scientific data

Check for updates

OPEN

DATA DESCRIPTOR

Dataset for Antenna-Based Detection of Fault Types in Covered Conductors for 22 kV Voltage Power Lines

Ondřej Kabot[1,3], Lukáš Klein[1,2,3] ✉, Lukáš Prokop[1], Stanislav Mišák[1] & Zdeněk Slanina[1,2]

This abstract presents a dataset for the detection of fault types in XLPE-covered conductors utilized in 22 kV medium voltage power distribution systems. We employed an antenna-based approach for detecting partial discharges. The dataset encompasses 12 distinct fault categories, ranging from ground phase faults to inter-phase faults, and no-fault case with steel or covered conductor as fault. We also used three different antennas. Each sample is a single measurement from antenna, consisting of 10^6 data points as floating numbers. The utilization of the antenna-based method offers the potential for a more cost-effective and straightforward installation for the detection of partial discharges. The objective of dataset is to enhance the identification of fault types, thereby promoting broader adoption of covered conductors in overhead power distribution lines. Such adoption proves particularly beneficial in confined areas, including natural parks, where safety is a prime concern. It is noteworthy that this dataset represents an original contribution, as no prior publication has addressed detection for this specific range of fault types and method of detection.

Background & Summary

Medium voltage power distribution lines play a pivotal role in ensuring the efficient and reliable distribution of electrical energy to consumers. In this context, the utilization of XLPE (Cross-Linked Polyethylene) covered conductors has emerged as a significant advancement due to several key advantages. Notably, XLPE covered conductors allow for the establishment of smaller safety zones around the power transmission infrastructure, mitigating potential hazards to the surrounding environment. Additionally, these conductors are designed to withstand short-term contact with vegetation without necessitating a complete shutdown of the power supply to customers. As this has potentially to save environment and finance of distributors operating the grid[1]. However, an inherent challenge arises when prolonged contact with vegetation occurs, giving rise to a phenomenon known as partial discharge (PD).

Partial discharge is characterized by low energy pulses within or on the surface the insulation of power distribution lines[2], and if left undetected or unaddressed, it can lead to significant insulation faults. The consequences of such faults are far-reaching and can include forest fires, power supply disruptions, and other potentially dangerous situations[3].

Efforts directed towards the detection of PDs have spurred the development of a multitude of systems, encompassing both contact galvanic methods[4] and contactless approaches utilizing antennas[5]. These systems hold considerable promise in terms of augmenting safety and bolstering reliability within various applications[6]. However, they frequently encounter limitations, such as elevated costs or incompatibility with large-scale production. One particularly promising avenue for PD detection is the utilization of radiometric antennas, as exemplified by the work of Kaziz *et al.*[7] and Martinovic *et al.*[8]. This approach stands out as potentially the most cost-effective method, obviating the need for power line shutdown during installation and maintenance.

Despite the ongoing development and scientific inquiry into the realm of PD detection and analysis, numerous opportunities for improvement persist[9,10]. The primary challenge lies in the detection of these discharges

[1]ENET centre - CEET, VSB - Technical University of Ostrava, Ostrava, Czech Republic. [2]Faculty of Electric Engineering and Computer Science, VSB - Technical University of Ostrava, Ostrava, Czech Republic. [3]These authors contributed equally: Ondřej Kabot, Lukáš Klein. ✉e-mail: lukas.klein@vsb.cz

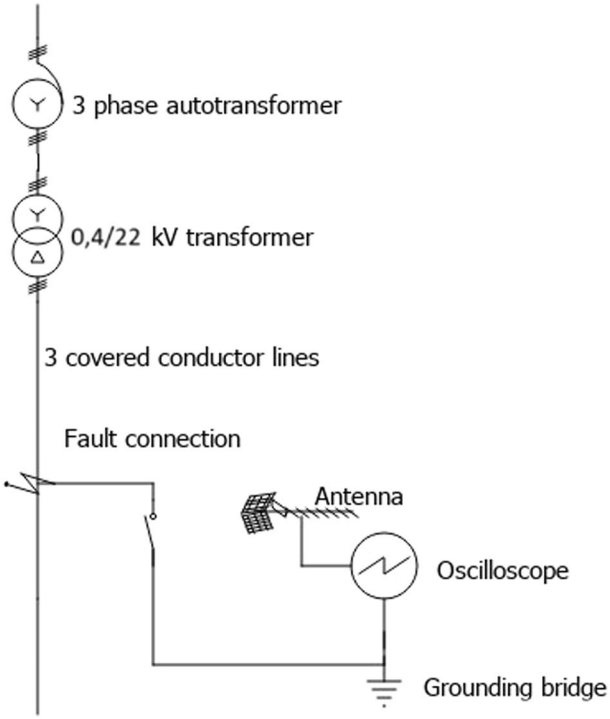

Fig. 1 1 pole schematic.

due to the inherently low energy of the pulses, which heightens the probability of some pulses being masked by ambient noise[11].

This research article introduces a valuable resource in the form of a meticulously annotated dataset tailored for the detection of various fault types in XLPE-covered conductors used in medium voltage power distribution lines. The dataset was specifically curated to support the enhancement of antenna-based detection systems. It encompasses comprehensive data collected using antennas, capturing five distinct fault classes as well as background measurements. We captured signal from antenna, that consisted of 20 ms, corresponds to utility frequency of power grid. Each sample is a single continuous measurement from antenna captured by spectral analyser with high sampling rate. Each sample has 10^7 data points as floating numbers. To ensure the dataset's robustness, precautions were taken to minimize the impact of background noise. This entailed systematically altering the fault type every seventh measurement and collecting data over two separate days.

The identification and classification of fault types constitute a crucial aspect in the evaluation of potential issues within the power transmission infrastructure. This undertaking serves a dual purpose: firstly, it aids in the real-time detection of faults as they manifest, and secondly, it provides a means to precisely characterize the nature of these faults. Consequently, this contributes to the development of more accurate and effective maintenance and repair strategies. This advancement holds paramount significance in the promotion of the utilization of XLPE-covered conductors in overhead power distribution lines, thereby concurrently enhancing both safety and operational efficiency.

It is worth noting that a different but similar data were utilized in a related study titled "Enhanced Fault Type Detection in Covered Conductors Using Stacked Ensemble and Novel Algorithm Combination"[12]. Also in mentioned article only subset of fault types were captured and only with one antenna. In the aforementioned study, a stacking ensemble methodology was applied, incorporating a meta-learner founded on logistic regression, along with five distinct individual components utilizing two diverse algorithms, namely, XGBoost and MiniRocket. This was done to further exemplify the practical applications and significance of the dataset and show validity of the approach in order to distinguish fault types in distribution power transsmision lines.

Furthermore, it is imperative to underscore the exceptional nature of this dataset. To date, there exists no comparable dataset featuring annotated fault type data acquired through antenna-based methodologies. In addition to this dataset, our research group has produced an alternate dataset[13] utilizing antenna-based methodologies. However, this alternative dataset is devoid of fault type annotations, and its data are directly sourced from real-world environments in a raw, unprocessed format. Nonetheless, there is potential for synergistic utilization of these two distinct datasets in the development of innovative algorithms for the detection of various fault types.

Methods

In this section, we will provide an extensive account of the methodologies utilized for data acquisition. Furthermore, we will present a detailed exposition of the procedures involved in the establishment of the experimental framework and the subsequent data collection process.

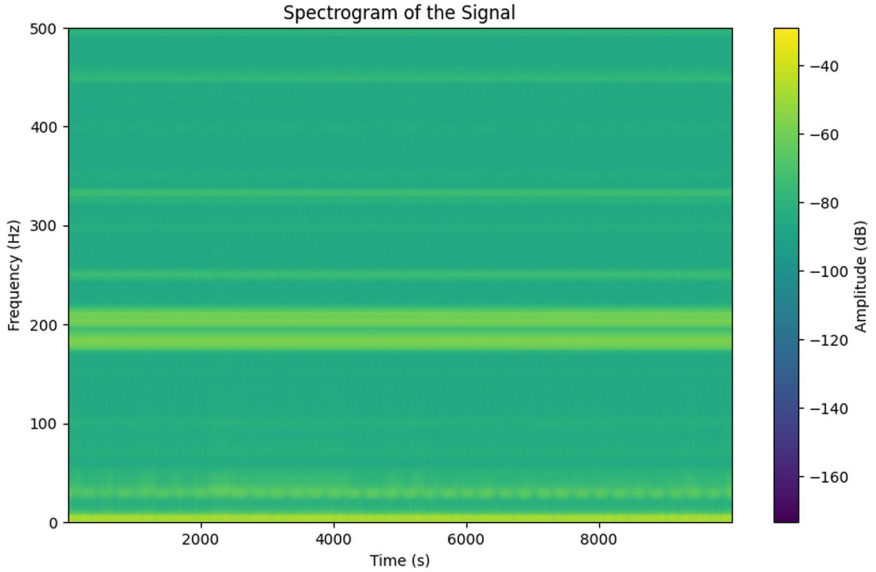

Fig. 2 Spectogram of a single sample with PDs.

Description of Experimental Setup. To facilitate the detection of PDs, an antennas were employed. The orientation of the antennas were perpendicular (as can be seen in Figs. 4 and 5) to the power lines and positioned at a distance of one meter from them. Subsequently, the antennas were integrated with an oscilloscope, serving as the platform for data acquisition. Three different antennas were used: the Bonito BONI-WHIP, the MiniWhip, and the 433MHz TX433-XP-200 antenna, to demonstrate the differences in PD pattern capture using each antenna. These antennas were subsequently linked to 50 Ohm inputs of an oscilloscope Siglent SDS-5034X (manufacturer Siglent, the Netherlands), serving as the data acquisition platform. The data acquisition platform can be seen in Figs. 12 and 6.

BONI-WHIP. The BONI-WHIP active antenna[14] succeeds the popular Mini-Whip antenna, offering remarkable reception capabilities across longwave, mediumwave, shortwave, and VHF frequencies up to 300 MHz, despite its compact 17 cm length. Compared to its predecessor, it introduces significant enhancements and improved performance. Enclosed in a weather-resistant housing, the BONI-WHIP features a robust active antenna circuitry and receives power through the antenna cable, which includes the necessary power supply module.

In summary, key specifications of the BONI-WHIP include a frequency range of 20 kHz - 300 MHz, a voltage supply of 12-15 V, a gain of 3 dB, an upper frequency limit (-1 dB) of 300 MHz, and IP3 (+32.5 dBm) and IP2 (+55 dBm) values. These specifications make the BONI-WHIP an adaptable and high-performance choice for various radio reception applications.

TX433-XP-200 antenna. The TX433-XP-200 antenna, known as a high-gain suction antenna, is a versatile and reliable choice for various applications requiring excellent performance from a compact antenna. This antenna is available for a wide range of frequencies including 170 MHz, 230 MHz, 433 MHz, 490 MHz, 780 MHz, 868 MHz, 916 MHz, and 2.4 GHz, making it a flexible solution for different usage scenarios.

Key specifications of the TX433-XP-200 include a gain of 3.5 dBi, ensuring enhanced signal strength. It features an SMA male connector with an external thread, internal thread, and external pin, providing secure and stable connectivity. The antenna itself measures 18.5 cm in length, and it comes with a 200 cm cable for convenient installation.

In terms of performance, the TX433-XP-200 boasts a standing wave ratio (SWR) of less than 1.5, ensuring efficient transmission and minimal signal loss. Operating at a frequency of 433 MHz, it has an impedance of 50 Ohms, which is standard for many RF applications. This combination of features makes the TX433-XP-200 an excellent choice for high-performance radio frequency applications.

MiniWhip. The MiniWhip active antenna amplifies the potential difference signal between a small metal plate and ground, then transmits the amplified signal to the receiver through the core and shield of a coaxial cable. This design ensures robust signal reception across a broad frequency range.

Key features of the MiniWhip include its good contact performance for stable operation, a mini design with an easy-to-connect SMA female output interface, and a frequency range of 10 kHz to 30 MHz. The power supply range is DC9-15 V, typically 12 V, with a maximum output power exceeding -15 dBm.

Technical Specifications

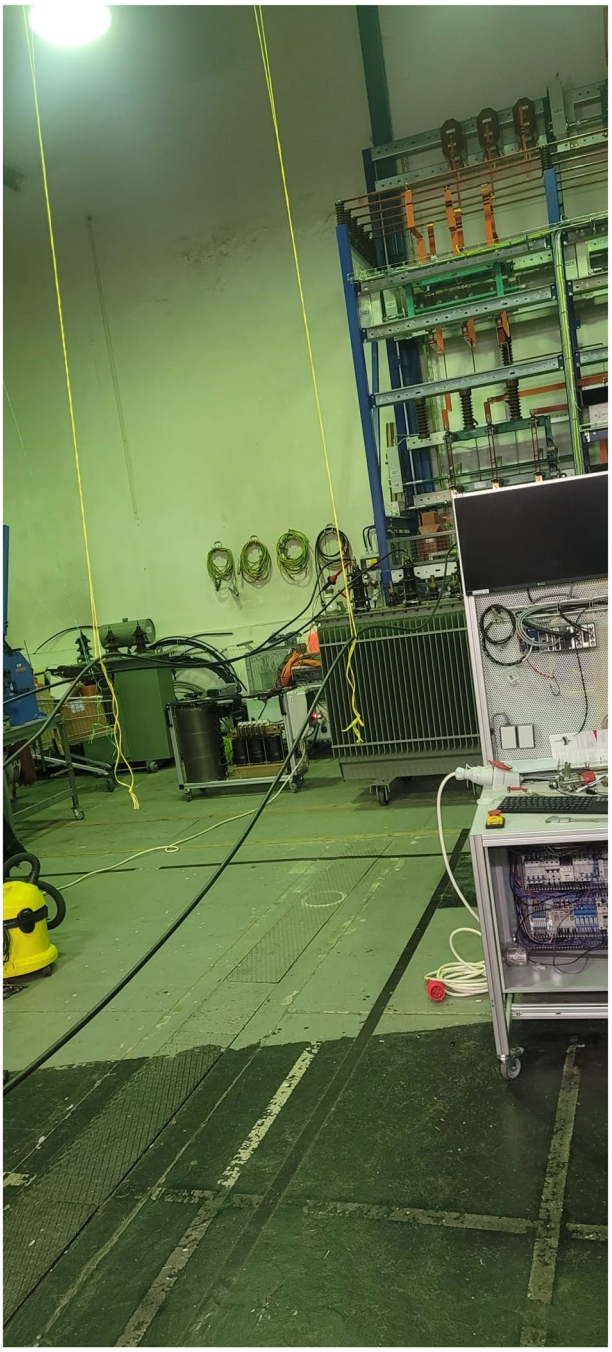

Fig. 3 Experimental setup featuring a transformer and the initial segments of electrical lines.

- **Name:** Active Antenna
- **Material:** Printed Circuit Board (PCB)
- **Working Frequency Range:** 10 kHz - 30 MHz
- **Power Supply Range:** DC9-15 V, typically 12 V
- **Maximum Output Power:** > -15 dBm
- **Output Interface:** SMA female

Measurements. The measurements for the intended purpose were conducted under conditions replicating real-life scenarios. The sole variable in this experimental setup pertained to the elevation of individual lines relative to the ground. The measurement apparatus comprised an adjustable 3-phase auto-transformer connected to the low-voltage side of a 0.4/22 kV oil insulated transformer. Following this, the high-voltage output of the transformer was connected to the individual insulated conductor lines, as depicted in Fig. 1. The voltage was set to 22 kV as in medium voltage power distribution lines. For this purpose, we employed XLPE insulated

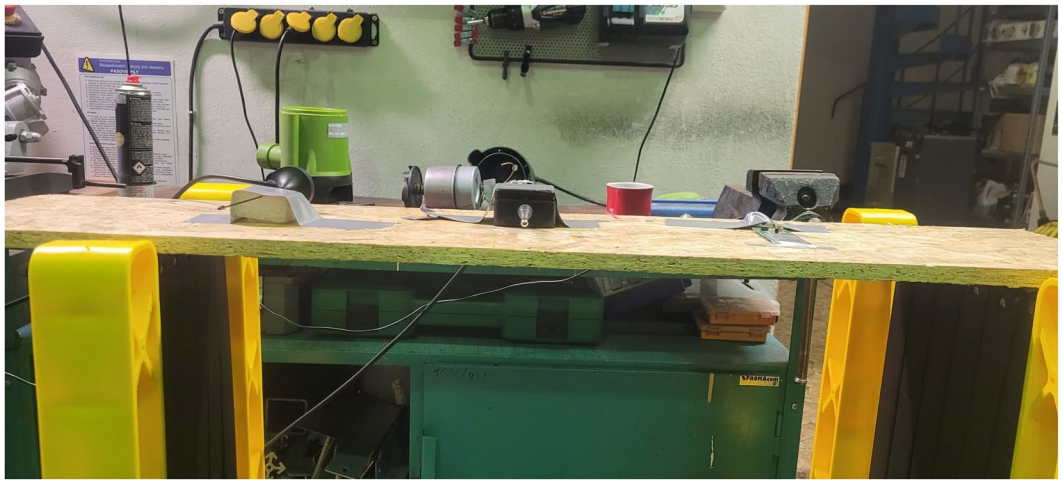

Fig. 4 Three antennas capturing signals are mounted on a wooden board supported by yellow brackets, with one antenna centrally placed and the other two positioned to the right.

conductors with crossection of 35 mm2 consistent with those utilized in actual power distribution lines in the Czech Republic.

To secure each line, we hanged the lines in air, and stainless steel spheres were positioned at both terminals of the line to establish a geometrically graded electrical field, thereby preventing undesired occurrences of partial discharges. Insulators employed at this measurement were porcelain pin type insulators that are used in real environment. A similar methodology was applied to the high-voltage terminals of the 22 kV transformer, meaning the screws which are the main connection points were equipped with semi-conductive tape and stainless steel spheres. To ensure accurate installation, an assessment was conducted utilizing a UVC-sensitive camera (sample measurment can be seen Fig. 7), revealing the absence of undesirable PD incidents during the experiments.

Fault Types. In this subsection, we delineate the primary categories of faults utilized in our experimental setup: using a stainless steel tube, employing a covered conductor, and connecting lines together. The incorporation of a tree branch was excluded from our experiments to maintain consistent and reproducible conditions, as natural elements introduce excessive variability at each contact point. To address the issue of partial discharges, which typically occur due to high electrical field intensities near sharp metal edges, we attached stainless steel spheres at both ends of the tube. This modification aims to geometrically grade the electrical field, thereby mitigating the risk of discharge. The assembled unit was then strategically placed across the selected phases and grounded, especially during tests that involved faults associated with ground potential. Figures 9, 3, and 13 provide visual representations of the setup and fault types. Different combinations of faults were tested and measured. With and without grounding:

1. With grounding - can not be done for joined together (CC) fault class
2. Without grounding - can not be done when only one power line is utilized

Different fault approach:

1. Stainless steel tube (LLS or LGS)
2. Covered conductor (LLCC or LGCC)
3. Joined together (CC)

Stainless Steel Tube. To simulate a specific fault scenario, we employed a stainless steel tube (LLS or LGS) (Fig. 9), configured to span sufficiently across all three phases, thereby establishing an electrical connection among them. This setup facilitates the study of faults that mimic conditions where conductive objects accidentally contact high-voltage lines, potentially leading to phase-to-phase or phase-to-ground faults.

Covered Conductor. The second fault type involved the use of a covered conductor (LLCC or LGCC), designed to emulate scenarios where insulated conductive materials come into contact with live wires. This configuration is particularly useful in studying faults that occur in densely populated or vegetated areas where insulated or partially insulated conductive objects might interact with electrical infrastructure. Details of the experimental setup for this fault type are illustrated in Fig. 10.

Joined Together. The final category of faults examined involves directly joining two or three lines together (CC shortname). This setup was exclusively tested without grounding, as grounding is not feasible in such configurations. By simulating these conditions, the experiments aim to replicate potential real-world occurrences where

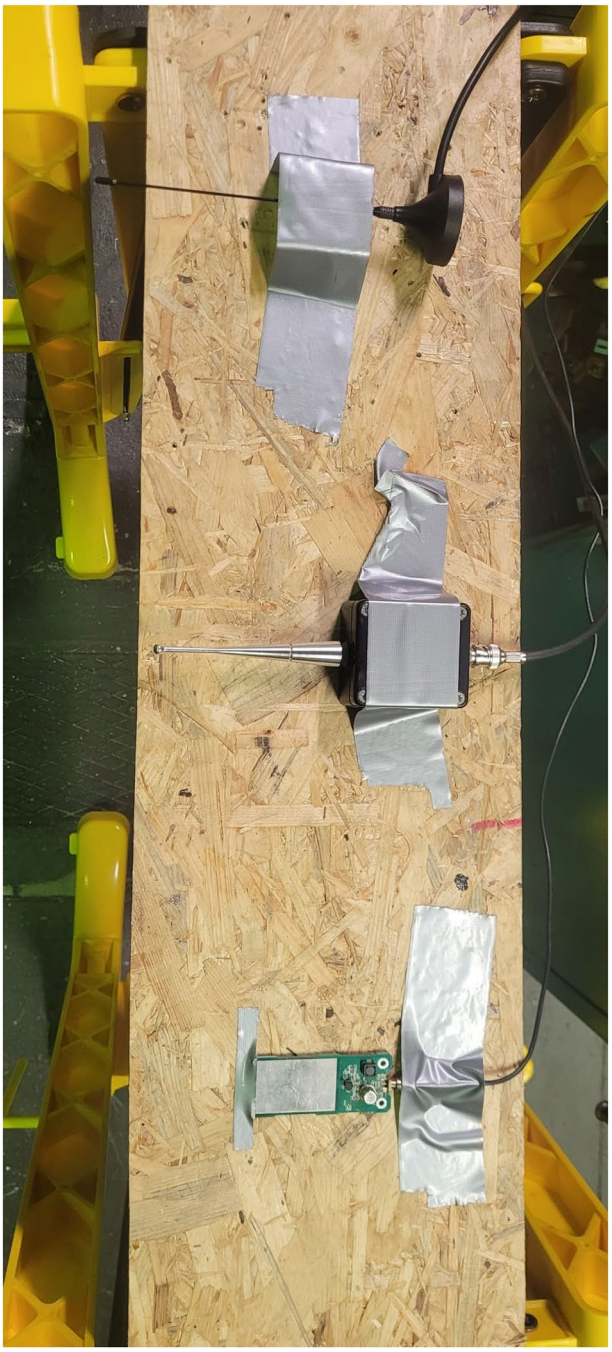

Fig. 5 Three antennas from above.

multiple power lines may come into contact due to environmental factors or mechanical failures. The outcomes of these experiments are critical for developing algorithms capable of effectively discerning between different types of electrical faults. Figures 11 and 8 depict the arrangements for these experiments.

In each scenario (besides joining power lines together), tests were conducted with and without grounding. The configurations varied—ranging from the stainless steel tube covering two or three lines, to similar arrangements for the covered conductor. The joined together fault type was consistently tested in configurations involving either two or three lines, emphasizing the diversity of fault scenarios that can occur in real-world power infrastructure. These experimental designs are intended to enhance the reliability and specificity of fault detection algorithms in power line monitoring systems.

Testing on Covered Conductors. A systematic approach was diligently employed to ensure the consistency and reliability of the results obtained. We captured the same set of 7 faults for each testing run, followed by the introduction of a different type of fault, while taking background measurements between these changes in fault types. This methodological choice was made with the aim of ascribing any incidental background noise or

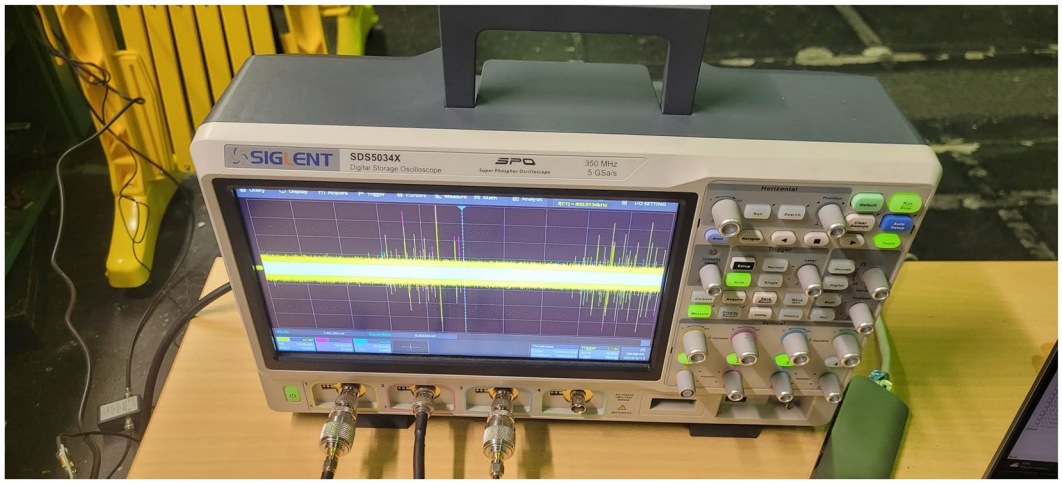

Fig. 6 The digital storage oscilloscope used for capture.

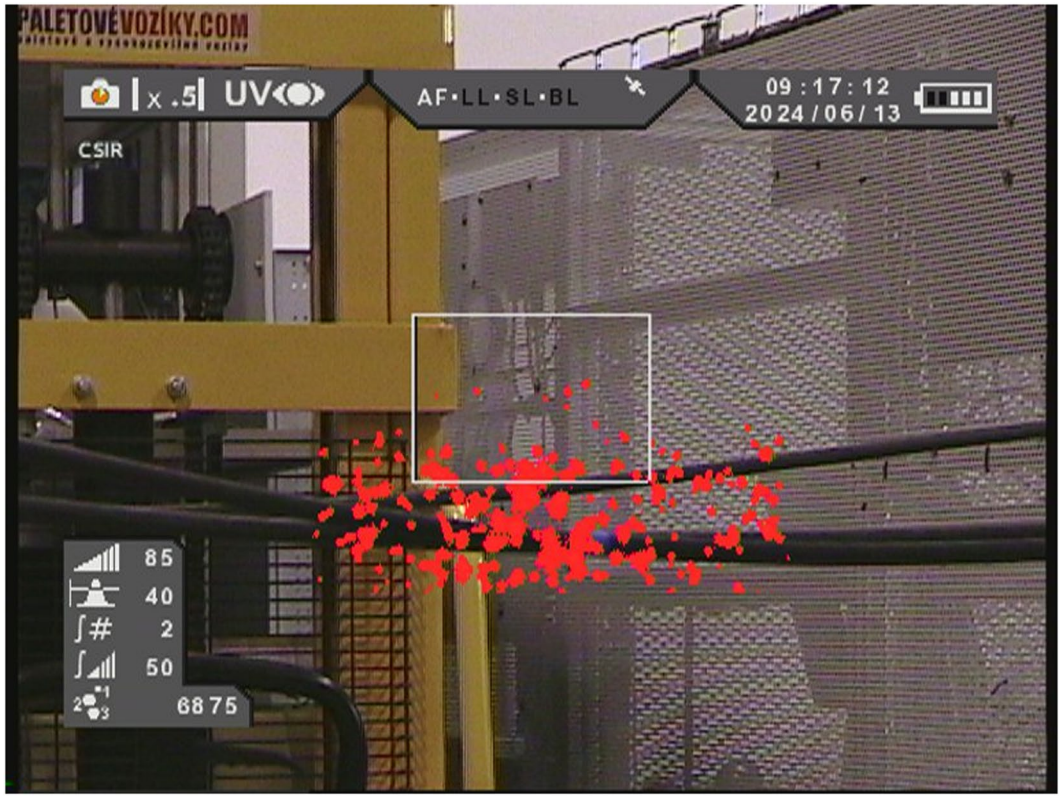

Fig. 7 Sample image illustrating the validation process (fault case) using a UVC camera to detect partial discharges, with red markers indicating PDs.

interference to a specific fault type, thereby mitigating the potential impact of random noise on the measured data, which could negatively influence results of our experiment.

To mitigate the potential influence of extraneous variables, such as variations in weather conditions and diurnal fluctuations, which have the potential to impact research outcomes, the testing protocol spanned a temporal window spanning half a day, distributed across two different days. This temporal partitioning was deliberately implemented to encompass data acquisition on two distinct days, thereby augmenting the robustness of the investigative study. The comprehensive examination encompassed the assessment of 12 distinct fault categories, concomitant with the acquisition of baseline measurements obtained under normal operational circumstances with and without power line powered.

Over the course of the experimental procedure, each individual sample underwent meticulous recording for a precisely defined duration of 20 milliseconds. This recording process was executed at a notably high sampling

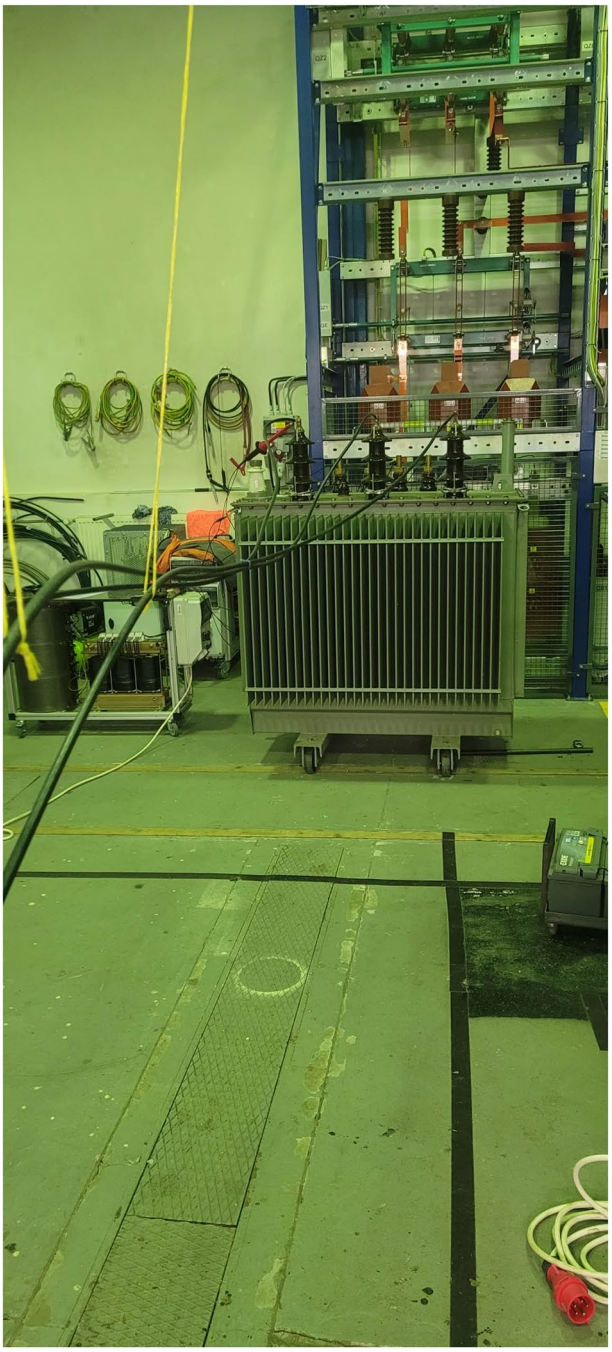

Fig. 8 The fault with two connected lines together.

frequency of 500 MHz, thereby affording the capacity to capture an entire cycle of the utility frequency. This elevated sampling frequency was instrumental in facilitating an intricate and precise analysis of the fault conditions under scrutiny.

Data Records

In this section we provide description of the data set.

Dataset Organization. The dataset is publicly available on the Fiqshare repository[15] and is presented as single file. The data file is stored in the form of compressed zip archive. Within archive, data is organized in a flat folder structure, containing measurements captured. Each file has it name consisting of channel (corresponding to specific antenna), type of the fault and sample number.

In Table 1, we present a comprehensive listing of the classes included in the dataset, along with their corresponding assigned names. Additionally, a single sample filename from the dataset is provided as an illustrative example.

Name	Description	Shortname	Example Filename (.bin)
Lines Together			
2 Covered Conductors	Configuration with 2 covered conductors together	2CC	C3_2CC_04.bin
3 Covered Conductors	Configuration with 3 covered conductors together	3CC	C3_3CC_05.bin
Lines and Steel Pole			
1 Line, Ground, Steel Pole	Configuration with 1 line, grounding, and a steel pole	1LGS	C3_1LGS_07.bin
2 Lines, Steel Pole	Configuration with steel pole over 2 lines	2LLS	C3_2LLS_01.bin
2 Lines, Ground, Steel Pole	Configuration with steel pole over 2 lines with grounding	2LGS	C3_2LGS_06.bin
3 Lines, Ground, Steel Pole	Configuration with steel pole over 3 lines with grounding	3LGS	C3_3LGS_21.bin
3 Lines, Steel Pole	Configuration with steel pole over 3 lines	3LLS	C3_3LLS_03.bin
Lines Connected and Covered Conductor			
2 Lines, Covered Conductor	Configuration with covered conductor 2 lines over them	2LLCC	C3_2LLCC_02.bin
3 Lines, Covered Conductor	Configuration with 3 lines using a covered conductor over them	3LLCC	C3_3LLCC_20.bin
2 Lines, Ground, Covered Conductor	Configuration with 2 lines, grounding, and a covered conductor over them	2LGCC	C3_2LGCC_08.bin
3 Lines, Ground, Covered Conductor	Configuration with 3 lines, grounding, and a covered conductor over them	3LGCC	C3_3LGCC_09.bin
1 Line, Ground, Covered Conductor	Configuration with 1 line, grounding, and a covered conductor over them	1LGCC	C3_1LGCC_10.bin
Background Conditions			
Background (no power)	Background condition without power	BG	C3_BG_11.bin
Background 2 (no power)	Another background condition without power	BG2	C3_BG2_12.bin
Background 3 (no power)	A third background condition without power	BG3	C3_BG3_13.bin
Background, High Voltage	Background condition with high voltage applied	BGHV	C3_BGHV_14.bin
Background, High Voltage 2	Another background condition with high voltage applied	BGHV2	C3_BGHV2_15.bin

Table 1. Description of Fault Classes for Partial Discharge Detection in High Voltage Power Lines.

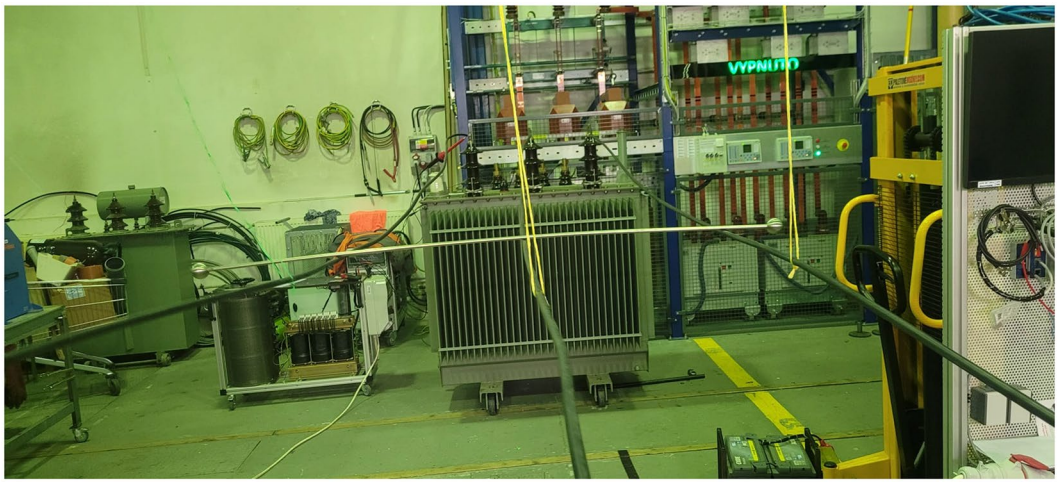

Fig. 9 Fault using steel pole.

File Naming Convention

Each file name follows the pattern: `C#_XYZ_NN.bin`

- **First Part (C#):** This indicates the type of antenna used. It can be one of the following:
  - `C1`: MiniWhip
  - `C2`: BONI-WHIP
  - `C3`: Passive Antenna TX433-XP-200
- **Middle Part (XYZ):** This indicates the fault type. The specific codes for fault types can vary.
- **Last Part (NN):** This is the sample number, which is a two-digit identifier for the specific sample <1-28>.

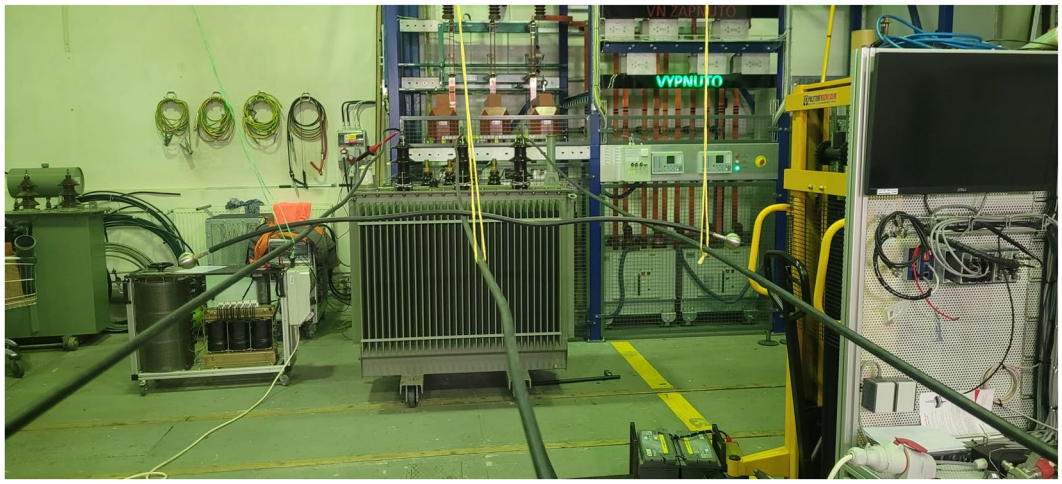

Fig. 10 Fault using covered conductor.

Examples

- `C3_2CC_04.bin`
  - `C3`: Passive Antenna TX433-XP-200
  - `2CC`: 2 Lines together
  - `04`: Sample number 4
- `C3_3LLS_03.bin`
  - `C3`: Passive Antenna TX433-XP-200
  - `3LLS`: Stainless steel tube over 3 power lines withou grounding
  - `03`: Sample number 3

Dataset Description. The dataset consists of samples from 14 different classes, comprising 12 types of faults and two types of background. The background class (BG) comprises 84 samples without powered power line and 56 with powered power line, while each of the fault classes contains 28 samples. The fault classes are categorized in Table 1. In the Table 2 can be seen the count of measures for each class.

Each sample within the dataset consists of 10,000,000 numerical data points, encompassing a range from -1 to 1. These numeric values represent measurements taken from the electrical system under various fault conditions. The dataset can be categorized into two primary subsets: fault samples, which consist of 336 samples for each antenna, and samples devoid of faults, referred to as background samples, totaling 140 samples for each antenna .

An illustration of a spectrogram derived from a signal is depicted in Fig. 2. Figure 14 displays a representative sample containing a fault, wherein characteristic peaks within the signal are discernible. Conversely, Fig. 15 exhibits a sample devoid of any faults.

The dataset has been formatted and stored in the original binary file format (`.bin`), providing a convenient means for data access and manipulation. Consequently, we highly recommend the preprocessing of this dataset to enhance its suitability for effective utilization. It is noteworthy that the entire uncompressed dataset spans a substantial size of 14.2 gigabytes.

Alternatively we also provide script to convert binary data to data in numpy array format (.npy), which is smaller (39 megabytes) and also can be more convenient to use with certain languages.

Technical Validation

In the present study, we have conscientiously devised an experiment to ensure the absence of any issues associated with the provided data. It is imperative to underscore that, to the best of our knowledge, no potential issues have been identified thus far.

To strengthen the credibility of our experiment, we drew from a previous study ([12]). While the prior study primarily centered on algorithmic approaches without extensive data analysis, our work reaffirms and extends these findings. Specifically, we found substantial differences among fault types using similar algorithmic techniques as those employed previously. This shows that the data are valid.

Our commitment to data integrity extended to a manual validation procedure, wherein we meticulously assessed the accuracy of the dataset. Notably, we ascertained that the majority of samples containing faults exhibited discernible peaks within the signal, indicative of faults within the covered conductors. Crucially, no

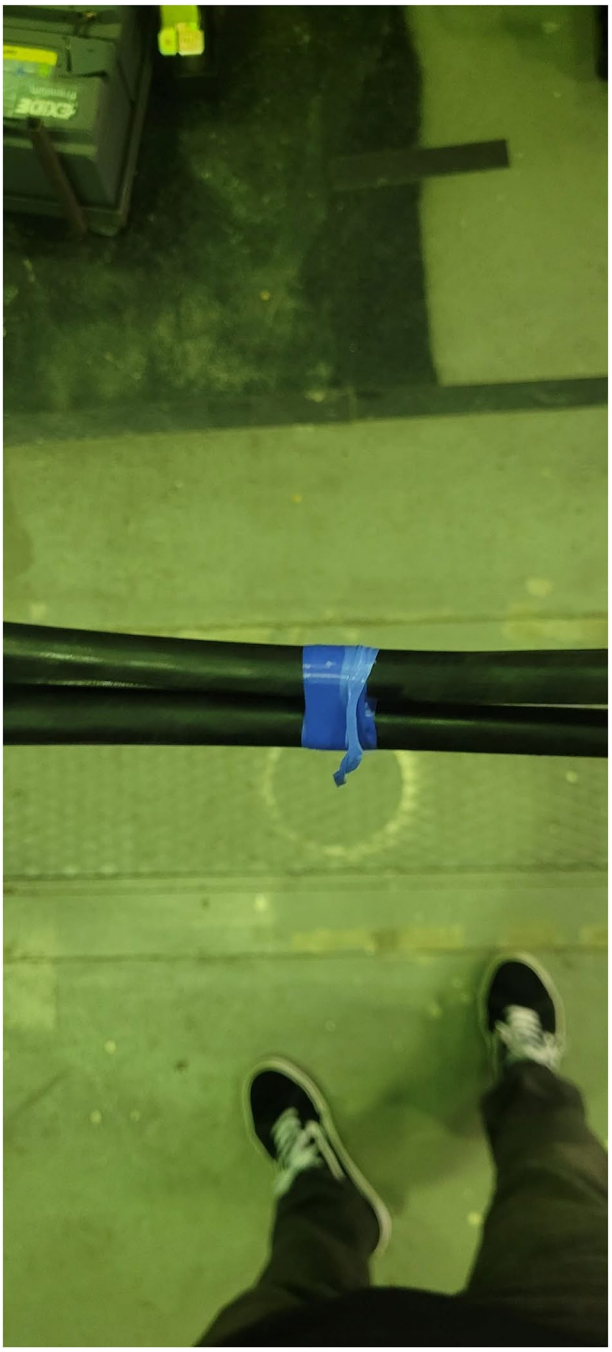

Fig. 11 The fault with two connected lines together from above.

such peaks were discerned in the background, a phenomenon signifying the absence of partial discharges or other extraneous sources during non-fault conditions.

Additionally, our investigative efforts encompassed an examination of potential sources of significant background noise, yielding the determination that no prominent artificial noise sources were detected within the dataset.

Usage Notes

This dataset is meticulously crafted to cater to supervised multi class classification tasks. In the realm of supervised learning, both the input data points and their corresponding class labels are supplied, thereby enabling the training of classification models. Multi class classification entails the assignment of a singular class label to each input sample drawn from a defined set of multiple classes. In the context of this dataset, these classes correspond to distinct fault types as well as the background class.

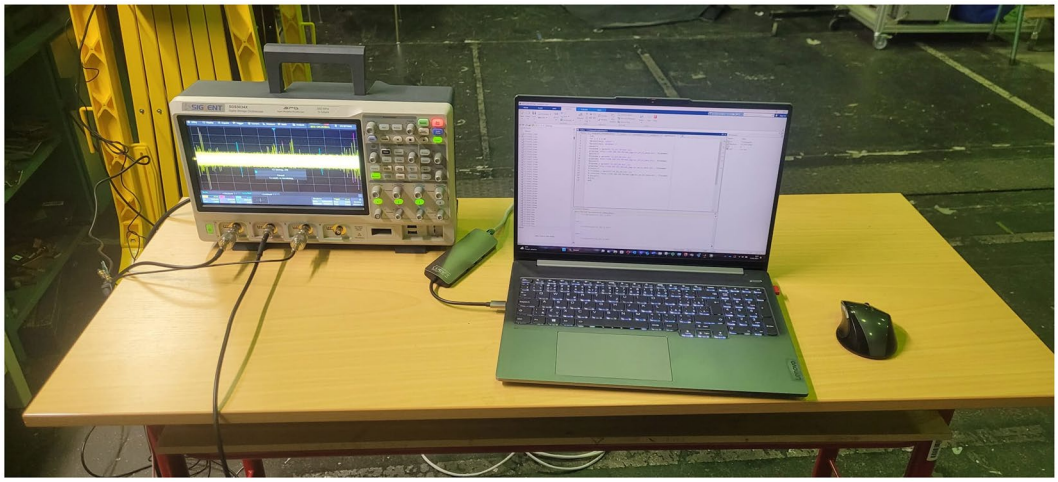

Fig. 12 The measurement setup.

Fig. 13 Experimental setup featuring the terminating segments of electrical lines.

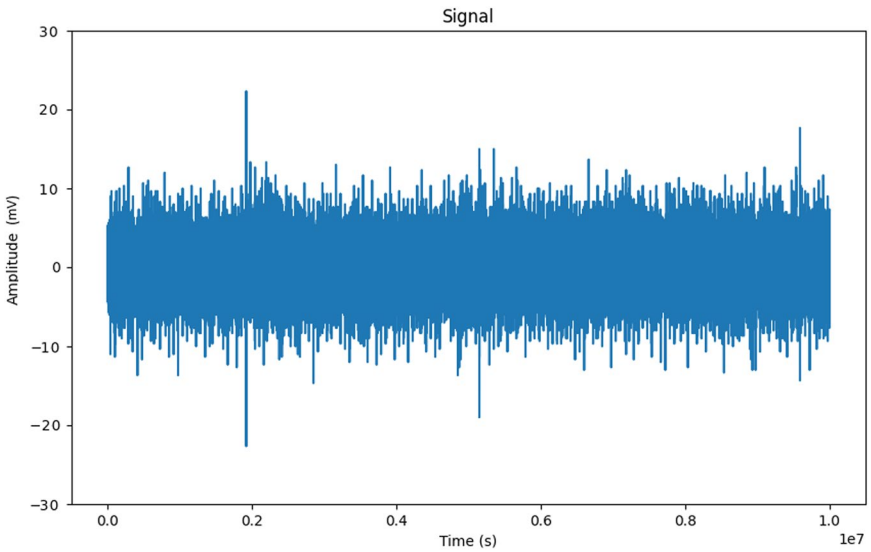

Fig. 14 Single sample with PDs (fault).

Class	Shortname	Number of Samples for single antenna
2 Lines, Steel Pole	2LLS	28
2 Lines, Covered Conductor	2LLCC	28
3 Lines, Steel Pole	3LLS	28
3 Lines, Covered Conductor	3LLCC	28
2 Covered Conductors	2CC	28
3 Covered Conductors	3CC	28
2 Lines, Ground, Steel Pole	2LGS	28
3 Lines, Ground, Steel Pole	3LGS	28
1 Line, Ground, Steel Pole	1LGS	28
2 Lines, Ground, Covered Conductor	2LGCC	28
3 Lines, Ground, Covered Conductor	3LGCC	28
1 Line, Ground, Covered Conductor	1LGCC	28
Background (no power)	BG	28
Background 2 (no power)	BG2	28
Background 3 (no power)	BG3	28
Background, High Voltage	BGHV	28
Background, High Voltage 2	BGHV2	28
Total		476

Table 2. Data Set Summary for Partial Discharge Detection.

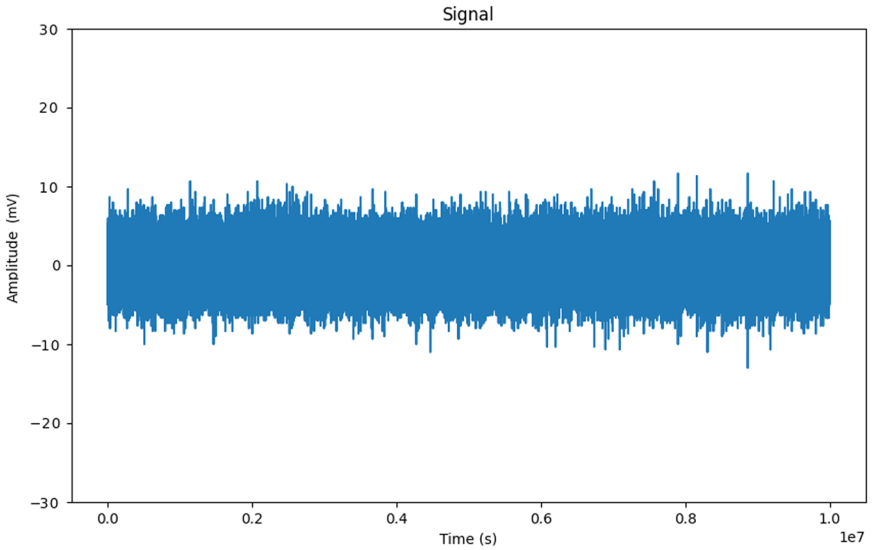

Fig. 15 Single sample without PDs (bgn).

The utilization of the dataset is straightforward, with each measurement provided as an individual file. For optimal data analysis, we recommend employing either MATLAB or Python in conjunction with the SciPy library. These software tools offer robust support for loading the dataset files.

Given the substantial size of the dataset, it is advisable to consider the implementation of downscaling techniques, such as max pooling. However, the specific choice of downsampling methods should be left to the discretion of the researcher, as these techniques can assist in managing the computational demands associated with the dataset.

Second option of using numpy format is more straight forward as user can use Python and Numpy library and load the data directly.

Data availability

We provide basic Python scripts for working with data, converting to .npy and summaries on the https://github.com/Lukykl1/dataset-fault-type.

Received: 31 July 2024; Accepted: 14 October 2024;
Published online: 04 December 2024

References

1. Talaat, M., El-Shaarawy, Z., Tayseer, M. & El-Zein, A. An economic study concerning the cost reduction of the covered transmission conductors based on different optimization techniques. *Results in Engineering* **11**, 100262 (2021).
2. Kabot, O., Fulneček, J., Mišák, S., Prokop, L. & Vaculík, J. Partial discharges pattern analysis of various covered conductors. In *2020 21st International Scientific Conference on Electric Power Engineering (EPE)*, 1–5 (IEEE, 2020).
3. Lehtonen, M. Fault rates of different types of medium voltage power lines in different environments. In *Proceedings of the 2010 Electric Power Quality and Supply Reliability Conference*, 197–202, https://doi.org/10.1109/PQ.2010.5549998 (2010).
4. Misak, S., Fulnecek, J., Jezowicz, T., Vantuch, T. & Burianek, T. Usage of antenna for detection of tree falls on overhead lines with covered conductors. *Advances in Electrical and Electronic Engineering* **15**, https://doi.org/10.15598/aeee.v15i1.1894 (2017).
5. Fulnecek, J. & Misak, S. A simple method for tree fall detection on medium voltage overhead lines with covered conductors. *IEEE Transactions on Power Delivery* **36**, 1411–1417, https://doi.org/10.1109/tpwrd.2020.3008482 (2021).
6. Chiu, B., Roy, R. & Tran, T. Wildfire resiliency: California case for change. *IEEE Power and Energy Magazine* **20**, 28–37 (2022).
7. Kaziz, S. *et al*. Radiometric partial discharge detection: A review. *Energies* **16**, 1978 (2023).
8. Martinovic, T. & Fulnecek, J. Fast algorithm for contactless partial discharge detection on remote gateway device. *IEEE Transactions on Power Delivery* 1–1, https://doi.org/10.1109/tpwrd.2021.3104746 (2021).
9. Chan, J. Q., Raymond, W. J. K., Illias, H. A. & Othman, M. Partial discharge localization techniques: A review of recent progress. *Energies* **16**, 2863 (2023).
10. Song, Y. *et al*. Online multi-parameter sensing and condition assessment technology for power cables: A review. *Electric Power Systems Research* **210**, 108140 (2022).
11. Uwiringiyimana, J. P., Khayam, U. & Montanari, G. C. *et al*. Comparative analysis of partial discharge detection features using a uhf antenna and conventional hfct sensor. *IEEE Access* **10**, 107214–107226 (2022).
12. Kabot, O., Klein, L., Prokop, L. & Walendziuk, W. Enhanced fault type detection in covered conductors using a stacked ensemble and novel algorithm combination. *Sensors* **23**, https://doi.org/10.3390/s23208353 (2023).
13. Klein, L. *et al*. A data set of signals from an antenna for detection of partial discharges in overhead insulated power line. *Scientific Data* **10**, https://doi.org/10.1038/s41597-023-02451-1 (2023).
14. Dennis. Boni-whip (2020).
15. Kabot, O., Klein, L., Prokop, L., Mišák, S. & Slanina, Z. Dataset for antenna-based detection of fault types in covered conductors for 22 kv voltage power lines, https://doi.org/10.6084/m9.figshare.26946910 (2024).

Acknowledgements

This article has been produced with the financial support of the European Union under the REFRESH - Research Excellence For Region Sustainability and High-tech Industries project number CZ.10.03.01/00/22_003/0000048 via the Operational Programme Just Transition.

Author contributions

The project was a collaborative effort, with each team member taking on specific roles and responsibilities. In terms of conceptualization, Lukáš Prokop and Ondřej Kabot played pivotal roles, shaping the project's overall direction. Methodology was primarily led by Lukáš Klein and Ondřej Kabot, who devised the research approaches. Lukáš Klein also took charge of the software development aspect. The validation process was a joint effort by Lukáš Prokop and Ondřej Kabot, ensuring the accuracy and reliability of the data. The writing and documentation tasks were undertaken by Lukáš Klein and Ondřej Kabot, crafting the project's reports and findings. Lastly, Lukáš Prokop provided supervision and guidance throughout the project. Additionally, Zdeněk Slanina contributed significantly to data validation and data gathering, further enhancing the project's robustness and credibility.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to L.K.