



Smart sheet metal forming: importance of data acquisition, preprocessing and transformation on the performance of a multiclass support vector machine for predicting wear states during blanking

Christian Kubik¹ · Sebastian Michael Knauer¹ · Peter Groche¹

Received: 7 December 2020 / Accepted: 15 May 2021 / Published online: 4 June 2021
© Springer Science+Business Media, LLC, part of Springer Nature 2021

Abstract

In consequence of high cost pressure and the progressive globalization of markets, blanking, which represents the most economical process in the value chain of manufacturing companies, is particularly dependent on reducing machine downtimes and increasing the degree of utilization. For this purpose, it is necessary to be able to make a real-time prediction about the current and future process conditions even at high production rates. Therefore, this study investigates the influence of data acquisition, preprocessing and transformation on the performance of a multiclass support vector machine to classify abrasive wear states during blanking based on force signals. The performance of the model was quantitatively evaluated based on the model accuracy and the separability of the classes. As a result, it was shown, that the deviation of time series represents the key parameter for the resulting performance of the classification model and strongly depends on the sensor type and position, the preprocessing procedure as well as the feature extraction and selection. Furthermore, it is shown that the consideration of domain knowledge in the phases of data acquisition, preprocessing and transformation improves the performance of the classification model and is essential to successfully implement AI projects. Summarizing the findings of this study, trustworthy data sets play a crucial role for implementing an automated process monitoring as a basis for resilient manufacturing systems.

Keywords Resilient manufacturing · Machine learning in blanking · Data driven process optimization

Introduction

Forming processes often represent the most economical part of the value-added chain and represent a key factor in times of energy and raw material shortages due to the optimal utilization of materials as well as the lower specific energy requirement compared to subtractive or additive manufacturing processes (Lange, 1985). Even though forming processes offer significant advantages, they are under high cost pressure. This is mainly caused by the low margins per part, which are directly influenced by short-term machine downtimes. Blanking processes are especially affected by this problem, as their high stroke speeds and required accuracy

can lead to unplanned downtimes of tool and machine components (Lange, 1985). In order to achieve the objective of high productivity under these circumstances, a reduction of machine downtimes and thus a maximization of the degree of utilization is to be aimed at. For this purpose, it is necessary to be able to make a real-time prediction about the current and future process conditions even at high production rates. Especially in the case of blanking, condition monitoring based on sensorial captured time series is difficult due to the complexity of the process. Blanking is influenced by more than forty parameters related to process-specific uncertainties (wear, varying specifications of semi-finished product, etc.) as well as machine- and tool-specific variables (clearance, cutting edge radii, etc.) (Hirsch et al. 2011). Furthermore, considering data availability in the context of progressive digitalization of manufacturing processes, resilient process control and a comprehensive condition monitoring by qualified personnel or conventional monitoring systems is difficult to achieve. In order to be able to make real-time predictions about the process and thus avoid unplanned

✉ Christian Kubik
kubik@ptu.tu.darmstadt.de

¹ Institute for Production Engineering and Forming Machines,
Technische Universität Darmstadt, Otto-Berndt-Straße 2,
64287 Darmstadt, Germany

machine downtimes, plan maintenance and identify deviations of product quality at an early stage, it is necessary to set up data-driven process monitoring. Especially in times of digitization of the value-added chain, a lack of qualified personnel, reduced physical contact with the machine caused by an ongoing automation and a required remote access, trustworthy data sets play a crucial role to guarantee a resilient model to describe or predict the state of process. Industry 4.0 (also known as Smart Manufacturing or Smart Factory) provides an approach to this, combining elements of artificial intelligence (AI), new types of sensors as well as state-of-the-art information technologies (Oztemel & Gursev, 2020). Central to this approach are data which are more and more available to companies due to high-performance processors and sensors and cross-linking of processes (Moyné & Iskandar, 2017). Industry 4.0 aims to use this data available over the entire product life cycle to improve the value chain and build manufacturing intelligence (O'Donovan et al. 2015). This steadily increasing amount of data offers the possibility of making predictions about process correlations that were not possible before. However, the value of this data is not only determined by its quantity, but also by the knowledge and information hidden in it. AI and the use of machine learning (ML) algorithms, as a key technology in Industry 4.0, offer exactly this knowledge for manufacturing process (Penumuru et al. 2020). ML algorithms provide pattern recognition and the identification of correlations to large datasets and the derivation of predictive models (Wuest et al. 2016). Combining high-performance computing with valid data sets, ML algorithms build a basis for data-driven process monitoring and thus the real-time prediction of process conditions (Tsai et al. 2015). While current literature focuses on the transfer of new ML algorithms to classify or predict the process condition, little work has been done so far to draw conclusions about the influence of the quality of the data set and related to this the influence of data acquisition, preprocessing and transformation on the performance of ML models. Especially in the blanking process the data set should be considered when using AI-supported condition monitoring. This is caused by the nonlinearity of the process as a result of transient elastic and plastic deformation of

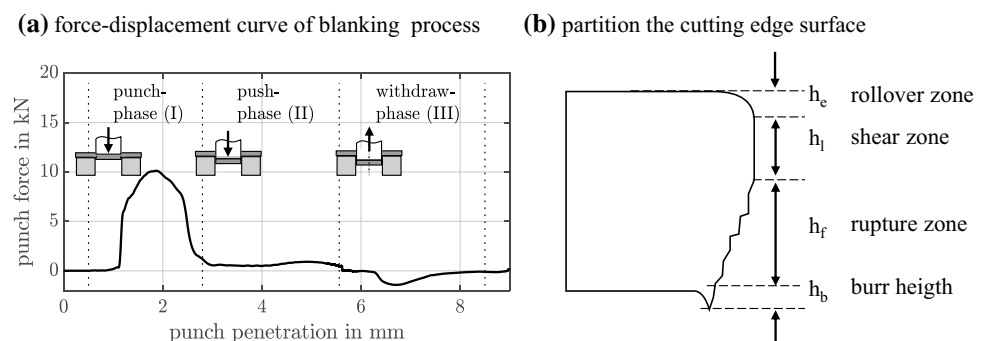
the sheet metal as well as the static and dynamic behavior of the press. Furthermore, over forty process variables can be identified that interact with one another and are related to uncertainties of the process (Jin & Shi, 2000). Even the smallest changes in these process variables can significantly influence the process. In contrast, small variations of process variables influence the time signals recorded by sensors less strongly. This leads to high demands on the quality of the acquired data even at high stroke rates in order to be able to physically differentiate effects in the process from noise even at minor variations (Groche et al. 2019).

Therefore, this work aims at describe the influence of data acquisition, preprocessing and transformation step on the performance of a multiclass support vector machine (mSVM) model. As a use case, the abrasive wear during blanking process is to be classified. The influence of these three steps on the performance of the mSVM is quantified by the accuracy and the separability between classes, related to the Mahalanobis distance. In order to transfer these findings to industrial blanking processes the resilience of the model to predict the current wear state is validated by industry-related data sets at different stroke speeds.

Data-driven monitoring of blanking processes

As shown in Fig. 1, a blanking process can be divided into three phases according to the force–displacement curve. In the punch-phase (I), the punch impacts on the sheet metal and starts to elastically deform the material. If the stresses that occur exceed the maximum shear strength of the material, it tends to deform plastically. When the shearing stress finally exceeds the shear fracture limit, the material tears and the stored elastic energy is released abruptly. In the following push-phase (II), the component or the grid-shaped discard is completely pushed out of the die and the punch passes through the bottom dead center. Finally, the punch is pulled out of the die in the withdraw-phase (III) and withdrawal forces occur as a result of jamming between the sheet and the punch.

Fig. 1 Force–displacement curve (Hoppe et al. 2019) (a) and partition of the cutting edge surface during blanking of sheet metals into characteristic areas according to the VDI 2906-2 (b)



The quality of blanked parts can be described by the characteristic areas of the cutting edge surface (see Fig. 1b). According to VDI 2906-2, these features are defined as a percentage of the blanked surface orthogonal to the cutting line and are divided into rollover zone h_e , shear zone h_s , rupture zone h_r as well as the burr height h_b . They are directly influenced by the wear state and reflect the current tool condition. Adhesive and abrasive wear occurs in the contact zone between punch, die and sheet metal and is classified into three types of wear, flank wear, face wear and tip wear (Hambli, 2001; Hernández et al. 2006). According to Hernández et al. flank wear especially occurs during the push- and withdraw-phase of a blanking cycle (Fig. 1 phase II and III). Due to the longer wear length, normal stresses between the lateral surface of punch, die and sheet metal lead to an increased flank wear, especially for large sheet thicknesses. Cutting soft steel grades ($R_m < 350$ MPa) reinforces this effect (Hohmann et al. 2017). For example, Hohmann et al. showed in their work that a cold-rolled steel (DC03) tends to adhesive flank wear. This leads to higher frictional forces between punch and sheet metal which is directly influencing the push- and withdraw-phase force signal of a blanking process. In contrast, high-strength steel grades ($R_m > 600$ MPa) tend to abrasive tip wear which leads to a rounded cutting edge radius (Klingenberg & Singh, 2004). In contrary, face wear occurs especially when blanking thin sheets, which is due to high resulting surface pressures between the face surface of the punch and the sheet metal. In addition, temperature peaks in the contact zone lead to local melting and microstructural changes of the punch material which enhances micro-machining processes (Toussaint, 2000). The combination of these three wear states finally changes the geometry of the punch and die, as shown in Fig. 2. In this work we focus on the rounding of the cutting edge radii by abrasive wear. Hambli, Klingenberg and Singh as well as Kubik et al. showed in their work that exactly this abrasive wear is the most common cause of wear in blanking and has a significant influence on the quality of the component (Hambli, 2001; Klingenberg & Singh, 2004; Kubik et al. 2021).

Over the last decade, many experimental and empirical studies have been conducted to understand wear in the blanking process. On the one hand, they investigated the influence of blanking parameters (clearance (Mucha, 2010), cutting edge radii (Klingenberg & Boer, 2008), sheet-metal thickness (Hambli, 2001) and the number of blanking cycles (Cheon & Kim, 2016)) on the wear evolution. On the other hand, the influence of occurring wear states on the quality of the component was investigated (Hernández et al. 2006; Makich et al. 2008). In addition to these empirical and experimental investigations, great progress has been made in the 2D and 3D numerical simulation of wear phenomena during blanking. Using computational methods,

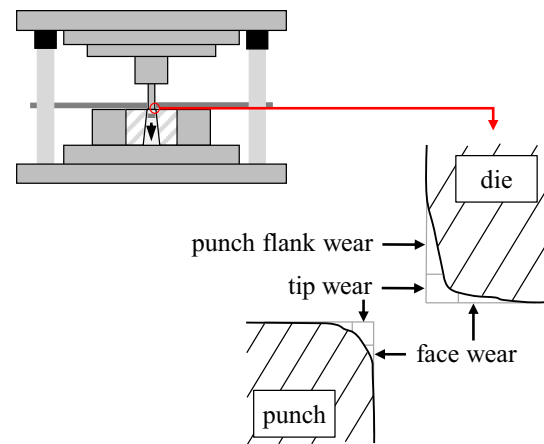


Fig. 2 Geometry of the active element worn surface in generic case of tool wear (Hernández et al. 2006)

several authors tried to predict tool wear and the resulting form errors on the blanked parts (Hambli et al. 2009) and to optimize process parameters to reduce these errors (Faura et al. 1997). According to Hernández et al., these numerical simulations are very complex, due to large deformations, ductile fractures and crack propagation occurring during the process of blanking (Hernández et al. 2006). Due to the highly non-linear behavior of blanking processes, a detailed description of the wear state by means of empirical or numerical models is only possible to a limited extent. For this reason, real-time monitoring of the wear state is difficult to implement in an industrial production environment and is based on diagnostic approaches that focus on the detection, isolation and identification of defects as they occur (Jemielniak et al. 2012). In these systems, empirical knowledge is necessary to determine thresholds and envelopes to detect erroneous process conditions. The potential of the data is not fully exploited and is limited to the identification of binary faults (Isermann, 2011). Identification of error-cause relationships or the quantification of the size of wear does not take place in the industrial production environment. In contrast, a large number of studies in the literature investigated data-driven approaches to fully exploit the potential of given data sets for detailed description of wear conditions in manufacturing processes. Especially in the field of machining authors have penalized ML models for predicting wear states in the last decades. They focus on predicting wear-related quality features of machined parts [27, 28], quantifying wear of machining tools (Shen et al. 2020; Zhou et al. 2020) or predicting tool life [31]. In contrast, ML models for blanking processes are mainly used to describe discrete fault states, while the prediction and quantification of wear conditions is hardly found in recent literature. Jin and Shi developed a wavlet transformation technique for feature-preserving data compression of tonnage information

during a stamping process. Based on these features a segmental thresholding strategy similar to a decision tree is presented to identify a change in thickness of the material (Jin & Shi, 2001). Lee et al. investigated an architecture of an automatic supervisory system for monitoring wear during a blanking process. The system employs an autoregressive time-series model to predict the online measured maximal blanking force. Based on the autoregressive model, coefficients wear-related state of the tool (sharp vs. worn) are predicted using a linear discriminant function (Lee et al. 1997). In a similar way, Ge et al. extracted features from a force signal of a blanking process by means of an autoregressive model and used them in a Hidden Markov Model to classify faulty conditions such as misfeeding of the work piece, slug pulling or variation of the sheet thickness (Ge et al. 2004). Bassiuny et al. investigated an empirical mode decomposition to extract the features from the strain signals of a sheet metal stamping process. Based on the features, a Hilbert marginal spectrum is calculated which is used to classify the faulty process conditions (misfeed and material too thick) by a learning vector quantization network (Bassiuny et al. 2007). Jin and Shi extracted features from the tonnage signal of a stamping process by a principal component analysis and performed a regression analysis to model the relationship between features and process variables. Hierarchical classifiers and a cross-validation method are used for root-cause determination to predict the binary state (normal / abnormal) of the process variable (Jin & Shi, 2000). Hambli presented a backpropagation neural network algorithm for predicting burr height formation on blanked parts, considering the tool clearance and its wear state. The inputs of the artificial neural network (ANN) are generated by a finite element analysis of the circular blanking process. Validation of the results by the ANN showed good agreement with a deviation of 10% between predicted and experimental burr height (Hambli, 2001). Next to the presented regression models and ANN, Support Vector machines are getting more and more attractive for data driven analysis and their application to manufacturing processes. In particular, for monitoring machinery (health) condition (Goyal et al. 2020; Liu et al. 2017) or for quality classification of machined, additive manufactured or welded parts (Baturynska & Martinsen, 2021; Çaydaş & Ekici, 2012; He & Li, 2016). In this regard, the support vector machine (SVM) is not only one of the most powerful and robust classification and regression algorithms, but has also significantly improved the handling with multi classification problems and unbalanced data sets (Cervantes et al. 2020). Even with these advantages, the use of SVM in sheet metal forming processes is mainly limited to the classification of discrete errors. Ge et al. presented in his work SVMs with different kernel functions to classify erroneous states during a blanking process. The experimental results showed that SVM can achieve a success rate of 99% detecting these

faults (Ge et al. 2002). He presented in his study a one class SVM for detecting abnormal health conditions (ok / nok) of a progressive stamping machine. To achieve a desirable performance considering a trade-off relationship between the false alarm rate and false detection rate hyper parameters of the SVM need to be optimized (Qiu et al. 2020). Another work is presented by Zhou et al., which uses an SVM to identify a missing part in one of the die stations of a progressive tool using tonnage signals (Zhou & Jin, 2005). Considering this literature, SVM approaches are only used for binary classification. The extension of this binary SVM to a multiclass scenario is not yet state of the art in metal forming and is also limited in other domains.

Common to all these ML approaches is the need for a valid data set that is used to train and validate the model based on experimental or simulatively generated data. No effort was invested into pre-processing the data as this is one of the most time-consuming steps in applying an ML model to scientific problems and how the data is acquired. Investigations into the extent to which sensor configuration and positioning, as well as the step of data pre-processing, influence the model performance for predicting the process state or the quality of manufactured parts are not to be found in the literature. Especially in the process of blanking, which takes place at production rates up to 1000 strokes per minute (spm), careful selection and integration of the sensor technology as well as the preparation of the data before the modeling step are of crucial importance. The high stroke rates lead to short tool engagement times, complex stresses and strains, and a nonlinear course of sensorial measured time series (Hambli, 2001). In addition, the process is affected by vibrations caused by dynamic effects. During a blanking cycle, impacts occur when the blank holder and the punch hit the sheet metal, the sheet metal breaks and the punch is pulled out of the sheet metal. Each impulse induces vibrations in the tool which directly influence the acquired time signal. These vibrations are superimposed on the time signal and their influence increases significantly with higher stroke speeds. Hirsch et al. as well as Slavič et al. identify an increase of the maximum process force of approximately 14% when the stroke speed difference is increased by 200 spm (Hirsch et al. 2011; Slavič et al. 2014).

A summary of the literature shows that ML models can fully exploit the potential of existing data to describe, evaluate and predict process states in manufacturing processes. However, due to the complexity and dynamics of blanking processes, as well as the nonlinear and transient characteristics of the measured time series, only a few studies applying ML models to the process can be found in the literature. In particular, for wear prediction, the literature is limited to the identification of binary states. Furthermore, in order to apply a robust ML model to describe wear states during blanking, it is essential to consider the steps of data acquisition (sensor

position and sensor type) data preprocessing and data transformation. Therefore, in this study an SVM is applied to classify the wear states of abrasives in a blanking process and the influence of the data acquisition and data preparation step on the performance to the model is quantified. To achieve this objective, Sect. 3 presents the procedure in this work. For this purpose, the experimental setup is presented, force signals are descriptively analyzed and (SVM) to classify the wear states in the presented blanking process are explained. In Sect. 4, the qualitative and quantitative evaluation of the SVM is carried out depending on the sensor configuration (type and position), the preprocessing (filtering) and the transformation (extraction and selection). The resilience of the model is validated using different stroke speeds and associated uncertainties due to dynamic effects.

Experimental setup and procedure

In order to quantify the influence of data acquisition, preprocessing and transformation on the performance of mSVM for predicting the wear states, the experimental procedure is shown Fig. 3. In the first step, the acquisition of force signals and the measuring of the different stages of the abrasive worn punches are conducted. The force signals are acquired during each stroke cycle varying sensor positions and types (Table 2). Afterwards, feature engineering and a principal component analysis (PCA) is conducted to extract features from these force signals. The wear states are quantified by optically measured cutting edge radii of the punches. As

demonstrated in Sect. 3.3, even two features derived from the force signals represent over 95% of the total variance in the data set. Therefore, the two major principal axes (f_1 and f_2) serve as input while the five wear states serve as output of the mSVM. The acquired process and quality data (features f_i and wear states r_i) are divided into test and training data sets. Finally, the training and validation of the mSVM takes place based on these data. Figure 3 shows the required steps for predicting the wear state during blanking including data collection, feature extraction, feature evaluation, and application of mSVM.

Press and experimental tool

All experiments were carried out on a high-speed press from Bruderer AG (BSTA 810). The press allows a nominal force of 810 kN and stroke rates of up to 1000 spm at a stroke height of 35 mm. The machine parameters for the experiments were set to a stroke distance of 35 mm and stroke speeds of 200 spm, 300 spm, 400 spm and 500 spm. The geometry of the punch was chosen to be a cylindrical with a diameter of 6 mm. The gap between punch and die is set to 0.15 mm and in relation to the sheet thickness of the semi-finished product of 2 mm, this results in a clearance of 7.5%. Table 1 shows the tool and press parameters as well as the properties of the sheet metal.

The multisensory tool consists of a lower and an upper part, which are connected to each other by four guide columns. The cylindrical punch is connected to the adapter plate in the upper tool via a plunger. A piezo electrical

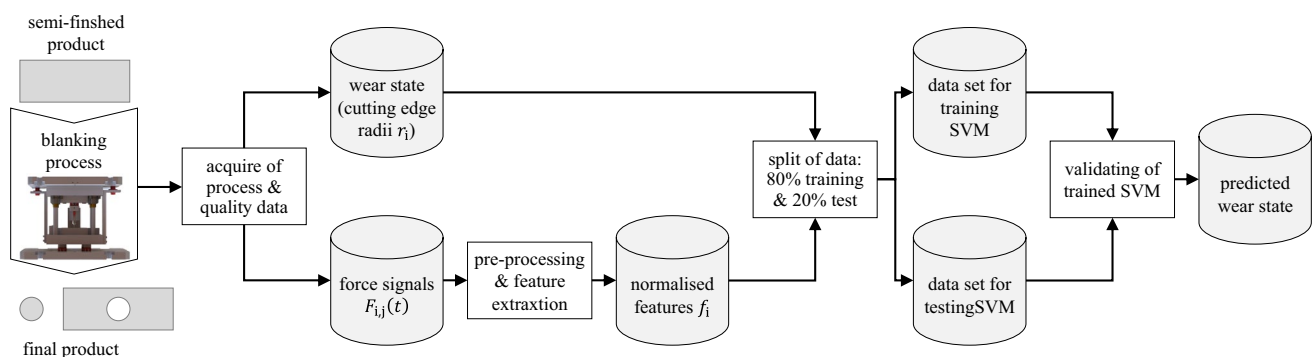


Fig. 3 Procedure for predicting abrasive wear states during a blanking process including the steps of data acquisition, feature extraction, feature selection and mSVM application

Table 1 Tool and press parameters and the properties of the sheet metal for conducting the experiments

Material		Press		Tool	
Material description	DC03 (1.0347)	Stroke height in mm	35	Diameter of punch in mm	6
Sheet thickness in mm	2 ± 0.02	Stroke speed in spm	200 ... 500	Clearance in %	7.5
Tensile strength in MPa	299.8 ± 2.9			Cutting edge radii (wear state) in μm	7 ... 604

force washer is integrated into the direct force flux. In the lower tool, the die is integrated into an adapter plate and connected to the base plate via plungers. Therefore, four force washers are integrated into the direct force flux. In addition, four linear strain gauges combined to a full bridge are taped on the plunger, a measuring pin is mounted in a hole of the adapter plate in the upper tool and a strain sensor is screwed into the indirect force flow to the columns of the press. Figure 4 shows the detailed design of the tool, its integration into the press and the positioning of the sensors.

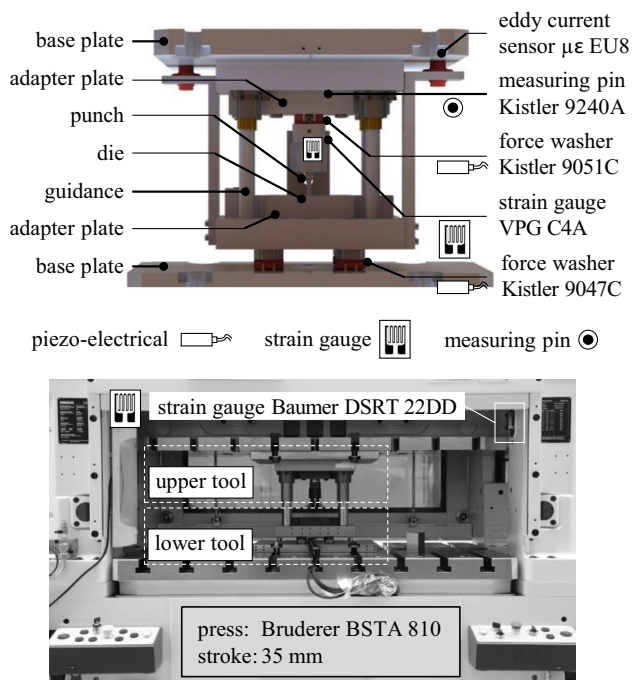


Fig. 4 Sectional view of a multisensory blanking tool with respect to the positioning of the force sensors

Acquisition of process and quality data

To measure the force signals, various sensor types were integrated at different positions in the tool and the press (Fig. 4). Table 2 summarizes these sensor types and positions. Four sensor types are differentiated, resistive strain gauges, piezo-electrical force washers and piezo-electrical measuring pins. In the upper tool an uniaxial piezo-electrical force washer (Kistler 9051C) with a nominal force of 120 kN was integrated into the direct force flux. In addition, the force signals are measured in the direct force flux of the lower tool by four triaxial piezo-electrical force washers (Kistler 9047C) with a nominal force of 80 kN. The symmetrical arrangement of the sensors allows the cutting force to be spatially resolved. For the further investigations in this work, these four force signals parallel to the stroke path of the ram are summed up to a total force. In addition to these piezo-electrical force washers, a piezo-electrical quartz transverse measuring pin (Kistler 9240 A) is integrated on the adapter plate in the upper tool. The size of $8 \times 14.4 \text{ mm}^2$ allows measuring of strains from 0 to $500 \mu\epsilon$ even with limited space in the tool. The charge distribution resulting from a change of the load on the force washers and the measuring pin is converted into a voltage signal by a charge amplifier (Kistler 5073A). In addition to the piezo-electrical sensors, low-cost strain gauges and strain sensors are integrated into the tool and the press. Four linear strain gauges (VPG C4A-06-125SL-350) combined to a full bridge are installed between the punch holder and the adapter plate. Strain sensors (Baumer DSRT 23DF) are already mounted to the columns of the press and are used to monitor the nominal forces to protect the press from overloading. The force washers, quartz transverse measuring pin and the charge amplifier as well as the strain gauges were calibrated according to DIN EN ISO 376. For this purpose, the tool was integrated into a universal testing machine (Zwick Roell Z100) and cylindrically loaded at different force levels. The calibration factor of each sensor is determined by calculating the slope of the calibration lines. Since the strain sensors were already integrated

Table 2 Various sensor types and their position inside the tool /press for acquiring the force signals

Specification	Position	Calibration factor	Type	NI module	Description in paper
Uniaxial force washer Kistler 9051C	Upper tool in direct force flow	446.34 N/V	Piezo-electrical	NI 9220	Force washer upper tool (reference)
Uniaxial force washer Kistler 9047C	Lower tool in direct force flow	1499.68 N/V	Piezo-electrical	NI 9220	Force washer lower tool
Quartz measuring pin Kistler 9240A	Adapter plate upper tool in indirect force flow	-7556.45 N/V	Piezo-electrical	NI 9220	Piezo pin
Linear strain gauge VPG C4A	Shell surface of the punch	$2.34 \cdot 10^5 \frac{\text{N}}{\text{V/V}}$	Resistive	NI 9220	Strain gauge punch
Foils strain gauge Baumer DSRT 23DF	Column of the press frame	$2.51 \cdot 10^4 \text{ N/V}$	Resistive	NI 9237	Strain sensor press frame

into the press, a calibration was already carried out by Bruderer AG. The voltage signals of the force sensors and the eddy current sensor are recorded by a CompactRIO system with integrated measuring modules NI 9220 (analog input ± 10 V), NI 9215 (analog input ± 10 V) and NI 9237 (analog input ± 25 mV/V).

Since the blanking process is strongly influenced by the dynamic effects resulting from the high stroke rates, the setting of the sampling frequency has to be set high enough. The required sampling frequency depends on the selected stroke rate. Thereby, a complete stroke cycle of the Bruderer press from top dead center to bottom dead center back to top dead center is determined by 360° . The actual tool engagement time only takes place in a limited angular range. With an assumed stroke frequency of 500 spm and an angular range of 160° to 210° , this results in a time of 0.024 s while the tool is in contact with sheet metal. As the tool engagement time is further divided into the three phases (Fig. 1), a time window of 3.3 ms results for the punch-phase. In order to be able to continue processing the dynamic effects under these circumstances, a sampling frequency of f_s 50 kHz was selected. To keep the amount of data as small as possible, only a part of the entire stroke is recorded. Therefore, data recording starts if the tool passes an inductive proximity sensor. During the reverse stroke, the tool again passes past the inductive proximity sensor and the measurement is stopped. This finally results in a measuring range of 120° – 240° or punch penetration of 13 mm and a time series with 2070 samples for the highest stroke speed of 500 spm. In addition to the acquired force signals, quality data (abrasive wear state of the punch) is optically measured by a confocal white light microscope (NanoFocus AG type μ surf explorer). It is assumed that abrasive wear causes the rounding of the cutting edge radii of the punch (Falconnet et al. 2015; Hohmann et al. 2017). In order to approximate the reproducible abrasive wear conditions without the need for time-consuming long-term experiments, the cutting edge

radii are mechanically rounded by a post machining process. The cutting edge is varied in five steps starting from a sharp cutting edged radii r_0 . Table 3 shows the results of the five optically measured the five wear states represented by the cutting edge radii r_i and the abrasive wear volume (Fig. 5).

Based on these five wear states, the experimental procedure is shown in Table 4. 100 stroke cycles at four different stroke speeds are acquired for each wear state, resulting in twenty single experiments and 2000 time series for all experiments. Considering the number of sensor types and positions, 10,000 time series must be evaluated.

Descriptive analyses of force signals

As mentioned above the time series characteristics of the measured force signals are nonlinear, transient and affected by dynamic effects. Therefore, in the following section the force signal quality is quantified by of descriptive statistics. Differences between the sensor positions and types, as well as the influence of increasing stroke speed, and the related dynamic effects have to be quantified. To compare

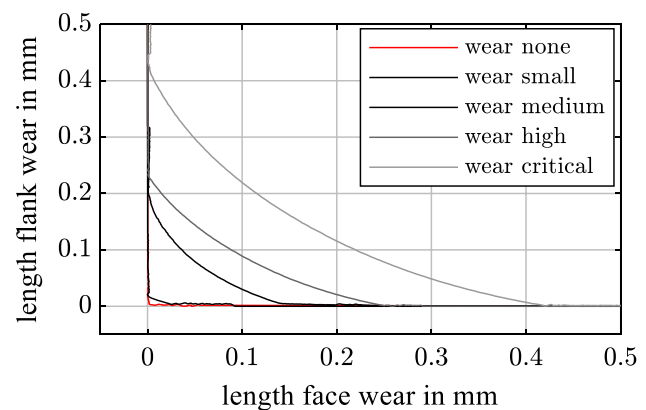


Fig. 5 Quantitative and qualitative characterization of the abrasive wear states by optically measuring the cutting edge radii r_i

Table 3 Quantitative characterization of the abrasive wear states by optically measuring the cutting edge radii r_i

Wear state	Cutting edge radii r_i in μm	Abrasive wear volume in mm^3
Wear state: no abrasive wear Cutting edge radii r_0 (reference)	7.624	0.003
Wear state: small abrasive wear Cutting edge radii r_1	66.717	0.005
Wear state: medium abrasive wear Cutting edge radii r_2	254.281	0.178
Wear state: high abrasive wear Cutting edge radii r_3	354.459	0.361
Wear state: critical abrasive wear Cutting edge radii r_4	604.826	1.046

Table 4 Experimental procedure for acquiring force signals considering sensor positions and types, stroke speeds v_i and wear states r_i

Parameter	Variations
cutting edge radii r_i in μm $r_0=7$, $r_1=66$, $r_2=254$, $r_3=354$ and $r_4=604$	5
Stroke speeds v_i in spm $v_0=200$, $v_1=300$, $v_2=400$ and $v_3=500$	4
Sensor position and sensor type See Table 2	5
Total number of experiments	1000
Number of strokes per experiment	100
Total number of time series	10,000

the influence of different sensor positions and types as well as stroke speeds on the quality of the acquired force signals this descriptive analysis is exemplarily conducted with an unworn punch (r_0) and a stroke speed of 200 spm. Figure 6 shows the scaled and offset-adjusted force signals of 100 single strokes for each sensor described in Table 2. During the further investigations, the force washer in the upper tool is used as a reference. In Fig. 6 the mean of 100 strokes of the reference sensor is visualized with a red line to compare it with the 100 force signals of each sensor type.

The qualitative analysis of the force signal shows that especially strain gauge sensors tend to noise, which is visualized by the wide scatter band. While the piezo-electrical sensors show less noise, especially the force washer in the lower tools is superimposed by vibrations. In addition, the strain sensor mounted to columns of the press shows superimposed by vibrations but with a lower frequency 80.4 ± 17.3 Hz for all stroke speeds caused by the eigenfrequency of the press. Looking at the time signals, it is obvious that the sensor types and position influence the deviation of the measured time series. In order to quantify this deviation as well as to explore the cause for it the standard deviation of the force signal at each sample point s is determined for each sensor type. Let x_i be the i -th sample $i \in \{1, \dots, n\}$ of the force signal of a specific sensor type and \bar{x} the mean of the force

signal of a specific sensor type, standard deviation of the sample is calculated as follows

$$s = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2} \quad (1)$$

Figure 7 shows the standard deviations calculated from 100 strokes cycles for each sensor type.

The red line visualizes the standard deviations of the reference sensor (force washer upper tool). The results show a high averaged deviation over the entire stroke cycles for the strains gauges (strain gauge punch and strain sensor). In contrast, the mean deviation of piezo-electrical sensors is negligible over the entire stroke cycle (force washer lower and piezo pin). Whereby the piezo pin shows a slightly higher mean standard deviation than the force washer in the lower tool. Furthermore, the piezo-electrical sensors show a peak of the standard deviation in the sampling range from 845 to 855 (P1) and 1150 to 1160 (P2).

Taking the sensor type and sensor position as well as the stroke speed into account allows the conclusion that the standard deviation of each experiment (including 1000 strokes) is mainly caused by electrical noise and dynamic effects. The dynamic effects result from the static and

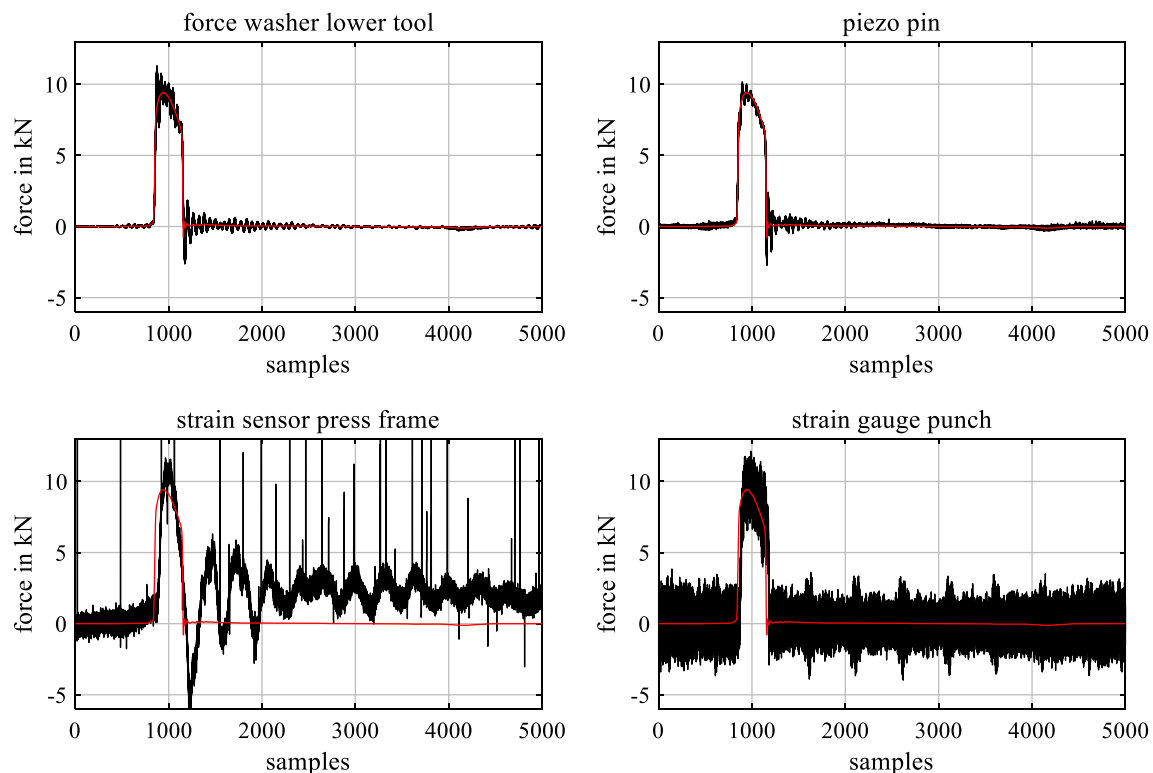


Figure 6 Force signals of the sensor types for 100 single strokes (wear state r_0 and stroke speed 200 spm) compared to the mean force signal of the reference sensor

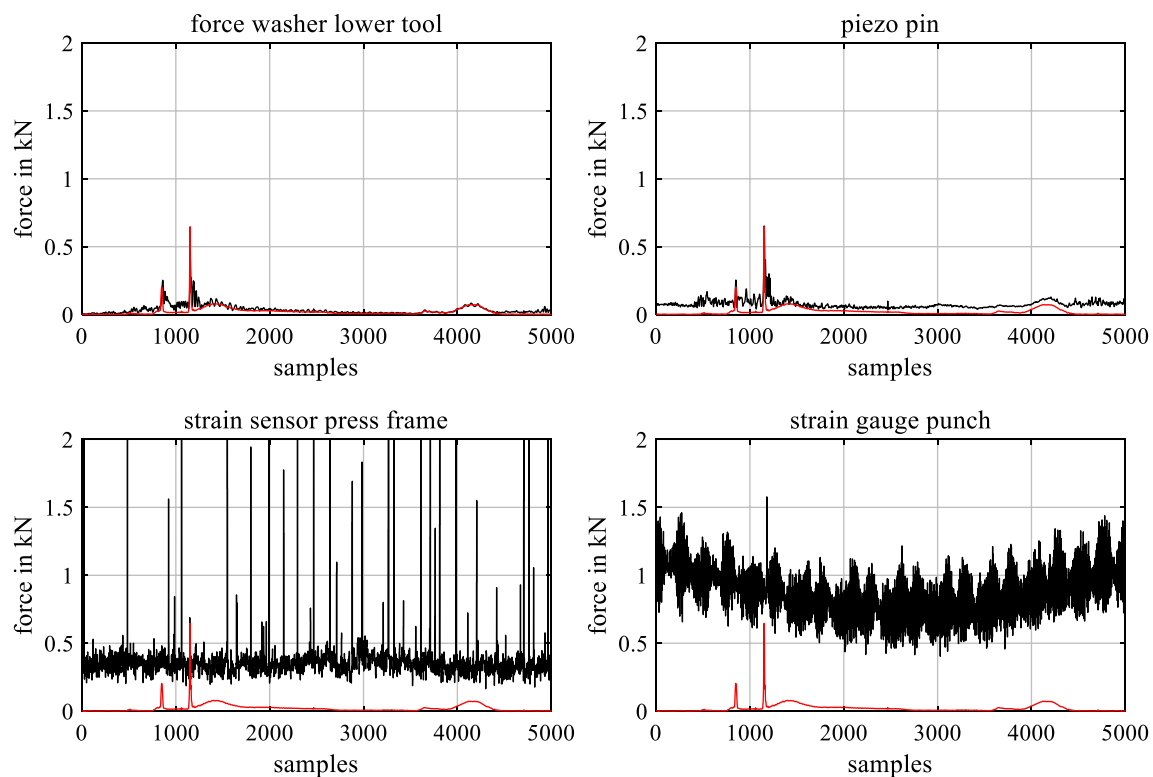


Fig. 7 Standard deviation of the sensor types for 100 single strokes (wear state r_0 and stroke speed 200 spm) compared to the standard deviation of the force signal of the reference sensor

dynamic behavior of the press as well as process-related vibrations (P1 and P2). At peak P2 the material breakage is initiated and elastic energy stored in the system is abruptly released. Caused by the masses located between the forming zone and the position of measurement, inertia forces are induced to the system and superimpose vibrations to the force signal. Especially the comparison between the piezo-electrical sensors in the upper and lower tool (Fig. 6) shows vibration superimposed to the force signal in the lower tool caused by an uneven mass distribution. In this case the mass of the components between the sensors and the forming are ten times higher in the lower tool than in the upper tool. In contrast, peak P1 results from the punch hitting the sheet metal and initiating an impact into the structure of the press. Since the impact of the punch on the sheet is a short impulse, the product of inertial forces and exposure time is approximated by a Dirac impulse. This Dirac impulse is superimposed on the actual force signal and leads to an infinite, continuous frequency spectrum. Especially this impulse is recognizable in the time series of the strain sensor (Fig. 6).

In addition to the dynamic effects the deviation of the force signal depends on the electrical noise. Especially strain gauges which are indirectly measuring the force related to a change in the resistance are affected by this. On the one hand, the change of resistance is caused by the elongation

of the measuring grid. On the other hand, the resistance is influenced by the electrical contacting of the sensors, the length of the cables and electrostatic as well as electrodynamic phenomena (e. g. grounded loops, thermal, shot, flicker, burst and transit-time noise) which are further summarized under the term electrical noise.

Next to the dynamic effects and the electrical noise, deviations in the force signals result from a time shift (already few samples) of the force signal. Although the signal between following strokes is absolutely repeatable, even a small offset of a few samples can cause large deviations, especially with strongly fluctuating signals. This effect is especially demonstrated by the shift of the point in time at which the material breaks and the punch hits the sheet metal caused by varying local properties of the material (tensile strength, elongation at break, crystallographic defects) or the semi-finished product (sheet thickness, lubrication, etc.). However, it is not only the material properties but also the hardware of the measuring chain (A/D converter, charge amplifier, measuring modules, performance CPU, etc.) that influences a time shift between following force signals. In particular, triggering the start and end point of the measurement is effected by uncertainties of the magnetic field detected by the eddy current sensor. Because of the uncertainties of this inductive proximity sensors, a shift on the

time axis also occurs even if the signal between following strokes is absolutely repeatable. To summarize, the deviations in the force signal are strongly influenced by vibrations or electrical noise which are amplified by a time shift (few samples) of the force signals from stroke to stroke due to the properties of the material as well as the design of the measuring chain. In order to quantify the influence of the sensor types and sensor position as well as the stroke speed on deviation of the signal in each stroke cycle by one characteristic value, the average standard deviation \bar{s} is determined. Equation (2) is used to calculate the confidential range with a statistical certainty of 99%. Error bars are estimated by the value μ which consists of the arithmetic mean \bar{x} and the confidential limits. These confidential limits are calculated from the student factor t , the standard deviation of the sample s number of observation n .

$$\mu = \bar{x} \pm \frac{t \cdot s}{\sqrt{n}} \quad (2)$$

Table 5 summarizes the averaged standard deviations \bar{s} for an unworn punch (r_0), the four stroke speeds and all sensor types. It shows qualitatively that the averaged standard deviations increase with higher stroke speeds, progressive distance from the forming zone, a direct measurement compared to an indirect measurement and the use of a resistive sensor compared to a piezo electrical sensor.

Data transformation

The experimental procedure shown in Table 4 results in a data set that has to be prepared for mSVM before application. A major step of this data transformation is the reduction of the dimension of the data set by removing the redundant data (Liu & Motoda, 1998). A large data set results in a high model complexity which leads to poor generalizability of the model and to high computational times. A trade-off between the accuracy of the model and the computational effort can take place by extracting features from the data set. Since the performance of the ML algorithm is significantly influenced by this data transformation, the feature extraction is a key factor for successful ML projects (Domingos, 2012). In

engineering applications, features are usually extracted from sensorial measured time series. According to Li, features can be extracted either from the time domain, the frequency domain or the time–frequency domain (Wang & Gao, 2006). The feature extraction from the time domain can be derived directly from the sensor signals. Mostly these are statistical parameters such as maximum values, mean values, standard deviations, skewness, kurtosis or the root mean square. Furthermore, Hoppe et al. as well as Kubik et al. were able to show in their studies that engineering feature from the time domain represent effective parameters for describing process conditions during blanking (see Fig. 8) (Hoppe et al. 2019; Kubik et al. 2021). Therefore, the force signal is initially divided into three phases and characteristic points which define the start and end points as well as extrema during each phase are identified. Finally, the features are derived from these characteristic points and can be described as the length $l_{j,i}$, maximum force $F_{j,i}$ and work done $W_{j,i}$ in each. In this case the index j describes the respective phase of the cutting process (punch-phase (p), push-phase (pu), withdraw-phase (w)) and i the number of variations in the experiments.

For extracting features from the frequency domain spectral and frequency analyses are performed. While the use of frequency domain analyses (Hilbert-Huang Transformation, Short-Time Fourier Transformation, etc.) are already widespread in domains like health and condition

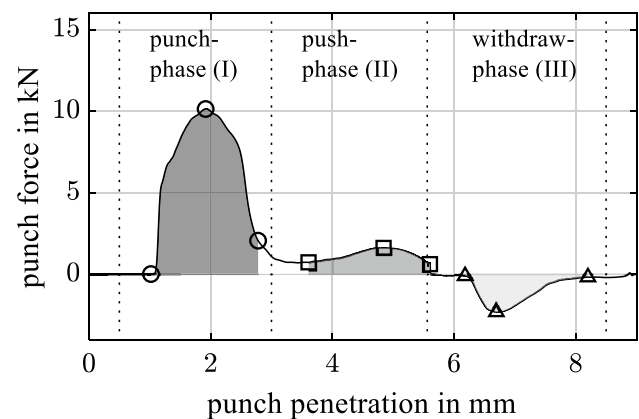


Fig. 8 Features extracted from force signal

Table 5 Averaged mean standard deviation of the model for all sensor types over all stroke speeds for an unworn punch (r_0)

Stroke speed in spm	Force washer upper tool in N	Force washer lower tool in N	Piezo pin in N	Strain gauge punch in N	Strain sensor frame in N
200	13 ± 1	36 ± 1	71 ± 1	888 ± 7	390 ± 15
300	14 ± 1	40 ± 2	74 ± 2	994 ± 7	446 ± 25
400	15 ± 1	45 ± 1	389 ± 9	1160 ± 11	531 ± 18
500	21 ± 2	47 ± 3	517 ± 5	1169 ± 9	799 ± 34

monitoring of machining processes, these methods are rarely used in forming technology production (Aydin et al., 2012). Representatives of time–frequency domain techniques are the wavelet transform, which are particularly suitable for the investigation of instationary time series (Fan et al. 2001). The most important technique for feature extraction in engineering applications is the principal component analysis (PCA) (Addison et al. 2003). In production engineering applications, PCA is used for the purpose of quality control (Wu & Chyu, 2004), condition monitoring (Rato & Reis, 2020) and predictive maintenance (Lewin, 1995). PCA is a multivariate statistical technique that handles large amounts of data via orthogonal projection. It reduces the dimensionality of a data set by projecting original data into a lower dimensional orthogonal space defined by a few significant eigenvectors. Applying this method to a blanking process means that we are searching for a compact representation of the measured force signals that still contains the information about relevant variations. In this paper PCA is used to extract features from experimentally acquired force signals, that denoted as a matrix $\mathbf{X} \in \mathbb{R}^{m \times p}$, in which each row vector \mathbf{x}_i is a complete cycle of a force signal with p measurement points, and m is the total number of observations. Therefore, the first step of the PCA is to compute the covariance matrix

$$\Sigma = \text{cov}(\bar{\mathbf{X}}) = \begin{bmatrix} \sigma_{11} & \cdots & \sigma_{1p} \\ \vdots & \ddots & \vdots \\ \sigma_{p1} & \cdots & \sigma_{pp} \end{bmatrix} \quad (3)$$

of a zero-mean set of NN measurement series

$$\bar{\mathbf{X}} = \mathbf{X} - \boldsymbol{\mu} \quad \text{assuming} \quad \mu_i = \frac{\sum_{j=1}^p x_{i,j}}{p} \quad (4)$$

to receive the variance at each point in time and also the joint variability with all other points in time. The principal components are computed by solving the eigenvalue problem of the covariance matrix $\Sigma \in \mathbb{R}^{p \times p}$

$$\Sigma \cdot \mathbf{v}_j = \lambda_j \cdot \mathbf{v}_j \quad (5)$$

The vector $\mathbf{v}_j \in \mathbb{R}^{p \times 1}$ ($j = 1, \dots, p$) is the normalized eigenvector of the sample covariance matrix Σ of $\bar{\mathbf{X}}$. Using these principal axes \mathbf{v}_j , each force signal \mathbf{x}_i ($i = 1, \dots, m$) can be transformed into a couple of features

$$f_{j,i} = \mathbf{v}_j^T \cdot \mathbf{x}_i \quad (6)$$

When a data set is projected to eigenvectors, it is often found that only the first few eigenvectors, corresponding to larger eigenvalues, are associated with the systematic process variations. All remaining eigenvectors reflect the variations of the process noise (Jin & Shi, 2000). This noise

is caused by random disturbances, electrical noise, uncontrollable process variations or dynamic effects. Therefore, dimension p of the original PCA can be reduced into a smaller dimension p_{opt} ($p_{\text{opt}} < p$). To do this, the principal axes are first sorted by their size of the eigenvalue and after this the dimension is reduced until the selection index η_{opt} becomes smaller than a defined value.

$$\eta_{\text{opt}} = \frac{\sum_{j=1}^{p_{\text{opt}}} \lambda_j}{\sum_{j=1}^p \lambda_j} \quad (7)$$

There is no methodical approach to define this selection index, η_{opt} set to 95%. This ensures that more than 95% of the variance in the force signal can be explained by the p_{opt} largest eigenvalues. Depending on the stroke speed every data set consists of N samples, being $N = f(v_i)$, and 500 dimensions which are related to the experiments conducted for each stroke speed. Since force signals from five different sensors are recorded, the data set is dimensioned as $\mathbf{X}_i \in \mathbb{R}^{500 \times N}$ with $i \in \{1, 2, 3, 4, 5\}$. The five corresponding wear states are captured in the output vector $\mathbf{Y} \in \mathbb{R}^{1 \times N}$ with the same dimensions. Looking at the selection index η_{opt} the first two eigenvectors explain 96.2% of the total variance in the signal. For further investigations only these two eigenvectors will be used as input variables for the mSVM. To investigate the influence on the performance of the mode, three additional features (f_3, f_4 and f_5) from PCA as well as nine engineered features (Fig. 8) are provided. To ensure the comparison of the models, even by combining PCA and Engineered Features, the features are normalized by the Z-score.

The machine learning method of SVM

As described in Sect. 2 there are many different types of machine learning methods that can be applied for classification tasks. Therefore, the most suitable method for classifying the wear states during blanking has to be determined during a grid search. In this study a twostep grid searches were executed. On the one hand to find the best machine learning method and on the other hand the optimize the parameter configuration of the chosen ML method. From the force signal of the strain gauge (VPG C4A) on the punch, the ML models were trained and validated. Based on the classification accuracies for predicting the wear conditions the performance of the model was evaluated. For training the ML algorithm, the first two principal axes (process data) are available as input data and the wear states (quality data) as output data for each time series. To keep the ML models as simple as possible.

To keep the model as simple as possible, mainly linear classification models were compared. It should be noted, however, that the use of non-linear classification functions can improve the model performance. Since the focus of this study is to show how data acquisition, data preprocessing and data transformation affects the model performance, a simple ML model is chosen. Therefore, a linear discriminant analysis (LDA), random forest (RF), mSVM with a linear kernel function, Naive Bayes classifier (NB), as well as k-nearest neighbor classifier (k-NN) for classifying the abrasive wear states were tested. In Fig. 9a the accuracy achieved with a fivefold cross validation for the best configuration of the machine learning methods are shown. The accuracy values of the classification results from the NB and k-NN are not good and only depicted for completeness. The best prediction results were obtained from the mSVM with a linear kernel function. In this simplest case of linearly separable classes, an optimal hyperplane is sought, that maximizes the distance between the best separating plane and the nearest data. Considering a dataset given as $[(x_1, y_1), (x_2, y_2), \dots, (x_i, y_i)]$ consisting of m training samples, where $x_i \in \mathbb{R}^m$ is an m -dimensional feature vector representing the i -th training tuple and $y_i \in \{-1, 1\}$ is the corresponding class label the optimal hyperplane can be found by the following optimization problem:

$$\begin{aligned} \text{minimize: } & L(w) = \frac{1}{2}w^2 \\ \text{subject to: } & y_i(w^T x + b) \geq 1 \\ & i = 1, \dots, m \end{aligned} \quad (8)$$

The hyperplane in the feature space can be described by the equation $w^T x + b$, where $w \in \mathbb{R}^m$ and $b \in \mathbb{R}$. In the case of linearly nonseparable classes, no hyperplane is found that is capable of correctly classifying every training sample. In this case the optimization problem can be described by setting a soft margin by including a slack variable ξ_i and a tuning parameter C , the box constraint. This parameter allows the training algorithm a certain misclassification in the training set and applies costs to this misclassification. The higher the box constraint, the higher the cost for the misclassified points, leading to a stricter separation of the data (Cherkassky & Ma, 2004). In this simplest type, SVM divides the data points linearly into classes. In real-world problems, however, we find more than two classes. In this paper we have to deal with five abrasive wear states (cutting edge radius r_i). Therefore, this mSVM is broken down to multiple binary SVMs. Two commonly used techniques one-vs-rest (OVR) and one-vs-one (OVO) can be found in the literature. While the OVO approach splits the dataset into one dataset for each class versus every other class, the OVR approach splits the dataset into one binary dataset for each class (Kijisirikul & Ussivakul, 2002). Considering an M -class problem, where $y_i \in \{1, \dots, M\}$ are the corresponding class labels with i -the SVM optimization problem can be described as follows:

$$\begin{aligned} \text{minimize: } & L(w_i, \xi_j^i) = \frac{1}{2}w_i^2 + C \sum_{j=1}^m \xi_j^i \\ \text{subject to: } & y_j(w_i^T x_i + b_i) \geq 1 - \xi_j^i \\ & \xi_j^i \geq 0 \end{aligned} \quad (9)$$

The mSVM used in this work deal with a classification of five different wear states by considering the problem as collection of binary classification problems using the OVO approach. As mentioned above, in the second step of the grid search the hyperparameter of the mSVM with a linear kernel are optimized and set to $\xi_i = 0.65$ for the slack variable and $C = 413$ for the tuning parameter. Figure 9b shows the optimized hyperparameter for the mSVM.

With principal axis extracted from the PCA and the engineered features fifteen features as input data and five wear states as output data are available for the mSVM. The model is trained for every combination of stroke speed and sensor type, resulting in 200 classification models. For each model the input data set is reduced to two significant parameters which always consist of the first two features from the PCA. Only for the investigation of the influence of data transformation on the model performance (Sect. 4.3) different

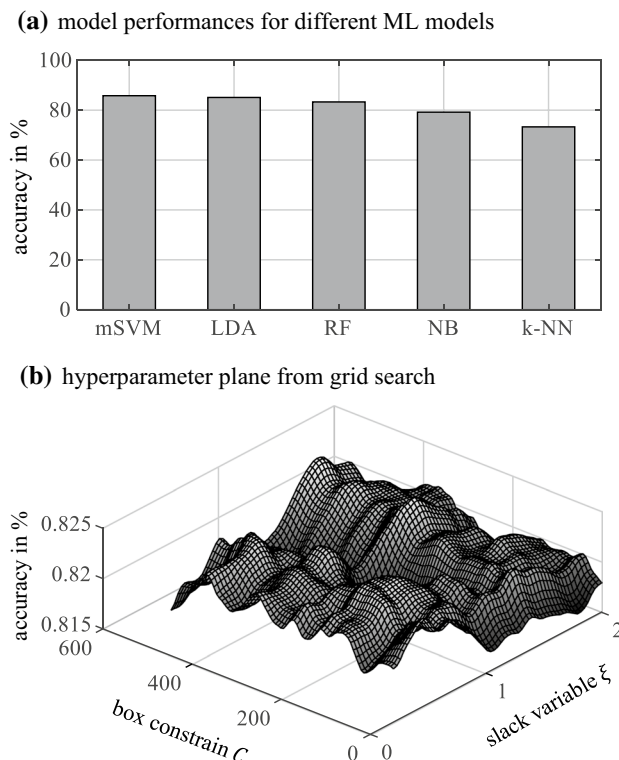


Fig. 9 Classifications results for each tested ML method (a) and optimized hyperparameter for the mSVM with linear kernel function (b)

feature combinations are used. To train the model 80% ($X^{2 \times 400}$) of the data is used while 20% ($X^{2 \times 100}$) is applied to ten-fold cross-validation of the model. For the final evaluation of the model and its performance accuracy as well as the overlap of the probability density function between classes separated by the SVM are considered. The accuracy describes the percentage of correct predictions and is defined as the quotient of number of observations correctly assigned to a class in relation to all observations. In addition to this accuracy, the performance of the model is determined by calculating the Mahalanobis distance d_M . Calculating this distance allows to quantify the separability of classes by considering the distance between the centroid of these classes as well as the variance inside each class. As the results of this study will show, the mSVM is suitable for the classification of abrasive wear states during blanking and achieves close to 100% accuracy, depending on the sensor position, the sensor type as well as the stroke speed, and the calculation of Mahalanobis distance is of interest for the separation of closely adjacent wear states. In the further investigations of this paper, the Mahalanobis distance is calculated only for the separation between the wear conditions ‘little (low) wear’ (r_2) and ‘medium wear’ (r_3).

Evaluation of model performance

In the following section the influence of data acquisition, preprocessing and transformation on the performance of the mSVM will be shown. The model performance is evaluated by the accuracy and the Mahalanobis distance. In order to quantify the resilience of the classification model, the performance of the mSVM is analyzed at different stroke speeds (see Table 4). To quantify the influence of data acquisition, preprocessing and transformation, the trained models are compared with a reference model for each experiment. This reference model is obtained by training the mSVM with the

force signals of the piezo electrical sensor in the upper tool (force washer upper tool) at a maximum stroke speed of 500 spm. The reference sensor in the upper tool was used to train the mSVM, since the sensor is placed in the direct force flux close to the forming zone, is slightly affected by superimposed vibrations in the system and generates high quality data reflecting the actual physical state of the process.

Model performance depending on sensor position and type

The performance of the classification model depends on the quality of the data set and related to this to their acquisition procedure (Calmano et al. 2013; Groche et al. 2019). It is decisive which sensor types are selected, which measuring method is used and which position is chosen inside the tool or the press. Figure 10 shows the result of the mSVM for a piezo electrical force washer (force washer lower tool: direct measurement), a piezo electrical measuring pin (piezo pin: indirect measurement) and a resistive strain gauge sensor (strain gauge punch: indirect measurement).

Considering these results qualitatively, it is shown that especially piezo electrical sensor types in direct and indirect force flux significantly improve the performance of the classification model. They allow to measure dynamic effects without losing the information stored in the high frequency domain and are able to classify the wear state even with stroke speeds of up to 500 spm which results in high model accuracy as shown in Table 6.

The accuracy of the model is close to 100% over the entire stroke speed range for both piezo electrical sensors. Although both sensors have comparable accuracy, they show differences in the separability of their classes. However, to compare a piezo electrical sensor with a resistive sensor, the resistive sensor, which is susceptible to electrical noise, must be filtered. Therefore, the resistive strain gauge sensor is low pass filtered ($f_{c,opt}=7.5$ kHz). The results show that this type

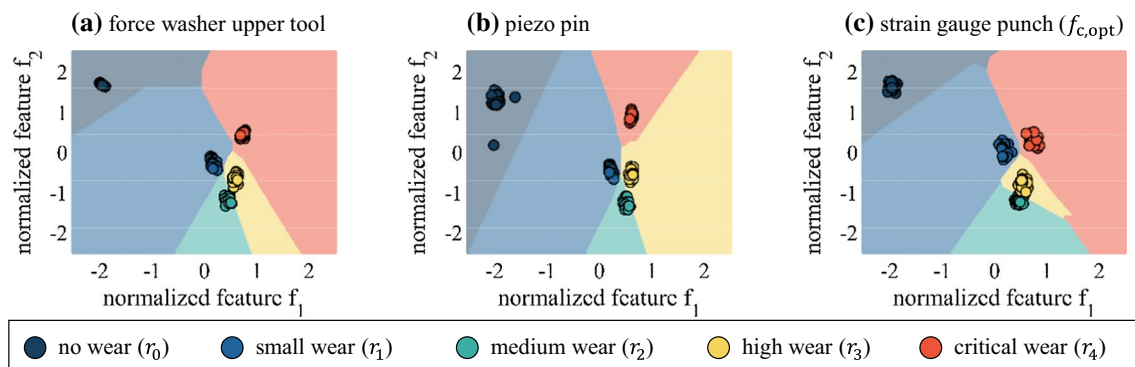


Fig. 10 Qualitative visualization of the mSVM classifying five wear states comparing piezo electrical force washer (a), piezo electrical measuring pin (b) and a resistive strain gauge on the punch (c)

Table 6 Influence of different sensor types and positions on accuracy to predicting wear states with a multiclass SVM and Mahalanobis distance to quantify the separability of class 2 and 3

Model performance	Sensortype	$v_0=200$ spm	$v_1=300$ spm	$v_2=400$ spm	$v_3=500$ spm
Accuracy θ_{acc} in %	Force washer upper tool	100	100	100	100
	Piezo pin	100	100	100	99
	Strain gauge punch	100	100	100	100
	Force washer lower tool	100	100	100	100
	Strain sensor frame	100	98	88	71
d_M (class 2/3)	Force washer upper tool	252.74	560.60	140.69	74.46
	Piezo pin	225.19	306.90	77.32	78.01
	Strain gauge punch	114.08	114.87	78.97	71.75
	Force washer lower tool	203.65	282.63	30.96	45.36
	Strain sensor frame	100.53	22.34	21.44	3.24

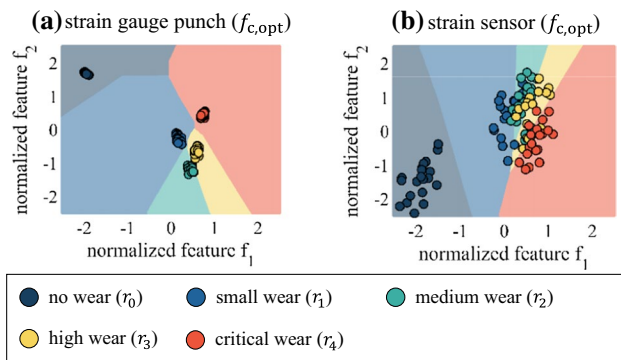


Fig. 11 Qualitative visualization of multiclass SVM classifying five wear states comparing different sensor position of strain gauge mounted at the punch (a) and strains sensor (b) integrated to the columns of the press

of sensor can correctly detect all wear states, with a 100% accuracy, even at high stroke speeds. In contrast, the resistive sensor provides a poorer separability of the classes which decreases with higher stroke speed, caused by the limited dynamic properties and the tendency to noise.

As described by Calmano et al., in addition to the sensor type, sensor position also affects the quality of the acquired data and thus the classification performance of the mSVM (Calmano et al. 2013). Comparing the performance of the classification models for two sensors of the same type at different positions in the process, this hypothesis can be confirmed. Therefore, Fig. 11 shows the classification results for two low pass filtered resistive sensors (strain sensor frame and strain gauge punch). The strain sensor in the columns has a comparable accuracy of 100% to the strain gauge in the punch for classifying the wear states at lower stroke speeds. However, as the stroke speed increases, a resilient description of the wear states is no longer possible and the accuracy drops to 71% at a stroke speed of 500 spm. A similar tendency can be seen in the separability, which decreases from 114.08 to 71.75 for the strain gauge punch and even from 100.53 to 3.24 for the strain sensor frame.

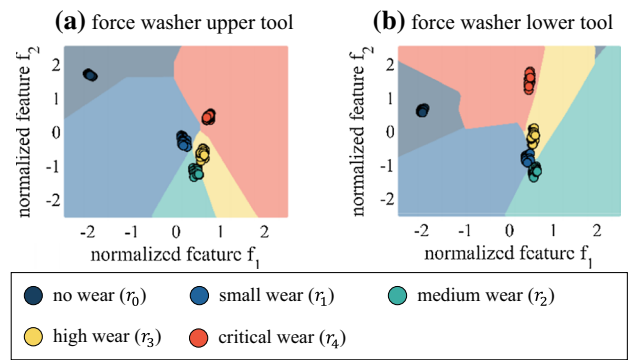


Fig. 12 Qualitative visualization of the multiclass SVM classifying five wear states comparing the integration of piezo electrical force washers in lower (a) and upper tool (b)

Comparing the performance of the classification models for two piezo electrical force washers (force washer upper tool and force washer lower tool) which have the same physical distance to the forming zone but are integrated in the upper and lower tool differences in the model performance become apparent. The classification model based on the force washer in the upper tool shows lower deviations within each class, which is also reflected by the Mahalanobis distance. The Euclidean distance between centroid from class 2 and class 3 is about $d_E = 0.65 \pm 0.05$ for both models. In contrast to this, the Mahalanobis variant takes the variances into account and shows a difference of 13.6% averaged over the stroke speeds between the force sensor in the upper tool and lower tool. The accuracy of the model is 100% for both piezoelectric-type sensors over all stroke speeds (see Fig. 12).

Model performance depending on data preprocessing

In addition to the data acquisition step and the related generation of valid data sets, the preprocessing of the data plays an important role in the application of ML models, since

time signals acquired by sensors are usually consisting of a signal superimposed by a noise level. The noise is characterized by a stochastic, unpredictable behavior of the acquired variable caused by electrical or external sources. Noise is a combination of electrical (thermal, shot, flicker, burst and transit-time noise) and background noise (electromagnetic and acoustic noise as well as mechanical vibrations) (Computing & Corporation, Data Acquisition Handbook, 2004, 2004). From the perspective of signal processing, force signals in this study are described as follows:

$$F(t) = F_{\text{sig}}(t) + F_{\text{n,el}}(t) + F_{\text{n,tool}} + F_{\text{n,process}} \quad (10)$$

$F_{\text{sig}}(t)$ describes a relevant part of the acquired force signal, which represents the physical state of the process. The electrical noise $F_{\text{n,el}}(t)$ is composed of internal, physical effects such as thermal or transient time noise and external sources like electrostatic noise (voltage is induced in a conductor that is exposed to a time-varying electric field) and electromagnetic noise (current is induced in a conductor that is exposed to a time-varying magnetic field). Resistance measurements in particular, (e. g. strain gauges) are influenced by varying voltages and currents caused by electrostatic and electromagnetic noise. In addition, the resistance is influenced by the hardware of the measuring chain which is influenced by the length of the connection cable, its shielding and insulation, the electrical contacting, the temperature and the integration process. In contrast to strain gauges, piezo electrical sensors are less affected by this electrical noise. They are mainly effected by a drift of the signal related to a time-dependent shift of charge in the piezo electrical crystal. In addition to the electrical noise, mechanically caused noise is superimposed on the actual force signal. It is divided into process noise $F_{\text{n,process}}$ and press and tool noise $F_{\text{n,tool}}$. Process noise is related to vibrations caused by the impact of the punch on the sheet metal, by the material breakage, and the pushing and withdrawing of the punch through the sheet metal. This process-related noise may contain valuable information about the state of the process and is related to physical effects in the process. In contrast, press- and tool-related noise is caused by the static and dynamic effects (inertial forces resulting from the high accelerations of the ram) of the press and the connected peripherals (inertial forces and vibrations of the feed unit which accelerate and decelerate the sheet metal strip during each stroke cycle). While the mechanically caused noise strongly depends on the dynamic of the blanking process and therefore is related to the stroke speed, the electrical noise mainly depends on the sensor type (piezo electrical or resistive), the measuring method (direct or indirect measurement) and the measuring position (distance to the forming zone). To reduce both noise effects,

filter operations are a common method (Jackson, 1996). In order to quantify the influence of filtering and the selection of the filter design on the performance of the classification model, the time signals of a resistive sensor and a piezo electrical sensor are analyzed. In particular, the influence of a selectable design filter parameter on model performance is demonstrated. A distinction is made between an optimally designed filter parameter, a poorly designed filter parameter and no filter. As a filter design, a third order Butterworth filter with a normalized cutoff frequency w_i is used. The normalized cutoff frequency is calculated by Eq. 12 with the sampling frequency f_s of 50 kHz and the cutoff frequency $f_{c,i}$ as design parameters.

$$w_i = \frac{2 \cdot f_{c,i}}{f_s} \quad (11)$$

However, special attention has to be paid to the selection of the designing filter parameter $f_{c,i}$. For filtering the force signal of the resistive strain gauge sensor on the punch an optimal cutoff frequency $f_{c,\text{opt}} = 7.5$ kHz was used, while the cutoff frequency for the deficient case $f_{c,\text{def}} = 8.5$ kHz was determined based on previous investigations (see Fig. 13). Comparing these filter designs, the optimal cutoff frequency shows a significantly better performance of the model. Looking at the frequency domain of the strain gauge sensor distinct frequencies in the range from 7.8 to 8.1 kHz and low-frequencies in the range up to 1.2 kHz are detected. While the high frequencies are related to electrical noise, the lower frequencies depend on the static and dynamic behavior of the press and the physics of the process.

An optimally designed filters achieve an accuracy of the model up to 100% and significantly increased the separability of the classes even at strokes speeds of 500 spm. In contrast, the cutoff frequency of the deficient filter design is above the frequency range $f_{c,\text{def}} = 8.5$ kHz where the electrical noise is highly dissipated. Therefore, an improvement of the signal quality and related to this the model performance is not expected. Instead, the deficient filtered force signal remove physically relevant parts of the data and worsens the accuracy of the classification model to 82% even at stroke low stroke speeds as shown in Table 7 and qualitatively visualized in Fig. 14. A similar effect can be seen in the class separability, which improved insignificantly with a deficient designing filter parameter compared to non-filter design.

Looking at the frequency domain of the force washer, signal frequencies of major impact are localized in a lower range up to 1.2 kHz (Fig. 15b). Electrical noise in a high frequency range, as it occurs with resistive sensors, cannot be identified. In order to reduce the influence of these mechanical vibrations, the cutoff frequency is set to 0.7 kHz. However, selecting the cut-off frequency in this case requires

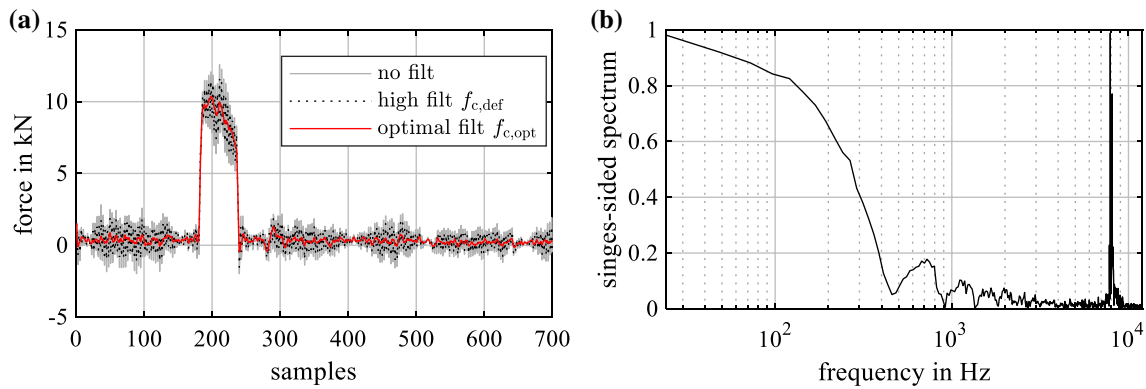


Fig. 13 Visualization of the force signal by the strain gauge on the punch filtered in the time domain **(a)** and frequency range of the raw signal **(b)**

Table 7 Influence of different filter operations on accuracy in predicting wear states with a multiclass SVM and Mahalanobis distance to quantify the separability of class 2 and 3

Model performance	Sensortype	$\nu_0=200$ spm	$\nu_1=300$ spm	$\nu_2=400$ spm	$\nu_3=500$ spm
Accuracy θ_{acc} in %	Strain gauge punch (no filter)	80	83	84	76
	Strain gauge punch ($f_{c,opt}=7.5$ kh)	100	100	100	100
	Strain gauge punch ($f_{c,def}=8.5$ kh)	82	86	81	97
	Force washer lower tool (no filter)	100	100	100	100
	Force washer lower tool ($f_{c,opt}=0.7$ kh)	100	100	100	100
	Force washer lower tool ($f_{c,opt}=0.3$ kh)	93	100	98	91
d_M (class 2/3)	Strain gauge punch (no filter)	5.09	9.50	5.76	12.68
	Strain gauge punch ($f_{c,opt}=7.5$ kh)	114.08	114.87	78.97	71.75
	Strain gauge punch ($f_{c,def}=8.5$ kh)	7.22	14.10	2.62	40.83
	Force washer lower tool (no filter)	252.73	282.62	30.95	45.36
	Force washer lower tool ($f_{c,opt}=0.7$ kh)	348.81	254.83	389.33	81.32
	Force washer lower tool ($f_{c,opt}=0.3$ kh)	2554.83	2987.23	9931.17	186.95

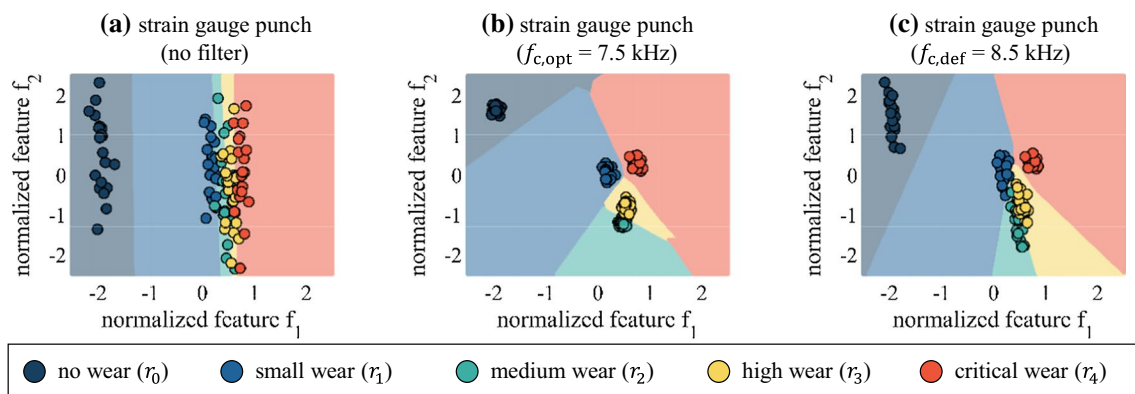


Fig. 14 Qualitative visualization of the multiclass SVM classifying five wear states comparing the non-filtered force signals from strain gauge on the punch **(a)** with the optimal $f_{c,opt}$ **(b)** and deficit $f_{c,def}$ **(c)** filtered signal

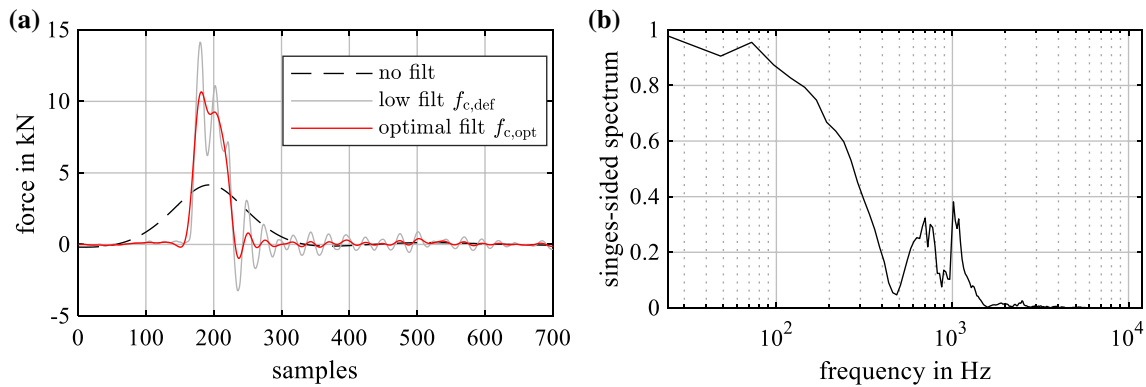


Fig. 15 Visualization of the force signal by the force washer in the lower tool filtered in the time domain (a) and frequency range of the raw signal (b)

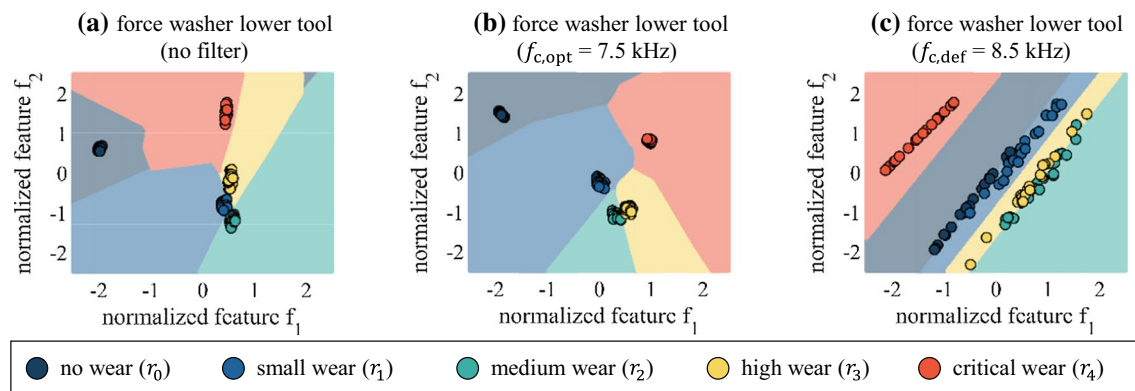


Fig. 16 Qualitative visualization of the multiclass SVM classifying five wear states comparing the non-filtered force signals from force washer in the upper tool (a) with the optimal $f_{c,opt}$ (b) and deficit $f_{c,def}$ (c) filtered signal

caution, since the physically relevant characteristics of the force signal are removed if the cutoff frequency is set to a value smaller or equal to 0.5 kHz. This can be seen in the qualitative visualization of the classification model (Fig. 16).

While the accuracy for both the unfiltered and optimally filtered design reaches 100%, the accuracy for the deficient filter design drops to 91%. In contrast to the accuracy, a better separability of the classes is achieved with a lower cutoff frequency of $f_{c,def} = 0.2$ kHz (see Table 7). Due to mechanical vibrations in connection with a stroke-related time shift, the deviation of the force signals increases, which worsens the separability. Filtering the force signal with a cutoff frequency lower than 0.2 kHz reduces these deviations and increases the Mahalanobis distance, especially for higher strokes speeds, as shown in Table 7.

Consequently, the results show that signal filtering can significantly improve the model quality, but the selection of the filter operation and design is of crucial importance. On the one hand, systematic investigations of the frequency spectrum help with the design. On the other hand, it is essential to include process knowledge in the design procedure.

Especially, for identifying the frequency range of the force signal which contains relevant information, process knowledge is needed. In addition, for the filter design procedure it could be shown that low frequencies contain important information for a resilient implementation of the classification model, while high frequency ranges caused by electrical noise are negligible (Table 8).

Model performance depending on feature selection

In addition to data acquisition and preprocessing, the transformation of data has a significant influence on the performance of ML algorithms (Blum & Langley, 1997). Therefore, we aim to achieve the accuracy of an ML model as represented by the entire data set by selecting relevant features from given times series with a minimal loss of information while improving the computational overhead. In the following part, the performance of the classification model is evaluated by comparing the features extracted by a PCA and a feature engineering approach (Sect. 3.3). These features are composed of the two major principal axes (f_1

and f_2) extracted by the PCA, and nine engineered features (F_i , w_i and l_i with $i \in \{\text{punch, push, withdraw}\}$). In the first step a correlations matrix was created to remove redundant features. The corresponding parameters and their Pearson correlation coefficients were determined for $N=500$ single strokes for each experiment and are shown in Fig. 17.

Evaluating the matrix shows a significant correlation between the maximal force (F_{punch}), the length (l_{punch}) and the work done (W_{punch}) in the punch-phase. As a result of progressive abrasive wear, the cutting edge radii (r_i) is rounded, stress peaks in the forming are reduced in contrast to a sharp cutting edge radii and the plastic deformation phase is extended. Caused by this extended deformation, the percentage of the shear zone and the related phenomena increase with the length of the punch-phase (Klingenberg & Boer, 2008). Assuming that the progressive wear has only an insignificant influence on the maximal force of the punch-phase, an extension of the length of the punch-phase leads to an increased amount of work in this phase. Redundancies between the work done and the length of the punch-phase are identified. In contrast, the work done in the push- and withdraw-phase mainly depend on the maximal force in each phase (F_{push} and F_{with}). Hohmann et al. prove in their work that progressive wear leads to increased frictional forces between the punch in the sheet metal and thus to higher maximal forces in the push- and withdraw-phase (Hohmann et al. 2017). Therefore, a redundancy between the maximal forces and the work done in both phases is detected, which is also confirmed by high correlation coefficients of 0.99 for the push- and 0.97 for the withdraw-phase. The work performed in all three phases is redundant to the maximal forces (F_{push} and F_{with}) and the length of the punch-phase (l_{punch}). This feature is negligible for the further training of the classification mode. Furthermore, the literature indicates

an abrasive wear on the cutting edge marginally effects the maximal force in the punch-phase. In addition, previous investigations within the scope of this work have shown that a resilient identification of the start and end points as well as the maximum force in the punch-phase is difficult to automate due to the limited formation of this phase. Since the determination of the minimum force in the withdraw-phase is more resilient than the determination of the length of this phase, feature l_{punch} and F_{with} are obtained for further investigations. This selection is confirmed in the coefficients of the correlation matrix, which provides a moderate correlation of -0.83 between both features. At this point it should be mentioned that the correlation matrix coefficient can be used for an unsupervised feature selection leading to an exclusion of highly correlating features. However, this does not exclude the possibility that important information may be contained in two correlating features, which are crucial for the further identification of the wear states (Hoppe et al. 2019). Figure 18 shows the qualitative influence of different feature combinations on the performance of the multiclass SVM. The training of the model is based on (a) two major principal axes of the PCA (f_1 and f_2), (b) two engineered features (l_{punch} and F_{with}) and a hybrid approach (f_1 and l_{punch}). Force signals are acquired with the force washer in the upper tool.

This is shown qualitatively in the visualization of the classification models and quantitatively in the calculated Mahalanobis distance which is 35.2% higher in the hybrid approach compared to the isolated PCA approach and 9.3% higher compared to the isolated engineered features approach. This leads to the conclusion, that feature selection can improve the performance of the model. The availability of domain knowledge and process expertise for the extraction and selection of the features, especially in the feature engineering approach, plays a crucial role (Table 8).

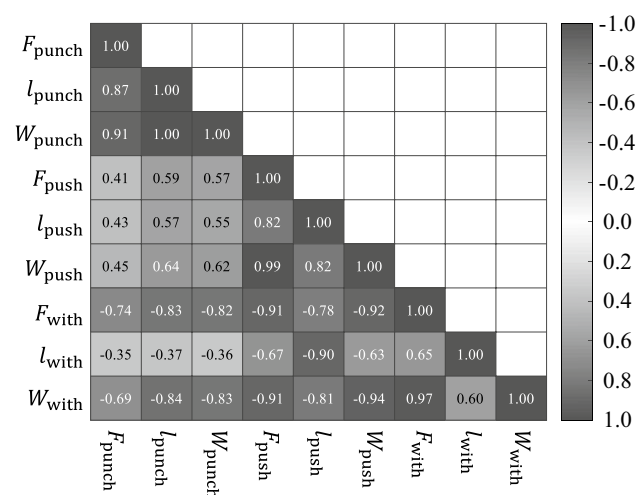


Fig. 17 Correlation matrix of the engineered features

Discussion

The results demonstrate that abrasive wear states can be classified by a mSVM but data acquisition, data preprocessing and data transformation affect the performance of the model. Depending on these three steps as well as the number of strokes, the accuracy can range from 100 to 71%. It was found that the variance within the acquired data sets, described by the mean standard deviation \bar{s} , correlates strongly with the performance of the classification model. Figure 19a shows the model performance as a double logarithmic function of the mean standard deviation for different sensor positions and sensor types and a stroke rate of 300 spm. With a decreasing mean standard deviation, as observed in the piezo electrical force sensors close to the forming zone, the accuracy of the model raises. A similar effect is seen for the separability represented by the

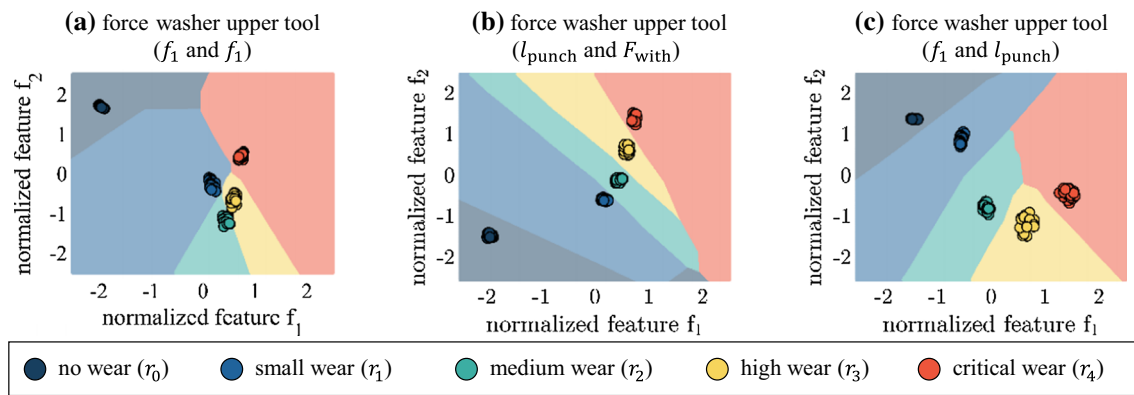


Fig. 18 Qualitative visualization of the multiclass SVM classifying five wear states comparing the feature selection approaches based on PCA features (a), engineered features (b) and a hybrid approach (c)

Table 8 Influence of different feature selections approaches on the accuracy in predicting wear states with a multiclass SVM and Mahalanobis distance to quantify the separability of class 2 and 3

Model performance	Sensortype	$v_0=200$ spm	$v_1=300$ spm	$v_2=400$ spm	$v_3=500$ spm
Accuracy θ_{acc} in %	Force washer upper tool (f_1 and f_1)	100	100	100	100
	Force washer upper tool (l_{punch} and F_{with})	100	100	100	100
	Force washer upper tool (f_1 and l_{punch})	100	100	100	100
d_M (class 2/3)	Force washer upper tool (f_1 and f_1)	203.65	560.60	140.69	74.46
	Force washer upper tool (l_{punch} and F_{with})	239.72	295.62	197.62	470.93
	Force washer upper tool (f_1 and l_{punch})	290.04	102.32	233.01	1592.21

Mahalanobis distance. In addition to the sensor position and the sensor type, the strokes speed influences the mean standard deviation. As a result of higher stroke speeds, dynamic effects are superimposed on the force signal. This increases the average standard deviation of the time signals, regardless of which sensor is used or where the sensor is positioned. As an example of this effect, Fig. 19b shows for the strain sensor in the press frame the influence of the stroke speed on the model accuracy and separability correlated with the mean standard deviation as a double logarithmic plot.

From a Machine Learning point of view, a correlation can be established between the variance in a time signal and the performance of mSVM. The origin of this variance in the data set is mainly caused by the sensor type and the sensor position (physical distance between sensor and forming zone). Looking at the commonly used sensor types for monitoring of forming processes (piezoelectric and resistive force sensors), the resistive sensors in particular show a significantly higher variance in the data set than the piezoelectric sensors. This is mainly due to the fact that resistive sensors tend to produce noise. In this process, resistive sensors measure forces indirectly by a change in resistance caused by an elongation of the measuring grid. This change in resistance is not only

determined by the elastic deformation of the measuring grid, but also by electrical noise caused by the electrical contacts, the length of the cable and the mounting process. In addition, resistive sensors are limited in representing high frequency ranges due to their mechanical design and inertia. In particular, using this sensor type for acquiring forces in a blanking process results in a loss of physical valuable information, as the resistive sensor is not capable of representing dynamic effects.

As mentioned in the "Data-driven monitoring of blanking processes" section, blanking processes run at over 1000 spm. Impacts are triggered when the punch hits the sheet metal and the elastic energy stored in the system is released abruptly. This causes oscillations in the process which, due to the Dirac-shaped impulse, distribute the physical information in the measured force signals over a wide frequency range. This effect is amplified as the stroke speed increases. In contrast, piezo electrical sensors are able to improve the representation of dynamic effects during blanking due to their higher stiffness. In addition, they tend to be less susceptible to noise due to their compact design (protected from environmental influences such as vibrations, heat transfer, mechanical shocks, etc.) and their good electrical shielding, including the wires. Furthermore, compared to resistive

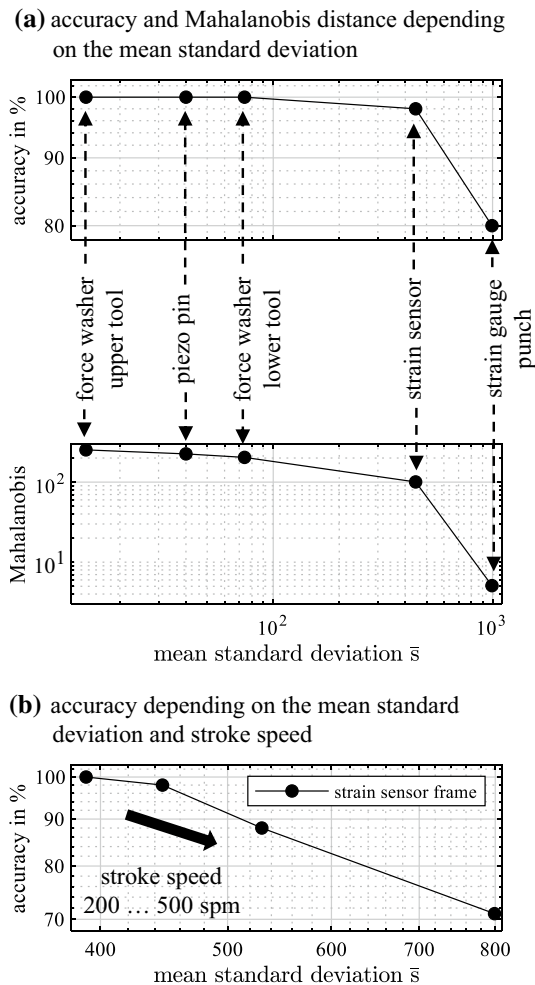


Fig. 19 Dependency of the model performance on the mean standard deviation for all sensors at 200 spm (a) and the correlation between stroke speed and model accuracy depending on the mean standard deviation (b)

sensors, piezoelectric sensors are less sensitive to temporal and spatial self-heating effects and the associated thermal noise.

In addition to the sensor type, it was shown that the performance of the mSVM is influenced by the physical distance of the sensor from the actual forming zone and its actual position in the tool. Depending on the distance between the sensor and the forming zone, the acquired force signal is influenced by the static and dynamic behavior of the press. The high accelerations generate inertial forces in the press frame, which have to be compensated by the mass balance of the press. Especially the mass balance system of the Bruderer press (BSTA 810) is designed for large tools (< 250 kg) causing unbalanced inertial forces using the experimental tool (~ 100 kg). As a result, an inertial force opposite to the movement of the ram is generated during the downward movement of the ram, which

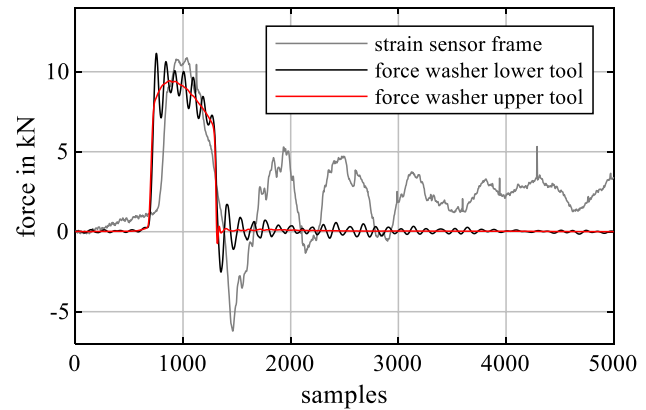


Fig. 20 Comparing force signal of piezo-electrical force washer integrated to the upper and lower tool as well as the frame sensor in the press frame at a stroke speed of 300 spm and an unworn punch (r_0)

is superimposed on the signal of the blanking process. These inertial forces are especially noticeable in the strain sensor of the press frame. Figure 20 shows a comparison of the time series of the force sensor in the press frame (strain sensor frame) and in the upper tool (force washer upper tool).

From a Machine Learning perspective, the data set is enriched with information about the dynamic behavior of the press, which is not useful for predicting abrasive wear states. In contrast, even a small change in the data set by adding or removing data points can significantly influence the results of ML models (Nguyen et al. 2015). This effect is also seen in the positioning of sensors in the blanking tool. With increasing distance to the forming zone, more and more unwanted information of the overall system is superimposed on the physical information of the blanking process (e. g. inertial forces). Not only the physical distance to the forming zone but also the position inside the tool plays a crucial role for the performance of an ML model. Looking at the force washers (force washer upper tool and force washer lower tool) in Fig. 4, it is evident that the physical distance to the forming zone is unchanged, while the position in the tool is different. Nevertheless, the mean standard deviation over the entire stroke rate range for the force washer in the lower tool is approximately 62% higher than the mean standard deviation of the force washer in the upper tool, as shown in Table 5. This is caused by the uneven distribution of the masses between the force washer and the forming zone. The mass of the lower tool components m_l (adapter plate, columns and die holder) is greater than the mass of the upper tool components m_u (adapter plate and punch holder), resulting in a mass ratio of $m_l/m_u \sim 10$. Due to the greater mass in the lower tool, vibrations caused by the impact of the punch on the sheet metal and the material breakage increase dynamic forces as shown in Fig. 20.

In summary, these findings show that the variance in a time signal depends on the sensor type and its positioning in the tool and correlates with the performance of ML models. In order to generate a high-performance ML model, the variance in the input data has to be minimized. In the case of blanking, dynamic effects caused by the high stroke speeds in combination with electrical noise and physically induced vibrations are the key factors influencing this variance. In order to minimize this variance in the signal, there are two ways of adjusting it. Firstly, the sensor position and type must be explicitly designed for the application in a blanking process to directly minimize the variance in the input data. It could therefore be shown that piezo electrical force washers near the forming zone in the upper tool are suitable to monitor the wear state in order to improve the quality of the input variables and thus the performance of the mSVM. Secondly, the variance in the input data can be indirectly affected by the step of data preprocessing and data transformation. Filtering offers the possibility to remove undesired characteristics from the force signal, but on the other hand it bears the risk of losing physically relevant information. Therefore, it is crucial to design filter operations based on expert knowledge, process knowledge and empirical experience in a way that unwanted properties are removed from the signal without losing physically relevant signal components (see “[Model performance depending on data preprocessing](#)” section).

Also the feature selection procedure can improve the model performance. Thereby it is necessary to select features that are robust against deviations in the time signal. As the results show, time domain features such as extreme values, first to fourth order statistical moments, and engineered features are affected by the variance in the signal. These findings are confirmed by Fig. 18, which shows a small value for the Mahalanobis distance for separating classes using only the engineered feature length of punch-phase $l_{p,i}$ and the work done in the withdraw-phase $W_{w,i}$. In particular, such features are very sensitive to shifts on the time axis. Even an offset of a few samples, caused by vibrations or uncertainties in the measurement chain (shift of the starting point of the measurement) lead to significant differences in the value of the feature. Figure 21 shows the deviation of the engineered feature $l_{p,i}$ depending on the start end point of the punch phase. Even if the end point varies around ± 2 samples, the length of the punch varies by ± 0.05 mm. A similar effect can be seen when using PCA features that do not distinguish between variance due to measured values (e. g. electrical noise) and variance due to real underlying signal variations. Even if an estimate of measurement variance is available, this information is not used in the construction of eigenvectors, e.g. by weighting noisy data. Therefore, selecting robust features and combining them, as shown in “[Model](#)

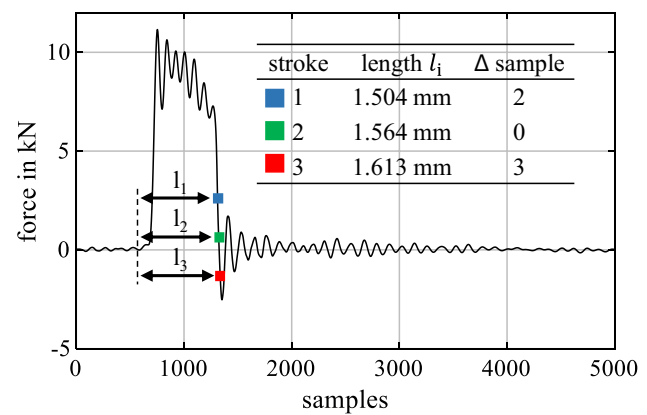


Fig. 21 Influence of small shifts on the time axis on the value of the engineered feature length of the punch-phase $l_{p,i}$

performance depending on feature selection” section, can improve the performance of ML models.

Summary and conclusion

This study examines the influence of data acquisition, preprocessing and transformation on the performance of a mSVM to classify abrasive wear states during blanking based on force signals. The performance of the model was quantitatively evaluated based on the model accuracy and the separability of the classes. Especially the extent of variance in the time signals significantly influences the performance of the classification model. This variance depends on the sensor types and positions as well as the measuring method (direct measurement vs. indirect measurement) and is improved by data preprocessing as well as a suitable extraction and selection of features. Therefore, the acquisition of a valid and trustworthy data set combined with a data preprocessing and transformation assisted by domain knowledge improves the successful implementation of AI projects.

Furthermore, this study shows for the first time a successful procedure predicting wear states during blanking even at high stroke speeds considering industrial conditions. Especially the use of piezo electrical sensors in the direct force flux close to the forming zone provides the suitable results from a technical and economical point of view. In combination, a following preprocessing step by filtering the data considering domain knowledge by designing the filter operation further improves the model performance. Also taking into account the data transformation can improve the model performance, especially the combination of different parameters (PCA features and engineered features) and selection of an optimal feature space. The methods investigated so far are limited to domains like machining or process engineering considering linear transformations in the time domain.

Since blanking is a nonlinear and highly dynamic process, the extension of extraction methods to nonlinear features from the frequency or time–frequency domain is necessary. In addition to force signals, further process variables (e.g. acceleration, AE, acoustic emissions, etc.) and sensor types have to be investigated with respect to their suitability for the identification of process states during blanking. Furthermore, the transferability of the quantified influences of data acquisition, preprocessing and transformation on the performance of ML algorithm to other manufacturing processes should be demonstrated.

Acknowledgements The results of this paper are achieved within the project “Mittelstand 4.0 - Kompetenzzentrum Darmstadt” funded by the German Federal Ministry for Economic Affairs and Energy (BMWi). The authors wish to thank for funding and supporting this project. Furthermore, the authors sincerely thank Bruderer AG for providing the high-speed press BSTA 810-145 and Kistler AG for providing measurement equipment to record sets of data even under challenging circumstances.

Funding Open Access funding enabled and organized by Projekt DEAL. The research leading to these results received funding from Federal Ministry of Economics and Energy – Collaborative Project: Mittelstand Kompetenzzentrum 4.0 Darmstadt under Grant Agreement No. 01MF15005A.

Declarations

Conflict of interest The authors report no declarations of interest.

References

- Addison, D. J. F., Wermter, S., & Arevian, G. Z. (2003). A comparison of feature extraction and selection techniques. In *Proceedings of the international conference on artificial neural networks* (pp. 212–215).
- Aydin, I., Karakose, M., & Akin, E. (2012). An adaptive artificial immune system for fault classification. *Journal of Intelligent Manufacturing*, 23, 1489–1499. <https://doi.org/10.1007/s10845-010-0449-5>
- Bassiuny, A. M., Li, X., & Du, R. (2007). Fault diagnosis of stamping process based on empirical mode decomposition and learning vector quantization. *International Journal of Machine Tools and Manufacture*, 47, 2298–2306. <https://doi.org/10.1016/j.ijmactools.2007.06.006>
- Baturynska, I., & Martinsen, K. (2021). Prediction of geometry deviations in additive manufactured parts: Comparison of linear regression with machine learning algorithms. *Journal of Intelligent Manufacturing*, 32, 179–200. <https://doi.org/10.1007/s10845-020-01567-0>
- Blum, A. L., & Langley, P. (1997). Selection of relevant features and examples in machine learning. *Artificial Intelligence*, 97, 245–271. [https://doi.org/10.1016/S0004-3702\(97\)00063-5](https://doi.org/10.1016/S0004-3702(97)00063-5)
- Bustillo, A., Pimenov, D. Y., Mia, M., & Kaplonek, W. (2021). Machine-learning for automatic prediction of flatness deviation considering the wear of the face mill teeth. *Journal of Intelligent Manufacturing*, 32, 895–912. <https://doi.org/10.1007/s10845-020-01645-3>
- Bustillo, A., Reis, R., Machado, A. R., & Pimenov, D. Y. (2020). Improving the accuracy of machine-learning models with data from machine test repetitions. *Journal of Intelligent Manufacturing*, 79, 1127. <https://doi.org/10.1007/s10845-020-01661-3>
- Calmano, S., Schmitt, S. O., & Groche, P. (2013). Prevention of overdimensioning in light-weight structures by control of uncertainties during production. In *Proceedings of the international conference on “new developments in forging technology”*.
- Çaydaş, U., & Ekici, S. (2012). Support vector machines models for surface roughness prediction in CNC turning of AISI 304 austenitic stainless steel. *Journal of Intelligent Manufacturing*, 23, 639–650. <https://doi.org/10.1007/s10845-010-0415-2>
- Cervantes, J., Garcia-Lamont, F., Rodríguez-Mazahua, L., & Lopez, A. (2020). A comprehensive survey on support vector machine classification: Applications, challenges and trends. *Neurocomputing*, 408, 189–215. <https://doi.org/10.1016/j.neucom.2019.10.118>
- Cheon, S., & Kim, H. (2016). Prediction of tool wear in the blanking process using updated geometry. *Wear*, 352–353, 160–170. <https://doi.org/10.1016/j.wear.2016.01.024>
- Cherkassky, V., & Ma, Y. (2004). Practical selection of SVM parameters and noise estimation for SVM regression. *Neural Networks*, 17, 113–126. [https://doi.org/10.1016/S0893-6080\(03\)00169-2](https://doi.org/10.1016/S0893-6080(03)00169-2)
- Domingos, P. (2012). A few useful things to know about machine learning. *Communications of the ACM*, 55, 78. <https://doi.org/10.1145/2347736.2347755>
- Falconnet, E., Chambert, J., Makich, H., & Monteil, G. (2015). Prediction of abrasive punch wear in copper alloy thin sheet blanking. *Wear*, 338–339, 144–154. <https://doi.org/10.1016/j.wear.2015.06.007>
- Fan, C., Dong, C., Zhang, C., & Wang, H. P. (2001). Detection of machine tool contouring errors using wavelet transforms and neural networks. *Journal of Manufacturing Systems*, 20, 98–112. [https://doi.org/10.1016/S0278-6125\(01\)80033-9](https://doi.org/10.1016/S0278-6125(01)80033-9)
- Faura, F., López, J., & Sanes, J. (1997). Criterion for tool wear limitation on blanking 18-8 stainless steel strips. *Revista De Metalurgia*, 33, 304–310. <https://doi.org/10.3989/revmetalm.1997.v33.i5.842>
- Ge, M., Du, R., & Xu, Y. (2004). Hidden markov model based fault diagnosis for stamping processes. *Mechanical Systems and Signal Processing*, 18, 391–408. [https://doi.org/10.1016/S0888-3270\(03\)00076-1](https://doi.org/10.1016/S0888-3270(03)00076-1)
- Ge, M., Du, R., Zhang, G., & Xu, Y. (2002). Fault diagnosis using support vector machine with an application in sheet metal stamping operations. *Mechanical Systems and Signal Processing*, 18, 143–159. [https://doi.org/10.1016/S0888-3270\(03\)00071-2](https://doi.org/10.1016/S0888-3270(03)00071-2)
- Goyal, D., Choudhary, A., Pabla, B. S., & Dhami, S. S. (2020). Support vector machines based non-contact fault diagnosis system for bearings. *Journal of Intelligent Manufacturing*, 31, 1275–1289. <https://doi.org/10.1007/s10845-019-01511-x>
- Groche, P., Hohmann, J., & Übelacker, D. (2019). Overview and comparison of different sensor positions and measuring methods for the process force measurement in stamping operations. *Measurement*, 135, 122–130. <https://doi.org/10.1016/j.measurement.2018.11.058>
- Hambli, R. (2001). Blanking tool wear modeling using the finite element method. *International Journal of Machine Tools and Manufacture*, 41, 1815–1829. [https://doi.org/10.1016/S0890-6955\(01\)00024-4](https://doi.org/10.1016/S0890-6955(01)00024-4)
- Hambli, R. (2002). Design of experiment based analysis for sheet metal blanking processes optimisation. *The International Journal of Advanced Manufacturing Technology*, 19, 403–410. <https://doi.org/10.1007/s001700200041>
- Hambli, R. (2002). Prediction of burr height formation in blanking processes using neural network. *International Journal of Mechanical Sciences*, 44, 2089–2102. [https://doi.org/10.1016/S0020-7403\(02\)00168-6](https://doi.org/10.1016/S0020-7403(02)00168-6)

- Hambli, R., Soulat, D., & Chamekh, A. (2009). Finite element prediction of blanking tool cost caused by wear. *The International Journal of Advanced Manufacturing Technology*, 44, 648–656. <https://doi.org/10.1007/s00170-008-1859-9>
- He, K., & Li, X. (2016). A quantitative estimation technique for welding quality using local mean decomposition and support vector machine. *Journal of Intelligent Manufacturing*, 27, 525–533. <https://doi.org/10.1007/s10845-014-0885-8>
- Hernández, J. J., Franco, P., Estrems, M., & Faura, F. (2006). Modeling and experimental analysis of the effects of tool wear on form errors in stainless steel blanking. *Journal of Materials Processing Technology*, 180, 143–150. <https://doi.org/10.1016/j.jmatprotec.2006.05.015>
- Hirsch, M., Demmel, P., Golle, R., & Hoffmann, H. (2011). Light metal in high-speed stamping tools. *Key Engineering Materials*, 473, 259–266.
- Hohmann, J., Schatz, T., & Groche, P. (2017). Intelligent wear identification based on sensory inline information for a stamping process. In *Proceedings of 5th international conference on advanced manufacturing engineering and technologies* (pp. 285–295). https://doi.org/10.1007/978-3-319-56430-2_21
- Hoppe, F., Hohmann, J., Knoll, M., Kubik, C., et al. (2019). Feature-based supervision of shear cutting processes on the basis of force measurements: Evaluation of feature engineering and feature extraction. *Procedia Manufacturing*, 34, 847–856. <https://doi.org/10.1016/j.promfg.2019.06.164>
- Isermann, R. (2011). *Fault-diagnosis applications*. Springer.
- Jackson, L. B. (1996). *Digital filters and signal processing*. Springer.
- Jemielniak, K., Urbański, T., Kossakowska, J., & Bombiński, S. (2012). Tool condition monitoring based on numerous signal features. *The International Journal of Advanced Manufacturing Technology*, 59, 73–81. <https://doi.org/10.1007/s00170-011-3504-2>
- Jin, J., & Shi, J. (2000). Diagnostic feature extraction from stamping tonnage signals based on design of experiments. *Journal of Manufacturing Science and Engineering*, 122, 360–369. <https://doi.org/10.1115/1.538926>
- Jin, J., & Shi, J. (2001). Automatic feature extraction of waveform signals for in-process diagnostic performance improvement. *Journal of Intelligent Manufacturing*, 12, 257–268. <https://doi.org/10.1023/A:1011248925750>
- Kijsirikul, B., & Ussivakul, N. (2002). Multiclass support vector machines using adaptive directed acyclic graph. In *International joint conference on neural networks* (pp. 980–985). <https://doi.org/10.1109/ijcnn.2002.1005608>
- Klingenberg, W., & de Boer, T. W. (2008). Condition-based maintenance in punching/blanking of sheet metal. *International Journal of Machine Tools and Manufacture*, 48, 589–598. <https://doi.org/10.1016/j.ijmachtools.2007.08.013>
- Klingenberg W, Singh UP (2004) Principles for on-line monitoring of tool wear during sheet metal punching. In *Proceedings of the 34th international MATADOR conference* (pp. 169–174). https://doi.org/10.1007/978-1-4471-0647-0_25
- Kubik, C., Hohmann, J., & Groche, P. (2021). Exploitation of force displacement curves in blanking—feature engineering beyond defect detection. *The International Journal of Advanced Manufacturing Technology*, 19, 403. <https://doi.org/10.1007/s00170-020-06450-z>
- Lange, K. (1985). *Handbook of metal forming*. New York: McGraw-Hill.
- Lee, W. B., Cheung, C. F., Chiu, W. M., & Chan, L. K. (1997). Automatic supervision of blanking tool wear using pattern recognition analysis. *International Journal of Machine Tools and Manufacture*, 37, 1079–1095. [https://doi.org/10.1016/S0890-6955\(97\)88104-7](https://doi.org/10.1016/S0890-6955(97)88104-7)
- Lewin, D. R. (1995). Predictive maintenance using PCA. *Control Engineering Practice*, 3, 415–421. [https://doi.org/10.1016/0967-0661\(95\)00015-M](https://doi.org/10.1016/0967-0661(95)00015-M)
- Liu, H., & Motoda, H. (1998). *Feature extraction*. Springer.
- Liu, S., Hu, Y., Li, C., Lu, H., et al. (2017). Machinery condition prediction based on wavelet and support vector machine. *Journal of Intelligent Manufacturing*, 28, 1045–1055. <https://doi.org/10.1007/s10845-015-1045-5>
- Makich, H., Carpentier, L., Monteil, G., Roizard, X., et al. (2008). Metrology of the burr amount—correlation with blanking operation parameters (blanked material—wear of the punch). *International Journal of Material Forming*, 1, 1243–1246. <https://doi.org/10.1007/s12289-008-0167-0>
- Measurement Computing Corporation. (2004). *Data acquisition handbook*.
- Moyné, J., & Iskandar, J. (2017). Big data analytics for smart manufacturing: case studies in semiconductor manufacturing. *Processes*, 5, 39. <https://doi.org/10.3390/pr5030039>
- Mucha, J. (2010). An experimental analysis of effects of various material tool's wear on burr during generator sheets blanking. *The International Journal of Advanced Manufacturing Technology*, 50, 495–507. <https://doi.org/10.1007/s00170-010-2554-1>
- Nguyen, A., Yosinski, J., & Clune, J. (2015). Deep neural networks are easily fooled: High confidence predictions for unrecognizable images. In *2015 IEEE conference on computer vision and pattern recognition (CVPR)* (pp. 427–436). IEEE, 07.06.2015–12.06.2015.
- O'Donovan, P., Leahy, K., Bruton, K., & O'Sullivan, D. T. J. (2015). An industrial big data pipeline for data-driven analytics maintenance applications in large-scale smart manufacturing facilities. *Journal of Big Data*, 2, 117. <https://doi.org/10.1186/s40537-015-0034-z>
- Oztemel, E., & Gursev, S. (2020). Literature review of Industry 4.0 and related technologies. *Journal of Intelligent Manufacturing*, 31, 127–182. <https://doi.org/10.1007/s10845-018-1433-8>
- Penumuru, D. P., Muthuswamy, S., & Karumbu, P. (2020). Identification and classification of materials using machine vision and machine learning in the context of industry 4.0. *Journal of Intelligent Manufacturing*, 31, 1229–1241. <https://doi.org/10.1007/s10845-019-01508-6>
- Pimenov, D. Y., Bustillo, A., & Mikolajczyk, T. (2018). Artificial intelligence for automatic prediction of required surface roughness by monitoring wear on face mill teeth. *Journal of Intelligent Manufacturing*, 29, 1045–1061. <https://doi.org/10.1007/s10845-017-1381-8>
- Qiu, T., Lai, X., & Ni, J. (2020). Machine learning based novelty detection methods for progressive stamping process health monitoring. In *Volume 2: Manufacturing processes; manufacturing systems; nano/micro/meso manufacturing; quality and reliability*. American Society of Mechanical Engineers, 09032020.
- Rato, T. J., & Reis, M. S. (2020). An integrated multiresolution framework for quality prediction and process monitoring in batch processes. *Journal of Manufacturing Systems*, 57, 198–216. <https://doi.org/10.1016/j.jmsy.2020.09.007>
- Shen, Y., Yang, F., Habibullah, M. S., Ahmed, J., et al. (2020). Predicting tool wear size across multi-cutting conditions using advanced machine learning techniques. *Journal of Intelligent Manufacturing*, 26, 213. <https://doi.org/10.1007/s10845-020-01625-7>
- Slavič, J., Bolka, Š., Bratuš, V., & Boltežar, M. (2014). A novel laboratory blanking apparatus for the experimental identification of blanking parameters. *Journal of Materials Processing Technology*, 214, 507–513. <https://doi.org/10.1016/j.jmatprotec.2013.10.006>
- Toussaint, A. (2000). *Einfluss des Werkzeugverschleißes auf die Teilequalität beim Scherschneiden von Elektroblechen*, Dissertation, Technische Universität München.
- Tsai, C.-W., Lai, C.-F., Chao, H.-C., & Vasilakos, A. V. (2015). Big data analytics: A survey. *Journal of Big Data*, 2, 1170. <https://doi.org/10.1186/s40537-015-0030-3>

- Wang, L., & Gao, R. X. (2006). *Condition monitoring and control for intelligent manufacturing*. Springer.
- Wu, F.-C., & Chyu, C.-C. (2004). Optimization of correlated multiple quality characteristics robust design using principal component analysis. *Journal of Manufacturing Systems*, 23, 134–143. [https://doi.org/10.1016/S0278-6125\(05\)00005-1](https://doi.org/10.1016/S0278-6125(05)00005-1)
- Wuest, T., Weimer, D., Irgens, C., & Thoben, K.-D. (2016). Machine learning in manufacturing: Advantages, challenges, and applications. *Production and Manufacturing Research*, 4, 23–45. <https://doi.org/10.1080/21693277.2016.1192517>
- Zhou, S., & Jin, J. (2005). Automatic feature selection for unsupervised clustering of cycle-based signals in manufacturing processes. *IIE Transactions*, 37, 569–584. <https://doi.org/10.1080/07408170590929036>
- Zhou, Y., Sun, B., Sun, W., & Lei, Z. (2020). Tool wear condition monitoring based on a two-layer angle kernel extreme learning machine using sound sensor for milling process. *Journal of Intelligent Manufacturing*, 61, 457. <https://doi.org/10.1007/s10845-020-01663-1>

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