Predicting Partial Discharges of Transformers: Decision Support System for Factory Acceptance Test

B. Gigerl[1], Y. Zhao[2], J. Raminger[3], J. Bakakeu[4], R. Kern[5], S. Thalmann[6]
[1]Department of Process Digitalization, Siemens Energy, Weiz, Austria, benjamin.gigerl@siemens-energy.com
[2]Department of Manufacturing Technology, Siemens Energy, Nürnberg, Germany, yang.zhao@siemens-energy.com
[3]Department of Test Field, Siemens Energy, Weiz, Austria, johann.raminger@siemens-energy.com
[4]Department of Sales Engineering, Alteryx, München, Germany, jupiter.bakakeu@alteryx.com
[5]Institute of Interactive Systems and Data Science, Technical University of Graz, Graz, Austria, rkern@tugraz.at
[6]Institute of Operations and Information Systems, University of Graz, Graz, Austria, stefan.thalmann@uni-graz.at

Abstract - **Partial discharges, mainly caused by an insufficient drying processes or different types of contamination, reduce the lifetime of a transformer, and thus lead to expensive rework costs. Herein, a decision support system for partial discharges of transformers occurring in oven and filling processes is introduced. Based on machine learning (ML), partial discharge results are evaluated in dependency of various manufacturing parameters and an automated prediction tool to guide the production with preventive actions is developed. The required data, obtained from sensors and manufacturing sources, is used to train supervised learning algorithms that aim to predict and classify partial discharges. To achieve adequate accuracy and reliability, multiple ML and data mining techniques are applied, including feature engineering, clustering, and final evaluation of performance by cost factors. The evaluation results show that the introduced ML models effectively detect and classify early test failures in oven and filling processes, resulting in a successful identification of key factors and consequently to more efficient action derivation. Overall, the potential of decision support systems as a valuable tool in the field of transformers is emphasized.**

Keywords - **Decision Support System, Predictive Quality Control, Machine Learning, Partial Discharges**

I. INTRODUCTION

Existing studies have highlighted the potential of machine learning (ML) in predictive quality control [1,2]. However, in industrial engineering the available data is limited, or the products are highly individualized. Our goal is to show in the context of transformers that a practical application is now possible. After all, damage to a main part of a transformer is associated with considerable costs.

In this study, critical and non-critical transformers were classified for the final partial discharge test, with an additional time series analysis of various sensor data on the product quality of the installed materials. Due to the complex relationships of production processes, our work highlights the importance of domain expert knowledge in the development of the ML models, which has also been noted as essential in previous studies [2,3]. Finally, we evaluate the strengths and weaknesses of common ML algorithms for oven and filling processes.

Recently, Suschnigg et al. [4] mentioned that industrial companies will be confronted with increasingly more complex production processes in the future, where a variety of factors can influence the final product quality. To identify patterns in quality data, various data mining as well as cleaning methods for general industry processes have been investigated in the past [3,5]. Building on this, Burggräf et al. [2] showed how these techniques could be successfully applied to sensor data from machines to improve preventive prediction of failure products in practice. Other scholars also see potential for improvements in overall business performance through predictive analytics [6,7]. However, to the best of our knowledge, no successful use case for partial discharge prediction exists in the literature.

The novelty is subject to complete automation for the production employee, whereby additional criteria, such as the planned delivery date, are also included in the final decision of the preventive quality action. This evaluation also extends the knowledge of existing studies in the application of supervised learning methods in quality prediction [8]. We also recognized that the interplay of data quality and explainable predictions is necessary for an effective integration in the process [9]. Finally, the results outlined that even with a small minority class of less than 150 partial discharge transformers, we were able to develop an accurate system for predicting test failures, while also allowing production employees to understand the decisions made by the artificial intelligence (AI).

II. USE CASE

Transformers are complex industrial products that are manufactured in various manual steps. The final partial discharge test in the test laboratory is decisive in determining whether the product meets the customer's required quality and can therefore be delivered. To narrow down the multitude of influencing factors, the oven and filling processes were chosen in this study, as these manufacturing steps have a high degree of digitalization. By implementing Industrial Internet of Things (IIoT), sensor data from the machines can thus be automatically stored in databases [6]. In our application, we use two oven systems, both of which are parameter controlled. Fig. 1 shows oven A, which can dry up to five active parts of transformers simultaneously using low-frequency heating and circulating air.

Fig. 1. Production plant oven A for the drying process

Dew point, pressure and temperatures in the plants are recorded in different sequences over the process time and evaluated in pressure temperature control (PTC) reports.

Afterwards, the installation of the active part in the tank begins and the final product is isolated from the atmosphere. After a historical data analysis, however, we found that seasonal influences can increase the partial discharge occurrence. Through internal experiments from the past, domain experts know that woods, e.g., can rewet during this tanking period. Therefore, we also include sensor data from our internal weather stations from the halls in the analysis to take external factors into account in the ML model. Downstream, the transformer is filled with oil, whereby different pump runs and circuits are conducted to ensure high quality. Fig. 2 shows the filling process.

Fig. 2. External filling system for filling processes

During the standing time until the partial discharge test, the oil impregnate the transformer and it is finalized by the production employees. Historical repair actions are divided into three classes, *no action*, *light action,* and *heavy action*, with a binary differentiation made for simplicity. The goal is thus to develop a system for improved decision-making that identifies critical transformers early enough so that quality actions can be initiated at the right time. The implementation of these actions is then illustrated by the automation of the decision support system.

III. MODEL OVERVIEW

Overall, the model incorporates data from various departments such as quality, engineering, manufacturing, test field and planning. The step-by-step implementation is illustrated by the architecture in Fig. 3.

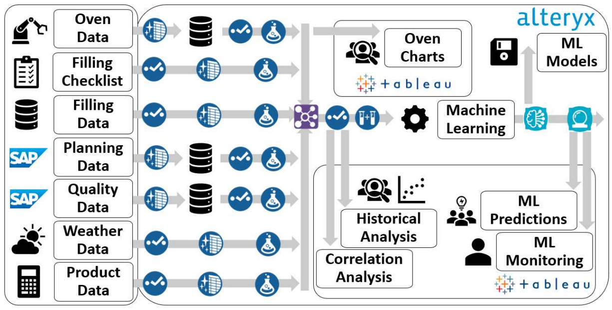

Fig. 3. Architecture of the ML methodology

The data integration is mainly done from different SQL databases and the entire extract, transform, load (ETL) process is done via Alteryx [10]. After all, a total of 12 raw tables were linked at transformer level, whereby initially the filling checklist had a data quality of 68% consistent available data on average for all features. During data pre-processing, a lot of time was spent on cleaning the data, which ultimately accounted for 70% of the entire Data Science Lifecycle (DSL). The biggest challenges were different measurement methods, which were resolved through process knowledge. This is in line with Schuetz et al. [6] that collaboration with domain experts is essential for the knowledge generation of the model. The production operator has the possibility to perform two quality predictions for partial discharges in the process. The first after the end of the oven process and the second after the end of the oil circulation. In our study, process records from January 2022 onwards are used, as various plant retrofits before this date can lead to false conclusions. In addition, this selection allows consistent data to be available for each month in 2022, which was used for the further correlation analysis of the weather data on product quality. In total, this analysis covers 447 oven processes and 406 external filling processes.

First, the target processes and interesting partial discharge results from the past were discussed in expert meetings. With the knowledge generated, the raw data was then automatically linked and cleaned, creating an appropriate data quality through plausibility checks. Overall, plant parameter checks enabled an absolute increase in ML accuracy of 7%. The feature engineering methodology here is based on hypotheses from process experts, whereby 32 additional parameters were considered in the oven and filling process for the evaluation of the ML model. Finally, 112 parameters were assessed for significance to evaluate the algorithms. To correctly classify transformers for partial discharges, we evaluated Decision Tree, Logistic Regression, Random Forest and XGBoost as part of the case study. The Naive Bayes classifier was also tested on the raw data, but due to dependent predictors, only discriminative models were used in the final algorithm selection. For ML training, we chose a 70/30 split of the data. We justify this choice by the fact that we use a larger dataset for the validation according to the non-conformity cost (NCC) criterion to evaluate the correct ML test predictions. To take the imbalance of the data into account in the classification, attention was paid to the minority class of true positives, i.e., real partial discharges. This different evaluation methodology thus helps to select the ML model that best identifies future total losses preventively in the manufacturing process. In addition, the selected models are retrained daily using a random sample so that performance over time can be monitored by the data scientist using dashboards. Following the visual analytics dashboard of Suschnigg et al. [4], we created interactive dashboards in our server structure, where specific correlations can be examined through filter methodologies from the entire data mining process. The decision support system is presented in Fig. 4.

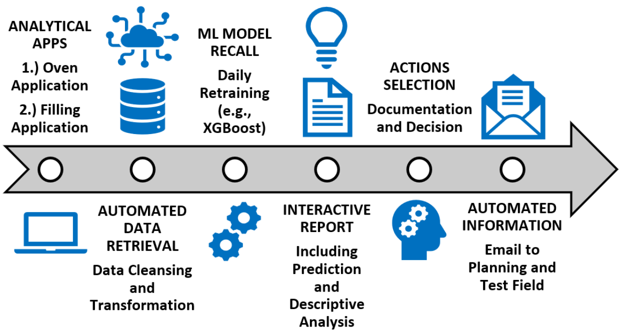

Fig. 4. Decision support system for partial discharges

In the production environment, the production operator accesses the current ML models through analytical apps. When a process is finished, the user can request actual manufacturing data with a call request and thus perform a live quality forecast. Automatically, the last process batch is identified, and a partial discharge failure probability is calculated for each transformer. Our focus is on explainable AI, so all-encompassing process diagrams, significant outlier parameters in the process and statistics of similar good and bad transformers are presented to the production operator as part of an interactive report. This helps the user to interpret quality-influencing factors from oven and filling processes and thus to implement preventive actions in the next step. To make the decision-making process transparent and efficient, predefined actions are suggested and details of the action can be documented by means of a comment field. For example, Fig. 5 shows the production operator's user interface for documenting the actions for the oven A model.

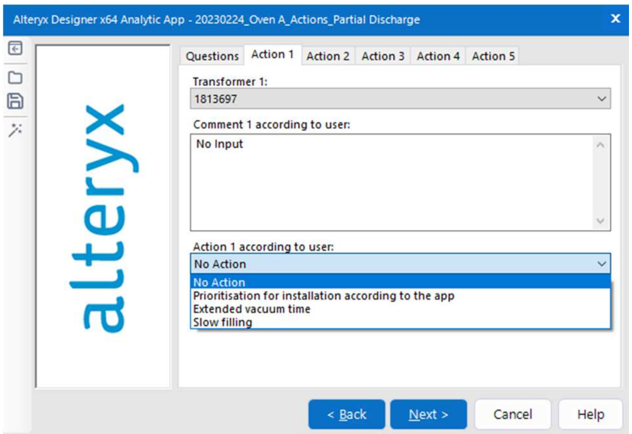

Fig. 5. User interface for preventive filling process actions

In the event of high failure probabilities for partial discharges, the planning and test field departments are informed directly by means of an automated email, including the report and the metadata for the action. Each action taken in turn helps the AI system to learn from the effects of the preventive actions taken.

IV. EVALUATION

To select the correct ML methodology, four different algorithms were trained on the existing process data. In the initial training, the oven model refers to 617 and the filling model to 394 transformers.

A. Selection of the Best Performing Algorithms

The algorithms of the initial training are shown in Table I.

TABLE I
RESULTS OF THE TESTED ALGORITHMS
Explanation: True negative (TN), False positive (FP), False negative (FN), True positive (TP), True positive non-conformity costs (TP-NCC)

	ML OVEN MODELS (N = 617 Transformers)					
Algorithms	Accuracy [%]	TN [%]	FP [%]	FN [%]	TP [%]	TP-NCC [€]
XGBoost	83.3	97	3	79	21	199,461
Random Forest	82.5	99	1	91	9	-
Logistic Regression	81.8	100	0	100	0	-
Decision Tree	76.5	88	12	77	23	74,473

	ML FILLING MODELS (N = 394 Transformers)					
Algorithms	Accuracy [%]	TN [%]	FP [%]	FN [%]	TP [%]	TP-NCC [€]
Random Forest	82.2	98	2	84	16	3,639
XGBoost	81.2	96	4	79	21	48,720
Logistic Regression	80.7	100	0	100	0	-
Decision Tree	70.8	83	17	80	20	-

In this study, we evaluate the performance of the algorithms as follows. On the one hand, we rank the algorithms according to their accuracy, which leads us to the first conclusion. Logistic Regression and Decision Tree did not provide good overall accuracy, but the Decision Tree in the oven model achieved the best value for predicting the actual percentage of partial discharges. If a transformer is assigned the value true positive, we could have correctly identified it early on. In this use case, however, the goal is to preventively identify the transformers in need of repair, whereby we also need algorithms with a low prevalence error. Here, the XGBoost method was shown to be the best choice in both processes, correctly classifying the largest number of NCC in each case, as well as providing stable prediction accuracy. To extend the findings, the Decision Tree for the oven process and the Random Forest for the application of the filling process were included in the further analysis to compare the productive algorithms over time.

The significance test of the features was done by the permutation importance of Gini impurity. This ranking provided new insights into the parameters influencing oven and filling processes on partial discharges. For example, the existing knowledge of maximum temperatures in the oven, process end values and tanking times could be optimized and extended by these findings. In general, the process times of individual sequences, the humidity in the halls, the temperature fluctuations in the plants, as well as the technical specification of the transformers showed a high feature importance after hyperparameter tuning. We were also able to develop a novel methodology for the automated time series analysis of the dew point curves, which highlights humidity indicators for the process experts and takes them into account in the ML model. Using explainable AI, it is thus possible to adjust future process parameters based on the new quality indicators.

B. Evaluation and Implementation in a Live Application

To check the effectiveness of the models, partial discharges from production were predicted and compared with model test forecasts. The key points for the evaluation are shown in Table II.

TABLE II
RETROSPECTIVE EVALUATION OF THE ML MODELS

PERIOD: 02/01/2023 - 02/03/2023 (8 ½ Weeks)	
Number of Transformers [#]	Hit Rate of the Target Variable for Actual Partial Discharges [%]
92	71

The result was that five out of seven transformers with partial discharges could be predicted correctly preventively in the process. Two transformers after the oven process and three transformers after the finalization of the filling process. The model calculated failure probabilities for insufficient oven processes from 70% to 82% and failure probabilities for critical filling processes from 56% to 75% for an actual test damage. The assessment of the predictive power was analyzed in detail by verifying the PTC reports with process experts, as well as retrospectively through physical quality sampling on the shop floor.

In addition, we subsequently show in our study the results when performing a live practice test of the preferred XGBoost algorithms from the ML model in later production (see Table III).

TABLE III
PRODUCTIVE INSTANCE OF THE ML MODELS

PERIOD: 03/03/2023 - 19/05/2023 (11 Weeks)	
Number of Transformers [#]	Hit Rate of the Target Variable for Actual Partial Discharges [%]
124	24

In the second step, the value of the true positive class decreased and approached the initial test model. It should be noted that in this use case, preventive actions are significantly more cost-effective compared to a total loss with a delivery delay, whereby the misclassified partial discharges do not significantly minimize the company's turnover. With a failure probability of more than 50%, production employees had to take an action directly after the end of the process, which in turn represents a causal relationship to the final test field result. Therefore, we consider it essential that the ML models are linked to the real quality actions taken and thus learn from them in the future.

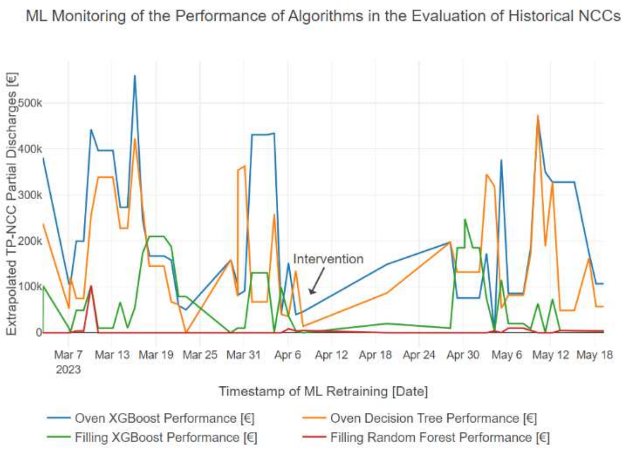

Fig. 6. Extrapolated NCC of the true positive (TP) partial discharges

The illustration shown in Fig. 6 visualizes how the ML algorithms behaved over time. The timestamp of the retraining of the models is plotted on the horizontal axis and the vertical axis shows the respective monetary sum of the correctly classified NCC from the past in Euros. The data scientist can thus use the NCC from the quality to monitor whether one of the classification algorithms deviates significantly to perform a new selection of the classification algorithm. For example, we documented that the oven Decision Tree exceeded the XGBoost twice, with the selected ML model reaching the threshold again within two days. This effect is shown by the random sample of the 30% of validation data used to calculate the extrapolated value for the accruing NCC of the past partial discharges since January 2022. On the other hand, our study shows that the oven models follow the same trend, but that the XGBoost is preferable for these specific data.

V. DISCUSSION OF RESULTS

The potential of broad data availability for predicting test results in production was demonstrated by our practical evaluation [11]. Our experience showed that domain knowledge is not only necessary at the beginning of the DSL but should be understood as an iterative process in each ETL step. It is important to note the causalities of the process data for the prediction in the ML model. If two events occur various times together, it cannot be assumed that the events are in a cause-effect relationship [6]. Therefore, we see it as necessary to check the plausibility of the predictions by means of explainable AI or to apply causal discovery [12]. In a further step, the influence parameters from upstream production processes, the storage conditions of purchased parts and particle measurements in the halls are also to be included in the analysis. In addition, the true positive rate could be further improved, but this could not be implemented at the time of the study because the databases have not yet been expanded across all processes. Thus, we see it as mandatory in the future that machines are comprehensively equipped by means of IIoT. Another point of discussion would be to conduct a quality prediction in advance after the technical design of the transformer. This would provide additional information for retrograde planning and thus further optimize the production flow. In relation to the specific study, however, the effectiveness of XGBoost in practical cases for test field results of industrial products could be emphasized, whereby we call for further research in other application domains.

VI. CONCLUSION

The results show that even with limited amount of data, accurate predictions of partial discharges in transformers are still possible. By evaluating different ML algorithms, it was shown that XGBoost achieved the best performance. The model was implemented in a productive environment where production employees access the current predictions and evaluate quality-influencing factors via interactive report to derive actions. The use of explainable AI is crucial to make the predictions plausible and to set the most appropriate actions. Evaluation of the models against real-world data showed that five out of seven transformers with partial discharges were detected preventively. A live test phase in production also visualized how the algorithms behaved over time according to the NCC decision criterion. It has been shown that predictive quality control could identify significant rework costs at an early stage. Moreover, we would like to draw attention to the need for collaboration between data scientists and domain experts to ensure that data cleaning and data interpretation are done correctly. In general, it can thus be shown that factory acceptance tests can be made more sustainable and cost-efficient through decision support systems with limited data and highly customized products.

REFERENCES

[1] T. Wuest, D. Weimer, C. Irgens, and K. Thoben, “Machine learning in manufacturing: advantages, challenges, and applications”, *Production and Manufacturing Research: An Open Access Journal*, 4(1), 2016, pp. 23-45.

[2] P. Burggräf, J. Wagner, B. Heinbach, F. Steinberg, A. R. Pérez M., L. Schmallenbach, J. Garcke, D. Steffes-Lai, and M. Wolter, “Predictive analytics in quality assurance for assembly processes: lessons learned from a case study at an industry 4.0 demonstration cell”, *Procedia CIRP*, 104(1), 2021, pp. 641-646.

[3] S. Xu, B. Lu, M. Baldea, T. F. Edgar, W. Wojsznis, T. Blevins, and M. Nixon, “Data cleaning in the process industries”, *Reviews in Chemical Engineering*, 31(5), 2015, pp. 453-490.

[4] J. Suschnigg, F. Ziessler, M. Brillinger, M. Vuković, J. Mangler, T. Schreck, and S. Thalmann, “Industrial Production Process Improvement by a Process Engine Visual Analytics Dashboard”, In *Proceedings of the 53rd Hawaii International Conference on System Sciences*, 2020, pp. 1320-1329.

[5] N. Laranjeiro, S. N. Soydemir, and J. Bernardino, “A Survey on Data Quality: Classifying Poor Data”, In *IEEE 21st Pacific Rim International Symposium on Dependable Computing*, 2015, pp. 179-188.

[6] C. G. Schuetz, M. Selway, S. Thalmann, and M. Schrefl, “Discovering Actionable Knowledge for Industry 4.0: From Data Mining to Predictive and Prescriptive Analytics”, In *Springer eBooks*, 2023, pp. 337-362.

[7] M. Gashi, and S. Thalmann, “Taking Complexity into Account: A Structured Literature Review on Multi-component Systems in the Context of Predictive Maintenance”, In *Lecture notes in business information processing*, 2019, pp. 31-44.

[8] Z. Ge, Z. Song, S. X. Ding, and B. Huang, “Data Mining and Analytics in the Process Industry: The Role of Machine Learning”, *IEEE Access*, 5(1), 2017, pp. 20590-20616.

[9] L. Bertossi, and F. Geerts, “Data Quality and Explainable AI”, *Journal of Data and Information Quality*, 1(1), 2020, pp. 1-9.

[10] Alteryx. (2023, May 25). Why Alteryx | Alteryx. https://www.alteryx.com/why-alteryx.

[11] Y. Cheng, K. Chen, H. Sun, Y. Zhang, and F. Tao, “Data and knowledge mining with big data towards smart production”, *Journal of Industrial Information Integration*, 9(1), 2018, pp. 1-13.

[12] M. Vuković, and S. Thalmann, “Causal Discovery in Manufacturing: A Structured Literature Review”, *Journal of Manufacturing and Materials Processing*, 6(1), 2022, pp. 2-17.