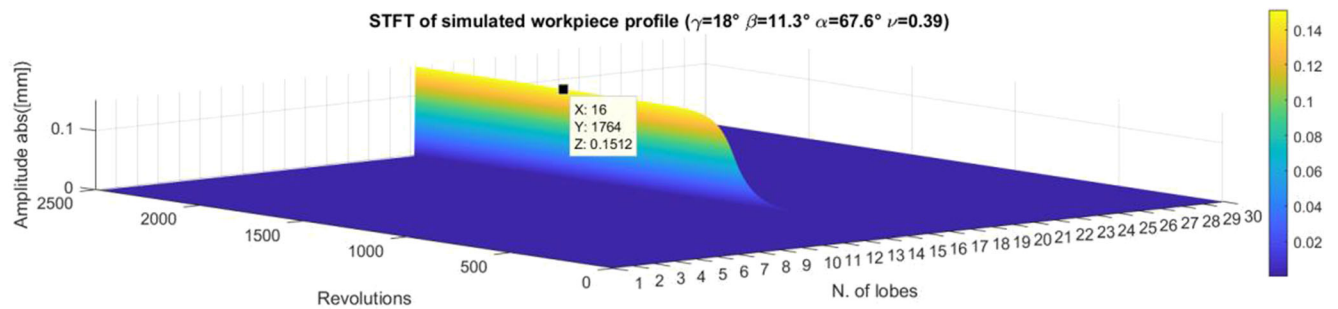

Fig. 2 Model block diagram

Fig. 3 Time-frequency analysis of the sample simulated WP profile

$$l_f = \sqrt{\frac{8 \cdot R_r^2 \cdot F_n' \cdot \left(d_w^{-1} + d_s^{-1}\right)^{-1}}{\pi E^*}} \tag{8}$$

where

- R_r is a roughness factor equal to 1 for a smooth cylinder but typically ranges from 5 to 15 for a grinding wheel.
- F_n' is the normal grinding force per unit width.
- E^* is the combined elastic properties of the grinding wheel and workpiece, i.e. $\frac{1}{E^*} = \frac{1-\nu_1}{E_1} + \frac{1-\nu_2}{E_2}$, where E_1, E_2, ν_1, and ν_2 are the Young modules and Poisson ratio, respectively.

2.1.3 Stiffness factor

Grinding machines static and dynamic compliance affects wheel-workpiece relative displacement. At very low frequencies—namely, lower than the first significant natural frequency—only static compliance can be considered; thus,

Fig. 4 Sample WP profile after 950 revolutions

deflection is approximately in phase with the depth of cut. Static deflection Δ_{el} is given by:

$$\Delta_{el} = K_S/(K_m I_e) \tag{9}$$

where K_S (process stiffness) is the ratio between normal grinding force and actual infeed I_e and K_m is the wheel-workpiece static stiffness. Whereas the actual infeed I_e depends on the nominal infeed I by $I_e = I - \Delta_{el}$, Eq. (12) yields:

$$I_e = KI \tag{10}$$

with $K \triangleq \dfrac{K_m}{K_m + K_S}$ (stiffness factor) \hfill (11)

2.1.4 Simulation model
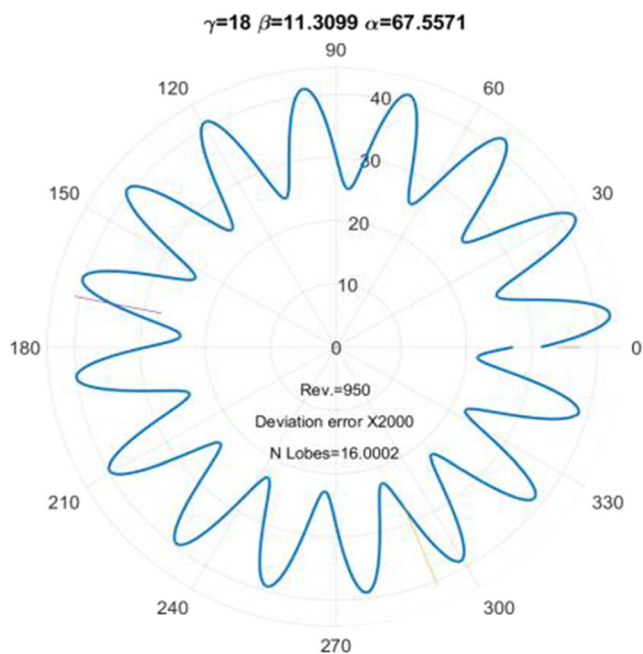
The obtained process model is represented by the block diagram of Fig. 2. Based on it, a numeric simulation code has been developed in MATLAB™ to estimate the WP profile, discretized by a circular array of 7200 elements, representing WP radial reduction at a given angular position. The contact filtering illustrated in Section 2.2 has been implemented by a 0-phase symmetric FIR filter: a high order of 361 has been selected to fit properly the $Z_{cs}(n)$ of Eq. (8). Contact length l_{cs} has been computed using Eqs. (9), (10) and (11). All the

Table 1 DOE-controlled parameters

	Control parameters	Ranges and discrete values
1	Workpiece height	From -5 mm to +30 mm
2	Workpiece diameter (WPD)	Ø5; Ø20; Ø60
3	Workpiece length (WPL)	50 mm; 100 mm;150 mm
4	Supporting blade angle (γ)	20°; 30°; 40°
5	Feed rate	From 0.010 mm/s to 0.004 mm/s
6	Regulating wheel velocity	From 10 rpm to 30 rpm
7	Infeed	From 0.15 mm to 20 mm

Table 2 DOE-fixed parameters

	Fixed parameters	Values
1	Grinding wheel linear velocity (Vs)	32 m/s
2	Regulating wheel diameter (nominal)	300 mm
3	Operating wheel diameter (nominal)	500 mm
4	Operating wheel type (ANSI B74 13–1977)	500X165X254A 24A80L4V19
5	Regulating wheel type (ANSI B74 13–1977)	300X157X127A 10A80RR
6	Workpiece material	C45

necessary parameters, such as R_r, $F_{n'}$, E_1, E_2, ν_1 and ν_2, have been taken from literature, given the WP material and grinding wheel type and status. Stiffness factor K has been estimated fitting an exponential decay on wheel spindle current signal during spark-out tests, as described in, par. 19.11.4.

Figure 3 depicts the Short-time Fourier transform (STFT) of the WP profile for a simulated unstable operation. It can be observed the transition from an initial linear unstable condition, with exponential growth, to a steady state condition due to clipping nonlinearity. Moreover, only one dominant harmonic component raises: the simulated profile, with "number of lobes" equal to 16 lobes, is plotted in Fig. 4 and used to compute the "simulated roundness". Additionally, a "detach index" is evaluated, indicating the processing time percentage when detachment occurs. Exploiting the conceptual process model of Fig. 2, an analytical stability analysis is performed by the Nyquist criterion, delivering the exponential growth rate that, multiplied by the number of WP revolutions, produces the so-called "stability index". The *simulated roundness*, *detach index*, *simulated number of lobes* and

stability index plus *dullness level of the grinding wheel*, which is calculated after each test, will be used as additional inputs by the hybrid model described in the following.

2.1.5 Experimental verification

To validate the obtained model, a series of experimental tests were done, varying seven independent control parameters (Table 1) by a randomly generated "latin hypercube" approach, while keeping 7 fixed parameters (Table 2), for a total of 100 samples. Ranges have been selected according to the industrial practice, by surveying technologists experience. The discrete variables (γ, WPD, WPL) have been chosen by discretizing the continuous random variables.

The test has been executed on a Monza 520 M6 century edition grinding machine by "Monzesi srl", with manual work rest blade adjustment (Fig. 5). The WP roundness has been measured with a "Mitutoyo Roundpack 400" system, at three different highs along the WP, producing then an *average roundness* for each WP.

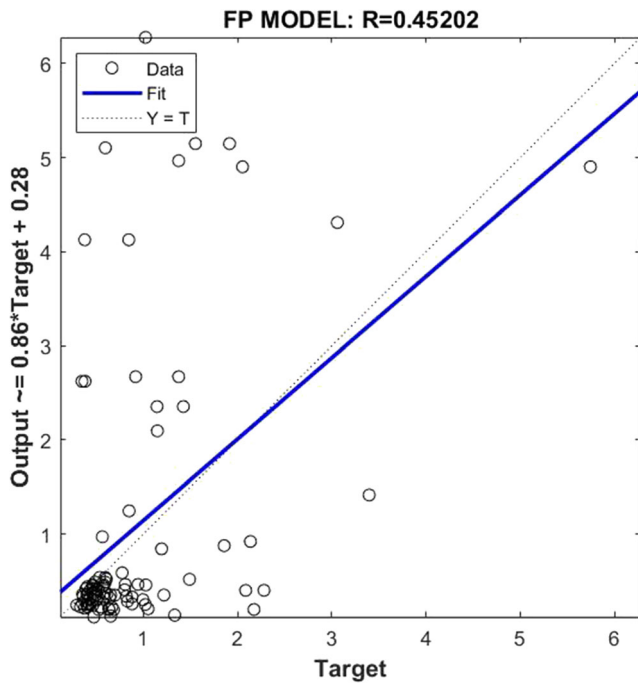**Fig. 5** Monza 520 M6 century edition (working zone)

Fig. 6 Correlation between simulated and experimental workpiece normalized roundness error R_N

When the process is unstable, waviness grows exponentially until grinding wheel/workpiece detachment occurs. Roundness final value, both experimentally and in simulations, is very sensitive to several system parameters. To effectively support industrial production, the analysis must be able to discriminate between acceptable and excessive roundness errors "RE". Instead of adopting a discrete classification approach, this judgement is reproduced computing a normalized roundness error "RE_N" by a pseudo-sigmoid function, adapted

from [29], that mitigates the disrupting effect of uncertainty and epistemic uncertainty:

$$RE_N = \frac{RE}{RE^{0.6} + 2} \tag{12}$$

The relationship between normalized roundness R_N predicted by the FP model and measured from the experiments is plotted, for all samples, in Fig. 6, showing an R correlation value around 0.45.

3 Machine learning techniques

3.1 Learning setup

In the pre-processing stage, data has been checked for missing values and outliers and, in case of existence, they were removed which in this case only one point was eliminated. For reducing the effect of different scales across input variables on numerical conditioning during training, input data normalization has been performed to have values between −1 and 1. Then, a random selection of 20% of dataset was separated and kept aside as final test data, while the remaining 80% of original data was used to train the models. Furthermore, since the dataset was small, to improve prediction generality and remove effects of the initial random values, all algorithms were trained separately 100 times. At each single model training, the training dataset was randomly divided into two parts, 70% of the data was used for training and the remaining 30% was used for testing of that model.
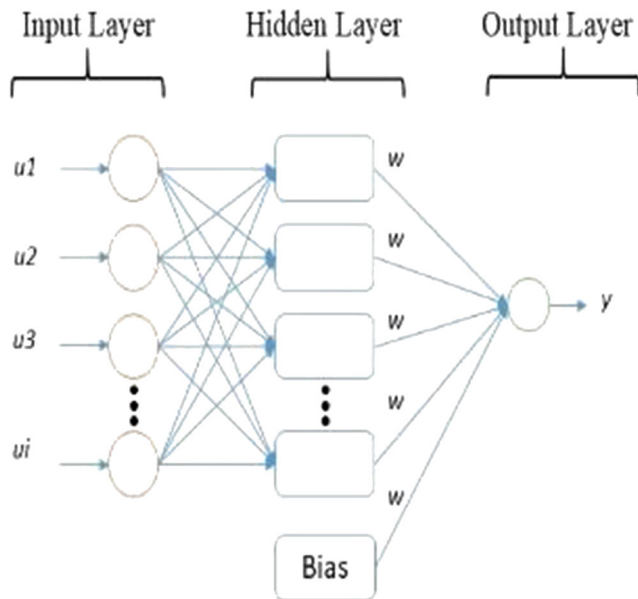
3.2 Neural network model

In this approach, the system was modelled using a pure NN method (Fig. 7). For this aim, Kayri [30] has suggested using a Bayesian regularized artificial neural network (BRANN), because they limit overtraining and overfitting. The input vector of independent variables u_i is linked to the target using the design represented in Fig. 8, with one hidden layer. Each single node is associated to those of the previous layer by compliant weights. In each neuron, inputs are multiplied by the corresponding weight and summed. Nevertheless, Alados et al. [31] propose to apply an activation function to the resulting sum, producing the neuron output, which is transferred to the next layer. After the feed forward step, the variance between the real output and predicted output of the network is defined as error. In the back-propagation phase, the error signal is transmitted from the output to the input layer via a layer-by-layer manner, adjusting the network weight. As Okut [32] explicates,

Fig. 7 The architecture of an artificial neural network

Fig. 8 Neural network architecture
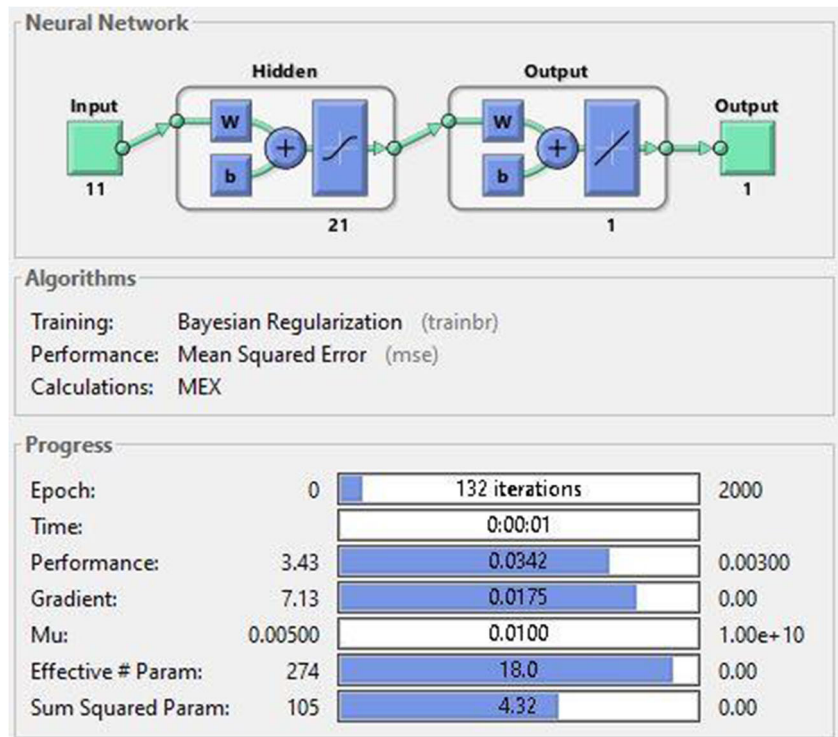
the objective function used in BRANN has an additional term that castigates large weights: because of this contraction, the effective number of parameters attained by BRANN can be less than the total number of available parameters. In this way, a smoother mapping is achieved, with reduce overfitting and improved model generalization ability.

3.2.1 Neural network-only model

The neural network model is defined via few hyperparameters, as the number of hidden layers and the number of neurons in each layer, etc. They must be selected before the training stage but there is no unique or predefined way of selecting them. In this study, a series of experiments has been done with different
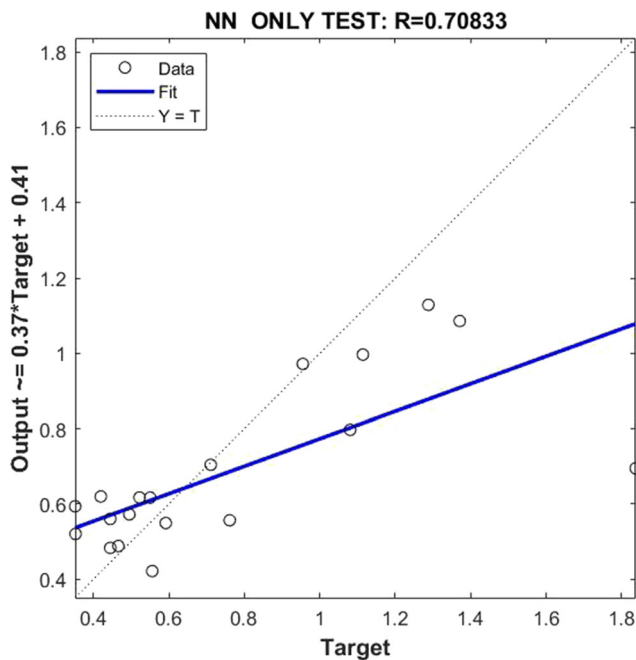
Fig. 9 NN test regression plot
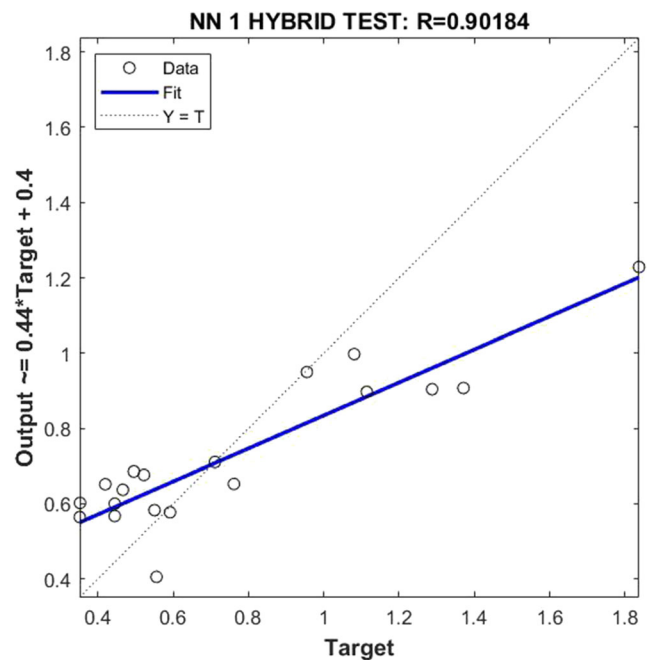
Fig. 10 NN hybrid test regression plot

Table 3 SVR optimal SVR hyperparameters

Kernel function	Box constraint	Epsilon (ε)
Polynomial order 3	0.32966	0.00055033

Table 4 GPR parameters

Kernel function	Sigma	Predict method	Optimizer
Ardexponential	0.02	Exact	fmincon

hyperparameters arrangement and it was observed that the best network performance was achieved with the parameters reported in the following. The performance of the networks is measured by the achieved minimum mean square error (MSE) and maximum correlation value R.

A network with one hidden layer has been considered since the number of effective parameters and overall performance of model in terms of its prediction precision did not change significantly by adding more hidden layers. The number of neurons in each hidden layer is selected accordingly to the number of input parameters (NI): $2 \cdot NI - 1$, according to the rule of thumb suggested by João [33]. As Kayri [30] mentioned, adding more neurons resulted in increasing computation time without performance improvements: the effective number of parameters is unchanged. During BRANN training, the less relevant weights are set to zero. In the hidden layer, a tangent hyperbolic activation function was chosen as suggested [34], since they reduce the number of training iterations and because the output of the system is a real number. On the other side, for the output layer, a "purelin" activation function is used.

MATLAB (2020a) Statistics and Machine Learning toolbox was employed for exploring the BR artificial neural network. In this study, as Reece [35] reported, the training process is halted if (a) it reaches the maximum number of iterations (2000); (b) the maximum amount of time is exceeded (no limit has been considered); (c) the estimation error is below

the target (0.003); (d) the performance gradient drops below minimum gradient; or (d) the Marquardt adjustment parameter (μ) becomes larger than 10^{10}.

To improve generalization, taking into account the small dataset used, Shaikhina [36] suggested to train multiple NNs, selecting a random training dataset (while preserving the initial 20% of the data for the final evaluation). Therefore, 100 neural networks, with the previously described architecture, were trained, with all seven independent parameters used in FP plus the dullness level of the grinding wheel. Then only top 10 models, based on their R value, were selected for final testing stage: the average output of these models was used for roundness predictions based on the 20% dataset kept as final test data. The resulting correlation between them and the actual roundness values for each test is represented in Fig. 9, with an R value around 0.7, i.e. a 52% increase over the FP model R value.

3.2.2 Neural network-hybrid model

It is clear that both the FP and NN approaches have their advantages and drawbacks. But as Driscoll et al. [37] reported, generally physic models tend to have higher bias but lower variance. On the other hand, machine learning models tend to have a high variance and low bias. It is therefore suggested to combine both methods to deliver a further robust system. In the proposed hybrid model structure by Ahmad [38] $g(\mathbf{z})$, the
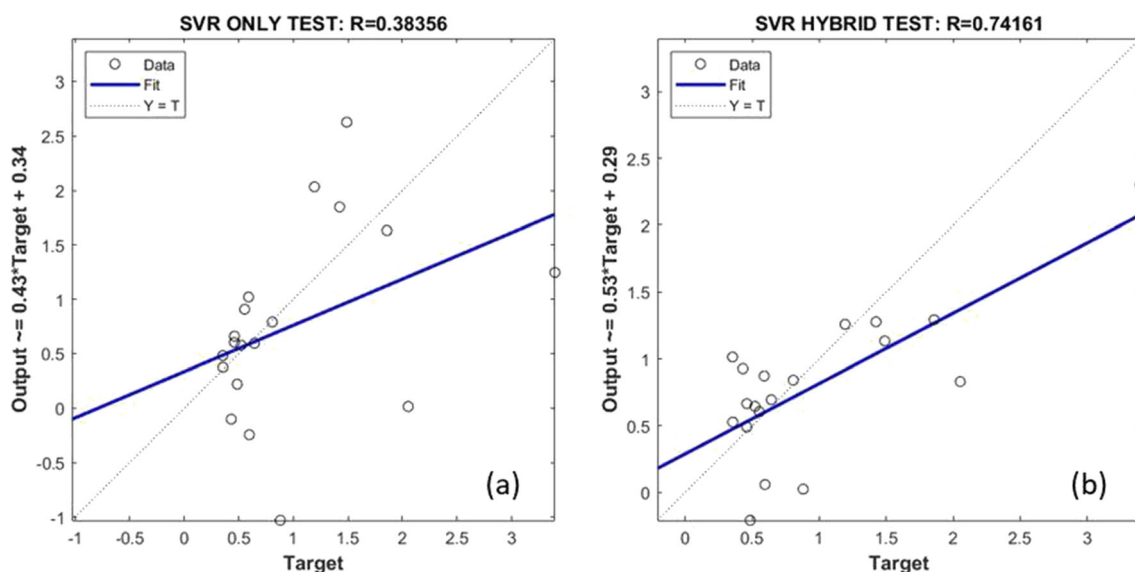
Fig. 11 SVRs test regression plots: **a** SVR only, **b** SVR hybrid
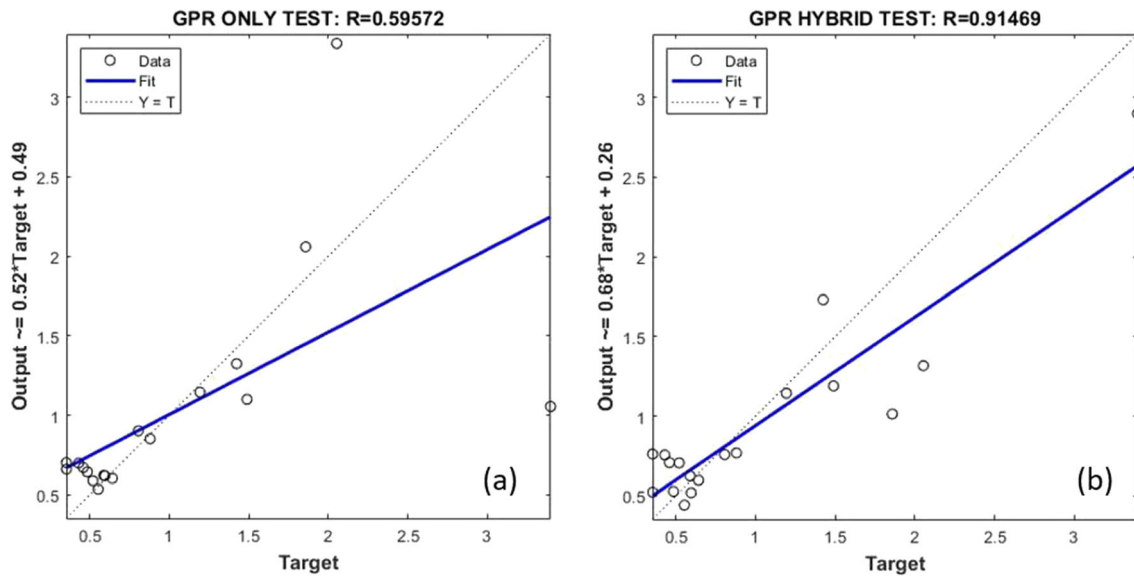
Fig. 12 GPRs test regression plots: **a** GPR only, **b** GPR hybrid

output $f(\mathbf{x})$ of the FP model, based on known parameters "$\mathbf{x}$", is exploited as an additional input for the NN model, together with the global inputs "$\mathbf{z}$" to the process. Parameters "$\mathbf{x}$" are a subset of "$\mathbf{z}$", as wheel dullness indicator (that is considered by the ML models) does not play any role in the adopted physical model:

$$y = \mathrm{g}(\mathbf{z}, f(\mathbf{x})) \tag{13}$$

For doing so, an NN model is considered as before, with one hidden layer and number of neurons selected, based on previously explained method. The input parameters of this new NN are the ones from NN in the previous section plus primary and secondary outputs from the FP (explained in Section 2.1.4.) such as expected roundness, detach index, expected number of lobes and stability parameter. As previously explained, inputs are normalized and a BRANN with the same activation functions was adopted. Then, 100 models were

trained and tested as before and consequently, top 10 models were selected, to perform the final test.

The results obtained from the hybrid model (Fig. 10) illustrate that with this approach, in the test dataset, the average correlation value R between data and predicted value reaches 0.9, with an improvement of 28% with respect to the NN-only method with same architecture and an improvement of 89% with respect to the sole FP method.

3.3 Support vector regression (SVR)

Support vector machine (SVM) analysis is a widespread machine learning tool for classification and regression, first proposed by Vapnik [39]. SVM regression is contemplated as a nonparametric technique because it depends on kernel functions. As Campbell [40] describes, SVMs have properties such as good generalization ability and a small number of free

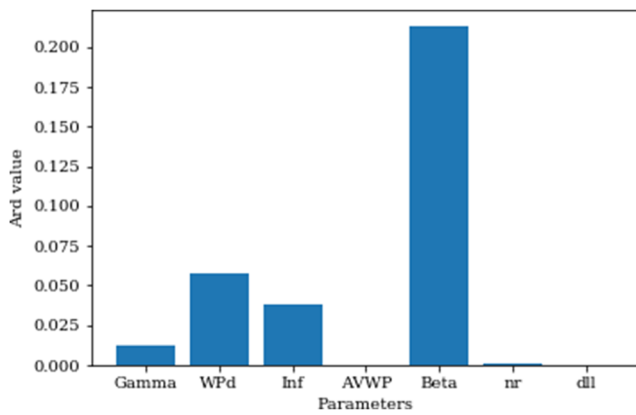**Fig. 13** Precision comparison between models
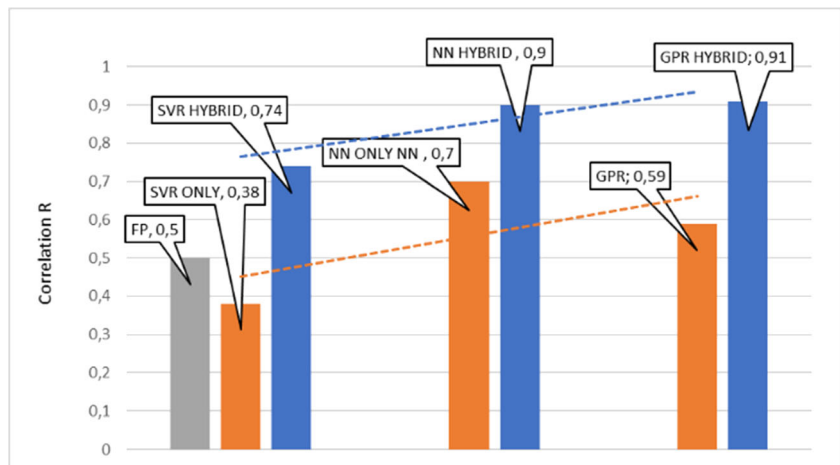
Fig. 14 Primary model inputs sensitivity

adjusting parameters, and unlike NN approach, it has no prerequisite for designing the architecture of the machine learning model. They were initially established for classification tasks, but they can be employed in regression problems as support vector regression (SVR) by including of a loss function based on a distance measure. SVR constructs a linear model after the input has been mapped into a higher dimensional feature space using some nonlinear mapping (usually by reproducing kernels). The estimation accuracy of SVR depends on three hyperparameters:

- Kernel function, such as linear function, polynomial function, radial basis function and sigmoid function.
- Box constraint: parameter that controls the maximum penalty imposed on margin-violating observations and aids in preventing overfitting (regularization). If the box constraint is increased, the SVM classifier assigns fewer support vectors. However, increasing the box constraint can lead to longer training times.
- Insensitive loss function (ε): the value of ε influences the number of support vectors used to form the regression function. If ε increases, fewer support vectors are chosen, and the smoothness of the regression function increases too, explains [26].
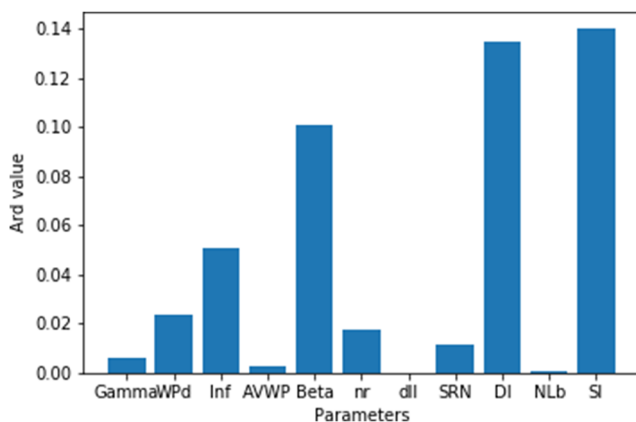
Fig. 15 Primary and secondary model inputs sensitivity

For finding those parameters, an automatic hyperparameters optimization was conducted on the dataset: the best combination based on model MSE is reported in Table 3.

3.3.1 SVR model

Once SVR hyperparameters are selected, the SVR model is ready for the learning process. As explained in Section 2.2.1, two different SVR models were created: one using the seven input parameters from FP model and one using primary and secondary inputs, by the hybrid method.

As before, the top 10 models from each approach were selected and averaged, to perform the final test. Figure 11 shows average R values achieved with these two models: in the SVR only approach, the correlation between predicted and actual roundness was less than in the FP model, but, by the hybrid approach, the correlation value was doubled and performed even better than FP model with 73% improvement.

3.4 Gaussian process regression (GPR)

Gaussian process regression (GPR) models are nonparametric kernel-based probabilistic models, which make them powerful tools for Bayesian supervised learning. They have in common with support vector machines and regularization networks the theory of regularization via reproducing kernels that allow the straight specification of the smoothness properties of the class of functions under consideration. As Smola [41] stated, this makes them popular for regression of small datasets. Through a series of experiments and automatic hyperparameters optimization bringing the least MSE, a GPR model with customized parameters (Table 4) is selected for roundness prediction.

Employing top 10 models from 100 trained and tested models and averaging results obtained from each of them for roundness prediction, the difference in terms of correlation between predicted and actual values is much more significant between two approaches. As shown in Fig. 12, the correlation value obtained by GPR hybrid method is raised by 100% with respect to FP model and 50% with respect to GPR only method.

4 Discussion

This paper presents results from attempting to combine first principle and machine learning techniques into hybrid models to forecast performance of a centreless grinding process in terms of workpiece final roundness. The FP model is established on theoretical calculations and assumptions that need prior understanding of the system under the study. The adopted FP model, while similar to models discussed in similar literature, exhibits an unsatisfactory accuracy, probably because of improperly treated phenomena, e.g. contact length

dependency on local compliance, on contact force and on grinding wheel roughness; variability in material properties of the grinding wheel and workpiece; and progressive grinding wheel wear and dullness. On the contrary, NN, SVR and GPR models are established straight from experimental data, engaging a range of statistical techniques to identify suitable mathematical models. However, precise estimates are possible only if a large training set is available, with accurate data. Therefore, for obtaining the benefits of both methods, they have been combined into a hybrid modelling approach. From NN approach, it was observed that the hybrid model with one hidden layer leads to significant reduction in prediction error. It was also noticed that, using the SVR and GPR approaches, the improvement from normal model to the hybrid model was more significant. Comparing these seven models, the best test results was obtained by hybrid GPR model that, in addition, takes considerably less time for training with respect to an NN hybrid with a similar precision. Figure 13 presents a brief comparison between all seven models in terms of precision.

For understanding what causes this difference between pure machine learning (ML) models and hybrid models, a sensitivity analysis has been conducted by automatic relevance determination (ARD) based on GPR model, on both input datasets [42]. Figure 14 illustrates that in standard data-driven models, the roundness value is most sensitive to the Beta angle and then to WP diameter, infeed rate (Inf) and Gamma angle. However, by adding primary and secondary outputs of FP model as secondary input in ML models, new parameters are more interrelated with WP roundness: Fig. 15 shows that simulated roundness (SRN) and simulated number of the lobes (NLb), estimated by the FP model, do not help ML model as expected. The secondary outputs of the FP model, such as the detach index (DI) and the stability index (SI), have the highest correlation with WP roundness.

5 Conclusion

A hybrid approach has been suggested to improve roundness predicted by the numerical models based on first principles. Data-driven models, based on nominal process parameters augmented by additional outputs provided by a Physics-based model, deliver an optimal estimation that alleviates the effect of the unavoidable uncertainties and epistemic errors. The results presented in this paper indicate that even a first principal model with mediocre primary output performance includes other secondary outputs which are beneficial to enhance machine learning techniques, compared to black box approaches with the same techniques.

Future activities will be aimed at improving the first principle model. A larger campaign on a commercial centreless grinding machine will be used to examine more combination of parameters and further improve prediction accuracy.

Compliance with ethical standards

References

1. Dhavlikar M, Kulkarni M, Mariappan V (2003) Combined Taguchi and dual response method for optimization of a centerless grinding operation. J Mater Process Technol 132(1–3):90–94. https://doi.org/10.1016/S0924-0136(02)00271-6

2. Zhou SS, Gartner JR, Howes TD (1996) On the relationship between setup parameters and lobing behavior in centerless grinding. CIRP Ann - Manuf Technol 45(1):341–346. https://doi.org/10.1016/S0007-8506(07)63076-5

3. Gallego I (2007) Intelligent Centerless grinding: global solution for process instabilities and optimal cycle design. CIRP Ann - Manuf Technol 56(1):347–352. https://doi.org/10.1016/j.cirp.2007.05.080

4. Klocke F, Friedrich D, Linke B, Nachmani Z (2004) Basics for in-process roundness error improvement by a functional Workrest blade. CIRP Ann 53(1):275–280. https://doi.org/10.1016/S0007-8506(07)60697-0

5. Dall A (1946) Rounding effect in centerless grinding. Mech Eng ASME 58:325–329

6. Yonetsu S (1959) Consideration of centerless grinding characteristics through harmonic analysis of out-of-roundness curves. Proc Fujihara Meml Fac Eng Keio Univ 12(47):184–202

7. Rowe WB, Barash MM (1964) Computer method for investigating the inherent accuracy of centreless grinding. Int J Mach Tool Des Res 4(2):91–116. https://doi.org/10.1016/0020-7357(64)90002-2

8. Marinescu ID, Hitchiner MP, Uhlmann E, Rowe WB, Inasaki I (2006) Handbook of machining with grinding wheels. CRC Press

9. Bueno R, Zatarain M, Aguinagalde JM, Le Maître F (1990) Geometric and dynamic stability in centerless grinding. CIRP Ann - Manuf Technol 39(1):395–398. https://doi.org/10.1016/S0007-8506(07)61081-6

10. Rowe WB, Richards DL (2016) Geometric stability charts for the centerless grinding process. J Mech Eng Sci 14(2):155–160

11. Bianchi G, Leonesio M, Safarzadeh H (2020) A double input describing function approach for stability analysis in centerless grinding under interrupted cut. Int J Adv Manuf Technol. https://doi.org/10.1007/s00170-020-05362-2

12. Lizarralde R, Barrenetxea D, Gallego I, Marquinez JI, Bueno R (2005) Practical application of new simulation methods for the elimination of geometric instabilities in centerless grinding. CIRP Ann 54(1):273–276. https://doi.org/10.1016/S0007-8506(07)60101-2

13. Zhou ZX, van Lutterwelt CA (1992) The real contact length between grinding wheel and workpiece - a new concept and a new measuring method. CIRP Ann - Manuf Technol 41(1):387–391. https://doi.org/10.1016/S0007-8506(07)61228-1

14. Liu H, Chen Q, Li B, Mao X, Mao K, Peng F (2011) On-line chatter detection using servo motor current signal in turning. Sci China Technol Sci 54(12):3119–3129. https://doi.org/10.1007/s11431-011-4595-6

15. Hashimoto F (2017) Model Development for Optimum Setup Conditions that Satisfy Three Stability Criteria of Centerless Grinding Systems. Inventions 2(4):26. https://doi.org/10.3390/inventions2040026

16. Zakharov OV, Datskovskaya EA (2010) Setup of centerless superfinishing machine tools. Russ Eng Res 30(12):1263–1267. https://doi.org/10.3103/S1068798X10120191

17. Barrenetxea D, Marquinez JI, Álvarez J, Fernández R, Gallego I, Madariaga J, Garitaonaindia I (2012) Model-based assistant tool for the setting-up and optimization of centerless grinding process. Mach Sci Technol 16(4):501–523. https://doi.org/10.1080/10910344.2012.729480

18. Sjöberg J, Zhang Q, Ljung L, Benveniste A, Delyon B, Glorennec PY, Hjalmarsson H, Juditsky A (1995) Nonlinear black-box modeling in system identification: a unified overview. Automatica 31(12):1691–1724. https://doi.org/10.1016/0005-1098(95)00120-8

19. Rowe WB, Yan L, Inasaki I, Malkin S (1994) Applications of artificial intelligence in grinding. CIRP Ann 43(2):521–531. https://doi.org/10.1016/S0007-8506(07)60498-3

20. Junkar M, Filipie B, Bratko I (1991) Identifying the grinding process by means of inductive machine learning

21. Filipic B, Junkar M (2000) Using inductive machine learning to support decision making in machining processes

22. Cherukuri H, Perez-Bernabeu J, Selles JA, Schmitz TL (2019) A neural network approach for chatter prediction in turning. Proc Manuf 34:885–892. https://doi.org/10.1016/j.promfg.2019.06.159

23. Khasawneh FA, Munch E, Perea JA Chatter classification in turning using machine learning and topological data analysis https://doi.org/10.1016/j.ifacol.2018.07.222

24. Zhang G, Li J, Chen Y, Huang Y, Shao X, Li M (2014) Prediction of surface roughness in end face milling based on Gaussian process regression and cause analysis considering tool vibration. Int J Adv Manuf Technol 75(9–12):1357–1370. https://doi.org/10.1007/s00170-014-6232-6

25. Aguiar PR, Cruz CED, Paula WCF, Bianchi EC (2008) Predicting surface roughness in grinding using neural networks. Adv Robot Autom Control 480

26. Lela B, Bajić D, Jozić S (2009) Regression analysis, support vector machines, and Bayesian neural network approaches to modeling surface roughness in face milling. Int J Adv Manuf Technol 42(11–12):1082–1088. https://doi.org/10.1007/s00170-008-1678-z

27. Rowe WB (Apr. 1979) Research into the mechanics of Centreless grinding. Precis Eng 1(2):75–84. https://doi.org/10.1016/0141-6359(79)90137-5

28. Rowe WB (2014) Principles of modern grinding technology (second edition). Elsevier Inc

29. Schütt HH, Harmeling S, Macke JH, Wichmann FA (2016) Painfree and accurate Bayesian estimation of psychometric functions for (potentially) overdispersed data. Vis Res 122:105–123. https://doi.org/10.1016/j.visres.2016.02.002

30. Kayri M (2016) Predictive abilities of Bayesian regularization and Levenberg–Marquardt algorithms in artificial neural networks: a comparative empirical study on social data. Math Comput Appl 21(2):20. https://doi.org/10.3390/mca21020020

31. Alados I, Mellado JA, Ramos F, Alados-Arboledas L (2004) Estimating UV erythemal irradiance by means of neural networks. Photochem Photobiol. https://doi.org/10.1562/2004-03-12-RA-111

32. Okut H (2016) Bayesian Regularized Neural Networks for Small n Big p Data. Artificial Neural Networks - Models and Applications, InTech

33. João NCCL, Rosa PS, Guerra DJD, Horta NCG, Martins RMF (2019) Using artificial neural networks for analog integrated circuit design automation. Springer Nature

34. Jurkovic Z, Cukor G, Brezocnik M, Brajkovic T (Dec. 2018) A comparison of machine learning methods for cutting parameters prediction in high speed turning process. J Intell Manuf 29(8):1683–1693. https://doi.org/10.1007/s10845-016-1206-1

35. Reece PL (2007) Progress in smart materials and structures. Nova Science Publishers, New York, p 372

36. Shaikhina T, Khovanova NA (2017) Handling limited datasets with neural networks in medical applications: a small-data approach. Artif Intell Med 75:51–63. https://doi.org/10.1016/j.artmed.2016.12.003

37. O'Driscoll P, Lee J, Fu B (2019) Physics Enhanced Artificial Intelligence. pp. 1–8

38. Ahmad I, Kano M, Hasebe S, Kitada H, Murata N (2014) Gray-box modeling for prediction and control of molten steel temperature in tundish. J Process Control 24(4):375–382. https://doi.org/10.1016/j.jprocont.2014.01.018

39. Vapnik VN (1999) An overview of statistical learning theory. IEEE Trans Neural Netw 10(5):988–999. https://doi.org/10.1109/72.788640

40. Campbell C (2002) Kernel methods: a survey of current techniques. Neurocomputing 48(1–4):63–84. https://doi.org/10.1016/S0925-2312(01)00643-9

41. Smola AJ, Bartlett P (2000) Sparse Greedy Gaussian Process Regression

42. Burden F, Winkler D (2008) Bayesian regularization of neural networks