Mechanical performance prediction for friction riveting joints of dissimilar materials via machine learning

Bock, Frederic Eberhard; Blaga, Lucian Attila; Klusemann, Benjamin

Published in:
Procedia Manufacturing

DOI:
10.1016/j.promfg.2020.04.189

Publication date:
2020

Document Version
Publisher's PDF, also known as Version of record

Available online at www.sciencedirect.com

ScienceDirect

Procedia Manufacturing 47 (2020) 615–622

Procedia MANUFACTURING

www.elsevier.com/locate/procedia

23rd International Conference on Material Forming (ESAFORM 2020)

Mechanical Performance Prediction for Friction Riveting Joints of Dissimilar Materials via Machine Learning

Frederic Eberhard Bock[a,*], Lucian Attila Blaga[a], Benjamin Klusemann[a,b]

[a]*Institute of Materials Research, Materials Mechanics, Department of Solid State Joining Processes, Helmholtz-Zentrum Geesthacht, Geesthacht, Germany*
[b]*Institute of Product and Process Innovation, Leuphana University of Lüneburg, Lüneburg, Germany*

* Corresponding author. Tel.: +49 (0) 4152 87 26 49; *E-mail address:* frederic.bock@hzg.de

Abstract

Solid-state joining techniques have become increasingly attractive for joining similar and dissimilar materials because it enables further optimization of lightweight components. In contrast to fusion-based joining processes, solid-state joining prevents the occurrence of typical defects such as pores or hot cracking. Machine learning algorithms are powerful tools to identify and quantify relationships between essential features along the process-property chain. In particular, different supervised machine learning algorithms can be used to perform regression analyses and establish correlations between process parameters as well as resulting properties. This can help to circumvent the demand for conducting a vast number of additional experiments to determine optimized process parameters for desired material properties. Additionally, this knowledge can be utilized to obtain a deeper understanding of the underlying mechanisms. In this study, a number of regression algorithms, such as support vector machines, decision trees, random forest and 2nd-order polynomial regression have been applied to correlate process parameters and materials properties for the solid-state joining process of force-controlled friction riveting. Experimental data generated via a central-composite Design of Experiments, serves as source of two separate data sets: one for training and one for testing the machine learning algorithms. The performances of the different algorithms are evaluated based on the determination coefficient R^2 and the standard deviation of the predictions on the test data set. The trained algorithms with the best performance measures can be used as predictive models to forecast specific influences of process parameters on mechanical properties. Through the application of these models, optimized process parameters can be determined that lead to desired properties.

Keywords: Solid state joining; process parameters; ultimate tesnile force; decision trees; random forests; support vector machines

1. Introduction

Continuous scientific, economical and societal demands towards the reduction of costs, fossil-fuel consumption and emission of greenhouse gases, enforce the need to improve lightweight structures for transportation industries and achieve weight-savings.

For joining dissimilar materials, recently developed joining processes, such as friction riveting [1], can contribute to weight-savings as well as economic and ecological use, when compared to conventional joining techniques, such as mechanical fastening and adhesive bonding, due to their respective disadvantages. For the latter, extensive surface treatments and long curing times are obligatory and they are sensitive to moisture and temperature. For the former, numerous parts are required, typically leading to a weight penalty. Additionally, these parts need to be prepared and, when assembled, can create locally concentrated stresses, which reduce fatigue performance and corrosion resistance.

Friction riveting (FricRiveting) is patented by the Helmholtz-Centre Geesthacht (HZG) [2] and was initially developed for joints of AA2024 rivets and polyetherimide (PEI) work pieces [3]. Various aspects of the process have been investigated since. Validation and characterization was

2351-9789
10.1016/j.promfg.2020.04.189

performed for multiple thermoplastic polymers and composites, using steel, aluminum and titanium rivets relevant for industrial applications. The main aspects studied were feasibility (joint formation), microstructural changes, process temperature development and local/global mechanical performance [4]. Moreover, the influence of process parameters on the prior mentioned characteristics were studied via statistical analysis, i.e. Design of Experiments (DoE), where the process was optimized mostly in terms of mechanical performance for the investigated material combinations. The experience gained, together with the development of new joining equipment, led to versatile process variants for FricRiveting, depending on desired joint configurations (point-on-plate joints, multi-material overlaps, etc.). The process can be controlled by time, force and position [1,5].

Pina Cipriano *et al.* [1,6] investigated the modified force-controlled FricRiveting joining process to identify relationships between process parameters, mechanical energies, geometrical joint formation and mechanical properties. A central-composite DoE and the response surface methodology were used to establish an analytical model for predicting joint formation geometries and ultimate tensile forces (UTF), respectively. An interpretation was discussed on how the different process parameters and their resulting mechanical energies imposed on the material, influence the joint formation and mechanical performance. The friction force and friction time were identified as the FricRiveting parameters to linearly correlate with the UTF by 32% and 21%, respectively. Through their linear regression model, the ultimate tensile strength was predicted with a determination coefficient R^2 of 77.9% and a standard deviation of 1065 N.

Machine learning (ML) can be used as a powerful alternative to building prediction models with better performance measures than analytical models, which can lead to further understanding of underlying physical mechanisms. The application of ML and data mining (DM) approaches in materials mechanics has been recently reviewed in [7], where numerous examples are shown how to identify and utilize relationships along the process-structure-property-performance chain. One example is the prediction of the geometry of additively manufactured layers deposited via gas metal arc welding based on the given process parameters (wire feed rate, welding speed, arc voltage and nozzle-to-plate distance) [8]. A simple 3-layered feed forward neural network was used to predict the two outputs – height and width of weld bead layers. Another example is provided by Reimann *et al.* [9], who used random forests and support vector machines for direct and inverse identification of relationships between micromechanical simulations and macroscopic material behavior. And for microstructure prediction, in dependence on process parameters of additive manufactured parts, Popova *et al.* [10] pursued a data-driven approach to perform tasks such as data preparation, microstructure quantification, dimensionality reduction as well as identification and validation of the parameter-microstructure-relationships via different ML-algorithms. However, in their comparison of prediction performances of decision trees, random forests, support vector machines and linear regression, the latter outperformed the former ones, overall.

ML is a new programming paradigm in comparison to classical programming. For classical programming, rules on how input data should be processed to obtain desired answers are explicitly defined through human input. The ML-algorithm is trained with the answers of the input data and creates the corresponding rules. When these rules are applied to new data, the produced answers can be new and original [11]. The DM process consists of six interrelated steps, according to the cross-industry standard [12]: (i) problem understanding, (ii) data understanding, (iii) data preparation, (iv) data modelling, (v) data evaluation and (vi) deployment of the trained algorithm. ML is typically employed in steps (iv) and (v). Thus, ML and DM are key enablers to accelerate the development of novel materials and their processes.

In this study, ML-algorithms are used to build an accurate predictive model to foresee the UTF of force-controlled FricRiveting AA2024-polyetherimide joints. To further improve the joining process, the aim was to build a prediction model of higher accuracy than the commonly used analytical model based on linear regression [6,8]. For this reason, we compare prediction models that are based on 1st- and 2nd-order linear regression, decision tree regression, random forest regression and support vector regression with each other, in order to evaluate and select the most accurate one for the current case study. For the prediction of the UTF, two different sets of input features are used. First, only the process parameters and second, process parameters as well as their corresponding mechanical energies introduced into the material during the process, are considered for training and testing of the predictive models.

2. FricRiveting

FricRiveting is a mechanical fastening technology for joining polymer-metal components, based on the principles of mechanical interlocking and friction welding, see Fig. 1. In the basic metallic-insert process configuration, a plain cylindrical metallic rivet is inserted into a thermoplastic plate under applied rotational speed (RS) and friction force (FF). With rivet insertion and consequent heat accumulation through friction, the polymer melts or softens around the rotating rivet's insertion path and is partially expelled as flash. The heat further concentrates around the rivet, due to the lower heat conductivity of the polymer. This leads to the subsequent plasticizing of the tip of the rivet. The rotation is stopped and the axial forging force (FoF) is applied, with the plasticized rivet tip being thus deformed and anchored under constant pressure within the polymer. The process time is divided into the two characteristic phases: friction time (FT) and forging time (FoT) [4]. The process variant studied in this work is the force-controlled FricRiveting limited by time, whereby the rivet insertion is a process procedure, as detailed by Cipriano *et al.* [1].

FricRiveting is overcoming some of the limitations of conventional joining techniques. It does not require any surface pre-treatment or post-processing. It is a single-step and single-sided joining process, leading to reduced joining cycles and inherent cost optimization (no hole-drilling, single-side accessibility, etc.), fast process speeds (process times between 1-10 seconds, depending on material combination and joint

configuration) and joints with high mechanical performance [4].

Major characteristics of the process are:

- Metallic rivet rotates and under axial force, leading to frictional heat between rivet and polymeric plate.
- Tip of the rivet is plasticized and deformed, enabling mechanical anchoring.
- Metal is processed below melting temperature;
- Polymer is processed above glass transition temperature/melting temperature.
- Bonding is achieved via mechanical anchoring and adhesion forces.

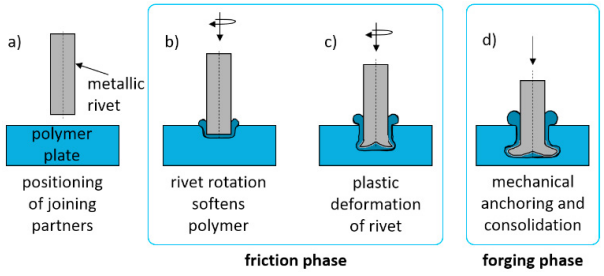

Fig. 1: Process phases of FricRiveting: a) starting position, b) polymer softening, c) rivet plastification, d) forging phase. Figure reprinted from [1], under the terms of the Creative Commons Attribution 3.0 license. http://creativecommons.org/licenses/by/3.0 Arrangement was narrowed and colors were slightly adapted.

The aim of using machine learning is to enable prediction of the UTF with higher accuracy and reduced experimental effort than based on DoE. Furthermore, the use of machine learning can eventually capitalize on previous existing experimental data, while DoE results are usually valid only for the investigated process parameter range and material combination.

3. Machine Learning Regression

Machine learning tasks can be divided into supervised, semi-supervised, unsupervised and reinforcement learning. For supervised learning, which is used in this study, the outcome is known and can be predicted either as a label via classification or as a value via regression analysis [13].

In this section, ML-algorithms utilized in this study are briefly explained. To perform regression analysis for the prediction of the UTF, 1st-order and 2nd-order linear regression (as simple benchmark examples), decision tree regression (DTR), random forest regression (RFR) and support vector regression (SVR) are utilized.

3.1. 1st and 2nd-order linear regression

In linear regression, relationships between independent input and dependent output variables are estimated by adjusting its parameters (or weights), which remain linear. The method is not directly associated as a machine leaning algorithm, but as weights are "learnt", it represents a simplified machine learning technique. In particular, the independent input variables are mapped to the dependent output variables by adjusting the weights with respect to minimizing the residual sum of squares $\sum_{i=1}^{N} \varepsilon_i^2$ between true outputs and predicted outputs approximated by a linear function. The dependent output variable $y(\vec{x})$ can be simply predicted by:

$$y(\vec{x}) = \vec{w}^T \Phi(\vec{x}) + \varepsilon \quad (1)$$

with $\vec{w}$ as the weight vector, ε as the error, $\Phi(\vec{x}) = [1, x_1, \dots, x_n]$ as the 1st-order basis function and n as the number of samples. Similar to 1st-order regression, for d-order regression, weights are adjusted based on the least-squares-error but with a basis function expansion by using a higher order polynomial such as $\Phi(\vec{x}) = [1, x_1, x_1^2, \dots, x_n^d]$, where the weights remain linear.

3.2. Decision tree regression

Decision trees contain simple decision rules defining as few as possible if/else questions for given inputs to arrive at the right output answer. These rules are organized in a hierarchical chain of nodes, where, for regression, values are distinguished based on being above or below a threshold [14], as shown in Fig. 2. DTR is built via two main steps:

- The available output answers/values in the predictor space for the given input values $(X_1, X_2, \dots, X_p)$, are divided into distinct J number of regions $(R_1, R_2, \dots, R_J)$ that do not overlap. The aim is to define regions $R_1, \dots, R_J$ to minimize the residual sum of squares as follows:

$$\min\left(\sum_{j=1}^{J} \sum_{i \in R_j} \left(y_i - \hat{y}_{R_j}\right)^2\right) \quad (2)$$

  where y_i denotes to the i-th output of the input that is used for prediction and $\hat{y}_{R_j}$ to the mean output of the training set outputs for the j-th region. Since consideration of every possible partition of the feature space into J regions is infeasible to compute, top-down recursive binary splitting is performed, where the predictor X_j and a cutpoint s are selected to split the predictor space into regions $\{X|X_j < s\}$ and $\{X|X_j \geq s\}$ which constitute the decision rules in order to minimize the residual sum of squares [15].
- Consequently, the same prediction $\hat{y}_{R_j}$ is made for any input value that is assigned to the same region R_j, which is the mean of output values of the training set that fell into that region [15].

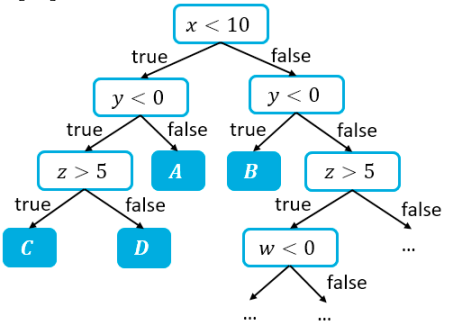

Fig. 2: Schematic of a decision tree, predicting output values for given inputs based on simple decision rules.

3.3. Random forest regression

An RFR is an ensemble method that consists of numerous decision trees whose individual prediction values are averaged to yield the random forest prediction value, as shown in Fig. 3. In particular, each tree node is split based on randomly selected features that are governed by an independently sampled random vector with identical distribution for all trees. The number of trees in a forest is increased until their average converges into a set range [16].

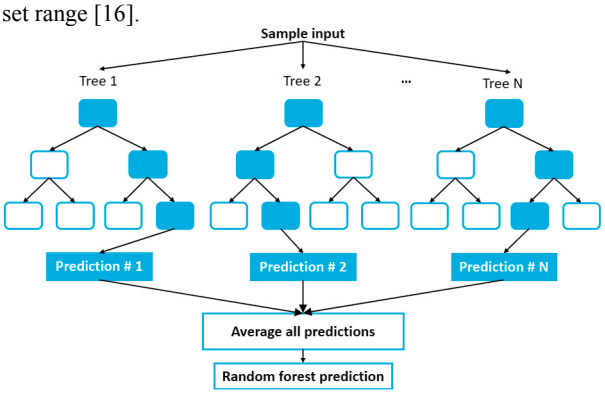

Fig. 3: Illustration of a random forest consisting of several decision trees whose predictions are averaged to the final random forest prediction.

3.4. Support vector regression

SVR is similar to support vector classification (SVC), with a few modifications [17]. For both, a non-linear mapping of input feature vectors into a higher-dimensional feature space is performed via the "kernel trick". Through this transformation into a higher dimensional space, a linear function can be defined that describes all data points, which would be unfeasible in the lower dimensional space. A simplified illustration of the kernel trick is shown in Fig. 4. The major difference of SVR compared to SVC is the output of a real value, as opposed to a classification label in SVC. The aim of SVR is to find a function that fits the target data with a minimal distance to the margins, denoted as ϵ. ξ_1, ξ_2 and ξ_3 are the support vectors.

Based on the work of Abu-Mostafa [18], a constraint minimization problem is solved, through the following term:

$$\min_{\vec{w},b,\xi} \left(\frac{1}{2}\vec{w}^T\vec{w} + C\sum_{i=1}^{N}\xi_i\right) \tag{3}$$

with weight vector $\vec{w}$, bias constant b, margin violation ξ and number of samples N for given inputs y_i and corresponding outputs x_i under the restriction of $\xi_i \geq 0$ and $1 \leq i \leq N$ as well as the constraint $y_i(\vec{w}^T x_i + b) \geq 1 - \xi_i$. Through adjustment of the penalty term C, the margin can be either soft or hard. When mapping the problem from the primal space to the dual space, the weight vector reads:

$$w = \sum_{i=1}^{N}\alpha_i\, y_i x_i \tag{4}$$

and the problem can be expressed as a maximization problem of α:

$$\alpha = \max\left(\sum_{i=1}^{N}\alpha_i - \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}\alpha_i\alpha_j y_i y_j \vec{x}_i^T\vec{x}_j\right) \tag{5}$$

with respect to the constraints $0 \leq \alpha_i \leq C$ and $\sum_{i=1}^{N}\alpha_i\, y_i = 0$. Based on this formulation, the space can be changed via substituting $\vec{x}_i^T\vec{x}_j$ with $\vec{z}_i^T\vec{z}_j$ and generalizing this inner product with a kernel function, which is known as the "kernel-trick":

$$\vec{z}_i^T\vec{z}_j = K(x_i, x_j) = \Phi(x_i)\cdot\Phi(x_j) \tag{6}$$

with $\Phi: X \to Z$. The particular kernel function can be individually defined. Popular kernel functions are the radial basis function $K(x_i, x_j) = \exp(-\gamma\|x_i - x_j\|^2)$ and the sigmoid function $K(x_i, x_j) = \tanh(ax_i^T x_j + r)$, where a, γ and r are respective kernel parameters. Ultimately, a prediction hypothesis is formulated via the following approximation function:

$$f(x) = \vec{w}^T x + b = \sum_{i=1}^{N}\alpha_i y_i \vec{x}_i^T\vec{x}_j + b. \tag{7}$$

Overall, SVR is valuable because of flexible representation of complex functions of non-linear problems, some resistance to overfitting, as well as robust generalization ability due to the margin separator and the constraint optimization being performed in a convex space with no local minima.

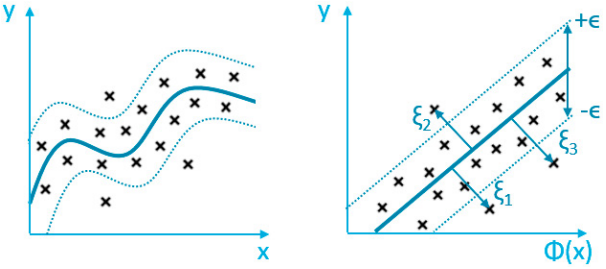

Fig. 4: Illustration of the "kernel trick" performed during support vector regression from (a) space where data points are hard to be linearly separated to (b) space where data points are easily linearly separable.

4. Methodology

In this section, the details of the DoE taken from Pina Cipriano *et al.* [1,6] are briefly reviewed, as this experimental data set represents the basis to be split into training and testing data sets for the ML-algorithms of this work. Afterwards, the ML-methodology in terms of data preparation and data modeling is explained, followed by a description on how the ML-model predictions were evaluated based on performance measures.

4.1. Design of Experiments for force-controlled FricRiveting

A DoE was performed to identify relationships between process parameters, mechanical energies and UTF for joining AA2024 rivets, 5 mm in diameter, with polyetherimide (PEI) plates of 13 mm nominal thickness. Both materials are frequently used in the automotive and aircraft industry [19]. According to the applied central composite design, 36 samples

were required to establish useful links between the UTF and the five process parameters, RS, FT, FoT, FF and FoF.

The work of Cipriano *et al.* [1,6] focused on several aspects concomitantly: joint formation, heat development, energy efficiency and process optimization, i.e. determining the process parameters leading to the maximum mechanical performance. By combining these aspects, the aim was to achieve a tailored approach for force-controlled FricRiveting joints, i.e. the capacity of predicting a certain mechanical behavior under the least amount of mechanical energy input. In Table 1, the DoE data is listed.

Table 1. DoE data-set according to central composite design, as published by Pina Cipriano et al. [1,6], which was split into training and tests sets.

#	RS (*rpm*)	FT (*s*)	FoT (*s*)	FF (*N*)	FoF (*N*)	E_f (*J*)	E_d (*J*)	E_M (*J*)	UTF (*N*)
1	18000	1.6	1.0	2000	5100	10	14	24	1776
2	20000	1.6	1.0	2000	3900	26	20	46	4943
3	18000	2.0	1.0	2000	3900	33	20	53	5427
4	20000	2	1	2000	5100	39	38	77	9619
5	18000	1.6	2	2000	3900	14	14	29	2202
6	20000	1.6	2	2000	5100	16	20	36	3897
7	18000	2	2	2000	5100	35	30	65	6256
8	20000	2	2	2000	3900	41	16	57	7829
9	18000	1.6	1	3000	3900	36	24	60	6391
10	20000	1.6	1	3000	5100	40	43	83	9004
11	18000	2	1	3000	5100	63	43	106	8192
12	20000	2	1	3000	3900	90	66	155	9362
13	18000	1.6	2	3000	5100	45	33	78	8251
14	20000	1.6	2	3000	3900	55	31	86	8046
15	18000	2	2	3000	3900	82	39	120	9106
16	20000	2	2	3000	5100	122	86	208	8996
17	19000	1.8	1.5	2500	4500	39	25	63	7290
18	19000	1.8	1.5	2500	4500	42	34	76	9304
19	19000	1.8	1.5	2500	4500	42	32	74	8824
20	19000	1.8	1.5	2500	4500	42	31	73	9033
21	19000	1.8	1.5	2500	4500	29	27	56	6068
22	19000	1.8	1.5	2500	4500	30	28	59	7663
23	19000	1.8	1.5	2500	4500	24	23	47	5041
24	19000	1.8	1.5	2500	4500	37	30	67	8701
25	19000	1.8	1.5	2500	4500	36	32	68	7741
26	19000	1.8	1.5	2500	4500	40	31	71	8461
27	17000	1.8	1.5	2500	4500	28	23	51	5689
28	21000	1.8	1.5	2500	4500	46	37	83	9049
29	19000	1.4	1.5	2500	4500	17	19	36	3166
30	19000	2.2	1.5	2500	4500	84	52	136	8643
31	19000	1.8	0.5	2500	4500	36	28	64	9098
32	19000	1.8	2.5	2500	4500	40	32	73	9029
33	19000	1.8	1.5	1500	4500	21	17	38	1096
34	19000	1.8	1.5	3500	4500	88	72	159	7864
35	19000	1.8	1.5	2500	3300	34	24	59	6811
36	19000	1.8	1.5	2500	5700	42	44	86	9668

The mechanical energy E_M introduced into the material was calculated according to the following formula:

$$E_M = E_f + E_d = \int M \cdot \omega \cdot dt + \int F \cdot \vartheta \cdot dt \quad (8)$$

with E_f as frictional energy, based on measured torque M and rotational speed ω, as well as with the estimated deformation energy E_d, which is determined based on axial force F and deformation rate ϑ.

4.2. Machine Learning Prediction

In this section, the workflow for implementing ML-predictions, based on the cross-industry standard on the data mining process [12] is described. Addressed are the following points: problem understanding, data understanding, data preparation, as well as data modelling and evaluation via ML, which are all inter-connected and dependent.

Two different sets of input features were chosen for the prediction of the output feature UTF (in N). First, the input feature space for training the ML-models contained only the five process parameters RS, FT, FoT, FF and FoF. Second, that input feature space was further enriched by the mechanical energies E_f, E_d and E_M (in J, respectively) as set-actual correctives since they are based on measurements during the process.

For data-preparation, the different samples of the DoE study can be categorized into axial points, center points and factorial points. The axial points contain the upper and lower limits of the parameter ranges, the factorial points are intermediate points within these ranges and the center points are replicates of the midpoint values within these ranges. The DoE data set was split into a training set and a test set with fractions of 80% and 20%, respectively, i.e. 80% of the complete data set was attributed to the training set, whereas the remaining 20% was separated and assigned to the test data set. Consequently, samples included in the training set are not included in the test set, and vice versa. The test data set was used for ML-performance evaluation. Each of the axial points was contained in the training set, in order to cover the complete value range of the data by including its limits during training and to perform an interpolating prediction instead of an extrapolating prediction. All testing samples were randomly selected among factorial points, as a result, and lie within the range limits of the training set. The randomization, for splitting samples into training and testing sets among the factorial points, is varied to evaluate the robustness of the prediction algorithms to differently sampled training and test data sets.

Moreover, the replicate samples, where process parameters are kept constant (samples # 18 to # 26), were neither used for training nor testing but exploited for uncertainty quantification of the experimental measurements. In summary, the remaining 27 samples were divided and separated into 21 samples for training and 6 samples for testing. The randomized split was implemented on the samples of the factorial points in three different ways by varying random state numbers for the split (9, 33 and 42) [20]. The workflow was executed in Python with the software library Scikit-Learn [21].

Data modelling is realized via training and testing the following ML-predictors: 1st-order and 2nd-order linear regression, DTR, RFR and SVR. For defining suitable ML-model parameters, i.e. hyper-parameters, a randomized-search and a grid-search, both with 5-fold cross-validation, were conducted for RFR and SVR, respectively [20]. Accordingly, RFR hyper-parameters were set with the number of features to consider while fitting to be 3 and 6, as well as the number of trees in a forest to be 122 and 53, for feature spaces excluding and including mechanical input energies, respectively. SVR hyper-parameters were set with $C = 1.8 \cdot 10^{26}$, $\epsilon = 9$, $\gamma = 1.6 \cdot 10^{-25}$ and a non-linear sigmoid kernel function, employed for both feature spaces. UTF predictions of ML-models on the test data set are evaluated based on performance measurements.

4.3. ML-Performance measurements

For the evaluation of prediction performances of the different ML-models, three performance measures were chosen: the determination coefficient R^2, the adjusted determination coefficient $R^2_{adj.}$ and the standard deviation s. The determination coefficient is determined by:

$$R^2 = 1 - \frac{\sum_{i=1}^{N}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{N}(y_i - \bar{y}_i)^2} \quad (9)$$

with the numerator as the total sum of squares and the denominator as the residual sum of squares, where y_i represents the data set value, $\hat{y}_i$ the predicted value, $\bar{y}_i$ the mean of the data set values and N the number of samples. The adjusted version of the determination coefficient is calculated as:

$$R^2_{adj.} = 1 - (1 - R^2)\frac{N-1}{N-p-1} \quad (10)$$

with p as the number of features. The standard deviation is calculated as:

$$s = \sqrt{\frac{\sum_{i=1}^{N}(\hat{y}_i - y_i)^2}{N-1}} \quad (11)$$

with the same definition of variables as in Eq. (9).

Based on the scatter of the afore excluded samples # 18 to # 26, a calculation of the standard deviation s of the UTF generated by the constant process parameter set (RS, FT, FoT, FF, FoF) was performed. The resulting experimental standard deviation $s_{exp} = 1373\ N$ is used as error bar for evaluating the performances of the ML-predictions of the UTF. The ML-models are evaluated with respect to R^2 and $R^2_{adj.}$ values being as close to 1.0 and s to be as below the experimental standard deviation as possible.

5. Results

UTF predictions of metal-polymer force-controlled FricRiveting T-pull-joints via trained ML-models are tested, compared and discussed. First, results of ML-models trained with only process parameters as input features are presented, followed by the predictions of ML-models trained with mechanical energies and process parameters in the input feature space.

5.1. ML-prediction with only process parameters as input features

The results of the different ML-models are shown in Fig. 5 (a)-(c) and summarized in Table 2. Overall, prediction results generated by 1st-order linear regression were most accurate, in comparison to the other ML-models. In particular, the R^2 value with 0.72 and the adjusted $R^2_{adj.}$ value with 0.63 are highest, whereas the standard deviation $s = 1364$ N, shows the second lowest value, for the data split by random state number 9. This trend is preserved on the two other randomly split training and test data sets, which indicates the highest robustness for 1st-order linear regression among the ML-models. Even though the performance measures R^2 and $R^2_{adj.}$ decline from data splits by random state numbers 42 to 33 and 9, respectively, the standard deviation is simultaneously decreased and remains below the experimental standard deviation of $s_{exp} = 1373$ N, for all predictions. However, on data split by random number 42, R^2 and $R^2_{adj.}$ of SVR amount to 0.65 and 0.53, respectively, which is equal to the values for 1st-order linear regression, but the standard deviation of 1726 N is inferior to the one of 1st –order linear regression with 1294 N. Moreover, the standard deviation of SVR on data split by random number 33, with $1168\ N$, is the lowest among all values but accompanied by inacceptable R^2 and $R^2_{adj.}$ of 0.25 and 0.01, respectively. These exceptions of SVR performance measures seem rather arbitrary and can partly be explained by potential over- and under-fitting of the SVR-Model to the data probably caused by inappropriate hyper-parameters leading to an inability of the trained SVR to generalize well.

The performance measurements for the other ML-models, DT and RFR as well as 2nd-order linear regression are also inferior compared to 1st-order linear regression and are not useful for reliable predictions of the UTF based on the process parameters alone. In general, the prediction performances of all ML-models were not sufficient, even for the best performing algorithm. A reason for the poor prediction performance could be the significant scatter of the UTF in the experimental data (s_{exp} = 1373 N) provoking uncertainties within the training and test data sets, which is challenging for the ML-models to generalize upon and to predict precisely. Because of the poor prediction performance of the presented ML-algorithms so far, and the relatively small data set available, the input feature space was expanded to include the experimentally determined mechanical energies that were introduced into the materials during the process, with the aim to improve the performance measures for ML-predictions of the UTF.

5.2. ML-prediction with process parameters and mechanical energies as input features

Once the input feature space contains mechanical energies E_f, E_d, E_M in addition to process parameters RS, FT, FoT, FS, FoS, some ML-prediction results are substantially improved, whereas others remain similar in terms of overall performance measures, as shown in Fig. 5 (d)-(f) and in Table 3.

The prediction performances of the 1st- and 2nd-order linear regressions remain analogous, as there are neither major improvements nor deteriorations in their performance measures, compared to the predictions with the reduced feature input space containing only process parameters. There seems to be no more significant linear correlation between the now-included mechanical energies and UTF than between input parameters and UTF.

This is confirmed by the improved predictions of DTR and RFR, as they can map non-linear relationships between inputs and outputs. In particular, RFR outperformed in terms of performance measures all other ML-predictors with R^2, $R^2_{adj.}$ and s showing values between 0.90 – 0.92, 0.87 – 0.90 and 691 N – 719 N, respectively. The performance measures of DTR are in the ranges of 0.86 – 0.92 and 0.81 – 0.89 for R^2 and $R^2_{adj.}$, respectively, accompanied by a standard deviation range of 719 N – 1056 N. Comparing RFR and DTR to the performance measures achieved by the linear regression model of Pina Cipriano [6]: $R^2 = 77.9\%$, $R^2_{adj.} = 79.4\%$ and $s = 1065\,N$, both DTR and RFR could also outperform their quality of UTF predictions. Several aspect are assumed to contribute to these good prediction performances. Due to the decision rules contained in DTR, the apparent non-linear relationships between process parameters in combination with mechanical energies and the UTF could be represented more appropriately by the DTR and most appropriately by RFR. Their performances differ, as in DTRs, best features among all features are searched for when splitting nodes, whereas for RFRs, the best features among a random subset of features are searched for. As a result, RFRs develop an increased diversity among the contained DTRs, i.e. an increased bias is traded for a decreased variance, leading overall to an improved prediction model [20]. RFR are also more robust to noise, compared to DTR [16]. Both is beneficial for predictions of the UTF based on the data set used in this study, especially in relation to the scatter among the UTF values.

Since SVR can also represent non-linear relationships between input and output features [17], a prediction capability related to the quality of DTR and RFR was expected to be utilized. However, for SVR, R^2 and $R^2_{adj.}$ are slightly decreased through increasing the input feature space, on all three randomly spilt data sets. As a possible explanation, it can be argued that SVRs have advantages in high dimensional space, when the dimensionality of the input space leads to a feature space that exceeds the number of samples [17], the decreased performance of SVR could be referenced to the low dimensional prediction space of the problem in comparison to the sample number. A more probable explanation for the poor prediction performance of SVR, despite a kernel function that can represent non-linear relationships, is an insufficient selection of a hyper-parameter combination. The grid-search algorithm was employed to search for a suitable combination of hyper-parameters enabling improved predictions. According to the low performance measures, the found hyper-parameter combination is not ideal, probably because a suitable one lies outside the considered space or resolution of the grid-search algorithm. Additionally, the choice of a sigmoid kernel function might not have been best to represent the specific non-linear relationships. Thus, it is assumed that there is potential for improvement, in the future.

Generally though, a comparison of the feature input space with only process parameters to the feature input space including both process parameters and experimentally determined mechanical energies, leads to the suggestion that the energies help to consider important non-linear relationships between set process parameters and actual process conditions, as their use improves the predictions of the manufacturing process. The standard deviation of the replicate experiments (samples #18-26) with $s_{exp} = 1373\,N$ indicates a substantial uncertainty in the experimental process and its control, reinforcing the need for having an accurate prediction model that accounts for the set-actual deviation by considering the experimentally determined mechanical energy input into the material. The enhanced RFR-model delivers better prediction performances for evaluation of whether or not joints withstand a required UTF prior to failure. Process parameters remain the only possible adjustments in experiments; therefore, this model can serve as predictive tool to assure that desired UTF of produced joints are met, based on not only process parameters but on additional consideration of experimentally determined mechanical energies. As a result, when the energy can be monitored and controlled more precisely, desired UTF could be tailored to the specific/anticipated FricRiveting application by employing the presented models.

Table 2. Prediction performance measures of ML-algorithms with only process parameters as input features.

Performance measure	Random state	1st-LR	2nd-LR	SVR	DTR	RFR
R^2	9	**0.72**	-0.05	0.57	0.36	0.56
	33	**0.56**	-0.26	0.25	-0.22	0.12
	42	**0.65**	0.52	**0.65**	0.02	0.23
R^2 adj.	9	**0.63**	-0.4	0.43	0.15	0.41
	33	**0.41**	-0.68	0.01	-0.62	-0.17
	42	**0.53**	0.36	**0.53**	-0.30	-0.03
Standard	9	**1364**	2727	1541	2136	1756
deviation	33	1319	2603	**1168**	2052	1686
in N	42	**1294**	2029	1726	2658	2488

Table 3. Prediction performance measures of ML-algorithms with process parameters and mechanical energies as input features.

Performance measure	Random state	1st-LR	2nd-LR	SVR	DTR	RFR
R^2	9	0.72	-0.15	0.48	**0.92**	**0.92**
	33	0.71	-0.2	0.11	0.90	**0.92**
	42	0.55	-0.13	0.21	0.86	**0.90**
R^2 adj.	9	0.65	-0.44	0.35	**0.90**	**0.90**
	33	0.64	-0.5	-0.11	0.88	**0.90**
	42	0.43	-0.41	0.02	0.82	**0.88**
Standard	9	1404	2839	1917	748	**702**
deviation	33	1356	2741	1384	719	**691**
in N	42	1860	3082	1358	1056	**719**

6. Conclusion

In conclusion, consideration of correlations between process parameters and mechanical energy input for training ML-

models to predict mechanical properties of produced parts is highly beneficial for the prediction performance. For force-controlled FricRiveting joints, in particular, RFR outperformed other ML-algorithms with respect to their ability to predict the UTF, based on the used performance measures. Additionally, the ML-model can predict the UTF of FricRiveting with higher precision than linear models previously used with DoEs.

In the future, the robustness of the ML-model prediction should be challenged by changing the material combinations. The use of artificial neural networks needs also to be evaluated, as the identification of an appropriate network architecture, the limited amount of available data and under/overfitting of the algorithm present some challenges.

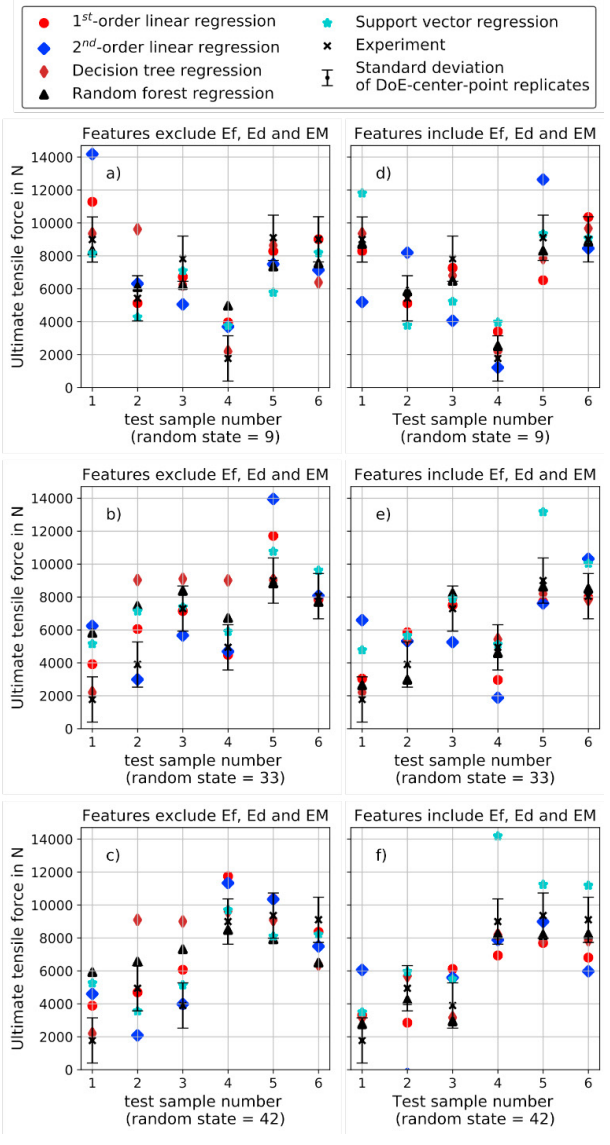

Fig. 5: ML-predictions of UTF (a),(b),(c): with only process parameters as input features; as well as (d),(e),(f): with process parameters and mechanical energies as input features: Data sets were split via random state numbers 9 (top), 33 (centre) and 42 (bottom).

Acknowledgements

The authors acknowledge funding from the Helmholtz-Association via an ERC-Recognition-Award (ERC-RA-0022).

References

[1] Pina Cipriano G, Blaga LA, F Dos Santos J, Vilaça P, Amancio-Filho ST. Fundamentals of Force-Controlled Friction Riveting: Part I-Joint Formation and Heat Development. Materials 2018;11(11).

[2] Amancio-Filho ST, Beyer M, dos Santos JF. Method of connecting a metallic bolt to a plastic workpiece: United States Patent(US 7,575,149 B2).

[3] Amancio-Filho ST, dos Santos JF. Influence of processing parameters on microstructure and properties of a polyetherimide joined by fricriveting: investigation of rotational speed. ANTEC 2009, 67th Annual Technical Conference of the Society of Plastics Engineers, Vol. 2, Chicago, IL, US, Jun 22-24, 2009 2009:750–6.

[4] Amancio S, Blaga L. Joining of polymer-metal hybrid structures: Principles and applications. New York, NY: John Wiley et Sons; 2018.

[5] Borba NZ, Blaga L, dos Santos JF, Amancio-Filho ST. Direct-Friction Riveting of polymer composite laminates for aircraft applications. Materials Letters 2018;215:31–4.

[6] Pina Cipriano G, Blaga LA, Dos Santos JF, Vilaça P, Amancio-Filho ST. Fundamentals of Force-Controlled Friction Riveting: Part II-Joint Global Mechanical Performance and Energy Efficiency. Materials 2018;11(12).

[7] Bock FE, Aydin RC, Cyron CJ, Huber N, Kalidindi SR, Klusemann B. A Review of the Application of Machine Learning and Data Mining Approaches in Continuum Materials Mechanics. Front. Mater. 2019;6:443.

[8] Xiong J, Zhang G, Hu J, Wu L. Bead geometry prediction for robotic GMAW-based rapid manufacturing through a neural network and a second-order regression analysis. J Intell Manuf 2014;25(1):157–63.

[9] Reimann D, Nidadavolu K, ul Hassan H, Vajragupta N, Glasmachers T, Junker P et al. Modeling Macroscopic Material Behavior With Machine Learning Algorithms Trained by Micromechanical Simulations. Front. Mater. 2019;6:605.

[10] Popova E, Rodgers TM, Gong X, Cecen A, Madison JD, Kalidindi SR. Process-Structure Linkages Using a Data Science Approach: Application to Simulated Additive Manufacturing Data. Integr Mater Manuf Innov 2017;6(1):54–68. https://doi.org/10.1007/s40192-017-0088-1.

[11] Chollet F. Deep learning with Python. Shelter Island, NY: Manning; 2018.

[12] Shearer C. The CRISP-DM Model: The New Blueprint for Data Mining. Journal of Data Warehousing 2000;5(4):13–22.

[13] Witten IH, Frank E, Hall MA. Data mining: Practical machine learning tools and techniques. 3rd ed. Amsterdam: Elsevier/Morgan Kaufmann; 2011.

[14] Quinlan JR. Induction of decision trees. Machine Learning 1986;1(1):81–106. https://doi.org/10.1007/BF00116251.

[15] James G, Witten D, Hastie T, Tibshirani R. An introduction to statistical learning: With applications in R. 8th ed. New York, Heidelberg, Dordrecht, London: Springer; 2017.

[16] Breiman L. Random Forests. Machine Learning 2001;45(1):5–32. https://doi.org/10.1023/A:1010933404324.

[17] Drucker H, Burges CJC, Kaufman L, Smola AJ, Vapnik V. Support vector regression machines. In: Advances in neural information processing systems; 1997, 9:155–161.

[18] Abu-Mostafa YS, Magdon-Ismail M, Lin H-T. Learning from data: A short course. [s. l.]: AML; 2012.

[19] Thomas S, P. M. V. Handbook of engineering and specialty thermoplastics. Hoboken, N.J.: Wiley; 2012.

[20] Géron A. Hands-on machine learning with Scikit-Learn and TensorFlow: Concepts, tools, and techniques to build intelligent systems. Beijing, Boston, Farnham, Sebastopol, Tokyo: O'Reilly; 2017.

[21] Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O et al. Scikit-learn: Machine Learning in Python. J Mach Learn Res 2011;12:2825–30.