Comparison of Machine Learning Approaches for Time-series-based Quality Monitoring of Resistance Spot Welding (RSW)

Baifan Zhou, Tim Pychynski, Markus Reischl and Ralf Mikut

Abstract In automatic manufacturing, enormous amounts of data are generated every day. However, labeled production data useful for data analysis is difficult to acquire. Resistance spot welding (RSW), widely applied in automobile production, is a typical automatic manufacturing process with inhomogeneous data structures as well as statistical and systematic dynamics. In resistance spot welding, an electric current flows through electrodes and the materials in between. The materials are first heated and melted, then congeal, forming what is known as a weld nugget, joining the materials together. The nugget size is an important quality indicator, but can only be precisely obtained by using costly destructive methods. This paper strives to address the issue of the scarcity of labeled data by using simulation data generated with a verified finite element model.

Baifan Zhou · Tim Pychynski
Robert Bosch GmbH
Robert-Bosch-Campus 1, 71272 Renningen, Germany
✉ baifan.zhou@de.bosch.com
✉ tim.pychynski@de.bosch.com

Markus Reischl · Ralf Mikut
Karlsruhe Institute of Technology (KIT), Institute for Automation and Applied Informatics (IAI)
Hermann-von-Helmholtz-Platz 1, 76344 Eggenstein-Leopoldshafen, Germany
✉ markus.reischl@kit.edu
✉ ralf.mikut@kit.edu

Archives of Data Science, Series A
(Online First)
KIT Scientific Publishing
Vol. 5, No. 1, 2018

DOI: 10.5445/KSP/1000087327/13
ISSN 2363-9881

Physics-based simulation enables large amounts of labeled data to be generated with fewer limits on sensors and costs. Based on the simulation data, this paper explores and compares multiple machine learning methods, predicts the nugget size with a high degree of accuracy, and conducts an analysis of the influence of feature number and amount of training data on prediction accuracy.

1 Introduction

Technological advances in computing, connecting and sensoring enables a new wave of productivity improvement. We are currently in the era of Industry 4.0, a term first coined by Kagermann et al (2011), which represents the fourth technological revolution. In manufacturing, data mining, as one of the key technologies of Industry 4.0 (Rüßmann et al, 2015), has created new intelligent tools for automatic information extraction and knowledge discovery (Wang, 2007; Nagorny et al, 2017). Automated analysis of sensor data from industrial environments is important for practices such as identifying root causes of exceptions or making correct decisions (Zhu et al, 2011).

Although enormous amounts of data are generated every day in automatic manufacturing, the acquired data is not always relevant (such as missing labels) and thus suitable for data analysis (Wuest et al, 2016). *Resistance Spot Welding* (*RSW*) is a typical fully automated manufacturing process, widely applied in the automobile industry, with 3000 to 6000 welding spots per car chassis. In RSW, the two electrode caps of the welding gun (see Figure 1) press two or three worksheets between the electrodes with force. An electric current then flows from one electrode, through the worksheets, to the other electrode, generating a substantial amount of heat as a result of electric resistance. The materials in a small area between the two worksheets, known as the welding spot, will melt, and form a weld nugget connecting the worksheets. The electrode caps directly touching the worksheets wear out easily due to high thermo-mechanical loads and oxidation, and need to be changed on a regular basis.

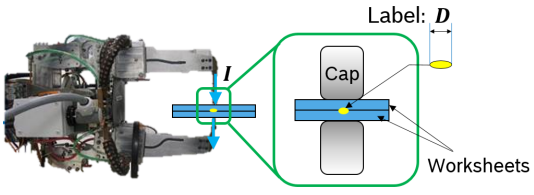


Figure 1: An example of a welding gun (I = electric current, D = welding cross-section diameter).

The quality of the welding spots is typically quantified by the weld nugget diameter as defined in the international standards (ISO, 2004) and German standard (DVS, 2016). The welding quality is then determined as accepted or rejected according to specific tolerance bands. This work strives to predict the spot diameter as a numeric value rather than a classification as accepted or rejected, because the exact diameter values are of great interest for process experts, who can potentially gain better insights by analyzing which input factors would influence the diameter values to what extent. After predicting the numeric values of spot diameters, a classification can still be made according to different welding conditions and user-defined tolerance bands.

Figure 2: Welded worksheets torn apart.

However, the nugget diameter is difficult to measure, and the volume of data with nugget diameter (or labeled data in data science) is low. The common practice is to tear the welded worksheets apart and measure the nugget directly, for example using a micrometer (see Figure 2). Performing quality control in the automobile industry usually requires the destruction of several car chassis, which is extremely expensive.

Many previous studies have adopted the approach of data-driven models. Leveraging the power of data science, data-driven models can assess the nugget diameters based on the recorded process data. These models are beneficial as they can potentially perform quality control for every welding spot reliably, ensuring process capability and reducing costs for quality control. Many authors treated the problem as classification (Martín et al, 2007; Sun et al, 2017), while others evaluated the predicted diameters as a regression problem (Ruisz et al, 2007; El Ouafi et al, 2010; Wan et al, 2016). Their dataset size ranges from approximately 10 (Cho and Rhee, 2004) to around 3000 labeled welding

spots (Boersch et al, 2016). Various methods were explored, including Random Forests (Sumesh et al, 2015), Neural Networks (Afshari et al, 2014), etc. Most of them used process curves, such as electrode displacement (Li et al, 2012), while others also used scalar process parameters, such as coating (Yu, 2015).

Most of the previous studies have mentioned that their data is collected from laboratory experiments. It is questionable whether models developed from laboratory data are applicable in real production, as the welding conditions (such as cooling time and wear) are usually different. Apart from the labeled data amount problem, other difficulties exist in building successful data-driven models. This paper summarizes these difficulties systematically as *three challenges* for data analysis in RSW and in automatic manufacturing:

- *Challenge* 1 is the *limit on labeled data amount*, which results from the difficulty in collecting labeled data in production or even in the laboratory.
- *Challenge* 2 is the *limit on features*. Additional sensors, like electrode displacement, or extra measurements, such as actual worksheet thickness, may be necessary for reliable quality prediction in manufacturing processes. However, more sensors mean higher costs, and an increased risk of expensive machine stops due to sensor failures. When deciding whether to install an extra sensor, its benefit needs to outweigh the disadvantages. It is, therefore, important to understand which sensors or measurements (or features in data science) are necessary and to quantify their benefit to justify the higher costs.
- *Challenge* 3 is the *limit on coverage of relevant situations*. Quality failures are very unusual in manufacturing data, because the welding quality is good under normal conditions, which is the most frequent case in production. Moreover, quality failures may be of different types, making it even more difficult to collect sufficient data for reliable quality prediction.

This work will address the first two challenges. They are interpreted as *three questions* in collecting costly labeled data:

- Question 1: "*How much labeled data to collect first*?" (Challenge 1)
- Question 2: "*Which sensors should be installed first*?" (Challenge 2)
- Question 3: "*Which precision level should the sensors and measurements of welding spot diameters have*?" (Challenge 2)

The authors of this paper suggest the use of physics-based simulation data generated with a verified Finite Element Method (FEM) model (Schif, 2017) to help overcome these three challenges. An established simulation model has *four advantages*:

1. Each simulation run produces one labeled data point. Running a lot of simulations can thus offer a large amount of labeled data.
2. There is almost no limit for obtaining features that are costly or difficult to realize through sensor measurements in laboratory or production.
3. Using astutely designed scenarios, rare situations in production can be studied in detail.
4. There is no measurement sensor precision issue in data generated from a simulation model.

This work gives a short introduction to the FEM simulation to cover some points that the readers may be interested in. The FEM simulation models mechanical effects (elastic-plastic part deformation and thermal expansion), thermal effects (temperatures and heat transfer) and electrical effects (electric current density and electric potential field), taking into account non-linear changes of material and contact properties (such as electrical and thermal contact conductivity). Strong interactions between all three fields result in a very dynamic process behavior and require a multi-field coupled simulation.

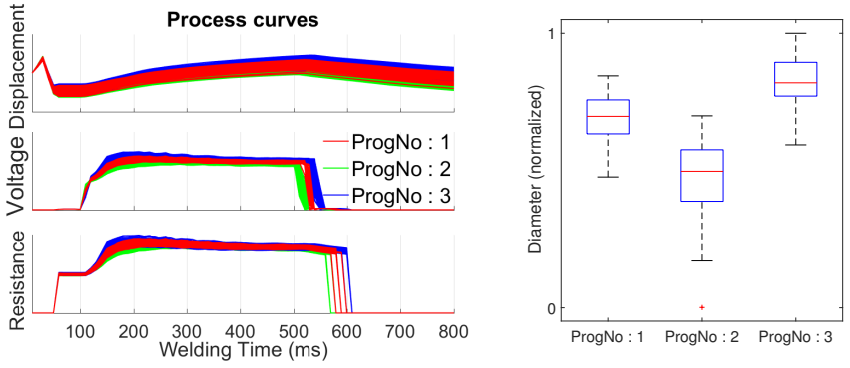
The model has been verified using measurements from lab tests under controlled conditions. Compared quantities were curves of electrical voltage drops, temperatures and electrode displacements measured during welding as well as spot diameters and electrode imprints determined by destructive or optical methods after welding. The main effort of the FEM simulation lies in:

- gathering all relevant information such as worksheet and electrode geometries, temperature-dependent material data, contact properties etc. These features are inputs to the simulation,
- setting up a parametric Finite Element Model in an automated simulation loop, including simulation preparation and data extraction, using commercial and open source tools, such as Python, software packages for the Finite Element Method, etc.

2 Data Description

The most normal conditions in production, i.e. only random variation without spatter or other disturbances, have been selected as the simulation scenario for the data studied in this paper. This simple scenario consists of one welding machine, one type of worksheet pair with identical nominal sheet thickness and material, and three welding programs for three different target spot diameters (see Figure 3). A total of 13,952 welding spots with diameter measurements were simulated. Two types of data exist for each welding spot.

- Time series: Process curves are series of values with time stamps, referred to as time series in data science. Among the extracted 20 time series, process input curves exist, such as electric current I, voltage U, resistance R, and process feedback curves, such as force of the electrode F, displacement s of the electrode, temperature T of certain measurement positions, etc.
- Single features: Numeric values or text strings that remain constant for a welding process. These are nominal and measured geometry or material properties of the caps and worksheets, the number of spots welded, positions of the welding spots, welding programs, etc. The simulation dataset can contain up to 235 single features.



(a) (b)

Figure 3: (a) Process curves for the three welding programs. ProgNo indicates the program numbers. The welding programs prescribe the way the welding process is performed, by specifying the process curves and some additional welding parameters. (b) Boxplot of the diameters for the three welding programs.

3 Methods and Experiments

The approach for exploring answers to the three questions proposed above centers on training machine learning models with different data subsets and comparing their performance:

- *Data splitting into training and test set* in this paper is not the conventional random splitting. In manufacturing, the historical data collected for training may have slightly different statistical properties than the data in the later application phase. As mentioned in Section 1, the electrode caps are regularly changed in RSW. The caps in the collected training dataset will therefore always be different from the caps in the test set. In this paper, 14 caps are simulated in total. Data generated with 9 caps (7973 data points) is used for training, while data from the remaining 5 caps (5979 data points) is used for testing.
- *Data is split into subsets of different training data numbers* for the purpose of answering question 1 ("*How much labeled data to collect at the start*?"). A series of training subsets is built by randomly selecting a different number of data points from the training dataset. To allow direct comparison of the testing results, the test dataset always contains the same 5979 data points. Machine learning methods are applied to build models using these subsets with different sizes of training data, and tested on the same test set (see Table 1).

Table 1: Data splitting to subsets of different training data size.

Set	**# of Data Points**						
Training sets	100	250	500	1000	2000	5000	7973
Test sets	5979	5979	5979	5979	5979	5979	5979

- *Data is split into subsets of different features* for the purpose of answering question 2 ("*Which features are important*?"). All available features, i.e. time series features as well as single features, are divided into four subsets as in the following enumeration. By comparing the performance of machine learning models trained using the four different feature subsets, it is possible to estimate the importance of these features.

- Production Set: Features that are always available in production, such as current, voltage, resistance (15 single features and 4 time series).
- LabLowCost Set: Production Set + features available in laboratory with relatively low cost (16 single features and 10 time series).
- LabHighCost Set: LabLowCost Set + features available in laboratory with higher cost. (29 single features and 14 time series)
- Complete Set: LabHighCost Set + other features that are difficult to realize or extremely costly (235 single features and 20 time series).

Combinations of possibilities of training data number and feature set result in $7 \times 4 = 28$ subsets.

- *Extra data sets with manually generated noise* are for the purpose of answering question 3 ("*Which precision level should the sensors have*?"). Adding noise to the aforementioned non-noisy 28 subsets results in another corresponding 28 subsets with noise. The noise levels are a best engineering guess derived from discussions with process and measurement experts.
  - Sensor data (time series plus single features): Gaussian noise with 2 % standard deviation (in the time series the noise is added for every single sample point).
  - Spot diameter measurement: Gaussian noise with 0.1 mm standard deviation.

In total, $7 \times 4 \times 2 = 56$ subsets are used to train and test machine learning models and compare their performance. The *workflow of experiments* on these subsets is standard (Mikut et al, 2007) and is displayed in Figure 4.

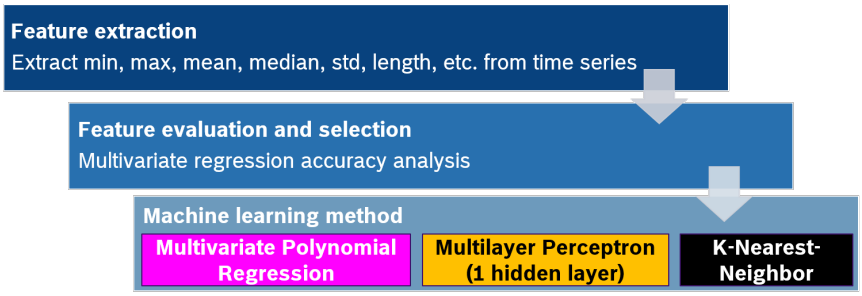


Figure 4: The workflow of data analysis in this paper.

First, some simple features are extracted from the time series. The extracted features are minimum, maximum, minimum position, maximum position, mean, median, standard deviation, length. We decided to extract only these features to obtain interpretable features. These extracted features combined with all other single features are then evaluated and selected using step-forward selection, starting with the evaluation of each single feature, and incrementally adding more features. Features that result in models with the best regression accuracy (RMSE) for the prediction of spot diameters are selected. After feature selection, three machine learning methods (*polynomial regression*, *neural networks* and *k-nearest neighbors*) are applied to build different data-driven models. These models are implemented using the MATLAB toolbox SciXMiner developed by Mikut et al (2017).

The feature extraction and machine learning models are intentionally kept simple, for better generalization and automation in other similar applications in automatic manufacturing. The characteristics of the three machine learning methods and the reasons for choosing them are explained below.

- *Polynomial regression* is an established method even in traditional engineering and statistics. The solving of polynomial regression is deterministic. This explains why the performance of polynomial regression (see Figure 8) is rather stable, after a certain amount of training data is fed into the model. The features selected by polynomial regression and the model formula provide transparency for understanding with engineering knowledge.
- *Neural networks* are a representative non-linear method. They have potentially unlimited power to approximate any functions, thus they are usually powerful enough to solve any supervised learning problem if adequate data is available. There are, however, two drawbacks. Neural networks are difficult to interpret, and they often require a large amount of training data. After a relatively large amount of data has been used for training, it can be seen in the results that the performance of neural networks (see Figure 9) even exceeds that of the polynomial model.
- The most substantial characteristic of *k-nearest neighbors* is that it is non-parametric, i.e. no parameters are learned through training. Instead, all training data is stored in the model. The process of prediction is merely

to use an average of similar points measured by a pre-defined distance in the training data. K-nearest neighbors would work well if the local structure of the data were linear, and becomes time-consuming if the training dataset is too large.

The hyper-parameters of these models are chosen with 5-fold cross validation with exemplary training datasets. For all machine learning models trained on subsets without noise, the hyper-parameters are chosen with the non-noisy subset of all training data (with 7973 training data points) and the Production Feature Set. For all machine learning models trained with subsets with noise, this is performed again with the noisy subset of all training data (with 7973 training data points) and the Production Feature Set.

The hyper-parameter selection results in a first order polynomial regression model without interaction terms (namely a linear regression model, referred to as Polynomial1), a multi-layer perceptron with one hidden layer of 16 neurons (referred to as MLP16), and a k-nearest neighbors model (referred as KNN3). Details are listed in Table 2.

Table 2: Machine learning models and hyper-parameters selected through 5-fold cross validation.

Machine Learning Model	**Hyper-parameter**	**Value**
Multivariate polynomial regression	Order of the polynomial	1
Multi-layer perceptron (one hidden layer)	Number of neurons in the hidden layer	16
K-nearest neighbor	k	3

The experiments show that many features are highly correlated and cause problems in feature selection due to random effects. For example, some features are selected for better performance in the feature selection stage, but could show better or worse performance in a later stage of performance comparison, or even in another repetition of training and testing. To circumvent this problem, for training machine learning models on subsets with more features, the selected features from smaller feature sets will always be used, and the models have the opportunity to add more features. It is worthy to note this issue of highly correlated features, for this phenomenon may be prevalent in manufacturing data analysis.

4 Results

To answer the three questions proposed in Section 1, it is necessary to compare the performance of the various models built with different methods and trained with different subsets. In order to compare the performance, an appropriate measure is important. Various performance measures were proposed in the literature. This section first discusses the various performance measures, then uses the selected measure to compare the models and provides answers to the proposed questions. Diameter prediction is treated as a regression problem in this work, as process experts can obtain more information from a predicted diameter than a mere good/bad classification.

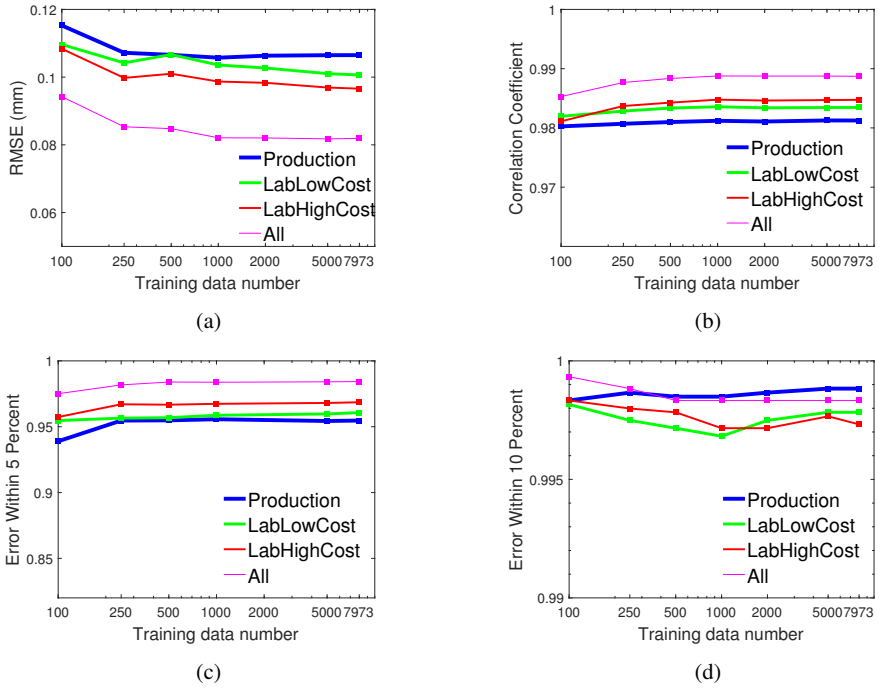


Figure 5: Comparison of performance measures by testing Polynomial1 on non-noisy datasets.

RMSE (or *MSE*; see Figure 5 a) stands for root mean squared error (or mean squared error) between the reference diameters and predicted diameters. It is a classic performance measure for regression tasks, and has been used

in Martín et al (2009) and Boersch et al (2016), etc. RMSE is also used in this paper for feature selection. The disadvantages of RMSE are that it is scale-dependent (Hyndman and Koehler, 2006), and not particularly intuitive compared to other measures.

Correlation Coefficient (or *R-squared*; (see Figure 5 b), used in Muhammad and Manurung (2012), is good for judging the models' ability to predict the general trend of the training data. Since one of the goals of diameter prediction is to minimize the number of product failures, only knowing the trend is not sufficient for this goal.

The most intuitive measures for process experts are perhaps User-Defined Errors (UDE, Boersch et al (2016); Figure 5 c, d). These are more suitable for helping to set the factor of safety, thus closer to the goal of minimizing product failures. Two User Defined Errors are adopted in this paper: *Error Within* 5 % and *Error Within* 10 %. Error Within 5 % means the percentage of predictions with relative errors smaller than 5 % of the reference value. Error Within 10 % is likewise.

Comparing models trained on different feature sets (see Figure 5) shows that, as expected, when more features are available in training, the models also have better performance in testing: All › LabHighCost › LabLowCost › Production.

Comparing models trained on different training data subsets shows that, when the training data number is greater than 250, the performance improves insignificantly as the training data number further increases.

From results of testing the first-order polynomial model (see Figure 5), trained on different numbers of training data and features, it can be seen that RMSE, Correlation Coefficient and Error Within 5 % have similar trends. This is also the case for the multi-layer perceptron model and k-nearest neighbor model. Error Within 10 % cannot differentiate the models trained with different data number and feature sets, since they always show very good results. Therefore, in the following discussions, only Error Within 5 % will be compared.

Comparing models built using different machine learning methods can be made by comparing Figure 5 c, Figure 6, and 7, or in Figure 8. The comparison indicates the performance of the models built with three machine learning methods is: Polynomial ≈ MLP > KNN. Figure 6 shows that MLP16 tends to overfit for the Feature Set LabLowCost and LabHighCost. In this regard, Polynomial1 is more advantageous.

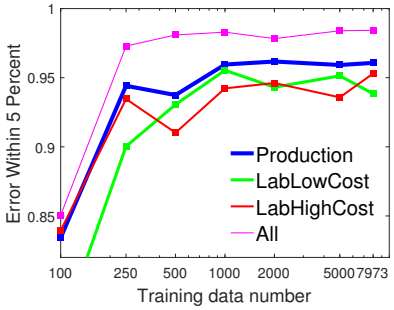


Figure 6: Testing results MLP16 on non-noisy datasets.

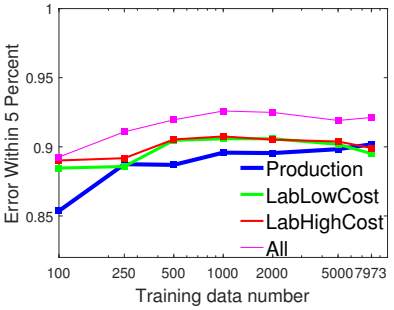


Figure 7: Testing results KNN3 on non-noisy datasets.

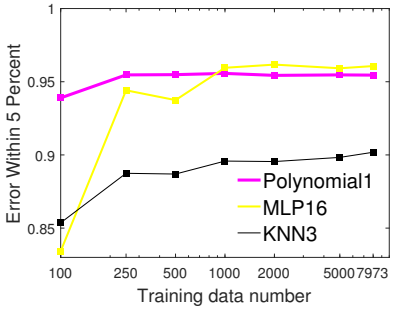


Figure 8: Testing models trained on non-noisy Production Set.

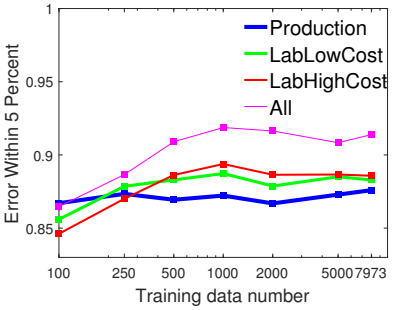


Figure 9: Testing Polynomial1 on noisy datasets.

Comparing models trained with non-noisy data and noisy data can be made by contrasting Figure 5 c and Figure 9. The comparison suggests the performance deteriorates strongly with noisy data. Generally speaking, the difference between models trained on these three feature subsets, Production, LabLowCost, and LabHighCost is insignificant.

5 Conclusion and Outlook

Simulation data can provide large amounts of economically obtained labeled data and features that would be otherwise unavailable. This data contains no or low measurement error. As a result, it is possible to hold a collection of valuable

data from production or laboratory. Reviewing the starting point of the three questions, the following conclusions are made based on the datasets generated by a simulation model in the selected simple scenario (see Section 2), serving as a guidance for a first plan of data collection:

- Answer to Question 1: Starting with a collection of 250 labeled data points.
- Answer to Question 2: Sensors available in Production can already be useful.
- Answer to Question 3: Measurement uncertainty can lead to significant deterioration of the prediction performance. In our case, precision better than Gaussian errors of 2 % standard deviation for sensors and 0.1 mm standard deviation for diameter measurements respectively is recommended.

It is important to note that the conclusions regarding the necessary number of data points and features are highly dependent on the dataset, and the simulation scenario from which the data is collected. It is therefore safe to say that the results can be generalized for the specific situation simulated. They can serve as a first guidance for data collection, but should not be assumed as valid in general. Since the selected simulation scenario of one welding machine and three welding programs tends to be oversimplified, these conclusions may not hold for other scenarios or for complex production data. The optimal data collection plan may vary depending on the complexity of the actual situations compared to the simulated situation.

In future research, several topics can be explored. We have plans for combining the simulated data and laboratory or production data.

- The first option is to use laboratory data or production data to further improve and fine-tune the simulation model so that realistic physical effects and authentic variance of influencing factors can be rebuilt in the simulation. This realistic simulation data can be used for further data analysis.
- After the first point has been realized, a further step is to apply the model trained on simulated data, and test on the laboratory or production data to see the extent to which the trained model can be generalized.

- A third option is to apply transfer learning, i.e. to pre-train a model on simulated data, then to fine-tune the model on laboratory data or production data, and to see how much improvement in performance is gained.
- The improved simulation model and generated data can be used for production condition evaluation before a real production process begins, so that the influence of change of welding parameters, such as welding gun properties, material properties, distance between the worksheets, can be systematically evaluated.
- We will also analyze alternative feature extraction methods (such as features based on Fourier Transformation, Wavelet Transformation or Principal Component Analysis) to evaluate the difference of prediction accuracy using different feature extraction methods.
- As actual production conditions may change due to a change in supplier, new engineering design, etc., it is always questionable to what extent a model trained on historic or lab data can be applied to production data in the future, which constitutes a general problem in manufacturing. This is also called model monitoring and remains an open question.

This paper tries to use data analysis as a guidance for data collection. The iterative process data-collection → data analysis → data-analysis guided data collection → data analysis is an attempt to combine data collection and analysis. A similar practice would be particularly necessary in fields where a limit on the amount of labeled data, features, and covering situations is a problem of interest.

Acknowledgements The authors would like to thank Simon Waczowicz for his help using the SciXMiner toolbox, Alexander Schif for providing the data, Friedhelm Günter, Florian Schmid, Martin Dieterle for their valuable input on improving the quality of the paper, and Katherine Quinlan-Flatter for proofreading this article.

References

Afshari D, Sedighi M, Karimi MR, Barsoum Z (2014) Prediction of the Nugget Size in Resistance Spot Welding with a Combination of a Finite-element Analysis and an Artificial Neural Network. Materiali in tehnologije 48(1):33–38. DOI: 10.26634/jms.3.1.3366.

Boersch I, Füssel U, Gresch C, Großmann C, Hoffmann B (2016) Data Mining in Resistance Spot Welding. The International Journal of Advanced Manufacturing Technology, pp. 1–15, Springer. DOI: 10.1007/s00170-016-9847-y.

Cho Y, Rhee S (2004) Quality Estimation of Resistance Spot Welding by Using Pattern Recognition with Neural Networks. IEEE Transactions on Instrumentation and Measurement 53(2):330–334. DOI: 10.1109/TIM.2003.822713.

DVS (2016) DVS 2902–3: Widerstandspunktschweißen von Stählen bis 3 mm Einzeldicke – Konstruktion und Berechnung. Standard, Deutscher Verband für Schweißen und verwandte Verfahren e.V., Düsseldorf (Germany).

El Ouafi A, Bélanger R, Méthot JF (2010) An On-line ANN-based Approach for Quality Estimation in Resistance Spot Welding. In: Advanced Materials Research, Vol. 112, pp. 141–148. DOI: 10.4028/www.scientific.net/AMR.112.141.

Hyndman RJ, Koehler AB (2006) Another Look at Measures of Forecast Accuracy. International Journal of Forecasting 22(4):679–688. DOI: 10.1016/j.ijforecast.2006.03.001.

ISO (2004) ISO 14327:2004: Resistance Welding – Procedures for Determining the Weldability Lobe for Resistance Spot, Projection and Seam Welding. Standard, International Organization for Standardization, Geneva (Switzerland).

Kagermann H, Lukas WD, Wahlster W (2011) Industrie 4.0: Mit dem Internet der Dinge auf dem Weg zur 4. industriellen Revolution. VDI Nachrichten 13.

Li Y, Zhao W, Xue H, Ding J (2012) Defect Recognition of Resistance Spot Welding based on Artificial Neural Network. Advances in Intelligent and Soft Computing 115(2):423–430. DOI: 10.1007/978-3-642-25349-2_56.

Martín Ó, López M, Martin F (2007) Artificial Neural Networks for Quality Control by Ultrasonic Testing in Resistance Spot Welding. Journal of Materials Processing Technology 183(2-3):226–233, Elsevier. DOI: 10.1016/j.jmatprotec.2006.10.011.

Martín O, Tiedra PD, López M, San-Juan M, García C, Martín F, Blanco Y (2009) Quality Prediction of Resistance Spot Welding Joints of 304 Austenitic Stainless Steel. Materials and Design 30(1):68–77. DOI: 10.1016/j.matdes.2008.04.050.

Mikut R, Burmeister O, Grube M, Reischl M, Bretthauer G (2007) Interaktive Auswertung von aufgezeichneten Zeitreihen für Fehlerdiagnosen und Mensch-Maschine-Interfaces. atp - Automatisierungstechnische Praxis 49(8):30–34. ISBN: 978-3-180919-80-5.

Mikut R, Bartschat A, Doneit W, Ordiano JÁG, Schott B, Stegmaier J, Waczowicz S, Reischl M (2017) The MATLAB Toolbox SciXMiner: User's Manual and Programmer's Guide. arXiv:1704.03298.

Muhammad N, Manurung YHP (2012) Design Parameters Selection and Optimization of Weld Zone Development in Resistance Spot Welding. In: Proceedings of World Academy of Science, Engineering and Technology, 71, p. 1220.

Nagorny K, Lima-Monteiro P, Barata J, Colombo AW (2017) Big Data Analysis in Smart Manufacturing: A Review. International Journal of Communications, Network and System Sciences 10(03):31, Scientific Research Publishing. DOI: 10.4236/ijcns.2017.103003.

Ruisz J, Biber J, Loipetsberger M (2007) Quality Evaluation in Resistance Spot Welding by Analysing the Weld Fingerprint on Metal Bands by Computer Vision. International Journal of Advanced Manufacturing Technology 33(9-10):952–960. DOI: 10.1007/s00170-006-0522-6.

Rüßmann M, Lorenz M, Gerbert P, Waldner M, Justus J, Engel P, Harnisch M (2015) Industry 4.0: The Future of Productivity and Growth in Manufacturing Industries. Boston Consulting Group.

Schif A (2017) Examination of Machine Learning Methods for the Evaluation of Resistance Spot Welds based on Simulated Process Scenarios. Master's thesis, University of Stuttgart.

Sumesh A, Rameshkumar K, Mohandas K, Babu RS (2015) Use of Machine Learning Algorithms for Weld Quality Monitoring using Acoustic Signature. Procedia Computer Science 50:316–322. DOI: 10.1016/j.procs.2015.04.042.

Sun H, Yang J, Wang L (2017) Resistance Spot Welding Quality Identification with Particle Swarm Optimization and a Kernel Extreme Learning Machine Model. The International Journal of Advanced Manufacturing Technology 91(5-8):1879–1887, Springer. DOI: 10.1007/s00170-016-9944-y.

Wan X, Wang Y, Zhao D (2016) Quality Monitoring based on Dynamic Resistance and Principal Component Analysis in Small Scale Resistance Spot Welding Process. The International Journal of Advanced Manufacturing Technology 86(9-12):3443–3451, Springer. DOI: 10.1007/s00170-016-8374-1.

Wang K (2007) Applying Data Mining to Manufacturing: The Nature and Implications. Journal of Intelligent Manufacturing 18(4):487–495. DOI: 10.1007/s10845-007-0053-5.

Wuest T, Weimer D, Irgens C, Thoben KD (2016) Machine Learning in Manufacturing: Advantages, Challenges, and Applications. Production & Manufacturing Research 4(1):23–45. DOI: 10.1080/21693277.2016.1192517.

Yu J (2015) Quality Estimation of Resistance Spot Weld based on Logistic Regression Analysis of Welding Power Signal. International Journal of Precision Engineering and Manufacturing 16(13):2655–2663, Springer. DOI: 10.1007/s12541-015-0340-6.

Zhu S, Kong L, Chen L (2011) Data Mining of Sensor Monitoring Time Series and Knowledge Discovery. In: Proceedings of the 2nd International Conference on Artificial Intelligence, Management Science and Electronic Commerce (AIMSEC), pp. 2914–2917. DOI: 10.1109/AIMSEC.2011.6010471.