Evaluation of Anomaly Detection Algorithms on Circuit Breaker Lifetime Test Data

Ralf Gitzel[1], Fabian Frech[1] and Aydin Boyaci[1]

[1] *ABB Corporate Research, Kallstadter Str. 1, 68309 Mannheim*

Abstract

Industrial condition monitoring based on AI is often hampered by the lack of properly labeled data. One possible solution to this problem is to use anomaly detection algorithms. In this paper we test four machine learning anomaly detection algorithms (DeepSVDD, GMM, KNN, ECOD) on endurance test condition monitoring data recorded from two medium voltage circuit breakers in a long-term laboratory experiment. The paper shows the suitability of these algorithms for the early detection of faults in a deteriorating mechanical system. We benchmark the results against the interpretation of traditionally-used metrics. The comparison leads to the conclusion that anomaly detection on the raw sensor data is a preferable alternative to those metrics.

Keywords

Machine learning, anomaly detection, switchgear, circuit breakers, condition monitoring

1. Introduction

Modern life is heavily dependent on a reliable and steady supply of electrical energy. One of the key components of the electrical grid is medium voltage (MV) switchgear [7]. Switchgear panels are found in substations around the world and are mainly responsible for the safety of devices connected to the grid by interrupting current flow in abnormal situations. Since only healthy switchgears can guarantee the availability of this critical safety function, it is prudent to monitor their condition. One condition monitoring approach is to use so-called travel curve sensors on the circuit breakers (CBs) housed in the switchgear panels.

Condition monitoring solutions need to be automated to be deployed cost-effectively. Most breaker monitoring solutions today use metrics and thresholds based on experience and engineering domain knowledge. Often, frequent human decisions are needed to interpret the current asset health state. In this paper we want to explore whether the use of AI, in particular anomaly detection algorithms, can improve on the state of the art of MV circuit breaker condition monitoring.

The paper has the following structure. We start with an explanation of the problem domain, the sensors used, and the data available. In section 3, we survey the related work, i.e. anomaly detection algorithms and their application to condition monitoring. We also explain

LWDA'24: Lernen, Wissen, Daten, Analysen. September 23–25, 2024, Würzburg, Germany

ralf.gitzel@de.abb.com (R. Gitzel); fabian.frech@de.abb.com (F. Frech); aydin.boyaci@de.abb.com (A. Boyaci)
https://www.abb.com/ (R. Gitzel); https://www.abb.com/ (F. Frech); https://www.abb.com/ (A. Boyaci)
0000-0001-7584-4329 (R. Gitzel); 0000-0001-7584-4329 (F. Frech); 0000-0001-7584-4329 (A. Boyaci)

CEUR Workshop Proceedings (CEUR-WS.org)

the traditional metrics that will be used as a benchmark. In section 4, we describe our proposed anomaly detection approach including the physical experiments conducted to obtain the data. We compare these results to traditional metrics and end with conclusions and future work in the section 5.

Our goal for this paper is to answer the following **research question:** Are anomaly detection algorithms suitable for MV breaker monitoring and how do they compare to traditional approaches?

2. Domain Background - Breakers and Monitoring

It is beyond the scope of this paper to discuss the details of medium voltage circuit breakers and their failure modes. However, to provide some context for the reader, we briefly explain the technology of both the asset monitored and the sensors used for their monitoring.

2.1. MV Circuit Breakers

A CB is an electrical safety device to protect electrical circuits from damage due to overcurrent. Its primary role is to interrupt the current flow protecting equipment and minimizing the risk of fire. The circuit breaker is resettable after tripping, unlike a fuse, which must be replaced after tripping.

A MV CB is capable of a very large number of opening and closing operations before the end of its useful life. In this study, the number of operations is non-dimensionalized with respect to a standard lifetime of MV CBs. Common ratings of MV CB are voltages of up to 36kV, currents of 3kA and short circuit currents of up to 63kA. MV CBs are installed in MV switchgear panels that protect our electrical gird.

2.2. Travel Curve Sensors

There are several ways to monitor the condition of CBs. A common sensor type used in CBs are mechanical position sensors such as linear potentiometers for travel curve monitoring. The positions sensors measure the travel curve at appropriate locations in the kinematic chain that represents the movement of the moving electrical contact for each phase. The measured analogue voltage signal is then sampled with a laboratory data collector (cDAQ). Due to the high speed and acceleration a high sample rate is necessary to get a decent amount of data in the operation time window.

In simpler terms, the travel curve represents the position (shown on the y-axis) of the moving contact of the CB at a given point in time (x-axis). The intermediate values are traversed as the contact “travels” between open and closed. Faults are very likely to affect the speed and positions of the opening and closing operation. An example of an opening as well as that of a closing travel curve will be shown and discussed in section 3.3.

3. Related Work

There are three areas of related work we need to discuss. First, we examine anomaly detection algorithms in general and the ones we have picked for this paper in particular. Next, we discuss related work that uses such algorithms for condition monitoring. To our knowledge, there are almost no examples of anomaly detection algorithms applied to MV CB travel curve data but there are some interesting papers for other equipment. Finally, we need to briefly explain the traditional metrics we use as a baseline for our algorithms.

3.1. Anomaly Detection Algorithms

According to Ruff et al., an anomaly is an observation that deviates from some concept of normality. Anomaly detection algorithms use initial training data to establish this concept and new samples are rated according to some kind of anomaly score [13]. A division into "normal" data previously seen and some future unknown data with unclear properties is quite suitable for condition monitoring where we cannot anticipate all possible failure modes.

Recent developments go towards deep models for anomaly detection. Choi et al. identify the following categories of deep models for time series anomaly detection: Autoencoders in different variants, Generative Adversarial Networks (GANs), Transformers, Graph Neural Networks (GNNs), Recurrent Neural Networks (RNNs), Temporal Convolutional Networks (TCNs), and Hierachical Temporal Memories (HTMs) [4]. Additionally, there are deep one-class classifiers and self-supervised methods [13].

The first big group are generative models. These models typically exploit the learned compression to a lower-dimensional latent space which results in some form of reconstruction error. The error can be used as an anomaly score, because previously unseen types of data are expected to be less easy to reconstruct than "normal" data. A good overview of GANs suitable for anomaly detection can be found in [5]. We did not use these models in this paper because they need large volumes of training data that we could not provide, which resulted in terrible performance during some superficial tests.

Self-supervised anomaly detection is a fascinating concept based on the idea of applying transformations on "normal" data to generate anomalies. These new samples can be labeled automatically as the type of transformation is known. Using this data, a classification algorithm can be trained to identify the transformations. The working hypothesis is that anomalies will be classified as one or more transformations [9]. There are similar approaches such as contrastive learning [15]. However, these approaches require a selection and implementation of transformations that might or might not allow good anomaly detection in the end. Also these models have mostly been used on image or video data so far (cf. [2]).

In this paper, we examined four anomaly detection algorithms. The first is a one-class classifier called Deep Support Vector Data Description (DeepSVDD) [14]. DeepSVDD is a deep model that uses a neural network to map the features tightly into to a lesser-dimensional space. Any anomaly will exhibit a high distance from the center of the hypersphere that encapsulates the features of the training data. Since our data consists of highly correlated data points, a deep model with dimensionality reduction seemed to be promising.

Next, we are using an algorithm derived from the K-Nearest Neighbors (KNN) classification algorithm. Like the original algorithm, this variant determines the distance to the K nearest neighbors. However, there are no classes, so the distances are used to compute an anomaly score (e.g. max or mean) [12].

Third, we use Gaussian mixture models (GMMs), which are based on the assumption that the data under observation is generated by a mixture of several Gaussian distributions. Each distribution represents one cluster in the data. Outliers have a low probability of belonging to any of those components (cf. [3]).

Our final model is Empirical-Cumulative-distribution-based Outlier Detection (ECOD), a recent non-deep anomaly detection algorithm. ECOD is an unsupervised outlier detection algorithm that identifies data points that deviate from a general data distribution. It is based on the idea that outliers are often the "rare events" that appear in the tails of a distribution. ECOD assumes that all features are statistically independent, which allows a non-parametric distribution estimation in the form of an empirical cumulative distribution function [10]. While this assumption is not true for our data, we decided to explore its suitability despite this potential limitation.

3.2. Condition Monitoring Anomaly Detection

There are many examples of using anomaly detection algorithms to interpret condition monitoring data, so we can give only a few illustrative examples.

The work closest to our experiment is by Malek Mahdavi who uses autoencoders to analyze the travel curves of low voltage CBs (as opposed to medium voltage in our case). In the publication, autoencoders and dynamic threshold techniques are used to detect anomalies [11]. Despite the superficial similarities, both the assets under observation as well as the algorithms are so different that a direct benchmarking is not possible.

There is one other paper that examines condition monitoring of switchgear using anomaly detection algorithms. The paper looks at partial discharge and identifies such scenarios with an autoencoder that reconstructs UHF sensor signals [16]. This is different from our use case which focuses on mechanical as opposed to electrical faults.

The remaining papers focus on different assets. Some papers apply condition monitoring anomaly detection to data from whole plants. Dix and Ganguly use various anomaly detection algorithms to detect the failure of the sensors and valves in a system to separate oil, water, and gas [6]. Zhang et al. use a Gaussian Process Ensemble Model to analyze power plant data to detect anomalies. Their input is multivariate time series data [17].

Hendrickx et al. apply anomaly detection to fleet data (electrical motors), postulating that most elements in the fleet would be normal and any fault an anomaly. The approach does not require historical data as comparisons are made within the existing fleet. Their input is electrical and vibration data [8]. Ahmad et al. explore the use of autoencoder-based anomaly detection algorithms on rotating machinery data. The input data are raw vibration signals without any feature engineering. The algorithm is compared to an alternative one based on IForest and using handcrafted features [1]. Additionally, Zhang et al. use a probabilistic regression model to detect anomalies in wind turbines [18].

It is interesting to note that most of these approaches do not use the algorithms the state of the art literature identifies as suitable anomaly detection tools, with the exception of autoencoders.

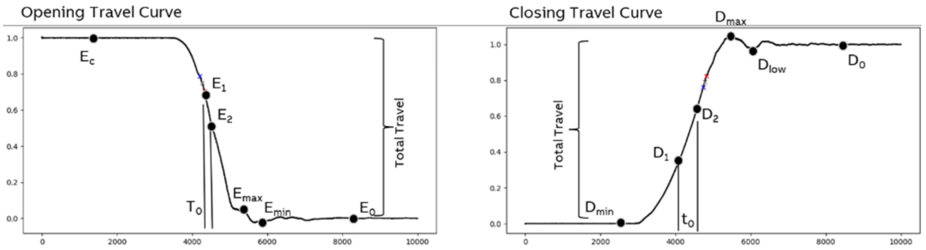


Figure 1: Traditional metrics for travel curves

For this reason, we did not try to adapt these algorithms to our problem. Furthermore, almost all papers we found are for different types of equipment. Essentially, the problem of detecting mechanical anomalies in medium voltage switchgear using state of the art anomaly detection algorithms is not solved yet.

3.3. Traditional Metrics

In applications today, travel curves are analyzed via a series of metrics computed from the curves. In this paper, we use the following metrics according to [11] as a baseline to compare our algorithms to (also see **Fig. 1**):

- Total travel: The distance between the closed and open position travelled during the opening or closing operation. It is the delta between E_c and E_0 for opening and between D_0 and D_{min} during closing.
- Time duration/period for the estimation of opening/closing speed: The time T_0 or t_0 that passes between two characteristic points along the slope of the curve. These points are fixed positions and are called E_1 and E_2 or D_1 and D_2 respectively.
- The distance between E_{max} and E_{min} during opening.
- The distance between D_{max} and D_{low} during closing.

4. Anomaly Detection Approach

In this section we describe our experiments. We only give a high-level description of the hardware setup as this paper is intended for an audience more interested in the software aspects.

4.1. Experimental Setup (Hardware)

In our lab, we ran several endurance tests on CBs of type VD4. The test setup is shown in **Fig. 2**. During the test, the CB is frequently opened and closed to simulate its use in the field at an accelerated pace. The last recorded travel curve is the one of the operation when the breaker finally failed. At the time of this publication, usable lifetime test data was available for two of the CBs, labeled CB4 and CB6.

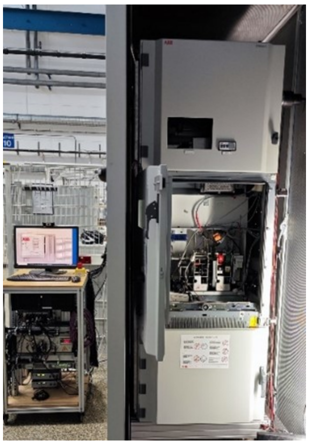


Figure 2: The experimental setup in the laboratory

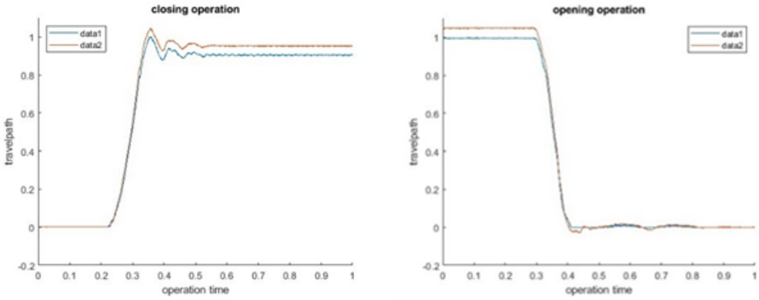


Figure 3: Examples of the data from the endurance test

For the data acquisition, a cDAQ HIL system was used. Its modular configurability makes it very convenient to process various sensor and control signals simultaneously.

The system is controlled using relays, switches, and external power supplies, which, in turn, are controlled with the cDAQ system by its digital output channels. A LabView program is running on a measurement computer for this purpose. Measurement data is only captured during operation time, triggered by control signals. This significantly reduces storage space compared to continuous data acquisition. On the sensor side, the travel curve sensors described in section 2.2 were used.

4.2. Experimental Setup (Software)

The result of the physical experiment are almost 26 GB of binary data. For each opening and closing operation, there is a single travel curve consisting of about 11000 values that represent the positions during the operation. The data are unlabeled but we can assume early data points to still be healthy. Examples of the actual data are shown below (**Fig. 3**). Both datasets were normalized in time and absolute value.

This means that there are four datasets available for use in this paper: CB4 open, CB4 close,

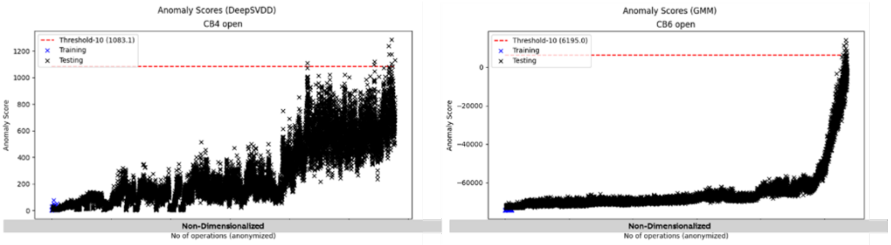


Figure 4: Examples of results

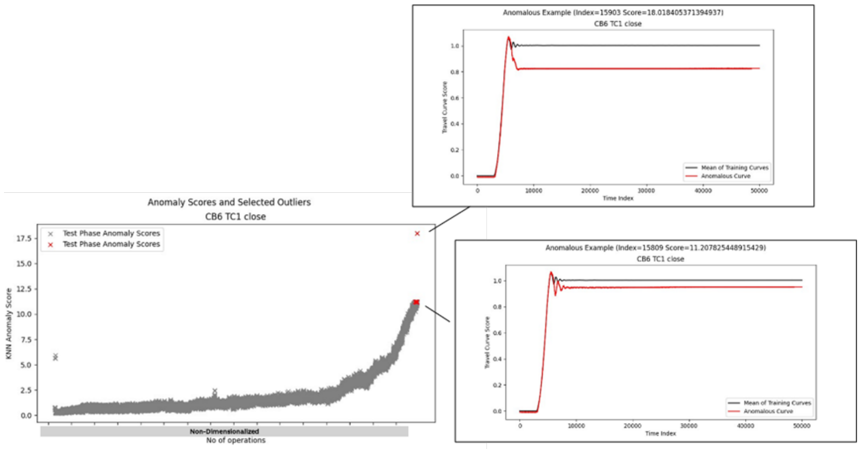


Figure 5: Deep-dive analysis of some KNN anomalies in CB6 close

CB6 open, and CB6 close. Each dataset was divided into training data (the first several hundred samples) and test data (the remainder). Next, four anomaly algorithms were applied: DeepSVDD, KNN, GMM, and ECOD. Some exemplary test results can be seen in **Fig. 4**. The curves show how anomaly scores develop over time. In both of these two examples, there is a trend that shows the effect of deterioration in the breaker. The CB4 data shows a system that is unstable with parts sliding in and out of the optimal position due to damage.

The example on the right of **Fig. 4** (CB6 open) has a strong trend in its anomaly score that is easy to see for human eyes. In **Fig. 5**, we explore the data for this case further. We have highlighted two late-stage travel curves (shown in red) in comparison to the average of early-stage travel curves (black). There is a clear change in the total travel, i.e., the breaker does not fully close any more. This is a first piece of evidence that the anomaly detection works with our data. However, a more elaborate evaluation strategy is needed.

4.3. Evaluation

Many papers exploring anomaly detection algorithms for condition monitoring use an unsupervised or self-supervised algorithm but for the test scenario have access to fully labeled data for an evaluation [6, 8]. For example, Dix and Ganguly use a simulator to test their algorithms [6].

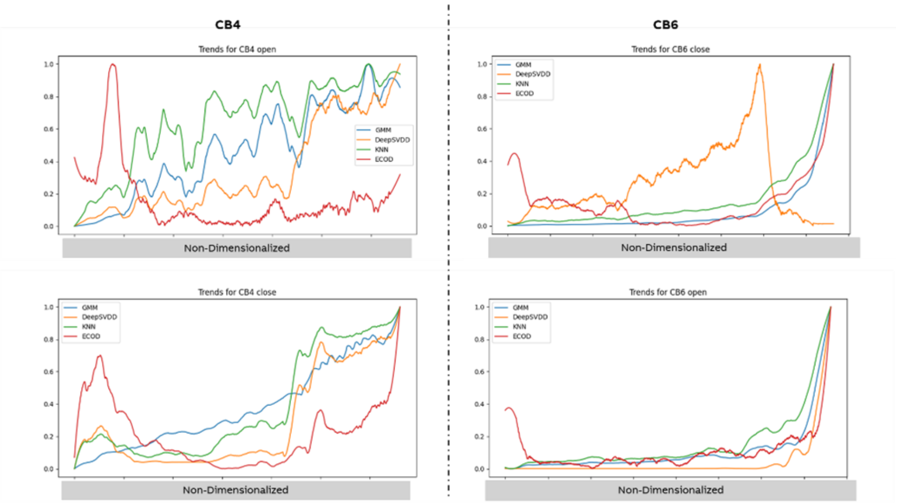


Figure 6: Normalized and smoothened anomaly scores for different algorithms

Others only evaluate behavior during clearly normal states (e.g. Zhang et al [17]).

The endurance test data in this paper is fully unlabeled. We can safely assume the first n data points to be healthy and the last m data points to be increasingly faulty but neither n nor m are known values. For the sake of the study, we define a relatively low n to identify a training set and know that the last value is the strongest anomaly as it represents the failure.

Therefore, we can only do a limited analysis of algorithm quality. First, based on the assumption that the data is from a deteriorating mechanical system, a good algorithm will show a clear trend that culminates in failure. Second, it should be possible to statistically identify a threshold that gives a warning that a failure would occur soon. The threshold should only be crossed a few operations before failure. Warnings given too early result in unnecessary cost as still working equipment is replaced too soon.

4.3.1. Presence of Clear Trends in Anomaly Scores.

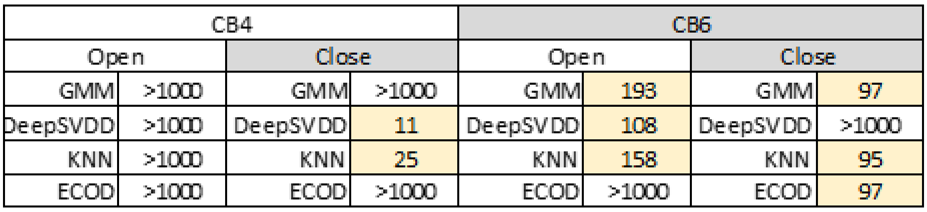
To make the raw anomaly scores of the different algorithms easier to interpret and compare, we have applied a filter to make the curves smoother and have normalized all results to a range from 0 to 1. These results are plotted in Fig. 6 for both breakers and their opening and closing operations. The comparison shows that many of the algorithms exhibit a clear trend that is indicative of deterioration. However, there are also cases that are not suitable for condition monitoring such as DeepSVDD applied to the closing operation of CB6, which behaves in a completely counterintuitive way after a major peak relatively close to failure.

A subjective, visual assessment whether the curves display a clear rising trend (important for monitoring of deterioration) yields the results shown in **Table 1**. The algorithms have a good performance except for ECOD which displays a strong peak in the beginning in all four cases.

Table 1
Trend in anomalies for different algorithms

Algorithm	Cases with strong trend	Percentage of all cases
GMM	4	100%
DeepSVDD	3	75%
KNN	4	100%
ECOD	0	0%

Table 2
Early warning in number of operations



CB4				CB6			
Open		Close		Open		Close	
GMM	>1000	GMM	>1000	GMM	193	GMM	97
DeepSVDD	>1000	DeepSVDD	11	DeepSVDD	108	DeepSVDD	>1000
KNN	>1000	KNN	25	KNN	158	KNN	95
ECOD	>1000	ECOD	>1000	ECOD	>1000	ECOD	97

4.3.2. Existence of a Potential Warning Threshold.

The anomaly scores of a perfect algorithm would increase monotonously as the equipment accumulates minor damage. If (in retrospect) we can take the 10^{th} largest anomaly score as a threshold, the ideal algorithm would give a warning 10 operations before failure. Using such a threshold, the following early warning times are achieved for our actual algorithms (Table 2). Some of the algorithms are foiled by early outliers that result in useless warnings.

4.3.3. Comparison to Traditional Metrics.

The traditional metrics can be used by a human to see developments and judge whether the overall trend is detrimental. The metrics can act as manually defined “anomaly scores” and thus be used as a benchmark for the anomaly detection algorithms.

For the first comparison, we normalize and smoothen the raw metrics. We also flip decreasing metrics (i.e. total travel) for consistency. The results as shown in **Fig. 7**. One striking parallel to **Fig. 6** is the shape of the metrics and anomaly scores for CB6. Clearly, the anomaly is dominated by these effects.

In **Table 3**, a subjective assessment of the visibility of trends is shown. Since opening and closing uses different metrics, the highest score achieved for each category is 2 as opposed to 4 for the anomaly detection algorithms. T0 shows a good trend for both breakers.

The early warning potential of the various metrics is shown in **Table 4**. When compared to **Table 2**, these values are inferior for the more complex failure seen in CB4. Total travel of CB6 is roughly on par with the anomaly detection algorithms.

Clearly, there are limitations to this comparison. **Fig. 8** shows the three metrics for CB6 open. Total travel is a solid metric. Delta Low Min/Max is not very useful for this fault, triggering a series of false alarms from the middle of the endurance test. T0, however, would be quite useful

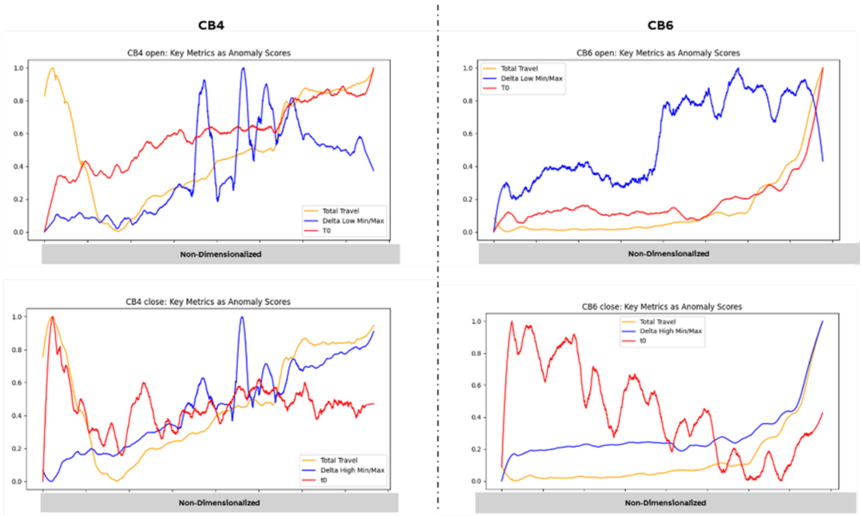
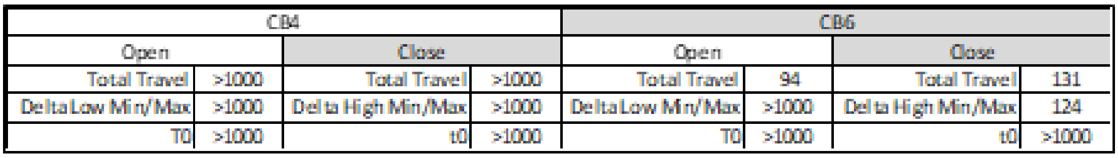


Figure 7: Normalized traditional metrics

Table 3
Trends in metrics-based anomaly detection

Metric	Cases with strong trend	Percentage of all cases
Open Total Travel	1	50%
Delta Low Min/Max	0	0%
T0	2	100%
Close Total Travel	1	50%
Delta High Min/Max	1	50%
t0	0	0%

Table 4
Metrics-based early warning in number of operations



CB4				CB6			
Open		Close		Open		Close	
Total Travel	>1000	Total Travel	>1000	Total Travel	94	Total Travel	131
Delta Low Min/Max	>1000	Delta High Min/Max	>1000	Delta Low Min/Max	>1000	Delta High Min/Max	124
T0	>1000	t0	>1000	T0	>1000	t0	>1000

if a few outliers in the beginning could be removed.

5. Conclusions and Future Work

In this paper, we have applied four anomaly detection algorithms to data from two endurance tests on medium voltage circuit breakers. We have compared the results to more traditional metrics. While there are limitations on the verification of unlabeled data, some assumptions have enabled us to show the strong performance of most of the algorithms we have tested.

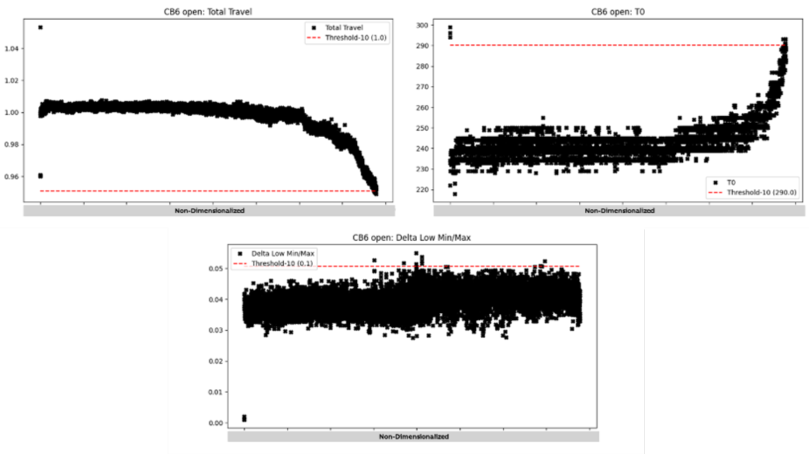


Figure 8: Metrics and thresholds for CB6 open

The best performance was given by KNN outperforming the deep model DeepSVDD. We come to the conclusion that anomaly detection is indeed quite suitable for condition monitoring of medium voltage breakers.

However, there is still room for improvement. Some early outliers on the anomaly scores indicate that more data cleaning might help to further improve the results. Also, the raw data as we used it is highly redundant. Techniques to reduce the dimensionality of the data by removing redundancy could be tested to better understand the unexpectedly poor performance of DeepSVDD. Another direction to explore is the combination of traditional metrics and anomaly detection algorithms.

Finally, we plan to run additional endurance tests and obtain test results from other research teams to further test the algorithms. We feel that more varies fault scenarios are needed to better assess and improve the algorithms.

References

1. Ahmad, Sabtain; Styp-Rekowski, Kevin; Nedelkoski, Sasho; Kao, Odej: Autoencoder-based Condition Monitoring and Anomaly Detection Method for Rotating Machines. In: Aluru, S. et al (Eds.): Proceedings of the IEEE Conference on Big Data, Atlanta, USA (2020)
2. Ali, R., Khan, M.; Kyung, C.: Self-Supervised Representation Learning for Visual Anomaly Detection: arXiv. Available online at doi:10.48550/ARXIV.2006.09654. (2020)
3. Charu C Aggarwal. Outlier analysis. In *Data mining*, 75–79. Springer (2015)
4. Choi, K., Yi, J., Park, C., & Yoon, S.: Deep Learning for Anomaly Detection in Time-Series Data: Review, Analysis, and Guidelines. IEEE Access, 120043–120065 (2021)
5. Di Mattia, F. Galeone, P., de Simoni, M., Emanuele, G.: A Survey on GANs for Anomaly Detection: arXiv. Available online at doi:10.48550/ARXIV.1906.11632. (2019)
6. Dix, M., Ganguly, S.: AI-based anomaly detection of asset failures in industrial process plants. atp magazin 64 (5), 60–67 (2022)

7. Hoffmann, Martin W., et al.: Integration of novel sensors and machine learning for predictive maintenance in medium voltage switchgear to enable the energy and mobility revolutions. In *Sensors* 20.7 (2020)
8. Kilian Hendrickx; Wannes Meert; Yves Mollet; Johan Gyselinck; Bram Cornelis; Konstantinos Gryllias; Jesse Davis: A general anomaly detection framework for fleet-based condition monitoring of machines. In *Mechanical Systems and Signal Processing* 139, p. 106585. (2020)
9. Li, C., Sohn, K., Yoon, J., Pfister, T.: CutPaste: Self-Supervised Learning for Anomaly Detection and Localization. In: Forsyth, D., Gkioxari, G., Tuytelaars, T., Yang, R., Yu, J. (eds.) Proceedings of the Conference on Computer Vision and Pattern Recognition (CVPR). Online Event (2021)
10. Li, Z.: ECOD: Unsupervised Outlier Detection Using Empirical Cumulative Distribution Functions. In: IEEE Transactions on Knowledge and Data Engineering (2023)
11. Malek Mahdavi, R,: Anomaly detection of angular travel curves of low voltage ABB circuit breakers. Master Thesis. Politecnico di Milano, Milan, Italy. Available online at https://www.politesi.polimi.it/handle/10589/213232 (2023)
12. Ramaswamy, Sridhar; Rastogi, Rajeev; Shim, Kyuseok: Efficient algorithms for mining outliers from large data sets. ACM SIGMOD Record Volume 29 Issue 2 (2000)
13. Ruff, L., Kauffmann, J. R., Vandermeulen, R. A., Montavon, G., Samek, W., & Kloft, M.: A Unifying Review of Deep and Shallow Anomaly Detection. Proceedings of the IEEE, 756–795 (2021)
14. Ruff, L., Vandermeulen, R., Goernitz, N., Deecke, L., Siddiqui, S., Binder, A., Müller, E., Kloft, M.: Deep One-Class Classification. In: Dy, J., Krause, A. (Eds.): Proceedings of the 35th International Conference on Machine Learning (ICML 2018). Stockholm, Sweden. IEEE, pp. 4393–4402. (2018)
15. Tack, J., Mo, S., Jeong, J., Shin, J. (2020): CSI: Novelty Detection via Contrastive Learning on Distributionally Shifted Instances.
16. Thi, Ngoc-Diem Tran; Do, The-Duong; Jung, Jae-Ryong; Jo, Hyangeun; Kim, Yong-Hwa: Anomaly Detection for Partial Discharge in Gas-Insulated Switchgears Using Autoencoder. In *IEEE Access* 8, pp. 152248–152257 (2020)
17. Zhang, Yuchen; Dong, Zhao Yang; Kong, Weicong; Meng, Ke: A Composite Anomaly Detection System for Data-Driven Power Plant Condition Monitoring. In *IEEE Transactions on Industrial Informatics* 16 (7), pp. 4390–4402 (2020)
18. Zhang, Yuchen; Li, Meng; Dong, Zhao Yang; Meng, Ke: Probabilistic anomaly detection approach for data-driven wind turbine condition monitoring. In *CSEE Journal of Power and Energy Systems* 5 (2), pp. 149–158. (2019)