



Data-driven prediction of the surface layer state in hard-turning for optimization of component quality

Felix Wittich¹ · Thomas Wegener² · Alexander Liehr² · Wolfgang Zinn² · Thomas Niendorf² · Andreas Kroll¹

Received: 29 August 2023 / Accepted: 8 December 2023 / Published online: 30 January 2024
© The Author(s) 2024

Abstract

Data-driven approaches are an effective solution for modeling problems in machining. To increase the service life of hard-turned components, it is important to quantify the correlation between the cutting parameters such as feed rate, cutting speed and depth of cut and the near-surface properties. For obtaining high-quality models with small data sets, different data-driven approaches are investigated in this contribution. Additionally, models that enable uncertainty quantification are crucial for effective decision-making and the adjustment of cutting parameters. Therefore, parametric multiple polynomial regression and Takagi–Sugeno models, as well as non-parametric Gaussian process regression as a Bayesian approach are considered and compared regarding their capability to predict residual stress and surface roughness values of 51CrV4 specimens after hard-turning. Moreover, a novel method based on optimization of data driven non-linear models is proposed that allows for identification of cutting parameter combinations, which at the same time lead to satisfactory surface roughness and residual stress states.

Keywords Surface integrity · Data-driven modeling · Hard-turning · Residual stress

1 Introduction

Surface finishing machining processes are well-known to strongly influence the near-surface properties of a component, i.e., topography, residual stress state, microstructure, strain hardening or crack pattern, among others [1]. As the surface usually represents the area being subject to highest stresses, especially in case of highly-loaded components, the state of the near-surface layer has a significant influence on the mechanical properties of the component. The correlation of surface characteristics and mechanical properties is referred to as surface integrity (SI) [2]. As a result, robust surface conditioning in manufacturing processes controlling both geometric features and near-surface properties is of great importance in order to improve the service life time and reliability of components [3]. Here, data-driven

modeling concepts represent a promising approach to overcome the challenges related to SI in surface finishing machining processes [2]. In this context, 12 projects within the Research Priority Program 2086 of the German Research Foundation (DFG) focused on the development of process control systems for surface machining using a combination of in-process soft sensors and process knowledge, which allows defined geometric features as well as near-surface properties to be adjusted simultaneously in metallic components [4]. Within the Research Priority Program 2086 several different surface machining processes such as turning (including hard-turning and external longitudinal turning) [5, 6], deep-hole drilling [7], grinding [8] and milling [9] were considered. In their project, the groups of the authors in this contribution focused on the prediction of the near-surface layer properties for optimization of a hard-turning process using in-process measurement technology and data-driven modeling methods. In this context, an online implementation of the Fraunhofer IZFP 3MA-II system in the hard turning process was realized [10]. A prediction model for post-process residual stress depth profiles was developed in [11] as well as for integral width in [12] including a comparison of different model approaches, i.e., global multiple polynomial regression models and locally affine Takagi

✉ Felix Wittich
felix.wittich@mrt.uni-kassel.de

¹ Department of Measurement and Control, University of Kassel, Mönchebergstraße 7, 34125 Kassel, Germany

² Institute of Materials Engineering, Metallic Materials, University of Kassel, Mönchebergstraße 3, 34125 Kassel, Germany

Sugeno models. For the latter one, bounded-error parameter estimation and nonparametric Gaussian process regression were applied and compared in [13] allowing for quantification of the uncertainty of the prediction.

Modeling of the surface layer (SL) state can be accomplished using analytical, numerical or empirical approaches [14]. Predictions using finite element models are computationally expensive and therefore, not suitable for real time application. Analytical models require in-depth knowledge of the process and can therefore, be difficult to develop [15]. Collection of experimental data for manufacturing processes is time and cost intensive. Therefore, empirical modeling approaches are advantageous, which achieve good model quality with small data sets. In [16], the relationship between cutting parameters and the resulting microhardness and surface roughness for hardened 100Cr6 (AISI 52100) was investigated using the response surface methodology. [17] used artificial neural networks (ANN) to predict residual stresses at five different depths for Ck 53 (AISI 1053) and Ck 67 (AISI 1070). Similarly, [18] used ANNs to model characteristic points of the residual stress profile for 100Cr6 (AISI 52100) depending on the cutting states and used FEM simulations to create additional training data since neural networks require a large number of training data. Existing literature mainly uses multiple regression approaches or data intensive ANNs. In this contribution nonlinear modeling approaches are investigated that can achieve better prediction performance with less data. The SL properties residual stress and surface roughness are investigated. Moreover, two initial hardness states are considered.

From a manufacturer's point of view near-surface properties of a component, i.e., surface roughness and residual stress state, are of great interest. This applies in particular for components being subject to cyclic loads. In this case, surface roughness promotes fatigue crack initiation due to notch effects whereas compressive residual stresses are well known to delay or prevent crack initiation and propagation. Nevertheless, cutting parameter sets leading to a low surface roughness do not necessarily result in an advantageous compressive residual stress state and vice versa. Here data driven modeling can help identifying cutting parameter values leading to satisfactory surface roughness and residual stress states.

2 Modeling and methods

2.1 Data-driven modeling

In this contribution, empirical modeling approaches are considered, that allow for modeling nonlinear systems given only small data sets. The goal is to find a functional

relationship between n input variables $\mathbf{x} \in \mathbb{R}^n$ and the output $y \in \mathbb{R}$ of interest

$$y(l) = f(\mathbf{x}(l)) + e(l), \quad l = 1, \dots, N \quad (1)$$

given the experimental data set $Z^N = \{\mathbf{x}(l), y(l)\}_{l=1}^N$ with N elements. The function $f : \mathbb{R}^n \rightarrow \mathbb{R}$ is the nonlinear function representing the system behavior to be modeled. For the hard-turning process, the model input \mathbf{x} consists of the cutting parameters v_c, f, a_p (cutting speed, feed rate and depth of cut) as well as the initial hardness HV_{init} . For models with a depth profile the depth under the surface d_s is also considered as an input. The model output $y \in \mathbb{R}$ is the respective SL state. For each SL state a multi input single output (MISO) model is estimated. The term $e(l)$ accounts for measurement errors and other disturbances. A more detailed discussion is given in Sect. 2.2 in the context of uncertainty modeling. Three model classes are considered: two parametric and one nonparametric approach. Firstly, multiple polynomial regression (MPR) is considered as a standard parametric approach in Sect. 2.1.1. Secondly, Takagi Sugeno (TS) models are utilized (Sect. 2.1.2), which provide for a more flexible parametric modeling approach by partitioning the system into local models. Thirdly, Gaussian process regression (GPR) is considered (Sect. 2.1.3), which is a very flexible nonparametric modeling concept.

2.1.1 Multiple polynomial regression

MPR of degree m is considered as baseline:

$$\begin{aligned} \hat{y}(l) = & \theta_0 + \sum_{i=1}^n \theta_i x_i(l) + \sum_{i_1=1}^n \sum_{i_2=i_1}^n \theta_{i_1 i_2} x_{i_1}(l) x_{i_2}(l) \\ & + \dots + \sum_{i_1=1}^n \dots \sum_{i_m=i_{m-1}}^n \theta_{i_1 \dots i_m} x_{i_1}(l) \dots x_{i_m}(l), \end{aligned} \quad (2)$$

with the model parameters $\theta_r \in \mathbb{R}$, $r = 0, \dots, M$, $M = (m+n)!/(m!n!) - 1$, corresponding to the $M+1$ model terms in (2). With

$$\boldsymbol{\varphi}(l) = [1 \quad x_1(l) \dots x_n(l) \quad x_1^2(l) \quad x_1(l)x_2(l) \quad \dots \quad x_{n-1}(l)x_n^{m-1}(l) \quad x_n^m(l)]^T \quad (3)$$

(2) can be written in linear regression form:

$$\hat{y}(l) = \sum_{r=0}^M \theta_r \varphi_r(l), \quad (4)$$

with $\varphi_r(l)$ being the r th element of $\boldsymbol{\varphi}(l)$ in (3). Model training consists of estimating all θ_r in (4), e.g. by minimizing the sum of squared errors for all N data, which is an ordinary least squares estimation problem. Because of the large amount of potential regressors in (4), stepwise regression

(SWR) is used in this contribution to successively add or remove parameters based on their statistical significance. The set of potential regressors in (2) is given by $K = \{\varphi_r\}$. The significance is evaluated by an F -test that compares the performance of two models built of different subsets of regressors starting from a constant model $\hat{y}(l) = \theta_0$ and following a greedy stepwise search strategy. In each step, the hypotheses

$$H_0 : \theta_i = 0 \text{ and } H_1 : \theta_i \neq 0 \tag{5}$$

are tested against each other for a confidence level α . When the null hypothesis is rejected for a new model term, the term will be added to the model, otherwise it will be discarded.

2.1.2 Takagi Sugeno models

Locally affine TS models are characterized by a high model flexibility while having an exploitable model structure and a compact parametric representation. They can be used for identifying nonlinear systems. A TS model consists of $c \in \mathbb{N}_+$ superposed local models $\hat{y}_i(l) = f(\theta_{i,\text{LM}}, \tilde{\varphi}(l)) : \mathbb{R}^n \rightarrow \mathbb{R}$ weighted by their corresponding fuzzy basis function $\phi_i(z(l)) : \mathbb{R}^{n_z} \rightarrow [0, 1]$, which depends on the scheduling variable $z(l) = [z_1(l) \dots z_{n_z}(l)]^T \in \mathbb{R}^{n_z}$:

$$\hat{y}(l) = \sum_{i=1}^c \phi_i(z(l)) \cdot \hat{y}_i(l). \tag{6}$$

The local models are of the form

$$\hat{y}_i(l) = \sum_{r=0}^n \theta_{i,r,\text{LM}} \cdot \tilde{\varphi}_r(l), \tag{7}$$

with $\tilde{\varphi}_r(l)$ being the r th element of the vector

$$\tilde{\varphi}(l) = [1 \ x_1(l) \ \dots \ x_n(l)]^T \tag{8}$$

and $\theta_{i,r,\text{LM}}$ being the r th element of the corresponding local model parameter vector $\theta_{i,\text{LM}} \in \mathbb{R}^n$. The use of local models with different structure is also possible, but the locally affine structure enables better interpretability.

The fuzzy basis functions $\phi_i(z(l))$ define the validity region of the corresponding local models. They are defined by

$$\phi_i(z(l)) = \frac{\mu_i(z(l))}{\sum_{j=1}^c \mu_j(z(l))}, \tag{9}$$

with the membership functions $\mu_i(z(l))$. In this contribution, fuzzy-c-means (FCM) type membership functions

$$\mu_j(z) = \left[\sum_{i=1}^c \left(\frac{\|z - v_j\|_2}{\|z - v_i\|_2} \right)^{\frac{2}{v-1}} \right]^{-1} \tag{10}$$

with the fuzziness parameter $v \in \mathbb{R}^{>1}$ are used. The prototypes $v_i \in \mathbb{R}^{n_z}$ are aggregated in the partitioning related parameter vector $\theta_{\text{MF}} = [v_1^T, \dots, v_c^T]^T$. In the remainder of the paper the scheduling variable is assumed to be $z = x$. Using a least squares type cost function, the identification of $\theta_{\text{LM}} = [\theta_{1,\text{LM}}^T \dots \theta_{c,\text{LM}}^T] \in \mathbb{R}^{n \cdot c}$ and θ_{MF} is carried out as follows:

1. θ_{MF} is estimated using FCM clustering.
2. Local model parameters θ_{LM} are estimated with ordinary least squares (OLS).
3. These estimates are used to initialize a simultaneous nonlinear optimization of θ_{MF} and θ_{LM} with a trust region reflective method using the MATLAB function `lsqnonlin`.

An alternative subsequent bounded error estimation method for θ_{LM} is described in Sect. 2.2.

2.1.3 Gaussian process regression

A nonparametric modeling approach is GPR. A detailed explanation and derivation of GPR can be found in [19]. The idea behind a Gaussian Process is to define a probability distribution over functions. For regression, the prior of the function to be learnt is defined as:

$$\hat{y}(l) = \text{GP}(m(x(l)), \kappa(x(l), x'(l))) \tag{11}$$

where $m(x)$ is the mean function and $\kappa(x, x')$ is the covariance function for two points x and x' in the input space (index l is omitted for brevity in the remaining section):

$$m(x) = \mathbb{E}[f(x)] \tag{12}$$

$$\kappa(x, x') = \mathbb{E}[(f(x) - m(x))(f(x') - m(x'))^T] \tag{13}$$

with a positive definite kernel $\kappa(x, x')$. The GP defines a joint Gaussian distribution for a finite set of points:

$$p(y|X) = \mathcal{N}(y|m, K) \tag{14}$$

with the design matrix $X \in \mathbb{R}^{N \times n}$ describing the input data, the covariance matrix $K \in \mathbb{R}^{N \times N}$, $K_{i,j} = \kappa(x(i), x(j))$, and the vector of means $m = [m(x(1)) \dots m(x(N))]^T \in \mathbb{R}^N$. The kernel plays a key role in GPs as it describes the dependencies between the data points. The choice of the kernel strongly influences the model's behavior. In this contribution a Matern 3/2 kernel with automatic relevance determination

(ARD) is used to assign individual length scale parameters to each variable:

$$\kappa_{M32}(\mathbf{x}, \mathbf{x}') = \theta_{f,M32} \left(1 + \sqrt{3}d_{M32}\right) \exp\left(-\sqrt{3}d_{M32}\right) \quad (15)$$

with

$$d_{M32} = \sqrt{\sum_{i=1}^n \frac{(x_i - x'_i)^2}{\theta_{f,SE+ARD,i}^2}} \quad (16)$$

with the individual length scales $\theta_{f,M32+ARD,i}^2$ and standard deviation $\theta_{f,M32}$.

To predict the output y_* for a new test input \mathbf{x}_* , the posterior conditional probability is computed by conditioning the multivariate Gaussian (zero mean and no noise are assumed for simplicity):

$$p(y_* | \mathbf{x}_*, \mathbf{X}, \mathbf{y}) = \mathcal{N}(y_* | m_*, \kappa_*) \quad (17)$$

with the new test input $\mathbf{x}_* \in \mathbb{R}^n$ and the vector of outputs $\mathbf{y} \in \mathbb{R}^N$. Usually, it is assumed that the observed output is affected by additive independent and identically, normally distributed noise i.e.

$$y(l) = f(\mathbf{x}(l)) + e(l), e(l) \sim \mathcal{N}(0, \sigma_y^2) \forall l. \quad (18)$$

The GPR model is trained on the data set by optimization of its hyperparameters, i.e. the kernel function parameters and the noise variance by maximizing its log marginal likelihood function, using a Quasi-Newton optimizer from the MATLAB GPR toolbox.

2.2 Uncertainty modeling

Besides point predictions, regression models can provide for quantification of the uncertainty of the prediction. In this contribution two approaches for uncertainty quantification are considered. Using GPR from Sect. 2.1.3, the probability distribution $p(f|Z^N)$ can be inferred directly. In this probabilistic setting the prediction error is described by a probability distribution. GPR assumes independent, normally distributed noise (18) and the noise variance is treated as a parameter that is estimated in model training.

An alternative error description, which is independent of probabilistic assumptions, is the bounded error (BE) approach. Here, the assumption is made that the prediction error $e(l)$ lies in an interval:

$$e_{\min} \leq e(l) \leq e_{\max}, l = 1, \dots, N. \quad (19)$$

Because BE methods use a set based framework, an exact determination of the BE estimation can become computationally expensive for a large number of parameters because of the "Curse of Dimensionality" [20]. An approximate BE

estimation method that can be used for TS models with a large number of parameters is the Ray Shooting Bounded Error (RSBE) method that is described in [21].

2.3 Model validation

To validate and compare the modeling results, two common evaluation criteria for regression models are used. The root-mean squared error (RMSE):

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{l=1}^N (y(l) - \hat{y}(l))^2} \quad (20)$$

and the coefficient of determination (R^2):

$$R^2 = 1 - \frac{\sum_{l=1}^N (y(l) - \hat{y}(l))^2}{\sum_{l=1}^N (y(l) - \bar{y})^2} \quad (21)$$

with the sample mean \bar{y} are considered.

To assess the generalization capabilities for new data, cross validation (CV) is used. In k -fold cross validation the available data set is randomly partitioned into k subsets $\{T_1, \dots, T_k\}$ of equal size. In each iteration $m \in 1, \dots, l$ the model is trained with $k - 1$ subsets $\{T_1, \dots, T_k\} \setminus T_m$ and validated on the remaining subset T_m . Here $k = 10$ is chosen to adequately evaluate the prediction quality [22]. The k -fold cross validated variants of the metrics (20) and (21) are determined by

$$\text{RMSE}_{\text{CV}} = \frac{1}{k} \sum_{m=1}^k \text{RMSE}(y^m, \hat{y}^m) \quad (22)$$

$$R_{\text{CV}}^2 = \frac{1}{k} \sum_{m=1}^k R^2(y^m, \hat{y}^m) \quad (23)$$

where \hat{y}^m is the prediction for T_m in each fold.

2.4 Experimental methods

Hard-turning experiments were conducted on cylindrical specimens made of a quenched and tempered (Q & T) steel 51CrV4 in different initial hardness levels, i. e. 400 HV30, 500 HV30 and 600 HV30. The focus of the present study will be on the hardness levels of 500 HV30 and 600 HV30. Prior to heat treatment, all specimens were manufactured with the same roughness requirements in order to ensure an almost identical surface finish before hard-turning. Three different sections of nine specimen of each hardness level (i.e. in total 27 section-specific areas per hardness level) were machined with varied cutting parameters feed rate (f), depth of cut (a_p) and cutting speed (v_c). Hard-turning of the specimens was carried out on a servo-conventional lathe of type Weiler C30 using polycrystalline boron nitride (PCBN)

inserts with a corner radius of 0.8 mm. After hard-turning, residual stress (σ_r) depth profiles and surface roughness R_a were determined using X-ray diffraction (XRD) and tactile roughness analysis, respectively for each specific area of a specimen. Residual stress measurements were conducted using a Pulstec μ -X360 diffractometer equipped with a 0.3 mm collimator and CrK α -radiation with an exposure time of 120 sec. Depth profiles were determined by successive removal of the material surface layer using electro-chemical polishing. The obtained data have been evaluated applying the cos α -method [23] without consideration of any mathematical stress correction. Surface roughness in axial direction was determined using a Mitutoyo SJ-210 tactile roughness measuring device. For more detailed information on the material and the experimental setup, i.e., chemical composition, specimen geometry, machines and parameters used for hard-turning operations and post process measurements, the reader is referred to [10].

3 Results for empirical modeling

3.1 Data base

The data base for empirical modeling was generated from the experiments described in Sect. 2.4. A full factorial experiment design was used with two levels for the initial hardness HV_{init} and three factors for the cutting parameters, with the levels given in Table 1. The resulting data set consists of $N = 54$ samples. For the residual stress depth profile modeling, measurements for 12 different depths are available: {0, 10, 20, 30, 40, 50, 60, 80, 100, 120, 150, 200} μm , providing for $N = 648$ samples in total.

3.2 Modeling results for residual stress depth profiles

For MPR modeling, as compromise between a flexible model and the number of potential model terms, a maximum model order of $m = 3$ was chosen. For d_s the data coverage is better, but tests with higher order model terms for d_s did not provide for better results. The threshold for the p -value to add a model term was set to $\alpha < 0.05$ and to remove a model term to $\alpha > 0.1$. The results of the variable selection

Table 1 Levels for experimental design

HV_{init} in HV30	f in mm	v_c in m/min	a_p in mm
500	0.05	100	0.05
600	0.25	175	0.25
	0.5	250	0.5

using SWR are shown in Table 2 in the right section of the table. Besides for the interaction terms, third order model terms were only selected for d_s . The model shows a significant influence of all cutting parameters on σ_r . A significant nonlinear influence of f can be inferred.

For the TS models, preliminary tests were conducted to determine the number of local models c and the fuzziness parameter ν , also see [11] for a systematic selection of these parameters for SL state modeling. Best results were achieved with $\nu = 1.3$ and $c = 3$.

For GPR models, the estimated lengthscale parameters can be used to evaluate the importance of the inputs [24]. Smaller lengthscales correspond to a higher relevance. Table 3 shows the estimated lengthscales for the GPR model, showing high importance for the cutting parameter f .

The performance of the resulting models for depth profiles of σ_r is summarized in Table 4.

The results for the training data (RMSE, R^2) as well as for cross-validation indicate that GPR achieves the best performance. While TS achieves better results on training data for the parametric models, MPR shows better generalization performance. The $R^2_{CV} = 0.94$ for GPR indicates a very good prediction performance of the GPR model. In Fig. 1 exemplary prediction surfaces for the three model approaches together with the training data are shown. The models in the figures are trained with the complete data set. Regions in the data with high gradients are best approximated by GPR. Due to limited model flexibility of the global modeling approach, MPR does not approximate the surface residual stress well. TS models show better approximation than MPR but are worse than the local modeling approach GPR.

3.3 Results for uncertainty modeling

Using the methods described in Sect. 2.2, prediction intervals were estimated for depth profile modeling. In Fig. 2 an exemplary prediction of a residual stress depth profile with the inputs $f = 0.25$ mm, $v_c = 175$ m/min, $a_p = 0.25$ mm, $HV_{init} = 500$ HV30 is shown. The uncertainty interval from the estimated TS model (RSBE) and the 95 % confidence interval for GPR are both shown. Because of the different distribution assumptions for the error, the intervals show different behavior. For the BE estimation the error bounds are not met for some data points because approximate estimation techniques were used. For the BE estimation the δ , i.e., the permissible error, was determined to be ± 200 MPa. To determine an appropriate error bound δ , an iterative search was used, see [13]. For GPR, the estimated σ_y is 37.2 MPa. This corresponds to a 95 % confidence value of $1.96 \cdot \sigma_y$ of 72.9 MPa. In comparison, δ is larger than σ_y , but the BE assumption is much stricter and the model fit of the TS multi models is worse than that for GPR models and

Table 2 Results of regressor term selection and estimated coefficients (non standardized, units dropped for coefficient estimates for better readability)

	Models for characteristic SL values from Sect. 3.4		R_a	Model for depth profile from Sect. 3.2		
	$\sigma_{t,max}$	$\sigma_{t,surf}$		σ_t	(Additional terms)	
(Intercept)	733.00	1917.76*	-3.81***	1637.15***	$HV_{init}:v_c:f$	-0.02***
HV_{init}	-0.72	-3.57*	0.01***	-2.81***	$v_c:f:a_p$	8.15***
v_c	-6.54*	-15.88**	0.02***	-6.70***	$v_c:a_p^2$	-10.66**
f	7385.94***	-5012.71	31.96***	3633.41*	d_s	-15.62***
a_p	-2503.98	1680.68**		-2705.57**	$HV_{init}:d_s$	0.01***
$HV_{init}:v_c$	0.01	0.02**		0.01***	$v_c:d_s$	-0.01
v_c^2	0.02**	0.01			$f:d_s$	12.64*
$HV_{init}:f$	-16.02***	14.94**	-0.06***	-5.27	$a_p:d_s$	-7.79***
$v_c:f$	1.88	75.37***	-0.18***	6.93	d_s^2	0.14***
f^2	-19052.72***	-13509.44***	18.32***	-12843.13***	$HV_{init}:f:d_s$	-0.03***
$HV_{init}:a_p$	5.55*			4.57**	$v_c:f:d_s$	-0.02*
$v_c:a_p$	19.36**			18.02***	$v_c:d_s^2$	0.01**
$f:a_p$	-5803.20***	-5036.04**		-6640.07***	$f:d_s^2$	0.06***
a_p^2	-1099.44	-2574.87*		1092.76	$a_p:d_s^2$	0.04***
$v_c^2:f$	-0.04**	-0.08**			d_s^3	-0.01***
$HV_{init}:f^2$	32.74***			-5.27		
$v_c:f^2$	26.91***	36.10***	-0.09***	17.95***		
$HV_{init}:v_c:a_p$	-0.03*			-0.02**		
$f:a_p^2$	7999.75***	7093.74*		-6640.07***		

Additional selected model terms for depth profile are recorded in the two right hand columns
p-value: **p* < 0.10; ***p* < 0.05; ****p* < 0.01

Table 3 Estimated length scale parameters for GPR (standardized for comparison)

$\theta_{HV_{init}}$	θ_f	θ_{v_c}	θ_{a_p}	θ_{d_s}
6.19	2.19	4.54	4.54	1.10

Table 4 Prediction performance of models for σ_t depth profiles

Output	Model	RMSE (MPa)	RMSE _{CV} (MPa)	R^2	R^2_{CV}
σ_t	MPR	150.9	160.2	0.73	0.69
	TS	116.0	169.3	0.84	0.65
	GPR	16.5	72.3	0.99	0.94

The best performing models are highlighted in bold

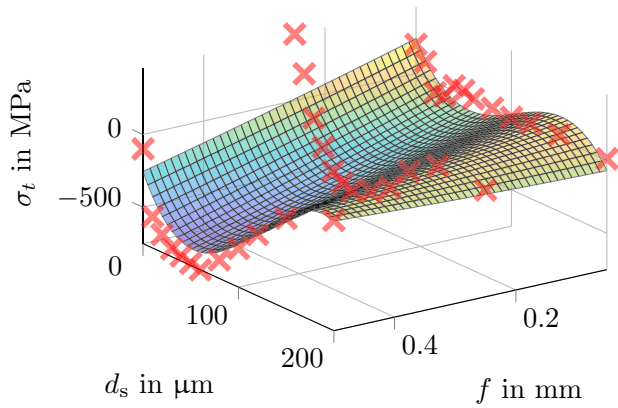
therefore, has a larger model error. However, a Shapiro-Wilk test on the residuals of the GPR model shows that assumption of normally distributed noise can not be accepted at 5% significance level. The BE approach on the other hand only makes the assumption that the error is bounded.

3.4 Modeling results for characteristic SL values

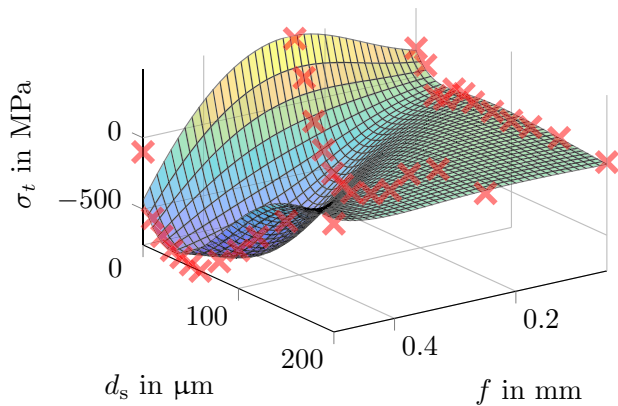
To obtain models that enable useful model-based analysis with respect to the cutting parameters, characteristic values of the depth profile data were extracted: (1) tangential

residual stress at the surface $\sigma_{t,surf}$ and (2) maximum tangential compressive residual stress $\sigma_{t,max}$. High maximum compressive residual stress in the depth and high compressive residual stress at the surface are beneficial for the components fatigue life [25]. Furthermore surface roughness R_a was modeled. The resulting performance criteria are summarized in Table 5.

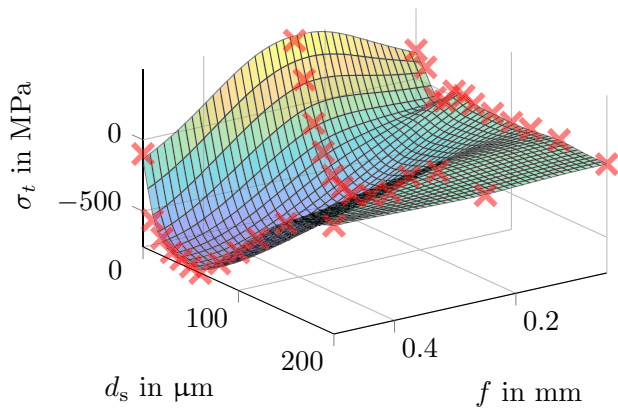
For all models GPR shows the best cross validated performance, whereby very good results were achieved for R_a with all model approaches, presumably because of less noise in the data and less nonlinear behavior. The performance on training data is better for TS models than for MPR but for cross validation this observation does not hold true, indicating overfitting of the TS models. For MPR again monomials and interaction terms up to the order of $m = 3$ were considered to capture the nonlinearities. Results for variable selection and parameter estimation are given in Table 2. For TS models analogous to Sect. 3.2 preliminary tests led to a choice of $c = 3$ and $v = 1.3$. The estimated length scale parameters in Table 6 give information about the importance of the input variables. For $\sigma_{t,max}$, $\sigma_{t,surf}$ and R_a the feed rate f has the lowest value, i.e., the highest importance. This can be explained based on the fact that the feed rate strongly influences the specific cutting pressure [26], thus resulting in different



(a) Multiple polynomial regression (MPR)



(b) Takagi Sugeno model (TS)



(c) Gaussian process regression (GPR)

Fig. 1 Prediction surfaces (surface) for three different model classes and measured data (x) for σ_t depth profiles depending on f (for constant $HV_{init} = 500$ HV, $v_c = 175$ m/min, $a_p = 0.25$ mm)

amounts of plastic deformation in front of the cutting edge. As the surface roughness and the residual stress state are known to have a strong correlation with the plastic deformation, the observed correlation can be rationalized. This is further in line with literature [27], as e.g., for

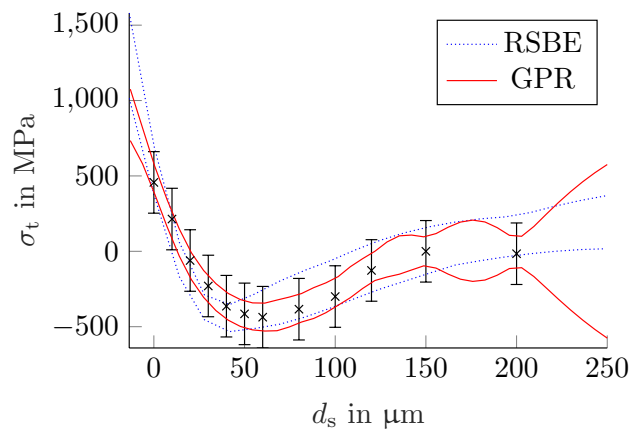


Fig. 2 Prediction error bounds for RSBE model (dotted) and 95% confidence interval for GPR model (solid) for σ_t depth profile models. Measured residual stresses (x) and errors bounds (bars) are shown

Table 5 Prediction performance of models of characteristic SL values

Output	Model	RMSE	RMSE _{CV}	R ²	R ² _{CV}
$\sigma_{t,max}$	MPR	95.3 MPa	121.2 MPa	0.873	0.83
	TS	58.4 MPa	119.0 MPa	0.95	0.85
	GPR	10.5 MPa	85.9 MPa	0.99	0.90
$\sigma_{t,surf}$	MPR	148.4 MPa	180.8 MPa	0.80	0.67
	TS	84.5 MPa	181.2 MPa	0.93	0.64
	GPR	74.7 MPa	144.3 MPa	0.95	0.82
R_a	MPR	0.14 μ m	0.19 μ m	0.97	0.94
	TS	0.05 μ m	0.16 μ m	0.99	0.96
	GPR	0.04 μm	0.11 μm	0.99	0.98

The best performing models are highlighted in bold

Table 6 Estimated length scale parameters for GPR (standardized for comparison)

	$\theta_{HV_{init}}$	θ_f	θ_{v_c}	θ_{a_p}
$\sigma_{t,max}$	7.21	1.61	4.26	4.04
$\sigma_{t,surf}$	6.15	0.42	1.45	8.39
R_a	3.48	1.46	4.48	51.10

hard-turning experiments in [28] the feed rate was found to be the dominant parameter for the shape of the residual stress depth distribution. It can therefore be concluded, that the results of the model used in the present study coincides with the experimental findings.

3.5 Discussion of prediction modeling results

In Table 7 the results for previous case studies of residual stress depth profiles using the hard-turning data set, e.g., described in [12] are shown. The experimental conditions, type of tool and type of steel were the same but a different

Table 7 Summarized prediction performance of previous depth profile modeling case studies [11–13]

	MPR		TS		GPR
	σ_t	σ_a	σ_t	σ_a	σ_t
R^2_{cv}	0.67	0.68	0.81	0.87	0.93

lathe was used. In addition, a different design of experiments was used. For sake of brevity, the reader is referred to [12] for details on the modeling process. The additional SL state axial residual stress σ_a was modeled in these case studies (not for GPR). These case studies confirm the results from this contribution in Sect. 3.2. GPR overall offers a better prediction performance. However, the parametric TS and MPR models offer a more compact model structure. For prediction, GPR inversion of the covariance matrix is needed, while TS and MPR only require algebraic operations. Furthermore, the local-affine structure of TS models can be exploited for control design based on established approaches for linear systems. Depending on the requirements of the application, it must therefore be decided which model properties are preferred. For example, if the models are to be used in situ for feedforward control methods, the computation time is of great importance in order to ensure fast model evaluation, especially, if the models are adapted to changing process conditions.

Based on the results the GPR models are superior in terms of prediction performance. Hence, for a model based analysis and parameter optimization of the hard turning process, the GPR models are used in the following.

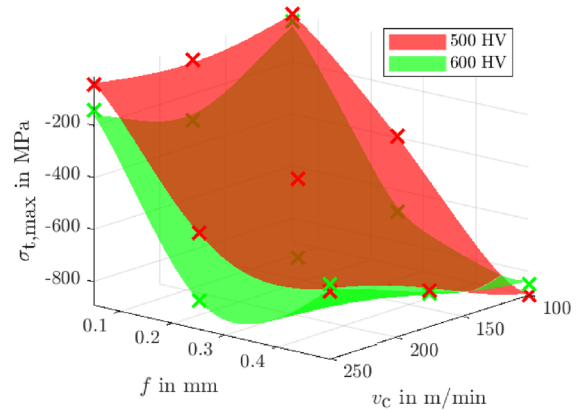
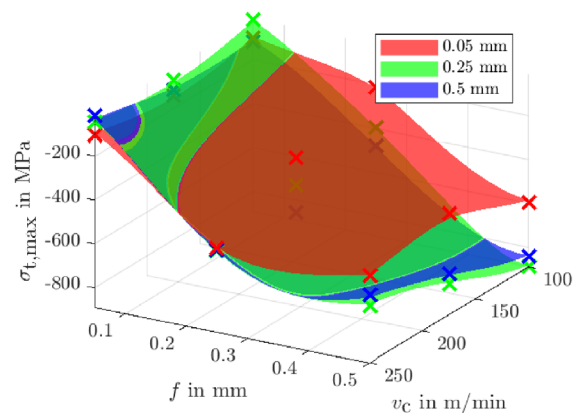
4 Model application

4.1 Model based process analysis

In this section, the identified prediction models are used to perform a process analysis of the correlations between cutting parameters and SL state. The SL properties residual stress and surface roughness show high correlation with machining parameters. For the analysis, the models for $\sigma_{t,max}$, $\sigma_{t,surf}$ and R_a are used. In Figs. 3, 4 and 5 prediction plots for the models from Sect. 3.4 are shown. Thereby, Fig. 3 focuses on the prediction of $\sigma_{t,max}$ for constant a_p in (a) as well as for constant HV_{init} in (b). (For the sake of brevity, in (b) only the case for $HV_{init} = 500$ HV30 is shown.)

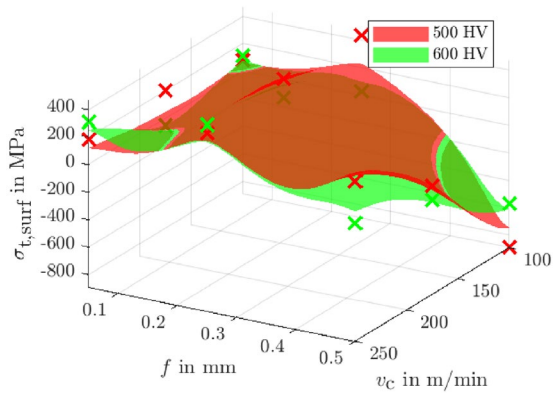
As can be seen from the illustration in (a) and as already explained in 3.4 the highest influence on the resulting $\sigma_{t,max}$ can be obtained for varying the feed rate f .

It can further be concluded that a higher HV_{init} provides a higher potential for compressive residual stresses as a consequence of the increased strength of this condition.

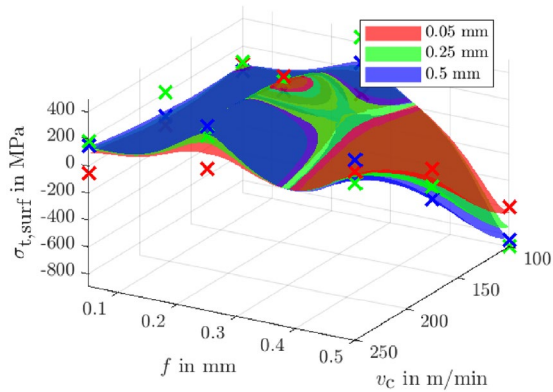
(a) For constant $a_p = 0.25$ mm and varied HV_{init} (b) For constant $HV_{init} = 500$ HV30 and varied a_p **Fig. 3** Prediction surfaces of GPR and measured data (x) for $\sigma_{t,max}$

This is in line with literature as higher compressive residual stresses after hard-turning of specimen with higher HV_{init} were likewise observed in [28]. Figure 3a moreover reveals that both initial hardness levels are characterized with $\sigma_{t,max}$ for use of nearly the same cutting parameter set including a feed rate and cutting speed of 0.3 mm^{-1} and 250 m/min , respectively. With increasing feed rates and decreasing cutting speeds the maximum stresses $\sigma_{t,max}$ converge to similar levels. However, it has to be noted that despite comparable $\sigma_{t,max}$ values for both considered hardness levels, these do not necessarily have to be present in the same depth as has been shown in [10].

Comparing the maximum compressive residual stresses $\sigma_{t,max}$ for one initial hardness level and varying cutting parameters as displayed in Fig. 3 (b), again the feed rate has the highest influence. Furthermore, for the lowest cutting depth as well as generally for decreasing the cutting speed, the maximum compressive residual stresses are decreasing due to the fact that the hard-turning process in this parameter range is dominated by plastic deformation in front of the



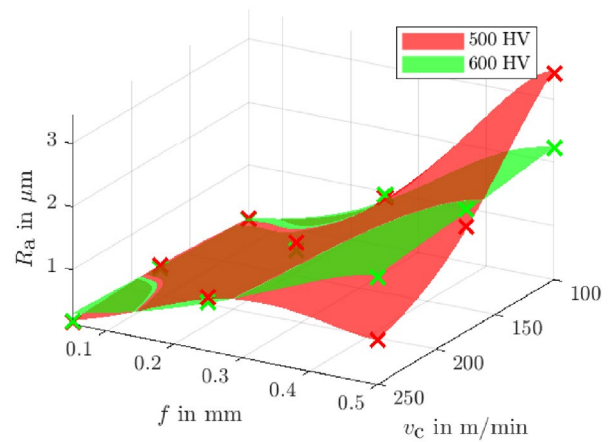
(a) For constant $a_p = 0.25$ mm and varied HV_{init}



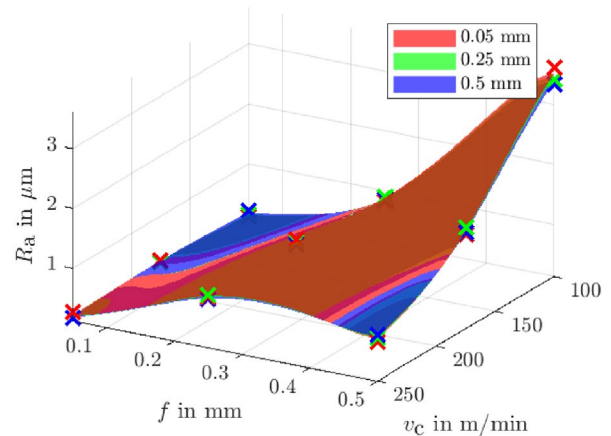
(b) For constant $HV_{init} = 500$ HV30 and varied a_p

Fig. 4 Prediction surfaces of GPR and measured data (x) for $\sigma_{t,surf}$

cutting edge rather than by cutting. In contrast, Fig. 4 illustrates the prediction of the surface residual stresses $\sigma_{t,surf}$. From Fig. 4 (a) it is clearly evident that unlike to Fig. 3 similar residual stress surface ($\sigma_{t,surf}$) values are obtained irrespective of the initial hardness. This observation is assumed to be a result of higher surface temperatures thus decreasing the influence of the initial hardness. Moreover, the proportion of the Hertzian pressure at the surface is not as pronounced. Differences in the surface residual stresses $\sigma_{t,surf}$ can only be obtained for low cutting speeds and high feed rates as the process temperatures are lower and increasing plastic deformation can be obtained for these cutting parameter values. Analyzing the surface residual stresses for one initial hardness level as displayed in Fig. 4b, the prediction surfaces of the used GPR and the experimental data show no pronounced differences for the range of the cutting parameters considered. In line with the maximum residual stresses in Fig. 3b, a trend for a decreasing absolute value of the surface residual stresses $\sigma_{t,surf}$ for very low cutting depth and decreasing cutting speeds can be observed.



(a) For constant $a_p = 0.25$ mm and varied HV_{init}



(b) For constant $HV_{init} = 500$ HV30 and varied a_p

Fig. 5 Prediction surfaces of GPR and measured data (x) for R_a

However, this trend is not as pronounced for the surface residual stresses.

Figure 5 focuses on the prediction surfaces of GPR for the surface roughness R_a . In contrast to the results displayed in Figs. 3 and 4, a combination of high feed rates and low cutting speeds does not result in similar roughness values for both initial hardness levels. From the prediction and the measured data shown in Fig. 5a it becomes obvious that the higher initial hardness level is more or less independent of the used cutting speed. Unfortunately for both initial hardness levels, the cutting parameter sets resulting in advantageous residual stress states (both maximum and surface) do lead to high surface roughness values. Furthermore, the variation of the cutting depth does not seem to have any influence on this fact as demonstrated by Fig. 5b. In line with the results of the parameter optimization in Sect. , the predictions

of the models in Figs. 3, 4 and 5 suggest that the models allow an exploration of the cutting parameter influences.

4.2 Model based cutting parameter optimization

To further demonstrate the usefulness and applicability of the models, multi objective optimization using the MATLAB optimization Toolbox with a variant of the NSGA-II genetic algorithm [29] was used to find optimal values for the cutting parameters. The optimization problem can be formulated as

$$\begin{aligned}
 & \min_{\mathbf{x}=[f, a_p, v_c]} \hat{y}_{\sigma_{t,max}}(\mathbf{x}), \hat{y}_{\sigma_{t,surf}}(\mathbf{x}), \hat{y}_{R_a}(\mathbf{x}) \\
 & \text{s.t. } 0.05 \text{ mm} < f < 0.5 \text{ mm} \\
 & \quad 0.05 \text{ mm} < a_p < 0.5 \text{ mm} \\
 & \quad 100 \text{ mm/min} < v_c < 250 \text{ m/min},
 \end{aligned} \tag{24}$$

with $\hat{y}_{\sigma_{t,max}}$, $\hat{y}_{\sigma_{t,surf}}$, \hat{y}_{R_a} the predictions of the GPR models for the respective SL state and the cutting parameters $\mathbf{x} = [f, a_p, v_c]$ as arguments, i.e., a minimization problem with 3 objective functions/target variables and 3 decision variables.

This multi objective optimization problem does not have a single best solution. The result of the optimization is the Pareto front, see Fig. 6 exemplary for $HV_{init} = 500$ HV30 and $R_{a,max} = 0.8 \mu\text{m}$. The surface roughness of a component is majorly determined by the design specification and a high roughness should be avoided. Therefore, the desired

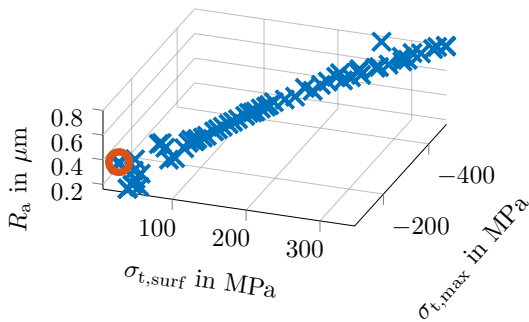


Fig. 6 Determined Pareto front by genetic algorithm for $HV_{init} = 500$ HV30 and $R_{a,max} = 0.8 \mu\text{m}$. Selected optimal point regarding the SL state marked by circle (o)

maximum value $R_{a,max}$ was considered by penalizing values $R_a > R_{a,max}$ during optimization.

As shown in Figs. 4 and 5, achieving compressive surface residual stress and low surface roughness are competing goals. While $\sigma_{t,max}$ is always in the compressive range, see Fig. 6, it is more critical to achieve a high level of compressive residual stress at the surface $\sigma_{t,surf}$. Therefore, from the solution set the solution with the minimal $\sigma_{t,surf}$ value was selected. Optimization was performed for both initial hardness levels. Results for the optimization are summarized in Table 8. The selected optimal cutting parameter values are given together with the resulting predicted SL state, taking into account the roughness specification. The maximal allowed surface roughness was exemplarily defined as $0.8 \mu\text{m}$ and $1.8 \mu\text{m}$, respectively. For 500 HV30 and $R_{a,max} = 0.8 \mu\text{m}$ a compressive residual stress is not achieved. Though for all other conditions high maximal residual stress and compressive surface residual stress can be achieved. As a consequence it can be concluded that depending on a possible application of a component and therefore on the desired SL state for a given initial hardness level, the model based optimization scheme can be used to obtain optimal cutting parameter values. However, it has to be noted that in order to analyze the accuracy of the results, further experimental evidence by hard-turning and post process determination of the SL state would be necessary and beneficial. Nevertheless, assuming a good prediction accuracy, as a next step an investigation on whether the methods presented in this study can be transferred to other machining finishing processes such as grinding, deep-hole drilling or milling could be conducted.

For the exemplarily selected roughness ranges the results of the cutting parameter optimization further demonstrate that comparatively high maximum as well as surface compressive residual stresses during hard-turning of the 51CrV4 in the cutting parameter range considered can only be obtained at the expense of surface roughness or vice versa.

If the load scenario of a component in use, i.e., tribological, mechanical or corrosive stresses is sufficiently known, from a manufacturer’s point of view the model presented could be used to chose appropriate cutting parameter values eventually resulting in increased service life and reliability. Depending on the load situation actually a low surface roughness, high residual compressive stress state or

Table 8 Results for cutting parameter optimization (not standardized)

$R_{a,max}$ in μm	HV_{init} in HV30	Cutting parameters			Surface layer state		
		f in mm	v_c in mm/min	a_p in mm	$\sigma_{t,max}$ in MPa	$\sigma_{t,surf}$ in MPa	R_a in μm
0.8	500	0.05	160.03	0.05	-147.53	3.91	0.23
	600	0.05	171.24	0.05	-346.23	-89.65	0.17
1.8	500	0.05	174.03	0.38	-772.97	-151.62	1.77
	600	0.39	100	0.5	-821.74	-157.20	1.77

a compromise solution may be most advantageous. Here, experimental evidence on the influence of different surface states for a given loading scenario, e.g., mechanical fatigue loading, could be of high interest and should therefore be investigated in future studies.

5 Conclusions and outlook

In this work, empirical modeling of the SL state depending on cutting parameters for hard-turning of 51CrV4 is investigated. Three data-driven modeling approaches for small data sets were compared and their generalization performance was evaluated. The comparison has shown that GPR overall achieves the best prediction performance. However, for this nonparametric method, the whole data set has to be used for inference. Parametric TS models and MPR models offer a more compact process description. Furthermore, TS models, in combination with the BE method, allow for a uncertainty estimation and provide for a compact model structure. The MPR models allow for easier interpretation of the regressor parameters and therefore, for better interpretability.

Data driven models can be used for cutting parameter optimization. An optimization scheme was developed to optimize the competing SL properties residual stress and surface roughness. Depending on the actual application and therefore, the loading scenario of a component, a cutting parameter set can be computed that leads to the desired SL state values.

In future work, experimental validation of the optimized cutting parameters and of the influence of the SL states on the fatigue life will be examined. An additional experimental validation of the models outside the specified cutting parameter limits is also planned in order to explore and identify new process windows. Furthermore, it is planned to use the models in a process control scheme with the goal of improving the obtained surface state after machining. Because SL properties such as residual stresses cannot be measured during the cutting process, standard feedback control methods cannot be used. Therefore, an approach combining feedforward control using the static models developed in this work and a workpiece-to-workpiece feedback control adapting the cutting parameters based on comparing pre- and post-process micromagnetic in situ measurements will be subject of future research.

In addition, the model approaches are to be applied to other manufacturing processes such as grinding, deep-hole drilling or milling and to different materials.

Acknowledgements The scientific work has been supported by the DFG within the research priority program SPP 2086 (Grant Number: KR 3795/8-1; NI 1327/22-1; ZI 1296/2-1). The authors thank the DFG for this funding and intensive technical support. Moreover, the authors

would like to thank Mr. R. Hunke, Mr. T. Sorge and Mr. C. Schott for technical and experimental support as well as for data analysis.

Funding Open Access funding enabled and organized by Projekt DEAL.

Data availability The data used in this study are available upon reasonable request from the authors.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

1. Grzesik W, Kruszynski B, Ruszaj A (2010) Surface integrity of machined surfaces. In: surface integrity in machining. Springer London, pp 143–179. https://doi.org/10.1007/978-1-84882-874-2_5
2. Stampfer B, González G, Gerstenmeyer M et al (2021) The present state of surface conditioning in cutting and grinding. *J Manuf Mater Process* 5(3):92. <https://doi.org/10.3390/jmmp5030092>
3. Schulze V (2020) Surface conditioning in machining processes. *tm - Technisches Messen* 87(12):743–744. <https://doi.org/10.1515/teme-2020-0071>
4. Uebel J, Ströer F, Basten S et al (2019) Approach for the observation of surface conditions in-process by soft sensors during cryogenic hard turning. *Procedia CIRP* 81:1260–1265. <https://doi.org/10.1016/j.procir.2019.03.304>
5. Denkena B, Breidenstein B, Dittrich MA et al (2021) Effects on the deformation-induced martensitic transformation in AISI 304 in external longitudinal turning. *Adv Ind Manuf Eng* 2:100044. <https://doi.org/10.1016/j.aime.2021.100044>
6. Böttger D, Stampfer B, Gauder D et al (2020) Concept for soft sensor structure for turning processes of AISI4140. *tm - Technisches Messen* 87(12):745–756. <https://doi.org/10.1515/teme-2020-0054>
7. Schmidt R, Strodtick S, Walther F et al (2020) Analysis of the functional properties in the bore sub-surface zone during BTA deep-hole drilling. *Procedia CIRP* 88:318–323. <https://doi.org/10.1016/j.procir.2020.05.055>
8. Heinzel J, Jedamski R, Epp J et al (2021) In-process measurement of Barkhausen noise and resulting productivity increase potential in grinding of case hardened steel. *CIRP J Manuf Sci Technol* 32:37–45. <https://doi.org/10.1016/j.cirpj.2020.11.011>
9. Wimmer M, Hameed MZS, Wölflé C et al (2020) The influence of the process parameters on the surface integrity during peripheral milling of ti-6al-4v. *tm - Technisches Messen* 87(11):721–731. <https://doi.org/10.1515/teme-2020-0052>

10. Wegener T, Liehr A, Bolender A et al (2022) Calibration and validation of micromagnetic data for non-destructive analysis of near-surface properties after hard turning. *HTM J Heat Treat Mater* 77(2):156–172. <https://doi.org/10.1515/htm-2021-0023>
11. Wittich F, Kahl M, Kroll A et al (2019) On nonlinear empirical modeling of residual stress profiles in hard turning. In: 2019 IEEE international conference on systems, man and cybernetics (SMC). IEEE. <https://doi.org/10.1109/smc.2019.8914272>
12. Schott C, Wittich F, Kroll A et al (2021) Prediction of near surface residual stress states for hard turned specimens using data driven nonlinear models. *Procedia CIRP* 101:1–4. <https://doi.org/10.1016/j.procir.2020.10.002>
13. Wittich F, Kistner L, Kroll A et al (2020) On data-driven nonlinear uncertainty modeling: methods and application for control-oriented surface condition prediction in hard turning. *tm - Technisches Messen* 87(11):732–741. <https://doi.org/10.1515/teme-2020-0057>
14. Arrazola P, Özel T, Umbrello D et al (2013) Recent advances in modelling of metal machining processes. *CIRP Ann* 62(2):695–718. <https://doi.org/10.1016/j.cirp.2013.05.006>
15. Ulutan D, Özel T (2011) Machining induced surface integrity in titanium and nickel alloys: a review. *Int J Mach Tools Manuf* 51(3):250–280. <https://doi.org/10.1016/j.ijmachtools.2010.11.003>
16. Shihab S, Khan ZA, Mohammd A et al (2014) Optimization of surface integrity in dry hard turning using RSM. *Sadhana* 39(5):1035–1053. <https://doi.org/10.1007/s12046-014-0263-4>
17. Zhang J, Liang S, Zhang G et al (2006) Modeling of residual stress profile in finish hard turning. *Mater Manuf Processes* 21(1):39–45. <https://doi.org/10.1081/amp-200060608>
18. Umbrello D, Ambrogio G, Filice L et al (2007) An ANN approach for predicting subsurface residual stresses and the desired cutting conditions during hard turning. *J Mater Process Technol* 189(1–3):143–152. <https://doi.org/10.1016/j.jmatprotec.2007.01.016>
19. Rasmussen CE, Williams CKI (2005) Gaussian processes for machine learning. The MIT Press, Cambridge. <https://doi.org/10.7551/mitpress/3206.001.0001>
20. Wittich F, Kroll A (2021) Evaluation of methods for feasible parameter set estimation of Takagi–Sugeno models for nonlinear regression with bounded errors. *at - Automatisierungstechnik* 69(10):836–847. <https://doi.org/10.1515/auto-2020-0157>
21. Wittich F, Kroll A (2022) Approximation of the feasible parameter set in bounded-error parameter estimation of Takagi–Sugeno fuzzy models for large problems by using a ray shooting method. In: 2022 IEEE international conference on fuzzy systems (FUZZ-IEEE). IEEE. <https://doi.org/10.1109/fuzz-ieee55066.2022.9882729>
22. Kohavi R et al (1995) A study of cross-validation and bootstrap for accuracy estimation and model selection. In: Proceedings of the 14th international joint conference on Artificial intelligence, Montreal, Canada, pp 1137–1145
23. Tanaka K (2019) The $\cos\alpha$ method for X-ray residual stress measurement using two-dimensional detector. *Mech Eng Rev* 6(1):18–00378. <https://doi.org/10.1299/mer.18-00378>
24. Liu K, Li Y, Hu X et al (2020) Gaussian process regression with automatic relevance determination kernel for calendar aging prediction of lithium-ion batteries. *IEEE Trans Ind Inf* 16(6):3767–3777. <https://doi.org/10.1109/tii.2019.2941747>
25. Dogra M, Sharma VS, Sachdeva A et al (2012) Surface integrity a key issue in hard turning—a review. *Int J Mach Mach Mater* 12(1/2):88. <https://doi.org/10.1504/ijmmm.2012.048560>
26. Bhemuni V, Rao S, Vinay P (2014) Effect of machining parameters on tool wear in hard turning of AISI d3 steel. *Procedia Eng* 97:338–345. <https://doi.org/10.1016/j.proeng.2014.12.257>
27. Dahlman P (2004) The influence of rake angle, cutting feed and cutting depth on residual stresses in hard turning. *J Mater Process Technol*. <https://doi.org/10.1016/j.matprotec.2003.12.014>
28. Heim HP, Biermann D, Homberg W (2013) Functionally graded materials in industrial mass production, vol 2. Verlag Wiss, Scripten
29. Deb K, Kalyanmoy D (2001) Multi-objective optimization using evolutionary algorithms. Wiley, Hoboken

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.