Received xx xxx 2020; | Revised xxx; | Accepted xxx

DOI: xxx/xxxx

arXiv:2010.06010v1 [cs.LG] 12 Oct 2020

ARTICLE TYPE

Machine Learning for Material Characterization with an Application for Predicting Mechanical Properties

Anke Stoll[1] | Peter Benner[2]

[1]Data-Mining and Machine Learning, Fraunhofer Institute for Machine Tools and Forming Technology IWU, Chemnitz, Germany

[2]Computational Methods in Systems and Control Theory, The Max Planck Institute for Dynamics of Complex Technical Systems, Magdeburg, Germany

Correspondence
*Anke Stoll Email: anke.stoll@iwu.fraunhofer.de

Present Address
Reichenhainer Straße 88, 09126 Chemnitz, Germany

Currently, the growth of material data from experiments and simulations is expanding beyond processable amounts. This makes the development of new data-driven methods for the discovery of patterns among multiple lengthscales and time-scales and structure-property relationships essential. These data-driven approaches show enormous promise within materials science. The following review covers machine learning applications for metallic material characterization. Many parameters associated with the processing and the structure of materials affect the properties and the performance of manufactured components. Thus, this study is an attempt to investigate the usefulness of machine learning methods for material property prediction. Material characteristics such as strength, toughness, hardness, brittleness or ductility are relevant to categorize a material or component according to their quality. In industry, material tests like tensile tests, compression tests or creep tests are often time consuming and expensive to perform. Therefore, the application of machine learning approaches is considered helpful for an easier generation of material property information. This study also gives an application of machine learning methods on small punch test data for the determination of the property ultimate tensile strength for various materials. A strong correlation between small punch test data and tensile test data was found which ultimately allows to replace more costly tests by simple and fast tests in combination with machine learning.

KEYWORDS:
Machine Learning, Material Characterization, Small Punch Test, Ultimate Yield Strength, Tensile Properties

1 | INTRODUCTION

The field of materials science relies on experiments and simulation-based models as tools for material characterization [1]. Material properties, such as their structure and behaviour, are critical to the potential application of the material of interest. More recently, the data generated by such experiments and simulations have created various opportunities for the application of data-driven methods. In addition to, e.g., the experimental trial and error approach or a physical metallurgy approach, machine learning (ML) methods for property prediction and material design have attracted a lot of attention in recent years, see e.g. [2–4].

While experimental investigations (the so called first paradigm of materials science) have been carried out since the stone and copper age, scientists of the 16th century started to describe physical relations by equations (second paradigm). Thus, analytical equations became a central instrument of theoretical physics which were able to complement the empirical and experimental

sciences. The 1950s marked the beginning of computational materials science and simulations, the third paradigm. Within this framework, computer experiments and simulations became possible, with the corresponding results being analyzed and interpreted like measured ones. It had to be recognized that many properties of materials cannot be described by a closed mathematical form as they are determined by several multi-level, intricate theoretical concepts. With the help of large amounts of data, hidden correlations, reflected in terms of structure and patterns in the data can be discovered that are not normally visible in small data sets. Thus, the fourth paradigm, data-driven science, of materials research was born[1,5].

However, it is not only an advantage to have a large data volume but it can also be a challenge to cope with tremendous amounts of data. Today, data are indeed more and more easily acquired and stored, due to huge progresses in sensors and ways to collect data on one side, and in storage devices on the other side. Nowadays, there is no hesitation in many domains in acquiring very large amounts of data without knowing in advance if they will be analyzed and how. The spectacular increase in the amount of data is not only found in the number of samples collected for example over time, but also in the number of attributes, or characteristics, that are simultaneously measured on a process. Data are gathered into vectors whose dimension correspond to the number of simultaneous measurements on the process. Growing dimensions result in high dimensional data, as each sample can be represented as a point or vector in a high-dimensional space. Working with high-dimensional data means working with data that are embedded in high-dimensional spaces[6] .The curse of dimensionality is the expression of all phenomena that appear with high-dimensional data, and that have most often unfortunate consequences on the behavior and performances of learning algorithms.

Contrary to the curse of dimensionality, databases in materials science are often limited in size due to expensive and time consuming data acquisition via experiments or simulations[7]. Then the insufficient data size for the training of a ML model compromises the learning success and suitable new approaches for small datasets have to be found.

This work contains a literature survey which covers an overview of ML for materials science and specifically for metallic material characterization. As the measurement of such parameters is often expensive and time consuming obtained via experiments, alternative basic tests, such as the small punch test (SPT) can be an option if it can be shown that the same material property information can be extracted.

There is a wide range of ML approaches based on SPT data which will be presented. Furthermore, in Section 3 an example is described which uses ML for the prediction of tensile properties of a insert-material-type based on SPT data. The objective of this study is to investigate whether it is possible to find a ML model which predicts/determines the tensile properties of a material from SPT data[8]. Section 4 concludes this paper by giving an outlook on further research perspectives.

2 | STATE OF THE ART

2.1 | Overview - machine learning for materials science

With ML, given enough data and a data-driven algorithm for rule discovery, a computer is able to determine physical laws which lead to the given data without human input[9,10]. Traditional computational approaches use the computer for the employment of a hard-coded algorithm provided by a human expert. By contrast, ML approaches learn the rules that underlie a dataset by assessing a portion of that data and building a model to make predictions[9]. However, the human still needs to choose suitable ML models which supposedly represent the data well and do manual (sub-)tasks in pre-processing and feature generation.

The existence of large amounts of data makes the use of ML models possible and enables data-driven knowledge to be obtained and patterns to be discovered. On the other hand, big data and their high dimensionality lead to difficult computational and statistical challenges, such as scalability and memory shortage, noise accumulation, interference correlation, incidental endogeneity and measurement errors[11].

Materials science is an interesting field of application for big data methods and ML approaches which is beginning to show enormous promise. Four primary elements are critical in materials science and engineering: processing, structure, properties, and performance[12,13]. There is no general agreement, however, on how these elements are interconnected. ML methods can be applied to the so called process-structure-property-performance chain for learning more about the intrinsic interrelations of these components. One main goal is the enabling, acceleration and simplification of the discovery and development of novel materials based on the convergence of high-performance computing, automation, and ML[14]. Another aim of using such approaches in the field of materials science is to achieve high-throughput identification and quantification of essential material properties[15].

Besides experimentally obtained datasets, numerous studies draw required information from simulation-based data mining. Altogether, it is shown that experiment- and simulation-based data mining in combination with machine leaning tools

provide exceptional opportunities to enable highly reliant identification of fundamental interrelations within materials for characterization and optimization in a scale-bridging manner[15].

For more detailed information on recent ML applications in materials science we refer to the general reviews of Mueller et al.[2], Wagner et al.[3], Dimiduk et al.[4] or Wei et al.[16]. Examples for successful applications of ML techniques in materials science are, for example, to represent inorganic materials[17–19], predict fundamental properties[20–22], create atomic potentials[23], identify functional candidates[24–26], analyze complex reaction networks[27], or guide experimental design[28,29], high-throughput phase diagram and crystal structure determination[30].

Open problem - interpretability

However, one of the major criticisms of ML algorithms in science is the lack of novel understanding and knowledge arising from their use. This is mostly because more complex ML algorithms are often treated as black boxes. Those machine-built models are hard to understand for humans[31].

For a better acceptance of ML models, data scientists aim to establish clear causal relations between materials structure defined broadly across length scales and properties. Especially scientific models have further constraints such as a minimal number of parameters and adherence to physical laws. It is the obligation of the data scientist to translate the results of their work into knowledge other scientists can use in aiding for example materials discovery or deployment[3]. Useful techniques for finding simple, reduced and interpretable models are for example Principal Component Analysis (PCA)[32], Cross-validation and regularization and a thoughtful choice of model.

PCA is a powerful technique for data dimensionality reduction. Large datasets are increasingly widespread. In order to interpret such datasets, PCA can be applied to drastically reduce their dimensionality in an interpretable way, such that most of the information in the data is preserved[33]. PCA extracts the orthogonal directions with the greatest variance from a dataset, the resulting principal components being linear combinations of the original variables. However, principal components are not necessarily simple to interpret physically but as the extracted features are linear combinations of the original variables they can still be intuitively explained. Moreover, it allows a very straight forward data visualization through data projection onto the main extracted components[34]. However, PCA might be the wrong choice if features are not covariant.

Another way to achieve interpretable ML models is intelligent feature selection for dimension reduction and thus easier interpretability. Regularization of a model entails adding a tunable penalty on model parameter size to the cost function being minimized leading to a reduced feature space[3].

Furthermore, the choice of ML model has an immediate impact on its explainability. Regressions lead to coefficients whose size gives information about the relative size effect of modifying an input on the output. Decision trees (DT) are set up like flow charts and therefore easy to read. More complex models such as Artificial Neural Networks (ANN) are missing a clear explanation of the machine's "thinking" due to complex node interactions. But methods such as feature visualization[35,36] or attribution[37–39] exist, which allow a better understanding and interpretability of black box models. However, sometimes it might be reasonable to trade model accuracy for better explainability.

Open problem - small data

Contrary to the curse of dimensionality, there is often the problem of small data in material characterization, since the experimental or simulative generation of data is complex and expensive. Also the availability of the testing units limits data generation. Special test environments and material properties that are difficult to implement in the laboratory are often required, such as the creation of a corroded workpiece.

Often models are fitted to extremely small training sets which does not play to the strength of ML and will not allow the replication of the success ML methods had in other fields. It is of course possible to use ML methods as a simple fitting procedure for small low-dimensional datasets[31].

A few approaches are known to tackle this problem. For example, a ML model can be constructed by restricting the configurational space of materials, such as predicting the band gaps of selected families of semiconductors with fixed composition or crystalline structure instead of modeling compounds spanning a wide chemical space[40–42]. Another approach by Zhang et al.[40] proposes to incorporate the crude estimation of the property in the feature space to establish ML models using small sized materials data, which increases the accuracy of prediction without the cost of higher degree of freedom.

Another approach for insufficient training data is the additional integration of prior knowledge into the training process, which leads to the notion of informed ML[43], or more specifically physics-informed ML[44–46]. Domain knowledge is often given as a

set of additional constraints[47]. The integration of additional domain knowledge generates a new hybrid formulation of the ML problem which then ideally leads to physically meaningful and significantly more accurate interpretations of the data[48]. Besides adding constraints, expert knowledge can be incorporated in different ways. Up to recently, it was mostly limited to labeling data for supervised learning and setting prior probabilities in Bayesian networks[49]. However, in semi-supervised clustering applications, user guidance can be given by partial labeling information, which can be incorporated using hard constraints[50]. The domain knowledge can also be integrated into the process of building a ML model by the application of data visualization which often improves the accuracy of the resulting model[49]. Another approach is the monotonization of ML functions based on known physical relationships[51].

In general, ML systems are rarely viewed in the context of small data, where an insufficient data size for the training model compromises the learning success. The bottleneck of the database size especially limits applications, in which the construction of a database via experiments is time consuming and costly[7]. Thus, the recent development of materials databases might be helpful in tackling the small data problem.

Existing databases

With the launch of the Materials Genome Initiative (MGI)[52] in 2011 and the coming of the big data era, a large effort has been made in the materials science community to collect extensive datasets of materials properties and to provide materials engineers with ready access to the properties of known materials. Existing databases are the materials project[53,54], the inorganic crystal structure database[55,56], the Materials genome initiative[52,57] ,the NOMAD archive[5,58], the Topological Materials Database[59], Supercon[60] or the National Institute of Materials Science 2011 (NIMS)[61] with many databases of material properties of metal alloys, or National Institute of Standards and Technology (NIST)[62] with databases of properties for material classes such as structural ceramics, oxide glasses, superconductors[63–68]. A more comprehensive list of material databases can be found in Correa-Baena et al.[14].

Traditionally, negative results are often discarded and left unpublished. However, negative data are often just as important for ML algorithms as positive results in order to prevent bias in the data. In some disciplines with a longer tradition of data-based research (like chemistry), such databases already exist. In a similar vein, data that emerges as a side product but are not essential for a publication are often left unpublished producing so-called publication bias[69].This eventually results in a waste of resources because other researchers then have to repeat the work in order to produce a balanced dataset for ML applications[70].

However, few of these available databases are ready to use with informatics techniques because they lack the uniform data formats or application programming interfaces required for informatics software[68].

Materials informatics

The databases mentioned above contain information on numerous properties of known materials and are essential for the success of materials informatics[70]. For more general information, the reader is referred to[71], which gives an introducing review on materials informatics, that aptly describes the concept of big data in materials science. Another publication on materials informatics is, for example,[72] which introduces four main research areas in materials informatics: standardization of representation and exchange of material data; organization, management, retrieval, filtration and correlation of material data; material graphics; and data mining and knowledge discovery of material data. Another review on materials informatics mostly focuses on atomic-scale modeling[68]. But it also promotes the idea to expand materials databases to make more data easily accessible to informatics. The widespread use of such data requires the digitalization and structuring of materials data. The data must also be easily sharable and acessible. Services that provide software interfaces to allow for automated data querying, processing, and access are evolving, for example, the Materials Data Facility[73] and Citrination[74,75]. Wagner et al.[3] propose a workflow for a materials informatics problem focussing on (1) the assembly of primary features, (2) the construction of an exporatory model, (3) refinement of the model to satisfactory accuracy and (4) final training and deployment. By proceeding in an iterative fashion upwards in complexity, the final model will be as simple as possible improving its explainability and interpretability. The shorter, but more philosophical report of Rajan et al.[76] focuses on the role of materials informatics that allows one to survey complex, multiscale information in a high-throughput, statistically robust, and yet physically meaningful manner.

2.2 | Machine learning for metallic material characterization

Mechanical material properties are characteristics to be precisely predicted and controlled as they are strongly linked to and highly affected by process parameters and resulting microstructures[15]. The basic idea of using ML methods for material property prediction is to analyze and map the relationships (nonlinear in most cases) between the properties of a material and their related characteristics by extracting knowledge from existing experimental or simulated data[70]. Mechanical behavior in simulations is often described by means of constitutive equations[15].

Research on the macroscopic performance of materials mainly focuses on the structure-activity relationship between the macroscopic (e.g., mechanical and physical) properties of a material and its microstructure[70]. Many material parameters can be estimated to within an order of magnitude using elementary physical ideas. Whenever these parameters cannot be reliably estimated as such, ML approaches can be helpful which then require experimental or simulative data[77].

Experimental testing methods which can be used on metals help understand materials and their properties better. Typical destructive tests are bend test, impact test, hardness test, tensile test, fatigue test, corrosion resistance test or wear test, see for example[78]. In the following efforts of the materials community to enhance such tests and their results with ML methods will be discussed.

Corrosion

Corrosion detection and monitoring are essential diagnostic and prognostic means for preserving material "health" and reducing life-cycle cost of industrial infrastructures, ships, aircraft, ground vehicles, pipelines etc.[79]. More recently, ML approaches have shown great potential to improve corrosion detection[80]. They will aid a human inspector and significantly cut down on the time and cost associated with inspecting for example civil infrastructure and eliminate the need for dependence on prior knowledge and human effort in designing features[81]. Popular approaches are the application of ANNs for image processing-based corrosion detection.

Convolutional Neural Networks (CNN) were used for corrosion detection by investigation of images and identification of rusty parts in the image[81]. The proposed CNN outperforms state-of-the-art vision-based corrosion detection approaches that are developed based on texture and color analysis using a simple multilayered perceptron network. Model input is an image of the material and region of interest, output is the classification information: corroded/not corroded for a sliding window over the image. Overfitting due to small datasets can be avoided by using pretrained networks.

Another example for the application of CNNs for corrosion detection is proposed by Bastian et al.[82]. Similar to Atha et al.[81] images are used as input and classified into any of the four classes: no corrosion, low-level corrosion, medium-level corrosion and high-level corrosion. More investigations of similar CNN-based publications for corrosion detection are listed in[82].

Fang et al.[83] proposed a novel hybrid methodology combining genetic algorithms (GA) and support vector regression (SVR), which is capable of forecasting the atmospheric corrosion depth of metallic materials such as zinc and steel. This hybrid approach is capable of solving nonlinear regression estimation problems in materials science. The GAs are adopted to automatically determine the optimal hyper-parameters for SVR. The inputs for the SVR are temperature, time of wetness (TOW), exposure time, sulphur dioxide concentration and chloride concentration, respectively. The outputs are the predicted corrosion depth for zinc or steel.

Another support vector machine (SVM) approach was applied by Hoang et al.[84] for image processing-based detection of pipe corrosion. The image texture including statistical measurements of image colors, gray-level co-occurrence matrix, and gray-level run length is employed to extract features of the pipe surface. SVM optimized by differential flower pollination is then used to construct a decision boundary that can recognize corroded and intact pipe surfaces by blockwise classification of the original image. Also with the application of a SVR and based on a much smaller database (trained on only 46 samples), the corrosion rate of 3C steel in different environments was predicted based on five different seawater environment factors, including temperature, dissolved oxygen, salinity, pH-value and oxidation–reduction potential. The prediction error was very small[85].

Jimenez et al.[80] compare various ML approaches (ANNs, SVMs, classification tree (CT) and k-nearest neighbour (kNN)) for automatic pitting corrosion detection in 316L stainless steel. Model input are environmental variables such as chloride concentration, pH and temperature while the output is an information about the material being corroded or not. The models based on ANNs and SVM with linear kernel were demonstrated to be a valuable tool to be applied for this purpose. The classification performance for ANN and SVM is much better compared to kNN and CT models for this application. The principal advantage

compared with the traditional techniques is that it is not necessary to apply a surface analysis technique to study corrosion behaviour of the material.

Fatigue

Fatigue as the weakening of a material caused by cyclic loading that results in progressive and localised structural damage and the growth of cracks can also be predicted with the incorporation of ML methods. The prediction of fatigue in welded structures for a wide range of structural materials by multi-scale FEM and ML was proposed by Shiraiwa et al.[86]. Two ML algorithms are applied: one is deterministic ML based on the traditional methods, and the other is model-based ML. Deterministic ML such as Multivariate Linear Regression (MLR) and ANNs use chemical composition, processing parameters (reduction ratio, heat treatment), inclusion sizes, and fatigue strength as input features to accurately predict fatigue strength. In the model-based ML, microstructures and stress-strain curves in 40 low carbon steels with different chemical compositions and heat treatment conditions were prepared to create the learning dataset which was used to train an ANN. This approach allows for incorporation of prior knowledge of structure and property, and it can account for uncertainty such as scattering of fatigue life.

Another attempt to identify novel connections between fatigue properties and a variety of material parameters with the help of ML algorithms has been made by Agrawal et al.[87] for the prediction of fatigue strength of steel from composition and processing parameters such as chemical composition, upstream processing details, heat treatment conditions, mechanical properties. Various ML methods such as basic regression, DTs, SVM, ANN were used. Most success showed ensemble methods and individualized methods for different types of materials.

Machine learning techniques have also been utilized to predict material fatigue life for P91 steel base metal based on the hold time in fatigue tests[88]. A combined approach of GA and SVM is used to predict the fatigue life with high accuracy.

Abdalla et al.[89] use an ANN radial basis function model, taking the maximum tensile strain and pressure ratio as input, and put forward the model of fatigue life of steel reinforcing bars.

Creep

Creep is a type of metal deformation that occurs at stresses below the yield strength of a metal, generally at elevated temperatures. Creep rupture is becoming increasingly one of the most important problems affecting behavior and performance of power production systems operating in high temperature environments and potentially under irradiation as is the case of nuclear reactors. Creep rupture forecasting and estimation of the useful life is required to avoid unanticipated component failure and cost ineffective operation[90]. The material behaviour is influenced by the multidimensional interdependencies between the individual elements of the chemical composition, the heat treatment parameters, product form, tensile properties and microstructure, which are difficult to describe using simple analytical methods. Modeling with ML techniques therefore seems to be an interesting alternative. Moreover, the application of ML takes away the requirement for long and expensive experiments[91].

For the design of materials, creep is considered an important material property. However, quite often such designs only focus on one objective (e.g. creep) without considering the comprehensive design of multi-property[92]. For the investigation of creep rupture life and rupture strength of austenitic stainless steels[93] once again ANNs are popular models. For the prediction of creep rupture life and the creep rupture stress for a given stress, the training database for the input parameters contains test conditions (stress and temperature), chemical composition, solution treatment temperature and time (the latter being available in a very limited number of cases), nature of the quench following, grain size, and logarithm of ruptured life for a broad variety of stainless steels.

Chatzidakis et al.[90] employ and compare General Regression Neural Networks, ANNs and Gaussian Processes to capture the underlying trends and provide creep rupture forecasting. Input parameters are experimental creep rupture data. However, the overall performance of the developed models was insufficient.

In the study of Shin et al.[94], the five different ML models random forest (RF), linear regression (LR), kNN, kernel ridge (KR), Bayesian ridge (BR) are applied for the prediction of Larson-Miller parameters which represent the creep behavior. From a wide range of available features (466), relevant ones are selected with optimization approaches and different set-ups of features and models evaluated. Highest accuracy was obtained by RF for a varying number of top ranking features between 5 and 21.

For the prediction of rupture and creep rupture stress of 9%Cr steels, a Multilayer Perceptron Neural Network was applied with the input parameters chemical composition, heat treatment information, geometrical form of the investigated components[91].

Flow behaviour and work hardening

In sheet metal forming operations, mechanical properties of the sheet material such as flow stress or stress–strain curves greatly influence metal flow and product quality[95]. The flow stress can be determined by tensile tests which provide uniaxial information about stress–strain behavior. Hardness tests measure the resistance of material to an indenter, and hardness correlates well with flow stress. Compression testing reveals flow stress but is complicated by friction and buckling. Bend testing and torsion testing provide a good measure of flow stress and fracture resistance but are complicated by radial stress and strain gradients[96].

One of the main advantages of applying ML approaches is that it is not necessary to postulate a mathematical model at first, which is quite difficult because of the nonlinearities in the response of the deformation behaviors of the materials under elevated temperatures and strain rates and the factors affecting the flow stress. Lin et al.[97] propose a feed forward back propagation ANN model to predict the constitutive flow behaviors of 42CrMo steel during hot deformation, and investigate the general nature of the influence of strain, strain rate and temperature on the compressive deformation characteristics of 42CrMo steel. The capability of the developed ANN model to predict the flow stress level, the strain hardening and flow softening stages is also investigated. The inputs of the ANN are deformation temperature, log strain rate and strain whereas flow stress level, the strain hardening and flow softening stages are the output. Low absolute relative errors for training and test data promise a good generalizing model.

Gupta et al.[98] predict flow stress based on a feed forward ANN trained with the back propagation algorithm. Input features were strain, strain rate and temperature. The implemented ANN produced more accurate results than conventional mathematical models such as Johnson Cook, modified Zerilli-Armstrong and modified Arrhenius. SVR based flow stress prediction for austenitic stainless steel 304 with strain, strain rate and temperature as inputs and the flow stress as output was proposed by Desu et al.[99] and provides more accurate results than the conventional mathematical models.

The identification of work hardening properties of steel and an aluminum alloy from indentation tests is described by Meng et al.[100]. The authors propose a material parameter identification protocol based only on the imprint shape of the indentation test using orthogonal decomposition and manifold learning for the prediction of the strain hardening exponent and yield stress.

ANN modeling for anisotropic mechanical properties and work hardening behavior of Inconel 718 alloy at elevated temperatures was proposed by Mahalle et al.[101]. Strain and temperature are used to successfully predict material properties such as ultimate strength, yield strength, elongation and strain hardening coefficient.

Tensile properties

Tensile properties indicate how the material will react to forces being applied in tension. Determining the tensile properties is crucial because it provides information about the modulus of elasticity, elastic limit, elongation, proportional limit, reduction in area, tensile strength, yield point, yield strength, and other tensile properties[102] which then define the state of the material, its longevity or its ability to perform in an application. Thus, the accurate prediction of tensile properties has great importance for the service life assessment of structural materials[103].

The ultimate tensile strength (UTS) of iron castings which gives information about the capacity of a metal to resist deformation when subject to a certain load was predicted based on a variety of input features such as composition, size of casting, cooling speed or thermal treatment (25 variables) to gain information about the mechanical properties of a foundry and thus predict foundry defects[104]. Estimating the value of UTS is one of the hardest issues in foundry production, due to many different circumstances and variables that are involved in the casting process. Bayesian networks, kNN and ANNs were used for classification of the UTS. All of the investigated approaches perform well, but ANNs outperformed the other classifiers. Similar to the aforementioned study, Sterjovski et al.[105] also used ANNs (of back propagation type ANN) to predict mechanical properties of steel, such as the impact toughness of quenched and tempered pressure vessel steel exposed to multiple postweld heat treatment cycles; the hardness of the simulated heat affected zone in pipeline and tap fitting steels after in-service welding; and the hot ductility and UTS of various microalloyed steels over the temperature range for strand or slab straightening in the continuous casting process. Input parameters were composition, cooling rate, temperature and thickness. It was shown that ANNs could successfully predict all mechanical properties investigated. Another example for the application of ANNs is proposed by Sankar et al.[106] who predict elongation, self-tempering temperature and yield strength for reinforcement steel bars subjected to thermomechanical treatment based on two input parameters (bar diameter and quenching duration). The numerical results of the ANN are compared with experimental results and are found to be in good agreement.

Pruning and predator prey algorithms were applied by Datta et al.[107] which were able to extract more knowledge from the input data, than typically possible with conventional ANN analysis. Alloy composition and the thermo-mechanical controlled processing parameters, deformation in different temperature zones, finish rolling temperature and cooling rate of high strength

steels have been taken as input parameters, whereas UTS, yield strength and percentage elongation were predicted. It was shown that in this type of steel the yield strength depends mostly on the solid solution hardening and the microstructural constituents while UTS is more influenced by the precipitation hardening, but all these strengthening mechanisms have a negative effect on the ductility of the steel.

Pattanayak et al.[108] investigated the role of the composition and processing parameters on the mechanical properties of API grade microalloyed pipeline steel, in respect to its strength, impact toughness and ductility. ANN models, capable of prediction and diagnosis in non-linear and complex systems, are used to obtain the relationship of composition and processing parameters with said mechanical properties. Then the models are used as objective functions for the multi-objective GAs for evolving the tradeoffs between the conflicting objectives of achieving improved strength, ductility and impact toughness. The Pareto optimal solutions were analyzed successfully to study the role of various parameters for designing pipeline steel with such improved performance.

For the ANN-based prediction of yield strength, UTS, ductility (elongation, and reduction of area) of ferritic steel weld metals appropriate for the welding of high strength low alloy steels, Metzbower et al.[109] use the chemical composition of as deposited weld beads, and the cooling rate. The various established ANN models are found to work well once combined to an ensemble of best models. The ensemble produced a more reliable prediction than an individual model and reproduced known metallurgical trends well. Poudel et al.[110] also compare different ML models and ensembles for the prediction of UTS, yield strength, elongation of steel bars instead of performing tensile tests. Input parameters were process parameters from the rebar manufacturing process such as for example material composition, temperatures, rod diameter, rod speed and cooling rate. Investigated classification algorithms were among others MLR, Principal Components Regression, Partial Least Squares Regression, ANNs and Locally Weighted Regression. The main conclusion from this research is that any single model cannot efficiently describe the complicated relationship of the input-output space of the rebar manufacturing process. Ensemble methods, however, as combination of ML approaches used together in conjunction with some model selection or model weighting techniques lead to robust prediction systems.

Applications of traditional ML algorithms for the prediction of tensile properties were proposed by[92,103,111–113]. Shigemori et al.[111,112] reported about the successful application of Locally Weighted Regression (LWR) in predicting the tensile strength for a certain type of steel product which is produced by hot rolling. As input, 18 items were selected from chemical composition, heating, rolling, and cooling temperature. These variables have a clear physical causal relationship with the output variable. Furthermore, least squares SVMs are suitable approaches for the prediction of the elastic modulus and yield stress of materials. In[113] FEM-simulated load-indentation curves of Al6061 and Al7075 are investigated for the determination of these material parameters based on a training set of large strain-large deformation FEM for the simulation of indentation tests. Characteristic features are extracted from the load-indentation curves and used in the ML model. The proposed least squares SVM model is capable of predicting reasonably accurately the elastic modulus and yield stress of materials based on the load-indentation curves of dual conical indenters with different half-angles. For material design, a RF model in combination with an optimization algorithm was found to relate yield strength, impact toughness and total elongation with material composition information and treatment parameters for the production of RAFM steels[92,103]. For yield strength, highly correlated features were tempering temperature and C content and tempering time and Cr content for elongation. The accuracy and generalization ability of the RF was acceptable ($R^2 > 85\%$).

An interesting study different from most approaches using material composition information as input features for material parameter prediction was proposed by Fragassa et al.[114] for the prediction of the tensile behaviour of cast alloys such as yield strength, ultimate strength, ultimate strain and Young's modulus, by a pattern recognition analysis on experimental data and the application of RF, ANN, kNN. All Information is directly taken from micrographs. For the prediction of UTS and yield strength ANNs show the best results.

Another application of ANNs is the interpretation of acoustic emission data for failure prediction. Christopher et al.[115] propose the prediction of the ultimate strength of aluminum/silicon carbide (Al/SiC) composites by using acoustic emission parameters through ANN analysis. This approach was earlier pursued for the prediction of the ultimate strength of unidirectional T-300/914 tensile specimens using acoustic emission response and an ANN back propagation algorithm[116].

To prevent surface cracks on cast steel, its hot ductility must be monitored. Experimental investigations are difficult to execute. Thus, ML models are proposed for the prediction of hot ductility. For example, a MLR[117] predicted hot ductility and grouped 12 chemical elements that had similar experimental effects on ductility. The cooling condition used here is different from actual continuous-casting conditions, so the model is difficult to apply in commercial setups. A back-propagation ANN method has been used[105] to predict hot ductility for various microalloyed steels over the temperature range for strand or slab straightening

in the continuous casting process. However, the recorded data are limited and do not lead to a generalizing model. Additionally, the NN used only one hidden layer, so it has a limited ability to describe the complex relationship between input and output. Also an ANN model was used to predict high-temperature ductility of various steel grades from their composition and thermal history (described by five experimental variables)[118]. The developed model can predict ductility for a wider composition range and thermal history than previous studies have achieved. Therefore, it can be used effectively in commercial production.

2.3 | Machine learning for the small punch test

As mentioned above, the estimation of tensile properties of materials such as elastic modulus, yield strength or strain hardening exponent is considered to be of fundamental importance. Conventional tests are destructive in nature and require reasonable specimen cross-section and volume. There are situations where limited volume of material is available for property assessment, like in material development, failure analysis and remaining life assessment of in-service components, materials for pressure vessels, turbines, thermal power plants or chemical processing industries[119,120]. In these situations, test techniques using small volume specimens become more attractive. Small specimen test techniques have been established as a reliable alternative to the traditional tensile test, as the results of these test techniques are in good agreement with the tensile test and are reproducible when tested under controlled conditions[8].

The extraction of mechanical properties of in-service materials through such small specimen test techniques has become more popular in recent years. Of the available small specimen test techniques, SPT methods have proven to be promising[8,121–123]. These test methods are basically non-destructive in nature and are proficient enough to extract the flow properties of the materials using small volume specimen. After more than 40 years' development, there are some standards in SPTs, such as ASTM-F2183[124], GB/T-29459[125], and CEN CWA-15627[126].

A prerequisite for using this test is to establish correlations between SPT and conventional tests[127,128], such as relations between the tensile test and the Erichsen cupping test for the determination of the critical damage value curves, the initiation time and location of fracture[129]. Other examples are the correlation between the SPT and tensile tests for the estimation of the tensile strength[130] or the construction of stress–strain curves only from indentation tests[131]. Here the link between mean pressure and the indent diameter, obtained from indentation tests, to the stress and strain determined from a tensile test was obtained by the empirical relation of Tabor[132].

Exploiting force-displacement curves (FDC) from spherical indentation simplifies the identification of mechanical properties. Parameters such as hardness, deformation mode, yield stress, Young's modulus of the indented material can also be extracted out of the imprint shape, which is a valid alternative for using the indentation curve[133]. For example, Milivcka et al. found linear correlations between the SPT maximum force and the tensile strength of a 9% Chromium steel for creep resistant high temperature applications[134].

Detailed analyses of stress and strain in the SPT disc have been performed by means of analytical elastic-plastic modelling, see for example[135,136] and by FEM to underpin the empirical correlations[137–139]. FEM simulations are mainly used to generate FDCs based on an assumed material constitutive law which are then compared with experimental FDCs. This approach is called inverse analysis[140,141]. For more information, see for example[16,142]. However, the solution to the inverse identification problem is non-unique. Furthermore, there are other problems such as insufficient accuracy, indentation frame/machine compliance, noisy input data, difficulty in determining the exact starting point of the load (force)-displacement curve, bending of the specimen in thin-sheet indentation and FE mesh dependence[133].

Various mechanical properties that can be extracted through SPT, are for example fracture mode[143–145], yield stress[137,146,147], UTS[121,148,149] (based on hydraulic bulge tests which are similar to SPT but high-pressure hydraulic oil is used instead of punch to cause specimen deformation); tensile strength[134], Young's modulus[150–153], fracture toughness[138,140,154–158], creep properties[159,160] or elastic plastic properties[161,162].

These examples show that the SPT is a widely used approach for the determination of various material parameters from a small amount of material. However, the approach also has a few disadvantages, which have to be kept in mind. To begin with, the selected sample size may not represent the bulk material. The sample size effect has an important influence on the mechanism of the fracture of SPT, because often the thickness of the SPT sample is only 5–8 times of the mean grain size. The changes in grain-boundary distribution and the grain orientation, that resulted from the sample size effect, may affect the mechanical behaviors of the SPT samples. These parameters may affect the mechanism of fracture on small scale material samples[163] In addition, SPT response is sensitive to various test parameters such as specimen shape, specimen thickness, test speed, ball

diameter, clamping force and material. For more information, the reader is referred to[127], where an extensive collection of SPT configurations is listed.

In recent years, using ML approaches on SPT data has become popular. Most approaches incorporate ANNs for the identification of material properties from SPT data. Abendroth et al.[164] identify ductile damage and fracture parameters from the SPT using ANNs. FEM is used to compute the load displacement curves. Via a systematic variation of the material parameters a data base is built up, which is used to train the ANNs. This neural network can be used to predict the load displacement curve of the SPT for a given material parameter set. The identified material parameters are validated by independent tests on notched tensile specimens. A similar approach was followed by Abendroth et al.[138,165] earlier for the estimation of the hardening and damage parameters of high strength steels. The combination of FEM, SPT data and ANNs was used to identify the parameters of the Gurson–Tvergaard–Needleman model for ductile damage and fracture parameter prediction.

Linse et al.[140] use synthetic load displacement curves generated via FEM for a variety of material parameters. This database is then used to train ANNs, which approximate the load displacement curves of the SPT as a function of the material parameters. The identification procedure itself consists of an optimisation algorithm, minimising the difference between the measured load displacement curves and its approximation by the neural networks until the true set of material parameters is found. Prerequisite is an accurately working FEM-Model. The approach was applied for the identification of hardening parameters and WEIBULL-parameters in the brittle and brittle–ductile transition region of two non-irradiated reactor vessel steels.

Another employment of FEM simulations of tensile, bulge, Erichsen tests is proposed by Abbassi et al.[166] for the calculation of damage parameters. The Gurson–Tvergaard–Needleman (GTN) model was employed. An identification procedure based on ANNs is used to determine the material parameters of the GTN damage model. The ANN was trained by using the FEM results of the notched tensile test with varying the damage parameters. A good efficiency for the identification of damage parameters was proven. The Poisson ratio can also be predicted using a similar approach (combination of SPT and FEM and ANN)[167]. Using FE simulations, the relation between the material parameters and the quantities characterizing the depth-load response is calculated. An approximate inverse function represented by an ANN is derived on the basis of these data.

In contrast to all the ANN-based approaches described so far, the study of Meng et al.[133] suggests the use of manifold learning for indentation-based material characterization. However, FEM was also used there. Input parameters were features describing the imprint shape of SPTs. The manifold learning approach was able to iteratively reduce the distance between the FE-simulated and the experimental imprint shapes in order to identify the material hardening parameters. The approach was successfully shown for three different materials: AISI1095 steel and two aluminum alloys EN AW-2017F and EN AW-5754F.

Another innovative approach suggests the application of digital video processing of the forming process created by the Erichsen cup test[168]. The OpenCV library[169] was used to develop a sheet-metal-forming image analysis program which identifies the indent shape and position to calculate the anisotropic coefficient which showed good agreement with tensile test results. However, results were biased by the environment around the testing area.

A code of practice on SPT was established in 2010 in order to harmonize the various test set-ups and to achieve a better comparability of the results of different labs[170].

3 | AN APPLICATION OF ML IN MATERIAL PROPERTY PREDICTION

Similar to the studies mentioned above, we want to apply ML models to extract material properties from SPT data since the SPT has long been recognised as a supportive means for the development and monitoring of structural materials.

Conventionally, the following empirical correlation has been used for the estimation of the UTS R_m from FDCs[137,148,177–179]

$$R_m = \beta_{R_m} \cdot F_m/(h_0 \cdot v_m) \tag{1}$$

with F_m being the maximum force, v_m the corresponding punch displacement, h_0 the initial specimen thickness and β_{R_m} an empirical coefficient. However, various researchers[120,178] propose the determination of R_m based on the force F_i instead of F_m for the correlation with the UTS in order to avoid a strong dependence of the correlation factor β_{R_m} on the tensile properties of the material:

$$R_m = \beta_{R_m} \cdot F(v_i)/({h_0}^2) \tag{2}$$

Material	Type	Product form	Heat treatment	Reference
P91	f/m steel	Hot rolled pipe ⌀ 360 × 50 mm	Normalization 1040–1100°C/30min tempering 730–780°C/60min	[171]
P92	f/m steel	Hot rolled pipe ⌀ 219 × 22 mm	AR: standard normalization + tempering HT1: AR + 800°C/2h HT2: AR + 760°C/2h HT3: AR + 750°C/2h HT4: AR + 740°C/2h	[172]
Eurofer97	f/m steel	Hot rolled plate 14.5 mm	Normalization 980°C/27min/air cooling tempering 760°C/90min/air cooling	[173,174]
22NiMoCr 37	Bainitic steel	Reactor pressure vessel Biblis C	890°C/4 h/water quenching 650°C/7h/air cooling	[175]
15Kh2MFA	Bainitic steel	Reactor pressure vessel Greifswald unit 8	Original RPV manufacturing technology	[176]

TABLE 1 Materials overview.

Material	Number of tests	Test temperatures
P91	23	-177...+331
P92-AR	3	Room temperature
P92-HT1	5	Room temperature
P92-HT2	4	Room temperature
P92-HT3	3	Room temperature
P92-HT4	5	Room temperature
Eurofer97	8	-24...+250
22NiMoCr37	31	-151...+332
15Kh2MFA	33	-150...+332

TABLE 2 Overview of SPTs.

with β_{R_m} being geometry and material dependent. More information about this empirical approach can be found in[120]. In addition to the empirical approach (1),(2), we propose two ML models to determine the UTS R_m from SPT data, shown in Figure 1 (b).

3.1 | Experiments

SPT data of three ferritic-martensitic (f/m) Cr-steels and two bainitic reactor pressure vessel steels[120] were used to predict the UTS R_m based on ML approaches. The UTS R_m was obtained by tensile tests. A materials overview can be found in Table 1 , which was adopted from Altstadt et al.[120]. The P92 was available in the, as received condition, and in four different heat treatments. The specimens were of 0.5 mm thickness. SPT were conducted with a punch diameter of 2.5 mm, a receiving hole diameter of 4 mm, edge size 0.2 mm and the edge type chamfer. All tests were performed at a displacement rate of 0.5 mm/min. The displacement v was measured by an inductive sensor with an accuracy of $\pm$ 1 μm. The FDCs $F(v)$ of SPTs at room temperature are shown in Fig. 2 (a) (mean of all available curves per material at room temperature). The $F(v)$ curves of the material P91 for different temperatures are shown in Fig. 2 (b). The curve for $T = -177°C$ represents a brittle failure, the other test curves represent ductile failure. By using 9 different materials and heats, and by testing selected materials at different temperatures, a wide variation of tensile properties is obtained. Table 2 gives the number of conducted tests and their test temperatures for each type of material. Figure 1 shows the UTS R_m for all investigated materials and temperatures.

For more detailed information on the experimental set-up the reader is referred to Altstadt et al.[120].

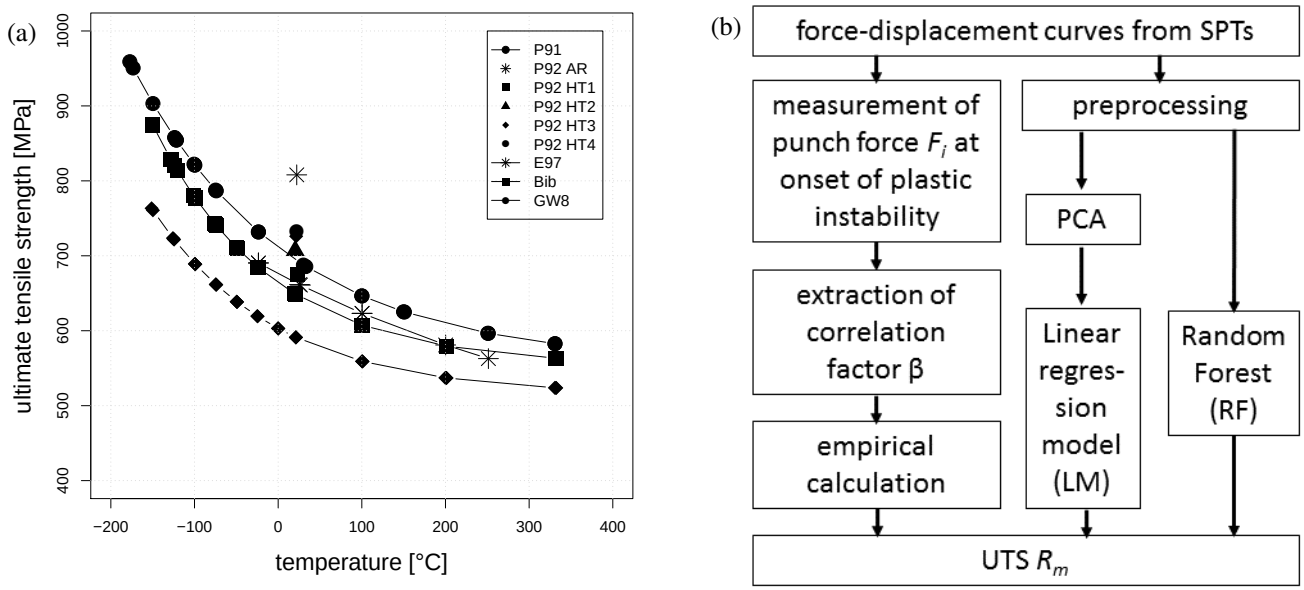

FIGURE 1 (a) UTS R_m for the investigated materials and temperatures. (b) scheme.

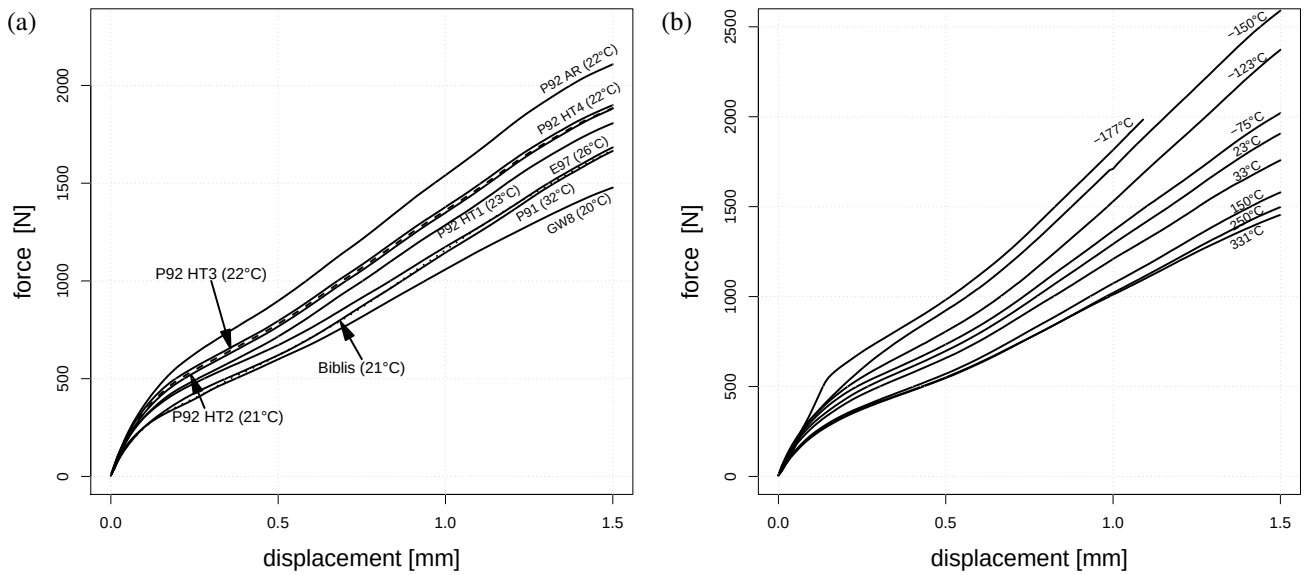

FIGURE 2 (a) SPT data at room temperature for materials listed in Table 1, (b) SPT data for P91 at different temperatures.

3.2 | Preprocessing

First, unified sampling points for the displacement were defined with a distance of 44 μm based on a sampling rate of roughly 1 μm in the raw data. Then each of the displacement values is defined as a feature, leading to 151 features for displacement values up to 1.5 mm. Failure occurred for displacements larger than 1.5 mm. An additional feature is the test temperature. Now, each time series is considered to be an observation. Dimensions correspond to the number of displacement sampling points + 1 (temp).

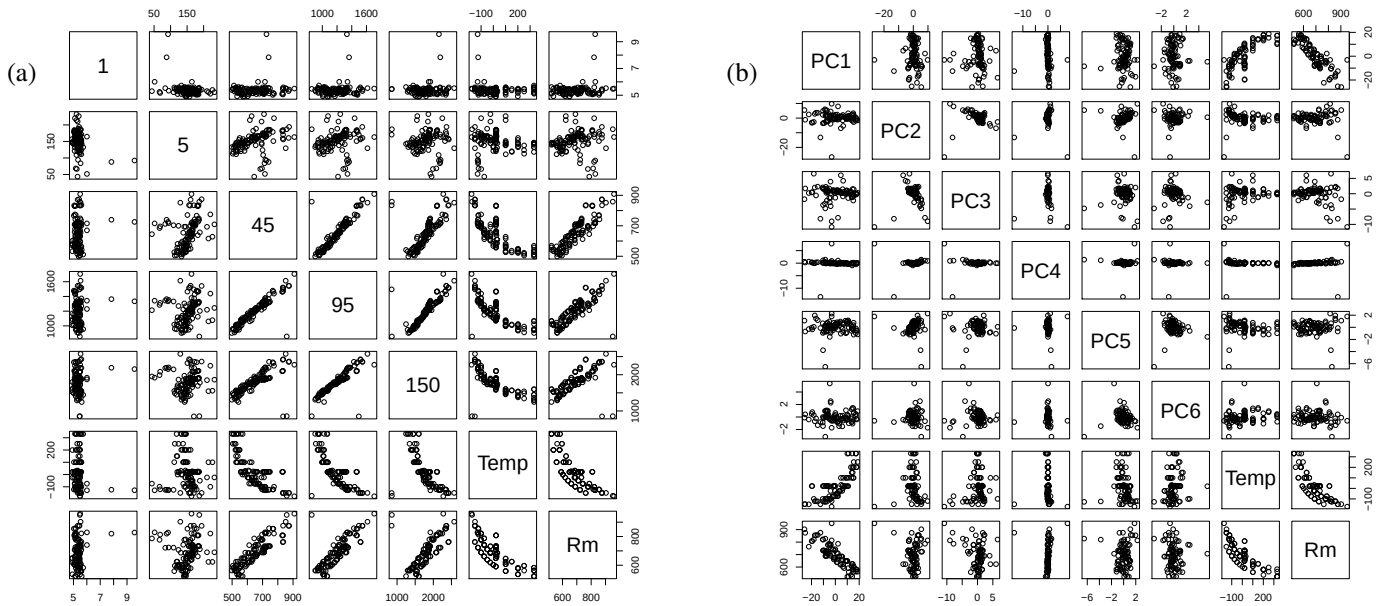

FIGURE 3 (a) Pairs plot of selected features, temperature and R_m, (b) pairs plot of first 5 principal components, temperature and R_m.

As the number of dimensions (p) increases, the volume of the domain increases exponentially. This, in turn, requires more samples (n) from the domain to provide effective coverage of the domain for a learning algorithm. This problem was introduced earlier as the curse of dimensionality[180]. ML algorithms overcome the curse of dimensionality by making assumptions about the data and structure of the mapping function from inputs to outputs which adds a bias and leads to a loss of the generalization power of the ML model[180]. Ways to approach this problem are feature selection, projection methods or the application of regularized algorithms.

Another problem when using FDC-data as input data for a ML algorithm is the multicollinearity of the features extracted from these curves. Exemplary correlations between features 1, 5, 45, 95, 150 and the temperature are shown in Figure 3 (a). Subsequently, principal component analysis (PCA) is performed on the FDC-data in order to eliminate co-linearity in the features. The broad idea behind this scheme is that, in contrast to the original features, the principal components will be uncorrelated. Furthermore, one expects that a small number of principal components will explain most of the variance and therefore provide an accurate representation of the dataset[31]. For the present dataset a total of 6 principal components explain 99% of the overall variability. Now, the pairs plot of the new features in Figure 3 (b) shows only horizontal or vertical lines between input features (principal components) which is typical for uncorrelated features.

3.3 | Machine learning models

For the estimation of tensile properties such as the UTS R_m from SPT data, empirical equations such as (1) or (2) can be used. The force F_i can be associated with the onset of plastic instability and is therefore well suited for a correlation with the UTS R_m of the uniaxial tensile test as proposed in[120]. With the number of available data points being rather small a linear regression model (LM) in combination with PCA and a RF based on the original SPT data are proposed for the prediction of R_m. The root mean squared error (RMSE) was used to compare the different approaches.

For the evaluation of the performance of the ML approaches 10-fold cross validation was conducted with the results shown in Table 3 . Figures 4 (a-c) show the accuracy of the empirical approach, LM and RF, respectively.

	empirical	LM	RF
RMSE	44.1	41.2	40.4

TABLE 3 RMSE comparison.

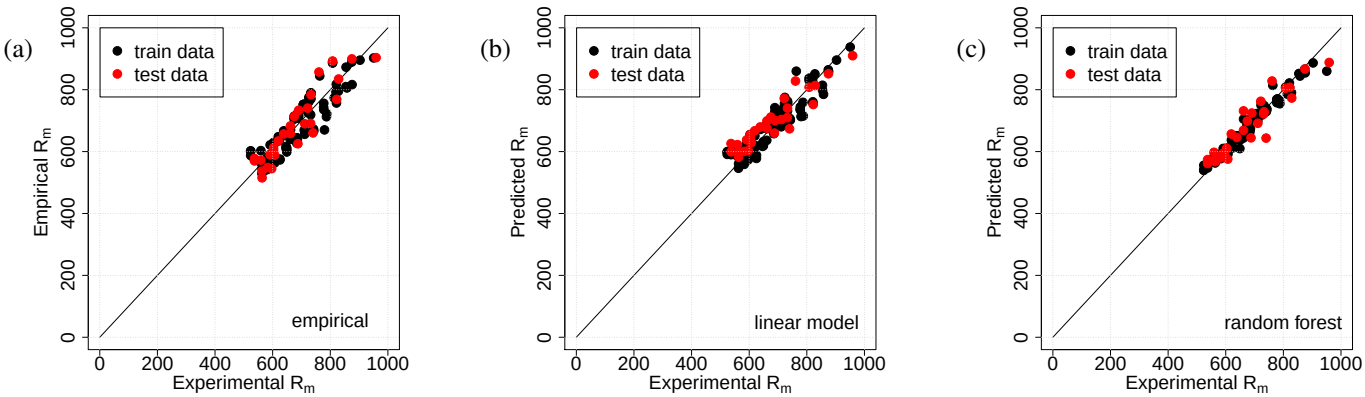

FIGURE 4 (a) Empirical R_m vs. experimental R_m, (b) LM based R_m vs. experimental R_m, (c) RF based R_m vs. experimental R_m.

3.4 | Application conclusions

The application presented here aims at the determination of the tensile property R_m from SPT data. Altstadt et. al[120] proposed an empirical equation which correlates the punch force at onset of plastic instability of the SPT with R_m. The correlation factor β for the estimation of R_m needs to be determined before the empirical equation can be applied and β is dependent on the sample geometry and the investigated material. The application of ML models on SPT data works well for the prediction of R_m, as Table 3 shows. However, the amount of available data is very limited and the generalization power of the LM and RF for other materials, temperatures, geometries has to be critically evaluated. RFs usually do not generalize well. Thus, the application of RF on SPT data might not be sensible for $R_m < 500$ MPa or $R_m > 1000$ MPa. More experimental data will have to be collected to verify the RF model further. However, the interpretation of the RF based on SPT data is straightforward and leads to relevant information about the importance of features for the performance of the model.

The LM approach is quite similar to the empirical approach of Altstadt et al. However, not only one specific force value of the SPT is used to correlate to R_m but the combination of several principal components generated from the SPT curves are employed which might justify the slightly higher accuracy of the LM approach. Thus, LM is a great tool to analyze the relationships among the variables but in combination with PCA the interpretability of the resulting model is not intuitive.

Overall, correlations between SPT curves and R_m clearly exist and SPT data can be used for the prediction of tensile properties such as R_m for five different steels (nine heats). Once the model is established, there is no further need for the conduction of time consuming tensile tests. In the future, the models might have to be re-established on a broader database. For the determination of R_m, FEM-simulations might not be necessary anymore. However, for more detailed information about the underlying constitutive laws via inverse analysis, FEM simulations will likely remain important an inevitable.

4 | DISCUSSION AND CONCLUSIONS

This survey addresses applications of machine learning strategies in materials science for material characterization. There exists a wide range of promising applications for ML in materials science, e.g. material discovery, molecular dynamics and global structural prediction. The demand for new approaches dealing with limited data is huge.

It was shown that data-driven approaches play a significant role in materials research in order to find relationships between the structure of a material and its properties. These relationships are often not linear. It is difficult to find generic patterns among multiple length scales and timescales. With experiments only, this cannot be achieved[76]. Therefore, data-mining techniques are

indispensable for the recognition of correlations in the (experimental and simulated) data. As the amount of publicly available materials data grows, ML techniques in particular will be able to extract from these data sets scientific principles and design rules that could not be determined through conventional analysis[181].

The majority of early ML applications to materials science employed straightforward and simple-to-use algorithms, like linear kernel models and DTs. Now, these proofs-of-concept exist for a variety of application even though there is a lack of benchmarking datasets and standards[31]. To date, ML often cannot realize the expected accuracy when applied to some tasks due to insufficient material data. Therefore, a more accurate model that was trained on a small but accurate data set is only meaningful within the input data space but does not generalize well while a less accurate model on a wide input data space is better at generalization but less exact. Therefore, accelerating the construction of publicly accessible material databases is highly important for the future development of ML in materials science[16]. Another issue that holds back the development of precise ML models is the absence of failure data. In this case, a cultural shift toward the publication of all valid data, may it be positive or negative, is required[16,31].

The majority of ML approaches in materials science is based on ANNs. However, conventional ANNs still suffer from several weaknesses such as the need for a large number of controlling parameters, the difficulty in obtaining stable solutions, the danger of overfitting and thus the lack of generalization capability[83]. However, ANNs have been enormously successful in understanding complex materials behavior, such as mechanical behavior (flow stress, hardness, tensile strength, fracture strength, and fatigue behavior) of metal alloys subjected to certain heat treatment and/or deformation procedures, as well as in the prediction of micro-structures and phases resulting from heat treatment and/or deformation processes. The most practical way to capture the complex dependence of a desired macroscopic property on the various process parameters is through such learning methods[2]. ANNs have the potential to minimize the need for expensive experimental investigation and/or inspection of structural materials used in various applications, hence resulting in large economic benefits for organisations[105].

In addition to ANN hybrid ML models or ensemble methods work well[83]. For this, multiple independent models are built and the final regression or classification result is usually obtained as an average over the ensemble. In this way, additional noise is introduced into the fitting process and overfitting is avoided[31]. However, there does not exist an overall solution that can be considered the best. The most appropriate model has always to be found specifically for the application and data situation.

Furthermore, this paper focuses on ML based material property prediction from SPT data. Such simple material tests have gained popularity over the last couple of years because even though they are cheap and simple to perform, they make accurate material characterization possible, especially for failure analysis and remaining life assessment of in-service components or structural parts. Nevertheless, a few disadvantages have to be taken into consideration. The small sample size might not represent the bulk material; the sample size effect influences the material properties; and the results of the SPT are sensitive to test parameters. However, also for SPT data, ML based model are popular for material parameter prediction. Most commonly found are ANNs, especially in combination with FEM for data generation. No application of traditional ML models to SPT data was found in the literature. The paper concludes with an application example which uses FDCs of structural materials as the basis for predicting the UTS. Simple ML approaches presented here, such as linear regression models or RFs provide good results for predicting the UTS based on SPT data, even for a very small database.

As a consequence, it is possible to confirm the benefit of simple ML techniques in predicting mechanical properties such as the UTS based on as simple material tests as the SPT and the authors are sure that ML will positively shape materials science for the years to come.

ACKNOWLEDGMENTS

This work is part of the Fraunhofer Lighthouse Project ML4P (Machine Learning for Production). The work profited from BiGmax, the Max Planck Society's Research Network on Big-Data-Driven Materials-Science. The data from Section 3 was kindly provided by Dr. Altstadt of the Helmholtz-Zentrum Dresden-Rossendorf, Germany.

Financial disclosure

None reported.

Conflict of interest

The authors declare no potential conflict of interests.

References

[1] A. Agrawal, A. Choudhary, *Apl Materials* **2016**, *4* (5), 053208.

[2] T. Mueller, A. G. Kusne, R. Ramprasad, *Rev Comput Chem* **2016**, *29*, 186–273.

[3] N. Wagner, J. M. Rondinelli, *Front Mater* **2016**, *3*, 28.

[4] D. M. Dimiduk, E. A. Holm, S. R. Niezgoda, *Integr Mater Manuf Innov* **2018**, *7* (3), 157–172.

[5] C. Draxl, M. Scheffler, *MRS Bull* **2018**, *43* (9), 676–682.

[6] M. Verleysen, D. François, in *International Work-Conference on Artificial Neural Networks*, Springer, **2005**, pp. 758–770.

[7] C. Shen, C. Wang, X. Wei, Y. Li, S. van der Zwaag, W. Xu, *Acta Mater* **2019**, *179*, 201–214.

[8] S. Arunkumar, Raghu V. Prakash, *Trans Indian Inst Met* **2015**, *69* (6), 1245–1256.

[9] K. T. Butler, D. W. Davies, H. Cartwright, O. Isayev, A. Walsh, *Nature* **2018**, *559* (7715), 547–555.

[10] M.I. Jordan, T.M. Mitchell, *Science* **2015**, *349* (6245), 255–260.

[11] J. Fan, F. Han, H. Liu, *Natl Sci Rev* **2014**, *1* (2), 293–314.

[12] National Academy of Sciences (US). Committee on the Survey of Materials Science, National Academy of Sciences (US). Committee on Science, Public Policy, *Materials and Man's Needs: Materials Science and Engineering; Supplementary Report*, *Vol. 1*, National Academies, **1974**.

[13] G.B. Olson, *Science* **1997**, *277* (5330), 1237–1242.

[14] J. Correa-Baena, K. Hippalgaonkar, J. van Duren, S. Jaffer, V. R. Chandrasekhar, V. Stevanovic, C.s Wadia, S. Guha, T. Buonassisi, *Joule* **2018**, *2* (8), 1410–1420.

[15] F. E. Bock, R. C. Aydin, C. J. Cyron, N. Huber, S. R. Kalidindi, B. Klusemann, *Front Mater* **2019**, *6*, 110.

[16] J. Wei, X. Chu, X. Y. Sun, K. Xu, H. X. Deng, J. Chen, Z. Wei, M. Lei, *InfoMat* **2019**, *1* (3), 338–358.

[17] A. Seko, H. Hayashi, K. Nakayama, A. Takahashi, I. Tanaka, *Phys Rev B* **2017**, *95* (14), 144110.

[18] K.T. Schütt, H. Glawe, F. Brockherde, A. Sanna, K.R. Müller, E.K.U. Gross, *Phys Rev B* **2014**, *89* (20), 205118.

[19] O. Isayev, D. Fourches, E. N. Muratov, C. Oses, K. Rasch, A. Tropsha, S. Curtarolo, *Chem Mater* **2015**, *27* (3), 735–743.

[20] B. Medasani, A. Gamst, H. Ding, W. Chen, K. A. Persson, M. Asta, A. Canning, M. Haranczyk, *npj Comput Mater* **2016**, *2* (1), 1–10.

[21] M. de Jong, W. Chen, R. Notestine, K. Persson, G. Ceder, A. Jain, M. Asta, A. Gamst, *Sci Rep* **2016**, *6* (1), 34256.

[22] F. Legrain, J. Carrete, A. van Roekeghem, S. Curtarolo, N. Mingo, *Chem Mater* **2017**, *29* (15), 6220–6227.

[23] Z. Li, J. R. Kermode, A. De Vita, *Phys Rev Lett* **2015**, *114* (9), 096405.

[24] Z. Li, S. Wang, W. S. Chin, L. E. Achenie, H. Xin, *J Mater Chem A* **2017**, *5* (46), 24131–24138.

[25] X. Ma, Z. Li, L. E.K. Achenie, H. Xin, *J Phys Chem Lett* **2015**, *6* (18), 3528–3533.

[26] A. Mannodi-Kanakkithodi, T. D. Huan, R. Ramprasad, *Chem Mater* **2017**, *29* (21), 9001–9010.

[27] Z. W. Ulissi, A. J. Medford, T. Bligaard, J. K. Nørskov, *Nat Commun* **2017**, *8* (1), 1–7.

[28] P. Raccuglia, K. C. Elbert, P. D.F. Adler, C. Falk, M. B. Wenny, A. Mollo, M. Zeller, S. A. Friedler, J. Schrier, A. J. Norquist, *Nature* **2016**, *533* (7601), 73–76.

[29] E. Kim, K. Huang, S. Jegelka, E. Olivetti, *npj Comput Mater* **2017**, *3* (1), 1–9.

[30] J. Graser, S. K. Kauwe, T. D. Sparks, *Chem Mater* **2018**, *30* (11), 3601–3612.

[31] J. Schmidt, M. R.G. Marques, S. Botti, M. A.L. Marques, *npj Comput Mater* **2019**, *5* (1), 1–36.

[32] S. Wold, K. Esbensen, P. Geladi, *Chemom Intell Lab Syst* **1987**, *2* (1-3), 37–52.

[33] I. T. Jolliffe, J. Cadima, *Phil Trans R Soc A* **2016**, *374* (2065), 20150202.

[34] A. Vellido, J. D. Martín-Guerrero, P. J.G. Lisboa, in *ESANN*, Citeseer, **2012**, pp. 163–172.

[35] D. Erhan, Y. Bengio, A. Courville, P. Vincent, *University of Montreal* **2009**, *1341* (3), 1.

[36] K. Simonyan, A. Vedaldi, A. Zisserman, *arXiv preprint arXiv:1312.6034* **2013**.

[37] R.C. Fong, A. Vedaldi, in *Proceedings of the IEEE International Conference on Computer Vision*, **2017**, pp. 3429–3437.

[38] M.D. Zeiler, R. Fergus, in *European Conference on Computer Vision*, Springer, **2014**, pp. 818–833.

[39] M. Sundararajan, A. Taly, Q. Yan, *arXiv preprint arXiv:1703.01365* **2017**.

[40] Y. Zhang, C. Ling, *npj Comput Mater* **2018**, *4* (1), 1–8.

[41] P. Dey, J. Bible, S. Datta, S. Broderick, J. Jasinski, M. Sunkara, M. Menon, K. Rajan, *Comp Mater Sci* **2014**, *83*, 185–195.

[42] G. Pilania, A. Mannodi-Kanakkithodi, B.P. Uberuaga, R. Ramprasad, J.E. Gubernatis, T. Lookman, *Sci Rep* **2016**, *6* (1), 19375.

[43] L. von Rueden, S. Mayer, K. Beckh, B. Georgiev, S. Giesselbach, R. Heese, B. Kirsch, J. Pfrommer, A. Pick, R. Ramamurthy, et al., *arXiv preprint arXiv:1903.12394* **2020**.

[44] M. Raissi, P. Perdikaris, G. Karniadakis, *arXiv preprint arXiv:1711.10561* **2017**.

[45] Y. Zhu, N. Zabaras, P.S. Koutsourelakis, P. Perdikaris, *J Comp Phy* **2019**, *394*, 56–81.

[46] E. Zhang, M. Yin, G.E. Karniadakis, *arXiv preprint arXiv:2009.04525* **2020**.

[47] J. Struyf, S. Džeroski, in *European Conference on Machine Learning*, Springer, **2007**, pp. 359–370.

[48] S. Ermon, R.L. Bras, S.K. Suram, J.M. Gregoire, C. Gomes, B. Selman, R.B. Van Dover, *arXiv preprint arXiv:1411.7441* **2014**.

[49] J. MacInnes, S. Santosa, W. Wright, *IEEE Comput Grap Appl* **2010**, *30* (1), 8–14.

[50] M. W. Chang, L. Ratinov, D. Roth, in *Proceedings of the 45th Annual Meeting of the Association of Computational Linguistics*, **2007**, pp. 280–287.

[51] H. Dette, N. Neumeyer, K.F. Pilz, *Bernoulli* **2006**, *12* (3), 469–490.

[52] The Materials Genome Initiative, https://www.mgi.gov/, Accessed: 2020-07-21.

[53] Materials Project, https://materialsproject.org/, Accessed: 2020-07-21.

[54] A. Jain, S. P. Ong, G. Hautier, W. Chen, W. D. Richards, S. Dacek, S. Cholia, D. Gunter, D. Skinner, G. Ceder, K. A. Persson, *APL Mater* **2013**, *1* (1), 011002.

[55] S. Rühl, *The Inorganic Crystal Structure Database (ICSD): A Tool for Materials Sciences*, https://icsd.products.fiz-karlsruhe.de/, **2019**, Accessed: 2020-07-21.

[56] G. Gergerhoff F. H. Allen, R. Sievers, *Crystallographic databases: Information content, software systems, scientific applications*, International Union of Crystallography. Data Commission, **1987**.

[57] A. White, *MRS Bull* **2012**, *37* (8), 715–716.

[58] Nomad Repository, https://nomad-coe.eu/, Accessed: 2020-07-21.

[59] Topological Materials Database, https://www.topologicalquantumchemistry.com/, Accessed: 2020-09-21.

[60] Materials Information Station National Institute of Materials Science, *SuperCon*, https://supercon.nims.go.jp/en/, Accessed: 2020-07-21.

[61] National Institute of Materials Science, https://www.nims.go.jp/eng/, Accessed: 2020-07-21.

[62] NIST Materials Data Repository, https://materialsdata.nist.gov/, **2020**, Accessed: 2020-07-21. https://doi.org/10.1002/9783527809080.cataz11525.

[63] J. E. Saal, S. Kirklin, M. Aykol, B. Meredig, C. Wolverton, *JOM* **2013**, *65* (11), 1501–1509.

[64] S. Kirklin, J. E. Saal, B. Meredig, A. Thompson, J. W. Doak, M. Aykol, S. Rühl, C. Wolverton, *npj Comput Mater* **2015**, *1* (1), 1–15.

[65] C. R. Groom, I. J. Bruno, M. P. Lightfoot, S. C. Ward, *Acta Crystallogr Sect B* **2016**, *72* (2), 171–179.

[66] B. Puchala, G. Tarcea, E. A. Marquis, M. Hedstrom, H.V. Jagadish, J. E. Allison, *JOM* **2016**, *68* (8), 2035–2044.

[67] A. Zakutayev, N. Wunder, M. Schwarting, J. D. Perkins, R. White, K. Munch, W. Tumas, C. Phillips, *Sci Data* **2018**, *5* (1), 180053.

[68] L. Ward, M. Aykol, B. Blaiszik, I. Foster, B. Meredig, J. Saal, S. Suram, *MRS Bull* **2018**, *43* (9), 683–689.

[69] A. Mlinarić, M. Horvat, V. Šupak Smolčić, *Biochem Medica* **2017**, *27* (3), 447–452.

[70] Y. Liu, T. Zhao, W. Ju, S. Shi, *J Materiomics* **2017**, *3* (3), 159–177.

[71] Q. Song, *Chinese Sci Bull* **2004**, *49* (2), 210.

[72] Q. Wei, X. Peng, X. Liu, W. Xie, *Chinese Sci Bull* **2006**, *51* (4), 498–504.

[73] B. Blaiszik, K. Chard, J. Pruyne, R. Ananthakrishnan, S. Tuecke, I. Foster, *JOM* **2016**, *68* (8), 2045–2052.

[74] J. O'Mara, B. Meredig, K. Michel, *JOM* **2016**, *68* (8), 2031–2034.

[75] M. W. Gaultois, A. O. Oliynyk, A. Mar, T. D. Sparks, G. J. Mulholland, B. Meredig, *APL Mater* **2016**, *4* (5), 053213.

[76] K. Rajan, *Mater Today* **2005**, *8* (10), 38–45.

[77] A. Lucas, *arXiv preprint arXiv:1402.2593* **2014**.

[78] M. A. Meyers, K. K. Chawla, *Mechanical Behavior of Materials*, Cambridge University Press, **2008**.

[79] V. S. Agarwala, P. L. Reed, S. Ahmad, et al., in *CORROSION 2000*, NACE International, **2000**.

[80] M.J. Jiménez–Come, I.J. Turias, F.J. Trujillo, *Mater Design* **2014**, *56*, 642–648.

[81] D. J. Atha, M. R. Jahanshahi, *Struct Health Monit* **2017**, *17* (5), 1110–1128.

[82] B. T. Bastian, Jaspreeth N., S. K. Ranjith, C.V. Jiji, *NDT&E Int* **2019**, *107*, 102134.

[83] S.F. Fang, M.P. Wang, W.H. Qi, F. Zheng, *Comp Mater Sci* **2008**, *44* (2), 647–655.

[84] N. D. Hoang, V. D. Tran, *Comput Intel Neurosc* **2019**, *2019*, 1–13.

[85] Y.F. Wen, C.Z. Cai, X.H. Liu, J.F. Pei, X.J. Zhu, T.T. Xiao, *Corros Sci* **2009**, *51* (2), 349–355.

[86] T. Shiraiwa, F. Briffod, Y. Miyazawa, M. Enoki, in *Proceedings of the 4th World Congress on Integrated Computational Materials Engineering (ICME 2017)*, Springer, **2017**, pp. 317–326.

[87] A. Agrawal, P. D. Deshpande, A. Cecen, G. P. Basavarsu, A. N. Choudhary, S. R. Kalidindi, *Integr Mater Manuf Innov* **2014**, *3* (1), 90–108.

[88] L. Zhang, J. Lei, Q. Zhou, Y. Wang, *Adv in Nat Sci* **2015**, *8* (1), 21–26.

[89] J. A. Abdalla, R. Hawileh, *J Franklin Inst* **2011**, *348* (7), 1393–1403.

[90] S. Chatzidakis, M. Alamaniotis, L. H. Tsoukalas, *IJMSTR* **2014**, *2* (2), 1–25.

[91] O. Frolova, E. Roos, K. Maile, W. Müller, *Trans MLDM* **2011**, *4* (1), 1–16.

[92] C. Wang, C. Shen, X. Huo, C. Zhang, W. Xu, *Nucl Eng Technol* **2020**, *52* (5), 1008–1012.

[93] T. Sourmail, H.K.D.H. Bhadeshia, D.J.C. MacKay, *Mater Sci Technol* **2002**, *18* (6), 655–663.

[94] D. Shin, Y. Yamamoto, M. P. Brady, S. Lee, J. Haynes, *SSRN Journal* **2018**, *168*, 321–330.

[95] G. Gutscher, H. C. Wu, G. Ngaile, T. Altan, *J Mater Process Technol* **2004**, *146* (1), 1–7.

[96] R. N. Wright, *Wire Technology*, Elsevier, **2016**.

[97] Y.C. Lin, Jun Zhang, Jue Zhong, *Comp Mater Sci* **2008**, *43* (4), 752–758.

[98] A. K. Gupta, H. N. Krishnamurthy, Y. Singh, K. M. Prasad, S. K. Singh, *Mater Design* **2013**, *45*, 616–627.

[99] R. K. Desu, S. C. Guntuku, B. Aditya, A. K. Gupta, *Procedia Mater Sci* **2014**, *6*, 368–375.

[100] L. Meng, P. Breitkopf, B. Raghavan, G. Mauvoisin, O. Bartier, X. Hernot, *Comput Methods Appl Mech Eng* **2015**, *297*, 239–257.

[101] G. Mahalle, O. Salunke, N. Kotkunde, A. K. Gupta, S. K. Singh, *J Mater Res Technol* **2019**, *8* (2), 2130–2140.

[102] R. Rahman, S. Zhafer Firdaus Syed Putra in *Mechanical and Physical Testing of Biocomposites, Fibre-Reinforced Composites and Hybrid Composites*, Elsevier, **2019**, pp. 81–102.

[103] C. Wang, C. Shen, Q. Cui, C. Zhang, W. Xu, *J Nucl Mater* **2020**, *529*, 151823.

[104] I. Santos, J. Nieves, Y. K. Penya, P. G. Bringas, in *2009 ICCAS-SICE*, IEEE, **2009**, pp. 4536–4541.

[105] Z. Sterjovski, D. Nolan, K.R. Carpenter, D.P. Dunne, J. Norrish, *J Mater Process Technol* **2005**, *170* (3), 536–544.

[106] I.Z. Sankar, K.M. Rao, B.V.R. Murhty, *IE (I) Journal-MM* **2009**, *90*, 3–6.

[107] S. Datta, F. Pettersson, S. Ganguly, H. Saxén, N. Chakraborti, *ISIJ Int* **2007**, *47* (8), 1195–1203.

[108] S. Pattanayak, S. Dey, S. Chatterjee, S. G. Chowdhury, S. Datta, *Comp Mater Sci* **2015**, *104*, 60–68.

[109] E.A. Metzbower, J.J. deLoach, S.H. Lalam, H.K.D.H. Bhadeshia, *Sci Technol Weld Joining* **2001**, *6* (2), 116–124.

[110] R. C. Poudel, T. Sakaguchi, Y. Shimizu, *J Chem Eng Japan* **2013**, *46* (4), 294–301.

[111] H. Shigemori, S. Kawamura, in *SICE Annual Conference 2007*, IEEE, IEEE, **2007**, pp. 810–815. https://doi.org/10.1109/sice.2007.4421095.

[112] H. Shigemori, M. Kano, S. Hasebe, *J Process Control* **2011**, *21* (2), 293–301.

[113] S. Swaddiwudhipong, K. K. Tho, Z. S. Liu, J. Hua, N. S. B. Ooi, *Modelling Simul Mater Sci Eng* **2005**, *13* (6), 993–1004.

[114] C. Fragassa, M. Babic, C. P. Bergmann, G. Minak, *Metals* **2019**, *9* (5), 557.

[115] L. Christopher, T. Sasikumar, C. Santulli, C. Fragassa, *FME Transaction* **2018**, *46* (3), 253–258.

[116] T. Sasikumar, S. Rajendraboopathy, K.M. Usha, E.S. Vasudev, *J Nondestruct Eval* **2008**, *27* (4), 127–133.

[117] Q. Liu, X. Zhang, B. Wang, B. Wang, *Sci Technol Cast Proc* **2012**, *26*.

[118] S. H. Kwon, D. G. Hong, C. H. Yim, *Ironmak Steelmak* **2019**, 1–12.

[119] A. Husain, R. Sharma, D.K. Sehgal, *Proc Eng* **2017**, *173*, 710–717.

[120] E. Altstadt, M. Houska, I. Simonovski, M. Bruchhausen, S. Holmström, R. Lacalle, *Int J Mech Sci* **2018**, *136*, 85–93.

[121] H. Wang, T. Xu, B. Shou, *Materials* **2016**, *10* (1), 23.

[122] M.P. Manahan, A.S. Argon, O.K. Harling, *J Nucl Mater* **1981**, *104*, 1545–1550.

[123] M. P. Manahan, Ph.D. thesis, Massachusetts Institute of Technology, **1982**.

[124] Standard Test Method for Small Punch Testing of Ultra-High Molecular Weight Polyethylene Used in Surgical Implants, *ASTM F2183-02(2008)*, https://www.astm.org/Standards/F2183.htm, Accessed: 2020-07-21.

[125] Small punch test methods of metallic materials for in-service pressure equipment, *GB/T 29459*, https://www.chinesestandard.net/PDF.aspx/GBT29459.2-2012, Accessed: 2020-07-21.

[126] CEN CWA-15627, *15627 Worskshop Agreement: Small punch test method for metallic materials*, **2006**.

[127] S. Arunkumar, *Met Mater Int* **2019**, *26* (6), 719–738.

[128] A. Janča, J. Siegl, P. Haušild, *J Nucl Mater* **2016**, *481*, 201–213.

[129] S. W. Kim, Y. S. Lee, *Metall and Materi Trans B* **2013**, *45* (2), 445–453.

[130] I. Klevtsov, A. Dedov, A. Molodtsov, in *6th International DAAAM Baltic Conference Industrial Engineering*, **2008**.

[131] M. Yetna N'Jock, D. Chicot, X. Decoopman, J. Lesage, J.M. Ndjaka, A. Pertuz, *Int J Mech Sci* **2013**, *75*, 257–264.

[132] D. Tabor, *J Inst Metals* **1951**, *79*, 1.

[133] L. Meng, B. Raghavan, O. Bartier, X. Hernot, G. Mauvoisin, P. Breitkopf, *Mech Mater* **2017**, *107*, 31–44.

[134] K. Milička, F. Dobeš, *Int J Press Vessels Pip* **2006**, *83* (9), 625–634.

[135] J. Chakrabarty, *Int J Mech Sci* **1970**, *12* (4), 315–325.

[136] T.S. Byun, E.H. Lee, J.D. Hunn, K. Farrell, L.K. Mansur, *J Nucl Mater* **2001**, *294* (3), 256–266.

[137] E.N. Campitelli, P. Spätig, R. Bonadé, W. Hoffelner, M. Victoria, *J Nucl Mater* **2004**, *335* (3), 366–378.

[138] M. Abendroth, M. Kuna, *Eng Fract Mech* **2006**, *73* (6), 710–725.

[139] I. Simonovski, S. Holmström, M. Bruchhausen, *Int J Mech Sci* **2017**, *120*, 204–213.

[140] T. Linse, M. Kuna, J. Schuhknecht, H.-W. Viehrig, *Eng Fract Mech* **2008**, *75* (11), 3520–3533.

[141] L. Lu, M. Dao, P. Kumar, U. Ramamurty, G.E. Karniadakis, S. Suresh, *Proceedings of the National Academy of Sciences* **2020**, *117* (13), 7052–7062.

[142] I. Peñuelas, C. Betegón, C. Rodríguez, J. Belzunce, *Numerical Simulations-Applications, Examples and Theory, INCTECH, Rijeka (Croatia)* **2011**, 311–330.

[143] J. Kameda, *Acta Metall* **1986**, *34* (12), 2391–2398.

[144] T. Misawa, T. Adachi, M. Saito, Y. Hamaguchi, *J Nucl Mater* **1987**, *150* (2), 194–202.

[145] J. McNaney, G.E. Lucas, G.R. Odette, *J Nucl Mater* **1991**, *179-181*, 429–433.

[146] J. Kameda, X. Mao, *J Mater Sci* **1992**, *27* (4), 983–989.

[147] T. Linse, M. Kuna, H.-W. Viehrig, *Mater Sci Eng A* **2014**, *614*, 136–147.

[148] X. Mao, H. Takahashi, *J Nucl Mater* **1987**, *150* (1), 42–52.

[149] R. Hurst, K. Matocha, *K. Matocha, R. Hurst, W. Sun, Determination of Mechanical Properties of Materials by Small Punch and other Miniature Testing Techniques. Ostrava, OCELOT sro* **2012**, 4–18.

[150] W.C. Oliver, G.M. Pharr, *J Mater Res* **1992**, *7* (6), 1564–1583.

[151] J.S. Field, M.V. Swain, *J Mater Res* **1993**, *8* (2), 297–306.

[152] B. Taljat, T. Zacharia, F.M. Haggag, *J Mater Res* **1997**, *12* (4), 965–974.

[153] N. Huber, D. Munz, Ch. Tsakmakis, *J Mater Res* **1997**, *12* (9), 2459–2469.

[154] P. Konopík, J. Dzugan, R. Prochazka, *Metal Brno Czech Republic* **2013**, *2013* (5), 15–17.

[155] Z. X. Wang, H. J. Shi, J. Lu, P. Shi, X. F. Ma, *Nucl Eng Des* **2008**, *238* (12), 3186–3193.

[156] J.H Bulloch, *Int J Press Vessels Pip* **1998**, *75* (11), 791–804.

[157] X. Mao, M. Saito, H. Takahashi, *Scr Metall Mater* **1991**, *25* (11), 2481–2485.

[158] T. Misawa, S. Nagata, N. Aoki, J. Ishizaka, Y. Hamaguchi, *J Nucl Mater* **1989**, *169*, 225–232.

[159] P. Dymáček, K. Milička, *Mater Sci Eng A* **2009**, *510-511*, 444–449.

[160] B. Gülçimen, P. Hähner, *Mater Sci Eng A* **2013**, *588*, 125–131.

[161] C. Sainte Catherine, J. Messier, C. Poussard, S. Rosinski, J. Foulds, *AM Soc Test Mater* **2002**, *4*, 350–370.

[162] E. Fleury, J.S. Ha, *Int J Press Vessels Pip* **1998**, *75* (9), 699–706.

[163] M. Song, K. Guan, W. Qin, J. A. Szpunar, *Nucl Eng Des* **2012**, *247*, 58–65.

[164] M. Abendroth, S. Soltysiak in *Recent Trends in Fracture and Damage Mechanics*, Springer International Publishing, **2015**, pp. 127–157.

[165] M. Abendroth, M. Kuna, *Comp Mater Sci* **2003**, *28* (3-4), 633–644.

[166] F. Abbassi, T. Belhadj, S. Mistou, A. Zghal, *Mater Design* **2013**, *45*, 605–615.

[167] N. Huber, A. Konstantinidis, Ch. Tsakmakis, *J Appl Mech* **2000**, *68* (2), 218–223.

[168] T. Sangkharat, S. Dechjarern, *Proc Manuf* **2019**, *29*, 390–397.

[169] Gary Bradski, Adrian Kaehler, *Learning OpenCV: Computer vision with the OpenCV library*, O'Reilly Media, Inc., **2008**.

[170] R. Hurst, K. Matocha, *Metall J* **2010**, *63*, 5–11.

[171] S. Kohlar, *Wissenschaftlich-Technische Berichte / Helmholtz-Zentrum Dresden-Rossendor* **2017**, *HZDR-082 2017*.

[172] M. Houska, E. Altstadt, *Eur Commiss JRC* **2017**, *SP tests on P92 (T70175) at RT Version 1.0*.

[173] C. Heintze, C. Recknagel, F. Bergner, M. Hernández-Mayoral, A. Kolitsch, *Nucl Instrum Methods Phy Res B* **2009**, *267* (8-9), 1505–1508, Proceedings of the 16th International Conference on Ion Beam Modification of Materials.

[174] A.-A.F Tavassoli, A. Alamo, L. Bedel, L. Forest, J.-M. Gentzbittel, J.-W. Rensman, E. Diegele, R. Lindau, M. Schirra, R. Schmitt, H.C. Schneider, C. Petersen, A.-M. Lancha, P. Fernandez, G. Filacchioni, M.F. Maday, K. Mergia, N. Boukos, Baluc, P. Spätig, E. Alves, E. Lucon, *J Nucl Mater* **2004**, *329-333*, 257–262, Proceedings of the 11th International Conference on Fusion Reactor Materials (ICFRM-11).

[175] C. Zurbuchen, *ASME Pressure Vessels and Piping Conference* **2009**, *Volume 5: High Pressure Technology; Nondestructive Evaluation Division; Student Paper Competition*, 511–517.

[176] H. W. Viehrig, M. Scibetta, K. Wallin, *Int J Press Vessels Pip* **2006**, *83* (8), 584–592.

[177] P. Dymáček, F. Dobeš, Petr Král, J. Dvořák, in *Acta Metallurgica Slovaca-Conference*, **2013**, pp. 57–64.

[178] K. Kumar, A. Pooleery, K. Madhusoodanan, R.N. Singh, J.K. Chakravartty, R.S. Shriwastaw, B.K. Dutta, R.K. Sinha, *J Nucl Mater* **2015**, *461*, 100–111.

[179] M. Bruchhausen, S. Holmström, I. Simonovski, T. Austin, J.-M. Lapetite, S. Ripplinger, F. de Haan, *Theor Appl Fract Mech* **2016**, *86*, 2–10.

[180] G. James, D. Witten, T. Hastie, R. Tibshirani, *An Introduction to Statistical Learning*, *Vol. 112*, Springer New York, **2013**.

[181] A. Jain, G. Hautier, S.P. Ong, K. Persson, *J Mater Res* **2016**, *31* (8), 977–994.

□