2nd Conference on Production Systems and Logistics

Feature Engineering for a Cross-process Quality Prediction of an End-of-line Hydraulic Leakage Test using an Experiment Sample

Christian Neunzig[1,2], Simon Fahle[1], Bernd Kuhlenkötter[1], Matthias Möller[2]

[1]*Chair of Production Systems, Ruhr-University Bochum , Universitätsstraße 150 , 44801 Bochum, Germany*
[2]*Bosch Rexroth AG, Bexbacher Straße 72, 66424 Homburg, Germany*

Abstract

The increasing availability of manufacturing data and advanced analysis tools are forcing the demand for data-driven approaches to improve the quality of workpieces and the efficiency of manufacturing processes. The analysis of real manufacturing data is challenging due to frequent changes in production circumstances. In this work, machine learning methods based on the data along the value chain of hydraulic valves are used to predict the leakage results during end-of-line testing. The leakage volume flow measurement results are very sensitive to changes in gap geometry and temperature level in the measurement cross-section. Additional measurements and experiments are required to interpret the systematic influences of the input data on the target variable and to introduce the missing information into the dataset. The design of a meta-model using experiment data supports the identification of statistical patterns to be applied to the real production dataset as a feature. This paper presents a systematic approach to hand-crafted feature engineering that improves the quality prediction of end-of-line hydraulic leakage testing.

Keywords

Feature Engineering; Hydraulics; Machine Learning; Quality Control

1. Introduction

In the last 20 years, the use of computational fluid dynamics (CFD) has produced remarkable advances and developments for optimizing geometries and understanding hydraulic phenomena [1], [2]. The application of machine learning is motivated by tremendous success in optimizing the product and process quality of industrial goods. Machine learning techniques support the continuous improvement process to reduce manufacturing costs by avoiding non-value-added production steps and to ensure product quality. [3], [4] Product quality include measures of the inner work piece state such as an internal leakage [5], [6]. The manufacturing of hydraulic control valves for agricultural peripherals consists of the machining process, assembly, and end-of-line inspection of 100% of the workpieces. In this use-case, the hydraulic testing is the most costly production step and the bottleneck in the manufacturing.

To relieve the test bench capacity, machine learning techniques are used to predict the end-of-line test results of the internal hydraulic leakage based on the cross-process data of the different series production steps. The internal leakage of the valve is intrinsic as the spool trembles at defined positions within the body to ensure the functionality of the valve, but it must be minimized due to the loss of performance [7]. The measurement results of the internal leakage are considered as labels of a supervised regression problem.

Exploiting existing manufacturing data to approximate the physical relationships of input and target variables is a promising strategy to add value from the necessary data management [3]. The requirement for minimum

production cost inhibits data collection of significant thermal and geometric information. The naïve approach of measuring component and fluid temperature is neither economically nor technically feasible for two reasons. First, simultaneous temperature measurement of the relevant components in the closed body of the valve is not possible without altering the physical bodies. Second, measuring the temperature of the fluid alters the fluid dynamics within the part. The sophisticated testing set-up requires the representation of the temperature and the gap geometry in the cross-section by other quantities. Extending the features through a systematic investigation with additional measurements, similar to experimental design methods, can provide significant benefits by adding knowledge to the predictive model [8]. Due to the high effort and lack of machine availability, the amount of experimental data is small. [10] The concept of meta-learning provides a solution to overcome the limited amount of labeled data. The use of multiple source tasks for prediction causes a beneficial initialization for learning on the target task [9]. A meta-regression model is developed to investigate the relationship between the gap geometry and the representative temperature variables based on an experimental dataset.

The paper begins with an overview of related work in meta-learning and applied machine learning and continues with an introduction to the use-case. This is followed by the design of an experiment dataset that acts as a meta-training dataset. A meta-regression model derives the coefficients c_i that best fit the meta-linear regression model. The resulting feature, called the t_{factor}, is applied to the production dataset to assess its predictive power and the functionality of the procedure. The paper concludes with a conclusion section and suggestions for further work.

2. Related work

2.1 Meta-Learning

Learning from a few examples is an essential aspect of human intelligence. One possibility that enables gain solutions to complex tasks from just a few examples is to use experience to learn a prior about tasks. [24] Meta-learning is an approach to learning from imbalanced datasets that are common in the real world. The challenge is to learn an accurate meta-model when a dominant task set has only a small number of training examples. [17] Meta-learning can be divided into mixed linear regression (MLR) and multi-task learning (MTL). The number of samples, the dimension of data points, and the number of clusters are defined as n, d, and k, respectively. In MLR, the sample and time complexity are either polynomially dependent on k or depends on the inverse of the k^{th} singular value of some moment matrix. In contrast, in MTL, all tasks have the same number of examples and performance is evaluated on the observed tasks used in training. [10] There are two perspectives on approaching meta-learning: optimization based ([18], [19], [20], [21]) and probabilistic ([22], [23], [24], [25]).

According to KONG ET AL. the formulation of a standard meta-training for supervised learning, motivated by a probabilistic view captured in formulas (1)–(4), assumes a given collection of n meta-training tasks $\{\boldsymbol{\mathcal{T}}_i\}_{i=1}^{\mathrm{n}}$ derived from some distribution P($\boldsymbol{\mathcal{T}}$). [10] Each task $\boldsymbol{\mathcal{T}_i}$ is associated with a dataset of size t_i that is defined as the meta-training dataset in (1):

$$\boldsymbol{\mathcal{D}}_{meta-train} = \left\{\left\{\left(\boldsymbol{x}_{i,j}, y_{i,j}\right) \in \mathbb{R}^d \times \mathbb{R}\right\}_{j\in[t_i]}\right\}_{i\in[n]} \quad (1)$$

Scooping some structural patterns in P($\boldsymbol{\mathcal{T}}$), the goal is to train a model for a new task $\boldsymbol{\mathcal{T}^{new}}$ that is part of P($\boldsymbol{\mathcal{T}}$), from a small amount of training dataset in (2)

$$\boldsymbol{\mathcal{D}} = \left\{\left(\boldsymbol{x}_j^{new}, y_j^{new}\right)\right\}_{j\in[\tau]} \quad (2)$$

Each task $\boldsymbol{\mathcal{T}_i}$ is associated with a model parameter c_i, where the meta-training data is taken from (3):

$$\left(\boldsymbol{x}_{i,j}, y_{i,j}\right) \sim P_{c_i}(y|\boldsymbol{x})P(\boldsymbol{x}) \text{ for all } j \in [t_i] \quad (3)$$

The prior distribution of the tasks and the model parameters are completely characterized by a meta-parameter θ that (4):

$$c_i \sim \mathrm{P}_\theta(c) \quad (4)$$

However, task ambiguity is a critical problem to deal with, if the goal of meta-learning is to learn solutions to new tasks from small datasets. If each task has only one sample, the problem is explored. Even with the best possible prior, there may simply not be enough information in the examples for a new task to solve that task with a high degree of confidence. Therefore, it is useful to develop meta-learning methods that can suggest multiple possible solutions to an ambiguous learning problem. Such a method can be used to evaluate uncertainty, perform active learning, or obtain direct human control over which pattern to prefer. [24] In some cases, the meta model cannot achieve vanishing error when there is noise [27].

2.2 Applied Machine Learning in Manufacturing

Machine learning methods can be distinguished by their learning strategy (e.g., supervised, unsupervised), prediction task (e.g., regression, classification), and structure (e.g., shallow, deep) [16]. The number of data points per input dimension into the model and the model complexity are strongly interrelated. For regression problems in manufacturing data with a ratio of data points to input dimensions smaller than 1000, WEICHERT ET AL. recommend the use of the following groups of machine learning algorithms with descending suitability: decision tree ensembles, SVM, linear regression, and artificial neural networks. For the selection of the suitable model, the relationship between the process complexity, the amount of data stored, and the model complexity is significant. If this is not taken into account, complex models that have been trained on small amounts of data can be trained with an increased risk of overfitting and/or lack of interpretability. [3] For this reason, the Extreme Gradient Boosting (XGBoost) is considered as a widely used and state-of-the-art representative of gradient boosting ensemble trees. [28], [29] In the sequential boosting process illustrated in Figure 1, each decision tree depends on the result of the previous tree to produce an improved prediction. The boosting algorithm is built incrementally by correcting the errors of the previous weak models. [30]

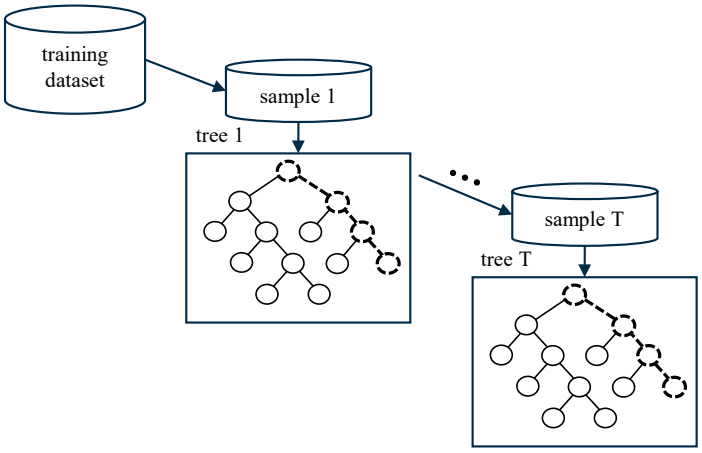

Figure 1: Boosting algorithm principle; cf. [29]

3. Use-Case: Hydraulic Leakage Flow of Directional Valves

The considered technical system is a control valve under test conditions. The electro-hydraulically actuated directional valve has four significant spool positions. The spool is actuated by a 4/3 directional pilot control valve which pressurizes the outer chambers to move the spool. The position of the control spool is fed back to the on-board electronics by means of a displacement transducer. The electronics control the position of

the spool depending on the requested set point using the current from the pilot valve. [11], [12] The desired flow is set at the consumer ports to operate the connected equipment, which is simulated by a test bench. The spool is characterized by various diameters in order to shut off different pressure spaces in the four defined spool positions. In general, the cross-section under consideration is simplified as an arrangement of an eccentric spool in a housing bore with the absolute eccentricity e of the spool, as shown in Figure 2.

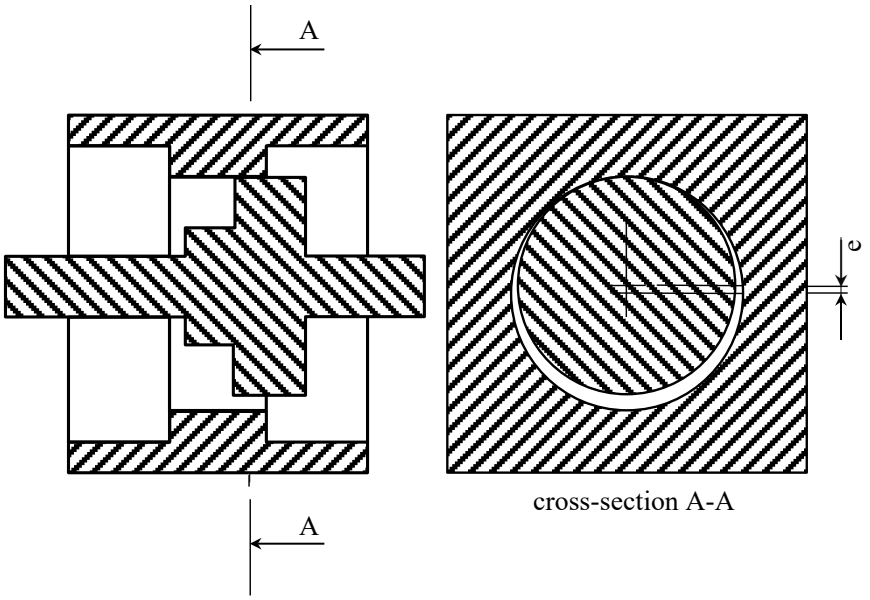

Figure 2: cross-section of an eccentric spool in a bore

Physically inevitable, the leakage flow between spool and housing bore takes the flow direction from the higher to the lower pressure chamber and is largely determined by the gap geometry. The gap width l is defined as the axial overlap of the spool and the bore. The clearance δ is defined as half the diameter difference of the spool and the bore in (7). The relative eccentricity ε is the ratio of the absolute eccentricity e and clearance δ, defined in (8). [12], [13]

$$\delta = \frac{D_{bore} - D_{spool}}{2} \tag{5}$$

$$\varepsilon = \frac{e}{\delta} \tag{6}$$

Under laminar, isothermal flow conditions and a large ratio of gap width l to gap height δ (0.001–0.02 mm) the leakage flow $q_{leakage,max}$ maximizes with the maximum of the eccentricity (ε=1) to 2.5 times the concentric arrangement and approximately follows the equation (7).

$$q_{leakage,max} = \frac{\Delta p \times \pi \times d_m \times \delta^3}{12 \times \eta \times l} \times (1 + 1.5 \times \varepsilon^2) \tag{7}$$

The quantitative calculation of the exact leakage flow is significantly dependent on the real-time viscosity of the fluid at the time of measurement. The adiabatic change of state cannot be assumed since the mass flow rate of the fluid is minor compared to the mass of the solids, so that the fluid temperature is supposed to converge with the temperature of the solids. Additionally, the different coefficients of thermal expansion sensitively affect the annular gap dimension during operation. Consequently, the temperature level of the environment and the connected components has an impact on the leakage measurement. [12], [13]

4. Meta-Regression Model

The underlying problem is to differentiate the impact caused by the gap geometry and by the thermal conditions of the valve in the test bench. A multilinear dependence of the variables on the temperature level is assumed. The goal of designing a meta-regression model is a hand-crafted feature that represents the temperature level of the technical system based on a given range of clearances δ expected to appear mostly in the production dataset and time variables. This approach is intended to eliminate or at least reduce the predictive influence of the clearance on the internal leakage flow, as the clearance is not included in the production dataset.

The test process has a temperature dependent impact. The directional valve is installed between two plates of the testing bench, which is implemented as a rotary indexing table with four workstations. The test program performs two relevant sets of automated testing steps at two successive workstations. The first relevant workstation is characterized by testing steps with high volume flows thus introducing high energy impact into the test part, the mounting plates and pipes. In contrast, testing steps with low volume flows are performed at the second relevant workstation including the label measurement under consideration in the neutral position. At both relevant workstations, a repetition or partial run of a test set is possible, resulting in different temperature scenarios.

4.1 Meta-Training Dataset

The meta-training dataset $\mathcal{D}$ of the directional control valve is provided by Bosch Rexroth AG for this investigation and consists of the shape and geometry data of the components, the pairing data and the metric sensor values measured on the test equipment of the series production. The information from about 20 sensor signals ranges from pressure, current, volume flow and position at around 70 different test steps. In sum, this results in approximately 1000 dimensions per test series. The following procedure is performed for 16 directional control valves. The sample size is limited to 16 at the time of the study due to the significant effort required for geometric measurement, manual bench assembly, and off-series hydraulic testing. Due to varying test durations, the target measurement is recorded with varying frequency for the same temperature scenarios, resulting in a total of 248 test series. In total, a data matrix with 248 rows of 1000 columns is used as the meta-training dataset. Due to a non-disclosure agreement, the dataset cannot be made available to the public.

To represent warm-up and cool-down phases in the system, two time-variables are inserted into the meta-training and production datasets. The warm-up phase is introduced as the runtime sum of the first functional test (FP) workstation with high energy load, while the cool-down phase is defined as the transition time between the end of the first and the beginning of the second relevant workstation with low energy load, including the rest time. In Table 1, the design of the experiment is described by the default setting and variation of the considered variables. The default setting of performing the first set of test steps only once is changed to two and three times to cover different temperature scenarios. The leakage measurement is performed several times for the same heating scenario, as the test part cools down during the measurement.

Table 1: Experiment structure

variable	standard setting	variation
clearance	tolerance centre	± 1.5 µm
number of testing runs	1	1, 2, 3
transition_time	23 sec	number of testing runs
fp_runtime_sum	continuous	number of testing runs

4.2 Meta-learning for Mixed Linear Regression

Following the recommendation of KONG ET AL., a meta-learning for mixed linear regression is developed with the time variables, the ring gap, and the resulting t_{factor} formulated in (8) [10]. The meta-dataset is split into a train and a test split with a split ratio of 85/15. In addition, a five-fold cross-validation is performed to obtain more reliable regression results. To apply this to a sample size of 16, the train dataset must contain 14 valves, while the validation and unseen test datasets must each contain one valve. The selection of valves for each dataset predetermines the result for the coefficients. Therefore, different combinations are iterated for the choice. To increase the repeatability of the results, an equal distribution of the target variables for the training, validation and test datasets is required. The Kolmogorov-Smirnov two-sample test checks for equal distribution for the datasets [14], [15]. Only 42 of the 120 possible combinations pass the Kolmogorov-Smirnov two-sample test with a p-value of 5% and thus only those 42 are considered.

$$t_{factor} = c_1 \times t_{transition} + c_2 \times t_{fp_{runtime}} + c_3 \times \delta^n \quad (8)$$

The median is applied to the probability density of the linear coefficients of the meta-regression model for exponent n = 1 and plotted as a vertical line in Figure 3 to include the most common ratio of c1/c2 for a given annular gap in the production dataset. Considering the heat direction, the transition time for heat dissipation consequently receives a negative sign and the runtime sum for heat input receives a positive sign.

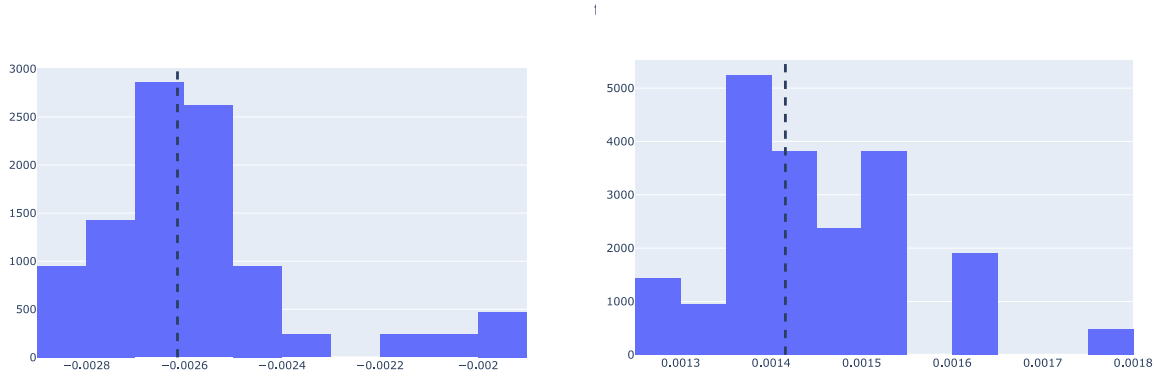

Figure 3: probability density of the linear coefficients

Since the clearance enters the leakage flow formula (7) polynomially and has the highest relevance, the medians from the distributions of the coefficients for different exponents of the clearance are examined and presented in Table 2.

Table 2: Overview of linear and polynomial regression coefficients

	linear	polynomial	
n (clearance)	1	3	50
$t_{transition}$	**-0.002614**	**-0.002580**	**-0,002458**
$t_{fp\_runtime\_sum}$	0.001416	0,001420	0,001182
δ	6.5	0,026222	2,5930

From the overview, it can be concluded that the polynomial regression approach on this meta-training dataset does not produce a significant change in the regression coefficients for the time variables. Since the Gini importance for the transition time in the production dataset is significantly higher compared for the runtime

sum, and the regression coefficient for the transition time remains nearly constant, polynomial regression does not produce an improvement in this case.

Since clearance does not occur in the manufacturing dataset, the ratio of the time variables is kept in the hand-crafted feature and follows formula (9) with the regression coefficients of the linear regression.

$$t_{factor} = -0.002614 \times t_{transition} + 0.001416 \times t_{fp_{runtime}} \quad (9)$$

5. Performance Evaluation with Real Manufacturing Data

The prediction model for forecasting leakage values based on series manufacturing data does not contain any geometry data for the bore under consideration. The geometric pairing of the components differs in the manufacturing tolerances of the housing bore and the control supplied. The daily deviation of the internal leakage measurement is caused, among other things, by systematic batch effects of the machining process. The pairing of spool and main housing bore is not cost-effective, so the internal leakage flow is checked as process validation.

First, the manufacturing dataset and the baseline comparison model (A) are described. The enhanced model also includes the time variables (B) and additionally the hand-crafted feature (C). The models are developed in an Anaconda environment using Python version 3.7.9 and XGBoost package version 1.2.1.

5.1 Manufacturing Dataset

With the exception of the geometry, the manufacturing data have the same structure as those of the meta-training dataset. The data for this product was created over the course of six months as part of a product development project. Accordingly, there are slight changes in the process sequence as well as practical inconveniences such as sensor jumps in the period under consideration. For this reason, a thorough cleaning of the data is necessary, during which approximately 3800 data series with about 1000 data columns can be analyzed. Within this dataset, there are numerous retests of the two test sequences so that different temperature levels occur in the system. In addition, tests were performed at all times of the day and different times of the year by different workers.

5.2 Baseline Model

For the baseline XGBoost model, a time series split with a split ratio of 80/20 and a 5-fold cross-validation is employed. A simple grid search is applied for the selection of hyperparameters and named in parentheses below. Hyperparameters not mentioned are set to default. XGBoost squared loss regression (reg:squarederror) is used as the optimization function. The subsample ratio for each boosting iteration was set to 0.8, and the learning rate was set to 0.05 to avoid overfitting. For the same reason, the default of the maximum depth of the trees is lowered from 6 to 4 (max_depth=4) and an early stopping of the training was introduced when the optimization metric does not improve after 15 rounds (early_stopping_rounds=15). Another way to make the model more conservative is to increase the ridge regularization (lambda=20) and the minimum loss reduction for another partition at a leaf node to 100 (gamma=100). Additionally, the minimum sum of instance weight needed in a child is set to 5 (min_child_weight=5). A random seed is implemented to allow comparison of the models.

5.3 Enhanced Model and Final Evaluation

For the enhanced XGBoost model, the baseline XGBoost model is first augmented with the time variables and then additionally with the hand-crafted feature so that the enhancement due to the additional information in the model can be separated from the enhancement due to feature engineering.

The performance of the predictive models using manufacturing data is evaluated using the following standard metrics: root mean square error (RMSE), mean absolute error (MAE), the mean absolute percentage error (MAPE), and coefficient of determination (R^2). Table 3 provides an overview of the metrics for each of the training and unseen test dataset for all three models. In this use case, the MAE and R^2 are most useful for interpretation.

When analyzing the performance evaluation from Table 3, it can be noted that, surprisingly, the metrics of Model B deteriorate in the training dataset compared to the baseline model, but improve slightly in the test dataset. All metrics improve from the baseline model to the extended model C in both the training and test datasets, which plead for the approach to hand-crafted feature engineering based on physical ideas.

Table 3:Performance evaluation

Model name	Model description	train-dataset metrics				test-dataset metrics			
		rmse-mean	mae-mean	mape-mean	r^2-mean	rmse-mean	mae-mean	mape-mean	r^2-mean
A	baseline-model	2.9432	**2.3123**	0.1345	**0.5643**	3.8232	**3.0176**	0.1758	**0.2621**
B	A + time variables	2.9558	2.3215	0.1346	0.5605	3.8104	3.0070	0.1747	0.2669
C	B + t_factor	2.8538	**2.2443**	0.1304	**0.5904**	3.8022	**2.9979**	0.1744	**0.2702**

Regardless of feature engineering, there is a large generalization gap across models between the training and test datasets, suggesting an information gap in each of the manufacturing and meta-training datasets. There are two perspectives to look at the change in metrics: First, the improvement is predominantly in the second decimal place, which certainly does not justify the effort to do so. Second, it is worth noting that the prediction can be improved with a ratio of 16 experiment components to over 3500 production components and despite the existence of a huge information gap in the datasets. The improvement in all metrics is not a coincidence and is a starting point to push this approach further.

Finally, we will mention three justified assumptions for missing information in the dataset. First, the variation of clearance in the production dataset was assumed to be ± 1.5 µm for the meta-training dataset. This range is assumed to be very small and needs to be increased. Second, the sample size for the meta-training dataset is too small compared to the size of the production dataset, as this underrepresents the influence with the greatest predictive power. Third, the time variables used to represent temperature levels are less substantial compared to direct temperature measurement of the components and fluid.

6. Conclusions

The results of the leakage volume flow measurement are very sensitive to changes in gap geometry and temperature level in the measurement cross-section. Using a meta-regression model, a hand-crafted feature based on thermal behavior assumptions can be derived that improves all metrics for predicting leakage volume flow in the post-commission range despite a large cross-model generalization gap between the training and test datasets.

7. Future Work

Referring to the assumptions on the information gap in the datasets, we will significantly increase the sample size for the meta-dataset with a larger span of clearances. Moreover, experiments will follow from which the saturation curves for the time variables can be derived. In addition to more precisely defined time variables,

the number of thermal features will be increased by installing more direct temperature measurements in the test bench, if possible, and including them in the manufacturing dataset.

Furthermore, the minimum data acquisition for sufficient prediction accuracy will be investigated. Finally, the quality prediction will be implemented in the serial production of hydraulic valves to predict leakage test results and contribute to a more efficient and competitive manufacturing.

Acknowledgements

The authors would like to additionally thank Dr.-Ing. Matthias Möller and Bosch Rexroth AG for their support and access to their real-world data.

References

[1] Posa, A.; Oresta, P., Lipplois, A.; 2012. Analysis of a directional hydraulic valve by a Direct Numerical Simulation using an immersed-boundary method, Energy Conversion and Management 65, pp. 497–506.

[2] Lisowski, E.; Czyzycki, W.; Rajda, J.; 2013. Three dimensional CFD analysis and experimental test of flow force acting on the spool of solenoid operated directional control valve, Energy Conversion and Management Volume 70, pp. 220–229.

[3] Weichert, D.; Link, P.; Stoll, A.; Rüping, S.; Ihlenfeldt, S.; Wrobel, S.; 2019. A review of machine learning for the optimization of production processes, Springer, London.

[4] Köksal, G.; Batmaz I.; Testik MC; 2011. A review of data mining applications for quality improvement in manufacturing industry. Expert System Applications 38 (10), pp. 13, 448–13, 467.

[5] Lisowski, E.; Rajda, J.; 2013. CFD analysis of pressure loss during flow by hydraulic directional control valve constructed from logic valves, Energy Conversion and Management 65, pp. 285–291.

[6] Wuest, T.; Klein; D.; Thoben; K.-D.; 2011. State of steel products in industrial production processes, Procedia Engineering 10, pp. 2220–2225.

[7] Wuest, T.; Irgens, C.; 2013. An approach to monitoring quality in manufacturing using supervised machine learning on product state data, Springer Science+Business, New York.

[8] Duboue, P.; 2020. The Art of Feature Engineering. Essentials for Machine Learning, pp. 79, 124–126, Cambridge University Press.

[9] Cai, D.; Sheth, R.; Mackey, L.; Fusi, N.; 2020. Weighted Meta-Learning, arXiv:2003.09465v1.

[10] Kong, W.; Somani, R.; Song, Z.; Kakade, S.; Oh, S.; 2020. Meta-learning for mixed linear regression, Proceedings of the 37th International Conference on Machine Learning.

[11] Lovrec, D.; Kastrevc, M.; Ulaga, S.; 2008. Electro-hydraulic load sensing with a speed-controlled hydraulic supply system on forming-machines, pp. 474, 537–543, Springer Vieweg, London.

[12] Findeisen, D.; Helduser, S.; 2015. Ölhydraulik. Handbuch der hydraulischen Antriebe und Steuerungen, pp. 66–68, 474, 537–543, Springer Vieweg, Berlin.

[13] Matthies, H.J.; Renius, K.T.; 2014. Physikalische Grundlagen ölhydraulischer Systeme, in: Einführung in die Ölhydraulik, pp. 48–50, Springer Vieweg, Wiesbaden.

[14] Hodges; J.; 1958. The significance probability of the smirnov two-sample test. Arkiv för Matematik, 3, No. 43, pp. 469–486.

[15] Lopes, R; Reid, I.; Hobson, P.; 2007. The two-dimensional Kolmogorov-Smirnov test. Proceedings of Science.

[16] Wang, J.; Ma, Y.; Zhang, L.; Gao, R.; Wu, D.; 2018. Deep learning for smart manufacturing: Methods and applications, Journal for Manufacturing Systems 48, pp. 144–156.

[17] Wang, Y.-X.; Ramanan, D.; Hebert, M.; 2017. Learning to model the tail; In: Advances in Neural Information Processing Systems, pp. 7029–7039.

[18] Li, Z.; Zhou, F.; Chen, F.; Li, H.; 2017. Meta-sgd: Learning to learn quickly for few-shot learning. arXiv preprint arXiv:1707.09835.

[19] Bertinetto, L., Henriques, J. F., Torr, P. H.; Vedaldi, A.; 2019. Meta-learning with differentiable closed-form solvers. International Conference on Learning Representations. arXiv:1805.08136, 2019.

[20] Zhou, F.; Wu, B.; and Li, Z.; 2018. Deep meta-learning: Learning to learn in the concept space, arXiv preprint arXiv:1802.03596.

[21] Zintgraf, L.; Shiarli, K.; Kurin, V.; Hofmann, K.; White-son, S.; 2019. Fast context adaptation via meta-learning. In: International Conference on Machine Learning, pp. 7693–7702.

[22] Rajeswaran, A.; Finn, C.; Kakade, S. M.; Levine, S.; 2019. Meta-learning with implicit gradients. In: Advances in Neural Information Processing Systems, pp.113–124.

[23] Grant, E.; Finn, C.; Levine, S.; Darrell, T.; Griffiths, T.; 2018. Recasting gradient-based meta-learning as hierarchical bayes, arXiv preprint arXiv:1801.08930.

[24] Finn, C.; Xu, K.; Levine, S.; 2018. Probabilistic model-agnostic meta-learning. In: Advances in Neural Information Processing Systems, pp. 9516–9527.

[25] Kim, T.; Yoon, J.; Dia, O.; Kim, S.; Bengio, Y.; Ahn, S.; 2018. Bayesian model-agnostic meta-learning. 32nd Conference on Neural Information Processing Systems, arXiv preprint arXiv:1806.03836.

[26] Harrison, J.; Sharma, A.; Pavone M.; 2020. Meta-learning Priors for Efficient Online Bayesian Regression. In: Morales M., Tapia L., Sánchez-Ante G., Hutchinson S. (eds) Algorithmic Foundations of Robotics XIII. WAFR 2018. Springer Proceedings in Advanced Robotics, vol 14. Springer, pp. 318–337.

[27] Chen, S.; Li, J.; Song, Z.; 2020. Learning mixtures of linear regressions in subexponential time via Fourier moments. Proceedings of the 52nd Annual ACM SIGACT Symposium on Theory of Computing, pp. 587–600.

[28] Friedman, J.; 2001. Greedy function approximation: a gradient boosting machine. Annals of Statistics, 29 (5); pp. 1189–1232.

[29] Kiangala, S.K.; Wang, Z.; 2021. An effective adaptive customization framework for small manufacturing plants using extreme gradient boosting-XGBoost and random forest ensemble learning algorithms in an Industry 4.0 environment, Machine Learning with Applications, 4.

[30] Chen, T.; Guestrin, C.; 2016. XGBoost: A Scalable Tree Boosting System. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. pp. 785–794.

Biography

Christian Neunzig has been a doctoral student in the Bosch PhD program since 2019. Previously, he completed a master's degree in production engineering and a bachelor's degree in mechanical engineering at RWTH Aachen University. His research focus is applied machine learning on real production data. The dissertation is supervised by Univ. Prof. Dr.-Ing. Bernd Kuhlenkötter.

Simon Fahle has been a member of the Chair of Production Systems (LPS) at Ruhr-Universität Bochum since 2019. He completed a bachelor's and master's degree in mechanical engineering at Ruhr-Universität Bochum. His primary research topics are machine learning, time series and radial-axial ring rolling.

Bernd Kuhlenkötter was responsible for product management and engineering at ABB Robotics Germany until 2009. In 2009, Bernd Kuhlenkötter took over the professorship for "Industrial Robotics and Production Automation" at the Technical University of Dortmund. Since 2015, Univ. Prof. Dr.-Ing. Bernd Kuhlenkötter has held the professorship for "Production Systems" at the Ruhr University Bochum.

Matthias Möller is head of the Technical Functions department at the Bosch Rexroth plant in Homburg. Mr. Möller received his doctorate (Dr.-Ing.) and diploma (Dipl.-Ing.) from the Technical University of Dortmund.