



Machine learning and deep learning based predictive quality in manufacturing: a systematic review

Hasan Tercan¹ · Tobias Meisen¹

Received: 22 December 2021 / Accepted: 5 May 2022 / Published online: 28 May 2022
© The Author(s) 2022

Abstract

With the ongoing digitization of the manufacturing industry and the ability to bring together data from manufacturing processes and quality measurements, there is enormous potential to use machine learning and deep learning techniques for quality assurance. In this context, predictive quality enables manufacturing companies to make data-driven estimations about the product quality based on process data. In the current state of research, numerous approaches to predictive quality exist in a wide variety of use cases and domains. Their applications range from quality predictions during production using sensor data to automated quality inspection in the field based on measurement data. However, there is currently a lack of an overall view of where predictive quality research stands as a whole, what approaches are currently being investigated, and what challenges currently exist. This paper addresses these issues by conducting a comprehensive and systematic review of scientific publications between 2012 and 2021 dealing with predictive quality in manufacturing. The publications are categorized according to the manufacturing processes they address as well as the data bases and machine learning models they use. In this process, key insights into the scope of this field are collected along with gaps and similarities in the solution approaches. Finally, open challenges for predictive quality are derived from the results and an outlook on future research directions to solve them is provided.

Keywords Industry 4.0 · Predictive quality · Machine learning · Deep learning · Manufacturing · Quality assurance · Artificial intelligence

Introduction

In the present era of Industry 4.0 and the digitization of the manufacturing industry, new technological possibilities are emerging that contribute to strengthen companies' competitiveness. Especially the combinations of new communication technologies with state-of-the-art methods from the fields of machine learning (ML) and deep learning (DL) enable promising applications for data-driven, smarter manufacturing (Shang & You, 2019; Tao et al., 2018). One area that can strongly benefit from these developments is quality assurance. It involves activities across the product lifecycle to ensure that the requirements on the quality of produced prod-

ucts are met (Hehenberger, 2020; Pfeifer & Schmitt, 2021). ML and DL methods offer ways to support these activities by enabling data-driven, automated quality analyses. Their application in this field is referred to as predictive quality (Nalbach et al., 2018; Schmitt et al., 2020b).

Predictive quality solutions are built upon data from the manufacturing process. By extracting recurring patterns from the data and relating them to quality measurements, predictive quality enables the data-driven estimation of the product quality based on process data. The estimations serve as a decision-making basis for quality enhancing measures, such as adjusting the process parameters for avoiding rejects (Schmitt et al., 2020b). The common approach to predictive quality includes four main steps: the formulation of the manufacturing process and target quality, the selection and collection of process and quality data, the training of a ML/DL model, and the use of the model for estimations as a basis for decisions (schematically illustrated in Fig. 1). In this context, predictive quality mainly comprises methods for supervised ML.

✉ Hasan Tercan
tercan@uni-wuppertal.de

Tobias Meisen
meisen@uni-wuppertal.de

¹ University of Wuppertal, Rainer-Gruenter-Strasse 21,
Wuppertal, Germany

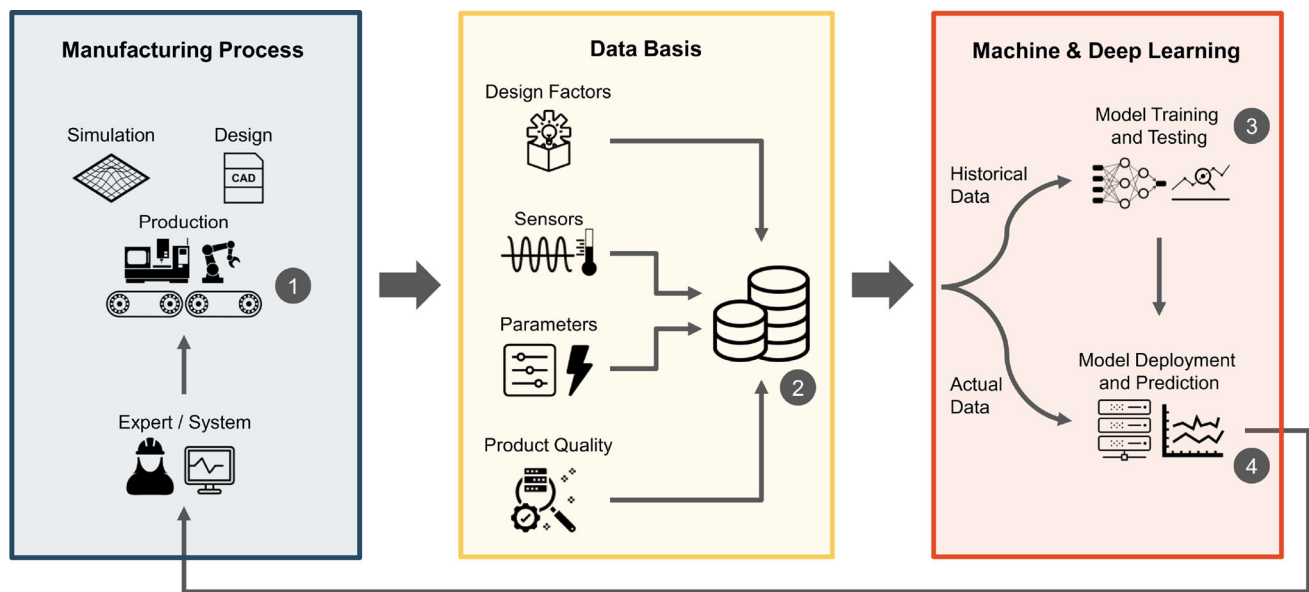


Fig. 1 Predictive quality approach: for a selected manufacturing process (1), relevant process and quality data is collected (2) and used as a basis for training a ML model (3). The trained model is used to perform quality estimations for decision support (4)

In the current manufacturing research, there exist many examples that demonstrate the feasibility of ML or deep learning based predictive quality, ranging from inline fault predictions (Mayr et al., 2019) to automated quality inspections (Schmitt et al., 2020a). The addressed manufacturing processes and quality criteria are manifold, such as the prediction of part cracks in deep drawing (Meyes et al., 2019), the estimation of roughness in laser cutting (Tercan et al., 2017; Zhang & Lei, 2017), or the detection of porosity defects in additive manufacturing (Zhang et al., 2019a). Though these use cases are different, their solution approaches have similarities in terms of the data and methods used. However, it is noticeable that often the use cases are considered in isolation, making it difficult to compare the proposed approaches. Hence, it is not clear where the overall predictive quality research currently stands, which methods are currently investigated, what the limitations are, and in which directions the research should go. In this paper, we address these issues by conducting a systematic review of the publications that address the field of predictive quality. We see the timing for such a review as appropriate because, on the one hand, there is a sufficiently large amount of published papers from which we can draw these insights. On the other hand, there are currently no adequate papers dealing with this topic in its entirety. Although studies with similar investigations exist (see “Related survey paper” section), they either cover a broader scope (e.g. applications of ML for the production context in general (Fahle et al., 2020; Sharp et al., 2018) or they are no longer up-to-date as their publication dates back several years (Köksal et al., 2011; Rostami et al., 2015).

Due to these observations, we define the primary goal of this review: providing a comprehensive overview of scientific publications from 2012 to 2021 that address predictive quality approaches in manufacturing. Our perspective on this field focuses on its common concepts depicted in Fig. 1. After collecting the relevant publications and building a corresponding corpus, we extract information about their use cases, the manufacturing processes and quality criteria they address and the data bases and ML methods they use. The goal is to categorize the publications along these concepts and to answer the following three driving questions

- **Q1: What are the addressed manufacturing processes and quality criteria?** Our goal is on the one hand to discover the scope of the field and the possible applications of predictive quality, and on the other hand to identify similarities and gaps in the domains.
- **Q2: What are the characteristics of the data used for model training?** Predictive quality is based on process data. We aim to find out the common data sources used in the publications, the variables selected for ML model training, and the modality of the data.
- **Q3: Which machine learning models of supervised learning are commonly trained?** We aim to discover which supervised learning problems are addressed, which models of ML and DL are trained for the quality estimations, and if they are compared with each other in the publications.

From these questions, we derive key insights and open challenges for predictive quality and provide an outlook on future

research directions that we believe will increase its prevalence in research and industry.

The paper is structured as follows. In “[Predictive quality](#)” section, we introduce predictive quality by first defining the term and then describing the underlying machine learning models and tasks. In “[Related survey paper](#)” section, we provide an overview of related survey papers that deal with similar topics. We then present the methodology underlying our study in “[Study methodology](#)” section. Here, we define the categories by which we evaluate the publications and formulate the search queries and criteria for selecting them. “[Results](#)” section presents the result of our study with regard to the stated questions. On the basis of these results, “[Key insights and challenges](#)” section provides key findings and identified gaps in predictive quality research. In “[Future research directions](#)” section, we envision future research directions to master them. “[Summary](#)” section summarizes this review.

Predictive quality

Regarding the use of data-driven methods with reference to product and process quality in production, different terms and terminologies are used to describe them in the state of research and technology. Examples are *data analytics*, *predictive analytics*, *machine learning for quality prediction* (Mayr et al., 2019) or *predictive model-based quality inspection* (Schmitt et al., 2020a). Predictive quality is another term that covers this broad field of research. The definitions by Schmitt et al. (2020b) and Nalbach et al. (2018) highlight two key aspects of predictive quality: product-related quality as well as the data-driven prediction of it. Schmitt et al. (2020b) define predictive quality as “enabling the user to make a data-driven prediction of product- and process-related quality” with the goal of “acting prescriptively on the basis of predictive analyses”. Nalbach et al. (2018) take a similar view – their definition of predictive quality includes methods that use data to “identify statistical patterns to foresee future developments concerning the quality of a product”. Although the term *prediction* is also very broad, it limits the range of use cases of data-driven methods. For example, it does not cover use cases for ML based quality inspection where the goal is to detect faults that have already occurred in the process rather than providing future-oriented quality predictions. With regard to data-driven methods, it is also the case that particularly machine learning and deep learning methods have been the central focus of research in recent years. In our consideration, we therefore also concentrate on these two fields.

Relying on these definitions, we define the term predictive quality as the basis for our literature search and specify the terminologies as well as their scope in the context of this paper:

Predictive quality comprises the use of machine learning and deep learning methods in production to estimate product-related quality based on process and product data with the goal of deriving quality-enhancing insights.

The following remarks provide further concretization of the definition:

- The term estimation includes prediction as well as classification of quality.
- The product-related quality can be a fixed quality parameter as well as known product faults.
- By process and product data, we mean product characteristics, process parameters, process states, and planning information

By our definition, predictive quality is not limited to the production phase of a product, but also is applied in the product or process planning phase - for example, in the design of a production process using test trials or a DoE-based feasibility analysis using process simulations. It should also be noted that our definition does not cover the notions of anomalies and anomaly detection. Anomalies in manufacturing are events that differ from normal behavior and as such are not initially associated with a known defect or quality degradation (Lopez et al., 2017). Anomaly detection therefore involves different methods than the ones considered in the context of this paper. For the sake of limiting the scope of our study, we exclude publications that deal with anomalies and anomaly detection.

As predictive quality involves the detection or prediction of quality, it mainly comprises methods of supervised learning. Supervised learning has the goal of estimating a numerical or categorical target variable on the basis of selected input variables (regression respectively classification). For supervised learning, methods of both fields machine learning and deep learning can be used:

- **Machine learning** methods such as regression analysis that detect linear and quadratic correlations and models that can handle complex and non-linear estimations such as support vector machine (SVM) and feed-forward artificial neural network (ANN) (also called multilayer perceptron (MLP)), and interpretable graphical models such as decision trees. In addition, a variety of ensemble methods (e.g. random forest) exist that comprise several individual models.
- **Deep learning** models that are based on deep ANNs and that have marked major milestones in the AI research in recent years. These include, for example, convolutional neural networks (CNNs) which are established in computer vision and image recognition (Redmon et al.,

2016; Szegedy et al., 2015), as well as recurrent neural networks such as long short-term memory (LSTM) (Hochreiter & Schmidhuber, 1997) or transformer networks, which represent the state of the art in natural language processing areas such as speech recognition or machine translation (Vaswani et al., 2017; Xiong et al., 2016).

The learning methods used for predictive quality are diverse and often depend on the purpose of the application. Following the discussions by Köksal et al. (2011) and Rostami et al. (2015) on the purposes of data-driven quality assurance, we list three tasks of machine and deep learning for predictive quality:

- **Quality description** The identification, evaluation and interpretation of relationships between process variables and product quality. The primary goal is to gain insights about interrelationships in the process.
- **Quality prediction** The model-driven estimation of a numerical quality variable on the basis of process variables. The goal here is the prediction of product quality, either for decision support or for automation.
- **Quality classification** Analogous to quality prediction, this involves model-driven estimation of categorical (binary or nominal) quality variables. An example is the prediction of certain product defect types.

In this paper, we study publications that primarily address these three tasks, focusing especially on quality prediction and classification. The use of such model estimations can be very diverse and range from pure knowledge gain for the user to automated feedback into the system. Generally, a process improvement is initiated which leads to the fulfillment of the specified product quality or its improvement. Possible improvements include the reduction of the reject rate through early intervention in the production process, the stabilization of the process for production in tighter tolerances (Schmitt et al., 2020a), or the optimization of process parameters (Weichert et al., 2019).

Related survey paper

A review similar to ours is provided by Köksal et al. (2011). The authors conducted an extensive literature review of data mining applications for quality improvement tasks in manufacturing and categorized them according the stated predictive quality tasks. Rostami et al. (2015) proposed a similar approach with a focus on applications of SVMs. Both studies date back several years and thus do not cover recent advances in this field. More recently, Weichert et al. (2019) reviewed machine learning applications for production pro-

cess optimization with regard to product- or process-specific metrics. The study shows that optimization approaches are mostly based on root-cause analysis and fault diagnosis in production plants or by the combination of ML methods with optimization methods. Although the study has overlaps with our work, the authors mainly address approaches to process optimization. In our study, we focus as previously described on approaches for quality estimation and evaluate them based on the data and methods used.

In the course of our literature search, we also identified survey papers that conduct similar investigations on machine learning applications but with a different scope. For example, there already exist extensive studies on the use of artificial intelligence and machine learning techniques in the production and manufacturing context (Fahle et al., 2020; Mayr et al., 2019; Shang & You, 2019; Sharp et al., 2018). Shang and You (2019) provided an overview of recent advances in data analytics for different application task areas such process monitoring and optimization. They also discussed the works in terms of usability and interpretability for control tasks. Fahle et al. (2020) and Mayr et al. (2019) studied machine learning applications in different task scenarios such as process planning and control. Sharp et al. (2018) focused on cross-domain applications in the product lifecycle. There are also surveys that deal with related research fields such as ML-based predictive maintenance (Dalzochio et al., 2020; Zonta et al., 2020), condition monitoring (Serin et al., 2020b) and machine fault diagnosis (Ademujimi et al., 2017). Our paper clearly differs from the mentioned papers as it reviews approaches that primarily address the quality of the products produced.

Study methodology

Our review is based on the guiding questions defined in “**Introduction**” section. First, we translated the questions into several categories which we used to categorize and summarize the publications. The categories are listed in Table 1.

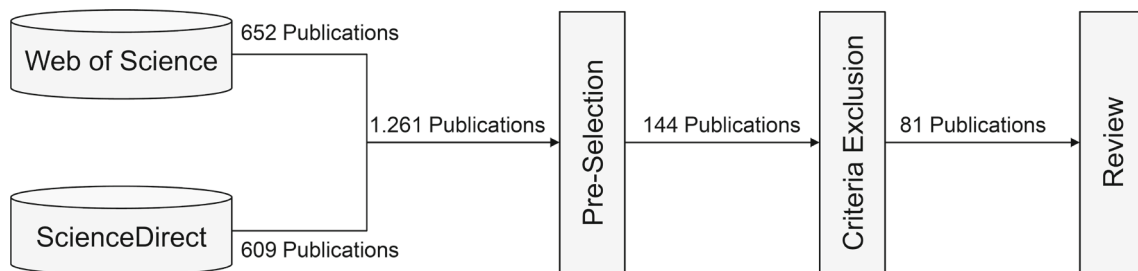
In the next step, we performed a literature search in the databases of Web of Science and ScienceDirect. As shown in “**Predictive quality**” section, there are different terms and terminologies for predictive quality that are used in publications. The same holds for the very broad application domain of manufacturing. In order to cover this broad scope in the search, we used different terms for the search queries and divided them into three categories: the predictive quality terminology, the machine learning field, and the manufacturing domain. Table 2 lists all defined search terms. We formulated search queries to find publications that contain at least one term from each of the three categories. In addition, we filtered the results based on the publication year from 2012 to 2021 (we performed the search on June 29, 2021). This

Table 1 Categories by which the publications are reviewed

Question	Category	Description
Q1	Use case	Main use case of paper and purpose of predictive quality
	Process	Addressed manufacturing process (e.g. laser welding, deep drawing)
	Category	Category of process according to DIN 8580 (e.g. cutting, forming)
	Criterion	Estimated quality criterion (e.g. product dimensions, OK/NOK quality)
Q2	Data source	Main source of process data (e.g. running production, simulation)
	Input variables	Data parameters used for the model training (e.g. sensor data)
	Data modality	Data types of gathered data (e.g. time series, categorical data)
	Data amount	Number of observations used for model training
Q3	Learning task	Formulated learning task (e.g. classification, regression)
	Prime model	Primarily used (or best performing) ML/DL model (e.g. SVM, CNN)
	Baselines	ML/DL-models used for comparison to prime model (e.g. SVM, CNN)

Table 2 Terms used for literature search

Category	Search terms
Predictive Quality	Predictive quality, predictive analytics, fault prediction, fault classification, defect prediction, quality prediction, smart manufacturing
Learning	Deep learning, neural network, machine learning
Domain	Manufacturing, production, industrial, engineering, automation, assembly

**Fig. 2** Methodology of the literature search

literature search resulted in a large list of 1,261 (potentially relevant) publications for our review (see Fig. 2).

With the aim of identifying only publications that fit the scope of the paper, we first screened the publications based on their title and abstract. In the course of this screening, many publications were discarded as they lay in the fields of predictive maintenance, fault diagnosis, remaining useful lifetime prediction, software defect prediction, water quality prediction, process engineering or civil engineering. The result of this pre-selection were 144 remaining publications.

We then read through the remaining publications in detail and categorized them according to the defined categories in Table 1. During this process, we excluded publications from our consideration that met the following exclusion criteria:

- Publications that do not contain information about the addressed manufacturing process or the data basis
- Survey papers and literature studies

- Publications that do not perform any development, implementation or evaluation of methods
- Publications that are not accessible to us

After all, there were 81 publications which were selected and considered for our study. Table 10 in Appendix A lists all of them, sorted by publication year. The majority (69%) are published in journals. 31% are published in the scope of scientific conferences. Figure 3 shows the number of publications per year. It clearly illustrates that the number of publications has been constantly increasing over the last years. It can also be assumed that this trend will continue in the further course of 2021 and 2022.

Results

In this chapter, we present the results of our study along the guiding questions formulated in “Introduction” section.

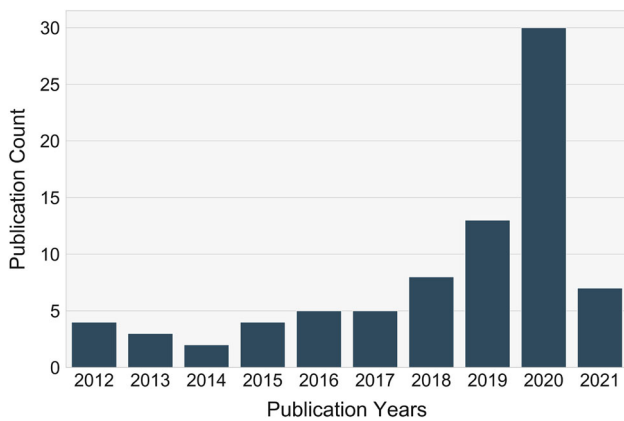


Fig. 3 Number of publications per year. The literature search was performed on June 29, 2021

First, we begin by categorizing the publications according to the manufacturing processes and quality criteria addressed (“[Manufacturing process types and quality criteria](#)” section). Then, we look at the data bases used for quality estimations (“[Data Bases and Characteristics](#)” section) and what learning models are trained and investigated (“[Machine learning methods](#)” section).

Manufacturing process types and quality criteria

Predictive quality approaches are being investigated in a wide variety of manufacturing processes. We therefore perform a categorization of the processes based on DIN 8580:2003-09 (2003). In the norm, the processes are divided into six main groups: *primary shaping*, *forming*, *cutting*, *joining*, *coating* and *changing of material properties*. Since there are also publications dealing with *additive manufacturing*, *assembly* processes or *multi-stage* processes, we add these three categories to our review. Considering how much each process category is represented in predictive quality research, there are significant differences and a large imbalance between them. Figure 4 shows the number of publication per category. While cutting comprises the largest group with 26 publications, there is no publication that primarily addresses processes for changing material properties. In the following, we will focus the categories in detail. For each of the categories, we also analyze which quality criteria are used in the publications as the estimated target variables.

Cutting Cutting includes a variety of manufacturing processes in which a commonly metallic workpiece is fractured by separating a portion from it. Table 3 lists the addressed cutting processes and quality criteria. It can be seen that most research aims to determine quality characteristics that reflect the shape of the finished product, while the surface roughness represents the overwhelming majority of them. In turning applications, for example, some approaches used ML methods such as multivariate regression or ANNs to

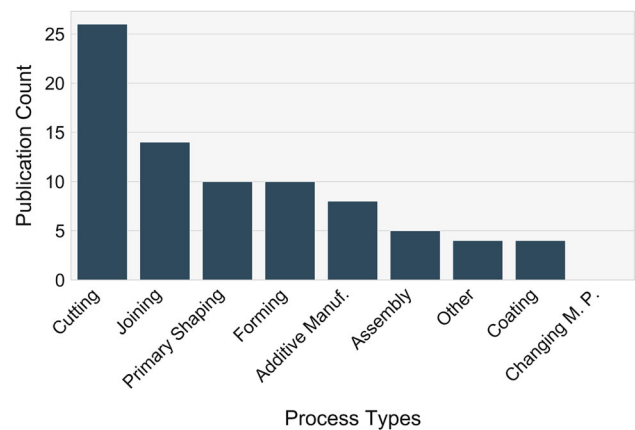


Fig. 4 Number of publications for each manufacturing process type

estimate the roughness based on gathered sensor data (Du et al., 2021; Elangovan et al., 2015; Moreira et al., 2019) and/or machine parameters (Acayaba & de Escalona, 2015). Tušar et al. (2017) developed an automated quality control of a turning and soldering process, predicting both the roughness as well as several soldering defects based on recorded camera images. Vrabel et al. (2016) also proposed an inline process monitoring of the surface roughness quality. Other applications of predicting the roughness are found in the field of milling (Hossain & Ahmad, 2014; Serin et al., 2020a), honing (Gejji et al., 2020; Klein et al., 2020), diamond wire saw cutting (Kayabasi et al., 2017), or laser cutting (Tercan et al., 2016, 2017; Zhang & Lei, 2017). An important quality criterion in drilling processes is the hole diameter, which can also be predicted based on sensor data (Neto et al., 2013; Schorr et al., 2020a, b). Other than that, Nguyen et al. (2020) predicted the waviness of the kerf in laser cutting by training an MLP on process parameters (e.g. gas pressure and laser power). Furthermore, it was shown that ML is capable of predicting the material removal rate in cylindrical grinding of hardened steel (Varma et al., 2017) and chemical mechanical polishing (Yu et al., 2019).

Joining Joining processes comprise the second largest field of addressed manufacturing processes (14 publications). Table 3 shows that mainly applications in welding were investigated. In laser welding, machine learning models (commonly ANNs) were trained on welding parameters (e.g. laser power, welding speed) or sensor data (e.g. light intensity) to predict quality values such as the tensile strength (Yu et al., 2016), weld bead dimensions (Ai et al., 2016; Lei et al., 2019), residual stress (Dhas & Kumanan, 2014) or to classify quality types captured with camera images (Yu et al., 2020). Similar approaches were conducted in spot welding (Hamidinejad et al., 2012; Martín et al., 2016) and ultrasonic welding (Li et al., 2020b; Natesh et al., 2019). Li et al. (2020a) and Gyasi et al. (2019) presented ANN-based inline quality control systems in welding processes. Goldman et al. (2021) conducted interpretability analysis of CNNs trained

Table 3 Considered cutting and joining processes and quality criteria (in parenthesis: number of publications)

Category	Process	Quality criteria	
Cutting	Turning (6)	Surface roughness (Acayaba & de Escalona, 2015; Du et al., 2021; Elangovan et al., 2015; Moreira et al., 2019; Tušar et al., 2017), machinability (Lutz et al., 2020)	
	Drilling (6)	Diameter (Neto et al., 2013), Schorr et al. 2020a, 2020b, surface roughness (Vrabel et al., 2016), hole defects (Jiao et al., 2020), surface gap (Bustillo et al., 2018)	
	Laser cutting (4)	Surface roughness (Tercan et al., 2016, 2017; Zhang & Lei, 2017), kerf waviness (Nguyen et al., 2020)	
	Milling (3)	Surface roughness (Hossain & Ahmad, 2014; Serin et al., 2020a), geometric deviation (de Oliveira Leite et al., 2015)	
	Honing (2)	Surface roughness (Gejji et al., 2020; Klein et al., 2020)	
	C. M. polishing (1)	Material removal rate (Yu et al., 2019)	
	Diamond wire cutting (1)	Surface roughness (Kayabasi et al., 2017)	
	Grinding (1)	Surface roughness (Varma et al., 2017)	
	Laser micro grooving (1)	Groove geometry (Zahrani et al., 2020)	
	Laser machining (1)	Dimensions (McDonnell et al., 2021)	
	Joining	Laser welding (4)	Weld bead dimensions (Ai et al., 2016; Lei et al., 2019), tensile strength (Yu et al., 2016), quality types (Yu et al., 2020)
		Resistance spot welding (3)	Tensile shear strength (Hamidinejad et al., 2012), tensile shear load bearing (Martín et al., 2016), welding deformation (Li et al., 2020a)
		Ultrasonic welding (3)	Quality types (Goldman et al., 2021; Li et al., 2020b), tensile strength (Natesh et al., 2019)
Gas metal arc welding (2)		Weld penetration (Gyasi et al., 2019), weld bead dimensions (Wang et al., 2021)	
Gluing (1)		Glue volume (Dimitriou et al., 2020)	
Welding (1)		Residual stress (Dhas & Kumanan, 2014)	

on welding sensor data. In contrast Wang et al. (2021) trained a CNN on line camera images for quality estimation. Other than welding, Dimitriou et al. (2020) estimated the glue volume based on 3D laser topology scans in a gluing process.

Primary shaping Primary shaping involves processes in which a body with a defined shape is produced from a shapeless material. Ten publications deal with this field (see table 4). The ones that lie in the field of casting proposed approaches for detecting casting defects on the product, such as by training CNNs on X-ray images (Ferguson et al., 2018) or MLPs on sensor data (Kim et al., 2018; Lee et al., 2018). In injection molding, data from machine parameters (e.g. temperature, packing pressure) was used for predicting quality values such as the product dimensions (Ke & Huang, 2020) or product weights (Ge et al., 2012). Alvarado-Iniesta et al. (2012) used a recurrent neural network to make warpage estimations for new parameter combinations. Garcia et al. (2019) predicted future product geometries of plastics tubes in a plastics extrusion process. Furthermore, two works lie in the field of spinning, where the goals were to predict the yarn quality in form of the count-strength-product (Nurwaha &

Wang, 2012) or the leveling action point (Abd-Elatif, 2013) with MLPs.

Forming Forming involves manufacturing processes in which raw parts are transformed into a different shape without material being added or removed. Among the 10 publications that lie in this field (see table 4), the ones addressing metal rolling processes noticeably differ from the works mentioned so far, as most of them aimed at in-line quality estimations in the rolling process. For example, Yun et al. (2020), Li et al. (2018) and Liu et al. (2021) proposed CNN-based quality inspection systems by detecting and classifying surface defects in line camera images. Ståhl et al. (2019) used inline geometry measurements to train LSTM networks and Lieber et al. (2013) made inline NOK quality predictions based on ultrasonic measurements. In sheet metal forming, Meyes et al. (2019) investigated LSTMs on sensor data for the inline prediction of part defects, while Essien and Gianetti (2020) trained them to estimate the machine speed. In contrast, Dib et al. (2020) made use of simulated experiments for ML-based part defect estimation. Other approaches were

Table 4 Considered primary shaping, forming and additive manufacturing processes and quality criteria (in parenthesis: number of publications)

Process type	Process	Quality criteria
Primary shaping	Casting (3)	Casting defects (Ferguson et al., 2018; Kim et al., 2018; Lee et al., 2018)
	Injection molding (3)	Product dimensions (Ke & Huang, 2020), product weight (Ge et al., 2012), warpage (Alvarado-Iniesta et al., 2012)
	Plastics extrusion (2)	Product geometry (Garcia et al., 2019), yield stress (Mulrennan et al., 2018)
	Spinning (2)	Yarn strength (Nurwaha & Wang, 2012), sliver evenness (Abd-Ellatif, 2013)
Forming	Metal rolling (5)	Surface defects (Li et al., 2018; Lieber et al., 2013; Liu et al., 2021; Yun et al., 2020), slab geometry (Ståhl et al., 2019)
	Sheet metal forming (3)	Part defects (Dib et al., 2020; Meyes et al., 2019), machine speed (Essien & Giannetti, 2020)
	Forging (1)	Process feasibility (Ciancio et al., 2015)
	Textile draping (1)	Shear deformation (Zimmerling et al., 2020)
Additive manuf.	Laser powder bed fusion (4)	Geometric deviation (Zhu et al., 2020), inherent strain (Li & Anand, 2020), structural defects (Bartlett et al., 2020), single-track width (Gaikwad et al., 2020)
	Direct metal deposition (1)	Volume porosity (Zhang et al., 2019a)
	Fused deposition modeling (2)	Tensile strength (Zhang et al., 2018, 2019b)
	PLA 3D printing (1)	Surface roughness (Li et al., 2019)

investigated in simulations of impression-die forging (Ciancio et al., 2015) or textile draping (Zimmerling et al., 2020).

Additive manufacturing Eight publications deal with predictive quality in additive manufacturing processes (see table 4). As additive manufacturing enables rapid prototyping, two of them are located in the design phase of products. Here, process simulations were used to train ANNs for fast predicting the inherent strain (Li & Anand, 2020) or geometric deviations (Zhu et al., 2020) of the product. Beyond that, ML can also be used in the realization phase, for example to make quality predictions based on optical measurements (Gaikwad et al., 2020; Bartlett et al., 2020) or machine sensors (e.g. IR, vibration) (Li et al., 2019; Zhang et al., 2018, 2019b). An in-line capability of quality monitoring in the process was demonstrated by Zhang et al. (2019a).

Assembly Regarding the 5 publications dealing with assembly (see Table 5), it becomes apparent that they are mainly concerned with ML-based classification of successful and unsuccessful assembly tasks. Examples are the detection of functioning products in manual assembly (Wagner et al., 2020) or correct positioning in SMT assembly (Schmitt et al., 2020a) by using virtual quality inspection systems. The assembled products as well as the data used may also be very different. While Sarivan et al. (2020) used acoustic signals to

make a quality prediction for the connection of wire plugs, Martinez et al. (2020) used line camera images for detecting correctly fastened screws. Lastly, Doltsinis et al. (2020) detected successful operations on the basis of robotic force signatures and machine sensors.

Coating In coating (4 publications, Table 5), manufacturing processes are involved to apply an adhesive layer of shapeless material to the surface of a certain workpiece. In dispensing, for example. Oh et al. (2019) proposed a SVM-based defect detection method for realtime visual quality inspection. Hsu and Liu (2021) trained CNNs on machine sensor data for OK/NOK classification of electric wafer quality. In contrast, some approaches are trained only on parameterizations of the process, which was shown for lacquering (Thomas et al., 2018) and car bodywork painting (Kebisek et al., 2020).

Multi-stage There are also publications which are not concerned with a single manufacturing process but multi-stage processes that comprise several types (4 publications, see Table 5). From a machine learning perspective, the challenge here is handling the increased complexity and number of data sources. For example, Liu et al. (2020b) investigated DL-methods (e.g. LSTMs) for quality prediction of a larger production line based on multimodal sensor data. The data

Table 5 Considered assembly, coating and multi-stage processes and quality criteria (in parenthesis: number of publications)

Process type	Process	Quality criteria
Assembly	Manual assembly (2)	Operation success (Wagner et al., 2020; Sarivan et al., 2020)
	Screw fastening (1)	Operation success (Martinez et al., 2020)
	SMT assembly (1)	Component position (Schmitt et al., 2020a)
	Snap-fit assembly (1)	Operation success (Doltsinis et al., 2020)
Multi-stage	Battery-cell manufacturing (1)	Battery capacity and state of health (Turetskyy et al., 2021)
	Metal forming and machining (1)	Product dimensions (Papananias et al., 2019)
	Production line (1)	Quality types (Liu et al., 2020b)
	Textile manufacturing (1)	Fabric defects (Jun et al., 2021)
Coating	Chemical vapor deposition (1)	quality types (Hsu & Liu, 2021)
	Lacquering (1)	Defect types (Thomas et al., 2018)
	Painting (1)	Paint structure (Kebisek et al., 2020)
	Primer-sealer dispensing (1)	Defect types (Oh et al., 2019)

was captured in different stages of the production. Similar work was also conducted in metal processing (Papananias et al., 2019) or multi-stage battery-cell manufacturing (Turetskyy et al., 2021), where the authors combined the model training with extensive feature selection to find suitable data inputs.

Data bases and characteristics

Data plays a central role in the context of predictive quality. Therefore, it is important to address the question of which data is used in the publications for training and evaluation of the predictive quality models. In the following, we aim to answer the second guiding question by categorizing the publications according to the source of the data, the amount of data used, the parameters used as input variables for the models, and the modality of the data.

Data set sources and amount

We have basically identified three source of data from which the publications obtain their data for model training: real data from a manufacturing process, virtual data from simulations, or from freely available data sets that have been generated for benchmarks or in the context of competitions. Real data is data that is gathered from the physical manufacturing process in which the corresponding product is produced on a machine or production line. Thus it is obtained during the manufacturing of a product. Depending on the use case and the manufacturing process, different measurement techniques and sensors are used for gathering real data. In contrast, virtual data is created during simulation runs of the manufacturing process. The data is generally created before the product is physically manufactured. Furthermore, in the case of real manufacturing data, we distinguish between two

types of publications: those that take their data from a running production, and those that generate their data experimentally using predefined experimental designs. Table 6 lists all publications according to the data sources. While few publications exist that primarily use simulation data (10%) or benchmark data (6%), the majority of them make use of real manufacturing data (65% with experimental data and 14% with data from running production). In addition, there are four publications (Abd-Ellatif, 2013; Ge et al., 2012; Nurwaha & Wang, 2012; Tušar et al., 2017) where the data source is not clear and therefore we cannot clearly assign them to one of the categories. In the following, we briefly discuss the data sources and also look at how large the data sets used for model training are.

Simulation Eight publications use simulations in which the respective manufacturing processes are simulated and the targeted quality variables are calculated. The generated data sets serve as a training basis for the machine learning models. The typical objective here is either to demonstrate the feasibility of ML-driven predictive quality using simulation data (Ciancio et al., 2015; Zhu et al., 2020), or to generate fast ML models from simulations in order to save cost and time in process design (Dib et al., 2020; Tercan et al., 2016, 2017). The most publications perform a certain set of experiments for data generation, either by varying process parameters (Alvarado-Iniesta et al., 2012; Ciancio et al., 2015; Dib et al., 2020; Tercan et al., 2016, 2017) and/or product design parameters (Li & Anand, 2020; Zhu et al., 2020; Zimmerling et al., 2020). On average, the publications use 9, 864 of simulated experiments, where the result of each experiment serves as a data sample for model training. The number of data samples vary widely across the publications: while Ciancio et al. (2015) used only 30 simulated data samples for training and testing, Tercan et al. (2017) generated over 22, 000. Zhu et al. (2020) additionally performed data

Table 6 Main sources of process data to train machine learning models

Data source	Publications
Simulation	Alvarado-Iniesta et al. (2012), Ciancio et al. (2015), Dib et al. (2020), Li and Anand (2020), Tercan et al. (2016, 2017), Zhu et al. (2020), Zimmerling et al. (2020)
Benchmark/competition	Ferguson et al. (2018), Jun et al. (2021), Liu et al. (2020b, 2021), Yu et al. (2019)
Real data (running production)	Essien and Giannetti (2020), Goldman et al. (2021), Kebisek et al. (2020), Lee et al. (2018), Li et al. (2018), Meyes et al. (2019), Oh et al. (2019), Schmitt et al. (2020a), Ståhl et al. (2019), Wagner et al. (2020), Yun et al. (2020)
Real data (experimental)	Acayaba and de Escalona (2015), Ai et al. (2016), Bartlett et al. (2020), Bustillo et al. (2018), Dhas and Kumanan (2014), Dimitriou et al. (2020), Doltsinis et al. (2020), Du et al. (2021), Elangovan et al. (2015), Gaikwad et al. (2020), Garcia et al. (2019), Gejji et al. (2020), Zahrani et al. (2020), Gyasi et al. (2019), Hamidinejad et al. (2012), Hossain and Ahmad (2014), Hsu and Liu (2021), Jiao et al. (2020), Kayabasi et al. (2017), Ke and Huang (2020), Kim et al. (2018), Klein et al. (2020), Li et al. (2019, 2020a, 2020b), Lutz et al. (2020), Mulrennan et al. (2018), Lei et al. (2019), de Oliveira Leite et al. (2015), Lieber et al. (2013), Martín et al. (2016), Martínez et al. (2020), McDonnell et al. (2021), Moreira et al. (2019), Natesh et al. (2019), Neto et al. (2013), Nguyen et al. (2020), Papananias et al. (2019), Sarivan et al. (2020), Schorr et al. (2020a, 2020b), Serin et al. (2020a), Thomas et al. (2018), Turetskyy et al. (2021), Varma et al. (2017), Vrabel et al. (2016), Wang et al. (2021), Yu et al. (2016, 2020), Zhang et al. (2018, 2019a, 2019b), Zhang and Lei (2017)

augmentation techniques to increase the data amount from originally 102 samples to 1980.

Benchmark We identified five publications that use freely available data sets as a training and evaluation basis for predictive quality. Three of them lie in the field of image-based defect classification, such as the GRIMA X-Ray casting data (Mery et al., 2015) used by Ferguson et al. (2018), the NEU-DET data set (He et al., 2020) for surface defects in hot-rolling (Liu et al., 2021), and the Xuelang manufacturing AI challenge data set (Tianchi, 2021) used for fabric defect classification (Jun et al., 2021). Besides that, Liu et al. (2020b) used the large-scale Kaggle Bosch Production Line Performance data set and Yu et al. (2019) used the 2016 PHM Data Challenge data set for material removal rate prediction (PHM Society, 2020). On average, the data sets contain 5, 722 samples or images, with the NEU-Det data set being the smallest with 1, 800 images and the Kaggle dataset being the largest with over 16, 000 samples. Yet again, the works which use image sets (Ferguson et al., 2018; Jun et al., 2021; Liu et al., 2021) performed data augmentation on the images to increase the training data sizes.

Real data (experimental) Most research on predictive quality is conducted on real manufacturing, meaning that the data basis for model training is retrieved from a physical process in which a certain product is produced on a machine or production line. The majority of these publica-

tions (83%) conduct predefined experiments on the process to generate the training data. Thereby a selected set of process parameters is varied under fixed boundary conditions and mostly for fixed workpieces. In some cases, well-known design of experiments are used for that, such as full factorial (Du et al., 2021; Elangovan et al., 2015; Gaikwad et al., 2020; Hamidinejad et al., 2012; Ke & Huang, 2020; Nguyen et al., 2020; Varma et al., 2017; Zhang et al., 2018, 2019b), half factorial (Gejji et al., 2020), fractional Box-Behnken (Hossain & Ahmad, 2014), central composite (Hamidinejad et al., 2012; Serin et al., 2020a), or Taguchi (Ai et al., 2016; Ciancio et al., 2015; Moreira et al., 2019) design method. The numbers of parameters that are varied (if known) are two (Bustillo et al., 2018; Gaikwad et al., 2020; Ke & Huang, 2020; Moreira et al., 2019; Vrabel et al., 2016), three (Du et al., 2021; Kayabasi et al., 2017; Mulrennan et al., 2018; Li et al., 2020a, 2019; McDonnell et al., 2021; Natesh et al., 2019; Nguyen et al., 2020; Varma et al., 2017; Yu et al., 2020; Zhang & Lei, 2017), four (Ai et al., 2016; Elangovan et al., 2015; Gejji et al., 2020; Hamidinejad et al., 2012; Jiao et al., 2020; Li et al., 2020b; Neto et al., 2013), five (Zhang et al., 2019a), six (Dhas & Kumanan, 2014), seven (Zahrani et al., 2020), or eight (Lei et al., 2019). Only a few publications vary other conditions besides process parameters, which are material batches (Lutz et al., 2020) or product specimens (Neto et al., 2013; Papananias et al., 2019). The

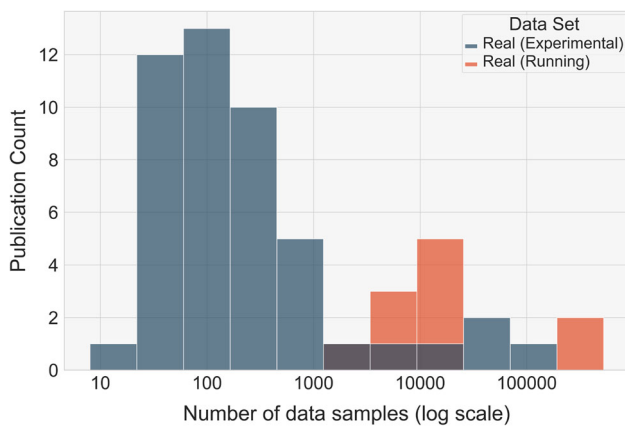


Fig. 5 Distribution of publications according to the number of data samples used for model training and evaluation

small number of parameter variations is also reflected in the small amount of data used for model training and evaluation. Figure 5 illustrates the distribution of the publications based on the number of data samples they use. The blue bars represent the publications with experimental data. It shows that the majority of publications use around 100 samples. The average number is about 5,600 and the median 144. Note that the number of data samples is not always equal to the number of experiments. For example, the top-4 publications that use more than 10,000 samples either perform data augmentation (Dimitriou et al., 2020; Hsu & Liu, 2021) or conduct multiple measurements per experiment (Wang et al., 2021; Zhang et al., 2019a):

Real data (running production) In contrast to experimentally generated training data, there are eleven publications that take data from a running manufacturing process. In these cases, the manufacturing of the products and the collection of the data took place over a longer period of time, such as three months in an aluminum casting process (Lee et al., 2018), six months of steel rolling production (Stahl et al., 2019), eleven months in a deep drawing process (Meyes et al., 2019), twelve months of aluminum can bodymaking (Essien & Giannetti, 2020), or 20 months in a car bodywork shop (Kebisek et al., 2020). Based on this data acquisition, historical data sets were generated on which the ML models are trained and evaluated. The data sets here are generally larger than in the publications that obtain the data experimentally (see red bars in Fig. 5). On average, the data sets contain 73,984 data samples. The largest data set used by Essien and Giannetti (2020) contains 525,600 samples from sensor data (machine speed) that is recorded every minute over the course of one year.

Input variables

The quality of a manufactured product depends on many factors. Among them are the design and parameterization of the

manufacturing process and the interplay of the manufacturing steps. A predictive quality model trained on process data correlates these input factors with the quality. We therefore pose the question of what types of input variables are used in the publications. We have identified three major types in the course of our research: process parameters, sensor data, and product measurements. As it turns out, in most publications only one type is selected as the basis for model training. Table 7 provides an overview of all publications. Four publications (Abd-Ellatif, 2013; Liu et al., 2020b; Wagner et al., 2020; Yu et al., 2019) are not included as we cannot clearly determine which types of variables they use.

Process parameters The first type of input variables are process and machine parameters. These are set for the production of a particular product and are usually not changed during production. A predictive quality model built on process parameters can be used to determine product quality under new, previously unknown parameter spaces. Depending on the manufacturing process, different parameters are essential for the product quality and thusly are selected in the publications. While in cutting processes the feed rate (Acayaba & de Escalona, 2015; Bustillo et al., 2018; Hossain & Ahmad, 2014; Jiao et al., 2020; Varma et al., 2017) or cutting and rotation speed (Acayaba & de Escalona, 2015; Bustillo et al., 2018; Hossain & Ahmad, 2014; Jiao et al., 2020; Nguyen et al., 2020; Serin et al., 2020a; Varma et al., 2017; Zahrani et al., 2020; Zhang & Lei, 2017) are frequently used, the laser power (Ai et al., 2016; Lei et al., 2019; Nguyen et al., 2020; Yu et al., 2020; Zahrani et al., 2020; Zhang & Lei, 2017; Zhu et al., 2020) or focal position (Ai et al., 2016; Tercan et al., 2016, 2017) are often selected in laser-based applications. In plastics manufacturing, on the other hand, process times and temperatures (Alvarado-Iniesta et al., 2012; Ge et al., 2012; Mulrennan et al., 2018) are frequently occurring in the data. Worth mentioning here is the work in additive manufacturing by Li and Anand (2020) who used design parameters such as hatch patterns as model inputs and Zhu et al. (2020) for adding product design parameters such as product size and geometry to the data basis.

Sensor data The second type of input data is sensor data, which is taken from the process or machine during the manufacturing process. Sensor data values therefore represent the actual state of the process or the condition of the machine. A model trained on sensor data can, for example, provide an estimate of the product quality as it is being manufactured. As with process parameters, sensor data is highly dependent on the particular manufacturing process. While in welding processes, for example, welding current is often captured (Goldman et al., 2021; Li et al., 2020a,b), in shaping processes it is temperature and pressure data (Garcia et al., 2019; Kim et al., 2018; Lee et al., 2018). In cutting processes, on the other hand, there are often vibration, torque or force sensors used (Du et al., 2021; Moreira et al., 2019; Neto et al., 2013;

Table 7 Types of input variables used for predictive quality models (in parenthesis: number of publications)

Variable type	Publications
Process parameters (30)	Acayaba and de Escalona (2015), Ai et al. (2016), Alvarado-Iniesta et al. (2012), Bustillo et al. (2018), Ciancio et al. (2015), Dhas and Kumanan (2014), Dib et al. (2020), Ge et al. (2012), Gejji et al. (2020), Zahrani et al. (2020), Hamidinejad et al. (2012), Hossain and Ahmad (2014), Jiao et al. (2020), Kayabasi et al. (2017), Kebisek et al. (2020), Mulrennan et al. (2018), Lei et al. (2019), Li and Anand (2020), Martín et al. (2016), McDonnell et al. (2021), Natesh et al. (2019), Nguyen et al. (2020), Serin et al. (2020a), Tercan et al. (2016, 2017), Varma et al. (2017), Yu et al. (2020), Zhang and Lei (2017), Zhu et al. (2020), Zimmerling et al. (2020)
Sensor Data (22)	Essien and Giannetti (2020), Doltsinis et al. (2020), Du et al. (2021), Garcia et al. (2019), Goldman et al. (2021), Gyasi et al. (2019), Hsu and Liu (2021), Kim et al. (2018), Lee et al. (2018), Li et al. (2019, 2020a, 2020b), Lieber et al. (2013), Meyes et al. (2019), Moreira et al. (2019), Neto et al. (2013), Nurwaha and Wang (2012), Papananias et al. (2019), Sarivan et al. (2020), Schorr et al. 2020a, 2020b, Turetskyy et al. (2021)
Sensor data + process parameters (9)	Elangovan et al. (2015); Ke and Huang (2020), Klein et al. (2020), Lutz et al. (2020), Thomas et al. (2018), Vrabel et al. (2016), Yu et al. (2016), Zhang et al. (2018, 2019b)
Product measurements (16)	Bartlett et al. (2020), Dimitriou et al. (2020), Ferguson et al. (2018), Gaikwad et al. (2020), Jun et al. (2021), de Oliveira Leite et al. (2015), Li et al. (2018), Liu et al. (2021), Martinez et al. (2020), Oh et al. (2019), Schmitt et al. (2020a), Ståhl et al. (2019), Tušar et al. (2017), Wang et al. (2021), Yun et al. (2020), Zhang et al. (2019a)

Schorr et al., 2020a,b). A special case are assembly processes. While Doltsinis et al. (2020) used sensor signals from the robots used in snap-fit assembly, Sarivan et al. (2020) used acceleration and audio measurements from wearables worn by operators.

Sensor data + process parameters There are also publications that obtain data from both process parameters and sensors (see Table 6). For example, Elangovan et al. (2015) showed for a turning process that the addition of sensor variables (e.g. vibration signals) can significantly improve the performance model in contrast to the sole use of process parameters (e.g. cutting speed). Similar approaches can be found in other work on cutting processes, where common process parameters are combined with sensor data such as force or flank wear width (Klein et al., 2020; Lutz et al., 2020; Vrabel et al., 2016), or additive manufacturing, where data from IR sensors are gathered together with temperature and speed parameters (Zhang et al., 2018, 2019b). Also mentionable are the investigations by Lutz et al. (2020) by additionally incorporating data from material batches and tool types and analyzing the deviations of prediction results across them.

Product measurements The third category is measurement data of the product itself during its production or in the course of visual inspection. Models trained on these measurements are often used to automatically detect product defects in them. In more than half of the publications that fall in this category, the measurements are images captured by a

(line) camera. They lie in the field of additive manufacturing (Bartlett et al., 2020; Gaikwad et al., 2020), assembly (Martinez et al., 2020), turning (Tušar et al., 2017), metal rolling (Li et al., 2018; Liu et al., 2021; Yun et al., 2020), welding (Wang et al., 2021) or textile manufacturing (Jun et al., 2021). Other measurement data than camera images are geometric measurements (Ståhl et al., 2019), thermal images (Oh et al., 2019), X-ray images (Ferguson et al., 2018), melt pool images (Zhang et al., 2019a) or laser topology scans (Dimitriou et al., 2020).

As mentioned, many publications focus on one type of input variables for predictive quality. This shows that even for the same manufacturing process different data can be used for quality estimations. A good example are turning processes, where on the one hand there are approaches based on setting parameters of the turning machine (e.g. cutting speed and cut depth) (Acayaba & de Escalona, 2015; Elangovan et al., 2015) and approaches using sensor data of the machine (e.g. vibration signals) on the other hand (Du et al., 2021; Lutz et al., 2020; Moreira et al., 2019). However, our analyses also show that there are also major differences between the respective manufacturing processes. Figure 6 shows how many times the three types of input variables appear in the publications for the nine most frequently addressed manufacturing processes. For instance, while turning and drilling applications use parameters and sensor data, laser cutting applications only use process parameters. It is also noticeable that in four out of five publications dealing with metal

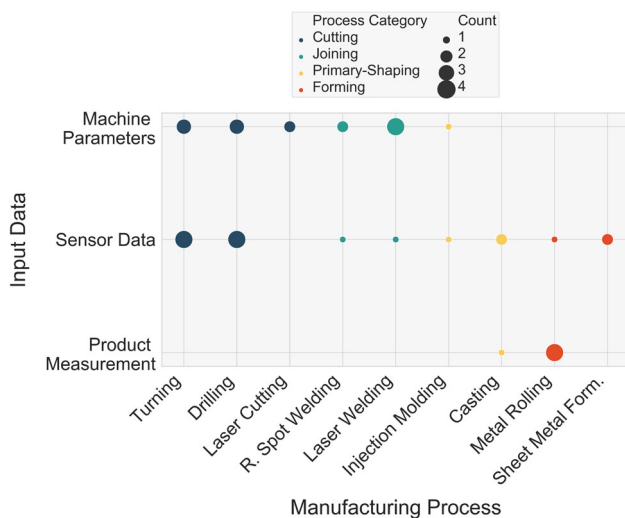


Fig. 6 Number of occurrences of the input variable types in the addressed manufacturing processes. Note that only publications using real data (experimental or running production) are considered here

rolling (Li et al., 2018; Liu et al., 2021; Ståhl et al., 2019; Yun et al., 2020), the quality estimation is solely based on product measurements.

Data modality

From the described input variables, the data sets for training the machine learning models are finally created. Thereby, the data is transformed into a form on which the models can be used. In the reviewed publications, four data types respectively data modalities are identified. Table 8 lists the data types and their occurrences in the publications.

The least common type is categorical data. These are mostly non-numeric representations of process entities, e.g., a tool type and material batch (Lutz et al., 2020), compositional elements of a batch (Lee et al., 2018), or component counts (Thomas et al., 2018). Eight of the nine publications that perform time-series-based quality estimation obtain it from sensor data (Essien & Giannetti, 2020; Goldman et al., 2021; Gyasi et al., 2019; Hsu & Liu, 2021; Meyes et al., 2019; Sarivan et al., 2020; Zhang et al., 2018, 2019b), one publication from product measurements (Ståhl et al., 2019). The data can be univariate time series (Essien & Giannetti, 2020; Meyes et al., 2019; Ståhl et al., 2019) or multivariate time series (Goldman et al., 2021; Hsu & Liu, 2021; Sarivan et al., 2020).

Another important data type used is image data. The majority of these are 2D images, which are drawn from the described product measurements (Bartlett et al., 2020; Dimitriou et al., 2020; Ferguson et al., 2018; Jun et al., 2021; Li et al., 2018; Liu et al., 2021; Martinez et al., 2020; Oh et al., 2019; Wang et al., 2021; Yun et al., 2020; Zhang et al.,

2019a). Dimitriou et al. (2020) used 3D point clouds which are generated by topology scan. In most cases, the generated images are enriched using data augmentation techniques such as adding noise or using random cropping (Dimitriou et al., 2020; Ferguson et al., 2018; Jun et al., 2021; Li et al., 2018; Liu et al., 2021; Martinez et al., 2020), resulting in a larger database for deep learning model training. In addition, Yun et al. (2020) used a variational auto encoder to generate synthetic image data.

The vast majority of publications almost exclusively use numerical/continuous data with scalar values for model training. These are values of parameter settings on the one hand, and quantities obtained from sensors and measurements on the other hand. In fact, 23 of the 31 publications (74%) that gather their data from sensors transform the values into scalar numerical quantities. Commonly, either statistical feature extraction (e.g., minimum value, maximum value, mean value) is conducted (Doltsinis et al., 2020; Du et al., 2021; Elangovan et al., 2015; Garcia et al., 2019; Lee et al., 2018; Li et al., 2019; Lieber et al., 2013; Papananias et al., 2019; Turetskyy et al., 2021; Vrabel et al., 2016; Yu et al., 2016) or expert-driven aggregation is performed (Ke & Huang, 2020; Klein et al., 2020; Schorr et al., 2020a, b).

Machine learning methods

In the third part of our publication review, we look at the learning tasks and machine learning models that are used for predictive quality. In general, the learning task results from the described quality goals and the quality variable to be estimated (e.g., error detection, quality forecasting). On the one hand, this involves classification, which accounts for 30 of 81 publications, and numerical prediction, i.e. regression, which is conducted in 51 publications. Besides from the learning task, the choice of an appropriate ML or DL model for predictive quality results from the data modality (e.g., images) and the complexity of the relationships between input and output (e.g., linear, non-linear). Many different models are trained and experimentally evaluated in the publications. The review reveals that in 49% of all publications (40 out of 81) the evaluation is performed only for a single model or for different variants of the same model (e.g. changes of hyperparameters, architectures, optimization methods). In contrast, 51% (41 of 81) of the publications compare several models in experiments. In the following, the model which is the focus of a publication (because it is the single model) or which performs best in a model comparison is referred to as the *prime model*. All other models used for comparison are called *baseline models*.

Table 9 lists all prime models used in the publications. On the one hand, it can be seen that some models are commonly used for both classification and regression, such as the MLP. On the other hand, the popularity of models based on artificial

Table 8 Occurring modalities of data sets used for training the ML and DL models

Variable Type	Publications
Categorical/discrete	Lee et al. (2018), Liu et al. (2020b), Lutz et al. (2020), Thomas et al. (2018)
Time Series	Essien and Giannetti (2020), Goldman et al. (2021), Gyasi et al. (2019), Hsu and Liu (2021), Meyes et al. (2019), Sarivan et al. (2020), Ståhl et al. (2019), Zhang et al. (2018), Zhang et al. (2019b)
Image	Bartlett et al. (2020), Dimitriou et al. (2020), Ferguson et al. (2018), Jun et al. (2021), Li et al. (2018), Liu et al. (2021), Martinez et al. (2020), Oh et al. (2019), Wang et al. (2021), Yun et al. (2020), Zhang et al. (2019a)
Continuous/numerical	Abd-Ellatif (2013), Acayaba and de Escalona (2015), Ai et al. (2016), Alvarado-Iniesta et al. (2012), Bustillo et al. (2018), Ciancio et al. (2015), Dhas and Kumanan (2014), Dib et al. (2020), Doltsinis et al. (2020), Du et al. (2021), Elangovan et al. (2015), Gaikwad et al. (2020), Garcia et al. (2019), Ge et al. (2012), Gejji et al. (2020), Zahrani et al. (2020), Hamidinejad et al. (2012), Hossain and Ahmad (2014), Jiao et al. (2020), Kayabasi et al. (2017), Ke and Huang (2020), Kebisek et al. (2020), Kim et al. (2018), Klein et al. (2020), Mulrennan et al. (2018), Lee et al. (2018), Lei et al. (2019), de Oliveira Leite et al. (2015), Li et al. (2019, 2020a, 2020b), Li and Anand (2020), Lieber et al. (2013), Liu et al. (2020b), Lutz et al. (2020), Martín et al. (2016), McDonnell et al. (2021), Moreira et al. (2019), Natesh et al. (2019), Neto et al. (2013), Nguyen et al. (2020), Nurwaha and Wang (2012), Papananias et al. (2019), Schmitt et al. (2020a), Schorr et al. (2020a, 2020b), Serin et al. (2020a), Tercan et al. (2017), Tercan et al. (2016), Thomas et al. (2018), Turetskyy et al. (2021), Tušar et al. (2017), Varma et al. (2017), Vrabel et al. (2016), Yu et al. (2016, 2019, 2020), Zhang and Lei (2017), Zhu et al. (2020), Zimmerling et al. (2020)

neural networks and CNNs becomes evident. This also holds true for the most recent publications. Looking at the share of prime models in the publications of 2020 and 2021 (see Fig. 7), both models account for 68% of the publications. In contrast, other models such as SVM, random forest, and decision tree are more often used as baseline models (see Fig. 8). In the following, we will categorize the publications based on the prime models used and also briefly discuss the baseline comparisons.

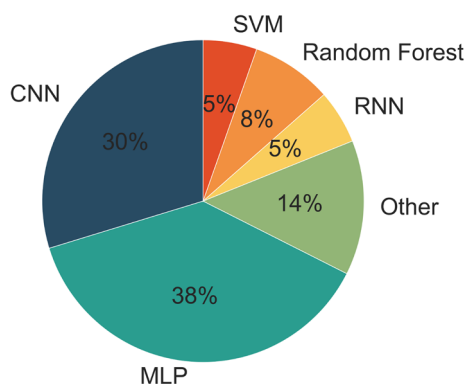
Multilayer perceptron (MLP) The most frequently used prime model is the MLP with 30 publications. The popularity of MLPs continues into 2020 and 2021, and their versatility is demonstrated in a wide variety of use cases. On the one hand, MLPs are used for classification, be it for prediction of certain fault types (Yu et al., 2020; Lee et al., 2018; Dib et al., 2020), binary classification of OK/NOK states (Wagner et al., 2020) or the classification of multiple quality classes (Bustillo et al., 2018; Kebisek et al., 2020; Ke & Huang, 2020). On the other hand, MLPs are used (much more frequently) for the regression of numerical quality metrics, such as the aforementioned estimations of surface roughness (Acayaba & de Escalona, 2015; Du et al., 2021; Kayabasi et al., 2017; Serin et al., 2020a; Vrabel et al., 2016), tensile strength (Hamidinejad et al., 2012; Natesh et al., 2019; Yu et al., 2016) or dimensions (McDonnell et al., 2021; Lei et al., 2019; Neto et al., 2013; Papananias et al., 2019). Regarding the baseline models with which MLPs are experimentally compared, the publications often use other popular ML models such as SVM (Ciancio et al., 2015; Dib et al., 2020; Lee et al., 2018; Lutz et al., 2020;

Nurwaha & Wang, 2012; Wagner et al., 2020), random forest (Dib et al., 2020; Lee et al., 2018; Li et al., 2020b; Lutz et al., 2020; Turetskyy et al., 2021), linear regression (Acayaba & de Escalona, 2015; Hamidinejad et al., 2012; Turetskyy et al., 2021), decision tree (Lee et al., 2018; Wagner et al., 2020), k-nearest neighbor (K-NN) (Dib et al., 2020; Wagner et al., 2020) and/or AdaBoost (Bustillo et al., 2018). In addition, MLPs are compared to methods such as adaptive neuro-fuzzy inference system (ANFIS) (Neto et al., 2013), gene expression programming (GEP) (Nurwaha & Wang, 2012), or an elman network (Papananias et al., 2019).

Convolutional neural network (CNN) CNNs are the second most frequently used models (14 publications), that is for numerical estimation or regression of quality values (Dimitriou et al., 2020; Wang et al., 2021; Zhu et al., 2020; Zimmerling et al., 2020), binary OK/NOK classification (Goldman et al., 2021; Hsu & Liu, 2021; Martinez et al., 2020; Sarivan et al., 2020) or multiclass classification (Ferguson et al., 2018; Jun et al., 2021; Li et al., 2018; Liu et al., 2021; Yun et al., 2020; Zhang et al., 2019a). They are well suited for pattern recognition in higher dimensional and spatial data and are therefore applied to 2D images (Ferguson et al., 2018; Jun et al., 2021; Li et al., 2018; Liu et al., 2021; Martinez et al., 2020; Wang et al., 2021; Yun et al., 2020; Zhang et al., 2019a), 3D point clouds (Dimitriou et al., 2020) and also time series data (Goldman et al., 2021; Hsu & Liu, 2021; Sarivan et al., 2020). Because of this characteristic, the CNN architectures used in the publications are often empirically compared with other CNN variations.

Table 9 Overview of addressed learning tasks and used prime models in all publications

Learning task	ML-model	Publications
Classification	CNN	Ferguson et al. (2018), Goldman et al. (2021), Hsu and Liu (2021), Jun et al. (2021), Li et al. (2018), Liu et al. (2021), Martinez et al. (2020), Sarivan et al. (2020), Yun et al. (2020), Zhang et al. (2019a)
	Decision tree	Tercan et al. (2016, 2017)
	Ensemble model	Gejji et al. (2020), Kim et al. (2018), Thomas et al. (2018)
	K-NN	Lieber et al. (2013)
	MLP	Bustillo et al. (2018), Dib et al. (2020), Ke and Huang (2020), Kebisek et al. (2020), Lee et al. (2018), Wagner et al. (2020), Yu et al. (2020)
	Naive Bayes	Bartlett et al. (2020)
	Random forest	Zahrani et al. (2020)
	RNN	Liu et al. (2020b), Meyes et al. (2019)
	SVM	Doltsinis et al. (2020), Oh et al. (2019), Schmitt et al. (2020a)
	Regression	ANFIS
CNN		Dimitriou et al. (2020), Wang et al. (2021), Zhu et al. (2020), Zimmerling et al. (2020)
Ensemble model		Li et al. (2019)
Extra tree		Schorr et al. (2020a)
EML		Nguyen et al. (2020)
GA-BPNN		Ai et al. (2016)
Linear regression		Elangovan et al. (2015)
MLP		Abd-Ellatif (2013), Acayaba and de Escalona (2015), Ciancio et al. (2015), Du et al. (2021), Gyasi et al. (2019), Hamidinejad et al. (2012), Jiao et al. (2020), Kayabasi et al. (2017), Lei et al. (2019), de Oliveira Leite et al. (2015), Li et al. (2020a, 2020b), Li and Anand (2020), Lutz et al. (2020), McDonnell et al. (2021), Natesh et al. (2019), Neto et al. (2013), Nurwaha and Wang (2012), Papananias et al. (2019), Serin et al. (2020a), Turetsky et al. (2021), Vrabel et al. (2016), Yu et al. (2016)
NN-GA-PSO		Dhas and Kumanan (2014)
Quadratic regression		Martín et al. (2016)
Random forest		Klein et al. (2020), Mulrennan et al. (2018), Schorr et al. (2020b), Tušar et al. (2017), Yu et al. (2019)
Relevance vector machine		Ge et al. (2012)
RNN		Alvarado-Iniesta et al. (2012), Essien and Giannetti (2020), Ståhl et al. (2019), Zhang et al. (2018), Zhang et al. (2019b)
SeDANN		Gaikwad et al. (2020)
SVM		Garcia et al. (2019)

**Fig. 7** Proportions of ML models (prime) used in publications in 2020 and 2021

For example, Jun et al. (2021) showed that a combination of CNNs with convolutional variational autoencoders (CVAE) provides better performances for class-imbalanced data than variations without CVAE. Yun et al. (2020) included well-known deep learning models such as AlexNet, VGG-16, and ResNet-50 in their comparisons. Liu et al. (2021) also developed their own architecture called TruingDet based on a Faster R-CNN and deformable convolutions and compared it to other state-of-the-art R-CNN models. In addition, other methods used for comparison with CNNs are MLP and SVM (Li et al., 2018), several computer vision algorithms (Jun et al., 2021), and shapelet forests (Hsu & Liu, 2021).

Recurrent neural network (RNN) Seven publication use a recurrent neural network architecture as a prime model for quality estimations. Since RNNs are suitable for application

on time-dependent or sequential data, they are commonly applied on gathered time series data (Meyes et al., 2019; Essien & Giannetti, 2020; Ståhl et al., 2019; Zhang et al., 2018, 2019b). Some of the publications use them for the binary classification of defects (Liu et al., 2020b; Meyes et al., 2019), others for the regression of quantities such as material warpage (Alvarado-Iniesta et al., 2012), machine speed (Essien & Giannetti, 2020), or tensile strength (Zhang et al., 2018, 2019b). Regarding the type of RNNs, the focus of the publications is clearly on LSTM network architectures (Essien & Giannetti, 2020; Liu et al., 2020b; Meyes et al., 2019; Ståhl et al., 2019; Zhang et al., 2018, 2019b). For example, Meyes et al. (2019) used bidirectional LSTM which allows a time series to be processed in a forward run and a backward run to gain better classification results. Essien and Giannetti (2020) used a convolutional LSTM encoder-decoder architecture to forecast future values of the series. The authors therefore compared their approach with another CNN architecture and with an ARIMA model. Other baseline models used for the comparison with RNNs are SVM (Liu et al., 2020b; Ståhl et al., 2019; Zhang et al., 2018, 2019b), random forest (Ståhl et al., 2019; Zhang et al., 2018, 2019b), XGBoost (Liu et al., 2020b), polynomial regression (Zhang et al., 2018) and/or logistic regression (Ståhl et al., 2019).

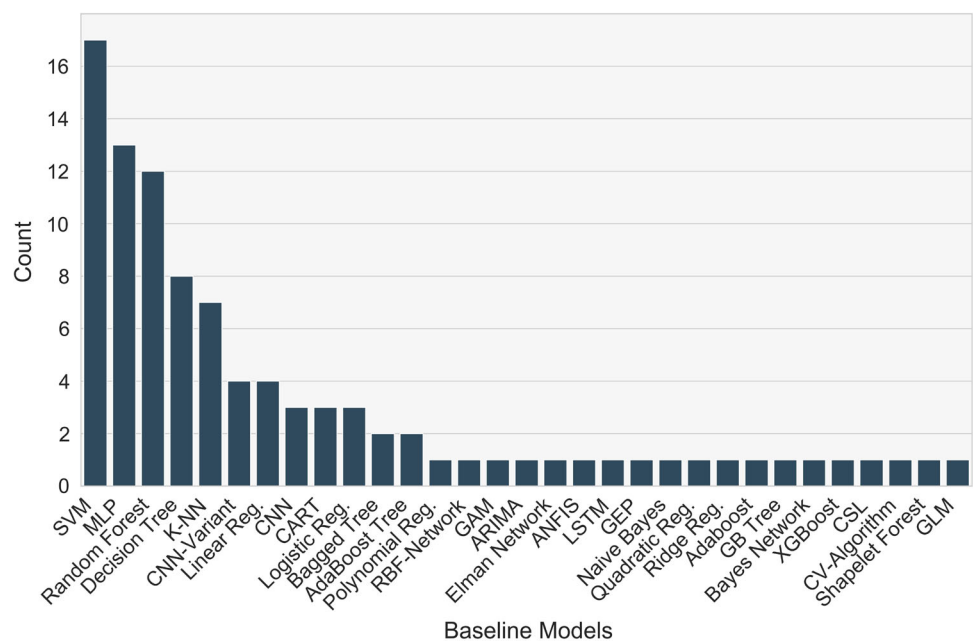
Non-linear ML models Some publications used traditional machine learning methods that are well suited for nonlinear decision making. These include SVMs for the classification of defects (Oh et al., 2019) and operation success (Doltsinis et al., 2020; Schmitt et al., 2020a) as well as for the numerical estimation of product geometries (Garcia et al., 2019), relevance vector machine (RVM) for estimating product weights (Ge et al., 2012), decision trees (Tercan et

al., 2016, 2017) and quadratic regression (Martín et al., 2016) for interpretable quality estimations, and both K-NN (Lieber et al., 2013) and naive bayes (Bartlett et al., 2020) for defect classification. With regard to their evaluation, these models are typically compared to each other (Garcia et al., 2019; Lieber et al., 2013; Oh et al., 2019; Schmitt et al., 2020a) and with other models such as MLPs (Oh et al., 2019; Garcia et al., 2019; Martín et al., 2016), gradient boosted trees (Schmitt et al., 2020a), or generalized additive models (GAM) (Martín et al., 2016).

Ensembles Ensemble methods involve the combination of multiple learning models, thereby aggregating their decisions to make a prediction. In some cases, extensive comparisons were conducted to show that ensembles can perform better than single models (Gejji et al., 2020; Kim et al., 2018; Thomas et al., 2018; Li et al., 2019). Probably the most popular ensemble method is the random forest, which is used for classification (Zahrani et al., 2020) and regression (Klein et al., 2020; Mulrennan et al., 2018; Schorr et al., 2020b; Tušar et al., 2017; Yu et al., 2019). The random forest is compared with single decision trees (Tušar et al., 2017), bagged trees (Mulrennan et al., 2018) and models such as MLP, CNN and SVM (Schorr et al., 2020b).

Variants and hybrid models with neural networks Some publications used methods that are hybrids or variants of artificial neural networks. These include ANFIS, which were proposed to estimate the surface roughness in cutting processes (Hossain & Ahmad, 2014; Moreira et al., 2019; Varma et al., 2017; Zhang & Lei, 2017), ANN variants such as sequential decision analysis neural network (SeDANN) (Gaikwad et al., 2020) and extreme machine learning (Nguyen et al., 2020), and hybrid models of neu-

Fig. 8 Occurrences of ML models as baselines in all publications



ral networks and evolutionary computation methods such as genetic algorithm optimized neural network (GA-BPNN) (Ai et al., 2016) and neuro evolutionary hybrid model with genetic algorithm and particle swarm optimization (NN-GA-PSO) (Dhas & Kumanan, 2014). Accordingly, the approaches were also frequently compared with regular MLPs (Hossain & Ahmad, 2014; Nguyen et al., 2020; Varma et al., 2017; Zhang & Lei, 2017). Gaikwad et al. (2020) additionally compared the SeDANN model with other established models such as CNN, LSTM, CART, and general linear model (GLM).

Key insights and challenges

The previous chapter presented the results of our extensive literature review on predictive quality in manufacturing, in which we categorized the publications along the three dimensions of manufacturing processes and criteria, data basis for model training, and learning models used. Based on the obtained results, we will provide in the following our main findings and identified gaps with respect to the posed driving questions.

Scenarios and manufacturing domains

Applications of predictive quality All in all, the reviewed publications show that machine and deep learning methods prove to be versatile and powerful tools for data-driven quality estimations. In this context, the methods are very often validated with respect to their prognostic quality and accuracy for the particular use cases. The results show that predictive quality has great potential value for quality assurance and manufacturing process improvement. Although many publications do not clearly formulate how the proposed methods are intended to be used in the manufacturing process, we have identified three application scenarios of predictive quality. The first is process design support and process optimization based on simulation data or process parameters. Here, predictive quality methods are used to estimate product quality based on setting parameters. The estimates could then be used either to gain knowledge for the process designer or, in combination with optimization methods, to automate the design of the process. The second scenario is in-line quality prediction during the manufacturing of a product based on process and sensor data. The predictions could then be used to initiate a quality-improving action in the manufacturing process to avoid faults or meet manufacturing tolerances. Third, predictive quality is being investigated for visual quality inspection using ML/DL methods. The methods are used to detect rare product defects in image data or to classify certain defect types. They therefore offer great potential to automate manual and costly visual inspections.

Manufacturing processes Looking at the study results in terms of the manufacturing processes and quality scenarios addressed, there are many similarities (e.g. prediction of the same quality criteria, similar research of approaches for defect detection), but also large imbalances between the manufacturing process groups. While cutting and joining processes account for half of the publications, there are process groups which are hardly dealt with (e.g. coating) or not at all (i.e. changing of material properties). Furthermore, there are also large differences within the process groups. While, for example, many different domains are covered in cutting (e.g. turning, drilling, milling, laser cutting), joining processes are largely covered only by welding processes. Other important branches such as riveting, gluing or soldering are hardly to be found. One reason for this imbalance may be the different degrees of digitization in the domains. The availability of process and quality data in the manufacturing process is an essential requirement for predictive quality. It is noticeable that many of the reviewed publications lie in manufacturing domains in which solutions for process and tool condition monitoring already exist, such as machining processes (Mohanraj et al., 2020; Serin et al., 2020b) or additive manufacturing processes (Lin et al., 2022; Montazeri & Rao, 2018). Accordingly, in these domains it is easier to collect data from the process.

Process integration As mentioned before, the learning models used in the publications show very promising results for their use in real manufacturing scenarios. In most cases, however, the approaches are not integrated into the manufacturing process. Though some publications use real production data for quality inspection (Li et al., 2018; Oh et al., 2019; Ståhl et al., 2019; Wagner et al., 2020; Schmitt et al., 2020a; Yun et al., 2020) or quality prediction (Goldman et al., 2021; Essien & Giannetti, 2020; Lee et al., 2018; Kebisek et al., 2020; Meyes et al., 2019), the training and evaluation as well as the use of the models mostly happen offline or away from the actual process. There are a few works that implement and deploy predictive quality approaches as part of a larger framework (Kebisek et al., 2020; Lee et al., 2018; Li et al., 2020a; Martinez et al., 2020; Oh et al., 2019; Schmitt et al., 2020a). In addition, some publications discuss aspects such as inline capability and real-time capability of the model estimations (Doltsinis et al., 2020; Li et al., 2018; Martinez et al., 2020; Moreira et al., 2019; Sarivan et al., 2020; Schmitt et al., 2020a; Zhang et al., 2019a). However, no discussions are given on the implementation of predictive quality in real quality assurance processes. Furthermore, there is a lack of evaluation of the approaches in terms of their impact on process quality by using quality-oriented metrics (e.g., reduction of rejections, yield rate).

Data bases and characteristics

Input variables The publications show that predictive quality models can be trained on very different data sources and types. The majority of publications use data taken from the physical manufacturing process. This can be process parameters, which are set for the manufacturing of the product, as well as measurements and sensor data, which are gathered during the manufacturing process. Many of the proposed approaches are based on the use of a few input variables of the same type. However, using parameters in combination with sensor data can significantly improve performance, which was shown by Elangovan et al. (2015). In addition, we noticed that other important variables that influence product quality in manufacturing, such as product design or material properties, are not considered in the publications at all. In most cases, ML models are studied for only one product type. The common approach is to train a model on data (parameters or sensors) for a single product with specific characteristics and material compositions. The question of how a model trained for a single product can be used for other products remains open.

Common data processing steps The pre-processing of raw data is an important step before training machine learning models. In particular when using measurement data and sensor data, many publications perform data cleansing (e.g., filling missing data, removing noise and outliers) and data scaling (e.g. normalization or standardization) methods prior to model training. In addition, we noticed two other processing steps that occur frequently. One is the transformation of temporal sensor data into scalar features using feature extraction methods (Doltsinis et al., 2020; Du et al., 2021; Gaikwad et al., 2020; Garcia et al., 2019; Li et al., 2019, 2020b; Lieber et al., 2013; Neto et al., 2013). This involves extracting statistical features from the data (e.g., minimum, maximum, mean values) or transforming the data using signal processing methods. On the other hand, when using image data, data augmentation methods are used to significantly enrich the data set (Dimitriou et al., 2020; Ferguson et al., 2018; Hsu & Liu, 2021; Li et al., 2018; Jun et al., 2021; Liu et al., 2021; Martinez et al., 2020; Yun et al., 2020; Zhu et al., 2020). These methods generate additional image variants by adding noise, rotating the images or randomly cropping them.

Data amount The availability of representative data in a sufficiently large quantity is a fundamental requirement for ML and DL and consequently also for predictive quality. This is a major challenge in a domain like manufacturing where generating data can be cumbersome and costly. The study results show that many approaches are developed and evaluated on a small amount of experimentally generated data, where the experiments often involve the variation of a few process parameters. While experiments offer the advantage over running productions that they can include boundary

conditions and edge cases, they provide in general a less representative data basis. Since many publications also use data sets that contain fewer than 100 data points, their results purely serve to demonstrate the potential and the feasibility of predictive quality. Therefore, solutions have to be researched and developed to improve the data representation and to increase the data quantity. Data augmentation is a promising approach for this. Though it used in some of the reviewed publications (as mentioned above), the focus here lies only on image data.

Benchmark data As mentioned, the generation and use of manufacturing process data is usually expensive and requires time and effort. The use of this data to investigate predictive quality is therefore an investment that cannot always be made. It is therefore all the more important to have freely available benchmark data sets that can be used for investigations. The literature review shows that the vast majority of the publications do not use freely available benchmark datasets, except from (Ferguson et al., 2018; Jun et al., 2021; Liu et al., 2021, 2020b; Yu et al., 2019), nor do they provide their own data base or source code. Thus, on the one hand, there is a lack of comparability between different approaches for similar predictive quality tasks. On the other hand, reproducibility and further development of existing research results is hardly possible.

Machine learning methods

Prime models 68% of publications from 2020 and 2021 use an MLP or CNN as their prime model. This clearly shows the popularity and potential of these models for predictive quality. They are well suited to identify complex patterns and relationships in process data. MLPs in particular are used in a wide variety of use cases and data sets. Many publications show that MLPs are superior to other machine learning models such as SVMs or random forests in terms of their prediction performance. CNNs are by far the most widely used deep learning models for predictive quality. As they are very well suited for pattern recognition in image data, they are often used in visual quality inspection. Regarding other deep learning models, only LSTMs are currently studied in a few publications. It is noteworthy that other popular models such as transformer networks are not found in the reviewed publications.

Baseline models Conducting experimental comparisons of different models or model variants is an important part in machine learning. About 50% of publications compare their prime model with other baseline models. These are mainly other machine learning models that are established for nonlinear learning problems, such as SVM, random forest, decision tree, and k-nearest neighbor. Although these models do not achieve the same performances as MLPs in

most experiments, they also show strong versatility for different predictive quality use cases.

Models and data modalities The selection of an appropriate ML or DL model for a particular predictive quality use case depends on many factors. One of them is the data modality. Looking at the publications with respect to which model is used for which data modality, some common approaches become apparent. If the data basis consists of images, most publications (9 out of 11) use a CNN-based approach (Dimitriou et al., 2020; Ferguson et al., 2018; Jun et al., 2021; Li et al., 2018; Liu et al., 2021; Martinez et al., 2020; Yun et al., 2020; Wang et al., 2021; Zhang et al., 2019a). For numerical/continuous data, an MLP is used in about half of the corresponding publications, followed by random forest, SVM, and other models. For time series data, no common model is established yet. As mentioned above, the common approach for raw time series data coming from sensors is to transform them first into numerical/continuous variables using feature extraction and training then a machine learning model on it. However, of the few publications that train models on the time series directly, five publications use LSTMs (Essien & Giannetti, 2020; Meyes et al., 2019; Ståhl et al., 2019; Zhang et al., 2018, 2019b) and three publications use CNNs (Goldman et al., 2021; Hsu & Liu, 2021; Sarivan et al., 2020).

Future research directions

Based on the obtained results and conclusions, we provide an outlook on future directions of predictive quality research. We believe that these research directions can address the identified gaps as well as boost the prevalence of predictive quality in research and industry.

Synthetic data generation Machine learning and especially deep learning models typically require large amounts of training data. Therefore, solutions have to be researched and developed to increase the data quantity and to overcome the identified data sparsity in predictive quality scenarios. One promising approach is to generate synthetic training data using generative deep learning models. It has been shown that they are suitable for generating realistic data in large quantities at low cost (Mao et al., 2019; Nikolenko, 2021; Pashevich et al., 2019). In the predictive quality context, they could be used with manufacturing simulations to synthetically replicate data for rare process variations and product defects. In addition, we propose the establishment and further development of data augmentation methods for manufacturing process data, in particular sensor and time series data. We thus refer to research on data augmentation for time series problems (Iwana & Uchida, 2021; Wen et al., 2021).

Benchmark data sets We see great potential in research to make results and data sets more accessible to other sci-

entists. Therefore, we recommend establishing benchmark data for predictive quality to be used for the evaluation of new approaches. This could be data for the classification of product defects or for the numerical prediction of quality values based on sensor data. In addition, we recommend scientists to use already existing data sets for their own research work. At this point, we refer to publications that provide an overview of datasets and repositories, such as for surface defect detection (Chen et al., 2021) or for machine learning in production (Krauß et al., 2019).

Novel deep learning methods The review results show that among the existing deep learning models, only CNNs and LSTM-based models are investigated for predictive quality. However, in the state of the art of deep learning, new types of methods have already been established. Among them is the Transformer network, an attention-based method that performs very well on sequential data (Vaswani et al., 2017) and image data (Khan et al., 2021). Though applications of Transformer networks for predictive quality are not yet known to us, there are already applications in the predictive maintenance context (Liu et al., 2020a; Mo et al., 2021). Also strongly researched in the deep learning field are graph neural networks (Zhou et al., 2020), which are suitable for graph data and can therefore be useful for pattern recognition in CAD or simulation data. We therefore see the potential for these methods to be equally well suited for predictive quality scenarios.

Time series classification and forecasting While deep learning on image data has gained acceptance in predictive quality (via CNNs), there is still a large research potential for training models on raw sensor data or time series data. In current deep learning research, there are already a number of different model approaches for performing time series classification (Ismail Fawaz et al., 2019) or forecasting (Lim & Zohren, 2021). Considering predictive quality scenarios, these approaches could be suitable to perform quality prediction based on temporal sensor data in the manufacturing process.

Transfer learning and continual learning The current works in predictive quality are mainly based on the assumption that the training data basis is representative for the given problem. However, this assumption is often not valid in industrial production, since manufacturing processes are subject to continuous changes (e.g. the production of new products). Process changes mean that previously trained learning models no longer work sufficiently well, which is why a lot of new training data has to be generated at great cost. The emerging fields of transfer learning and continual learning can overcome this challenge by training data-efficient and cost-effective models over process variants. In the current state of research, some efforts for the use of transfer learning (Maschler & Weyrich, 2021b; Maschler et al., 2021a; Tercan et al., 2018, 2019) and continual learning (Tercan et

al., 2021) in manufacturing exist. We see great potential in further research of these fields for predictive quality.

Integration and deployment With regard to an establishment of predictive quality in industrial manufacturing systems, further research work has to be done. On the one hand, we see the need to evaluate predictive quality solutions in real quality assurance processes. This includes the development of strategies for automated feedback of model decisions to humans and systems as well as the evaluation of approaches in terms of their impact on process quality by using quality-oriented metrics (e.g. yield rate). Another important aspect is the operationalization and automation of model training and model use. In the current state of research, first approaches of MLOPs (machine learning operationalization) exist to continuously monitor, integrate and deliver ML models in business environments (Cardoso Silva et al., 2020; Garg et al., 2021). We see great potential of MLOPs strategies for continuous integration of predictive quality models. Furthermore, there is a need for certification of predictive quality processes that guarantees the reliability of the ML models and thus enables their use in industrial manufacturing processes. To the best of our knowledge, there are currently no established approaches for certifying machine learning methods in industries.

Summary

This review paper provided a comprehensive overview of 81 scientific publications between 2012 and 2021 that address predictive quality in manufacturing processes. The publications were categorized and evaluated based on three guiding questions. The first question was to discover which manufacturing processes and quality criteria are addressed in the publications. The categorization was done according to the DIN 8580. On the one hand, the results show that predictive quality is used in a wide variety of manufacturing processes, estimating various quality metrics or defect types. On the other hand, an imbalance in the process groups is evident. While a lot of research is done in process groups such as cutting and joining, hardly any publications lie in the fields of coating and changing material properties.

The second question dealt with the data used in the publications for training and evaluating the learning models. Here it was shown that for a large part of the publications, process parameters or sensor data were gathered from a real manufacturing process and merged with the quality values. The generation of the data was often carried out experimentally by varying a few parameters or boundary conditions. With regard to data modality, it appeared that numerical quantities as well as image data (e.g. from product measurements) are playing an increasingly important role.

The third question addressed the machine learning and deep learning models used for predictive quality. The results of the review showed that especially models based on artificial neural networks (MLPs) and deep learning (mainly CNNs) were very much in focus. While the MLP was used in versatile ways, CNNs were commonly used on image data. In about half of the publications, the models were also compared experimentally with other models. Here, it was observed that popular machine learning methods such as SVMs and random forests were frequently used for comparison.

Based on the obtained results, central challenges for predictive quality research were derived, which need to be addressed in future work. On the one hand, these include the tackling of sparse data sets in the manufacturing context and the generation of benchmark data for more comparability in research. On the other hand, there is a lack of approaches to integrate and deploy sustainable and robust predictive quality solutions in real production processes.

In conclusion, predictive quality is a very heterogeneous and highly researched field in the manufacturing world. The relevance and popularity of the field will likely continue to increase in the coming years. The current state of research highlights the great potential that data-driven methods of machine learning and deep learning bring to quality assurance and inspection. Yet the use cases, approaches, and results are still viewed in a very isolated way. As such, there is still much to be done in research to overcome this isolated view and enable greater prevalence of predictive quality in the future.

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/>.

Appendix A: Literature

See Table 10

Table 10 List of all selected publications for the review

Publication	Type	Name
Alvarado-Iniesta et al. (2012)	Journal	Journal of Applied Research and Technology
Ge et al. (2012)	Journal	Industrial Engineering & Chemistry Research
Nurwaha and Wang (2012)	Journal	Fibres & Textiles in Eastern Europe
Hamidinejad et al. (2012)	Journal	Materials and Design
Lieber et al. (2013)	Conference	CIRP Conference on Manufacturing Systems
Neto et al. (2013)	Conference	CIRP Conference on Intelligent Computation in Manufacturing Engineering
Abd-Ellatif (2013)	Conference	Alexandria Engineering Journal
Dhas and Kumanan (2014)	Journal	Applied Soft Computing
Hossain and Ahmad (2014)	Journal	International Conference on Mechanical Engineering
Ciancio et al. (2015)	Conference	CIRP Conference on Intelligent Computation in Manufacturing Engineering
Acayaba and de Escalona (2015)	Journal	CIRP Journal of Manufacturing Science and Technology
Elangovan et al. (2015)	Conference	International Symposium on Big Data and Cloud Computing
de Oliveira Leite et al. (2015)	Journal	Applied Soft Computing
Vrabel et al. (2016)	Conference	CIRP Conference on Manufacturing Systems
Martín et al. (2016)	Journal	Materials Science & Engineering: A
Tercan et al. (2016)	Conference	Changeable, Agile, Reconfigurable & Virtual Production Conference
Yu et al. (2016)	Journal	Journal of Mechanical Science and Technology
Ai et al. (2016)	Journal	Optics and Lasers in Engineering
Kayabasi et al. (2017)	Journal	Solar Energy
Varma et al. (2017)	Journal	Material Today: Proceedings
Tušar et al. (2017)	Journal	Applied Soft Computing
Tercan et al. (2017)	Journal	Production Engineering
Zhang and Lei (2017)	Conference	Global Congress on Manufacturing and Management
Bustillo et al. (2018)	Journal	Journal of Manufacturing Systems
Ferguson et al. (2018)	Journal	Smart and Sustainable Manufacturing Systems
Li et al. (2018)	Conference	IFAC Workshop on Mining, Mineral and Metal Processing
Zhang et al. (2018)	Conference	International Conference on Through-life Engineering Services
Kim et al. (2018)	Journal	International Journal of Computer Integrated Manufacturing
Lee et al. (2018)	Journal	Sensors
Mulrennan et al. (2018)	Journal	Polymer Testing
Thomas et al. (2018)	Journal	Computers in Industry
Zhang et al. (2019a)	Journal	Additive Manufacturing
Gyasi et al. (2019)	Conference	International Conference on Flexible Automation and Intelligent Manufacturing
Garcia et al. (2019)	Journal	Journal of Intelligent Manufacturing
Zhang et al. (2019b)	Journal	Computers in Industry
Li et al. (2019)	Journal	Robotics and Computer-Integrated Manufacturing
Natesh et al. (2019)	Journal	Measurement
Moreira et al. (2019)	Journal	Computers & Industrial Engineering
Papananias et al. (2019)	Conference	CIRP Conference on Modelling of Machining Operations

Table 10 continued

Publication	Type	Name
Ståhl et al. (2019)	Journal	Applied Mathematical Modelling
Oh et al. (2019)	Journal	Reliability Engineering and System Safety
Meyes et al. (2019)	Conference	North American Manufacturing Research Conference
Yu et al. (2019)	Journal	Wear
Lei et al. (2019)	Journal	Journal of Manufacturing Processes
Gejji et al. (2020)	Conference	International Conference Interdisciplinarity in Engineering
Gaikwad et al. (2020)	Journal	Additive Manufacturing
Lutz et al. (2020)	Conference	CIRP Conference on Manufacturing Systems
Zimmerling et al. (2020)	Conference	International Conference on Material Forming
Dib et al. (2020)	Journal	Neural Computing and Applications
Dimitriou et al. (2020)	Journal	IEEE Transactions on Industrial Electronics
Doltsinis et al. (2020)	Journal	IEEE Transactions on Automation Science and Engineering
Zahrani et al. (2020)	Conference	CIRP Conference on Photonic Technologies
Essien and Giannetti (2020)	Journal	IEEE Transactions on Industrial Informatics
Serin et al. (2020a)	Conference	International Conference on Flexible Automation and Intelligent Manufacturing
Sarivan et al. (2020)	Conference	International Conference on Flexible Automation and Intelligent Manufacturing
Bartlett et al. (2020)	Journal	Materials Science & Engineering: A
Jiao et al. (2020)	Journal	Applied Sciences
Yun et al. (2020)	Journal	Journal of Manufacturing Systems
Ke and Huang (2020)	Journal	Polymers
Li et al. (2020a)	Journal	Scanning
Li et al. (2020b)	Journal	Materials and Design
Liu et al. (2020b)	Journal	Journal of Intelligent Manufacturing
Li and Anand (2020)	Journal	Journal of Manufacturing Processes
Kebisek et al. (2020)	Journal	IFAC-PapersOnLine
Martinez et al. (2020)	Journal	International Journal of Advanced Manufacturing Technology
McDonnell et al. (2021)	Journal	Journal of Intelligent Manufacturing
Wagner et al. (2020)	Conference	CIRP Conference on Computer Aided Tolerancing
Schmitt et al. (2020a)	Journal	Advanced Engineering Informatics
Schorr et al. (2020a)	Conference	CIRP Conference on Manufacturing Systems
Schorr et al. (2020b)	Conference	North American Manufacturing Research Conference
Klein et al. (2020)	Conference	CIRP Conference on Manufacturing Systems
Nguyen et al. (2020)	Journal	Optics and Lasers in Engineering
Yu et al. (2020)	Journal	Metals
Zhu et al. (2020)	Conference	CIRP Design
Turetsky et al. (2021)	Journal	Energy Storage Materials
Goldman et al. (2021)	Conference	International Conference on Industry 4.0 and Smart Manufacturing
Du et al. (2021)	Journal	Advances in Manufacturing
Hsu and Liu (2021)	Journal	Journal of Intelligent Manufacturing
Jun et al. (2021)	Journal	Textile Research Journal
Wang et al. (2021)	Journal	Journal of Manufacturing Processes
Liu et al. (2021)	Journal	Optics and Lasers in Engineering

References

- Abd-Ellatif, S. A. M. (2013). Optimizing sliver quality using artificial neural networks in ring spinning. *Alexandria Engineering Journal*, 52(4), 637–642. <https://doi.org/10.1016/j.aej.2013.09.007>
- Acayaba, G. M. A., & de Escalona, P. M. (2015). Prediction of surface roughness in low speed turning of AISI316 austenitic stainless steel. *CIRP Journal of Manufacturing Science and Technology*, 11, 62–67. <https://doi.org/10.1016/j.cirpj.2015.08.004>
- Ademujimi, T. T., Brundage, M. P., & Prabhu, V. V. (2017). A review of current machine learning techniques used in manufacturing diagnosis. In H. Lödging, R. Riedel, K.-D. Thoben, G. von Cieminski, & D. Kiritsis (Eds.), *Advances in production management systems*. IFIP advances in information and communication technology (Vol. 513, pp. 407–415). Springer. https://doi.org/10.1007/978-3-319-66923-6_48
- Ai, Y., Shao, X., Jiang, P., Li, P., Liu, Y., & Liu, W. (2016). Welded joints integrity analysis and optimization for fiber laser welding of dissimilar materials. *Optics and Lasers in Engineering*, 86, 62–74. <https://doi.org/10.1016/j.optlaseng.2016.05.011>
- Alvarado-Iniesta, A., Valles-Rosales, D. J., Garcia-Alcaraz, J. L., & Maldonado-Macias, A. (2012). A recurrent neural network for warpage prediction in injection molding. *Journal of Applied Research and Technology*, 10, 912–919.
- Bartlett, J. L., Jarama, A., Jones, J., & Li, X. (2020). Prediction of microstructural defects in additive manufacturing from powder bed quality using digital image correlation. *Materials Science and Engineering: A*, 794, 140002. <https://doi.org/10.1016/j.msea.2020.140002>
- Bustillo, A., Urbikain, G., Perez, J. M., Pereira, O. M., de Lacalle, Lopez, & Luis, N. (2018). Smart optimization of a friction-drilling process based on boosting ensembles. *Journal of Manufacturing Systems*, 48(C, SI), 108–121. <https://doi.org/10.1016/j.jmsy.2018.06.004>
- Cardoso Silva, L., Rezende Zagatti, F., Silva Sette, B., Nildaimon dos Santos Silva, L., Lucredio, D., Furtado Silva, D., & de Medeiros Caseli, H. (2020). Benchmarking machine learning solutions in production. In *2020 19th IEEE international conference on machine learning and applications (ICMLA)* (pp. 626–633). IEEE. <https://doi.org/10.1109/ICMLA51294.2020.00104>
- Chen, Y., Ding, Y., Zhao, F., Zhang, E., Wu, Z., & Shao, L. (2021). Surface defect detection methods for industrial products: A review. *Applied Sciences—Basel*, 11(16), 7657. <https://doi.org/10.3390/app11167657>
- Ciancio, C., Citrea, T., Ambrogio, G., Filice, L., & Musmanno, R. (2015). Design of a high performance predictive tool for forging operation. *Procedia CIRP*, 33, 173–178. <https://doi.org/10.1016/j.procir.2015.06.032>
- Dalzocho, J., Kunst, R., Pignaton, E., Binotto, A., Sanyal, S., Favilla, J., & Barbosa, J. (2020). Machine learning and reasoning for predictive maintenance in industry 4.0: Current status and challenges. *Computers in Industry*, 123, 103298. <https://doi.org/10.1016/j.compind.2020.103298>
- de Oliveira Leite, W., Campos Rubio, J. C., Duduch, J. G., & de Almeida, P. E. M. (2015). Correcting geometric deviations of CNC machine-tools: An approach with artificial neural networks. *Applied Soft Computing*, 36, 114–124. <https://doi.org/10.1016/j.asoc.2015.07.014>
- Dhas, J. E. R., & Kumanan, S. (2014). Neuro evolutionary model for weld residual stress prediction. *Applied Soft Computing*, 14, 461–468. <https://doi.org/10.1016/j.asoc.2013.08.019>
- Dib, M., Oliveira, N. J. C., Marques, A. E., Oliveira, M. C., Fernandes, J. V., Ribeiro, B., & Prates, P. (2020). Single and ensemble classifiers for defect prediction in sheet metal forming under variability. *Neural Computing & Applications*, 32(16), 12335–12349. <https://doi.org/10.1007/s00521-019-04651-6>
- Dimitriou, N., Leontaris, L., Vafeiadis, T., Ioannidis, D., Wotherspoon, T., Tinker, G., & Tzovaras, D. (2020). Fault diagnosis in microelectronics attachment via deep learning analysis of 3-D laser scans. *IEEE Transactions on Industrial Electronics*, 67(7), 5748–5757. <https://doi.org/10.1109/TIE.2019.2931220>
- DIN 8580:2003-09 (2003). Manufacturing processes—terms and definitions, division.
- Doltsinis, S., Krestenitis, M., & Doulgeri, Z. (2020). A machine learning framework for real-time identification of successful snap-fit assemblies. *IEEE Transactions on Automation Science and Engineering*, 17(1), 513–523. <https://doi.org/10.1109/TASE.2019.2932834>
- Du, C., Ho, C. L., & Kaminski, J. (2021). Prediction of product roughness, profile, and roundness using machine learning techniques for a hard turning process. *Advances in Manufacturing*, 9, 206–215. <https://doi.org/10.1007/s40436-021-00345-2>
- Elangovan, M., Sakthivel, N., Saravanamurugan, S., Nair, B. B., & Sugumar, V. (2015). Machine learning approach to the prediction of surface roughness using statistical features of vibration signal acquired in turning. *Procedia Computer Science*, 50, 282–288. <https://doi.org/10.1016/j.procs.2015.04.047>
- Essien, A., & Giannetti, C. (2020). A deep learning model for smart manufacturing using convolutional ISTM neural network autoencoders. *IEEE Transactions on Industrial Informatics*, 16(9), 6069–6078. <https://doi.org/10.1109/TII.2020.2967556>
- Fahle, S., Prinz, C., & Kuhlenkötter, B. (2020). Systematic review on machine learning (ML) methods for manufacturing processes—identifying artificial intelligence (AI) methods for field application. *Procedia CIRP*, 93, 413–418. <https://doi.org/10.1016/j.procir.2020.04.109>
- Ferguson, M., Ak, R., Lee, Y.-T.T., & Law, K. H. (2018). Detection and segmentation of manufacturing defects with convolutional neural networks and transfer learning. *Smart and Sustainable Manufacturing Systems*, 2(1), 137–164. <https://doi.org/10.1520/SSMS20180033>
- Gaikwad, A., Giera, B., Guss, G. M., Forien, J.-B., Matthews, M. J., & Rao, P. (2020). Heterogeneous sensing and scientific machine learning for quality assurance in laser powder bed fusion—a single-track study. *Additive Manufacturing*, 36. <https://doi.org/10.1016/j.addma.2020.101659>
- Garcia, V., Salvador Sanchez, J., Alberto Rodriguez-Picon, L., Carlos Mendez-Gonzalez, L., & de Jesus Ochoa-Dominguez, H. (2019). Using regression models for predicting the product quality in a tubing extrusion process. *Journal of Intelligent Manufacturing*, 30(6), 2535–2544. <https://doi.org/10.1007/s10845-018-1418-7>
- Garg, S., Pundir, P., Rathee, G., Gupta, P. K., Garg, S., & Ahlawat, S. (2021). On continuous integration / continuous delivery for automated deployment of machine learning models using mlops. In *2021 IEEE fourth international conference on artificial intelligence and knowledge engineering (AIKE)* (pp. 25–28). IEEE. <https://doi.org/10.1109/AIKE52691.2021.00010>
- Ge, Z., Song, Z., & Gao, F. (2012). Statistical prediction of product quality in batch processes with limited batch-cycle data. *Industrial & Engineering Chemistry Research*, 51(35), 11409–11416. <https://doi.org/10.1021/ie202554r>
- Gejji, A., Shukla, S., Pimparkar, S., Pattharwala, T., & Bewoor, A. (2020). Using a support vector machine for building a quality prediction model for center-less honing process. *Procedia Manufacturing*, 46, 600–607. <https://doi.org/10.1016/j.promfg.2020.03.086>
- Goldman, C. V., Baltaxe, M., Chakraborty, D., & Arinez, J. (2021). Explaining learning models in manufacturing processes. *Procedia Computer Science*, 180, 259–268. <https://doi.org/10.1016/j.procs.2021.01.163>

- Gyasi, E. A., Kah, P., Penttilä, S., Ratava, J., Handroos, H., & Sanbao, L. (2019). Digitalized automated welding systems for weld quality predictions and reliability. *Procedia Manufacturing*, 38, 133–141. <https://doi.org/10.1016/j.promfg.2020.01.018>
- Hamidinejad, S., Kolahan, F., & Kokabi, A. (2012). The modeling and process analysis of resistance spot welding on galvanized steel sheets used in car body manufacturing. *Materials & Design*, 34, 759–767. <https://doi.org/10.1016/j.matdes.2011.06.064>
- He, Y., Song, K., Meng, Q., & Yan, Y. (2020). An end-to-end steel surface defect detection approach via fusing multiple hierarchical features. *IEEE Transactions on Instrumentation and Measurement*, 69(4), 1493–1504. <https://doi.org/10.1109/TIM.2019.2915404>
- Hehenberger, P. (2020). Qualitätsmanagement in der Produktion. In P. Hehenberger (Ed.), *Computerunterstützte Produktion* (pp. 197–228). Springer. https://doi.org/10.1007/978-3-662-60876-0_8
- Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural Computation*, 9(8), 1735–1780. <https://doi.org/10.1162/neco.1997.9.8.1735>
- Hossain, M. S. J., & Ahmad, N. (2014). A neuro-fuzzy approach to select cutting parameters for commercial die manufacturing. *Procedia Engineering*, 90, 753–759. <https://doi.org/10.1016/j.proeng.2014.11.809>
- Hsu, C.-Y., & Liu, W.-C. (2021). Multiple time-series convolutional neural network for fault detection and diagnosis and empirical study in semiconductor manufacturing. *Journal of Intelligent Manufacturing*, 32(3), 823–836. <https://doi.org/10.1007/s10845-020-01591-0>
- Ismail Fawaz, H., Forestier, G., Weber, J., Idoumghar, L., & Muller, P.-A. (2019). Deep learning for time series classification: A review. *Data Mining and Knowledge Discovery*, 33(4), 917–963. <https://doi.org/10.1007/s10618-019-00619-1>
- Iwana, B. K., & Uchida, S. (2021). An empirical survey of data augmentation for time series classification with neural networks. *PLoS ONE*, 16(7). <https://doi.org/10.1371/journal.pone.0254841>
- Jiao, A., Zhang, G., Liu, B., & Liu, W. (2020). Prediction of manufacturing quality of holes based on a BP neural network. *Applied Sciences*. <https://doi.org/10.3390/app10062108>
- Jun, X., Wang, J., Zhou, J., Meng, S., Pan, R., & Gao, W. (2021). Fabric defect detection based on a deep convolutional neural network using a two-stage strategy. *Textile Research Journal*, 91(1–2), 130–142. <https://doi.org/10.1177/0040517520935984>
- Kayabasi, E., Ozturk, S., Celik, E., & Kurt, H. (2017). Determination of cutting parameters for silicon wafer with a diamond wire saw using an artificial neural network. *Solar Energy*, 149, 285–293. <https://doi.org/10.1016/j.solener.2017.04.022>
- Ke, K.-C., & Huang, M.-S. (2020). Quality prediction for injection molding by using a multilayer perceptron neural network. *Polymers*, 12(8), 1812. <https://doi.org/10.3390/polym12081812>
- Kebisek, M., Tanuska, P., Spendla, L., Kotianova, J., & Strelec, P. (2020). Artificial intelligence platform proposal for paint structure quality prediction within the industry 4.0 concept. *IFAC-PapersOnLine*, 53(2), 11168–11174. <https://doi.org/10.1016/j.ifacol.2020.12.299>
- Khan, S., Naseer, M., Hayat, M., Zamir, S. W., Khan, F. S., & Shah, M. (2021). Transformers in vision: A survey. arXiv Preprint. [arXiv:2101.01169v4](https://arxiv.org/abs/2101.01169v4)
- Kim, A., Oh, K., Jung, J.-Y., & Kim, B. (2018). Imbalanced classification of manufacturing quality conditions using cost-sensitive decision tree ensembles. *International Journal of Computer Integrated Manufacturing*, 31(8, SI), 701–717. <https://doi.org/10.1080/0951192X.2017.1407447>
- Klein, S., Schorr, S., & Bähre, D. (2020). Quality prediction of honed bores with machine learning based on machining and quality data to improve the honing process control. *Procedia CIRP*, 93, 1322–1327. <https://doi.org/10.1016/j.procir.2020.03.055>
- Köksal, G., Batmaz, İ., & Testik, M. C. (2011). A review of data mining applications for quality improvement in manufacturing industry. *Expert Systems with Applications*, 38(10), 13448–13467. <https://doi.org/10.1016/j.eswa.2011.04.063>
- Krauß, J., Dorißen, J., Mende, H., Frye, M., & Schmitt, R. H. (2019). Machine learning and artificial intelligence in production: Application areas and publicly available data sets. In J. P. Wulfsberg, W. Hintze, & B.-A. Behrens (Eds.), *Production at the leading edge of technology* (pp. 493–501). Springer Vieweg. https://doi.org/10.1007/978-3-662-60417-5_49
- Lee, J., Noh, S. D., Kim, H.-J., & Kang, Y.-S. (2018). Implementation of cyber-physical production systems for quality prediction and operation control in metal casting. *Sensors*, 18(5), 1428. <https://doi.org/10.3390/s18051428>
- Lei, Z., Shen, J., Wang, Q., & Chen, Y. (2019). Real-time weld geometry prediction based on multi-information using neural network optimized by PCA and GA during thin-plate laser welding. *Journal of Manufacturing Processes*, 43, 207–217. <https://doi.org/10.1016/j.jmapro.2019.05.013>
- Li, J., Su, Z., Geng, J., & Yin, Y. (2018). Real-time detection of steel strip surface defects based on improved yolo detection network. *IFAC-PapersOnLine*, 51(21), 76–81. <https://doi.org/10.1016/j.ifacol.2018.09.412>
- Li, L., & Anand, S. (2020). Hatch pattern based inherent strain prediction using neural networks for powder bed fusion additive manufacturing. *Journal of Manufacturing Processes*, 56, 1344–1352. <https://doi.org/10.1016/j.jmapro.2020.04.030>
- Li, L., Liu, Di., Liu, J., Zhou, H.-G., & Zhou, J. (2020). Quality prediction and control of assembly and welding process for ship group product based on digital twin. *SCANNING*, 2020,. <https://doi.org/10.1155/2020/3758730>
- Li, Y., Yu, B., Wang, B., Lee, T. H., & Banu, M. (2020). Online quality inspection of ultrasonic composite welding by combining artificial intelligence technologies with welding process signatures. *Materials & Design*, 194,. <https://doi.org/10.1016/j.matdes.2020.108912>
- Li, Z., Zhang, Z., Shi, J., & Wu, D. (2019). Prediction of surface roughness in extrusion-based additive manufacturing with machine learning. *Robotics and Computer-Integrated Manufacturing*, 57, 488–495. <https://doi.org/10.1016/j.rcim.2019.01.004>
- Lieber, D., Stolpe, M., Konrad, B., Deuse, J., & Morik, K. (2013). Quality prediction in interlinked manufacturing processes based on supervised and unsupervised machine learning. *Procedia CIRP*, 7, 193–198. <https://doi.org/10.1016/j.procir.2013.05.033>
- Lim, B., & Zohren, S. (2021). Time-series forecasting with deep learning: A survey. *Philosophical Transactions: Series A, Mathematical, Physical, and Engineering Sciences*, 379(2194), 20200209. <https://doi.org/10.1098/rsta.2020.0209>
- Lin, X., Zhu, K., Fuh, J. Y. H., & Duan, X. (2022). Metal-based additive manufacturing condition monitoring methods: From measurement to control. *ISA Transactions*, 120, 147–166. <https://doi.org/10.1016/j.isatra.2021.03.001>
- Liu, H., Liu, Z., Jia, W., Lin, X., & Zhang, S. (2020). A novel transformer-based neural network model for tool wear estimation. *Journal of Intelligent Manufacturing*, 31(6). <https://doi.org/10.1088/1361-6501/ab7282>
- Liu, Z., Tang, R., Duan, G., & Tan, J. (2021). Truingdet: Towards high-quality visual automatic defect inspection for mental surface. *Optics and Lasers in Engineering*, 138, 106423. <https://doi.org/10.1016/j.optlaseng.2020.106423>
- Liu, Z., Zhang, D., Jia, W., Lin, X., & Liu, H. (2020). An adversarial bidirectional serial-parallel LSTM-based QTD framework for product quality prediction. *Journal of Intelligent Manufacturing*, 31(6), 1511–1529. <https://doi.org/10.1007/s10845-019-01530-8>
- Lopez, F., Saez, M., Shao, Y., Balta, E. C., Moyne, J., Mao, Z. M., Barton, K., & Tilbury, D. (2017). Categorization of anomalies in smart manufacturing systems to support the selection of detection

- mechanisms. *IEEE Robotics and Automation Letters*, 2(4), 1885–1892. <https://doi.org/10.1109/LRA.2017.2714135>
- Lutz, B., Kisskalt, D., Regulin, D., & Franke, J. (2020). Ai-based approach for predicting the machinability under consideration of material batch deviations in turning processes. *Procedia CIRP*, 93, 1382–1387. <https://doi.org/10.1016/j.procir.2020.04.100>
- Mao, W., Liu, Y., Ding, L., & Li, Y. (2019). Imbalanced fault diagnosis of rolling bearing based on generative adversarial network: A comparative study. *IEEE Access*, 7, 9515–9530. <https://doi.org/10.1109/ACCESS.2018.2890693>
- Martín, Ó., Ahedo, V., Santos, J. I., De Tiedra, P., & Galán, J. M. (2016). Quality assessment of resistance spot welding joints of AISI304 stainless steel based on elastic nets. *Materials Science and Engineering A*, 676, 173–181. <https://doi.org/10.1016/j.msea.2016.08.112>
- Martinez, P., Al-Hussein, M., & Ahmad, R. (2020). Intelligent vision-based online inspection system of screw-fastening operations in light-gauge steel frame manufacturing. *International Journal of Advanced Manufacturing Technology*, 109(3–4), 645–657. <https://doi.org/10.1007/s00170-020-05695-y>
- Maschler, B., Braun, D., Jazdi, N., & Weyrich, M. (2021a). Transfer learning as an enabler of the intelligent digital twin. *Procedia CIRP*, 100, 127–132. <https://doi.org/10.1016/j.procir.2021.05.020>
- Maschler, B., & Weyrich, M. (2021b). Deep transfer learning for industrial automation. *IEEE Industrial Electronics Magazine*. <https://doi.org/10.1109/MIE.2020.3034884>
- Mayr, A., Kibkalt, D., Meiners, M., Lutz, B., Schäfer, F., Seidel, R., Selmaier, A., Fuchs, J., Metzner, M., Blank, A., & Franke, J. (2019). Machine learning in production—potentials, challenges and exemplary applications. *Procedia CIRP*, 86, 49–54. <https://doi.org/10.1016/j.procir.2020.01.035>
- McDonnell, M. D. T., Arnaldo, D., Pelletier, E., Grant-Jacob, J. A., Praeger, M., Karnakis, D., Eason, R. W., & Mills, B. (2021). Machine learning for multi-dimensional optimisation and predictive visualisation of laser machining. *Journal of Intelligent Manufacturing*. <https://doi.org/10.1007/s10845-020-01717-4>
- Mery, D., Riffo, V., Zscherpel, U., Mondragón, G., Lillo, I., Zuccar, I., Lobel, H., & Carrasco, M. (2015). GDxray: The database of X-ray images for nondestructive testing. *Journal of Nondestructive Evaluation*, 34(4), 42. <https://doi.org/10.1007/s10921-015-0315-7>
- Meyes, R., Donauer, J., Schmeing, A., & Meisen, T. (2019). A recurrent neural network architecture for failure prediction in deep drawing sensory time series data. *Procedia Manufacturing*, 34, 789–797. <https://doi.org/10.1016/j.promfg.2019.06.205>
- Mo, Y., Wu, Q., Li, X., & Huang, B. (2021). Remaining useful life estimation via transformer encoder enhanced by a gated convolutional unit. *Journal of Intelligent Manufacturing*, 32(7), 1997–2006. <https://doi.org/10.1007/s10845-021-01750-x>
- Mohanraj, T., Shankar, S., Rajasekar, R., Sakthivel, N. R., & Pramanik, A. (2020). Tool condition monitoring techniques in milling process—a review. *Journal of Materials Research and Technology*, 9(1), 1032–1042. <https://doi.org/10.1016/j.jmrt.2019.10.031>
- Montazeri, M., & Rao, P. (2018). Sensor-based build condition monitoring in laser powder bed fusion additive manufacturing process using a spectral graph theoretic approach. *Journal of Manufacturing Science and Engineering*, 140(9), 216. <https://doi.org/10.1115/1.4040264>
- Moreira, L. C., Li, W. D., Lu, X., & Fitzpatrick, M. E. (2019). Supervision controller for real-time surface quality assurance in CNC machining using artificial intelligence. *Computers & Industrial Engineering*, 127, 158–168. <https://doi.org/10.1016/j.cie.2018.12.016>
- Mulrennan, K., Donovan, J., Creedon, L., Rogers, I., Lyons, J. G., & McAfee, M. (2018). A soft sensor for prediction of mechanical properties of extruded PLA sheet using an instrumented slit die and machine learning algorithms. *Polymer Testing*, 69, 462–469. <https://doi.org/10.1016/j.polymertesting.2018.06.002>
- Nalbach, O., Linn, C., Derouet, M., & Werth, D. (2018). Predictive quality: Towards a new understanding of quality assurance using machine learning tools. In W. Abramowicz & A. Paschke (Eds.), *Business information systems. Lecture Notes in business information processing* (Vol. 320, pp. 30–42). Springer. https://doi.org/10.1007/978-3-319-93931-5_3
- Natesh, M., Yun, L., Vendan, S. A., Kumar, K. R., Gao, L., Niu, X., Peng, X., & Garg, A. (2019). Experimental and numerical procedure for studying strength and heat generation responses of ultrasonic welding of polymer blends. *Measurement*, 132, 1–10. <https://doi.org/10.1016/j.measurement.2018.09.043>
- Neto, F., & d. C., Gerônimo, T. M., Cruz, C. D. L., Aguiar, P. R., and Bianchi, E. (2013). Neural models for predicting hole diameters in drilling processes. *Procedia CIRP*, 12, 49–54. <https://doi.org/10.1016/j.procir.2013.09.010>
- Nguyen, T. H., Lin, C.-K., Tung, P.-C., Nguyen-Van, C., & Ho, J.-R. (2020). An extreme learning machine for predicting Kerf waviness and heat affected zone in pulsed laser cutting of thin non-oriented silicon steel. *Optics and Lasers in Engineering*, 134. <https://doi.org/10.1016/j.optlaseng.2020.106244>
- Nikolenko, S. I. (2021). *Synthetic data for deep learning*. Springer optimization and its applications (Vol. 174). Springer. <https://doi.org/10.1007/978-3-030-75178-4>
- Nurwaha, D., & Wang, X. H. (2012). Using intelligent control systems to predict textile yarn quality. *Fibres & Textiles in Eastern Europe*, 20(1), 23–27.
- Nikolenko, S. I. (2021). *Synthetic data for deep learning*. Springer optimization and its applications (Vol. 174). Springer. <https://doi.org/10.1007/978-3-030-75178-4>
- Papananias, M., McLeay, T. E., Mahfouf, M., & Kadirkamanathan, V. (2019). An intelligent metrology informatics system based on neural networks for multistage manufacturing processes. *Procedia CIRP*, 82, 444–449. <https://doi.org/10.1016/j.procir.2019.04.148>
- Pashevich, A., Strudel, R., Kalevatykh, I., Laptsev, I., & Schmid, C. (2019). Learning to augment synthetic images for Sim2Real policy transfer. In *2019 IEEE/RSJ international conference on intelligent robots and systems (IROS)* (pp. 2651–2657). IEEE. <https://doi.org/10.1109/IROS40897.2019.8967622>
- Pfeifer, T., & Schmitt, R. (2021). *Masing Handbuch Qualitätsmanagement*. Carl Hanser Verlag, München and Wien, 6, überarbeitete auflage edition.
- PHM Society. (2020). *PHM data challenge*. PHM Society.
- Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You only look once: Unified, real-time object detection. In *29th IEEE conference on computer vision and pattern recognition* (pp. 779–788). IEEE. <https://doi.org/10.1109/CVPR.2016.91>
- Rostami, H., Dantan, J.-Y., & Homri, L. (2015). Review of data mining applications for quality assessment in manufacturing industry: Support vector machines. *International Journal of Metrology and Quality Engineering*, 6(4), 401. <https://doi.org/10.1051/ijmqe/2015023>
- Sarivan, I.-M., Greiner, J., Álvarez, D. D., Euteneuer, F., Reichenbach, M., Madsen, O., & Bøgh, S. (2020). Enabling real-time quality inspection in smart manufacturing through wearable smart devices and deep learning. *Procedia Manufacturing*, 51, 373–380. <https://doi.org/10.1016/j.promfg.2020.10.053>
- Schmitt, J., Boenig, J., Borggraeve, T., Beiting, G., & Deuse, J. (2020). Predictive model-based quality inspection using machine learning and edge cloud computing. *Advanced Engineering Informatics*, 45. <https://doi.org/10.1016/j.aei.2020.101101>
- Schmitt, R. H., Kurzhals, R., Ellerich, M. Nilgen, G., Schlegel, P., & Dietrich, E. (2020b). Predictive quality—data analytics in produzierenden unternehmen. *Internet of Production Turning*

- Data into Value*, 2020, 226–253. <https://doi.org/10.24406/IPT-N-589615>
- Schorr, S., Möller, M., Heib, J., & Bähre, D. (2020). In-process quality control of drilled and reamed bores using NC-internal signals and machine learning method. *Procedia CIRP*, 93, 1328–1333. <https://doi.org/10.1016/j.procir.2020.03.020>
- Schorr, S., Möller, M., Heib, J., & Bähre, D. (2020). Quality prediction of drilled and reamed bores based on torque measurements and the machine learning method of random forest. *Procedia Manufacturing*, 48, 894–901. <https://doi.org/10.1016/j.promfg.2020.05.127>
- Serin, G., Sener, B., Gudelek, M. U., Ozbayoglu, A. M., & Unver, H. O. (2020). Deep multi-layer perceptron based prediction of energy efficiency and surface quality for milling in the era of sustainability and big data. *Procedia Manufacturing*, 51, 1166–1177. <https://doi.org/10.1016/j.promfg.2020.10.164>
- Serin, G., Sener, B., Ozbayoglu, M., & Unver, H. O. (2020). Review of tool condition monitoring in machining and opportunities for deep learning. *International Journal of Advanced Manufacturing Technology*, 109(3–4), 953–974. <https://doi.org/10.1007/s00170-020-05449-w>
- Shang, C., & You, F. (2019). Data analytics and machine learning for smart process manufacturing: Recent advances and perspectives in the big data era. *Engineering*, 5(6), 1010–1016. <https://doi.org/10.1016/j.eng.2019.01.019>
- Sharp, M., Ak, R., Hedberg, Jr., & Thomas, (2018). A survey of the advancing use and development of machine learning in smart manufacturing. *Journal of Manufacturing Systems*, 48(C, SI), 170–179. <https://doi.org/10.1016/j.jmsy.2018.02.004>
- Ståhl, N., Mathiason, G., Falkman, G., & Karlsson, A. (2019). Using recurrent neural networks with attention for detecting problematic slab shapes in steel rolling. *Applied Mathematical Modelling*, 70, 365–377. <https://doi.org/10.1016/j.apm.2019.01.027>
- Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., & Rabinovich, A. (2015). Going deeper with convolutions. In *2015 IEEE conference on computer vision and pattern recognition (CVPR)* (pp. 1–9). IEEE. <https://doi.org/10.1109/CVPR.2015.7298594>
- Tao, F., Qi, Q., Liu, A., & Kusiak, A. (2018). Data-driven smart manufacturing. *Journal of Manufacturing Systems*, 48, 157–169. <https://doi.org/10.1016/j.jmsy.2018.01.006>
- Tercan, H., Deibert, P., & Meisen, T. (2021). Continual learning of neural networks for quality prediction in production using memory aware synapses and weight transfer. *Journal of Intelligent Manufacturing*, 35(3), 617. <https://doi.org/10.1007/s10845-021-01793-0>
- Tercan, H., Guajardo, A., Heinisch, J., Thiele, T., Hopmann, C., & Meisen, T. (2018). Transfer-learning: Bridging the gap between real and simulation data for machine learning in injection molding. *Procedia CIRP*, 72, 185–190. <https://doi.org/10.1016/j.procir.2018.03.087>
- Tercan, H., Guajardo, A., & Meisen, T. (2019). Industrial transfer learning: Boosting machine learning in production. In *2019 IEEE 17th international conference on industrial informatics (INDIN)* (pp. 274–279). IEEE. <https://doi.org/10.1109/INDIN41052.2019.8972099>
- Tercan, H., Khawli, T. A., Eppelt, U., Büscher, C., Meisen, T., & Jeschke, S. (2016). Use of classification techniques to design laser cutting processes. *Procedia CIRP*, 52, 292–297. <https://doi.org/10.1016/j.procir.2016.08.001>
- Tercan, H., Khawli, T. A., Eppelt, U., Büscher, C., Meisen, T., & Jeschke, S. (2017). Improving the laser cutting process design by machine learning techniques. *Production Engineering*, 11(2), 195–203. <https://doi.org/10.1007/s11740-017-0718-7>
- Thomas, P., El Haouzi, H. B., Suhner, M.-C., Thomas, A., Zimmermann, E., & Noyel, M. (2018). Using a classifier ensemble for proactive quality monitoring and control: The impact of the choice of classifiers types, selection criterion, and fusion process. *Computers in Industry*, 99, 193–204. <https://doi.org/10.1016/j.compind.2018.03.038>
- Tianchi. (2021). Xuelang manufacturing AI challenge data set on Tianchi.
- Turetskyy, A., Wessel, J., Herrmann, C., & Thiede, S. (2021). Battery production design using multi-output machine learning models. *Energy Storage Materials*, 38, 93–112. <https://doi.org/10.1016/j.ensm.2021.03.002>
- Tušar, T., Gantar, K., Koblar, V., Ženko, B., & Filipič, B. (2017). A study of overfitting in optimization of a manufacturing quality control procedure. *Applied Soft Computing*, 59, 77–87. <https://doi.org/10.1016/j.asoc.2017.05.027>
- Varma, N. S. K., Rajesh, S., Raju, K. S. R., & Raju, V. M. K. (2017). Neural network and fuzzy logic based prediction of surface roughness and MRR in cylindrical grinding process. *Materials Today: Proceedings*, 4(8), 8134–8141. <https://doi.org/10.1016/j.matpr.2017.07.154>
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, U., & Polosukhin, I. (2017). Attention is all you need. In *Proceedings of the 31st international conference on neural information processing systems (NIPS'17)* (pp. 6000–6010). Curran Associates. <https://doi.org/10.5555/3295222.3295349>
- Vrabel, M., Maňková, I., & Beňo, J. (2016). Monitoring and control of manufacturing process to assist the surface workpiece quality when drilling. *Procedia CIRP*, 41, 735–739. <https://doi.org/10.1016/j.procir.2015.12.092>
- Wagner, R., Fischer, J., Gauder, D., Haefner, B., & Lanza, G. (2020). Virtual in-line inspection for function verification in serial production by means of artificial intelligence. *Procedia CIRP*, 92, 63–68. <https://doi.org/10.1016/j.procir.2020.03.126>
- Wang, Q., Jiao, W., Wang, P., & Zhang, Y. (2021). A tutorial on deep learning-based data analytics in manufacturing through a welding case study. *Journal of Manufacturing Processes*, 63(SI), 2–13. <https://doi.org/10.1016/j.jmapro.2020.04.044>
- Weichert, D., Link, P., Stoll, A., Rüping, S., Ihlenfeldt, S., & Wrobel, S. (2019). A review of machine learning for the optimization of production processes. *International Journal of Advanced Manufacturing Technology*, 104(5–8), 1889–1902. <https://doi.org/10.1007/s00170-019-03988-5>
- Wen, Q., Sun, L., Yang, F., Song, X., Gao, J., Wang, X., & Xu, H. (2021). Time series data augmentation for deep learning: A survey. In M. Gini & Z.-H. Zhou (Eds.), *Proceedings of the 30th international joint conference on artificial intelligence*, California (pp. 4653–4660). <https://doi.org/10.24963/ijcai.2021/631>
- Xiong, W., Droppo, J., Huang, X., Seide, F., Seltzer, M., Stolcke, A., Yu, D., & Zweig, G. (2016). Achieving human parity in conversational speech recognition. arXiv Preprint. [arXiv:1610.05256v2](https://arxiv.org/abs/1610.05256v2)
- Yu, J., Lee, H., Kim, D.-Y., Kang, M., & Hwang, I. (2020). Quality assessment method based on a spectrometer in laser beam welding process. *Metals*, 10(6), 839. <https://doi.org/10.3390/met10060839>
- Yu, J., Sohn, Y., Park, Y. W., & Kwak, J.-S. (2016). The development of a quality prediction system for aluminum laser welding to measure plasma intensity using photodiodes. *Journal of Mechanical Science and Technology*, 30(10), 4697–4704. <https://doi.org/10.1007/s12206-016-0940-9>
- Yu, T., Li, Z., & Wu, D. (2019). Predictive modeling of material removal rate in chemical mechanical planarization with physics-informed machine learning. *Wear*, 426–427, 1430–1438. <https://doi.org/10.1016/j.wear.2019.02.012>
- Yun, J. P., Shin, W. C., Koo, G., Kim, M. S., Lee, C., & Lee, S. J. (2020). Automated defect inspection system for metal surfaces based on deep learning and data augmentation. *Journal of Manufacturing Systems*, 55, 317–324. <https://doi.org/10.1016/j.jmsy.2020.03.009>
- Zahrani, E. G., Hojati, F., Daneshi, A., Azarhoushang, B., & Wilde, J. (2020). Application of machine learning to predict the product

- quality and geometry in circular laser grooving process. *Procedia CIRP*, 94, 474–480. <https://doi.org/10.1016/j.procir.2020.09.167>
- Zhang, B., Liu, S., & Shin, Y. C. (2019). In-process monitoring of porosity during laser additive manufacturing process. *Additive Manufacturing*, 28, 497–505. <https://doi.org/10.1016/j.addma.2019.05.030>
- Zhang, J., Wang, P., & Gao, R. X. (2018). Modeling of layer-wise additive manufacturing for part quality prediction. *Procedia Manufacturing*, 16, 155–162. <https://doi.org/10.1016/j.promfg.2018.10.165>
- Zhang, J., Wang, P., & Gao, R. X. (2019). Deep learning-based tensile strength prediction in fused deposition modeling. *Computers in Industry*, 107, 11–21. <https://doi.org/10.1016/j.compind.2019.01.011>
- Zhang, Y.-L., & Lei, J.-H. (2017). Prediction of laser cutting roughness in intelligent manufacturing mode based on ANFIS. *Procedia Engineering*, 174, 82–89. <https://doi.org/10.1016/j.proeng.2017.01.152>
- Zhou, J., Cui, G., Hu, S., Zhang, Z., Yang, C., Liu, Z., Wang, L., Li, C., & Sun, M. (2020). Graph neural networks: A review of methods and applications. *AI Open*, 1(1), 57–81. <https://doi.org/10.1016/j.aiopen.2021.01.001>
- Zhu, Z., Ferreira, K., Anwer, N., Mathieu, L., Guo, K., & Qiao, L. (2020). Convolutional neural network for geometric deviation prediction in additive manufacturing. *Procedia CIRP*, 91, 534–539. <https://doi.org/10.1016/j.procir.2020.03.108>
- Zimmerling, C., Poppe, C., & Kärger, L. (2020). Estimating optimum process parameters in textile draping of variable part geometries—a reinforcement learning approach. *Procedia Manufacturing*, 47, 847–854. <https://doi.org/10.1016/j.promfg.2020.04.263>
- Zonta, T., da Costa, C. A., da Rosa Righi, R., de Lima, M. J., da Trindade, E. S., & Li, G. P. (2020). Predictive maintenance in the industry 4.0: A systematic literature review. *Computers & Industrial Engineering*, 150, <https://doi.org/10.1016/j.cie.2020.106889>

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