



Perspectives on data-driven models and its potentials in metal forming and blanking technologies

Mathias Liewald¹ · Thomas Bergs² · Peter Groche³ · Bernd-Arno Behrens⁴ · David Briesenick¹ · Martina Müller² · Philipp Niemietz² · Christian Kubik³ · Felix Müller⁴

Received: 22 September 2021 / Accepted: 4 February 2022 / Published online: 2 March 2022
© The Author(s) 2022

Abstract

Today, design and operation of manufacturing processes heavily rely on the use of models, some analytical, empirical or numerical i.e. finite element simulations. Models do reflect reality as best as their design and structure may appear, but in many cases, they are based on simplifying assumptions and abstractions. Reality in production, i.e. reflected by measures such as forces, deflections, travels, vibrations etc. during the process execution, is tremendously characterised by noise and fluctuations revealing a stochastic nature. In metal forming such kind of impact on produced product today in detail is neither explainable nor supported by the aforementioned models. In industrial manufacturing the game to deal with process data changed completely and engineers learned to value the high significance of information included in such digital signals. It should be acknowledged that process data gained from real process environments in many cases contain plenty of technological information, which may lead to increase efficiency of production, to reduce downtime or to avoid scrap. For this reason, authors started to focus on process data gained from numerous metal forming technologies and sheet metal blanking in order to use them for process design objectives. The supporting idea was found in a potential combination of conventional process design strategies with new models purely based on digital signals captured by sensors, actuators and production equipment in general. To utilise established models combined with process data, the following obstacles have to be addressed: (1) acquired process data is biased by sensor artifacts and often lacks data quality requirements; (2) mathematical models such as neural networks heavily rely on high quantities of training data with good quality and sufficient context, but such quantities often are not available or impossible to gain; (3) data-driven black-box models often lack interpretability of containing results, further opposing difficulties to assess their plausibility and extract new knowledge. In this paper, an insight on usage of available data science methods like feature-engineering and clustering on metal forming and blanking process data is presented. Therefore, the paper is complemented with recent approaches of data-driven models and methods for capturing, revealing and explaining previously invisible process interactions. In addition, authors follow with descriptions about recent findings and current challenges of four practical use cases taken from different domains in metal forming and blanking. Finally, authors present and discuss a structure for data-driven process modelling as an approach to extend existing data-driven models and derive process knowledge from process data objecting a robust metal forming system design. The paper also aims to figure out future demands in research in this challenging field of increasing robustness for such kind of manufacturing processes.

Keywords Metal forming · Sheet metal blanking · Data science · Machine Learning · Explainable artificial intelligence

✉ Mathias Liewald
mathias.liewald@ifu.uni-stuttgart.de

¹ Institute for Metal Forming Technology (IFU), University of Stuttgart, Holzgartenstraße 17, 70174 Stuttgart, Germany

² Laboratory for Machine Tools and Production Engineering (WZL), RWTH Aachen University, Campus-Boulevard 30, 52074 Aachen, Germany

³ Institute for Production Engineering and Forming Machines, Technical University of Darmstadt, Otto-Berndt-Straße 2, 64287 Darmstadt, Germany

⁴ Institute of Forming Technology and Machines (IFUM), Leibniz University Hannover, An der Universität 2, 30823 Garbsen, Germany

1 Introduction

A model is a consciously constructed reproduction of reality and is based on structures, functions or analogies. Models are used to solve tasks, where direct operations on the original are possible only with difficulty, are too costly or not expedient [1]. In production technology, the design and operation of manufacturing processes relies on different types of models for process layout, observation, improvement and control. However, solving these tasks does require a detailed multi-level procedure starting with engineering and designing manufacturing equipment allowing to integrate measuring devices, sensors and additional actuators to a certain extent. Also, suitable sensors and signal amplifier have to be specified to ensure the sufficient quality of raw process data. Using comprehensive data management, a reduction of the data dimension usually must be carried out in order to generate suitable data sets enabled by modern feature engineering techniques [2].

In the following literature review, we cover these aspects by displaying current trends in developing data-driven models for different objectives in sheet and bulk metal forming. Also, recent research is briefly discussed to illustrate and to disclose current use of white-, grey- and black-box models in this field of manufacturing. Also feature extraction and feature selection plays an important role in data science, since feature extraction can be performed by time, frequency or mixed domains. The presented overview tries to introduce current trends in using neural networks and other tools in artificial intelligence (AI) in the field of process control and prediction of metal formed or blanked part quality. It is motivated by the fact that using state-of-the-art models such as numerical process simulation in engineering and operating metal forming processes are limited. Unfortunately, a sophisticated explainability of effects and impacts of uncritical process noise and critical process fluctuations on the component quality appear unreachable. Literature in that field unfortunately disclose limited and only incremental progress in performing numerical process simulation codes during last two decades in terms of prediction capabilities of part quality and explainability of transient process interactions. For that very reason, the authors recognise a future-oriented approach in supplementing existing numerical process simulations and sensitivity calculations by data-driven (*black-box*) models based on heuristic empirical knowledge. This is due to particular challenge in metal forming, that numerous physical effects (e.g. wear, elastic deflections, thermal expansion) often show a simultaneous and spatially resolved nonlinear effect on forming stages or the process sequence at all and thus also on the part quality [3]. These complex interrelationships are accompanied by

various impacts of the process noise, e.g. batch variations of the ingoing semi-finished product, which is sequentially processed in one or more forming tools. Currently, these so far unexploited effects can be observed in large amounts of process data captured by multiple sensors [4]. The main chapter of this contribution for that reason contains four use cases to highlight specific challenges in gaining sufficient process data, an efficient feature extraction and reliable interpretation of process outcome. Thus, the final discussion in this context raises the question to which extent process data gained from previous production runs and process simulations can be used for data-driven modelling by informed machine learning algorithms [5]. Next to the integration of domain knowledge, newly available explainable artificial intelligence (XAI) may help to increase transparency, interpretability and explainability [6] to derive an optimised design of forming tools and its active surfaces to be manufactured for the next serial production.

2 State of the art

2.1 Classification

Volk et al. reviewed the classification and characterisation of models, focusing on the field of metal forming technology [7]. To identify and evaluate proper models for specific purposes, an evaluation of capability of existing models in terms of the following criteria supports the selection procedure. *Accuracy* describes the deviation between real and modelled output and *precision* refers to consistent results and low spread [8]. *Execution* and *response time* of models do evaluate the computational efforts and latency. *Model robustness* refers to a constant precision of results under different boundary conditions. *Transferability* and *adaption of models* for different and non-intended tasks is described by *flexibility*. *Explainability* and *transparency* of models are evaluated by the degree of knowledge gain. Production engineers therefore mainly are interested to increase the degree of process knowledge gain, which may base on physically influenced models.

A *white-box model* bases completely on physical and mathematical correlations and therefore provides a high degree of knowledge gain and transparency. Those deterministic models, based on valid physical correlations, may often reduce complexity with the aid of simplifying assumptions to reach suitable response time and often accompanied by lack of accuracy. In the field of metal forming, such closed mathematical descriptions and semi-analytical methods are limited for simple load cases, e.g. the estimation of bending processes [9]. Nevertheless, these models do serve as a basis for the development of numerical process simulations by e.g.

finite element method (FEM). Hybrid modelling involving physical and empirical knowledge, so called *grey-box models* are commonly used and further developed over the last decades in many use cases in the tool and process design of metal forming applications. Therefore, recent developments of grey-box models in this field do consider the description of general physical phenomena like friction or impact of varying material properties in combination with fundamental physical experiments. Current and future developments address the modelling of series production effects by stochastic methods and an increased computational effort [10]. On contrary, empirical *black-box models* are based exclusively on data [2]. Correlations between input and output can be determined by stochastic or physically supported methods or self-learning algorithms. This results into high performant data-driven models regarding accuracy of the prediction of process outcome, execution and response time but appears disadvantageous with respect to transparency and gain of knowledge [2]. In the field of metal forming, black-box models based on process data may help to describe complex phenomena in series production like wear or material scatter and other superimposed effects without increased effort in [modelling](#) and computation.

2.2 Data acquisition, transformation and modelling

During last few years, great efforts have been made to digitise metal forming manufacturing processes and to analyse process characteristics in real time or based on historical data to predict their outcome and to improve process understanding [11]. However, in a metal forming process the complexity in many cases is increased by the variety of operating principles, which hinders the procedure of knowledge discovery from one process to another.

In manufacturing industries, the quality of a produced product is often characterised by data with respect to material and geometric specifications. Also, the product quality can be associated with corresponding process data. Recent developments were made in the combination of product and process data and the data acquisition gained from sensors, actuators and the forming machine itself. Their transformation onto mathematical spaces for data modelling may allow in the future a reduction of error propagation within process sequence and to increase process robustness. [12]

2.2.1 Data acquisition

Data sets are usually captured by sensors gained from actuators, drives, deformation or force gauges and other electrical components consuming energy. Measurement normally is based on direct or indirect measurement principles [13]. Direct measurement requires capability of the sensor to directly capture the desired physical value. In

indirect measurement, sensors are used to determine indirect values of corresponding physical measures, whose correlation to the actual desired physical value is known through a non-deterministic transfer function [14]. In manufacturing processes, a multitude of sensors, e.g. pressure, temperature, speed, position sensors and others are used to acquire real-time status information from different physical actions and effects [15], being observed in the e.g. forming machine or tool. Considering the technological context of metal forming technologies, each forming process is characterised by specific measurement conditions, which pose technology-related challenges for data acquisition that have to be taken into account. For instance, hot forming processes such as drop forging can be characterised by high temperature and forming forces as well as long motion paths of workpiece volume on the active die surface during the process [16]. Therefore, the sensors to be used must be suitable for corresponding temperature or force ranges. The occurring temperatures as well as high mechanical loads during forging are applied in a cyclic manner resulting into alternating loads. High stroke rates and correspondingly high frequencies are often also associated [17] to a harsh forging manufacturing environment, which requires high speed and high volume data recording [18]. In addition, the potential spray cooling or dust poses yet further challenges to capture the exact process status in [hot forging](#) technologies [19]. Comparable constraints in [data acquisition](#) are found also in other bulk forming technologies. Commonly and in another field of forming technologies, metal blanking, forces result from peak loads when the tool hits the workpiece and the material breaks are measured. This leads to nonlinear and transient time series that represent the physics of these processes. Thereby, high accelerations and short tool engagement times have to be considered by integrated sensors into blanking tools as well as designing the measurement chain [20]. In this context, the sensor type as well as its location in the tool or process should be taking into account [21].

Referring to the term “garbage in–garbage out” in the context of data processing, high demands should be placed on data acquisition, validity check of data as well as on data quality [22]. Data quality is often comprised of multiple dimensions [23] covering accuracy, consistency, time series data und completeness as being the most commonly referenced and widely respected dimensions [24] of data featuring. Considering those dimensions, it needs to be taken into account that, having a poor data quality results into poor data output [25]. This leads to the necessity of an evaluation of collected data for quality levels [18]. In addition to the challenge of ensuring good data quality, three main aspects in terms of variety, volume and velocity data acquisition must be addressed [23].

- *Variety* The data acquired originates from a multitude of sources and has a heterogeneous character [26]. In general, data can be divided into structured, semi-structured and unstructured types of data. While structured data is tagged and ready to be sorted as well as analysed, unstructured data is random, which makes it difficult to process [27].
- *Volume* The sensors and other measurement devices normally do generate time-dependent series data during operation, which increasingly accumulate into very large data sets. Therefore, distributed storage of massive sensor data is required [15].
- *Velocity* The expression velocity depends primarily on the speed of data acquisition as well as the reliability of data transmission. Besides that, the efficiency of data storage and the speed of discovery of useful knowledge also needs to be considered when use of data is aimed on advanced feature engineering objectives [18].

Given these aspects, the vast amount of data sets, which also include a great amount of data contents having many attributes, in fact do make manual data analysis unfeasible [28]. Automated methods and approaches for data transformation and modeling are needed, which are presented hereafter.

2.2.2 Data transformation

Following recent trends in digitisation of manufacturing processes and the related optimisation of sensor systems as well as the development of high-performance measurement software, companies in the field of metal forming technology are faced with heterogeneous nature of data captured in manufacturing due to different data types, properties or sources.

Bellmann [29] describes this problem from the perspective of machine learning as the "curse of high dimensionality" in order to point out the fact that algorithms that work well in low dimensions become intractable with high dimensional inputs. The complexity of models grow exponentially with higher dimensionality (number of features), which obviously hinders the generalisability of such models [30]. Furthermore, an advanced data acquisition procedure results into a data space including redundant features adding no further information to the model while drastically increasing computational effort of the learning algorithm. Thus, feature reduction techniques aim to reduce the dimensionality as well as remove noise and redundancy of given data without removing relevant information [31]. For practical application of black-box modelling approaches, this data transformation is divided into the two steps of feature extraction and feature selection (see Fig. 1).

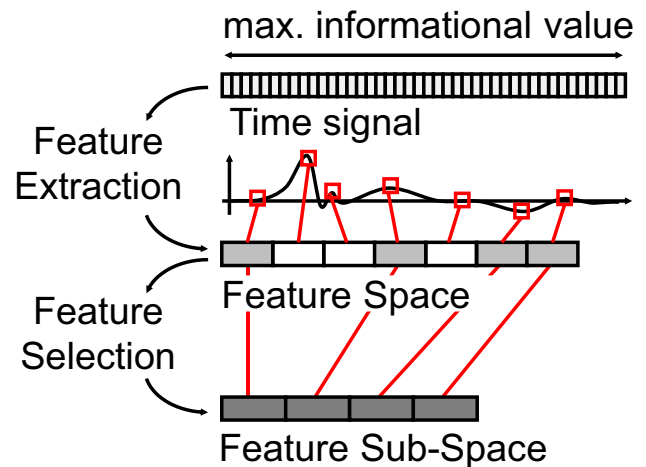


Fig. 1 Reduction of dimensionality by feature extraction and feature selection

During feature extraction, the dimensionality of data is automatically reduced by a transformation operation. However, this leads to a reduction of the engineering interpretability of extracted features and a loss of information. According to Li [32], extracting features from time signals is performed either from the time domain, the frequency domain, the time–frequency domain or it is based on a model approach. Features in the time domain do not require a special transformation operation and can be determined directly from the given data set. Mostly, these features are statistical parameters such as extreme values, the statistical moments of first to fourth order such as mean values, standard deviations, skewness, kurtosis or the root mean square [33]. Spectral and frequency analyses can be used to transform the time signals into the frequency domain, where spectral features such as the power spectrum, maximum frequencies, spectral entropy can be determined. Since in forming technology mainly sensor-based data such as forces, accelerations or torques etc., which have a transient signal characteristic, are acquired, the use of conventional spectral analysis techniques is limited [34]. In this context, advanced transformation techniques in the frequency domain such as the Haar and Hilbert–Hung transformation techniques are becoming increasingly important [35, 36]. Since spectral transformations in the time domain show poor resolution and information is lost when transforming to the frequency domain, approaches from the time–frequency domain such as the wavelet transform or the Wigner–Ville distribution do offer the possibility to consider features from both domains simultaneously [37]. In addition, model-based approaches for the feature extraction in manufacturing processes, which transform a high-dimensional data for visualisation into a lower two- or three-dimensional space, can more and more be found in literature. The most important examples

are autoregressive models, principal component analysis (PCA) and t-distributed stochastic neighbour embedding approach (t-SNE). One of the most promising learning techniques for dimensionality reduction is the Uniform Manifold Approximation and Projection, which is based on the t-SNE approach, and receives increasing attention in literature due to its low computational cost and good graphical separability of classes [38].

In the following step of feature selection, irrelevant features are removed and an optimised feature space for the model is derived. In addition to the use of automated algorithms, feature selection techniques often use a heuristic approach and consequently challenge the engineer with high demands. Therefore, selecting relevant features manually requires a deep process knowledge [39]. In principle, techniques of feature selection approaches can be divided into filter, wrapper as well as hybrid methods. When applying the filter method, a criterion is defined which quantifies the information value of single features for the performance of the model. The selection of the relevant features is performed by statistical evaluation of the correlation between the used feature and the performance of the model. In contrast, wrapper methods do not quantify the informational value of each feature for the model, but estimate the performance of the model selecting a certain feature subset. Hybrid methods do combine the advantages of filter and wrapper methods. While filter methods exclusively investigate the influence of specific features on model performance and therefore are computationally efficient. Wrapper methods establish correlations between model performance and a specific feature space, which needs high computational effort [40].

Table 1 gives an overview of feature extraction and selection approaches in metal forming technology which currently belongs to the state of the art. This survey shows that in a manufacturing context the transformation of data is mainly

based on statistical features. Thereby, these features are used for monitoring the actual process state by thresholds, envelopes or flat lines. Since transformation approaches in the time or time–frequency domain are mainly suitable for periodic or transient signals, they are rarely used for applications in the forming technology. Although the use of model-based transformation approaches has a great potential to reduce the amount of data with very little loss of information, they are also rarely implemented in real production environment. This is mainly due to the fact that the generated features are not interpretable by qualified personnel, but represent the physics of the process through abstract features. Here, techniques of explainable artificial intelligence (XAI) may help to derive correlations between these abstract features and the physical state of the process and thus provide acceptance for the increased use of such transformation approaches.

2.2.3 Modelling

After data transformation, models solely based on data are used to identify correlations between the transformed data (features) and the system output, which is typically experimentally measured. Especially the recently emerging *black-box models* generally show a performant capability in determining and representing those correlations. However, the disadvantage of this approach is often given by a high computational effort in finding the most useful model and, depending on its type, a gradual loss of explainability [2]. Here, explainability refers to the amount of reliable gain of knowledge and allows deeper insights into the phenomena acting in processes [7]. While common black-box models like regressions and support vector machines (SVM) are comprehensible and therefore widely used in the production engineering environment, the results generated by a neural network training are usually no longer transparent for the

Table 1 Sample of feature extraction and selection methods in forming technology

Feature extraction			
Time domain	Frequency domain	Time–frequency domain	Model-based
Blanking (force and displacement) [41]	Deep drawing (AE) [44]	Blanking (acceleration and force) [47]	Forging (force) [50] Stamping (force) [51]
Roll forming (temperature and force) [42]	Blanking (acceleration and force) [45]	Progressive stamping (force) [48]	Fine blanking (force) [52]
Blanking (force and AE) [43]	Deep drawing (AE) [46]	Blanking (acceleration and force) [49]	Blanking (force) [53]
Feature selection			
Wrapper method			Filter method
Blanking and bending (force) [54]			Blanking (force) [55]
Forging (force) [50]			Roll forming (force and rotational speed) [56]

Table 2 Selection of use cases for black-box models in forming technology

classifying
use cases – support vector machine
<ul style="list-style-type: none"> • fine blanking – roll height prediction based on FE-calculated training data [65] • stamping – identification of critical process states (poor workpiece quality) by monitoring press force signals [48]
prognostic
regression
<ul style="list-style-type: none"> • sheet metal forming – forming force prediction based on other process parameters and flexible rolling – quality prediction based on sheet thickness measurement [56] • sheet metal forming – spring back control by evaluation of press force and tool geometry [66] • deep drawing and bending – real-time process model based on historical force data to predict part quality of cutting, operations [67] • metal forming – estimating product-to-product variations in metal forming using force measurements [68]
neural network
<ul style="list-style-type: none"> • hot steel rolling – defect prediction by evaluation of self-organizing maps [69] • bulge testing – identification of material parameters using pressure displacement curves [70] • incremental sheet metal forming – machine learning-based parameterization of local support in robot-based incremental sheet forming [71] • flow forming – wear prediction based on experimental block-on-ring testing [57]

user [2], which opens future demands in extended research in the field of an explainable black-box-modelling.

A selection of typical examples for use cases of black-box models in the field of metal forming production engineering are given in Table 2. These so-called ‘deep learning approaches’ are indispensable with regard to further performance improvements, since they allow to determine correlations for universal problems without specific proximity functions. Belfiore et al. [57] for example presented a novel approach for the calculation of abrasive wear in grinding processes using neural networks. Analogous to the Archard model, the variables contact pressure, sliding velocity and temperature are employed to determine a functional relationship for wear estimation but without presuming the initial equation of Archard. This approach was able to achieve a high prediction quality with regard to wear evaluation without the restriction of an analytic model given by the need of proximity functions. However, it must be mentioned that the neural network has already been provided with pre-selected data channels for training, which are known to correlate highly with wear. Yet this circumstance generally encounters the typical problem in industrial applications that already today ‘any’ data is often recorded, which may not offer any significance for the modelling with regard to the result variable. For this reason, common scientific approaches in metal forming for other investigation objectives in production

engineering often include the evaluation of highly relevant process parameters, for example the press force or a measurement of produced parts or tool geometries. With models derived this way, it was possible in individual cases to achieve a high level of predictive quality for the considered problem. Due to the specific use of the black-box models, however, the transferability to other specific applications is hindered, since efforts have so far been limited to demonstrate the general applicability. Following this statement, there is a lack of a commonly accepted understanding of which practical guidelines black-box modelling approaches have to follow in order to be firmly adopted into the repertoire of knowledge-building techniques.

For future investigations, the importance of explainability appears extremely high for targeting a wide acceptance of models and new dimensions of process related calculations. A more far-reaching but practical example, in which traceability of calculations is required by law, is given by the development of the open accessible *code_aster* FE-calculation kernel, which was initially intended for the design of nuclear facilities [58]. In doing so, the authors see this example as transferable to manufacturing processes of the future, which with increasing complexity will also very likely have to withstand increasingly complex legal testing processes. To overcome this general issue, two main solutions appear conceivable. Initially, novel modelling approaches can be

supported by existing domain knowledge, but this limits the expected development steps. The research field of XAI-algorithms do present promising potential by trying to transform the results of neural networks back into an explainable dimension space by means of further processing techniques [59].

Recently developed XAI models like the SHAP Model introduces the possibility to automatically generate new "surrogate parameters" from the individual inputs [60]. Other models such as the DEEPLIFT use a scoring approach to identify relevant features for dimension reduction [61], while the LIME approach aims for an explanation of the initial black-box prediction through representation of local approximations [62]. However, according to literature, the evaluation of time series or sensor data sets presents a major challenge, since a time signal itself is unsuitable as an input parameter for a learning process and thus requires segmentation (feature engineering) [63]. Hu et al. [64] concluded that error deviations of less than 10% can be achieved when using perfectly time-synchronised data sets. In the field of metal forming technology, this time segmentation is well implementable via a time trigger synchronised e.g. to the top dead center of the press ram. In this way, individual features can be extracted from the data signals and related to a defined number of cycles or strokes, whereby the time series challenge described in the literature can be circumvented. Despite the steadily increasing number of new model approaches, *XAI modeling* is considerable as a relatively new approach in the scientific world and no prominent examples of its application in the field of metal forming are known up to date. However, there are high expectation to cover location and time dependent phenomena in complex forming processes with the help of prospectus XAI models. Compared to the capabilities of the models shown so far, high model prediction accuracies can be expected for the evaluation of serial forming processes, which is supported by the following use cases in the field of forging and blanking technology.

3 Practical approaches of data science in metal forming

As a result of the state of the art, the authors recognise in the long term a high potential of data-driven models in answering questions in metal forming technology far beyond numerical process simulation. The traditional and commonly used approach for gaining engineering knowledge about metal forming process interactions so far consists of building a feed forward model, starting with domain knowledge based on experience, physical analysis and numerical simulation code. Following conventional capability of

numerical models and calculation methods for achieving low-scrap production in sheet and bulk metal forming technologies, respective improvements in prediction accuracy of the expected part quality in fact will be needed in future. This target can be achieved by the use of data-driven models based on large quantities of process data recorded from continuously running series production processes. Furthermore, the incorporation of existing knowledge by informed machine learning technologies ensures an enhanced model quality and explainability. By following this approach, a reusable knowledge gain that does not require a time consuming and experienced based representation of formalised knowledge is enabled. However, the challenge here is to find a new generalised methodology with which the described modelling approaches can be implemented. In the following section, authors do present four exemplary use cases in which multiple aspects, current limitations and potentials of data-driven modelling were applied to achieve mentioned goals on the long term by focusing on conventional hot forming and metal cutting processes.

3.1 Hot forging

Bulk forming is one of the oldest disciplines in production and metal forming technology. New process management strategies address a higher forging process capability by means of reduced press force levels and preheating of the components. However, the additional temperature dependent process control results in conflicting objectives arising during the manufacturing of components. On the one hand, preheating of the material and tools lead to a significant thermo-mechanical load onto the tool active surface, which during series production leads to an increase of wear. On the other hand, the controlled temperature management in the components volume offers a high potential for the specific adjustment of e.g. lifetime properties and thus for an overall increase of quality of component. The following chapter therefore presents two current research trends, in which digitisation concepts are used to improve the prediction and assurance of component quality and tool wear under consideration of thermo-mechanical loading.

3.1.1 Digitisation in hot forging of processes of aluminum

The joint research project 'Increase of Performance of forging through development and integration of digital technologies—EMuDig 4.0' being performed at the Institute for Metal Forming Technology Stuttgart together with partners addressed the implementation of digitisation in bulk forming technologies. The publicly funded project aims to create a self-learning database for improved end-to-end product engineering and a significant increase in drop forging process capability. High requirements on forged components

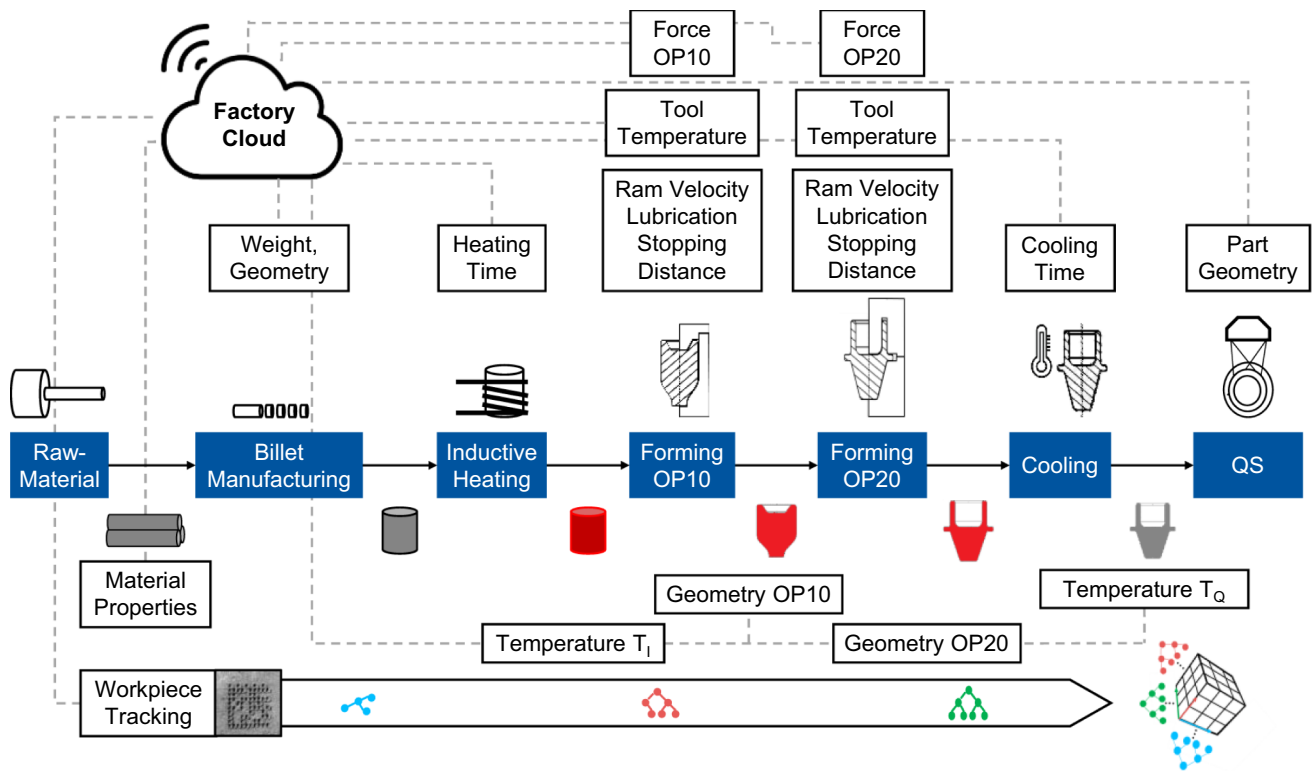


Fig. 2 Schematic representation of the controlled hot forging process (laboratory scale) showing flow of material and process data according to [72] and [73]

and corresponding tools under varying conditions in production like in-going material composition or lubrication during pressing set high demands on the quality and explainability of a suitable model. To meet those requirements, a feed forward controllable and digitised process sequence of exemplary aluminum hot forging was initially designed and operated successfully under laboratory conditions. Afterwards, the elaborated data-driven modelling methodology was transferred to real series production processes of project partners in forging companies, in particular to improve process control and failure detection during manufacturing of steel components and aluminum wheel forgings efficiently.

Figure 2 provides a schematic representation of realised laboratory forging process chain showing also the flow of material and data. Delivered raw material was separated into billets while geometric and material properties were manually captured. After a controlled and supervised inductive heating stage, the aluminum billet subsequently was formed by a two-step hot forging process and conclusively annealed. The part handling between forming and optical measurement operations as well as lubrication was realised by a 6-axis robot. Real time process data management was supported by a factory cloud based work piece tracking aided by a newly developed online analytical processing (OLAP) data base [72].

Besides the development and implementation of sensors, actuators and automation components, a valid numerical modelling of realised two-stage forging process served as a knowledge and data base of tool design and process control. Here, both forming operations were modelled by 2D-finite element method, considering simplifications like part symmetry and rigid tool structure. Afterwards, results of more than 1,200 stochastic numerical experiments were used as the initial database of data-driven modelling both forging stages to identify relevant actuated variables and first estimations of most sensitive parameter correlations.

With the help of the numerical gained synthetic forging domain knowledge, a recurrent neuronal network with auto-encoder structure (RNN) was pre-structured for each forming operation (see Fig. 3). This informed machine learning technology provided explainable correlations between input and output parameters such as in-going material properties, heating time, dimension of workpiece and microstructure as an advantage on the one hand. On the other hand, batch variations of raw materials, process fluctuations and numerical modelling assumptions, data-driven model accuracy was not sufficient enough for proper part quality prediction and process control. Therefore, real heterogeneous process data gained from production experiments were recorded and unified by a programmable logic controller (PLC). Figure 3

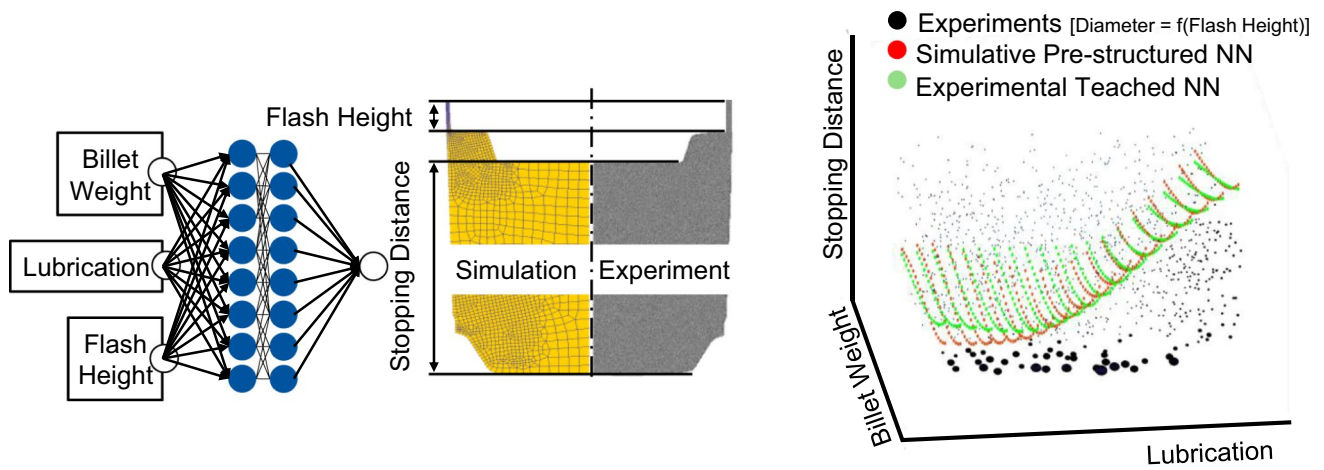


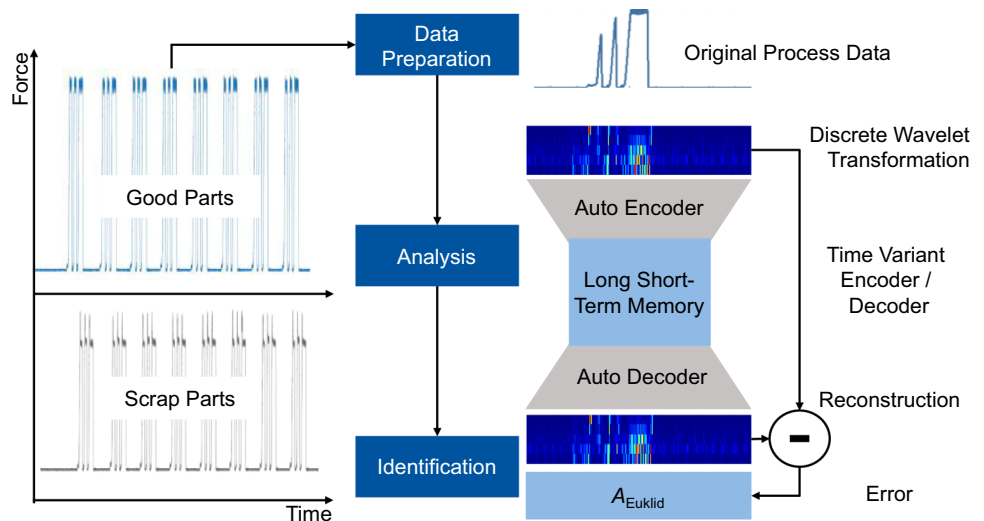
Fig. 3 Simulative pre-structured and afterwards experimental refined recurrent neuronal network (RNN) of preform stage according to [74]

represents an iteratively refinement of the knowledge based RNN by experimental data, thereby increasing model predictive quality after only a short time and number of strokes [74]. Now, the data-driven model overcomes existing lack of accuracy. Furthermore, implemented control logic on laboratory scale suggested autonomous decisions for each part and different materials and a change of the process sequence according to specific situations in order to keep the window of tolerances of the final product as small as possible [73].

The previously described and elaborated method developed within the framework of the joint research project EMuDig 4.0 was transferred for quality assurance and control objectives to two different forging companies. Both industrial partners belonged to the project consortium in order to support the transfer of elaborated method into real

forging processes such as car steering components made of steel and a hot forging process sequence of aluminum wheel rims. To detect process anomalies on data-driven modelling such as underfilling and other kinds of scrap, a long short-term memory network (LSTM) for time dependent behavior combined with a sequence-to-sequence network for expected part quality prediction objectives was applied. By means of feature extraction on process force signals with discrete-wavelet-transformation (DWT) and Autoencoder, a high data volume reduction and minimum loss of information for various recorded heterogeneous process signals was achieved. In doing so, the numerical pre-structured data-driven model was capable to judge between nominal and abnormal process signals (see Fig. 4), an automated suggestion of countermeasures for part quality robustification and to additionally predict wear of active tool surfaces as a correlation of data reconstruction error [75].

Fig. 4 Automated scrap detection by a long short-term memory in series production based on process anomalies [75]



Conclusively, this project carried out an enhanced data-driven model for quality prediction of hot forged parts by an informed machine learning, based on domain knowledge of numerical simulation and enhanced by process data. This enhancement exhibits a typical black-box characteristic which makes comprehensibility difficult and impede the use in forging tool and process design. This raises research questions in terms of transparency and explainability of data-driven models. The contained information in collected process data like correlations, anomalies and interactions may be uncovered by newly emerged data science technologies like XAI. In future work, process and material fluctuations might be considered without such an increased effort in numerical modeling, which makes data-driven modelling accessible for a robust tool and process design in metal forming.

3.1.2 Tool life prediction in hot forging determined by digital process data enhanced by FE-simulation

The dimensional precision of forged component mainly is joined with the durability of active surface of forging tools, process condition and scatter of material properties. The large scale production of forgings until today severely do suffer under unavoidable wear and structural damage due to extreme hot forging process conditions such as high cyclical mechanical as well as thermo-mechanical load applied onto active surface of forging tool. Subsequently the service life of the tools in **hot forging** is typically subjected to significant fluctuations [76] as depicted in Fig. 5. Reasons for that can be found in various causes of failure, e.g. crack formation propagating from active surface of the tool into its depth, and abrasive wear, which occurs unpredictably due to fluctuations of process parameters in general.

In manufacturing, there are no in-situ measurable parameters available for proper defect identification or prognosis.

Tool changes or tool rework are carried out often in advance and based on experience. On contrary, FE-simulations previously were used to calculate individual types of wear in isolation and rely on specific and mostly empirical models. A well-known white-box model is given by the Archard model for calculating abrasive wear. This model enables a comprehensible calculation of wear locations by evaluating local contact stresses, material sliding velocity and local tool hardness [77]. While the conventional Archard model considers the tool hardness as constant, current literature agrees that cyclic thermomechanical loading to the surface leads to extensive changes of microstructure and hardness of the layer during the running production [78, 79]. With regard to a data-driven modelling of this phenomenon, a test methodology was developed at the Institute of Metal Forming and Forming Machines (IFUM) to realistically replicate this load accumulation using a forming dilatometer to later analyze the resulting changes in the hardness of forging tool material [80, 81]. These measurements generate detailed and applicable domain knowledge that reduces uncertainties of existing models to predict local wear effects in the forging tool more reliable. Gained results finally were processed into hardness evolution curves ($H=f(T,\sigma,n)$), so that a realistic changes of hardness values can be estimated over the course of the forging operation. Process planners in industry do need such a function considering calculated process peak temperature T and comparative stress σ in conjunction with a specified number of forging cycles n . This function was successfully implemented into common FE codes by means of self-developed user subroutines and now can be directly applied to assess wear in more detail after a typical process calculation of a hot forging process [82].

However, we must acknowledge, that the results of a FE calculation model always represent one particular outcome or set of results based on a specific set of model and process parameters originating from a complex system of

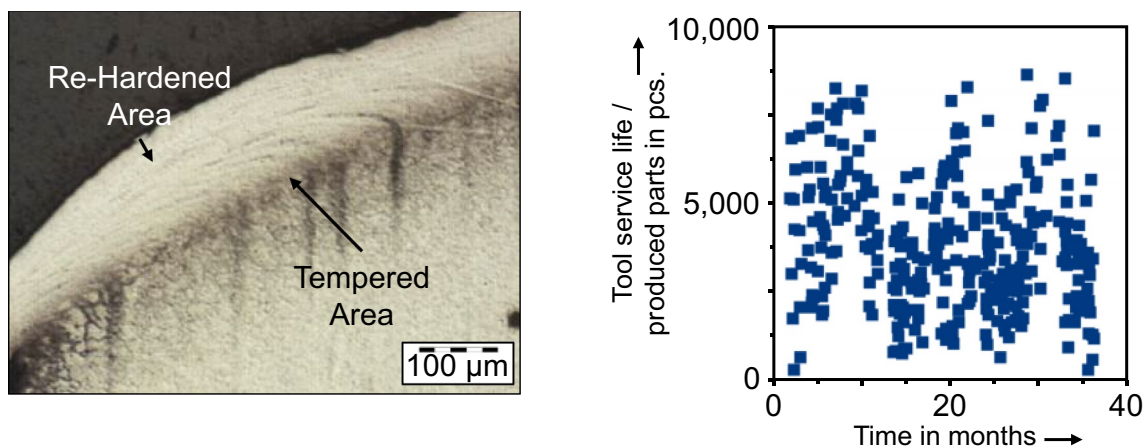


Fig. 5 Microstructural changes in the die edge layer during hot forging (left) and fluctuations in the tool life over a production period (right)

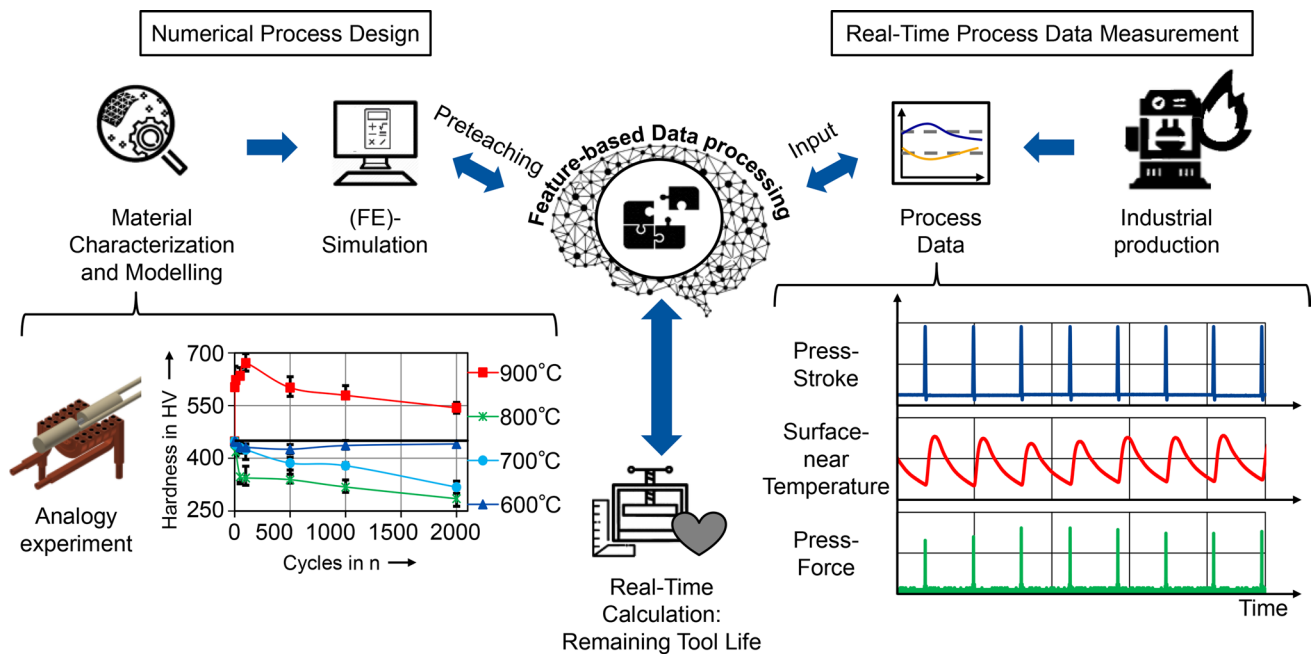


Fig. 6 Concept of a feature-based data processing for real time tool life calculation based on numerical process design and real time data measurement

interactions in reality. Due to lack of data coverage and fluctuations in the process (see Fig. 5) Errors or deviations from the calculated results are hardly predictable and explainable by FE-models. A new approach to overcome such problems is expected by the authors to be achieved in future research by using data-driven models, which can lead to a significant improvement of wear prediction capabilities in forging process engineering. To generate suiting input data, the continuous recording of extensive sensor data (for instance of stroke paths, press forces, temperatures, etc.) is of high interest. These data could serve as input for a *dynamic* feature processing model (Fig. 6) that provides a continuously updated tool life based on a pre-taught prediction data processing model. Pre-teaching is carried out on the one hand by incorporating the domain knowledge available for the specific forming operation, which is specifically integrated by means of FE simulations or material charts experimentally gathered as demonstrated by a recent examples in the literature [83]. On the other hand, XAI algorithms could be used to transform the features and mechanisms of the data models into a comprehensible domain, which in turn will enable a deeper understanding for further optimisation of tooling surfaces or process settings. Following this approach, a basis for future process enhancements is generated, which for example allows the implementation of a continuously updated tool life counter directly at the press control center. Therefore, the wear mechanism has to be precisely modelled considering the full range of process conditions as well as material of tool and part. The future approach of data-driven

modelling may fulfill the requirements of tool life prediction and enable an increased explainability of wear mechanisms and apparently stochastic tool life.

3.2 Digitisation in sheet metal blanking technologies

Shear cutting in general as well as fine-blanking of sheet metal both belong to the group of material separating processes that share a similar operating principle. Beyond that, both processes do have in common that the economic efficiency of the process is significantly influenced by the wear of the blanking tool components, which directly impacts the punched workpiece quality. However, in situ wear detection itself poses a challenge, since the tool remains completely closed during the shearing process of the sheet metal, thus emerging physical effects and technological interactions cannot be observed visually and assessed directly. Recent research proposes to measure the wear indirectly during the running process. Indirect measures of signals such as forces or vibration show considerable variations on a stroke-to-stroke basis as well as long-term trends, but are influenced not only by effect of wear, but also by e.g. inhomogeneous in-going material properties or dynamics of the machine tool [84]. Together with the sensitivity of the sensor opposes challenges to potential wear monitoring systems that need to filter important wear related information from auxiliary signals and derive estimators for specific wear effect in complex tool geometries. The first presented study shows

a visualisation of long-term trends in force and mechanical vibration signals captured during running fine blanking processes. This is followed by a further study presenting a new approach to predict the current wear state in shear cutting within an experimental environment.

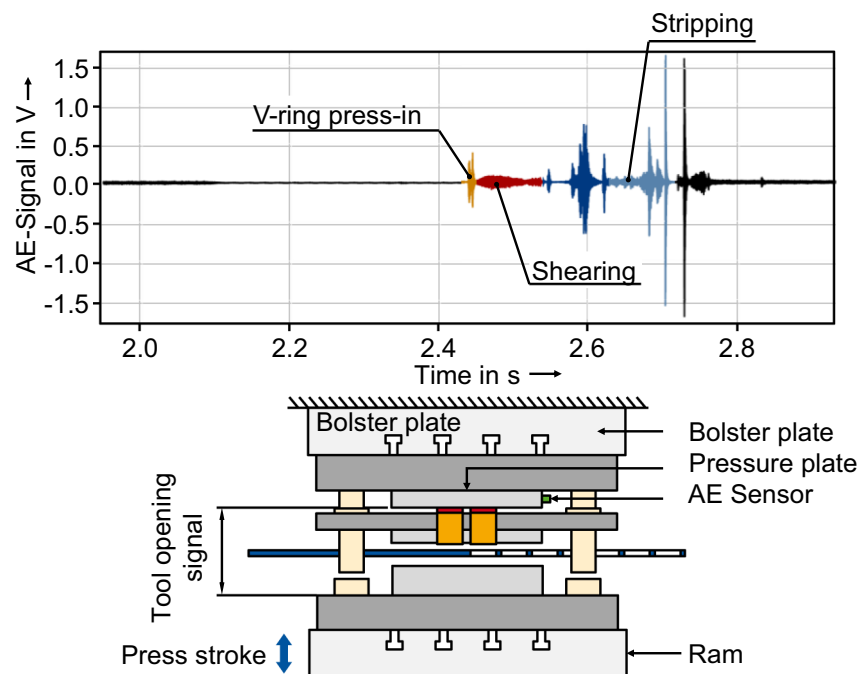
3.2.1 Long- and short-term variations in fine-blanking process data

Fine-blanking as well as shear cutting processes in general can, once setup correctly, be regarded as stable processes that can produce consistent high-quality outcomes in mass production. However, when executing extended stroke series, quality and process variations occur due to varying material properties and increasing wear of active tool components, both affecting the process performance. To understand occurring variations of the signals, the quantification and individual analysis of the fluctuations on a stroke-to-stroke basis as well as long-term trends is a prerequisite to discover their causal relationship to specific physical conditions defining the process, e.g. wear of tool components.

In the used fine-blanking press machine of this study, nine piezoelectric sensors have been integrated into the tool structure to measure the relevant process forces, namely punch, counterpunch and part holder force [52]. In addition, acoustic emission sensors have been applied to the upper pressure plate close to the punch positions [85]. These sensors detect elastic structural deflections as complex vibration signals, which are emitted from various points in the tool structure and press machine (Fig. 7). Force [86] as well

as acoustic sensors [87] have proven to capture important information about development of wear of tool components such as rounding emerging at cutting edges of punch and matrices during the life time of tool. Each fine-blanking operation was represented by a number of time series representing sensor signals during the operation from start to end. In this study each time series consisted of at least 10,500 data points that were regarded as input features originating from the blanking operation for subsequent analysis purposes. To reduce the number of features and to increase data point density, feature engineering describes the process of manipulating the raw sensor signals to extract, select and to condense relevant features for further processing steps. In the presented case, the raw sensor signal was first cleaned and subsequently segmented into phases that represent the stages during the shearing operation, such as blanking and stripping. In a next step, features of different feature domains, namely spectral, temporal and statistical features, were extracted using feature templates from the time series feature extraction library (TSFEL) [33]. Using the predefined feature templates, a total of 360 features, partly highly correlated, have been extracted for each stage of the forming process. To further reduce the redundancy and the amount of data, the dimensionality reduction technique uniform manifold approximation and projection (UMAP) [88] was utilised. UMAP is based on manifold learning techniques and ideas from topological data analysis and is basically to project the dimension of large dimensional data into lower dimensions while preserving the local structure of the signal.

Fig. 7 Mounting position of sensors in fine-blanking and the corresponding acoustic emission signal



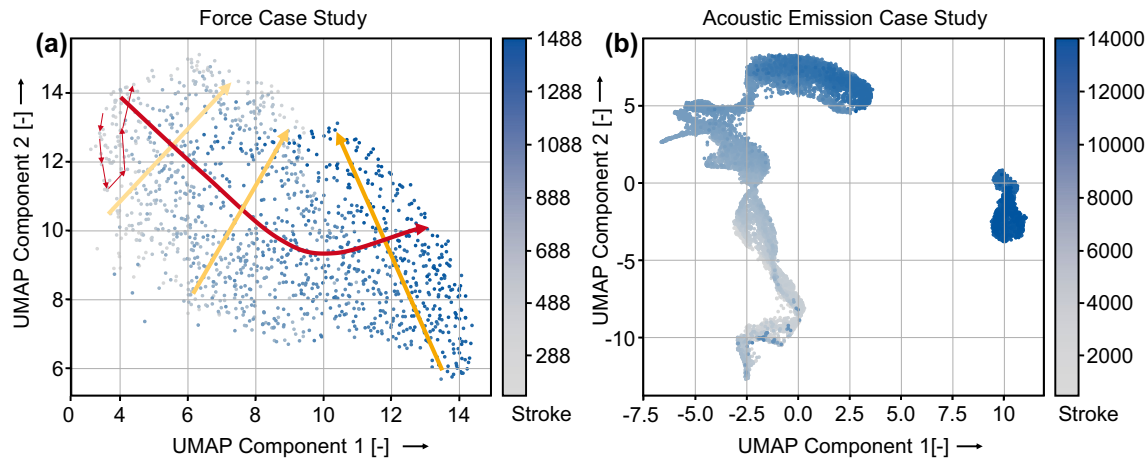


Fig. 8 UMAP Models trained on **a** force signals of 1488 consecutive strokes, with stroke-to-stroke based differences (thin-red), long-term trends (red) and local variance (yellow), **b** acoustic emission signals of 14,000 consecutive strokes

The presented feature engineering approach has been applied to two case studies on fine blanking including force measurements of 1488 consecutive strokes as well as the variation in acoustic emission (AE) signals over the course of 14,000 strokes in an industrial setting of a blanking process. The data sets were modeled by use of an unsupervised non-linear projection algorithm UMAP [88]. In this context, UMAP embeds the signals extracted features into two-dimensional Euclidean space for visualisation purposes by computing two dimensionless components. Results achieved in this study are presented in the two-dimensional plots in Fig. 8 that show the embedding of the force and acoustic emission at 1488 and 14,000 consecutive strokes, respectively. Both signals show a long-term trend of the derived two-dimensional representations, correlating with the stroke index (specifically features from time, statistical and frequency domain of the acoustic emission data), as well as stroke-to-stroke based and short-term variations (specifically force data). This embedding was also found in higher dimensions and is researched to understand the presented variations, to quantify them, e.g. by stroke-wise Euclidean distance, and to correlate this derived metric with increase of wear or other phenomena observable during process execution. Recent studies already indicate that this variation indeed can be connected to tool wear phenomena [89].

The presented plots are found representative for several series of experiments conducted with the fine-blanking process. For further analysis, reliable metrics that can grasp the amount of variation on a stroke-to-stroke basis [85] as well as long-term trends [90] need to be researched in detail.

The current results gained in this study are limited due to the following reasons: (1) The UMAP approach as well as other non-linear projection methods are only marginally explainable. Thus, an automated extraction of knowledge about which specific data points are important for the

projection is missing and limiting the value of the approach. (2) Neural networks do offer the possibility to automatically learn and extract compact features and representations of process signals but require large quantities of sample data for reliable training. The current data sets are large, but still do not fit to the amount needed for deep neural network training. This can be countered by using data augmentation techniques together with domain knowledge to support neural networks during the learning process and lower the amount of data needed for training. (3) The projections presented above utilize only one sensors data, while several more sensor data are available. Potentially, each forming operation can potentially be represented by sets of sensors based on force, mechanical vibration, tensions or power consumption. Additionally, data from the process itself can be further complimented by sensors detecting varying material properties online during process execution that further increases the complexity to the analysis. Finally, iv) the results have to be validated by measuring process outcome and wear increase of active tool components and complimented by white-box model that are already available, instead of relying only on (explainable) black-box models. In order to meet mentioned challenges and to overcome listed limitations in the near future, the future potential of XAI models to understand, and eventually reduce the amount of variations and trends in the execution of sheet metal forming processes can be leveraged.

3.2.2 Wear prediction during blanking using a multiclass support vector machine

Almost every product manufactured by sheet metal forming involves one or more cutting operations within its manufacturing process. Especially due to increasing requirements for processing high-strength materials combined with increasing production rates, wear is a phenomenon that significantly

influences the quality and productivity of this process [91]. In this context, workpiece quality is influenced by severe wash out of sheared edges as well as outline of workpiece, productivity of process is influenced by tool break down time, increased maintenance and high risk of breakage of highly loaded punch corners.

Currently, this challenge is addressed by conventional monitoring systems that provide binary control of the process state (e.g. scrap part/good part) [92]. However, predicting the amount of wear at cutting tools during series production is currently not possible. In this context and in a preliminary work, a black-box model based on a support vector machine (SVM), which allows for prediction of current wear state based on ‘in-process force-signals’ was developed [53]. Therefore, five abrasive wear states were characterised by the increase of cutting edge radii r_i of the blanking tool during processing and are estimated by a classification model based on a SVM. Force signals were acquired with a sampling rate of 90 kHz by multisensory blanking tool on a BRUDERER mechanical blanking press operated at a stroke rate of 200 strokes per minute using two different sensor types, a strain gauge applicated at the press frame **S** and a piezo electrical force washer integrated to the upper part of the tool (**P**). For each wear state, 100 experiments (time series with 4950 data points) were conducted per wear state leading to an aggragated force matrix F whereby $F \in \mathbb{R}^{600 \times 4950}$. Figure 9 shows the experimental setup and a representative time series of each sensor type. In addition, it was necessary to link gained process data with a label quantified by the cutting edge radii r_i (quality data) of the blanking tool. Since the wear state was not measured for each captured time series and a small amount of labeled data is combined with a large amount of unlabeled data to train the model, the case shown is considered as a semi-supervised learning approach. A cold rolled steel (1.0347) with a thickness of 2 ± 0.002 mm was used for the experiments.

Fig. 9 Experimental setup for the data acquisition within a multisensory blanking cutting tool [53]

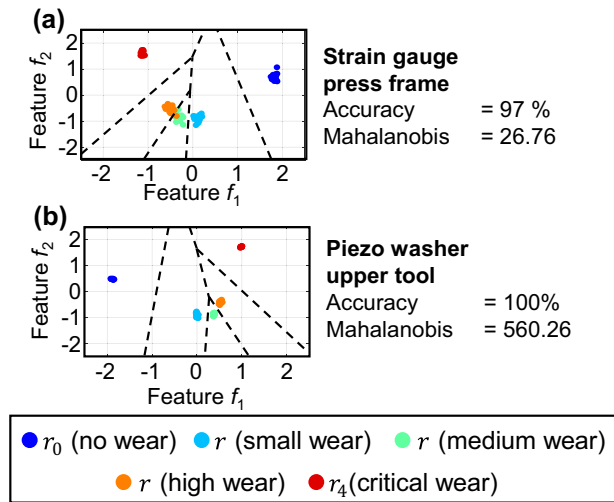
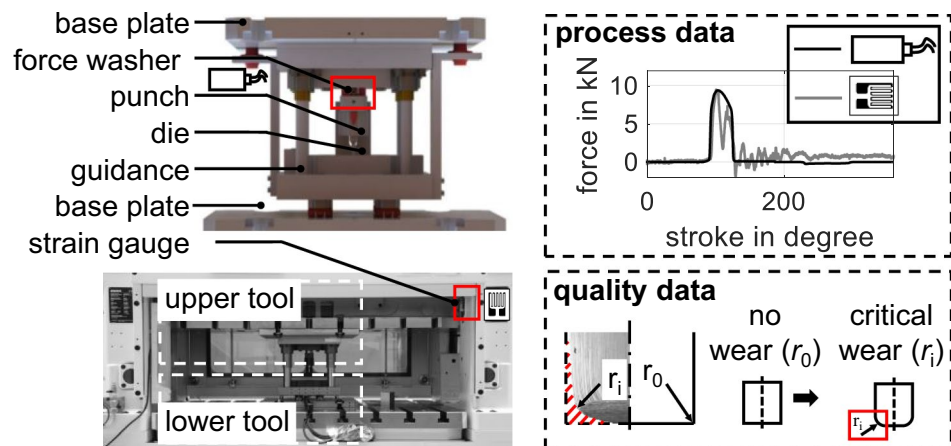


Fig. 10 Qualitative visualisation of multiclass SVM classifying five wear states comparing different sensor position

Due to the high stroke rates of industrial blanking processes the amount of data provided for modeling increases rapidly. To reduce the amount of data with as little loss of information as possible the model-based transformation technique PCA was used. For further investigations only these two principals were used as input for the SVM. Finally, to show that a classification of the wear state during blanking is possible, the SVM model was trained based on these two principal axis from the force signals. A grid search afterwards was used to optimise the hyper parameters of the SVM and to find a suitable kernel function (kernel function: linear; regularisation parameter: 4.9321; margin of tolerance: 0.0042). In order to validate the model, the data set was split into a training data set (80%) and a test data set (20%).

Figure 10 shows the results of the classification model delivering reliable results of both types of sensors. To quantify the performance of the number of correctly predicted wear classes in relation to the total number of predictors, as well as the separability of the classes quantified by the “Mahalanobis distance” was determined [93]. Thereby, the wear state classified by the model based on the signal of the strain gauge sensor shows a twenty times worse separability than the model based on the signal of the piezo electrical sensor. This is due to the physical distance of strain gauge sensor towards the actual forming zone as well as the tendency of this sensors type to electrical noise that is superimposes the physical relevant component of the forces signal. Despite the worse signal quality gained from the strain gauge sensor, an intelligent transformation of the data in combination with optimised SVM improves the accuracy of the model to 97%.

However, it must be conceded that the black-box models are able to quantify process states based on sensorial acquired data, but fail to derive knowledge for a system improvement. Especially, an abstract description of non-linear interlinkages through black-box models between varying process states such as wear, tool and press elasticity or material fluctuation and acquired process variables is difficult to interpret from an engineering point of view. However, to be able to identify system improvements for the reduction of varying process states it is necessary to explain the used black-box models as well as the transformation techniques and their system knowledge that is stored in terms of abstract models or features. Especially, these correlations between features or the model itself to physical states of the blanking process are enabled by XAI. This allows to generalise black-box models by creating an understanding of the interdependencies between the process state, extracted features and the structure of the model. In addition, XAI can help to select robust data based on features to ensure a model-based description of production systems via black-box models even with varying process parameters and increasing uncertainty.

4 Research objectives for an explainable data-driven modelling in metal forming

The complexity of metal forming processes is essentially caused by the contact conditions in the deformed zone between the workpiece and the tool. Furthermore, the variation within time and location, depending on the geometry, the state of the process and the material properties increase the level of complexity. As a result of those process conditions, the microstructure of the component and its entire geometry in fact do change after each forming step. In addition, high force densities in bulk forming or large area loads in sheet metal forming, for example, can cause considerable

elastic deformation of the acting structure. An accurate physical modelling of these deformations is limited, but in reality, a detrimental effect on the surfaces between the die and the workpiece under load must be acknowledged. In many forming operations, these interactions are always superimposed by unavoidable process noise, i.e., transient process effects, so that their overall stochastic effects on the final component quality are often not comprehensible and their causality remains unclear.

Indicated by the aforementioned use cases, the current challenge to grasp and to model the stochastic nature of forming process by using black-box approaches yield to promising intermediate results. This motivates further investigations on the combination of existing domain knowledge with new data-driven models. Specifically, presented results lack the interpretability to validate the plausibility of black-box model behaviour. A combination of white-, grey- and black-box approaches, together with cutting edge methods to enhance the explainability of data-driven models based on production as well as experimental data sets may overcome the current limits in modelling stochastic effects. Furthermore, it may improve the understanding and design of forming processes and tools. This can be achieved by taking explicitly available heuristic expert knowledge, numerical calculation results and digitized information on the semi-finished product, the condition of the forming tool and press machine as well as transient process conditions into account.

In this context, the following scientific questions contain consecutive hypotheses and subsequently define important milestones along future work in this field:

1. How can sufficiently accurate digital representations of metal forming operations be described mathematically and combined with domain-specific knowledge in an explicit manner?
2. How can the combined use of explicit knowledge and process data isolate process noise, decompose it into short- and long-term variations, quantify and relate it to key characteristics for influencing factors of the forming process and tool surface geometries?
3. How can the integration of the derived models into real production environments provide new and more explainable knowledge for a more efficient, demand-oriented determination of the process parameters and the effective surface of the die, and how can the explanatory power be increased?
4. How can the newly developed method and derived model be used to draw conclusions about process design to adjust system parameters based on a domain knowledge-supported, data-driven evaluation of process data?

For this purpose, use cases in Sect. 3 do represent those first working steps towards answering the above-mentioned

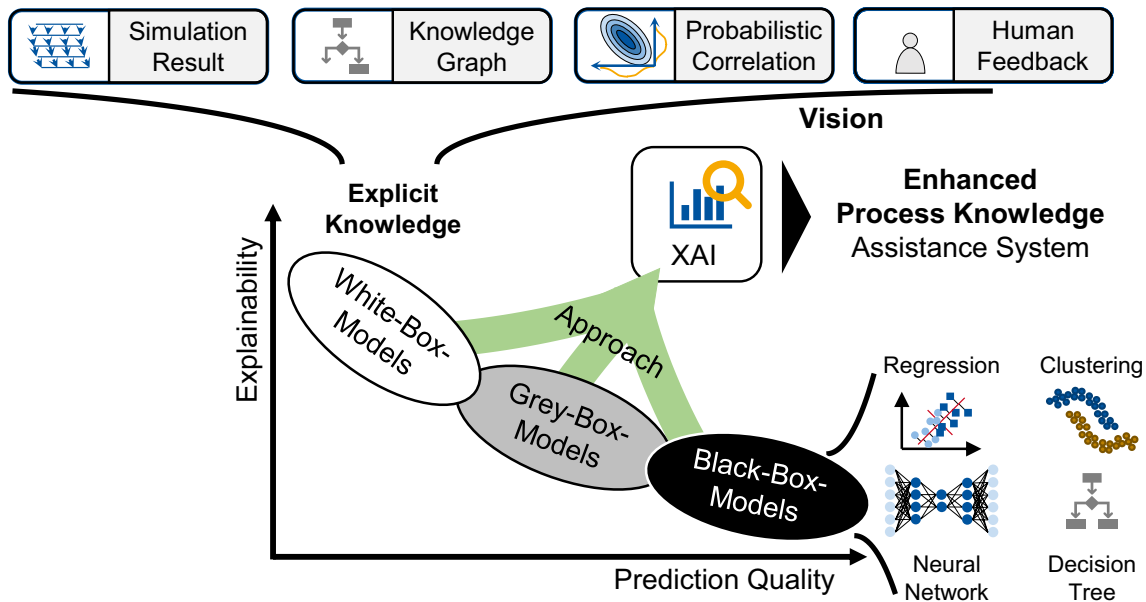


Fig. 11 Approach for an explainable and precise data-driven modelling to enhance process knowledge in forming technology

research questions on a laboratory scale. While the first and second research questions can be tackled by interdisciplinary research between mathematics, computer science and engineering departments alone, the third and fourth require extensive involvement of industry partners to gather data and sharpen the studied cases. For production engineering researcher, the claim is to disclose new scientific cause-effect relationships in industrial relevant metal forming processes. At the same time, characteristics of acquired process data can be explored and used to develop and compare new methods for the evaluation of these data. Finally, the model quality and transferability are assessed by a comparison with existing, state of the art methods for tool and process design as well as an application on new and unknown data sets of comparable forming processes. Thus, the need for research is to decode the interrelationships and isolate the scatter of different process stages, while at the same time exploring new, transparent, mathematical modelling approaches that exhibit high *interpretability* and *accuracy* (see Fig. 11). New data science techniques like XAI can be used for decoding and interpretable modelling of previously unknown dependencies. In doing so, new approaches for the optimization of the processes within one or several forming stages can be derived and new statements about process stability and tool design can be made. Only with the combination of explicit domain knowledge in various forms derived by years of research, aided by modern emerging tools in mathematics as well as computer science the potential of the increasing amount of field and experimental data can be leveraged.

Acknowledgements The authors would like to thank the German Federal Ministry for Economic Affairs and Energy within the framework of the Programme for Competition Digital Technologies for Business via the research project “EMuDig 4.0” and the project “Mittelstand 4.0 - Kompetenzzentrum Darmstadt”. Funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy–EXC-2023 Internet of Production–390621612. Furthermore, funded within the project DFG-397768783

Funding Open Access funding enabled and organized by Projekt DEAL.

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

References

1. Klaus G, Buhr M (1987) Philosophisches Wörterbuch. Bibliographische Institut, Westberlin
2. Loyola-Gonzalez O (2019) Black-box vs. white-box: understanding their advantages and weaknesses from a practical point of view. IEEE Access 7:154096–154113. <https://doi.org/10.1109/ACCESS.2019.2949286>

3. Fischer P, Heingärtner J, Renkci Y, Hora P (2018) Experiences with inline feedback control and data acquisition in deep drawing. *Procedia Manuf* 15:949–954. <https://doi.org/10.1016/j.promfg.2018.07.401>
4. Havinga J, Mandal PK, van den Boogaard T (2020) Exploiting data in smart factories: real-time state estimation and model improvement in metal forming mass production. *Int J Mater Form* 13(5):663–673. <https://doi.org/10.1007/s12289-019-01495-2>
5. von Rueden L et al (2021) Informed machine learning—a taxonomy and survey of integrating prior knowledge into learning systems. *IEEE Trans Knowl Data Eng.* <https://doi.org/10.1109/TKDE.2021.3079836>
6. Roscher R, Bohn B, Duarte MF, Garcke J (2020) Explainable machine learning for scientific insights and discoveries. *IEEE Access* 8:42200–42216. <https://doi.org/10.1109/ACCESS.2020.2976199>
7. Volk W et al (2019) Models and modelling for process limits in metal forming. *CIRP Ann* 68(2):775–798. <https://doi.org/10.1016/j.cirp.2019.05.007>
8. Selvik JT, Abrahamsen EB (2017) On the meaning of accuracy and precision in a risk analysis context. *Proc Inst Mech Eng Part O J Risk Reliab* 231(2):91–100. <https://doi.org/10.1177/1748006X16686897>
9. Heller B (2002) Halbanalytische Prozess-Simulation des Freibiegens von Fein- und Grobblechen. Dissertation, Universität Dortmund
10. Chenot J-L et al (2014) Recent and future developments in finite element metal forming simulation. In: 11th international conference on technology of plasticity
11. Lee J, Bagheri B, Kao H-A (2015) A cyber-physical systems architecture for industry 4.0-based manufacturing systems. *Manuf Lett* 3:18–23. <https://doi.org/10.1016/j.mfglet.2014.12.001>
12. Wang K-S (2013) Towards zero-defect manufacturing (ZDM)—a data mining approach. *Adv Manuf* 1(1):62–74. <https://doi.org/10.1007/s40436-013-0010-9>
13. Stavropoulos P et al (2013) Monitoring and control of manufacturing processes: a review. *Procedia CIRP* 8:421–425. <https://doi.org/10.1016/j.procir.2013.06.127>
14. Xu K et al (2020) Advanced data collection and analysis in data-driven manufacturing process. *Chin J Mech Eng.* <https://doi.org/10.1186/s10033-020-00459-x>
15. Bao Y et al (2012) Massive sensor data management framework in Cloud manufacturing based on Hadoop. In: IEEE 10th international conference on industrial informatics, pp 397–401. <https://doi.org/10.1109/INDIN.2012.6301192>
16. Klocke F (2013) *Manufacturing processes 4*. Springer, Berlin
17. Subramonian S, Altan T, Campbell C, Ciocirlan B (2013) Determination of forces in high speed blanking using FEM and experiments. *J Mater Process Technol* 213(12):2184–2190. <https://doi.org/10.1016/j.jmatprotec.2013.06.014>
18. Zhong RY, Newman ST, Huang GQ, Lan S (2016) Big data for supply chain management in the service and manufacturing sectors: challenges, opportunities, and future perspectives. *Comput Ind Eng* 101:572–591. <https://doi.org/10.1016/j.cie.2016.07.013>
19. Hawryluk M, Ziemba J (2017) Possibilities of application measurement techniques in hot die forging processes. *Measurement* 110:284–295. <https://doi.org/10.1016/j.measurement.2017.07.003>
20. 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(2):122–130. <https://doi.org/10.1016/j.measurement.2018.11.058>
21. Durand C et al (2021) Comparison of different sensor technologies to monitor a forging process. *ESAFORM 2021*. <https://doi.org/10.25518/esaform21.1475>
22. Mueller B (2012) Additive manufacturing technologies—rapid prototyping to direct digital manufacturing. *Assembly Autom.* <https://doi.org/10.1108/aa.2012.03332baa.010>
23. Song Z, Sun Y, Wan J, Liang P (2017) Data quality management for service-oriented manufacturing cyber-physical systems. *Comput Electr Eng* 64:34–44. <https://doi.org/10.1016/j.compeleceng.2016.08.010>
24. Wang RY, Strong DM (1996) Beyond accuracy: what data quality means to data consumers. *J Manag Inf Syst* 12(4):5–33. <https://doi.org/10.1080/07421222.1996.11518099>
25. Kilkenny MF, Robinson KM (2018) Data quality: “Garbage in—garbage out”. *Health Inf Manag J* 47(3):103–105. <https://doi.org/10.1177/1833358318774357>
26. Sagioglu S, Sinanc D (2013) Big data: a review. international conference on collaboration technologies and systems (CTS)
27. Zikopoulos P et al (eds) *Understanding big data—analytics for enterprise class hadoop and streaming data*. McGraw-Hill, New York
28. Braha D (ed) *Data mining for design and manufacturing—methods and applications*. Springer, Boston
29. Bellman R (1966) Dynamic programming. *Science* 153(3731):34–37. <https://doi.org/10.1126/science.153.3731.34>
30. Domingos P (2012) A few useful things to know about machine learning. *Commun ACM* 55(10):78–87. <https://doi.org/10.1145/2347736.2347755>
31. Kacprzyk J et al (2006) *Feature extraction*. Springer, Berlin
32. Li CJ (2006) Signal processing in manufacturing monitoring. In: Wang L, Gao RX (eds) *Condition monitoring and control for intelligent manufacturing*. Springer, London, pp 245–265
33. Barandas M et al (2020) TSFEL: time series feature extraction library. *SoftwareX* 11(10):100456. <https://doi.org/10.1016/j.softx.2020.100456>
34. Hoffmann R, Wolff M (2014) *Intelligente Signalverarbeitung 1*. Springer, Berlin
35. Huang NE et al (1998) The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. *Proc R Soc Lond A* 454(1971):903–995. <https://doi.org/10.1098/rspa.1998.0193>
36. Koh CKH, Shi J, Williams WJ, Ni J (1999) Multiple fault detection and isolation using the Haar transform, part 1: theory. *J Manuf Sci Eng* 121(2):290–294. <https://doi.org/10.1115/1.2831218>
37. Li X, Dong S, Yuan Z (1999) Discrete wavelet transform for tool breakage monitoring. *Int J Mach Tools Manuf* 39(12):1935–1944. [https://doi.org/10.1016/S0890-6955\(99\)00021-8](https://doi.org/10.1016/S0890-6955(99)00021-8)
38. Allaoui M, Kherfi ML, Cheriet A (2020) Considerably improving clustering algorithms using UMAP dimensionality reduction technique: a comparative study. In: El Moataz A, Mammass D, Mansouri A, Nouboud F (eds) *Image and signal processing*. Springer International Publishing, Cham, pp 317–325
39. Hira ZM, Gillies DF (2015) A review of feature selection and feature extraction methods applied on microarray data. *Adv Bioinform.* <https://doi.org/10.1155/2015/198363>
40. Aggarwal CC (2014) *Data classification*. CRC Press, Boca Raton
41. Cao J, Kinsey B, Solla SA (2000) Consistent and minimal spring-back using a stepped binder force trajectory and neural network control. *J Eng Mater Technol* 122(1):113–118. <https://doi.org/10.1115/1.482774>
42. Kubik C, Hohmann J, Groche P (2021) Exploitation of force displacement curves in blanking—feature engineering beyond defect detection. *Int J Adv Manuf Technol* 113(1–2):261–278. <https://doi.org/10.1007/s00170-020-06450-z>

43. Wadi I, Balendra R (1999) An intelligent approach to monitor and control the blanking process. *Adv Eng Softw* 30(2):85–92. [https://doi.org/10.1016/S0965-9978\(98\)00067-2](https://doi.org/10.1016/S0965-9978(98)00067-2)
44. Ubhayaratne I, Xiang Y, Pereira M, Rolfe B (2015) An audio signal based model for condition monitoring of sheet metal stamping process. In: *Industrial electronics and applications (ICIEA 2015)*, pp 1267–1272. <https://doi.org/10.1109/ICIEA.2015.7334303>
45. Zhang G, Li C, Zhou H, Wagner T (2018) Punching process monitoring using wavelet transform based feature extraction and semi-supervised clustering. *Procedia Manuf* 26(4):1204–1212. <https://doi.org/10.1016/j.promfg.2018.07.156>
46. Liang SY, Dornfeld DA (1990) Characterization of sheet metal forming using acoustic emission. *J Eng Mater Technol* 112(1):44–51. <https://doi.org/10.1115/1.2903185>
47. Ge M, Zhang GC, Du R, Xu Y (2002) Feature extraction from energy distribution of stamping processes using wavelet transform. *J Vib Control* 8(7):1023–1032. <https://doi.org/10.1177/107754602029577>
48. Zhou C et al (2016) An automatic process monitoring method using recurrence plot in progressive stamping processes. *IEEE Trans Autom Sci Eng* 13(2):1102–1111. <https://doi.org/10.1109/TASE.2015.2468058>
49. Zhang GC et al (2002) Bispectral analysis for on-line monitoring of stamping operation. *Eng Appl Artif Intell* 15(1):97–104. [https://doi.org/10.1016/S0952-1976\(02\)00007-6](https://doi.org/10.1016/S0952-1976(02)00007-6)
50. Lei Y, Zhang Z, Jin J (2010) Automatic tonnage monitoring for missing part detection in multi-operation forging processes. *J Manuf Sci Eng* 132(5):83. <https://doi.org/10.1115/1.4002531>
51. Jin J, Shi J (2000) Diagnostic feature extraction from stamping tonnage signals based on design of experiments. *J Manuf Sci Eng* 122(2):360–369. <https://doi.org/10.1115/1.538926>
52. Niemietz P et al (2020) Stamping process modelling in an internet of production. *Procedia Manuf* 49(15):61–68. <https://doi.org/10.1016/j.promfg.2020.06.012>
53. Kubik C, Knauer SM, Groche P (2021) 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. *J Intell Manuf* 23:1489. <https://doi.org/10.1007/s10845-021-01789-w>
54. Hoppe F et al (2019) Feature-based supervision of shear cutting processes on the basis of force measurements: evaluation of feature engineering and feature extraction. *Procedia Manuf* 34(5):847–856. <https://doi.org/10.1016/j.promfg.2019.06.164>
55. Lee WB, Cheung CF, Chiu WM, Chan LK (1997) Automatic supervision of blanking tool wear using pattern recognition analysis. *Int J Mach Tools Manuf* 37(8):1079–1095. [https://doi.org/10.1016/S0890-6955\(97\)88104-7](https://doi.org/10.1016/S0890-6955(97)88104-7)
56. Kirchen I et al (2017) Data-driven model development for quality prediction in forming technology. In: *15th international conference on industrial informatics*, pp 775–780. <https://doi.org/10.1109/INDIN.2017.8104871>
57. Belfiore NP et al (2007) A hybrid approach to the development of a multilayer neural network for wear and fatigue prediction in metal forming. *Tribol Int* 40(10–12):1705–1717. <https://doi.org/10.1016/j.triboint.2007.01.008>
58. Levesque JR (1998) The Code Aster: a product for mechanical engineers. <https://inis.iaea.org/search/searchsinglerecord.aspx?recordsFor=SingleRecord&RN=30019283>. Accessed 11 June 2021
59. Payrovnaziri SN et al (2020) Explainable artificial intelligence models using real-world electronic health record data: a systematic scoping review. *J Am Med Inform Assoc* 27(7):1173–1185. <https://doi.org/10.1093/jamia/ocaa053>
60. Lundberg S, Lee S-I (2017) A unified approach to interpreting model predictions. <http://arxiv.org/pdf/1705.07874v2>. Accessed 11 June 2021
61. Shrikumar A, Greenside P, Shcherbina A, Kundaje A (2016) Not Just a black box: learning important features through propagating activation differences. <http://arxiv.org/pdf/1605.01713v3>. Accessed 11 June 2021
62. Ribeiro MT, Singh S, G. C (2016) "Why Should I Trust You?": explaining the predictions of any classifier. <http://arxiv.org/pdf/1602.04938v3>. Accessed 11 June 2021
63. Schlegel U et al (2019) Towards a rigorous evaluation of XAI methods on time series. <http://arxiv.org/pdf/1909.07082v2>. Accessed 11 June 2021
64. Hu B, Chen Y, Keogh E (2015) Time series classification under more realistic assumptions. *Data Min Knowl Disc* 30:403–437. <https://doi.org/10.1137/1.9781611972832.64>
65. Stanke J et al (2018) A predictive model for die roll height in fine blanking using machine learning methods. *Procedia Manuf* 15:570–577. <https://doi.org/10.1016/j.promfg.2018.07.279>
66. Frayman Y, Rolfe BF, Webb GI (2002) Improving an inverse model of sheet metal forming by neural network based regression. In: *22nd computers*, pp 129–136
67. Ryser M et al (2021) Surrogate model-based inverse parameter estimation in deep drawing using automatic knowledge acquisition. *Int J Adv Manuf Technol* 117:997–1013. <https://doi.org/10.1007/s00170-021-07642-x>
68. Havinga J, van den Boogaard T (2017) Estimating product-to-product variations in metal forming using force measurements. In: *AIP conference proceedings* 1896. Penang, p 70002
69. Haapamäki J, Tamminen S, Rönning J (2005) Data mining methods in hot steel rolling for scale defect prediction. In: *International conference on artificial intelligence and applications (AIA 2005)*, pp 90–94
70. Chamekh A, Bel Hadj Salah H, Hamblil R, Gabbiche A (2006) Inverse identification using the bulge test and artificial neural networks. *J Mater Process Technol* 177(1–3):307–310. <https://doi.org/10.1016/j.jmatprotec.2006.03.214>
71. Möllensiepe D, Ohm M, Störkle DD, Kuhlenkötter B (2019) Experimental validation of smoothed machine learning-based parameterization of local support in robot-based incremental sheet forming. In: *Wulfsberg JP, Hintze W, Behrens B-A (eds) Production at the leading edge of technology*. Springer, Berlin, pp 483–492
72. Liewald M et al (2018) On the tracking of individual workpieces in hot forging plants. *CIRP J Manuf Sci Technol* 22(3):116–120. <https://doi.org/10.1016/j.cirpj.2018.04.002>
73. Liewald M, Karadogan C, Felde A, Ludwig R (2017) Development and integration of digital technologies in the forging process sequence. *New Dev Forging Technol* 2017:245–256
74. Lindemann B et al (2018) Cloud-based control approach in discrete manufacturing using a self-learning architecture. *Int Fed Autom Control Online* 51–10:163–168
75. Lindemann B, Jazdi N, Weyrich M (2019) Detektion von Anomalien zur Qualitätssicherung basierend auf sequence-to-sequence LSTM Netzen. at - utomatisierungstechnik 67(12):1058–1068. <https://doi.org/10.1515/auto-2019-0076>
76. Jarfors A, Castagne S, Danno A, Zhang X (2017) Tool wear and life span variations in cold forming operations and their implications in microforming. *Technologies* 5(1):3. <https://doi.org/10.3390/technologies5010003>
77. Archard JF (1956) The wear of metals under unlubricated conditions. *Proc R Soc Lond A* 236(1206):397–410. <https://doi.org/10.1098/rspa.1956.0144>
78. Krawczyk J, Widomski P, Kaszuba M (2021) Advanced complex analysis of the thermal softening of nitrided layers in tools during hot die forging. *Materials* (Basel, Switzerland). <https://doi.org/10.3390/ma14020355>
79. Singh K, Khatirkar RK, Sapate SG (2015) Microstructure evolution and abrasive wear behavior of D2 steel. *Wear* 328–329:206–216. <https://doi.org/10.1016/j.wear.2015.02.019>

80. Klassen A, Bouguecha A, Behrens BA (2014) Wear prediction for hot forging dies under consideration of structure modification in the surface layer. *AMR* 1018:341–348. <https://doi.org/10.4028/www.scientific.net/AMR.1018.341>
81. Behrens B-A et al (2020) Multi-layer wear and tool life calculation for forging applications considering dynamical hardness modeling and nitrided layer degradation. *Materials* (Basel, Switzerland). <https://doi.org/10.3390/ma14010104>
82. Behrens B-A et al (2020) A combined numerical and experimental investigation on deterministic deviations in hot forging processes. *Procedia Manuf* 47:295–300. <https://doi.org/10.1016/j.promfg.2020.04.231>
83. de Gooijer BM, Havinga J, Geijselaers HJM, van den Boogaard AH (2021) Evaluation of POD based surrogate models of fields resulting from nonlinear FEM simulations. *Adv Model Simul Eng Sci* 8(1):129. <https://doi.org/10.1186/s40323-021-00210-8>
84. Unterberg M et al (2019) In-situ material classification in sheet-metal blanking using deep convolutional neural networks. *Prod Eng Res Devel* 13(6):743–749. <https://doi.org/10.1007/s11740-019-00928-w>
85. Bergs T, Niemietz P, Kaufman T, Trauth D (2020) Punch-to-punch variations in stamping processes. In: *IEEE 18th world symposium on applied machine intelligence and informatics (SAMI)*, pp 213–218. <https://doi.org/10.1109/SAMI48414.2020.9108761>
86. Voss BM, Pereira MP, Rolfe BF, Doolan MC (2017) Using stamping punch force variation for the identification of changes in lubrication and wear mechanism. *J Phys Conf Ser* 896:12028. <https://doi.org/10.1088/1742-6596/896/1/012028>
87. Ubhayaratne I, Pereira MP, Xiang Y, Rolfe BF (2017) Audio signal analysis for tool wear monitoring in sheet metal stamping. *Mech Syst Signal Process* 85:809–826. <https://doi.org/10.1016/j.ymssp.2016.09.014>
88. McInnes L, Healy J, Melville J (2018) Umap: Uniform manifold approximation and projection for dimension reduction. *arXiv:1802.03426*
89. Niemietz P, Unterberg M, Trauth D, Bergs T (2021) Autoencoder based wear assessment in sheet metal forming. In: *40th international deep-drawing research group conference (IDDRG 2021)* (**accepted**)
90. Bergs T et al (2021) Wear monitoring in fine blanking processes using feature based analysis of acoustic emission signals. In: *Procedia CIRP 2021* (**accepted**)
91. Xing Z et al (2018) Simulated analysis and experimental investigation on edge qualities of high strength steels hot blanking parts. *Procedia Manuf* 15:619–626. <https://doi.org/10.1016/j.promfg.2018.07.286>
92. Jemielniak K (1999) Commercial tool condition monitoring systems. *Int J Adv Manuf Technol* 15(10):711–721. <https://doi.org/10.1007/s001700050123>
93. de Maesschalck R, Jouan-Rimbaud D, Massart DL (2000) The Mahalanobis distance. *Chemom Intell Lab Syst* 50(1):1–18. [https://doi.org/10.1016/S0169-7439\(99\)00047-7](https://doi.org/10.1016/S0169-7439(99)00047-7)

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