DISCLAIMER

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof. Reference herein to any social initiative (including but not limited to Diversity, Equity, and Inclusion (DEI); Community Benefits Plans (CBP); Justice 40; etc.) is made by the Author independent of any current requirement by the United States Government and does not constitute or imply endorsement, recommendation, or support by the United States Government or any agency thereof.

PNNL-34821
M3LW-23OR0404033

Light Water Reactor Sustainability Program

Spread Spectrum Time Domain Reflectometry (SSTDR) and Frequency Domain Reflectometry (FDR) for Detection of Cable Anomalies Using Machine Learning

September 2023

U.S. Department of Energy

Office of Nuclear Energy

PNNL-34821
M3LW-23OR0404033

Spread Spectrum Time Domain Reflectometry (SSTDR) and Frequency Domain Reflectometry (FDR) for Detection of Cable Anomalies Using Machine Learning

S.W. Glass, J.R. Tedeschi, M.P. Spencer,
M. Elen, J. Son, M. Taufique, D. Li, L.S. Fifield
Pacific Northwest National Laboratory

J.A. Farber and A.Y. Al Rashdan
Idaho National Laboratory

September 2023

Prepared for the
U.S. Department of Energy
Office of Nuclear Energy

SUMMARY

Cables are initially qualified for nuclear power plant use for 40 years. As plants extend their operating license to 60 and 80 years, continued use of these cables must shift to a performance-based approach since it is cost prohibitive to completely replace cables that are likely still capable of performing their design function. A variety of cable tests are available and are commonly applied during outages when the cables can be taken out of service.

Frequency domain reflectometry (FDR) is one of these test methods that is being more broadly accepted and used because it not only detects anomalies along the cable with a low-voltage signal that does not stress the cable insulation, but the technique also locates the anomalies. This supports follow-up local inspection and local repair or partial replacement of a damaged cable segment. Currently, FDR testing is only applied to cables that are taken out of service since the test instrument would be damaged by operational voltages.

A related technology that has found some acceptance in the aircraft and rail industry is spread spectrum time domain reflectometry (SSTDR). This technology has been implemented with a custom commercial instrument by LiveWire Innovation Inc. that is designed to operate on live cables up to 1000 volts.

PNNL performed a test of the LiveWire SSTDR system and an FDR system in 2022 using their Accelerated and Real-time Environmental Nodal Assessment (ARENA) cable motor test bed, that allowed many conditions to be evaluated under realistic and controlled conditions without risking safety critical systems of an actual power plant. The LiveWire system clearly showed its best response at the highest frequency bandwidth of 48Mhz. The FDR test system was most sensitive to cable anomalies from 100 to 500 MHz. Although the SSTDR and FDR bandwidths are not directly comparable, both the SSTDR and FDR show similar trends in their frequency bandwidth responses. Lower frequencies propagate further along the cable but have poorer resolution capability that can, for example, reveal the start and stop of cable damage segments. Higher frequencies have better spatial resolution but the signal attenuates more with propagation distance along the cable and so may be less suitable for long cable inspections. Although the FDR tests seemed less noisy than the SSTDR, both systems showed indications of thermal damage, low-resistance faults between phases, and for unshielded cable – sensitivity to the presence or absence of water as the cable was passed through a water bath. The data suggested that a higher frequency SSTDR may be more sensitive to cable anomalies and could better resolve the start and stop of damage segments. A PNNL SSTDR was developed to examine higher bandwidth frequencies based on a software-controlled laboratory instrument without requiring changes to the commercial SSTDR system.

One of the main conclusions of the previous effort was that cable reflectometry plots can be difficult for humans to analyze due to baseline noise, low or noisy anomaly response peaks, or large responses from cable ends. Detection of cable anomalies for many of these frequencies and test conditions was challenging for manual analysis. This presented an ideal opportunity for ML analysis to distinguish undamaged cable indications from anomalous cable indications. This research discusses application of machine learning (ML) to reflectometry cable test methods. The goal was to assess feasibility to distinguish undamaged cable reflectometry responses from damaged or anomalous cable reflectometry responses. The assessment considered the 3 instruments, multiple frequency bandwidths from each instrument, multiple cable anomalies and test conditions, and both supervised and unsupervised ML approaches.

The PNNL software-controlled laboratory instrument with expanded bandwidth capabilities was set to test cables at 50, 100, 200, 300, 400, and 500 MHz. The FDR was also set to acquire data at 50, 100, 200, 300, 400, and 500 MHz. The LiveWire SSTDR collected data at 6, 12, 24, and 48 Mhz. For unsupervised ML, all 6 PNNL SSTDR and FDR frequencies were used and only the 48 MHz

LiveWire data was considered. For supervised data, the PNNL SSTDR and FDR 300 and 500 MHz data was excluded plus all 4 LiveWire SSTDR frequencies were included. The test matrix included 42 test conditions although 3 were identically labeled and so reduced the number of supervised cases accordingly. The three instruments were evaluated over the test matrix conditions. The data was distributed to two independent teams – one applying unsupervised learning methods and the other applying supervised methods to the data.

Although approaches and analysis methods were not identical or directly comparable, both outputs were encouraging. The unsupervised prediction weighted accuracy was assessed by instrument and by frequency. It performed better at high frequencies with the highest prediction accuracy of 0.84 for the higher frequency FDR, 0.79 for the 48-MHz LiveWire SSTDR, and 0.77 for 300-MHz PNNL SSTDR. The initial weighted accuracy average across all frequencies for using supervised ML was 0.56 to 0.68. The supervised analysis was repeated with noisier training data removed resulting in weighted accuracies of 0.69 to 0.87. These weighted accuracies are not directly comparable due to differences in the supervised and unsupervised analysis details but do indicate an encouraging trend. Even with limited and unbalanced data, strong prediction accuracies seem encouraging for further work including more data under a wider range of conditions.

ACKNOWLEDGEMENTS

This work was sponsored by the Department of Energy, Office of Nuclear Energy, for the Light Water Reactor Sustainability (LWRS) Program Materials Research Pathway. The authors extend their appreciation to Pathway Leads Dr. Xiang (Frank) Chen and Dr. Leo Fifield for LWRS programmatic support. Thanks also to LiveWire Innovation who supported this work by lending their SSTDR instrument to the program for data acquisition. This work was performed at the Pacific Northwest National Laboratory (PNNL) and at the Idaho National Laboratory (INL). PNNL is operated by Battelle for the U.S. Department of Energy under contract DE-AC05-76RL01830, and INL is operated by Battelle Energy Alliance, LLC, for the U.S. Department of Energy under Contract DE-AC07-05ID14517. This report is submitted as the milestone deliverable M3LW-23OR0404033.

CONTENTS

FIGURES

TABLES

ACRONYMS

Ada	Adaptive Boosting Classifier (ML algorithm)
ARENA	Accelerated and Real-Time Environmental Nodal Assessment
AWG	arbitrary waveform generator
BPSK	binary phase shift keying
BW	bandwidth
et	Extra Trees Classifier (ML algorithm)
dB	Decibel
dt	Decision Tree Classifier (ML algorithm)
Fc	carrier frequency
FDR	frequency domain reflectometry
FFT	fast Fourier transform
gbc	Gradient Boosting Classifier (ML algorithm)
gpc	Gaussian Process Classifier (ML algorithm)
KNN	K Nearest Neighbors Classifier (ML algorithm)
LDA	Linear Discriminant Analysis (ML algorithm)
lightgbm	Light Gradient Boosting Machine (ML algorithm)
lr	logistics regression (ML algorithm)
LWRS	Light Water Reactor Sustainability
ML	Machine Learning
MLP	Multi-Layer Perceptron Classifier (ML algorithm)
nb	Naive Bayes classifier (ML algorithm)
NPP	nuclear power plant
PN code	pseudo-random noise code
PNNL	Pacific Northwest National Laboratory
rf	Random Forest Classifier (ML algorithm)
SMOTE	synthetic minority over-sampling technique
SSTDR	spread spectrum time domain reflectometry
SVM	Support Vector Machine linear kernel (ML algorithm)
V	volt
VNA	vector network analyzer
Xgboost	Extreme Gradient Boosting classifier (ML algorithm)

1. INTRODUCTION

The Light Water Reactor Sustainability (LWRS) program has a cable nondestructive evaluation (NDE) effort that aims to evaluate and advance promising NDE methods for cable inspection in support of nuclear power plant (NPP) owner and operator interest to improve cable management reliability and reduce cost. NDE techniques are designed to detect and locate aging and operational damage and degradation prior to failures. Although most safety critical cables are initially qualified in accordance with NRC Guide 1.1211 (NRC 1977) and IEEE 383 (IEEE 2015) for 40 years, most US plants have applied for or have been granted life extensions beyond their initial 40-year license. Part of the life extension program is to justify continued service of cables based on performance tests rather than the initial qualification that only justified service to 40 years. This extension is needed because completely replacing NPP cables would be cost prohibitive and experience has shown that cables can continue to function safely long beyond their initial qualification. Testing is normally performed during outages when the cables can be taken out of service, de-energized, and tested using a variety of test methods. This approach is costly to the utility and subjects the cables to risks of damage during re-termination and the stresses associated with some of the test methods.

Reflectometry techniques offer a promising improvement to cable testing. A test instrument is connected to the cable end and a voltage wave is applied that travels along the cable. If an impedance change is encountered, part of the energy is reflected back to the reflectometry test instrument that can sense the reflected wave. The time delay between wave initiation and reflection detection is related to the distance along the cable by the wave velocity of propagation. Two types of reflectometry are considered in this research – Frequency Domain Reflectometry (FDR) and Spread Spectrum Time Domain Reflectometry (SSTDR). FDR is becoming more commonly used in NPPs to identify and locate damage. SSTDR is being used in the rail and aircraft industry and is being evaluated for NPPs. A particular advantage of SSTDR is that it can be applied to low voltage energized cables as an online technique. Implementing this in NPPs offers significant potential advantages in that cable condition monitoring need not be limited to outage testing, does not require human error-susceptible de-termination/re-termination, and offers the possibility of improved sensitivity by monitoring for changes in the cable condition that can minimize noise influences of tight-radius bends, proximity to metal (for unshielded cable), cable end responses, and manufacturing anomalies.

A 2022 program (Glass et al. 2022) evaluated a commercial SSTDR system in the Pacific Northwest National Laboratory (PNNL) Accelerated and Real-time Environmental Nodal Assessment (ARENA) cable/motor test bed (Glass, Fifield, and Prowant 2021). Test conditions included FDR at 50-, 100-, 200-, and 400-MHz bandwidths, and SSTDR at 6, 12, 24, and 48 MHz. The actual frequency bandwidths are not equivalent but the general trend of higher or lower frequency bandwidths are common among all instruments. Lower frequency bandwidths propagate further along the cable but with lower spatial resolution. Higher bandwidths attenuate more and propagate shorter distances. Higher frequencies are generally noisier but offer better spatial resolution, which is helpful for many cable analysis situations. Conclusions from (Glass et al. 2022) included an observation that frequency bandwidth limitations of the commercial SSTDR may be limiting to fault detectability and speculation that a wider frequency bandwidth may offer improved performance.

Another conclusion of the previous effort was that cable reflectometry plots can be difficult for humans to analyze due to baseline noise, low or noisy anomaly response peaks, or large responses from cable ends. Although some faults produced large and clear responses, many of the anomalies produced subtle indications that would be difficult to detect for both FDR and SSTDR measurements. This presented an ideal opportunity for machine learning (ML) analysis to distinguish undamaged cable indications from anomalous cable indications. This research assesses feasibility of ML applicability to reflectometry cable test methods.

A software-controlled laboratory instrument-based system was developed that allowed higher frequency bandwidths. Development of this PNNL SSTDR is documented in (Glass et al. 2023), and an initial evaluation of the PNNL SSTDR compared to FDR and the LiveWire SSTDR for 42 different cable test conditions is documented in (Glass et al. 2023). Test conditions included phase-to-phase mid-cable shorts and low-resistance faults, thermal damage, and water exposure.

In this effort, data from (Glass et al. 2023) was used to evaluate ML ability to distinguish undamaged cable from cable with mid-span anomalies. Similar data was analyzed by to two independent teams. The Idaho National Laboratory (INL) team focused on unsupervised ML while the Pacific Northwest National Laboratory (PNNL) focused on supervised ML. The performance of each approach is discussed in this report. Before the approaches and results are discussed, an introduction to the ARENA test bed and the FDR and SSTDR methods is presented.

1.1 Arena Cable Motor Test Bed

To evaluate the degradation of electrical cables and particularly the interaction of electrical cable test technologies with various damage mechanisms, PNNL developed the ARENA test bed (Glass, Fifield, and Prowant 2021), shown in Figure 1. The vision behind the creation of this facility is to establish a modular test facility that allows for the implementation of a broad range of test methods to detect faults and anomalies in a variety of cables and systems in a controlled environment. ARENA includes:

- A motor controller for 3-phase 480 VAC motor control.
- A ½ horsepower 3-phase 480 VAC motor.
- Remote start capability and barriers to protect operators from arc-flash exposure should cable fail.
- Circuit breaker protection that guards and isolates the building power from test system failure.
- A thermal aging oven that allows up to 10 m of cable to be spooled and thermally aged.
- A water bath that allows cables to be submerged.
- Cable trays like those found in NPPs that allow cables to be spread and protected from operators standing on the cables or moving them during testing.

The system can be operated with either shielded or unshielded cable. For these tests, cable samples were ~ 100 ft in length with a 25-ft test lead.

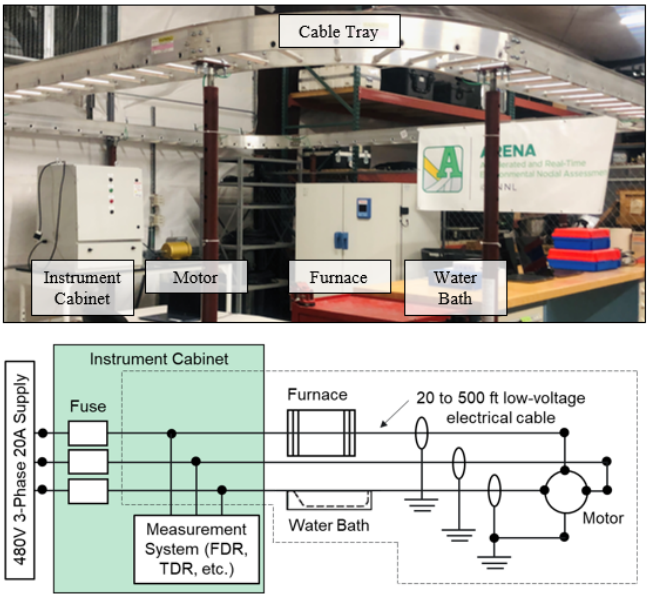


Figure 1. The ARENA Test Bed (top) digital image and (bottom) schematic (Glass et al. 2023)

1.2 FDR OVERVIEW

FDR is being used in nuclear plants, particularly to locate areas of concern. The FDR instrument, typically based on a vector network analyzer (VNA), is connected to two cable conductors, one considered the primary conductor under test and the other considered the system ground, as shown in Figure 2**,** or to a parallel conductor within the cable bundle (Glass et al. 2017). The instrument directs a swept-frequency chirp along the conductor and then listens for any reflection caused by an impedance change along the cable length. By listening and detecting the reflections in the frequency domain, eliminating noise susceptible frequencies, then transforming to the time domain with an inverse Fourier transform (IFT), significant noise immunity and sensitivity to subtle impedance changes can be achieved (Glass et al. 2017). The signal is then transformed to time/distance domain using an inverse Fourier transform and the velocity of propagation (VOP) (Glass et al. 2017). The bandwidth for the FDR is software adjustable up to 1.3 GHz, but experience shows the best responses from 50 MHz to 500 MHz. Higher FDR bandwidths produce sharper peaks capable of spatially resolving more closely spaced impedance changes, but the higher frequencies do not propagate as far along the cable length. Typical tests are performed at multiple bandwidths (for this study, 50, 100, 200, 300, 400, and 500 MHz were used) providing the analysts both high and low frequency bandwidths to consider in dispositioning tests. FDR instruments are restricted to relatively low voltages and cannot tolerate testing on energized cable systems.

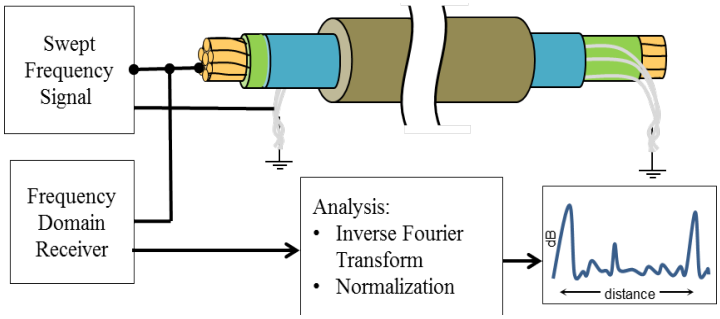


Figure 2. FDR cable test introduces a swept frequency chirp onto a conductor then listens for any reflection from any impedance change along the cable length.

1.3 LiveWire SSTDR OVERVIEW

The LiveWire commercial SSTDR produces a similar plot to the FDR, but all processing is in the time domain. A pseudo-random noise code (PN code) modulated with a square wave carrier through a high pass filter is input onto the cable conductor, and the instrument listens for any reflected response from cable anomalies (Figure 3). The SSTDR processes the signal as an autocorrelation, comparing the input PN code to any reflected signal detected. It can be used for measurements on energized cable up to 1 kV. The autocorrelation algorithm produces a robust noise-tolerant signal response and, as with the FDR, responses are quite different as a function of the bandwidth. The LiveWire SSTDR instrument produced responses for this study at 6, 12, 24, and 48 MHz. The definition of bandwidth for FDR is different it is for SSTDR. The 2022 study (Glass et al. 2022) indicated that some higher bandwidth SSTDR responses could help with analysis and the probability of detection.

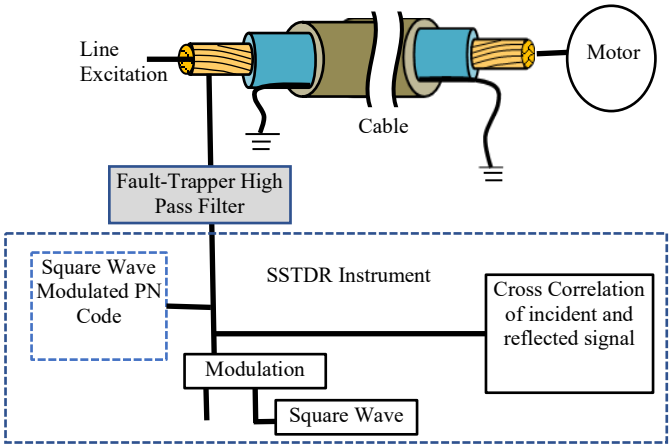


Figure 3. LiveWire SSTDR cable test applies a PN code to the conductor for cross correlation analysis.

1.4 PNNL SSTDR OVERVIEW

In response to interest in a higher bandwidth SSTDR, PNNL developed a flexible software-controlled laboratory-based SSTDR system to investigate SSTDR sensitivities to different cable impedance discontinuities as a function of bandwidth and particularly addressing higher frequency bandwidths than possible with the LiveWire instrument. The PNNL SSTDR system consists of the components in the function block diagram shown below in Figure 4 and the physical configuration shown in Figure 5. In the block diagram, the SSTDR signal generation is created by an arbitrary waveform generator (AWG) that creates a broadband excitation chirp. The AWG provides two outputs for a single waveform as a differential pair: the (+), or 0-degree waveform, is used as the signal injected down the cable line, and the (-), or 180-degree waveform, is used as a reference to correlate against the (+) signal as it is received. This method provides a phase and time synchronous copy of the SSTDR waveform and is ideal for cross correlations. The PNNL SSTDR was software-controlled for this test to produce reflectometry measurements with bandwidths of 50, 100, 200, 300, 400, and 500 MHz.

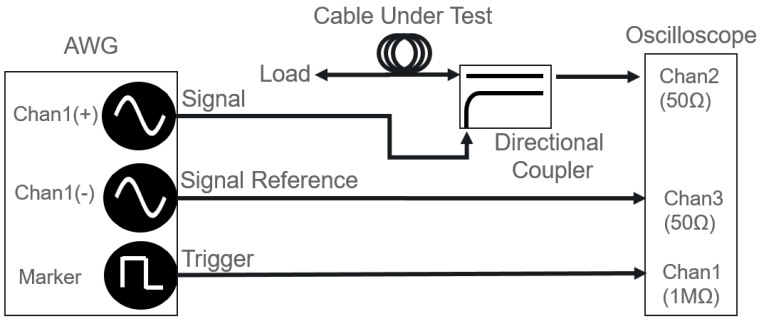


Figure 4. Block diagram of PNNL SSTDR system

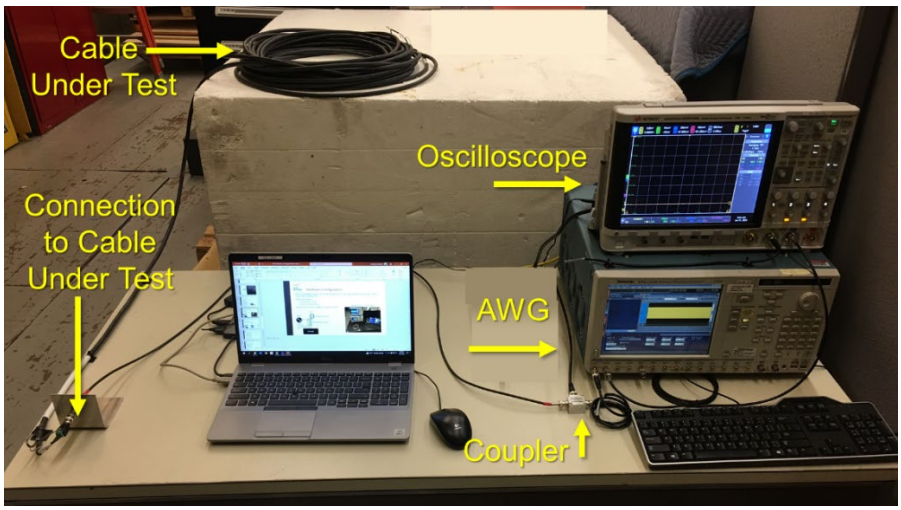


Figure 5. PNNL SSTDR initial test configuration

2. TEST CASES

Cables were tested in the ARENA cable and motor test bed (PNNL 2022). Evaluated cables were low voltage, tri-core, shielded and non-shielded cables manufactured by General Cable® (Catalog number 383830). The overall specifications for the selected cables are given in Table 1. Both shielded and non-shielded cable variants were comprised of three 14 AWG conductor wires insulated by ethylene propylene rubber (EPR) and protected by a chlorinated polyethylene (CPE) jacket. Additionally, the shielded cable contains an aluminum foil shield layer between the jacket and insulated wires. The cables have a voltage rating of 600 V and an operating temperature rating of 90 °C. All FDR and SSTDR tests were performed on cables of around 100-ft length.

Table 1. Manufacturer information for the cables selected for evaluation.

Manufacturer	P/N	Jacket	Insulation	Type
General Cable	354800	CPE	EPR	Shielded
	6-903-SH 14AWG-3/C FR-EP 600V FR-EPR/CPE Foil Shielded 600V E-2			
General Cable	383830	CPE	EPR	Non-Shielded
	6-903-G 14AWG-3/C FR-EP 600V FR-EPR/CPE Non-Shielded 600V E-2			

EPR = ethylene-propylene rubber; CPE = chlorinated polyethylene.

Tests were conducted using (1) an undamaged cable with open, short, 100 ohm, 312 ohm, and 2k ohm terminations at the distal cable ends, (2) thermally aged cable that entered and exited an oven at 45 to 75 ft, (3) phase-phase and phase-shield low-resistance faults at 26 ft, and (4) water immersion in a water bath from 61 to 67 ft.

Data can be viewed in either dB or magnitude arbitrary units (a.u.) and each display has its own advantages. The dB response compresses the full data set and allows some visibility of the lower-level changes between the responses from the two cable ends but still allows viewing the cable end response. The magnitude plots can show more subtle responses but may require the higher peaks associated with the cable ends to be truncated. Examples of this are shown in Figure 6 and Figure 7. These plots are of an anomalous cable response with a 100-ohm phase to phase fault at 26 ft. This is a relatively obvious indication with a substantial amplitude response above the noise floor at 50 ft between the peaks from the cable ends at 25 ft and 125 ft. The cable start peak is at 25 ft because there is a 25-ft test lead between the instrument and the cable start. The dashed vertical lines mark the beginning of the cable after the test lead and the peak corresponding to the cable end. One example of an undamaged cable is shown in Figure 8. Note that there are no peaks between the cable end responses. Seventeen cases of undamaged cable were presented with different cable end impedance terminations. The 100-ohm fault anomaly is rather easily distinguished from the undamaged condition but many of the cases were much more challenging to identify an anomaly response above the noise.

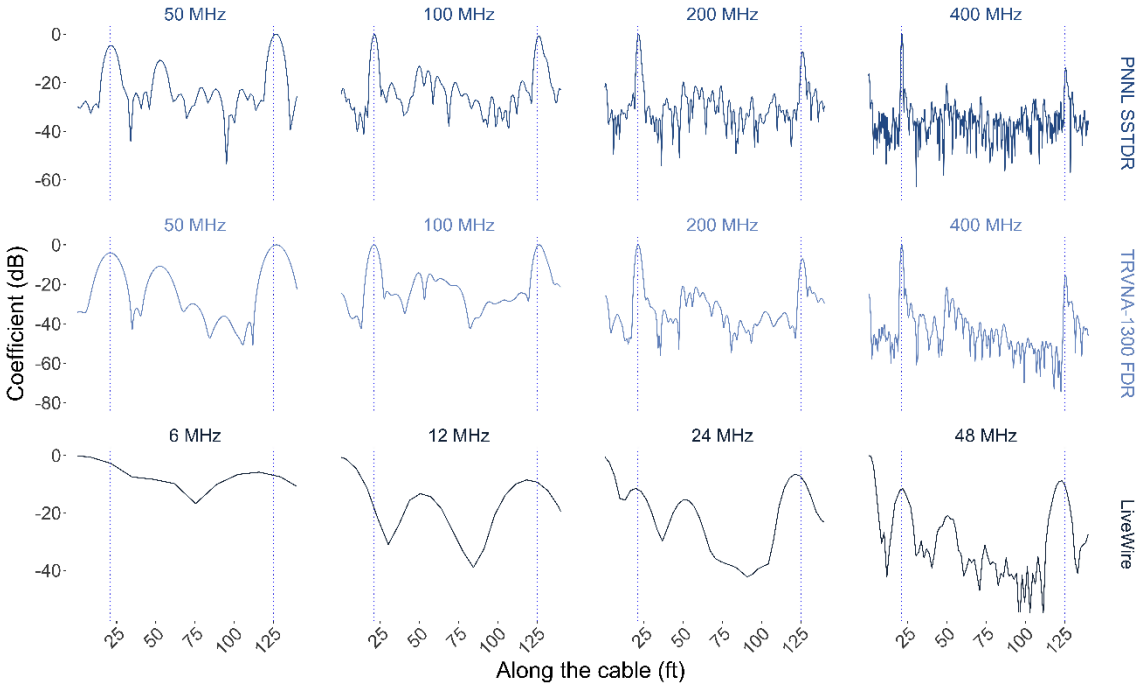


Figure 6. dB response for 100-ft shielded and undamaged cable with 100 ohms at 26-ft (shown as approximately 50-ft on plot due to test lead). The dashed vertical line marks the beginning of the cable after the test lead and the peak corresponding to the cable end.

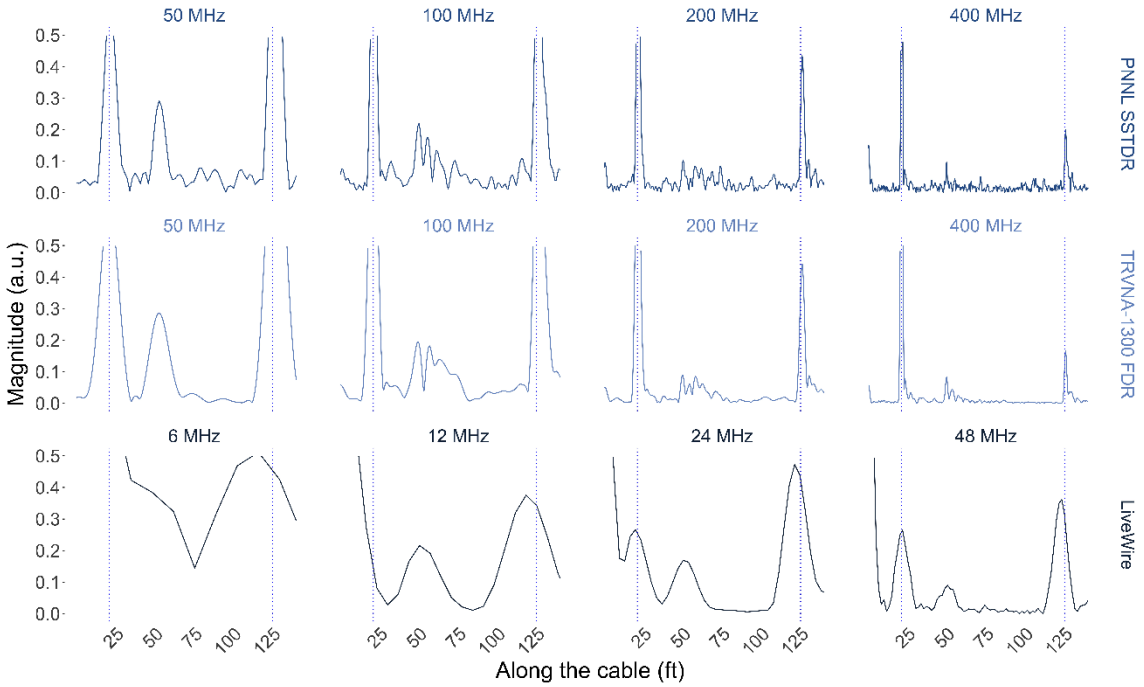


Figure 7. Magnitude response for 100-ft shielded and undamaged cable with 100 ohms at 26-ft (shown as approximately 50-ft on plot due to test lead).

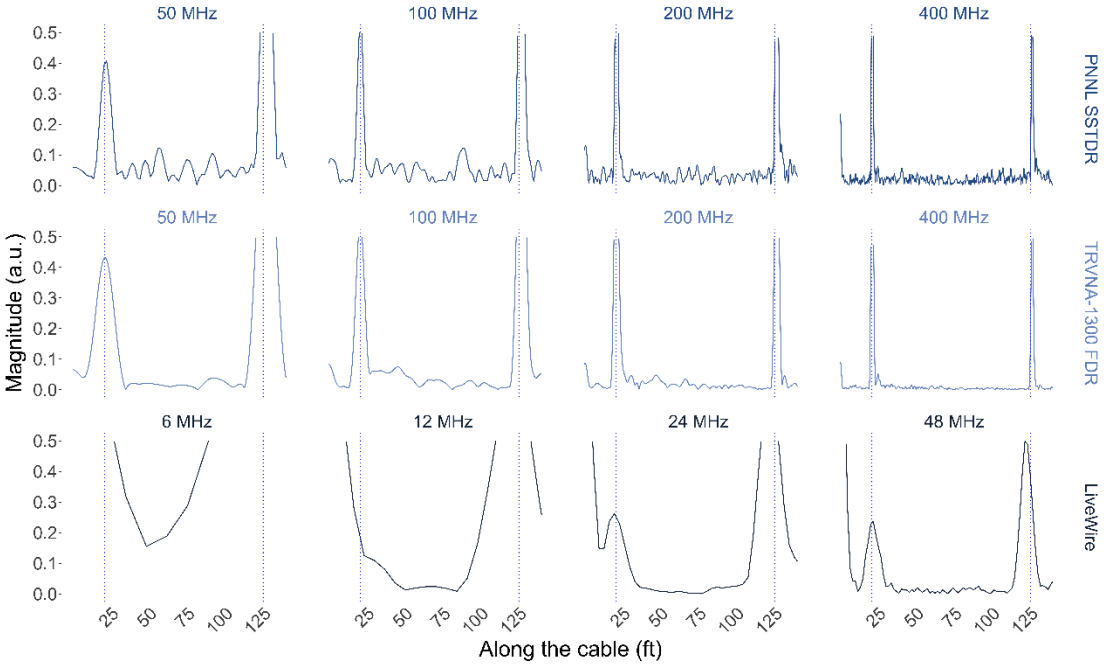


Figure 8. 100-ft shielded undamaged cable with distal end phase-phase short

The full test matrix is shown in Table 2. The test condition 2nd column is coded as either green for undamaged cables or yellow for cables with an anomalous condition that the reflectometry test is targeted to detect. For each of the 4 frequency bandwidths, a subjective ranking (-1, 0,1) and a color code (orange, yellow, or green) are assigned in accordance with guidelines of Table 3. There are 17 undamaged conditions and 25 anomalous conditions. This constitutes a slightly unbalanced distribution of data but still supports the ML effort. Note that there are a significant number of orange or yellow cells in indicating that manual interpretation of the data is challenging.

Table 2. Test matrix shows undamaged and damaged cable conditions and a subjective ranking of each test condition by instrument and frequency bandwidth.

	Key: Green conditions are for undamaged cable	PNNL SSTDR				FDR				LiveWire			
	Key: Yellow conditions are for anomalous cable condition	50	100	200	400	50	100	200	400	6	12	24	48
#	Test condition												
1	100ft Shielded Cable, 100 ohm phase - shield at distal end	1	1	1	1	1	1	1	1	-1	-1	1	1
2	100ft Shielded Cable, 2k ohm phase - shield at distal end	1	1	1	1	1	1	1	1	-1	0	1	1
3	100ft Shielded Cable, 312 ohm phase - shield at distal end	1	1	1	1	1	1	1	1	-1	-1	1	1
4	100ft Shielded Cable, Undamaged Cable (short distal end)	1	1	1	1	1	1	1	1	-1	1	1	1
5	100ft Shielded Cable, Short phase-shield at distal end	1	1	1	0	1	1	1	1	-1	1	1	1
6	100ft Shielded Cable, Undamaged Cable (short distal end)	1	1	1	1	1	1	1	1	-1	0	1	1
7	100ft Shielded Cable, Undamaged Cable 100 ohms at 26-ft	1	1	0	1	1	1	0	1	-1	1	1	1
8	100ft Shielded Cable, Undamaged Cable 2kohms at 26-ft	1	1	1	1	1	1	1	1	-1	-1	1	1
9	100ft Shielded Cable, Undamaged Cable 2kohms at distal end	1	1	1	1	1	1	1	1	-1	1	1	1
10	100ft Shielded Cable, Undamaged Cable 312 ohms at 26-ft	1	1	1	1	1	1	0	1	-1	1	1	1
11	100ft Shielded Cable, Undamaged Cable 312 ohms at distal end	1	1	1	1	1	1	1	1	-1	1	1	1
12	100ft Shielded Cable, Undamaged Cable connected to motor	1	1	1	1	1	1	1	1	-1	1	1	1
13	100ft Shielded Cable, Undamaged Cable short at 26-ft	0	0	0	0	0	0	0	0	-1	-1	0	0
14	Additional Testing, cable wirenut splices	0	0	0	0	0	0	0	0	-1	0	0	0
15	100ft Shielded Cable, 100 ohm phase - shield at distal end	1	1	1	1	1	1	1	1	-1	1	1	1
16	Phase-Phase Faults, Undamaged Cable 100 ohm at distal end	1	1	1	1	1	1	1	1	1	1	1	1
17	Phase-Phase Faults, Undamaged Cable 100 ohms at 26-ft	-1	-1	1	1	-1	1	1	1	-1	-1	-1	-1
18	Phase-Phase Faults, Undamaged Cable 2kohms at distal end	1	1	-1	-1	1	1	1	1	-1	-1	-1	1
19	Phase-Phase Faults, Undamaged Cable 2kohms at26-ft	1	1	1	-1	0	1	1	1	-1	-1	0	1
20	Phase-Phase Faults, Undamaged Cable 312 ohms at 26-ft	-1	-1	1	1	-1	1	1	1	-1	-1	-1	-1
21	Phase-Phase Faults, Undamaged Cable 312 ohms at distal end	1	1	-1	-1	1	0	0	0	-1	1	1	1
22	Phase-Phase Faults, Undamaged Cable connected to motor	1	1	1	1	1	1	1	1	-1	1	1	1
23	Phase-Phase Faults, Undamaged Cable short at 26-ft	1	1	1	1	1	1	1	1	-1	1	1	1
24	Thermally aged cable in oven, Good Undmgd Cable 1-pos 2-neg end1 open ends	1	1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
25	Thermally aged cable in oven, Good Undmgd Cable 1-pos 2-neg end2 open ends	0	0	-1	-1	-1	1	-1	-1	-1	-1	-1	-1
26	Thermally aged cable in oven, Good Undmgd Cable 2-pos 3-neg end1 open ends	1	1	-1	-1	0	0	-1	-1	-1	-1	-1	-1
27	Thermally aged cable in oven, Good Undmgd Cable 2-pos 3-neg end2 open ends	-1	0	-1	-1	-1	1	0	-1	-1	-1	-1	0
28	Thermally aged cable in oven, Good Undmgd Cable 3-pos 1-neg end1 open ends	1	1	-1	-1	0	0	-1	-1	-1	-1	-1	-1
29	Thermally aged cable in oven, Good Undmgd Cable 3-pos 1-neg end2 open ends	-1	0	-1	-1	-1	1	-1	-1	-1	-1	-1	0
30	Thermally aged cable in oven, Good Undmgd Shielded Cable 1-pos 2-neg end1 open ends	0	-1	-1	-1	1	1	-1	-1	-1	-1	-1	-1
31	Thermally aged cable in oven, Good Undmgd Shielded Cable 1-pos 2-neg end2 open ends	0	0	-1	-1	0	0	-1	-1	-1	-1	-1	-1
32	Thermally aged cable in oven, Good Undmgd Shielded Cable 2-pos 3-neg end1 open ends	0	-1	-1	-1	0	0	-1	-1	-1	-1	-1	-1
33	Thermally aged cable in oven, Good Undmgd Shielded Cable 2-pos 3-neg end2 open ends	0	0	-1	-1	1	-1	1	-1	-1	-1	-1	-1
34	Thermally aged cable in oven, Good Undmgd Shielded Cable 3-pos 1-neg end1 open ends	0	-1	-1	-1	0	0	-1	-1	-1	-1	-1	-1
35	Thermally aged cable in oven, Good Undmgd Shielded Cable 3-pos 1-neg end2 open ends	0	0	-1	-1	1	0	-1	-1	-1	-1	-1	-1
36	Thermally aged cable in oven, Thermally Aged Cable short ends(1,2) at ambient	-1	0	0	-1	1	1	1	-1	-1	-1	-1	1
37	Thermally aged cable in oven, Thermally Aged Cable connected to motor at ambient	-1	1	0	-1	1	1	0	-1	-1	-1	-1	-1
38	Thermally aged cable in oven, Thermaly Aged Cable 1-pos 2-neg end 1 open ends at ambiant	-1	0	0	1	1	1	0	-1	-1	1	1	1
39	Water Bath Tests, shielded cable with 3-ft section in bath	-1	-1	-1	-1	-1	1	-1	-1	-1	-1	-1	-1
40	Water Bath Tests, shielded cable with 3-ft section out of bath	1	1	1	1	1	1	1	1	-1	1	1	1
41	Water Bath Tests, unshielded cable with 3-ft section in bath	1	1	-1	-1	1	1	-1	-1	-1	-1	1	1
42	Water Bath Tests, unshielded cable with 3-ft section out of bath	1	1	1	1	1	1	1	-1	-1	0	1	1
	Relative ranking sum by bandwidth	17	17	16	14	16	20	18	20	-21	4	14	17
	Relative ranking sum by measurement system	**64**				**74**				14			

Table 3. Subjective evaluation guidelines for reflectometry measurement data

Condition assessed for FDR/BW, LiveWire SSTDR/BW, PNNL SSTDR/BW	Value/Color
Undamaged cable – ends visible, no anomalous peaks	1
Undamaged cable – ends visible, unexplained peaks observed	0
Undamaged cable – ends not visible, reflectometry response not apparent	-1
Anomalous cable – ends visible and damage response clearly visible	1
Anomalous cable – ends visible and damaged response weakly or possibly visible	0
Anomalous cable – ends visible but no clear anomalous response noted	-1
Anomalous cable – ends not visible and no anomalous damage response noted	-1

3. MACHINE LEARNING FEASIBILITY FOR CABLE REFLECTOMETRY TESTS

The objective of this section is to assess the feasibility of applying ML analysis to automatically differentiate between normal or undamaged and anomalous cable reflectometry data. This is different than determining what type of issue exists, which is usually the step following identification of the existence of an issue. The cable anomaly detection algorithm is trained on historical reflectometry data and is then asked to decide about the health status of the cables using new and unseen data. Automating this process is particularly important for two reasons: 1) the machine could be better in identifying anomalies than the human. Algorithms have been shown to have high degrees of accuracy in anomaly and change detection applications (Nassif et al. 2021). 2) the machine can perform the analysis instantly. This benefit is particularly important for SSTDR reflectometry with the LiveWire instrument that is designed to support online measurements. Data may be monitored continuously to detect damage that can progress quickly so that it may not be practical for a human to react and analyze reflectometry responses quickly enough.

When considering the ML process, it was assumed that baseline normal reflectometry data for a range of time periods and operating conditions would be available to train an algorithm. Additionally, a limited amount of anomalous reflectometry data would be available. This assumption seems reasonable for an online test scenario where the measurement will start monitoring a cable while it is healthy, enabling ample accumulation of normal data. This also means the normal historical data would represent that exact cable in normal conditions over a range of power levels and operating conditions.

ML paradigms are often divided into unsupervised and supervised learning (Harley, Zafar, and Tran 2023). With unsupervised learning, the model is trained with input data features but without known data labels. Supervised learning trains the algorithm using known data labels. The algorithm then tries to fit any new test data to one or more of the labeled training data. Approximately the same reflectometry data was provided to two teams – one that applied an unsupervised ML approach at the INL and a second applying a supervised learning approach at the PNNL. The teams worked independently, and both the specific data considered (see Table 4) and the analysis approach were slightly different as discussed below.

Table 4. Available and analyzed data for supervised and unsupervised ML analysis applied to cable reflectometry test data.

Instrument\Frequencies	6	12	24	48	50	100	200	300	400	500
Unsupervised Data										
PNNL SSTDR					Y	Y	Y	Y	Y	Y
FDR					Y	Y	Y	Y	Y	Y
LiveWire				Y						
Supervised - initial analysis										
PNNL SSTDR					Y	Y	Y		Y	
FDR					Y	Y	Y		Y	
LiveWire	Y	Y	Y	Y						
Supervised -2nd analysis										
PNNL SSTDR					Y	Y				
FDR					Y	Y	Y			
LiveWire			Y	Y						

3.1 Unsupervised Machine Learning Applied to Cable Reflectometry

The unsupervised anomaly detection approach was taken where no labeled instances of anomalies or undamaged cable were used in training the algorithm. The idea behind this type of approach is to learn patterns and features that represent normal behavior. Based on these patterns, the algorithm provides an anomaly score, with the goal that normal data should have low anomaly scores and anomalous data should have high scores. An example of this anomaly score concept is shown in Figure 9, where the curves are probability density functions showing the relative prevalence of seeing data with those scores for both normal and anomalous data. In anomaly detection applications, the amount of separation between the normal and anomalous classes is both problem and algorithm dependent.

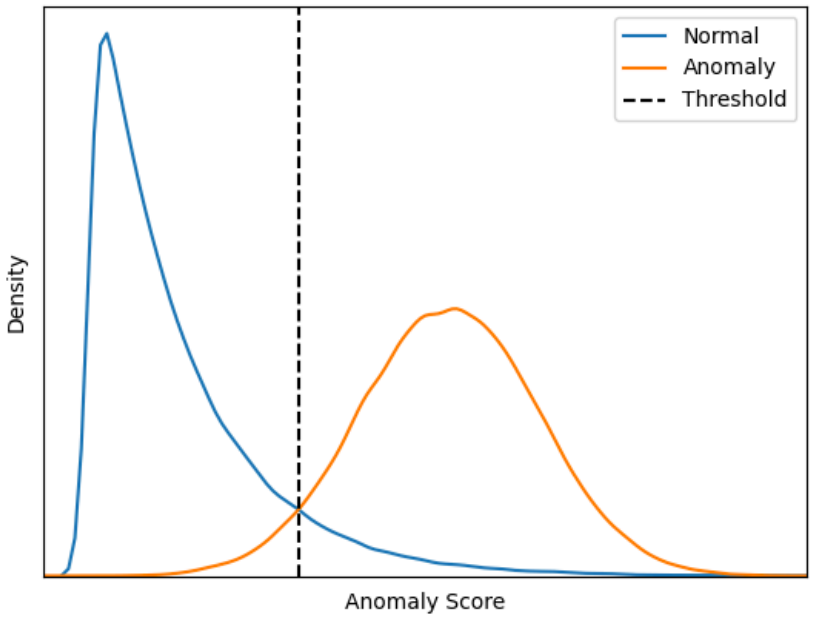


Figure 9. Probability density functions for hypothetical normal and anomalous classes, with a selected threshold to differentiate them.

As discussed earlier, the collected reflectometry data included three methods (FDR, LiveWire SSTDR, and PNNL SSTDR) and multiple frequencies (ranging from roughly 50-500 MHz). An important question was how to make use of the data across the different frequencies. Because this was a feasibility study, this effort treated each frequency dataset as independent, meaning that frequencies from the same method were not combined in making a single decision. The primary benefit from this was that it enabled comparisons to be made of the different frequencies to see which were most powerful in differentiating normal from anomalous conditions.

3.1.1 Preliminary Visual Evaluation

During preliminary investigations, one of the first questions was whether it was possible to visually see a difference between normal and anomalous data. To answer this question, each anomaly was plotted against the set of all normal data. Two examples of PNNL SSTDR data are shown in Figure 10 and Figure 11; in these figures, the dashed lines are the set of normal data and the solid blue line is the anomaly being analyzed. The data are plotted over the range of available frequencies. The x-axis is in ft and the y-axis is in dB. Starting with Figure 10, the anomalous data set is clearly visually different, so is a good candidate for an automated approach to detect this anomaly. By contrast looking at Figure 11, the anomalous data set is far more similar to the set of normal data, meaning an algorithm will have a more difficult time differentiating this from normal conditions. These visual plots provided some reference for guiding development of the detection methods used.

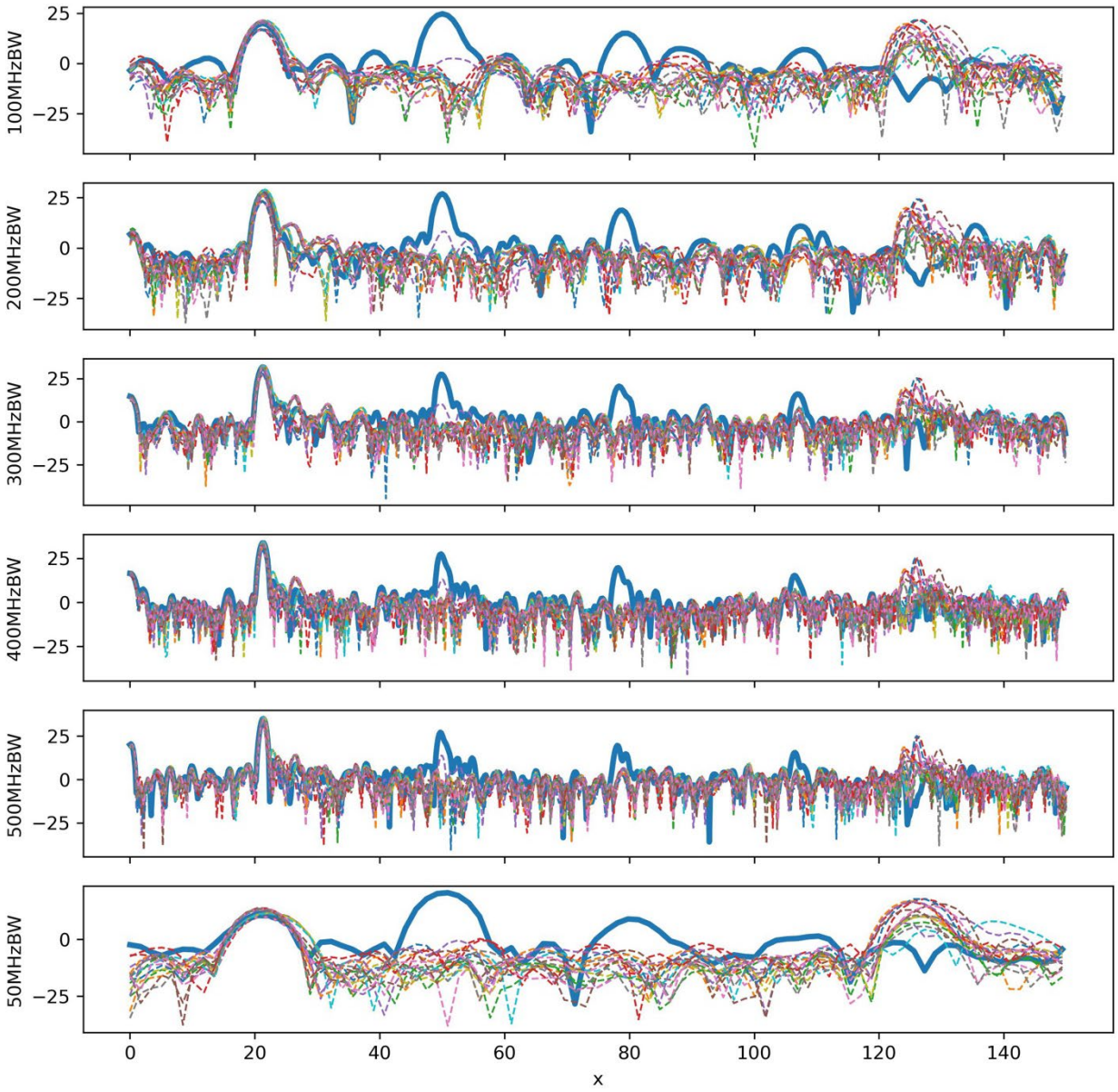


Figure 10. Visual comparisons between the anomalous 100-ft shield cable undamaged cable short at 26-ft condition PNNL SSTDR spectrometry data and the set of normal PNNL SSTDR spectrometry data.

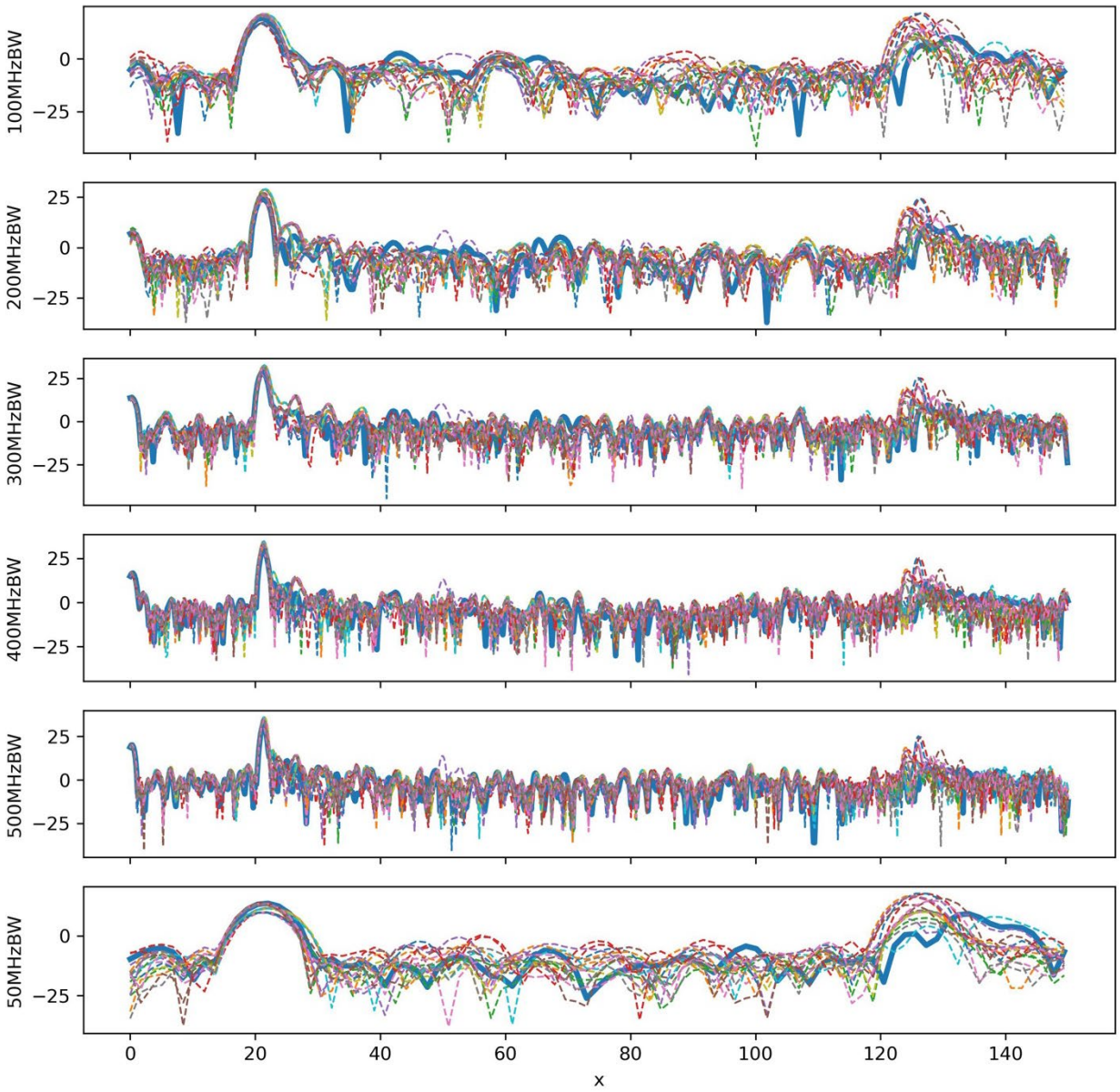


Figure 11. Visual comparisons between an example anomalous thermally aged cable PNNL SSTDR spectrometry data and the set of normal PNNL SSTDR spectrometry data.

3.1.2 Approach

For ML applications in general, data are separated into training and testing sets, where the training set is used to train the algorithm, and the test set (unseen by the algorithm) is used to assess the performance of the algorithm. This is a necessary step to avoid bias that would exist if the algorithm assessed performance with data that was used to train the algorithm. For this feasibility study, there were limited data with which to train and test a model. In total, there were 17 normal cable experiments and 25 anomalous cable experiments. With such a limited number of samples, the complexity of the algorithms was limited to avoid overfitting. Overfitting is the idea of extracting features that are specific only to the training samples and thus do not generalize well, providing poor performance on data not included in the training set.

Using the unsupervised framework described earlier, the approach taken in this effort was to compare unseen normal and anomalous data to a normal only training set using a distance metric that measured how far the unseen data was from the training set. This distance metric was directly used as the anomaly scoring algorithm, with small values indicating a high degree of similarity to the training set, and large values indicating a small degree of similarity (i.e., ideally indicating an anomaly). Two distance metrics were used: one that calculated the minimum point-wise distance between the test sample and the training set, and one that calculated the minimum vector-wise distance. For either of these to work, the spectrometry data needed to be sampled to contain the same number and spacing of x-axis values. This was done by interpolating each spectrometry dataset using a fixed set of values. It is important to note that for this effort, the process does work because all the cables were the same length; a more complicated sampling process may be necessary for cables of different lengths. If used in the field with historical baseline readings of a given cable compared with unseen data of that same cable, the approach used here should also be valid.

Starting with the point-wise distance metric, for a given test sample, the metric calculated the Euclidian distance at each point along the x-axis between the test sample and all the training samples and selected the minimum value. Then, the average value over the x-axis was selected as the anomaly score. In other words, this metric calculated the minimum distance between the test sample and the set of training samples. An example of this metric is shown in Figure 12; the test sample is shown in blue, and two training samples are shown using a dashed line in orange and green. The shaded area represents the minimum point-wise distance, and the color shows which of the two training samples is used to calculate the minimum.

By contrast, using the vector-wise distance metric, the metric calculated the Euclidian distance between the test sample (treated as a vector of values) and each training sample (also treated as vectors) and selected the minimum value. Here, the metric calculated the minimum distance between the test sample and the closest training sample. An example of this metric is shown in Figure 13. Here, the shaded area represents the minimum vector-wise distance, and the color is always the same because it selects the training sample that is closer to the test sample.

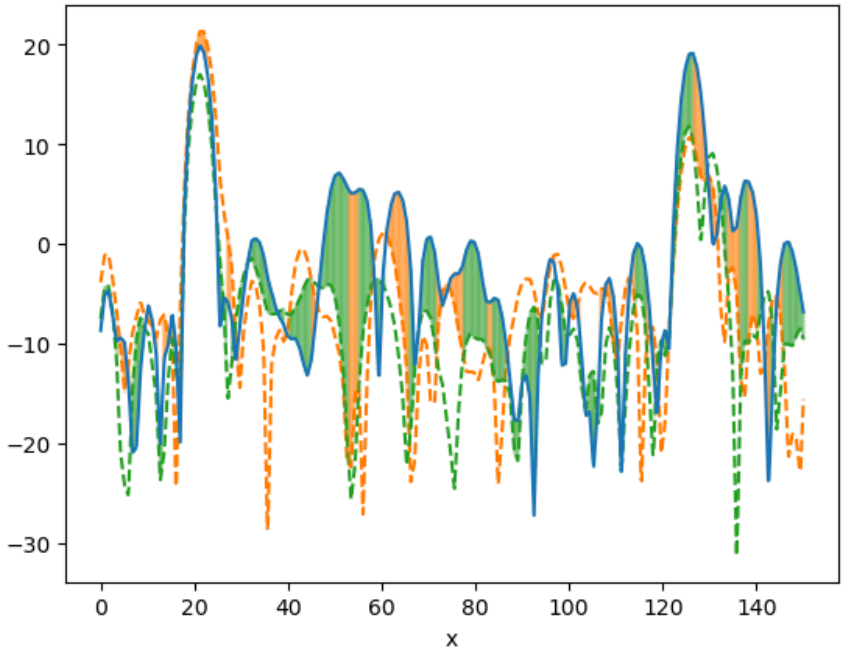


Figure 12. Example of the point-wise distance metric between a test sample and a training set.

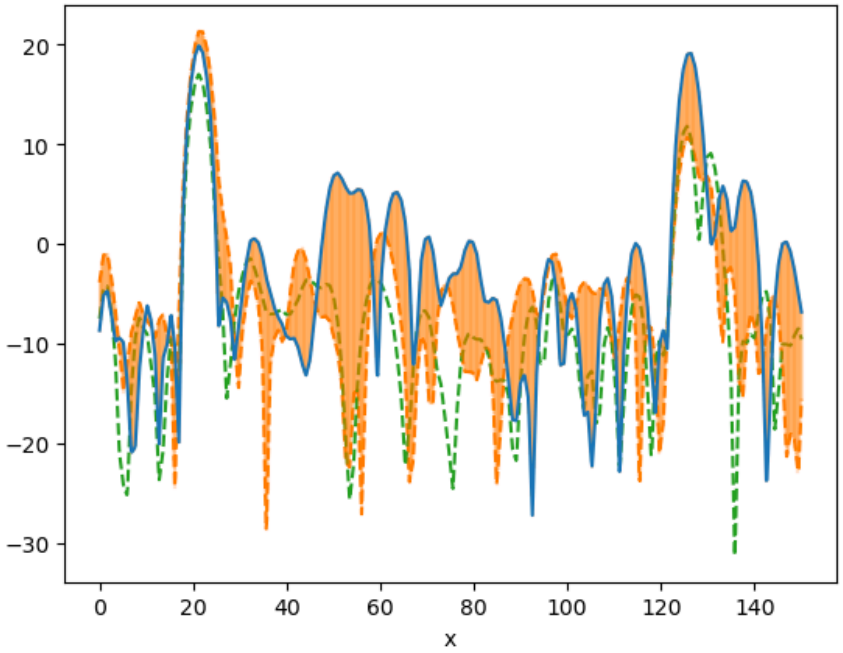


Figure 13. Example of the vector-wise distance metric between a test sample and a training set.

In addition to the distance metrics, two additional preprocessing options were tested and compared. The first option was to use the reflectometry data in either magnitude or decibels scale, to evaluate whether there was a performance difference between the two. The visual plots above are all in decibels, but that does not suggest which would work better. The second option was to include the synthetic minority over-sampling technique (SMOTE) originally developed for imbalanced classification problems (Chawla et al. 2002). At a high level, this is a data augmentation approach that takes two training samples and combines them to form a new synthetic training sample. In this effort, every possible pair of training samples was used, and the combination was done through averaging. Between the two distance metrics and preprocessing options, this resulted in eight approach permutations (Table 5).

Table 5. The different approach permutations tested in this effort.

Number	Distance Metric	Scale	Include SMOTE
1	Point-wise	Magnitude	Yes
2	Point-wise	Magnitude	No
3	Point-wise	Decibels	Yes
4	Point-wise	Decibels	No
5	Vector-wise	Magnitude	Yes
6	Vector-wise	Magnitude	No
7	Vector-wise	Decibels	Yes
8	Vector-wise	Decibels	No

Based on the limited data, one consideration in this effort was how to calculate anomaly scores for both the normal and anomalous data. This was necessary because if the normal data were separated into fixed training and testing sets, the testing set would contain too little normal data to generate reliable results. This was overcome using a cross-validation approach. Conceptually, this is a strategy of training and testing multiple times with different training and testing sets in order to generate anomaly scores for all of the data while avoiding overlap between the training and testing sets.

The cross-validation approach used was called leave-one-out cross validation. Based on the unsupervised framework, the data were split 17 times (equal to the size of the normal samples). During each split, one rotating normal sample and all the anomalous samples were placed in the testing set, the remaining 16 normal samples were placed in the training set, and the algorithm used the training set to calculate anomaly scores for each sample in the testing set. At the end of the 17 splits, a single anomaly score was calculated for each normal sample and 17 scores were calculated for each anomalous sample, which were then averaged, resulting in a single anomaly score for each data sample. This leave-one-out approach is shown in Figure 14.

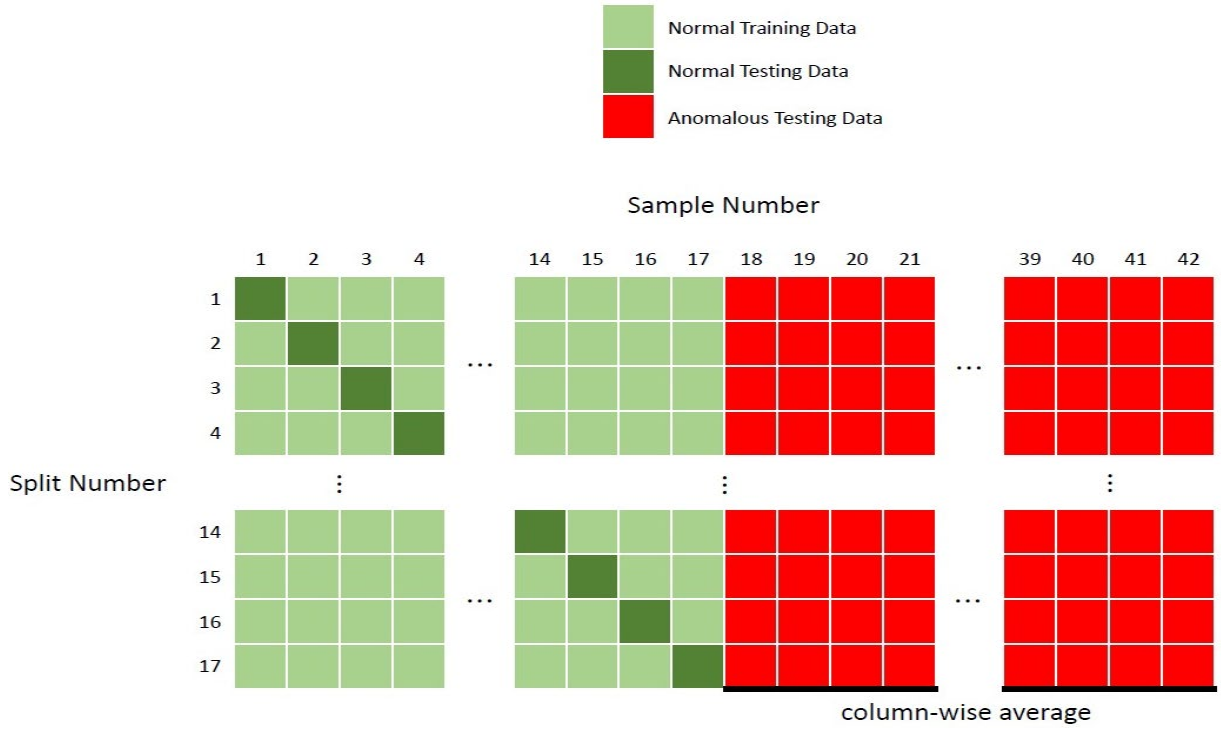


Figure 14. The leave-one-out cross validation approach used to calculate anomaly scores for all the data (17 split number undamaged cases and 42 sample number test cases)

3.1.3 Evaluation Metric

This effort used the weighted accuracy to report and compare the performance of different approach options, spectrometry methods, and frequencies. Accuracy was selected because it is highly interpretable compared to some other classification metrics, making it a clear choice, particularly for a feasibility study. Weighted accuracy was specifically chosen because it averages the accuracy for each class individually, which is beneficial when there are unequal numbers of samples in the different classes. As an example, if there are 99 normal samples and one anomalous sample and all are labeled as normal, the accuracy would be $\frac{99}{100} = 0.99$, but the weighted accuracy would be $\frac{\frac{99}{99}+\frac{0}{1}}{2} = 0.5$.

Using the developed approaches, anomaly scores for each sample and combination of approach options were calculated. To convert these to weighted accuracies, thresholds had to be selected. For this feasibility study, thresholds were selected to maximize the weighted accuracies. While these are likely higher than would be seen in real conditions, they provide appropriate approximations of the weighted

accuracies to determine whether the algorithms are working and be able to compare the algorithms against each other.

3.1.4 Performance

The first comparisons were between the point-wise and vector-wise distance metrics, and the magnitude and decibels scale. The weighted accuracies for these cases averaged over all spectrometry methods, frequencies, and SMOTE combinations are shown in Table 6. The averaging was done because it provides a better measure of just the effects of interest averaged over multiple possible implementations of that effect. Based on the results, interestingly the point-wise distance metric and magnitude scale did better than point-wise and decibels combination, but the vector-wise distance metric and decibels scale did better than the vector-wise and magnitude combination. As such, only these two better combinations were included in further analyses.

Table 6. Weighted accuracies for combinations of distance metric and scale, averaged over spectrometry methods, frequencies, and SMOTE combinations.

Distance Metric	Scale	Weighted Accuracy
Point-wise	Magnitude	0.729
Point-wise	Decibels	0.705
Vector-wise	Magnitude	0.681
Vector-wise	Decibels	0.743

Next, the effects of using the SMOTE data augmentation algorithm were investigated. For just the point-wise and magnitude, and vector-wise and decibels combinations, the results are shown in Table 7, again averaging over all spectrometry methods and frequencies. These results show that the SMOTE algorithm appears to have benefited both approaches, putting their weighted accuracies very close to each other (although as seen later, the individual accuracies over different methods and frequencies are still different). Similar to the last analysis, only those combinations that included SMOTE were analyzed further.

Table 7. Weighted accuracies for combinations of distance metric, scale, and SMOTE, averaged over spectrometry methods and frequencies.

Distance Metric	Scale	Include SMOTE	Weighted Accuracy
Point-wise	Magnitude	Yes	0.748
Point-wise	Magnitude	No	0.709
Vector-wise	Decibels	Yes	0.749
Vector-wise	Decibels	No	0.737

Based on these analyses, only two combinations of distance metric, scale, and SMOTE were selected for final results: point-wise, magnitude, and including SMOTE (first row in Table 5); and vector-wise, decibels, and including SMOTE (row 7 in Table 5). The results separated by method and frequency are shown in Figure 15 and Figure 16 for point-wise and vector-wise distance metrics, respectively. Based on these results, higher frequencies for the PNNL SSTDR and FDR generally showed better performance than lower frequencies, FDR generally performed better than PNNL SSTDR, and the LiveWire SSTDR, although only included one frequency, did significantly better than PNNL SSTDR and FDR at its respective frequency. Overall, the best, worst, and median classifier weighted accuracies were 0.83, 0.57, and 0.76, respectively, indicating that half of the classifiers had accuracy greater than or equal to 0.76.

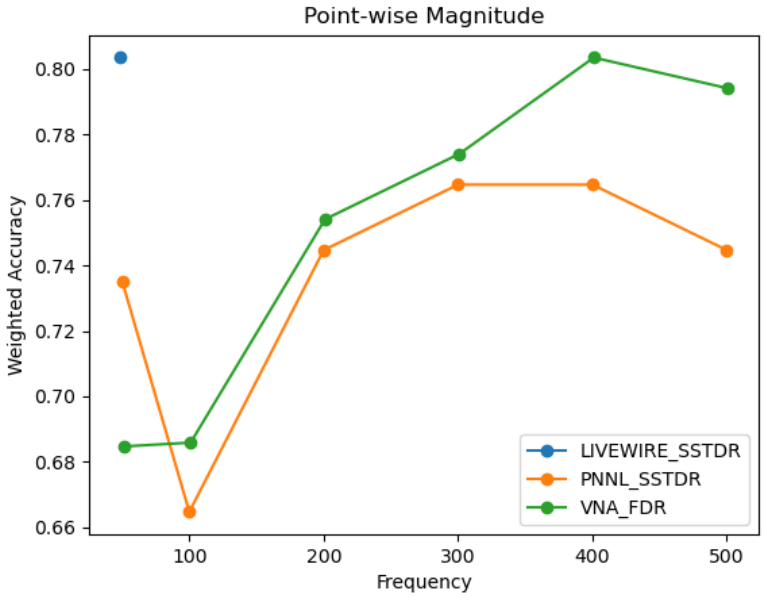


Figure 15. Weighted accuracies as a function of frequency for the point-wise distance metric and magnitude scale approach applied to the LiveWire SSTDR, PNNL SSTDR, and FDR datasets.

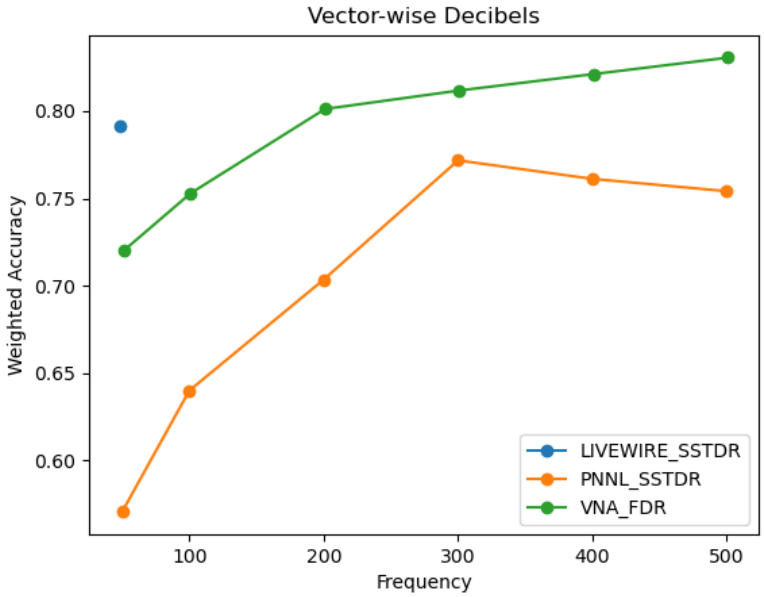


Figure 16. Weighted accuracies as a function of frequency for the vector-wise distance metric and decibels scale approach applied to the LiveWire SSTDR, PNNL SSTDR, and FDR datasets.

3.2 Supervised Machine Learning Applied to Cable Reflectometry

Supervised ML applies an algorithm to learn from labeled training data to make predictions or decisions without human intervention. In contrast to unsupervised anomaly detection, both labeled normal and anomalous data are used to train the supervised ML algorithm. One challenge with this approach for detection problems is that if data for a novel anomaly (i.e., one that the supervised algorithm has not seen before) is presented, it may not be obvious which class the algorithm should classify it as because the

algorithm has not seen that label before. By contrast, the unsupervised approach simply learns normal, so all anomaly classes are novel anomalies seen as different from normal.

3.2.1 Evaluation Metric

In this effort, an F1 score (described below) is used as a metric for model performance. The F1 score is considered a good metric for imbalanced data because it balances the trade-off between precision and recall. The definition of the F1-score, recall, precision, and accuracy are given below.

$$\text{Accuracy} = \frac{Correct\ Predictions}{All\ Predictions}$$

$$\text{Precision} = \frac{True\ Positives}{True\ Positives + False\ Positives}$$

$$\text{Recall} = \frac{True\ Positives}{True\ Positives + False\ Negatives}$$

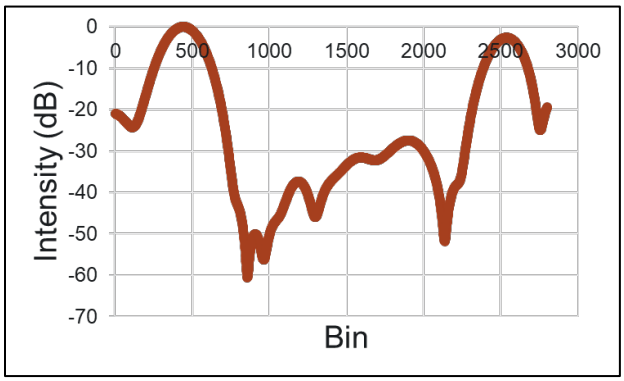
$$\text{F1-score} = 2 \times \frac{Precision \times Recall}{Precision + Recall}$$

A confusion matrix is used in this section. The confusion matrix is used in ML and statistics to evaluate the performance of a classification algorithm. It provides a summary of the predicted and actual classifications of a classification model to assess the model's accuracy.

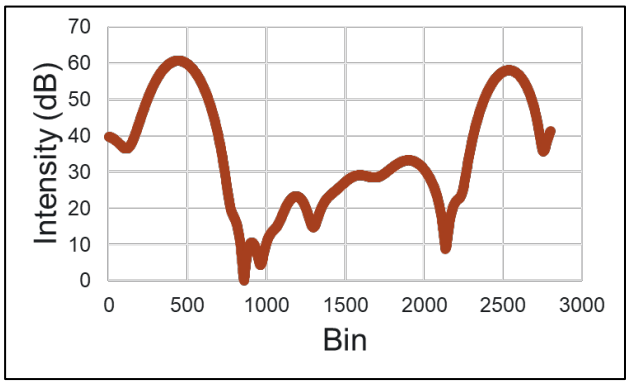
3.2.2 Approach

In supervised learning, the algorithm is provided with a dataset consisting of input-output pairs. The algorithm learns to map inputs to outputs. The data from four different bandwidths from each of FDR, LiveWire SSTDR, and PNNL SSTDR were analyzed similarly to the unsupervised ML. Data Pre-processing was applied to all the data to align the data in time and distance along the cable.

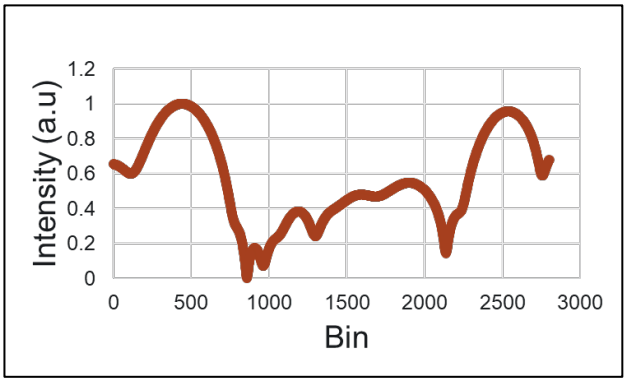
An equal number of aligned data points (0.0-ft bin width for 140-ft cable length) for all the frequencies was acquired without changing the shape of the curve as presented in Figure 17 (a). Since the intensity values in the raw data were both in +Y and -Y axis, in the data processing pipeline, all the values along +Y axis were shifted (b) and normalized (c) so that all the intensity values fell between 0 to 1. The intensity of the peaks was used as an anomaly indicator.



(a) Pre-processed data after making it equal bin of 2801



(b) Shift the intensity to +Y axis



(c) Normalized the intensity

Figure 17. Data preprocessing steps for the ML models.

Each of the reflectometry data samples were labeled either 0 (undamaged) or 1 (damaged). There were 14 uniquely labeled undamaged samples (3 of 17 labels were identical) where each test case had 4 frequency bandwidths leading to 56 of 164 undamaged samples (34%). There were 27 damaged or anomalous test conditions also with 4 frequencies leading to 108 of 164 damaged or anomalous cases (66%). The PyCaret Python Library code (Python 2023) was then used to apply classification algorithms including:

Logistic Regression (lr): Logistic Regression is a statistical technique used for binary classification tasks, where the goal is to predict one of two possible outcomes based on a set of input features. Unlike linear regression, which predicts continuous values, logistic regression employs the logistic function to constrain predictions between 0 and 1. This function transforms the linear combination of features and their respective coefficients into probabilities, representing the likelihood of belonging to the positive class. By setting a threshold, typically 0.5, these probabilities are then used to make binary predictions (Wright 1995). Logistic Regression is widely used due to its simplicity, interpretability, and effectiveness in a variety of fields such as healthcare, marketing, and natural language processing.

Ridge Classifier (ridge): The Ridge Classifier is a variant of logistic regression, primarily used for binary and multiclass classification tasks. It incorporates L2 regularization, also known as Ridge regularization, into the logistic regression model. This regularization method adds a penalty term based on the squared magnitudes of the coefficients of the model, discouraging large coefficient values, and thus

preventing overfitting. The Ridge Classifier constructs a linear decision boundary that separates different classes while mitigating multicollinearity issues and stabilizing the model (Hoerl and Kennard 1970).

Linear Discriminant Analysis (LDA): Linear Discriminant Analysis (LDA) is a supervised classification technique used to find a linear combination of features that best separates different classes in a dataset. LDA aims to maximize the distance between the means of different classes while minimizing the variance within each class, resulting in a robust decision boundary (Fisher 1936). By transforming the original feature space into a lower-dimensional space, LDA not only improves classification accuracy but also reduces the dimensionality of the data, making it valuable for dimensionality reduction in addition to classification. This method is particularly effective when dealing with multi-class classification problems and is widely used in applications such as image recognition, medical diagnosis, and pattern recognition.

K Neighbors Classifier (KNN): The K Nearest Neighbors (KNN) Classifier is a simple and intuitive supervised machine learning algorithm used for classification tasks. It works by finding the K nearest data points in the training dataset to a given test data point based on a distance metric (usually Euclidean distance) in the feature space. The class of the majority of these K nearest neighbors is then assigned to the test data point, effectively determining its classification (Peterson 2009). It is versatile and can handle both binary and multiclass classification problems, although its performance can be sensitive to the choice of K and the distance metric.

Naive Bayes (nb): The Naive Bayes classifier is a probabilistic and simple yet powerful machine learning algorithm used for classification tasks, particularly in natural language processing and text analysis. It is based on Bayes theorem and the assumption of feature independence, which simplifies calculations. The algorithm calculates the probability of a data point belonging to a particular class by considering the conditional probabilities of the individual features given that class. Then, it selects the class with the highest probability as the prediction (Al-Aidaroos, Bakar, and Othman 2010). Despite its "naive" assumption of feature independence, Naive Bayes often performs remarkably well in practice and is computationally efficient. It is widely used for tasks like spam email detection, sentiment analysis, and document categorization.

SVM - Linear Kernel: A Support Vector Machine (SVM) with a linear kernel is a machine learning algorithm that aims to find a straight-line (in two dimensions) or a hyperplane (in higher dimensions) that best separates different classes in a dataset. It accomplishes this by maximizing the margin, which is the distance between the separating hyperplane and the nearest data points of each class, known as support vectors (Cortes and Vapnik 1995). This approach results in a robust classification model that generalizes well to new data. SVMs with linear kernels are particularly effective when the data exhibits a clear linear separation, making them a valuable tool in various applications, including text classification, image recognition, and bioinformatics.

Gaussian Process Classifier (gpc): A Linear Kernel: A Support Vector Machine (SVM is a probabilistic machine learning model used for classification tasks. It extends the idea of Gaussian processes from regression to classification problems. The gpc models the relationship between input features and class labels as a probability distribution, capturing both the predicted class labels and the associated uncertainty (Rasmussen 2004). It provides a powerful framework for uncertainty quantification in classification tasks, making it useful in scenarios where not only accurate predictions but also confidence estimates are essential, such as medical diagnosis or fraud detection. The gpc can be employed with different kernel functions to capture various patterns and complexities in the data, and it naturally handles multiclass classification and imbalanced datasets while offering a principled way to balance model complexity and generalization.

MLP Classifier (mlp): A Multi-Layer Perceptron (MLP) Classifier is a type of artificial neural network used for supervised classification tasks. It consists of multiple layers of interconnected neurons (nodes) that process input data, apply nonlinear transformations, and make predictions. An MLP typically includes an input layer, one or more hidden layers, and an output layer. Each neuron in a layer receives

inputs, applies a weighted sum, and passes the result through an activation function, enabling it to model complex, nonlinear relationships in the data (Haykin 1999). MLPs are known for their ability to handle intricate patterns and are widely used in various domains, including image recognition, natural language processing, and predictive modeling, thanks to their capacity to learn intricate features and adapt to different types of data. However, they require careful hyperparameter tuning and sufficient data to prevent overfitting.

Decision Tree Classifier (dt): A Decision Tree Classifier is a popular machine learning algorithm used for both classification and regression tasks. It operates by recursively partitioning the dataset into subsets based on the values of input features, creating a tree-like structure of decisions. At each node of the tree, the algorithm selects the feature that best separates the data, typically using metrics like Gini impurity or information gain for classification tasks. The decision tree continues to split the data until it reaches a predefined stopping criterion, such as a maximum tree depth or a minimum number of data points per leaf. To make predictions, data points traverse the tree from the root node to a leaf node, where they are assigned the majority class label (for classification) or the mean value (for regression) of the training examples in that leaf (Wu et al. 2008). Decision Trees are interpretable and intuitive, capable of capturing complex interactions in the data, but they are prone to overfitting when not properly pruned or regularized.

Random Forest Classifier (rf): A Random Forest Classifier is an ensemble learning method that builds upon the concept of decision trees. It constructs a multitude of decision trees during training, each based on a random subset of the training data and a random subset of the input features. The randomness introduced in both data and feature selection helps to reduce overfitting and increase the model's robustness (Ho 1995). During classification, the predictions from all individual trees are combined, typically by majority voting for classification tasks, to make a final prediction. Random Forests are known for their high predictive accuracy, ability to handle large and complex datasets, and resistance to overfitting.

Extra Trees Classifier (et): An Extra Trees Classifier, short for Extremely Randomized Trees Classifier, is an ensemble learning method closely related to Random Forests. Like Random Forests, it builds multiple decision trees during training to make predictions. However, what sets Extra Trees apart is the additional randomness introduced in the tree construction process. In Extra Trees, decision trees are created with random splits that are not based on the best possible split points, as is the case in standard decision trees or Random Forests. Instead, Extra Trees selects random thresholds for feature splits, making the algorithm computationally efficient and less prone to overfitting (Geurts, Ernst, and Wehenkel 2006). Despite this extra randomness, Extra Trees often achieves similar or even superior performance to Random Forests, especially when dealing with noisy or high-dimensional data. It is a valuable choice for tasks where robustness and computational efficiency are crucial considerations.

Ada Boost Classifier (ada): Ada Boost (Adaptive Boosting) Classifier is an ensemble learning technique that combines multiple weak classifiers to create a strong classifier. It assigns weights to training instances, giving more weight to misclassified data points in each iteration to improve their classification. Weak learners, often simple decision trees or shallow models, are trained sequentially, with each focusing on the previously misclassified data. The final prediction is a weighted sum of the individual weak learners, where their weights are determined based on their accuracy (Schapire 2013) AdaBoost is effective for improving the classification performance of weak models and can handle both binary and multiclass classification problems.

Gradient Boosting Classifier (gbc): Gradient Boosting Classifier is a powerful ensemble method that builds an additive model by sequentially fitting new weak learners to the residual errors of the previous model. It starts with an initial prediction (usually the mean of the target variable) and then constructs decision trees in subsequent iterations to predict the remaining errors. The predictions from each tree are

combined, with each tree correcting the errors of the previous ones (Friedman 2001). Gradient Boosting is known for its strong predictive performance and ability to capture complex relationships in the data.

Extreme Gradient Boosting (xgboost): Xgboost classifier is an optimized and scalable implementation of gradient boosting that has gained popularity for its efficiency and high performance. It incorporates several techniques, such as regularized learning objectives and parallel processing, to enhance training speed and reduce overfitting (Chen and Guestrin 2016). Xgboost is widely used in ML competitions and real-world applications due to its excellent predictive accuracy, feature selection capabilities, and ability to handle large datasets.

Light Gradient Boosting Machine (lightgbm, LGBM) classifier is another gradient boosting algorithm that is designed for efficiency and speed. It uses a histogram-based approach for tree building, which significantly reduces memory usage and speeds up training (Ke et al. 2017). Lightgbm is well-suited for large datasets and has become a popular choice in various domains, including recommendation systems, fraud detection, and image analysis. It offers options for both classification and regression tasks and includes features like categorical variable support and early stopping for efficient model training.

During model training, the data was split into 80% (131 out of 164 data) for training and 20% (33 out of 164 data) for validation. Since the data set was somewhat imbalanced, SMOTE was also applied as it was originally developed for imbalanced classification problems (Chawla et al. 2002). Each model was trained both with and without the SMOTE processing technique.

3.2.3 Performance

The performance metric of the best performing models among all the above-mentioned ML models for different data set is provided in Table 8. The F1 score was chosen to select the best performing model and accuracy. Recall and precision was also calculated for the models. Table 8 demonstrates that the best-performing model varies across different datasets. This highlights the importance of training multiple models for such data sets, rather than relying on a single model. To understand the classification strength, a confusion matrix is generated for the three reflectometry test methods in Figure 18.

For the damaged or anomalous class without SMOTE, all three reflectometry test method prediction accuracies were better than with SMOTE calculations as presented in Figure 18. With SMOTE, the prediction accuracy for the un-damaged class slightly improved with a decrease in the accuracy for the damaged/anomalous class.

Table 8. Performance metrics for the best ML models for each data set.

Machine Name	Best ML Model	SMOTE (Yes/No)	Accuracy	Recall	Precision	F1
FDR	lightgbm	No	0.7929	0.8806	0.8241	0.8434
FDR	et	Yes	0.8011	0.8931	0.8254	0.8497
PNNL SSTDR	et	No	0.7775	0.9417	0.7744	0.8481
PNNL SSTDR	ada	Yes	0.7709	0.8403	0.8255	0.8283
LW SSTDR	lr	No	0.8385	0.9042	0.8610	0.8764
LW SSTDR	lr	Yes	0.7923	0.8111	0.8636	0.8289

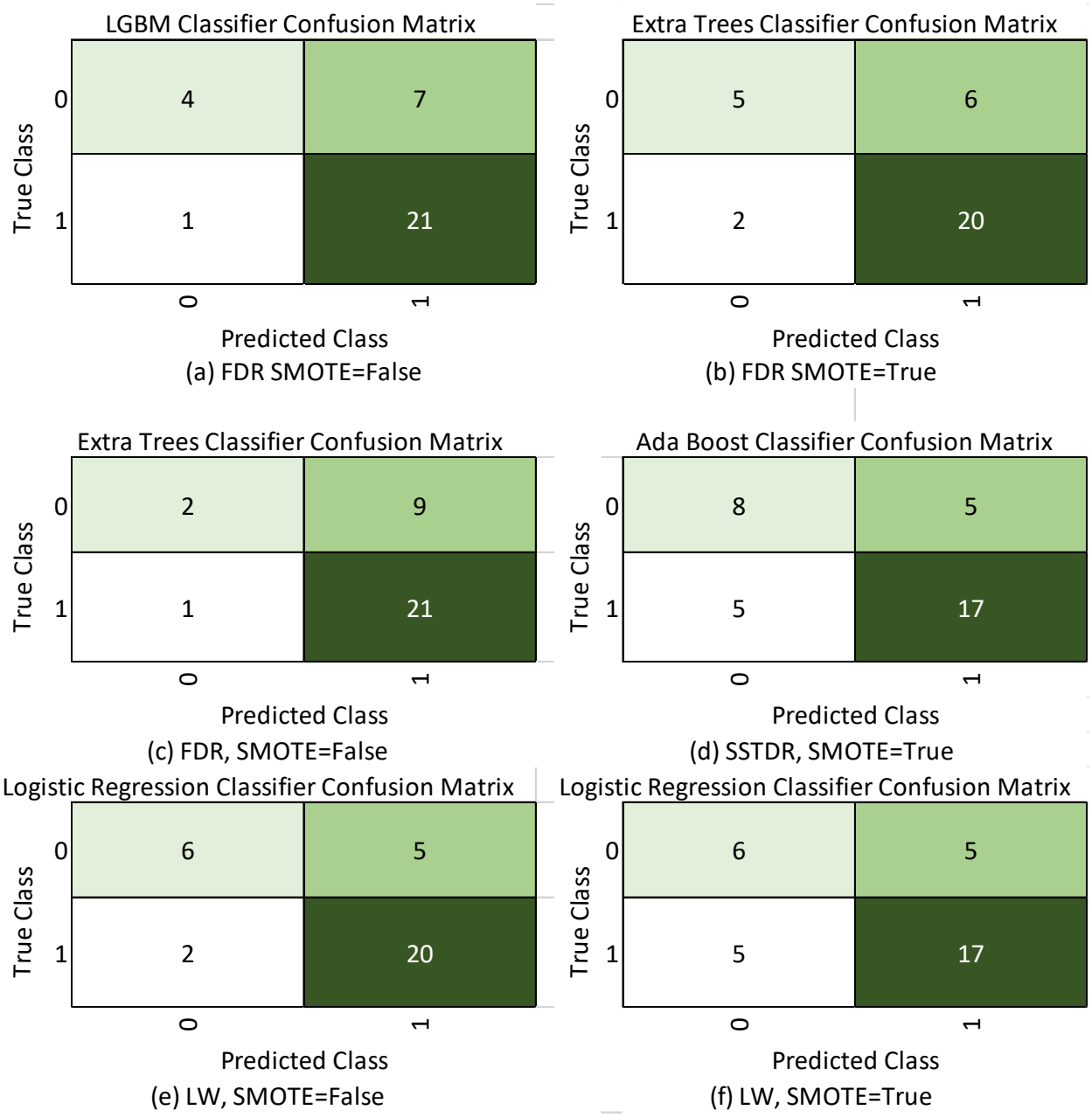


Figure 18. Confusion matrix for different ML models for different dataset.

In order to capture both the anomalous and undamaged class predictive accuracy, a weighted accuracy was calculated as discussed earlier. Table 9 presents the weighted accuracy for the ML models for all the data sets. Table 9 implies that using SMOTE increases the accuracy slightly even though there is a small increase in incorrect predictions for the major or anomalous class except for the LiveWire SSTDR data set. Additionally, overall model performance is slightly better in LiveWire SSTDR and FDR data set compared to the PNNL SSTDR data set probably due to the presence of high noise in the higher frequencies of the PNNL SSTDR data.

Table 9. Weighted accuracy for the ML models for all the data set.

Machine Name	Best ML Model	SMOTE (Yes/No)	Weighted Accuracy (%)
FDR	lightgbm	No	65.90
FDR	et	Yes	68.18
PNNL SSTDR	et	No	56.81
PNNL SSTDR	ada	Yes	65.90
LW SSTDR	lr	No	72.72
LW SSTDR	lr	Yes	65.90

One hypothesis was that due to some noisy data, particularly higher frequency FDR and PNNL SSTDR and lower frequency LiveWire SSTDR that was considered all together for each instrument, the overall weighted accuracies were lower than they would be with only the cleaner data. To test this hypothesis, some of the noisy data was omitted for model training. The 400 MHz data set was omitted for FDR, 200 and 400 MHz sets were omitted for PNNL SSTDR, and the 6 MHz and 12 MHz data sets were omitted for the LiveWire SSTDR data. Therefore, for the new calculation FDR had 123 samples, PNNL SSTDR and LiveWire SSTDR had 82 samples. It is important to mention the 80/20% data split was implemented along with the same model training protocols for the down-selected data as for the full data set. Table 10 presents the weighted accuracy for the new calculations. From the table, it is clear that for all cases, weighted accuracy increased, which signifies the importance of high-quality noise free/less noisy data in order to get a reliable ML prediction accuracy. For instance, the PNNL SSTDR weighted prediction accuracy improved from 56.81% to 69.69% for the case of without SMOTE calculation. For the LiveWire SSTDR, the low-noise weighted accuracy was 87.12% with the SMOTE calculation. This compares favorably to a weighted accuracy of 65.90% with the noisy data included.

Table 10. Weighted accuracy for the ML models for the down-selected data set.

Machine Name	Best ML Model	SMOTE (Yes/No)	Weighted Accuracy (%)
FDR	svm	No	73.95
FDR	ridge	Yes	70.83
PNNL SSTDR	xgboost	No	69.69
PNNL SSTDR	et	Yes	69.69
LW SSTDR	lightgbm	No	83.33
LW SSTDR	lr	Yes	87.12

4. CONCLUSIONS

Cable reflectometry techniques like FDR, LiveWire SSTDR, and PNNL SSTDR show promise to test cables and identify anomalous behavior before cable failure thereby minimizing unplanned outages and allowing repair to be scheduled and managed. Online cable test methods are expected to become more important as isolation technologies mature and utilities move to reduce manual outage related test programs. Cable reflectometry is a promising cable test method, but plots can be difficult for humans to analyze due to baseline noise, low or noisy anomaly response peaks, or large responses from cable ends. The goal of this task was to assess feasibility of ML methods to distinguish undamaged from damaged or anomalous reflectometry data. The assessment considered the 3 reflectometry instruments, multiple frequency bandwidths from each instrument, multiple cable anomalies and test conditions, and both supervised and unsupervised ML approaches.

The PNNL software-controlled laboratory instrument with expanded bandwidth capabilities was set to test cables at 50, 100, 200, 300, 400, and 500 MHz. The FDR was also set to acquire data at 50, 100, 200, 300, 400, and 500 MHz. The LiveWire SSTDR collected data at 6, 12, 24, and 48 Mhz. For unsupervised ML, all 6 PNNL SSTDR and FDR frequencies were used and only the 48 MHz LiveWire data was considered. For supervised data, the PNNL SSTDR and FDR 300 and 500 MHz data sets were excluded and all 4 LiveWire SSTDR frequencies were included. The test matrix included 42 test conditions, although 3 were identically labeled and so reduced the number of supervised cases accordingly. Except as noted above, approximately the same data was used by two independent teams – one applying unsupervised learning methods and the other applying supervised methods to the data. Although analysis methods were not identical or directly comparable, both the supervised and unsupervised ML outputs were encouraging. The primary performance metric was the weighted prediction accuracy.

The unsupervised prediction weighted accuracy was assessed by instrument and by frequency. It performed better at high frequencies with the highest prediction accuracy of 0.84 for the higher frequency FDR, 0.79 for the 48MHz LiveWire SSTDR, and 0.77 for 300MHz PNNL SSTDR.

For supervised ML, the initial analysis considering all the data and combining all frequency bandwidths, the weighted accuracy average across all frequencies for using supervised ML was 0.56 to 0.68. The supervised analysis was repeated with noisier training data removed resulting in improved weighted accuracies of 0.69 to 0.87. The supervised and unsupervised analysis performances were not directly comparable due to differences in the input data and analysis details but they do indicate an encouraging trend. Even with limited and unbalanced data, strong prediction accuracies seem encouraging for further work that should including more data under a wider range of conditions.

5. REFERENCES

Al-Aidaroos, K.M., A.A. Bakar, and Z. Othman. 2010. "Naïve bayes variants in classification learning." 2010 International Conference on Information Retrieval & Knowledge Management (CAMP), 17-18 March 2010.

Chawla, Nitesh V, Kevin W Bowyer, Lawrence O Hall, and W Philip Kegelmeyer. 2002. "SMOTE: synthetic minority over-sampling technique." *Journal of artificial intelligence research* 16:321-357.

Chen, T., and C. Guestrin. 2016. "XGBoost: A Scalable Tree Boosting System." *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. doi: https://doi.org/10.1145/2939672.2939785.

Cortes, Corinna, and Vladimir Vapnik. 1995. "Support-vector networks." *Machine Learning* 20 (3):273-297. doi: 10.1007/BF00994018.

Fisher, E.M. 1936. *Linear Discriminant Analysis. Statistics & Discrete Methods of Data sciences*.

Friedman, J.H. 2001. "Greedy Function Approximation: A Gradient Boosting Machine." *The Annals of Statistics* 29 (5).

Geurts, Pierre, Damien Ernst, and Louis Wehenkel. 2006. "Extremely randomized trees." *Machine Learning* 63 (1):3-42. doi: 10.1007/s10994-006-6226-1.

Glass, S W, A M Jones, L S Fifield, T S Hartman, and N Bowler. 2017. *Physics-Based Modeling of Cable Insulation Conditions for Frequency Domain Reflectometry (FDR).* PNNL-26493, Pacific Northwest National Laboratory, Richland, Washington

Glass, S.W., L S. Fifield, and M Prowant. 2021. *PNNL 31415 PNNL ARENA Cable Motor Test Bed Update.*

Glass, S.W., A. Sriraman, M. Prowant, M.P Spencer, L.S. Fifield, and S. Kingston. 2022. *PNNL-33334 Nondestructive Evaluation (NDE) of Cable Anomalies using Frequency Domain Reflectometry (FDR) and Spread Spectrum Time Domain Reflectometry (SSTDR).*

Glass, S.W., J R Tedeschi, M.P.; Spencer, J. Son, M. Elen, and L.S. Fifield. 2023. *PNNL-34511 Laboratory Instrument Software Controlled Spread Spectrum Time Domain Reflectometry for Electrical Cable Testing.* Richland Washington

Glass, S.W., J. Tedeschi, M.P. Spencer, J. Son, D. Li, M. Elen, and L S Fifield. 2023 *M3LW- 23OR0404025, Extended Bandwidth Spread Spectrum Time Domain Reflectometry Cable Test for Thermal Aging, Low Resistance Fault, and Water Detection.*

Harley, J.B., S. Zafar, and C. Tran. 2023. "Tips for Effective Machine Learning in NDT/E." *Materials Evaluation* 81 (7).

Haykin, S.S. 1999. "Neural Networks: A Comprehensive Foundation." *SPIE*.

Ho, T.K. 1995. "Random decision forests." Proceedings of the 3rd International Conference on Document Analysis and Recognition, 14-16 Aug. 1995.

Hoerl, Arthur E., and Robert W. Kennard. 1970. "Ridge Regression: Biased Estimation for Nonorthogonal Problems." *Technometrics* 12 (1):55-67. doi: https://doi.org/10.2307/1267351.

IEEE. 2015. *{IEEE 383 Standard for Qualifying Electric Cables and Splices for Nuclear Facilities},.*

Ke, G., Q. Meng, T. Finley, T. Wang, W. Chen, M. Weidong, Q. Ye, and T Liu. 2017. "Lightgbm: A highly efficient gradient boosting decision tree." *Advances in neural information processing systems*.

Nassif, A. B., M. A. Talib, Q. Nasir, and F. M. Dakalbab. 2021. "Machine Learning for Anomaly Detection: A Systematic Review." *IEEE Access* 9:78658-78700. doi: 10.1109/ACCESS.2021.3083060.

NRC. 1977. "Regulatory Guide 1.131 Qualification Tests of Electric Cables, Field Splices, and Connections for Light-Water-Cooled Nuclear Power Plants."

Peterson, Leif E. 2009. "K-nearest neighbor." *Scholarpedia* 4:1883.

PNNL. 2022. "Accelerated Real-Time Environmental Nodal Assessment (ARENA) Test Bed" accessed 09192023. https://www.pnnl.gov/arena

Python Pycaret AI/ML Library.

Rasmussen, Carl Edward. 2004. "Gaussian Processes in Machine Learning." In *Advanced Lectures on Machine Learning: ML Summer Schools 2003, Canberra, Australia, February 2 - 14, 2003,*

Tübingen, Germany, August 4 - 16, 2003, Revised Lectures, edited by Olivier Bousquet, Ulrike von Luxburg and Gunnar Rätsch, 63-71. Berlin, Heidelberg: Springer Berlin Heidelberg.

Schapire, R.E. 2013. "Explaining AdaBoost Empirical Inference." In *Empirical Inference*.

Wright, R.E. 1995. *Logistic Regression*. Edited by L.G.Grimm & P.R Yarnold. Washington DC: American Psychological Association.

Wu, Xindong, Vipin Kumar, J. Ross Quinlan, Joydeep Ghosh, Qiang Yang, Hiroshi Motoda, Geoffrey J. McLachlan, Angus Ng, Bing Liu, Philip S. Yu, Zhi-Hua Zhou, Michael Steinbach, David J. Hand, and Dan Steinberg. 2008. "Top 10 algorithms in data mining." *Knowledge and Information Systems* 14 (1):1-37. doi: 10.1007/s10115-007-0114-2.