



Gathering of Process Data through Numerical Simulation for the Application of Machine Learning Prognosis Algorithms

Mathias Jäckel, Tobias Falk, Julius Georgi, Welf-Guntram Drossel

► To cite this version:

Mathias Jäckel, Tobias Falk, Julius Georgi, Welf-Guntram Drossel. Gathering of Process Data through Numerical Simulation for the Application of Machine Learning Prognosis Algorithms. *Procedia Manufacturing*, 2020, 47, pp.608 - 614. <10.1016/j.promfg.2020.04.186>. <hal-03491079>

HAL Id: hal-03491079

<https://hal.science/hal-03491079v1>

Submitted on 22 Aug 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons CC BY-NC 4.0 - Attribution - Non-commercial use - International License



23rd International Conference on Material Forming (ESAFORM 2020)

Gathering of Process Data through Numerical Simulation for the Application of Machine Learning Prognosis Algorithms

Mathias Jäckel^{a*}, Tobias Falk^a, Julius Georgi^a, Welf-Guntram Drossel^a

^a*Fraunhofer Institute for Machine Tools and Forming Technology IWU, Nöthnitzer Straße 44, Dresden 01187, Germany*

* Corresponding author. Tel.: +49 351 4772-2434; fax: +49 351 4772-32434. E-mail address: mathias.jaeckel@iwu.fraunhofer.de

Abstract

In recent years, FEA simulation of forming processes has increasingly developed as a good alternative to complex experimental work in the determination of process parameters and product properties. However, detailed material data (e.g. flow curves) are necessary for the execution of FEA simulations, which are often not available to manufacturers and users in the early stages of the product development. In this paper, a method is shown by which application it is possible, that only on the basis the general mechanical properties (e.g. tensile strength, sheet thickness) and the use of data-based prognosis models of supervised machine learning to predict directly a result regarding suitable process parameters as well as expected forming result properties. Thereby an extensive technological database was generated for the joining by forming process self-pierce riveting (SPR) by means of numerical simulation. Subsequently, different learning algorithms are trained using these numerical data and their prediction quality is compared.

© 2020 The Authors. Published by Elsevier Ltd.

This is an open access article under the CC BY-NC-ND license <https://creativecommons.org/licenses/by-nc-nd/4.0/>

Peer-review under responsibility of the scientific committee of the 23rd International Conference on Material Forming.

Keywords: Machine Learning; Forming Simulation; Self-pierce riveting

1. Introduction

The digitalization of production is currently one of the most important fields of action to ensure growth and employment in the future. Mechanical joining technology plays an important role as an interlinked cross-sectional technology and must continue to develop within the framework of digitalization in order to remain competitive. The application of machine learning methods has great potential to enable digital business models for system manufacturers and users of mechanical joining technology.

An essential development effort in the integration of mechanical joining technology into production is the design of the joining tools and the determination of the resulting joint properties. This development work is mainly experience-based and experimental. First, the process parameters are selected on the basis of experience and the resulting joint properties are verified experimentally by

metallographic analyses. If the joint properties do not meet the expectations, further iteration loops are performed until an acceptable solution is found. The Ford F-150, in which semi-tubular self-pierce riveting was able to replace spot welding as the main joining process in large-scale production for the first time, is an example of the use of approx. 60 rivet-die combinations. The process parameters were determined purely experimentally [1]. Especially with constantly changing design statuses, which users of mechanical joining technology are confronted with, e.g. in automotive engineering, the experimental procedure is time-consuming and cost-intensive. In addition, the components to be joined and the corresponding equipment must be available for parameter determination, which is often the case only relatively shortly before the start of production, especially with smaller suppliers. In addition, the necessary know-how for the selection of suitable joining parameters must be available.

In the research described here, a methodology was applied by whose application it is possible to predict directly a statement about suitable joining parameters (e.g. rivet, die) as well as expected joint properties (e.g. interlock) using data-based calculation models of supervised machine learning and only general information (thickness, tensile strength) about the sheets to joined as input data.

In research work on mechanical joining technology, methods of machine learning have already been used. In [2] neural networks were used to predict process parameters and joint strength in solid self-pierce riveting. In the investigations [3] and [4] first neural networks and later k-nearest neighbor regressions were used to predict the load bearing capacity of clinch joints. In [5] neural networks were used to classify errors in process monitoring for radial clinching. According to [5], inadmissible process states could be classified well with the help of neural networks that were trained by means of the backpropagation method. Lambiasi used in [6] 27 clinching simulations to train an Artificial Neural Network (ANN) with one hidden layer with 100 Neurons and three output layers to create an optimization tool for the selection of clinching parameters which lead to a high joint strength. According to [6], despite the small amount of data, good forecast results of ANN could be observed.

2. Considered Process: Self-pierce riveting

2.1. Process description

Self-pierce riveting with semi-tubular rivet (SPR-ST) is the most used mechanical joining technology for car bodies when using aluminum and steel combinations. This joining method can be divided in three steps, shown in Fig. 1. The first step is characterized by positioning the rivet and the sheets between the punch, blank holder and die (a). When the punch presses the rivet in the punch-sided sheet, the rivet pierces a slug out of the material, which remains inside the cavity (b). The contour of the die forces the rivet to expand and an interlock is created (c). [7]

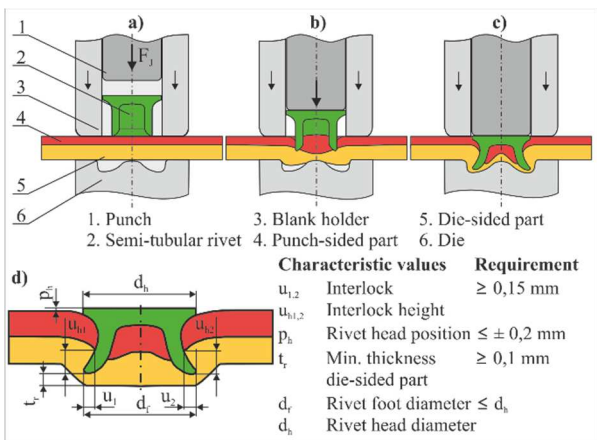


Fig. 1. Self-pierce riveting: a) – c) Process steps, d) Characteristic values [7]

SPR joints are evaluated by certain geometrical criteria (Fig. 1 d). These criteria correlate with the strength properties of the joint and must fulfill specific values in order that the joint can be considered as OK. The required values stated are based on standards from the automotive industry. No specified values are known for the interlock height.

2.2. Material combinations

In the study described here, the SPR-ST process of steel and aluminum sheets is investigated. Thereby, the materials and thicknesses described in Table 1 are considered. The steel sheets with $R_m > 650 \text{ MPa}$ were only placed on the punch-side, otherwise there would have been unwanted deformations of the considered rivets.

Table 1. Considered materials

Arrange-ment	Material	$R_m / R_{p0,2}$ in MPa	Thicknesses in mm			
Punch- and die- sided	EN AW-6016 T4	240 / 129	0.8	1.15	1.5	2.0
	EN AW-5182	287 / 144	1.15	1.25	1.5	2.0
	CR210IF	363 / 241	0.8			
	DX53D	287 / 144	1.0	1.25	1.5	1.75
	HCT600XD	627 / 390	1.25	1.5	1.75	2.0
Only punch- sided	HCT780XD	1035 / 622	1.25			
	HCT980XG	1022 / 865	1.0			

3. Experimental reference joints

In order to validate the simulation models of the SPR process and to generate a separate experimental database for the following considerations, with the materials of table 1 125 combinations were identified on the basis of a partial factorial experimental design according to the optimized latin hypercube sampling method [8] and subsequently experimentally joined and evaluated. The process parameters like rivet length and die design were chosen experience based.

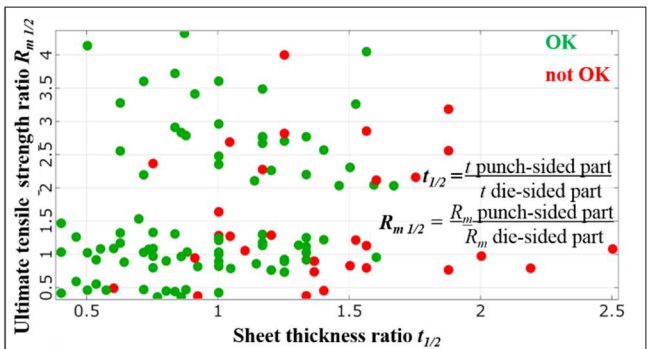


Fig. 2. Distribution of OK and not OK joints in relation to the tensile strength and sheet thickness ratio of the joined parts

Fig. 2 visualizes the results of experimental joining investigations. Thereby green dots represent the joints, which fulfill all necessary criteria (Fig. 1 d) and the red dots the joints which do not meet at least one of the criteria in relation

to the tensile strength and sheet thickness ratio of the punch- and die-sided sheets.

Thereby it can be noted, that sheet thickness ratios above 1.7 lead to critical results, which is in good correspondence with the general advice for SPR that it is better to join thinner in thicker sheets [7]. The average scatter of the geometrical characteristic values interlock and minimum thickness of the die-sided part in the experiments was approx. 20 % with three evaluated experiments per test series. This value is cited in the later investigations as a quality characteristic for the accuracy of the numerical or data-based prognoses. Interlock and minimum thickness of the die-sided part are considered, as these are by far the most important characteristic values and show the most correlation to the joint strength.

4. Data Mining via Numerical Simulation

4.1. Simulation Models

Following for all 125 experimental joints 2D simulation models in Simufact V15 with the in Fig. 3 shown structure were build up. Due to the following numerical sensitivity analysis the simulation models are designed with the background in mind to achieve a good balance between forecasting accuracy and computing time. An average computing time of 13 minutes per SPR process simulation is achieved on a workstation with 14 cores.

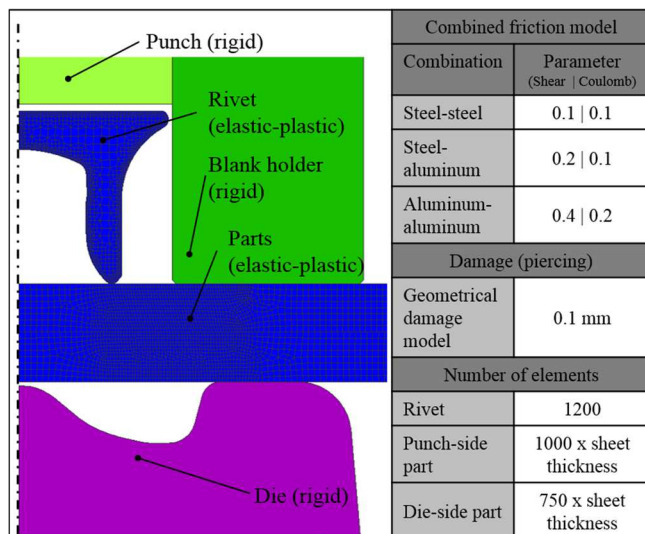


Fig. 3. 2D-simulation model for the SPR-process in Simufact V15

The chosen parameters of the combined friction model are a result of a numerical sensitivity analysis by fitting the calculated with the experimental joint contour and force-displacement curve of the SPR process. Flow curves for the sheet and the rivet material are determined by stack compression tests (SCT) due to the good comparability in terms of stress state between SPR and SCT [9].

Fig. 4 shows an overview of the validation of the SPR simulation models with the experimental data. All numerically calculated joining results between the red dashed lines represent connections in which the deviation

between simulation and experiment is at most 20 % and thus, also in retrospect to the scattering of the experiments, have sufficient accuracy. In total, 71 of the 125 simulation models calculated meet these accuracy criteria and can be used for the following sensitivity analysis.

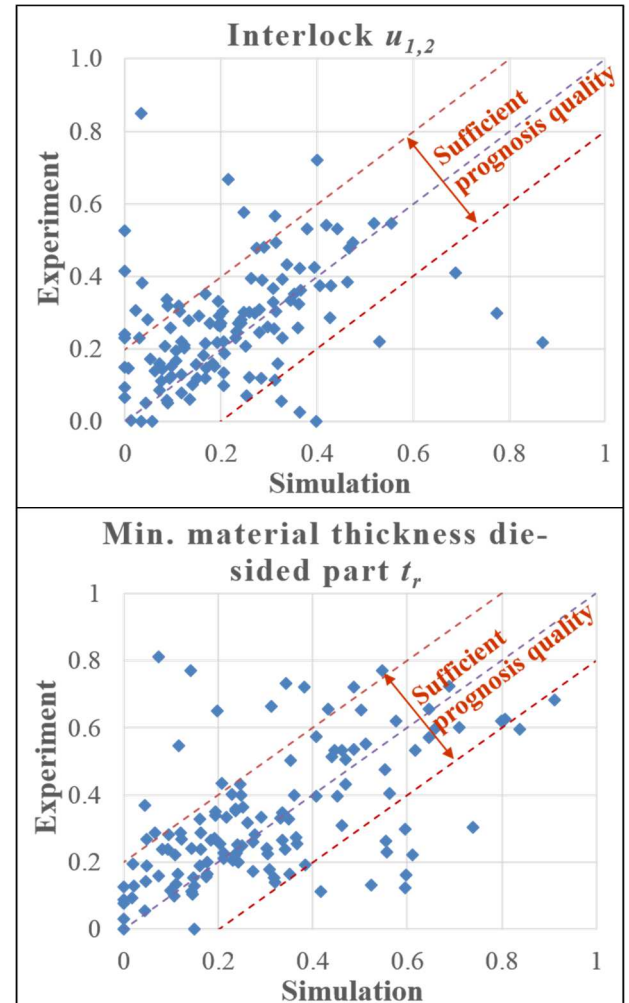


Fig. 4. Deviation of the simulated to the experimental result for interlock and minimum thickness of the die-sided part for the SPR process of the 125 considered material combinations

4.2. Numerical sensitivity analysis

Goal of the numerical sensitivity analysis is to generate an extensive database for the SPR process of the steel and aluminum sheets under consideration. In numeric, both process parameters and material properties can be varied much more flexibly than in experiments. 150 parameter variations are performed for each of the validated 71 simulation models, which will result in a total number of 10650 SPR joining results. For each material combination, the following input variables are individually varied in relation to the original values:

Material properties of the parts:

- Sheet thickness: -0.15; -0.1; -0.05; 0; +0.05; +0.1; +0.15 mm
- Flow curve variation: -10; -5; 0; +5; +10 %

Rivet properties (Fig. 5, right):

- Geometry: C-form; P-form
- Length (l_N): -0.25; 0; +0.25 mm

Die geometry (Fig. 5, left):

- Contour diameter (d_m): -0.5; 0; +0.5 mm
- Contour depth (h_m): -0.25; 0; +0.25 mm
- Thorn height (h_d): 0; 0.5; 1.0; 1.5 mm
- Thorn width (t_w): 0; 2.0; 3.0; 4.0 mm
- Contour edge radius (r_i): 0.1; 0.25; 0.5 mm

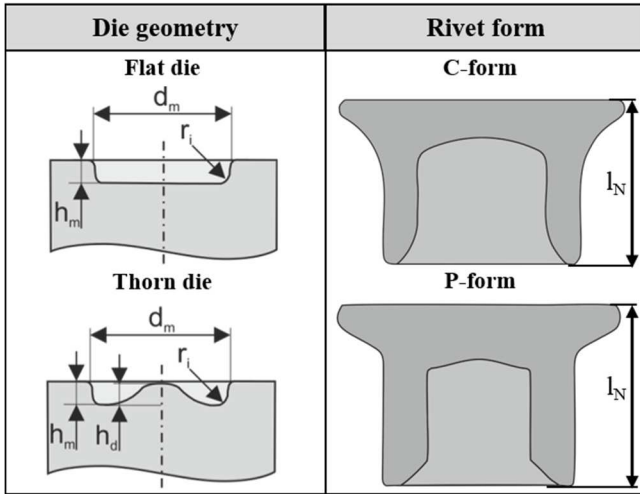


Fig. 5. Definition of die and rivet geometry

Due to the high number of possible parameter variations, a partial factorial design of simulations according to the optimized latin hypercube sampling method was generated. All the required material and geometry data were integrated into the Simufact Joining Optimizer, which is a special tool for the automatization of mechanical joining simulations inside the Simufact Joining software allowing the multitude of simulations to be built up, carried out and evaluated automatically. All the 10650 process simulations were calculated in a total time of approx. 630 h on a workstation with 14 cores.

5. Available databases for training of algorithms

Two databases are available in the origin for the evaluation of the process data of the SPR of the considered material combinations. The experimental database with 125 joining results (DB 1) and the numerically calculated database with 10650 joining results (DB 2). Since the numerical database is based on a statistical design it contains a series of simulations, which are incorrect from a technological point of view. Therefore, this database was additionally adjusted in order to improve the later databased prognosis. The following technological filters were included:

- Interlock $u_{1,2} \geq 0$ mm
- Min. thickness die-sided part $t_r \geq 0$ mm
- Rivet foot diameter $d_f \leq 8.2$ mm
- Rivet head position $-0.5 \text{ mm} \leq p_h \leq 0.5 \text{ mm}$

- Max. joining force $30 \text{ kN} \leq F_J \leq 85 \text{ kN}$

A total number of 2376 joining results fulfill these criteria, which are considered in the following as a separate database (DB 3). Fig. 6 visualizes the material combinations of all considered databases for the SPR process. From this, it becomes clear, that the considered range of material thicknesses and strength variants are well covered by all databases.

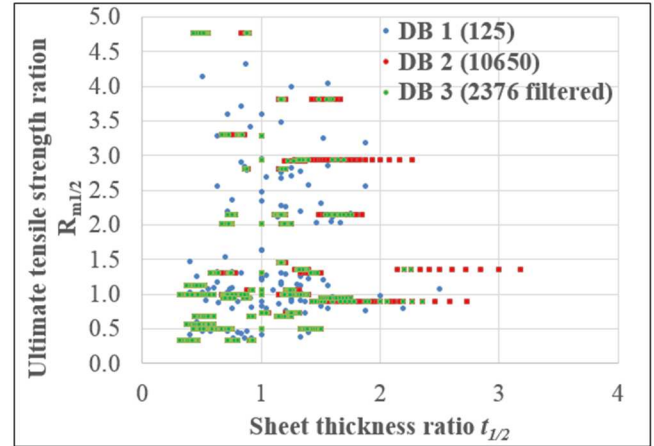


Fig. 6. Overview about the material combinations of all considered databases for the SPR process

6. Prognosis algorithms

6.1. Quality criteria

Following six different prognosis algorithms (models) are tested. Therefor the data set is split randomly in a trainings set (80 % of all data) to train the models and a test set (remaining 20 %). With these test data a comparison of original (y_i) and predicted (\hat{y}_i) values can be executed and with the coefficient of determination R^2 (1) all models can be compared and rated (\bar{y} mean value).

$$R^2 = 1 - \frac{\sum (y_i - \hat{y}_i)^2}{\sum (y_i - \bar{y})^2} \quad (1)$$

The closer original and predicted values are, the higher the percentage of R^2 and the better the quality of prediction model.

6.2. Considered prognosis algorithms

6.2.1. Linear regression

Linear models make the assumption that the target is a linear combination, so the predicted value $\hat{y} = (w_0 + w_1x_1 + \dots + w_nx_n)$. Linear Regression (LR) is the easiest and fastest model, that is part of this paper. The coefficients w are determined in the way, that the residual sum of squares between original and predicted data by the linear approximation is minimized (2).

$$\min_w |\hat{y} - y|^2 \quad (2)$$

6.2.2. Huber regression

Robust regression fits a regression model of corrupt data. The values are either outliers or errors in the model. Huber Regression (HR) is a robust algorithm, that applies a linear loss function to samples that are classified as outliers. Samples are classified as inliers if the error of it is less than a defined critical value. The aim of the Huber Regressor is to minimize (3), where the back part is the loss function [10]:

$$\min_{w, \sigma} \sum \left(\sigma + H_{\varepsilon} \left(\frac{x_i w - y_i}{\sigma} \right) \sigma \right) + \alpha |w|^2, \quad (3)$$

with

$$H_{\varepsilon}(z) = \begin{cases} z^2, & \text{if } |z| < \varepsilon, \\ 2\varepsilon|z| - \varepsilon^2, & \text{otherwise.} \end{cases}$$

In the present study the regularization parameters α , the value of samples that are classified as outliers and the maximum number of iterations were varied and the results of the best parameter setting with regard to R^2 is shown in the following.

6.2.3. Support Vector Regression

A Support Vector Machine (SVM) divides a set of objects into classes in such a way that a large area around the class boundaries remains free of objects. Advantages of this supervising learning method are the effectiveness in high dimensional spaces as well as the cases where the number of samples is less than the number of dimensions and the efficiency in memory. The main disadvantage of SVM is the possibility of over-fitting that means the predicted values are close to the tested ones, but may fail to fit additional data. Usually used for classification, the method can be extended to the in this paper used regression models, so called Support Vector Regression (SVR). [11]

Here, for the SVR two different kernel types are compared with various regularization values and distance epsilon setting.

6.2.4. k-nearest-neighbor

A k-nearest-neighbor (k-NN) algorithm classifies a sample by taking its k nearest neighbors into account. For regression or continuous numbers/values, the average of the values of the k nearest neighbors is determined [12]. There are four different activation algorithms for k-NN implemented:

- Brute Force: The most naive neighbor search implementation involves the brute-force computation of distances between all pairs of points in the dataset. It is very efficient for small data sets, but with a growing number of samples it becomes infeasible.
- K-D(imensional) Tree: To avoid the inefficiency of the brute force algorithm, tree-based approaches were invented. The main idea of this one is to reduce the

distance calculations by the use of transitivity. For example if point 1 is far away from point 2 and point 2 is close to point 3, then the distance between point 1 and point 3 has not to be calculated, because of the relation to point 2 it is obvious, that point 1 is far away from point 3.

- Ball Tree: Because the K-D tree algorithm partitions the data set along Cartesian axes, the ball tree method was developed to partition in hyper-spheres. This is more costly, but efficient with regard to high dimensions.
- Auto: This setting attempts to find the best approach for the actual data set.

In addition to the four described activation algorithms the number of neighbors is varied as well as the weight function used in the prediction.

6.2.5. Gradient Boosted Decision Tree

The goal of ensemble methods is to combine the predictions of several base estimators built with a given learning algorithm in order to improve generalizability / robustness over a single estimator. In boosting methods, such as Gradient Boosted Decision Trees (GBDT), base estimators are built sequentially and one tries to reduce the bias of the combined estimator. The motivation is to combine several weak models to produce a powerful ensemble. GBDT is a generalization of boosting to arbitrary differentiable loss functions. [13]

The two main parameter settings that are investigated in this paper are the number of boosting stages (a large number should lead to better results) and maximum number of nodes in the tree (best value depends on the interaction of the training data).

6.2.6. Multilayer perceptron – artificial neural network

A multilayer perceptron (MLP) is a class of feedforward artificial neural network (ANN). An MLP consists of at least three layers of nodes: an input layer, a hidden layer and an output layer. Except for the input nodes, each node is a neuron that uses a nonlinear activation function. Each node in one layer connects with a certain weight to every node in the following layer. Learning occurs in the perceptron by changing connection weights after each piece of data is processed, based on the amount of error in the output compared to the expected result. [14]

In this study the number of hidden layers as well as the number of neurons in each layer are varied. Furthermore three different functions to activate the hidden layers are implemented:

- rectified linear unit function, $f(x) = \max(0, x)$
- hyperbolic tangens function, $f(x) = \tanh x$
- logistic sigmoid function, $f(x) = 1/(1 + e^{-x})$

Each of the presented algorithms is tested with different parameter settings, such as number of neighbors (k-NN), depth of the trees (GBDT) or activation function of neurons (MLP). The number of tested parameters settings is shown in

Table 2. The best-possible parameter setting was chosen by the resulting R^2 .

Table 2. Number of tested parameter sets for each of the considered algorithms

Algorithm	Number of parameter sets
LR	1
HR	36
SVM	40
k-NN	40
GBDT	9
MLP	27

6.3. Prognosis results

Following, the different databases (DB 1, DB 2, DB 3) were used for training and evaluating or testing the considered regression algorithms. Thereby the implementation is based on the open source python library scikit-learn [15]. The inputs are the following parameter: Material properties of the parts to be joined:

- Ultimate tensile strength
- Sheet thickness
- Type of material (steel, aluminum)

SPR process parameter:

- Die contour (diameter, depth, thorn geometry)
- Rivet type and length
- Rivet head position

Output parameters are the joining results:

- Interlock $u_{l,2}$
- Interlock height $u_{h,2}$
- Min. thickness die-sided part t_r
- Rivet foot diameter d_f
- Max. joining force F_J

Fig. 7 shows the average R^2 for all forecasted output parameters with the considered algorithms trained with the different databases.

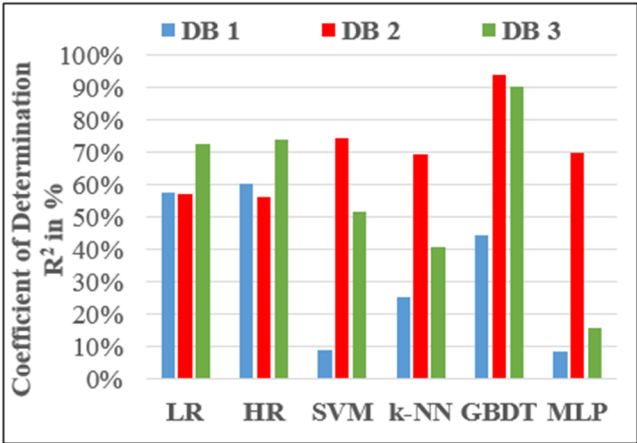


Fig. 7. Comparison of the Coefficient of prognosis R^2 for all considered regression algorithms trained with different databases

When using simpler regression models like LR and HR a positive influence from the use of a larger amount of data is more likely to be small. The best prognosis quality for these two models is achieved by using the filtered numerical database DB 3. With the more complex regression models SVM, k-NN, GBDT and MLP the prognosis quality is quiet low when using only a few datasets (DB 1) and increases significantly using the complete numerical database (DB 2). The most significant influence of the amount of data on the prognosis quality can be observed by using the neuronal network (MLP). Overall, the GBDT algorithm with coefficients of determination R^2 above 90 % provides the best predictive performance for the data sets under consideration.

7. Conclusion

The biggest challenge in using machine learning algorithms in an industrial environment is to provide comprehensive, suitable data that describes the material and process properties well. In this paper a methodology was described how a relatively small experimental database could be extended to a comprehensive database by means of numerical sensitivity analyses. The validation and automation of the simulation models plays a central role in numerical data acquisition. The comparison of different learning algorithms shows the influence of different large data sets on the respective forecast quality, whereby it emerges that the availability of larger valid data sets results in a significant increase of the forecast quality when using more complex models. For the described case of the self-pierce riveting of steel and aluminum sheets, the Gradient Boosted Decision Trees algorithm with the extensive numerical data basis enabled a very good prediction quality. This allows an accurate data-based prediction of joining results for new material combinations that have not yet been considered either experimentally or numerically, thus significantly reducing the effort required for the design of SPR joints.

Acknowledgement

The presented results are part of the research project "Data-based joining parameter prognosis at mechanical joining" of the European Research Association for Sheet Metal Working (EFB) funded by the program for "Industrial Research" (IGF) of the Federal Ministry of Economic Affairs and Energy and the AiF (19853 BR).

References

- [1] Briskham P. Self-pierce riveting of high strength aluminium in thick stack joints. Joining in Car Body Engineering. Bad Nauheim 2015
- [2] Thoms V, Kalich J. Prozessvorhersage beim Stanznieten mit neuronalen Netzen, EFB-Forschungsbericht Nr. 179, EFB, Hannover, 2002
- [3] Hahn O, Tan Y, Voigt H.-M. Maschinelles Lernen zur Vorhersage der Tragfähigkeit von Clinchverbindungen. UTF Science. 2006
- [4] Tan Y, Hahn O, Du F, Voigt H. Non-destructive determination of the loading capacities of clinched DC 04 joints by kNN-Regression. Materials Testing, 2006; 48(5): 211–216.
- [5] Breckweg A. Automatisiertes und prozessüberwachtes Radialclinchen höherfester Blechwerkstoffe. Dissertation. Stuttgart, 2006
- [6] Lambiasi F, Di Ilio A. Optimization of the Clinching Tools by Means of Integrated FE Modeling and Artificial Intelligence Techniques. Procedia CIRP, 2013; 12: 163–168.
- [7] DVS Gemeinschaftsausschuss, Mechanisches Fügen. *DVS/EFB Merkblatt 3410 Stanznieten-Überblick*. DVS-Verlag: Düsseldorf, Germany, 2014.
- [8] McKay M D, Beckman R J, Conover W J. Comparison of three methods for selecting values of input variables in the analysis of output from a computer code, Technometrics, 1979, 21(2): 239-245
- [9] Coppieters S, Jäckel M. et al. Large Strain Flow Curve Identification for sheet metal: Process informed method selection. Forming Technology Forum 2019, September 19-20, Hersching am Ammersee, Germany
- [10] Ronchetti, E M, Huber P J. Robust statistics. Wiley, 2013
- [11] Drucker H, et al. Support vector regression machines. Advances in neural information processing systems. 1997
- [12] Navot, Amir, et al. "Nearest neighbor based feature selection for regression and its application to neural activity." Advances in neural information processing systems. 2006
- [13] Friedman J H. Stochastic gradient boosting. Computational statistics & data analysis 2002, 38.4: 367-378.
- [14] Rosenblatt F. Principles of neurodynamics. perceptrons and the theory of brain mechanisms. No. VG-1196-G-8. Cornell Aeronautical Lab Inc Buffalo NY, 1961.
- [15] https://scikit-learn.org/stable/supervised_learning.html#supervised-learning, December 12, 2019