



Acoustic Emissions Monitoring Study

**Application of Airborne Acoustic Sensing for Motor Generator
Fault Detection in a Laboratory Environment**

2022 TECHNICAL REPORT

Acoustic Emissions Monitoring Study

Application of Airborne Acoustic Sensing for Motor Generator Fault Detection in a Laboratory Environment

All or a portion of the requirements of the EPRI Nuclear Quality Assurance Program apply to this product.

YES



EPRI Project Manager
S. Lopez



3420 Hillview Avenue
Palo Alto, CA 94304-1338
USA

PO Box 10412
Palo Alto, CA 94303-0813
USA

800.313.3774
650.855.2121

askepri@epri.com

www.epri.com

3002023744

Final Report, June 2022

DISCLAIMER OF WARRANTIES AND LIMITATION OF LIABILITIES

THIS DOCUMENT WAS PREPARED BY THE ORGANIZATION(S) NAMED BELOW AS AN ACCOUNT OF WORK SPONSORED OR COSPONSORED BY THE ELECTRIC POWER RESEARCH INSTITUTE, INC. (EPRI). NEITHER EPRI, ANY MEMBER OF EPRI, ANY COSPONSOR, THE ORGANIZATION(S) BELOW, NOR ANY PERSON ACTING ON BEHALF OF ANY OF THEM:

(A) MAKES ANY WARRANTY OR REPRESENTATION WHATSOEVER, EXPRESS OR IMPLIED, (I) WITH RESPECT TO THE USE OF ANY INFORMATION, APPARATUS, METHOD, PROCESS, OR SIMILAR ITEM DISCLOSED IN THIS DOCUMENT, INCLUDING MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE, OR (II) THAT SUCH USE DOES NOT INFRINGE ON OR INTERFERE WITH PRIVATELY OWNED RIGHTS, INCLUDING ANY PARTY'S INTELLECTUAL PROPERTY, OR (III) THAT THIS DOCUMENT IS SUITABLE TO ANY PARTICULAR USER'S CIRCUMSTANCE; OR

(B) ASSUMES RESPONSIBILITY FOR ANY DAMAGES OR OTHER LIABILITY WHATSOEVER (INCLUDING ANY CONSEQUENTIAL DAMAGES, EVEN IF EPRI OR ANY EPRI REPRESENTATIVE HAS BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGES) RESULTING FROM YOUR SELECTION OR USE OF THIS DOCUMENT OR ANY INFORMATION, APPARATUS, METHOD, PROCESS, OR SIMILAR ITEM DISCLOSED IN THIS DOCUMENT.

REFERENCE HEREIN TO ANY SPECIFIC COMMERCIAL PRODUCT, PROCESS, OR SERVICE BY ITS TRADE NAME, TRADEMARK, MANUFACTURER, OR OTHERWISE, DOES NOT NECESSARILY CONSTITUTE OR IMPLY ITS ENDORSEMENT, RECOMMENDATION, OR FAVORING BY EPRI.

THE FOLLOWING ORGANIZATION, UNDER CONTRACT TO EPRI, PREPARED THIS REPORT:

Pennsylvania State University

THE TECHNICAL CONTENTS OF THIS PRODUCT WERE **NOT** PREPARED IN ACCORDANCE WITH THE EPRI QUALITY PROGRAM MANUAL THAT FULFILLS THE REQUIREMENTS OF 10 CFR 50, APPENDIX B. THIS PRODUCT IS **NOT** SUBJECT TO THE REQUIREMENTS OF 10 CFR PART 21.

NOTE

For further information about EPRI, call the EPRI Customer Assistance Center at 800.313.3774 or e-mail askepri@epri.com.

Together...Shaping the Future of Energy®

© 2022 Electric Power Research Institute (EPRI), Inc. All rights reserved. Electric Power Research Institute, EPRI, and TOGETHER...SHAPING THE FUTURE OF ENERGY are registered marks of the Electric Power Research Institute, Inc. in the U.S. and worldwide.



Acknowledgments

The following organization, under contract to the Electric Power Research Institute (EPRI), prepared this report:

Pennsylvania State University
The Applied Research Laboratory
P.O. Box 30
State College, PA 16804

Principal Investigators
K. Reichard
G. Eckenrode
S. Nixon

This report describes research sponsored by EPRI.

This publication is a corporate document that should be cited in the literature in the following manner:

*Acoustic Emissions Monitoring Study:
Application of Airborne Acoustic
Sensing for Motor Generator Fault
Detection in a Laboratory
Environment.*
EPRI, Palo Alto, CA: 2022.
3002023744.



Abstract

This report describes the application of stationary airborne acoustic sensing to provide area-wide monitoring of the health of electrical motors and generators. Airborne acoustic sensing can provide many of the same detection and monitoring capabilities as contact vibration sensing for measuring machinery health and detecting and classifying faults. The project analyzed the design and implementation of airborne acoustic monitoring systems using single microphones and microphone arrays as discussed in EPRI report 3002015880, *Area-Wide Acoustic and Electromagnetic (AEM) Signature Health Monitoring: Phase 2*. The project also analyzed the design of microphone arrays to improve source localization and reduce interference from other noise sources. Data are presented herein which demonstrate the detection of motor faults from laboratory testing of motors with seeded electrical and mechanical faults using machine learning (ML) techniques applied by EPRI and the principal investigators.

Keywords

Acoustic sensing
Condition-based maintenance
Fault detection
Machine learning
Microphone

Deliverable Number: 3002023744

Product Type: Technical Report

Product Title: Acoustic Emissions Monitoring Study: Application of Airborne Acoustic Sensing for Motor Generator Fault Detection in a Laboratory Environment

PRIMARY AUDIENCE: Research and development personnel with an interest in developing, implementing, or commercializing airborne acoustic monitoring systems for electrically-driven rotating machinery.

SECONDARY AUDIENCE: Utility personnel involved in equipment online monitoring, and vendors developing acoustic sensing systems

KEY RESEARCH QUESTION

Can area-wide airborne acoustic monitoring be effective to identify equipment degradation or failure in motors and generators?

RESEARCH OVERVIEW

This project used airborne acoustic sensors consisting of a single microphone and multi-microphone arrays for the detection and classification of faults in electric motors and generators. The project investigated the effects of distance and angle on acoustic signatures to examine the differences between independent “signatures” taken from faulty equipment and baseline data from healthy equipment of similar form, fit, and function. Experiments compared the detection capabilities of a small, 3-element, triangular array to larger linear acoustic arrays. Furthermore, the project examined the theoretical and practical limits of linear arrays to localize acoustic signals to specific points of origination and developed recommendations on array geometry and the number of arrays or subarrays to use based on machinery layout. Artificial intelligence (AI) and machine learning (ML) techniques were also developed for array processing/source detection and localization, and for the detection and classification of mechanical and electrical faults in electric motors. Data was collected from a healthy electric motor and identical motors with seeded electrical and mechanical faults to test the ability of airborne acoustic sensors to detect and classify the fault conditions.

KEY FINDINGS

- Airborne acoustic sensors (microphones) can be used to detect and diagnose mechanical faults such as damaged bearings, damaged shafts, and load imbalance, and electrical faults such as in bad electrical connections, and damaged rotor or stator laminations and windings.
- Multi-element microphone arrays can be used to locate the direction of origination of acoustic signals from motors or other equipment and reject noise from other sources.
- Acoustic array design is affected by the acoustic frequencies being detected which are based on the type of machinery faults and layout of the machinery being monitored.
- The use of airborne acoustic sensors is better suited for mechanical faults, but also shows appreciable success with certain electrical faults.
- Traditional spectral analysis can be used to detect fault frequencies in the measured acoustic data.
- AI and ML techniques can distinguish between different types of faults detected in the acoustic data using trained models.

WHY THIS MATTERS

The results of this research provide guidelines for the design of airborne acoustic sensing systems and the design of multi-microphone arrays. The use of airborne acoustic arrays can aid in the localization of acoustic signatures in an operational environment with multiple machines but can also be affected by the size of the space and the spacing between the machines. AI and ML techniques can be applied to help distinguish between healthy and faulted motors, as well as distinguish between the different types of faults encountered. The ability to detect faults with an airborne acoustic sensor can minimize the cost and complexity of the instrumentation required to monitor the state of machinery through a reduction in the number and type of sensors located on each machine. This may be a preferred approach for certain machinery in the generation environment.

HOW TO APPLY RESULTS

This project focused on the use of airborne acoustic sensors consisting of single microphones and multi-microphone arrays for the detection and classification of faults in electric motors and generators. Earlier reports described the design of microphone arrays and the use of acoustic signal frequency spectra to detect faulted bearings in motors. This report describes the design of a 3-element microphone array and the application of AI and ML to the detection and classification of mechanical and electrical faults in a series of test motors. The number of sensors and the array design (if used) are dependent on the frequencies associated with the faults of interest and the relative locations of the machinery and the acoustic sensors within the space. Utilized data acquisition systems do not need to support full audio bandwidth but can be tailored to the bandwidth of the faults of interest. The AI and ML techniques required to process the acoustic data are suitable for implementation on standard computing platforms such as a laptop computer once the algorithms are sufficiently trained.

Research and development personnel with an interest in developing, implementing, or commercializing airborne acoustic monitoring systems for electrically-driven rotating machinery can follow the example described in this report to design microphone arrays capable of isolating the source of an acoustic signal or reduce the influence of noise from an interfering or distracting source on the measured acoustic signal. Data from a single microphone, or the output from a microphone array, can be used to train machine learning algorithms capable of detecting and classifying motor faults. The report described the processing steps used to create the spectra and preprocess the spectra to reduce the complexity of the ml models. The models used in the fault detector/classifier are built using techniques available in standard ML libraries.

LEARNING AND ENGAGEMENT OPPORTUNITIES

- Fleetwide Monitoring Interest Group
- Integrated Monitoring and Diagnostics Technical Advisory Committee
- I&C Reliability Technical Advisory Committee

EPRI CONTACT: Stephen Lopez, Principal Technical Leader, slopez@epri.com

PROGRAM: Instrumentation and Control, P41.05.03

IMPLEMENTATION CATEGORY: Reference

Together...Shaping the Future of Energy®

EPRI

3420 Hillview Avenue, Palo Alto, California 94304-1338 • PO Box 10412, Palo Alto, California 94303-0813 USA

800.313.3774 • 650.855.2121 • askep@epri.com • www.epri.com

© 2022 Electric Power Research Institute (EPRI), Inc. All rights reserved. Electric Power Research Institute, EPRI, and TOGETHER...SHAPING THE FUTURE OF ENERGY are registered marks of the Electric Power Research Institute, Inc. in the U.S. and worldwide.

Table of Contents

Abstract	V
Executive Summary.....	VII
Section 1: Introduction	1-1
Project Background.....	1-1
Condition Based Maintenance for Electric Motors	1-1
Scope of Work.....	1-3
Prior Research	1-4
Report Organization.....	1-5
Section 2: Application of AI and ML for Fault Detection and Classification.....	2-1
Manual Fault Detection and Classification	2-1
Supervised Learning of Acoustic Frequency Domain Features.....	2-2
Tree-Based Fault Classification.....	2-3
Regularized Regression Methods for Feature Selection in the Acoustic Frequency Domain	2-7
Section 3: Experimental Demonstration and Results.....	3-1
Motor Test Bed.....	3-1
Array Design and Installation.....	3-3
Acoustic Detection and Classification Performance	3-6
Section 4: Summary and Conclusions	4-1
Bearing Fault Identification with Acoustic Emissions.....	4-1
Recommendations for Future Work	4-2
Section 5: References.....	5-1

Appendix A: Glossary of Terms	A-1
Definitions and Nomenclature	A-1
Units of Measurement.....	A-2
Purpose Acronyms	A-2

List of Figures

Figure 2-1 Example classification and regression tree for bearing fault classification.....	2-5
Figure 2-2 Comparison of decision tree classifier performance.....	2-7
Figure 2-3 Decision tree accuracy by "number per segment" for both validation options.....	2-8
Figure 2-4 Coefficient values from Kept LASSO regression by classifier and NPS for severe faults for scaled (true) or non-scaled (false) data.....	2-8
Figure 2-5 Example acoustic spectra for two measurements.....	2-10
Figure 2-6 Spectra showing significant frequencies identified through LASSO regression	2-10
Figure 3-1 Photograph of motor, brake and blower used in the seeded fault testing	3-2
Figure 3-2 Acoustic array installed in the test chamber	3-4
Figure 3-3 Nominal array beam pattern with no steering	3-5
Figure 3-4 Array beam pattern with first null steered at load-end motor bearing.....	3-6

List of Tables

Table 2-1 Performance comparison of classifiers and validation approaches.....	2-5
Table 3-1 Bearing fault frequencies at 1760 RPM motor speed, matching specifications for similar bearings from SKF corporation	3-2
Table 3-2 Acoustic array characteristics	3-3
Table 3-3 Motor faults.....	3-7
Table 3-4 Motor test assets and conditions.....	3-7
Table 3-5 Final classifier results	3-9



Section 1: Introduction

Project Background

This project investigated the use of airborne acoustic sensors consisting of a single microphone and multi-microphone arrays for the detection and classification of faults in electric motors and generators. The project investigated the effects of distance and angle on defect signatures to examine the correlation between independent acoustic “signatures” in the frequency-domain, taken over a variety of distances and angles, to compare them to baseline data from healthy equipment of similar form, fit, and function. Experiments compared results using a small, 3-element, triangular array to larger linear acoustic arrays. The project also examined the theoretical and practical limits of linear arrays to localize acoustic signals to specific points of origin and developed recommendations on array geometry and the number of arrays or subarrays to use based on machinery layout. Lastly, the project examined the application of artificial intelligence (AI) and machine learning (ML) techniques to both the array processing/source detection and localization problem, and the problem of classifying collected acoustic signatures to diagnose specific machinery faults.

Condition Based Maintenance for Electric Motors

Condition based maintenance has become a critical part of the energy industry’s ability to reliably deliver power to the public at an optimal cost. Its implementation can lead to significant increases in uptime and safety, as well as reductions in operations and maintenance costs. All equipment is maintained one of four ways – reactively, preventatively, condition-based, or predictively. Reactive, or break-down maintenance, is the practice of waiting until an asset or machine has completely failed to fix it. One common example of reactive maintenance is changing burned-out light bulbs. Preventative, or time-based maintenance, is the practice of performing specific maintenance actions at specified intervals, based often on calendar time or run time. Condition-Based Maintenance (CBM), is the practice of periodically measuring the health of an asset and performing maintenance only when the health of the asset begins to degrade. An example of CBM is the practice of taking vibration readings on a high-value motor, and only changing the motor bearings when the vibration readings show that there is a significant bearing fault present in the motor. Lastly, predictive maintenance uses condition-based maintenance techniques to assess current health and predicts when maintenance is needed based the

expected degradation due to intended future use. An example of this is using vibration data to predict that there is an incipient bearing fault in a high value motor and then scheduling bearing replacements to be performed during a scheduled outage.

CBM relies on the ability to accurately measure the health of an asset. When dealing with electromechanical machines such as electric motors or generators, faults can originate from either the electrical or mechanical components. Examples of mechanical components include bearings and shafts; examples of electrical components include wiring, windings, and electronic components. Mechanical faults, generally, create vibrations and can be detected with vibration or displacement sensors. Electrical faults typically cause changes in the signal currents or voltages which can be detected using the associated measurement probes. Additionally, there is interaction between the electrical and mechanical forces in an electromechanical system which means that electrical faults may produce mechanical effects and vice versa.

There are several common ways to measure health in electromechanical systems, including vibration analysis (VA), thermography, oil analysis, and motor current signature analysis (MCSA). VA and MCSA can reveal fine details regarding motor health but may involve costly data collection systems, invasive installations, and specific subject matter knowledge to get the full benefit. Both also require one or more physical connections between transducer(s) (either accelerometer(s) for VA or current measuring device(s) for MCSA) and the motor under test. Common solutions involve either a portable data collection system or a continuous monitoring solution. Portable data collectors require a technician to visit each asset under consideration and collect data during rounds, which can significantly limit the amount of data that can be processed. These monitoring solutions require up-front costs to purchase and install multiple sensors, possibly cable, and database solutions. Recent advancements in wireless technology have reduced the costs of monitoring, but they also present other risks, including cyber security and electromagnetic interference.

Mechanical vibrations, such as those measured by VA, have the potential to interact with a surrounding medium (for example, air or water) and produce sound waves. Theoretically, these acoustic signals will carry the information necessary to perform limited VA at some distance if they are collected and analyzed. The conversion efficiency of mechanical vibrations into airborne acoustic signals depends on the frequency of the vibration, the materials properties of the vibrating component, the size of the component, and the properties of the surrounding medium. These factors will determine the amplitude and directivity of the sound produced by an object.

This report focuses on using airborne acoustic measurement without physically contacting the motor under test, which is one of the most substantial limiting factors of standard VA implementation. Depending on the characteristics and intensity of the fault, the distance may be great enough that multiple motors can be monitored from a single collection point. Once emissions from multiple acoustic sources can be collected, a direction localization technique becomes

useful to capitalize on the ability to detect degradation or a fault. Using the known geometric details of the microphone array used to collect the acoustic emissions, several techniques which utilize the phase differences between microphones exist which may be able to identify the direction of a specific acoustic source in relation to the array. Broadly, these terms fall under the signal processing category of beamforming. Combining the acoustic vibration analysis concept with beamforming may allow one data collection point, measuring and processing acoustic signals emitted by multiple sources, to monitor several motors or generators simultaneously. Understanding the potential of this technology, including the analysis required and fault detection capability, is the end goal of this project performed by EPRI.

Scope of Work

The early detection of anomalies and unusual performance prior to the occurrence of equipment faults and failures is critical to cost effective maintenance. However, comprehensive monitoring can be difficult, costly, and in some cases impractical when using sensors that are embedded or placed in direct contact with actual systems or equipment. It may require equipment modifications, as well as the purchase, installation, and maintenance of numerous sensors. A potentially more cost-effective approach is to develop a non-invasive, area-wide monitoring system that, from a distance, can collect signals emitted from a variety of electrical and mechanical equipment and perform data analysis that can reveal potential anomalies or provide an early indication of developing malfunctions.

This project collected and analyzed data from two sets of experiments. The first set of experiments was conducted in the Condition Based Maintenance development laboratory at the Pennsylvania State University Applied Research Laboratory (PSU-ARL). This effort focused on collecting repeatable datasets that represented different stages of degradation for a bearing fault in a specially modified motor-generator test bench. The study focused on using spectral analysis to find potential differences in the acoustic signatures collected with the motor/generator in different conditions (healthy and faulted), statistical analysis to estimate the differences observed in collected spectra, and beamforming analysis to determine the direction of various sources relative to the microphone array. Beamforming analysis was also used to study noise rejection from secondary, interfering, noise sources.

The second set of experiments were conducted by EPRI to expand the dataset and investigate detection capabilities for several additional electrical and mechanical faults. Data were collected from ten, 1-horsepower (hp) (0.74 kW) 460VAC 3-phase induction motors individually subjected to ten common motor faults at increasing levels of severity. All three phases of motor current and voltage were measured to investigate several electrical signature analysis (ESA) systems, and airborne acoustic data was collected using a 3-element microphone array mounted on the wall above the motor. Data was also collected from an

accelerometer mounted on the motor to provide ground truth for the source of the radiated acoustic signals. AI and ML techniques were applied to the recorded acoustic data to discriminate between the healthy motor(s) and the motors with each of the faulted conditions.

Prior Research

Earlier phases of this research project focused on the collection of data to show the utility of acoustic sensing for monitoring mechanical and electrical faults in motors. The results are described in EPRI Report 3002015880, *Area-Wide Acoustic and Electromagnetic (AEM) Signature Health Monitoring: Phase 2* [1]. Extensive acoustic data was collected from a test stand at PSU-ARL using three different bearings to simulate the progression of a mechanical fault. One bearing was unaltered, one featured extensive damage to both races as well as numerous rolling elements, and the third featured limited damage to the outer race. Data was collected with the target machine running alone and with an external noise source present in the room to provide extraneous acoustic emissions to further test the algorithm. Spectral and statistical analysis was performed on all collected data with the purpose of determining if the data collected in both levels of severity were significantly and consistently different than the data collected in the assumed healthy condition, and related only to the bearing condition, not the presence of the external source. Further signal processing was done to test the ability to identify the direction from which acoustic signals were collected, from the extraneous noise source and from the target machine.

Previous work also explored using arrays of microphones to isolate noise sources of interest (e.g., the motor being monitored) from other sources of radiated noise, or to localize the sound from one of potentially many machines within a single facility. Depending on the characteristics and intensity of the fault, the distance may be sufficient, such that multiple motors can be monitored from a single collection point. Once emissions from multiple acoustic sources can be collected, a direction localization technique can be used to capitalize on the ability to detect degradation or a fault. Using the known geometric details of the microphone array, several techniques which utilize the phase differences between microphones were investigated. Broadly, these techniques fall under the signal processing category of beamforming. Combining the acoustic vibration analysis concept with beamforming allowed one collection point, which persistently measures, and processes acoustic signals emitted by numerous sources, to simultaneously monitor the sources for changing acoustic signatures.

Report Organization

This report is organized into the following sections:

- Section 2 describes the application of AI and ML techniques for the detection and classification motor faults in acoustic data
- Section 3 describes the design and installation of acoustic sensors used to collect acoustic data from a series of 10 test motors with seeded faults. This section also describes the performance of AI and ML fault detection and classification algorithms on the acoustic data from the seeded fault testing.
- Section 4 summarizes the findings and conclusions from the research.



Section 2: Application of AI and ML for Fault Detection and Classification

Manual Fault Detection and Classification

Bearing faults generate vibrations at frequencies determined by the bearing geometry and the motor speed at levels correlated with the severity of the vibration. The vibrations generated by the bearing fault radiate into the air and are detectable using airborne acoustic sensors (microphones). Note, however, that different structures have different radiation efficiency and the ability to detect surface vibrations with airborne acoustic sensors must be experimentally verified.

EPRI report 3002015880, *Area-Wide Acoustic and Electromagnetic (AEM) Signature Health Monitoring: Phase 2* [1], described the use of a microphone array and beamforming to improve the signal-to-noise ratio of the airborne acoustic measurements and isolate the source of acoustic radiation. In general, the process for detecting and classifying faults based on airborne acoustic measurements is the same regardless of whether the system uses a single acoustic sensor or an array of sensors. Beamforming and source localization can be separated from fault detection and classification processing, which decouples the two processing steps. The output of the beamforming processing is presented as a single input signal to the fault detection and classification process the same as a single microphone output.

For simple mechanical faults, such as a bearing raceway spall, fault detection relies on the detection of energy at a single frequency. In a simple system, the condition monitoring software can detect a fault by identifying a frequency (or set of frequencies) in the spectrum where the energy exceeds the baseline and background "noise" by a predefined threshold. In cases where the background noise exceeds the level of the signal generated by the fault at a particular frequency, the signals can be averaged in time to improve the signal-to-noise ratio. There is also an inherent improvement in the signal-to-noise ratio resulting from calculating the frequency spectrum. The underlying assumption is that faults present themselves as narrowband (single frequency) components in the broader acoustic signal. The same assumptions are generally used in electrical signature analysis (ESA). The exception to this model is when the fault causes modulation of an underlying frequency resulting in amplitude or frequency

modulation, which distributes the fault-generated signal energy across a range of frequencies. Nevertheless, even in the case of amplitude or frequency modulation, the detection of the fault (used to trigger further analysis) is typically based on detecting fault-related frequency components in the measured signal.

The statistical characteristics of the signals and noise, characterized by their respective probability density functions, can be used to calculate the threshold used to detect the signal in the presence of broadband noise and maintain a desired probability of detection (the probability that a signal is present when the energy at a particular frequency exceeds the background noise by the threshold) and probability of false alarm (the probability that the exceedance of the threshold is due to a random large contribution of the noise at the frequency of interest). This process is known in the signal processing literature as Constant False Alarm Rate (CFAR) detection [2].

This approach to signal detection and classification works well for detecting a single type of fault, an assessment of the fault severity and ultimately a prediction of the remaining useful life or time until the equipment must be taken out of service. When a piece of equipment may be subject to several different types of faults, as is usually the case, monitoring algorithms must be applied that can distinguish between the different types of faults or employ a sequence of fault detection algorithms that attempt to detect one fault at a time. For example, in the case of an electric motor, the condition monitoring software may need to process the data once for each fault: bearing inner race, bearing outer race, stator winding, rotor winding, etc. Even if all of the fault detection algorithms use the same sensors and pre-processing (e.g., calculation of the frequency spectrum), the fault detection and identification process must be repeated for the specific frequencies associated with each fault.

Multiple faults can be detected within a single measurement (e.g., spectrum) by applying templates to the spectra matching the fault's spectral signature. This is where it is beneficial to leverage the power of ML to detect and classify faults from a particular type of measurement.

Supervised Learning of Acoustic Frequency Domain Features

There are many ML techniques for machinery fault detection. This project developed, implemented and tested several different techniques. Each was based on processing acoustic spectra with the goal of identifying whether the data were from a healthy motor (no fault) or from a motor with one of several known types of faults. The fault detection and classification steps were combined into a single process by including the no-fault condition as a class. Alternatively, ML can be applied to time-series data, however, using spectra as the input makes it easier for subject matter experts in machinery vibration to compare data from different motor conditions and provide an assessment of how much one signature differs from another.

Supervised learning refers to techniques in which the ML models are trained using data that are labeled with the associated fault. Supervised learning is generally considered "expensive" because the data must be labeled with the correct class. In most cases, labeling the data must be done by human subject matter experts, based on an independent assessment of the machinery condition. Unsupervised learning techniques, by comparison, learn or separate the data into different classes without knowledge of what those classes are. In some cases, the unsupervised learning process will separate the data into the same classes as a human subject matter expert. In other cases, the unsupervised learning may identify more or fewer classes in the data. Unsupervised learning is generally considered less expensive because it doesn't require independent analysis by human experts. Hybrid learning techniques attempt to use unsupervised learning to separate the data into classes and then use experts to label the learned classes with appropriate labels corresponding to the classes of interest.

In addition to determining the type of fault affecting a machine, it is often desired to determine the fault severity. Sometimes, it is possible to determine fault severity by simply measuring the strength of the signal components attributed to the fault (the fault frequency). In other cases, the frequency content of the signal may change as the severity of the fault increases, as may be the case when modulation increases. Another approach to determining the severity of the fault, is to include examples of the same fault at different severity levels as separate classes in the data (e.g., no fault, mild inner race fault, severe inner race fault, etc.), or the data can be processed again to determine the severity after the fault class has been determined (e.g., inner race fault). This was the approach used in this research.

Tree-Based Fault Classification

This project developed a tree-based fault classifier using acoustic spectra as the input. Regression analysis was applied to the acoustic spectra to identify the subset of frequencies within the data that provided the best separation between fault classes. The data analysis begins with a time recording of the measured acoustic signal. The spectra are generated by calculating the Fast Fourier Transform (FFT) of the time-series data and using Welch's Power Spectral Density (PSD) method to estimate the power spectral density [3,4]. To train the ML models, the PSD values were flattened. That is, each frequency component in the spectra became a column in a data table or matrix, and one row was added to the matrix for each data collected spectrum.

ML models were trained with spectra whose magnitudes were represented in their native linear amplitudes (e.g., Pascals, for microphone data) or in decibels. One general trait of machine learning models is that they work better on data whose range of values is compressed; consequently, data are often normalized or scaled to a predefined range. This is because models are often trained, through which model parameters are adapted, proportional to an error between input and output parameters. If all of the input values have the same normalized range,

then it prevents the model adaptation from being biased toward inputs with larger values (and therefore potentially larger errors). Because of this, better results were found using spectra whose amplitude was expressed in decibels instead of being linear.

From hyper parameters to the decision of scaled data, there are many features that need to be accounted for in the training of a ML model. In the models used in this project, the data was broken up for all permutations that were characteristic to the dataset. All methods, whether including the occurrence of scaling using a standard scaling approach, feature engineering, decibel scaling, and the number of samples per segment (typically 2560 samples per segment), were used in identifying fault health with the decision tree or random forest algorithms. The decision tree algorithm was chosen due to its ability to accept variable data types like categorical (discrete) or continuous [6] and its ability to make accurate predictions if handed a large volume of high-quality data [7]. The random forest algorithm is usually perceived as an extension of the decision tree algorithm since, as a forest has many trees, the random forest algorithm is comprised of many decision tree algorithms. Similar to the decision tree algorithm, the random forest algorithm takes on the properties of the decision tree algorithm and denotes each decision tree to evaluate the different classes or targets mentioned below.

To ensure that all proposed classifiers were judged fairly, data sets were carefully managed to ensure that all different classifiers were presented with identical, paired, sets of train and test data. The target/response column was the health status of the system, which has three possible values - healthy, mild, and severe. Training and test data sets were selected by randomly extracting data sets from each class.

Figure 2-1 shows an example classification tree for bearing health based on data collected for the test motor/generator described in EPRI Report 3002015880, *Area-Wide Acoustic and Electromagnetic (AEM) Signature Health Monitoring: Phase 2*. Starting with the block at the top, the classifier assumes the bearing is healthy, then compares the data to decision criteria to decide if the data indicate a severe fault or a healthy condition. At the second level, the severe case is separated into severe and mild fault cases, while the healthy case is separated again into healthy and severe. The colors of the blocks represent the classification at that level in the tree: purple → healthy, green → severe, orange → mild. The shade of the color in a box correlates to the Gini index for the decision. The Gini index is a measure of how well a decision tree is split [12]. The Gini index ranges from 0 to 0.5; a lower score reflects greater certainty in the decision between classes at the next level (darker shade), while a higher index value indicates lower confidence (lighter shade).

The actual decision criteria in the decision tree are learned by training the model on data representing the different states or classes. The criteria listed in the blocks in Figure 2-1 refer to indices and values in the acoustic spectra and are not meant to be human readable. The decision trees are trained from the top down (relative to the picture in Figure 2-1) or from the base to the branches and leaves. Additional branches are added during training to improve the accuracy of the classifier. One thing that is interesting to note is that there are several branches which can lead to the same final health assessment. For example, Figure 2-1 shows three possible paths to the final decision of a healthy bearing, four paths for a mild fault, and three paths for a decision that the fault is severe. A traditional, case-based decision tree, on the other hand, would have a single set of criteria for each decision and just three branches.

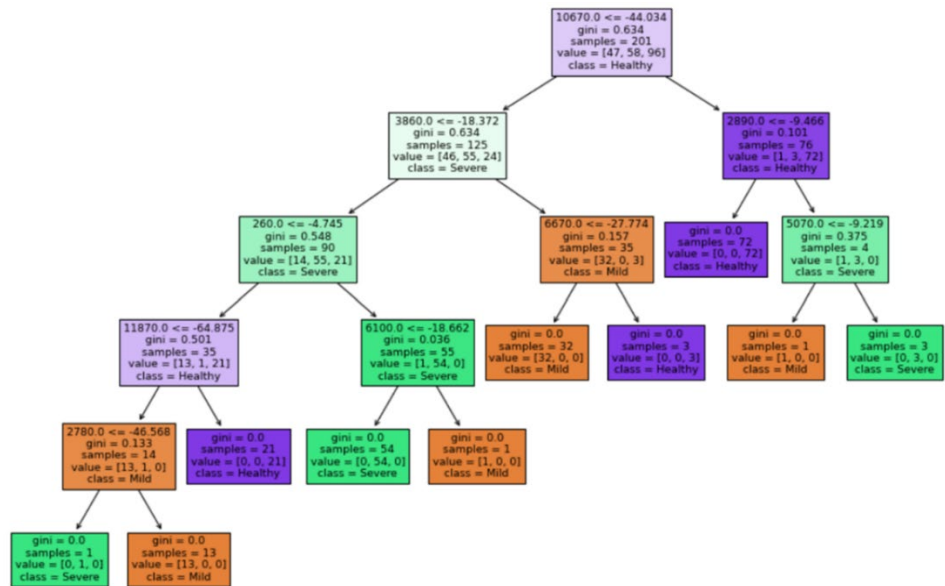


Figure 2-1
Example classification and regression tree for bearing fault classification

Table 2-1
Performance comparison of classifiers and validation approaches

Scale	NperSeg	Return (dB)	Classifier	Normal/Validate	Accuracy	RMSE
TRUE	2560	TRUE	Decision Tree	Normal	0.92157	0.2549
TRUE	2560	TRUE	Decision Tree	Validate	0.95635	0.10318
TRUE	2560	TRUE	Random Forest	Normal	1	0
TRUE	2560	TRUE	Random Forest	Validate	0.97222	0.0754
TRUE	2560	FALSE	Decision Tree	Normal	0.94118	0.11765
TRUE	2560	FALSE	Decision Tree	Validate	0.96429	0.08333
TRUE	2560	FALSE	Random Forest	Normal	0.92157	0.2549
TRUE	2560	FALSE	Random Forest	Validate	0.89683	0.26984
TRUE	5120	TRUE	Decision Tree	Normal	0.98039	0.07843
TRUE	5120	TRUE	Decision Tree	Validate	0.948413	0.146825

The experiments defined in Table 2-1 were performed to exhaustively compare the decision tree algorithm to the random forest algorithm under different validation approaches. The column "Normal/Validate" denotes two separate approaches for training - for observations denoted as "normal", the full training data set was used to fit one, single, classifier, while rows marked "validate" used a 5-fold cross-validation approach to estimate the optimal classifier given the training data. Cross-validation partitions the training data set into $k=5$ equal segments, and actually fits k different classifiers before deciding on optimal parameters given a particular set of training data. Cross-validation is intended to reduce the risk of overfitting a model, which is a significant concern for tree-based classifiers. Given the results in Table 2-1, the additional effort of cross-validation does not generally improve performance, which indicates that overfitting is not a significant concern for this analysis.

Usually, accuracy is a good metric to gauge how well your model performed, but accuracy only justifies the short-term negligence of the model. Error is a strong indicator of how a model will perform in the long run; hence, the reduction of error or the observance of a small error is key to a model's long-term performance. In this case, the testing of error was done using residual mean square error (RMSE), and the first 10 values provide an interesting insight. One would expect an inverse or Bayesian relationship (if accuracy goes up, RMSE must go down and vice versa). This is shown to be true in the third row where accuracy is 1 and error is 0, hence showing the tradeoff. A good example of a bad data point, where the classifier should not be utilized, is in the seventh row where accuracy is 0.92157 but the RMSE is 0.2549. This illustrates the difference between the two metrics, including the weighting importance of incorrect predictions on the RMSE.

Figure 2-2 shows that no method completely outperforms the others. Each bar is designated by its classifier and the presence of validation during training. For example, "DecisionTree_Validate" represents scores of the decision tree classifier, when trained with cross validation. For each classifier and validation technique, models were fit using ten different values of the number of points per segment parameter in the Welch's PSD estimate (2560, 5120, 7680, etc.), both with and without the results scaled to decibels. This comparison was independent of whether the training was done using cross validation. Figure 2-3 demonstrates this type of relationship, where NPS refers to the "points per segment" or "number (of points) per segment" in the PSD. For two different NPS values in Figure 2-3 and Figure 2-4, the models which did not use cross validation outperformed the models that did use cross validation. The results in Table 2-1 indicate that overfitting, which often occurs with the normal validation approach, is not expected to be a problem in this case.

Regularized Regression Methods for Feature Selection in the Acoustic Frequency Domain

Regularization is a process that aims to improve a model's performance throughout training, testing, and prediction. The regularization technique used for feature engineering during this effort was the Least Absolute Shrinkage and Selection Operator (LASSO), which combines variable selection and regularization [8]. In the current case, LASSO's regularization and feature selection methods are crucial to the identification of outlying points since it shrinks the regression coefficients, for unnecessary frequency features, to zero. If the LASSO algorithm determines that a features coefficient can be shrunk to zero, then the feature is removed. LASSO utilizes the L_1 least squares penalty for optimal loss. LASSO also minimizes an objective function with a parameter, λ , that controls the strength of the L_1 penalty. By default, λ is set to one so that the loss penalty is not influenced by strength. The application of LASSO regression reduces the number of inputs (frequencies in the PSD) for the ML model instead of using the full spectrum.

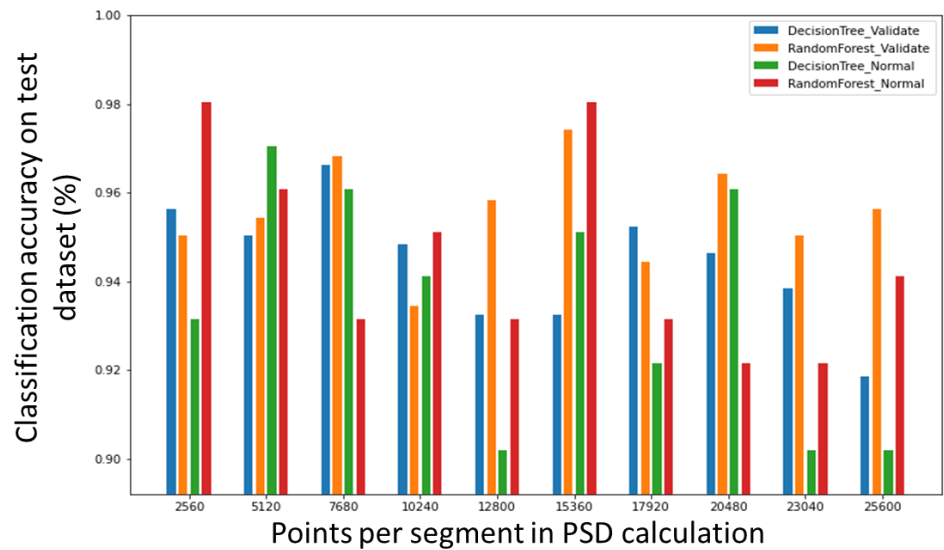


Figure 2-2
Comparison of decision tree classifier performance

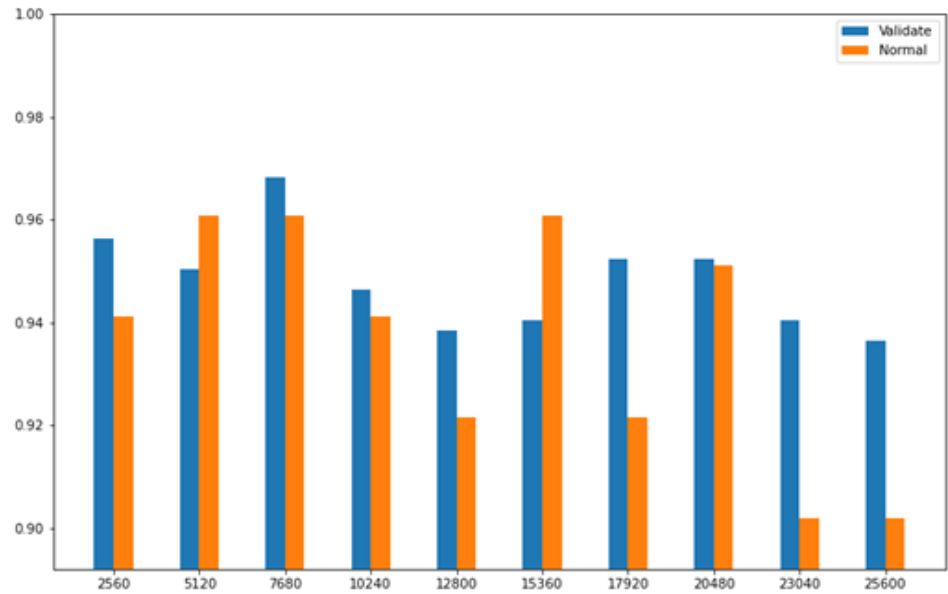


Figure 2-3
Decision tree accuracy by "number per segment" for both validation options

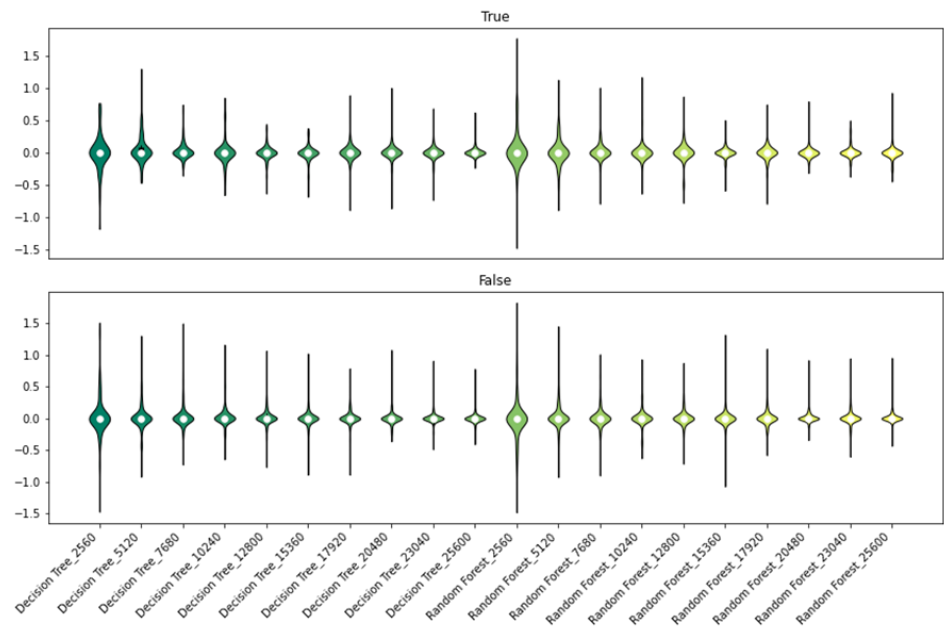


Figure 2-4
Coefficient values from Kept LASSO regression by classifier and NPS for severe faults for scaled (true) or non-scaled (false) data

In Figure 2-4, violin plots were constructed to understand the distribution between the scaled and non-scaled data. As described above, the points that were shrunk to zero as a result of LASSO regression were removed, and the points that were kept revolve around zero in each plot, either negative or positive. These values were kept indicating their relative importance. Typically, the farther away from zero a dataset is, either negative or positive, the more important the feature is since fewer values are closer to zero. Figure 2-4 shows that, “Random Forest_2560” and “Decision Tree_2560” were the most important classifier and NPS combinations for both scaled and non-scaled data. While Figure 2-4 shows that there are other classifier and NPS combinations worth consideration, these two designations clearly stand out amongst the others. Ultimately, the “Random_Forest_2560” designation proved to be much more valuable than any other combination for scaled and non-scaled data.

Larger numbers of NPS correspond to more data and more inputs to the classifier. An interesting observation is that there is a gradual collapse in distribution size for each classifier following 2560 NPS which was the highest (largest spread in Figure 2-4). Generally, this phenomenon appears to be true, especially for the random forest classifier with random sampling for data redundancy.

Figure 2-5 and Figure 2-6 illustrate the results of LAASSO regression applied to the acoustic spectra from the motors. The purple boxes in Figure 2-5 denote the frequency range of interest to which the regression was applied (0-2000 Hz). Having fewer features (frequencies) has two primary benefits. First, it helps reduce the tendency to overfit the model during classifier training. Second, reducing the number of features reduces the amount of data required to train the classifier and the time required to train the model.

Using LASSO regression to reduce the number of features used in the final classifier resulted in a negligible difference in classification accuracy compared to using the full spectrum but reduced computational complexity. For all experiments conducted, less than 10% of features (i.e., frequencies) were ever retained as significant. The percentage of features retained decreases logarithmically as the absolute number of features increases (controlled by number of points per segment). Of the 2,000 points per segment, typically ~2% of presented features were retained.

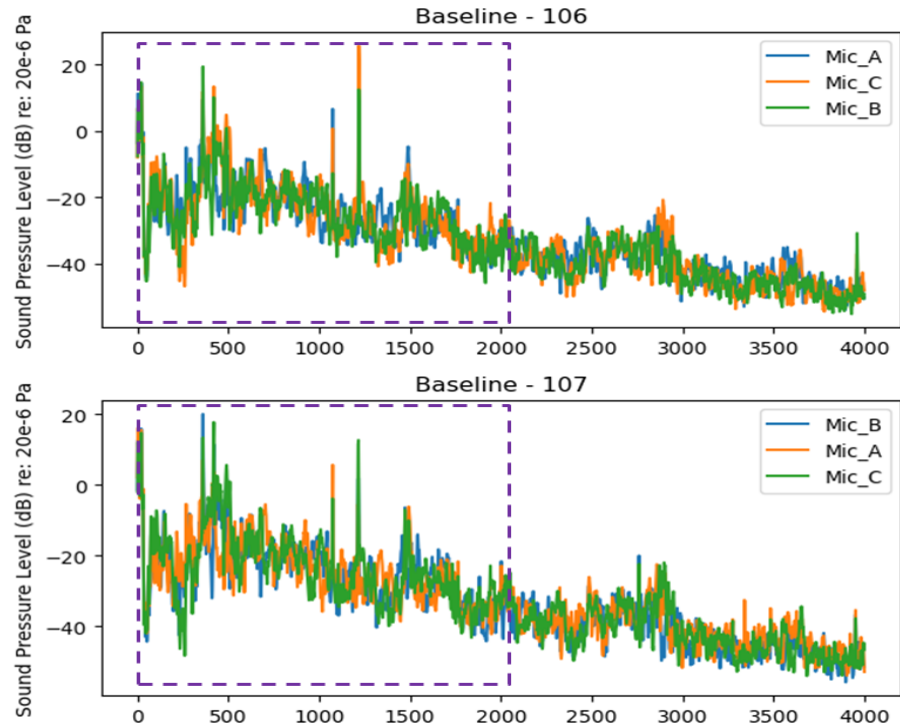


Figure 2-5
Example acoustic spectra for two measurements

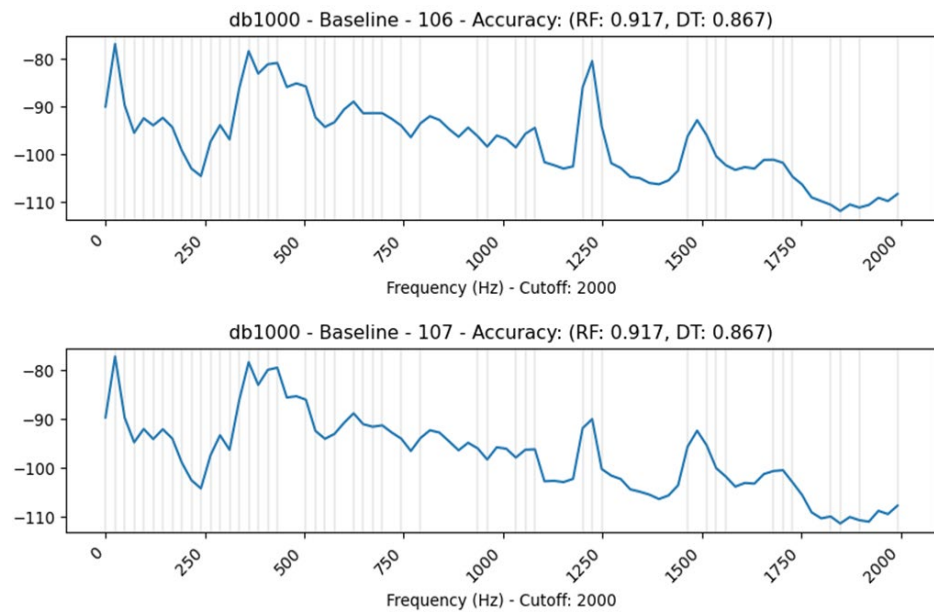


Figure 2-6
Spectra showing significant frequencies identified through LASSO regression (gray vertical lines)

Section 3: Experimental Demonstration and Results

Motor Test Bed

The final phase of the project included the deployment and testing of a microphone array during motor testing conducted on EPRI's behalf by Analysis and Measurement Services Corporation (AMS). The purpose of the testing was to evaluate electric motor condition monitoring technology through laboratory testing of commercially available ESA systems. A custom motor test chamber was designed and constructed to perform fault testing of 1 horsepower (hp) induction motors. Ten common types of faults were artificially induced in the motors at increasing levels of severity while current and voltage data was collected by several ESA systems to determine if the analysis can detect the fault. A three-element microphone array and accelerometer were also installed to supplement the voltage and current data collected for each motor. Data from the microphone array were recorded and post-processed to assess classifier performance. Details of the test setup and procedures are described in EPRI Report 3002020869, *Electrical Signature Analysis (ESA) for On-Line Equipment Condition Monitoring: Fault Detection Testing* [5].

The motor testbed consisted of a 1 hp (0.74 kW), 460 volt three-phase motor and a 25 ft-lb (33.9 Nm) magnetic brake with forced air. Fifteen of these motors were procured for this project: eleven for fault testing and four spares. The brake was controlled by a DC power supply with variable output current for adjusting the braking force on the motor shaft to ensure desired loading. The motor and brake were mounted on an 80/20 frame connected to a wooden base structure measuring 2 x 6 x 0.5 ft (0.61 x 1.8 x 0.15 m) and containing over 350 lbs (158.8 kg) of sand to dampen vibration.

Figure 3-1 shows a photo of the motor, brake, and fan. The investigators used information provided by AMS, regarding motor specifications, to identify expected acoustic features and critical frequencies associated with common bearing faults. The motor manufacturer provided specifications for the bearings mounted on the drive and non-drive ends. The manufacturer's catalog provided the necessary information to calculate the bearing fault frequencies at the motor operating speed (1760 RPM). Table 3-1 lists the calculated frequencies at 1760 RPM for common bearing faults: outer race (BPFO), inner race (BPFI), fundamental train or cage (FTF), and ball (BSF). Although the bearings differ in size, they have proportional dimensions, resulting in the same fault frequencies.

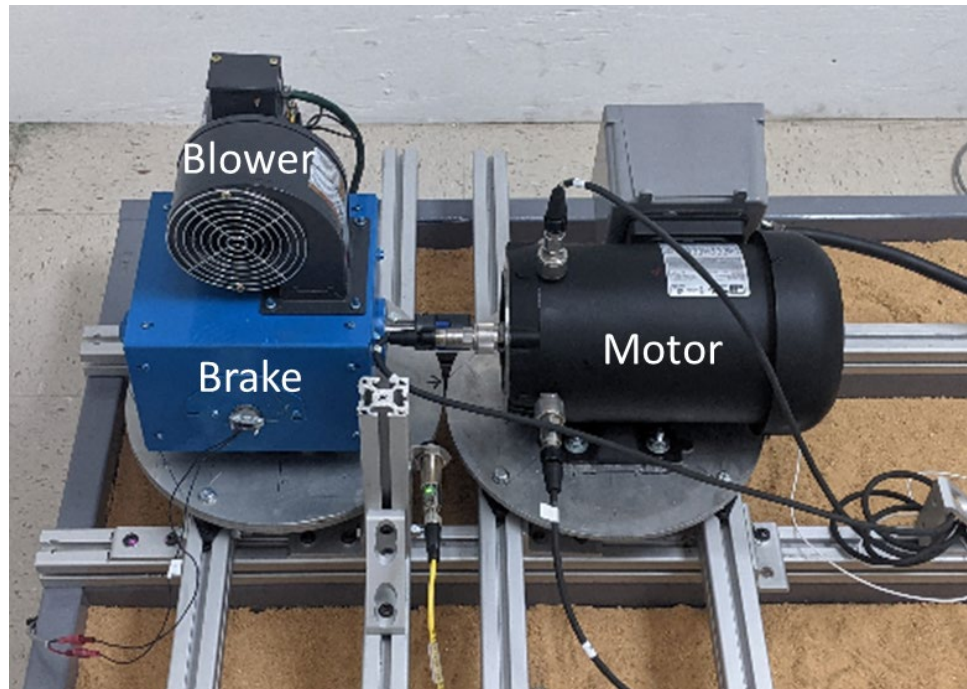


Figure 3-1
Photograph of motor, brake and blower used in the seeded fault testing

Table 3-1
Bearing fault frequencies at 1760 RPM motor speed, matching specifications for similar bearings from SKF corporation [9]

SKF 6204-2Z (Drive-end Bearing)	
BPFO	89 Hz
BPFI	145 Hz
FTF	11 Hz
BSF	58 Hz
SKF 6202-2Z (Non-drive-end Bearing)	
BPFO	89 Hz
BPFI	145 Hz
FTF	11 Hz
BSF	58 Hz

Array Design and Installation

As described in the previous project report [1], the design of an acoustic beamforming array depends on the frequencies of interest and the desired angular resolution of the array. The range of frequencies over which the array can localize a source is determined by the spacing of the array elements and the highest frequency of interest. The bearing fault frequencies show that the maximum frequency of interest is 145 Hz. If detectable, electrical motor faults are expected to present themselves at multiples of the motor speed (29 Hz).

A microphone array was designed to collect airborne acoustic data from the motors under test. Figure 3-2 shows a photograph of the acoustic array installed in the test chamber. The array was a three-element linear array with the microphones spaced 2 ft (0.61 m) apart. The overall size was limited by the dimensions of the room, while the spacing was determined by the upper range of the frequencies of interest. Acoustic foam was mounted on the wall below the array and above the microphones to reduce reflections of the acoustic signals from the walls and ceilings which can introduce multi-path effects and distort the phase of the acoustic signals at each microphone. Additional details on the array design and performance specifications are listed in Table 3-2. In addition, details on microphone array design can be found in standard acoustics texts [10] and can be calculated using online tools or analysis software such as Matlab® [11].

Table 3-2
Acoustic array characteristics

3-element linear array specifications	
Number of elements	3
Spacing	2 ft (0.61 m)
Total length	6 ft (1.83 m)
Bandwidth	290 Hz (2x bearing BPFI)
First-null beam width	80°
Main lobe half-power beam width (HPBW)	3.8 ft (1.16 m) at 6 ft (1.83 m)

Figure 3-3 shows the nominal beam pattern from the array (all microphones simply added together), and Figure 3-4 shows the array beam pattern with first null in the array steered toward the load-end motor bearing. The main lobe beam width gets narrower at higher frequencies and the nulls in the beam pattern disappear at lower frequencies. Theoretically, the main lobe of the array beam pattern can be steered toward one end of the motor or the other to help determine whether a bearing fault frequency is coming from the drive end or the non-drive end, or to help localize sounds from the motor or the blower (if they had the same frequency content). Steering 40 degrees will aim null center at the motor. However, because of the small space and reflections from the walls, there was little difference in the array output with any beam steering.



Figure 3-2
Acoustic array installed in the test chamber

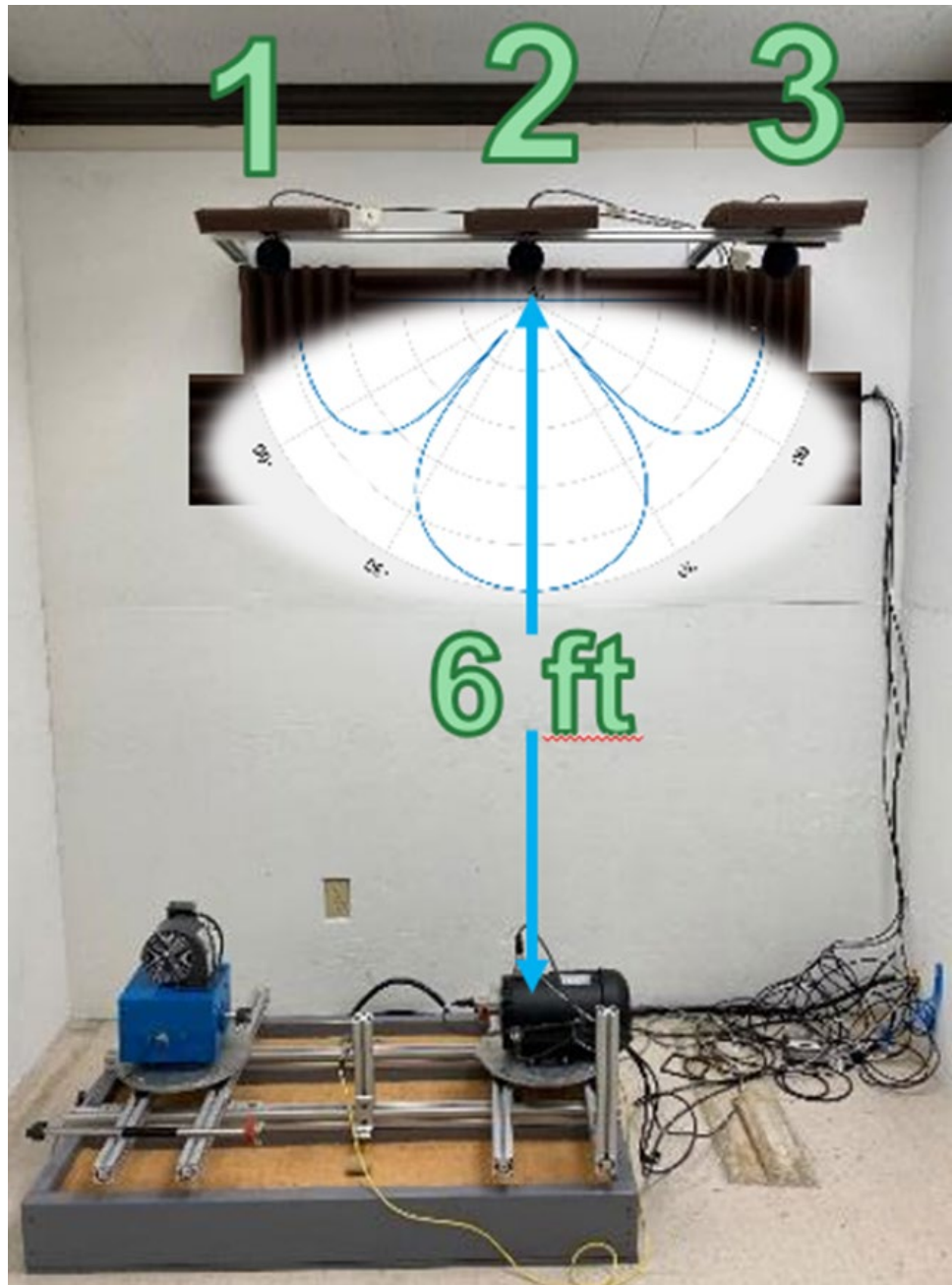


Figure 3-3
Nominal array beam pattern with no steering



Figure 3-4
Array beam pattern with first null steered at load-end motor bearing

Acoustic Detection and Classification Performance

The goal of the test was to collect data from identical motors with ten common faults and several levels of degradation, listed in Table 3-3. It can be difficult to recreate naturally occurring faults in test motors, and in some cases, it took several tries to create the desired fault conditions; hence the test included multiple instances of the electrical connection and stator winding faults. Table 3-4 lists the actual test assets and their associated condition, which resulted in a total of thirty-three test conditions across eighteen test assets. The investigators

collected 357 data sets over the course of the testing. EPRI Report 3002020869, *Electrical Signature Analysis (ESA) for On-Line Equipment Condition Monitoring: Fault Detection Testing* [5] contains additional details on the motor selection, methods for inducing the faults, and data collection.

Table 3-3
Motor faults

Item	Fault	Degradation Description
1	Bearings (front and rear)	Wear and tear
2	Rotor Fan	Looseness and Improper Cooling
3	Electrical Connections	Loose Connections and High Resistance
4	Winding Connections	Brazed Windings and High Resistance
5	Stator Windings	Insulation Degradation
6	Stator Laminations	Lamination Degradation
7	Rotor Laminations	Stress Fatigue Cracks
8	Rotor Bar	Cracked or Loose Joints
9	Shaft	Cracking of Shaft material
10	Load Imbalance	Imbalance on Shaft at the Load

Table 3-4
Motor test assets and conditions

Item	Fault	Degradation Description
1	Rotor Fan	<ol style="list-style-type: none"> Baseline Bend/break a blade Remove retaining clip, making the fan blades loose Split the plastic between the blades so it is not so tight around the bar, making it slip / with no pin
2	Electrical Connections	<ol style="list-style-type: none"> Baseline
3	Electrical Connections 2	<ol style="list-style-type: none"> Baseline
4	Electrical Connections Actual	<ol style="list-style-type: none"> Baseline Bad Crimp on one of the phases Pool Cleaner Diminished Connector Pool Cleaner Diminished Wire Aged Cable 1601 Hours Aged Cable 2250 Hours
5	Winding Connections	<ol style="list-style-type: none"> Baseline Bad solder with HCL corroded cable
6	Stator Windings Bad 1	<ol style="list-style-type: none"> Baseline
7	Stator Windings Bad 2	<ol style="list-style-type: none"> Baseline

Table 3-4 (continued)
Motor test assets and conditions

Item	Fault	Degradation Description
8	Stator Windings Official	1. Baseline 2. Turn-to-turn short
9	Stator Laminations	1. Sand down laminations 2. Bend laminations
10	Rotor Laminations	1. Baseline 2. Sand down laminations 3. Bend laminations
11	Front Bearing	1. Corrosion (HCL introduced to bearing)
12	Rear Bearing	1. Heat grease out of bearing
13	Rotor Bar	1. Baseline
14	Front Bearing	1. Heat grease out of bearing
15	Stator Laminations	1. Acid attack
16	Front Bearing	1. Introduce blasting sand
17	Rotor Bar	1. Crack bar 2. Undisclosed
18	Shaft	1. Baseline 2. Shaft 1

In general, acoustic monitoring is expected to perform best on faults which induce vibrations in the motor. Faults that induce imbalance can also generate noise from gross motion of the motor on its mounts. Faults of a purely electrical nature do not generate noise as efficiently, but if they change the rotational speed of the motor or introduce modulation of the speed, the effect may be detectable as changes in the vibrations generated by the rotation rate.

The principal investigators used random sampling of 300 of the 357 data collection samples for training and testing the classifier. The classifier was trained on a random subset of 75% of the selected data and then tested against the remaining 25% (73 samples). Figure 3-5 shows the confusion matrix of the classifier results, for one training and testing cycle. The confusion matrix compares the predicted condition to the true condition. The classifier used dB scaling of the spectra with 3000 frequencies in the PSD.

Overall performance for the case shown in Figure 3-5 was 97%. The results represented in the confusion matrix correspond to a total of seventy-one cases representing thirteen different faults. The system correctly classified eight of the thirteen conditions with 100% accuracy. There was some confusion in one of the data sets from the fan fault between a fan fault and an electrical connection fault, and confusion between the stator lamination and stator winding faults; however,

all stator related faults were classified as stator faults. One additional thing to note in the results is that there was no healthy motor case. Therefore, it is possible that one of the electrical faults could be confused with a healthy motor if it did not result in significant acoustic radiation.

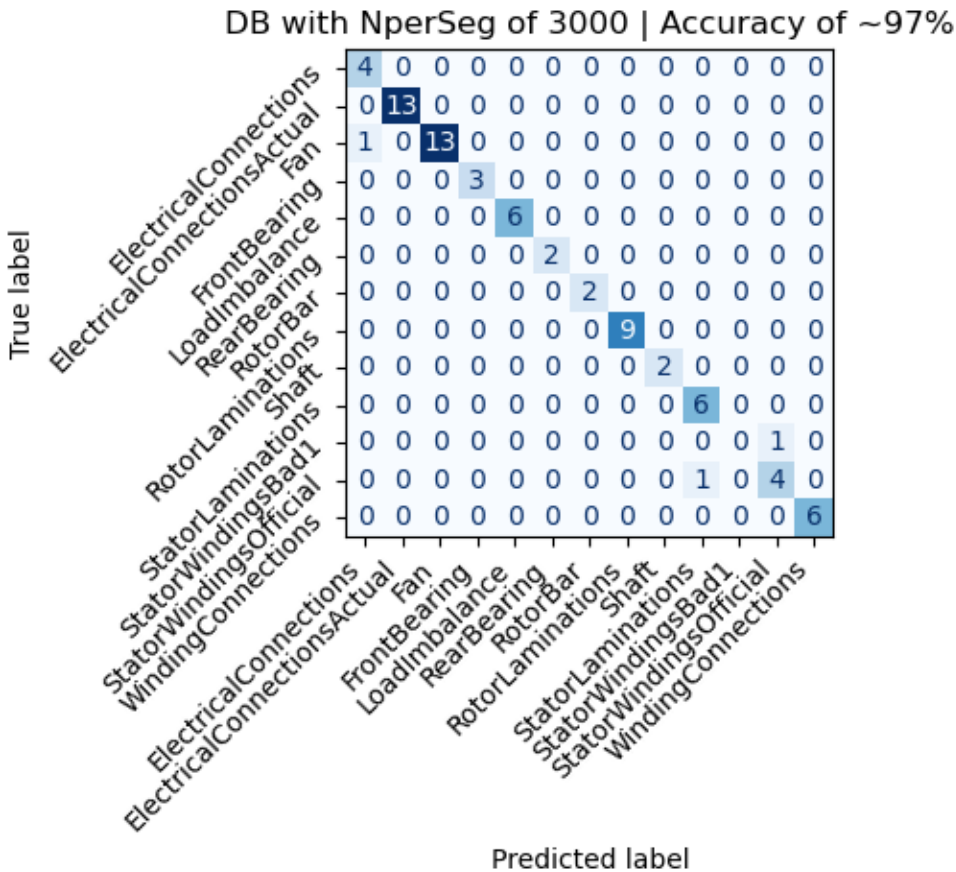


Table 3-5
Final classifier results



Section 4: Summary and Conclusions

Bearing Fault Identification with Acoustic Emissions

This project used airborne acoustic sensors consisting of a single microphone and multi-microphone arrays for the detection and classification of faults in electric motors and generators. The project investigated the effects of distance and angle on acoustic signatures to examine the differences between independent “signatures” taken from faulty equipment and baseline data from healthy equipment of similar form, fit, and function. Experiments compared the detection capabilities of a small, 3-element, triangular array to larger linear acoustic arrays. Linear arrays (both uniformly spaced and incorporating subarrays) are easier to model and scale their design for different machinery layouts. The project also examined the theoretical and practical limits of linear arrays to localize acoustic signals to specific points of origination and developed recommendations on array geometry and the number of arrays or subarrays required based on machinery layout.

Earlier phases of the project demonstrated the use of spectral analysis to detect and identify motor faults from acoustic signatures, based on knowledge of vibrational frequencies associated with the faults. The frequencies which were used to identify these differences in spectral activity had physical ties to the faults, further validating that the differences between the collected acoustic emissions were due to changes in the bearing states. The earlier results demonstrated success at identifying mechanical faults and limited success at identifying electrical faults: rectifier faults at the output of a generator and load imbalance were detectable, but internal turn-to-turn shorts in the stator or rotor were not.

The also project demonstrated improvements in fault detection and localization using arrays of microphones and array beamforming to process the signals from multiple microphones. Analysis of acoustic data collected from an array of three microphones, demonstrated that different bearings, featuring matching seeded faults, could be identified from a bearing which had no seeded fault.

Artificial intelligence and machine learning techniques were developed for array processing/source detection and localization, and for the detection and classification of mechanical and electrical faults in electric motors. The techniques used acoustic power spectra as the inputs to the AI/ML models, but did not rely on a priori knowledge of the faults or the frequencies in the spectra associated with the different faults.

A tree-based learning approach was studied that directly uses the output of Fourier analysis as columns in a matrix, which the model then uses as training or testing data. This general approach was studied with a wide variety of optimizations known as hyperparameter tuning. Different segmentation and averaging parameters were used in Welch's PSD to generate FFT's with differing levels of resolution. Decibel scaling was applied to test the effect of amplifying lower energy spectral peaks. Signal-to-noise ratio calculations were applied, to scale each spectral value against its surrounding peaks. Both standard decision trees and random forests of decision trees were tested. Regularized logistic regression (LASSO regression), using the L_1 norm penalty factor, was used to test the ability to reduce the number of frequencies used as input features to the tree models.

Based on the results of this project, the recommended approach is to use a single random forest classifier with inputs based on PSD frequencies selected by a LASSO regression model with a default λ parameter of one. This architecture can be implemented on data with as few as 2560 points per segment for averaging during Welch's PSD calculations. All of these models were highly successful, reporting accuracies greater than 90%, indicating that this approach is widely applicable and powerful, and it can further be concluded that tree-based learning based upon Welch's PSD can be highly successful in classifying collected acoustic emissions.

The AI/ML based classifiers were tested using data collected from 33 identical 3-phase AC motors. The motors tested included healthy motors and motors with seeded electrical and mechanical faults. A 3-element microphone array was used to collect radiated, airborne acoustic signatures from the motors during testing. The acoustic data was used to test the ability of the AI/ML models to detect and classify the fault conditions. The AI/ML based classifier correctly identified the correct fault condition with 97% accuracy for the faults tested. The AI/ML models successfully detected and classified both mechanical and electrical faults, indicating that the acoustic spectra contain information from electrical faults that may not have been apparent to human subject matter experts looking for single-frequency fault content.

Recommendations for Future Work

This project demonstrated successful detection and classification of common mechanical and electrical fault conditions in electrical motors using airborne acoustic data recorded with an array of microphones located above the motor. While further research is recommended, these results show the potential for deploying acoustic arrays to monitor electrical machinery in power plants and other industrial settings. The approaches used to process the sound recordings are common in the signal processing literature and are widely available in commercial software packages or on-line code repositories. The techniques used to train, test and implement the AI/ML models for classifying the faults are also widely available; however, the key to successful development of AI/ML models is the processes used for training and testing the models as described in this report.

Although this project demonstrated success in detecting and classifying the faults in the tested motors, the testing was conducted under controlled, laboratory conditions. Earlier phases of the project demonstrated the use of multiple acoustic measurements to steer the sensitivity of the microphone array in different spatial directions to increase the sensitivity to sound produced by a machine of interest and reduce the sound (noise) from unwanted sources. However, the physical limitations of the test setup, discussed in Section 3, did not permit testing of this feature.

Further development and testing are recommended to extend the applicability of these results to a wider range of power plant applications and confirm the results in a prototypical scenario. This includes real-time implementation of the data acquisition and signal processing, optimization of the microphone array deployment, and additional testing in power plant environments (on live equipment) to study the effects of an uncontrolled noise environment and a wider range of fault conditions. Several suggestions for future research and development are provided below:

- Identify which fault categories (mechanical and electrical) generate emissions significant enough to be reliably detected using acoustic sensing beyond those included in this study.
- Characterize the sensitivity of acoustic measurements to early-stage faults as a function of distance between the source and the acoustic sensors
- Study how general, across electric motor or generator sizes, are the amplitudes of emissions generated by various faults
- Determine how general, across different equipment types (electrical motor vs centrifugal pump vs etc.), are the amplitudes and transmission of bearing fault frequencies by acoustic emissions
- Characterize the performance of the AI/ML fault detection and classification models under different operating conditions (load and speed)
- Identify what other factors (including facility/space details, other active equipment in the area, machine specifics) contribute to the maximum distance at which machine health can be determined based on acoustic emissions

Section 5: References

1. *Area-Wide Acoustic and Electromagnetic (AEM) Signature Health Monitoring: Phase 2*. EPRI, Palo Alto, CA: 2019. 3002015880.
2. Mark Richards, *Fundamentals of Radar Signal Processing*, McGraw Hill, 2005
3. Hayes, Monson H. *Statistical Digital Signal Processing and Modeling*. New York: John Wiley & Sons, 1996.
4. Stoica, Petre, and Randolph Moses. *Spectral Analysis of Signals*. Upper Saddle River, NJ: Prentice Hall, 2005.
5. *Electrical Signature Analysis (ESA) for On-Line Equipment Condition Monitoring: Fault Detection Testing*. EPRI, Palo Alto, CA: 2021. 3002020869.
6. Bhargava, Neeraj, et al. "Decision tree analysis on j48 algorithm for data mining." *Proceedings of International Journal of Advanced Research in Computer Science and Software Engineering* 3.6 (2013).
7. Caruana, Rich, and Alexandru Niculescu-Mizil. "An empirical comparison of supervised learning algorithms." *Proceedings of the 23rd International Conference on Machine Learning*. 2006.
8. Tibshirani, Robert (1996). "Regression Shrinkage and Selection via the lasso". *Journal of the Royal Statistical Society. Series B (methodological)*. Wiley. 58 (1): 267–88. JSTOR 2346178.
9. <https://www.skf.com/group/products/rolling-bearings/ball-bearings/deep-groove-ball-bearings#cid-493604>
10. Lawrence Kinsler, Austin Frey, Alan Coppens, James Sanders, *Fundamentals of Acoustics*, 3rd Edition, John Wiley & Sons, 1982.
11. Mathworks Matlab R2020b Sensor Array Analyzer App.
12. https://en.wikipedia.org/wiki/Gini_coefficient



Appendix A: Glossary of Terms

This glossary provides definitions for technical terms used in this report, in the acoustics and electromagnetics field, or the wider CBM field.

Definitions and Nomenclature

Acoustic Emissions: Sound waves generated by rotating equipment (motors/generators for the purposes of this report).

Electromagnetic Emissions: A magnetic field emitted by rotating equipment (motors/generators for the purposes of this report).

Spectra/Spectral Analysis: The process of studying the frequency content of one or more time series signals. Spectral analysis is commonly used to find and identify different sources within one or more signals.

Linear Model: A statistical technique which estimates the effect different potential explanatory variables have on a single response variable.

T-test: A statistical technique for comparing a value estimated from a sample to a hypothesized population parameter using students T-distribution where the variance between the two standard deviations is unknown.

Harmonic(s): A peak or set of peaks in a spectrum that are integer-number multiples of a “base” fundamental frequency, typically the natural frequency. They are often evidence of a repeated step change in the time series which occurs at the fundamental frequency

Ultrasonic: Acoustic emissions/sound waves at a frequency above the standard human hearing range. For the purposes of this report, ultrasonic emissions are considered those above 10 kHz, which was the cutoff frequency used for the standard frequency analysis.

Phase: The relative difference between two coherent waves at the same frequency. Expressed in units of degrees in this report, waves that are in phase have a difference of 0°, while waves that are perfectly out of phase have a phase difference of 180°.

Beamforming: A signal processing technique that combines the signals either generated or received by elements of an antenna array for the purposes of directional signal reception or transmission.

Units of Measurement

Frequency – Hertz (Hz): Cycles per second. Often modified with “kilo”, meaning 1000 cycles per second.

Acoustic Amplitude – Sound Pressure Level (SPL): The deviation from ambient atmospheric pressure caused by acoustic emissions. Often measured on a log scale and represented by decibels.

General Amplitude – Decibels (dB): A unit representing the logarithmically scaled ratio between two values. Often used to describe units of value above or below a standardized value

Purpose Acronyms

AC	Alternating Current
AI	Artificial Intelligence
CBM	Condition Based Maintenance
DC	Direct Current
EM	Electromagnetic
ESA	Electrical Signal Analysis
FFT	Fast Fourier Transform
MCSA	Motor Current Signature Analysis
ML	Machine Learning
NPS	Number of points per segment
PSD	Power Spectral Density
VA	Vibration Analysis



Export Control Restrictions

Access to and use of this EPRI product is granted with the specific understanding and requirement that responsibility for ensuring full compliance with all applicable U.S. and foreign export laws and regulations is being undertaken by you and your company. This includes an obligation to ensure that any individual receiving access hereunder who is not a U.S. citizen or U.S. permanent resident is permitted access under applicable U.S. and foreign export laws and regulations.

In the event you are uncertain whether you or your company may lawfully obtain access to this EPRI product, you acknowledge that it is your obligation to consult with your company's legal counsel to determine whether this access is lawful. Although EPRI may make available on a case by case basis an informal assessment of the applicable U.S. export classification for specific EPRI products, you and your company acknowledge that this assessment is solely for informational purposes and not for reliance purposes.

Your obligations regarding U.S. export control requirements apply during and after you and your company's engagement with EPRI. To be clear, the obligations continue after your retirement or other departure from your company, and include any knowledge retained after gaining access to EPRI products.

You and your company understand and acknowledge your obligations to make a prompt report to EPRI and the appropriate authorities regarding any access to or use of this EPRI product hereunder that may be in violation of applicable U.S. or foreign export laws or regulations.

About EPRI

Founded in 1972, EPRI is the world's preeminent independent, non-profit energy research and development organization, with offices around the world. EPRI's trusted experts collaborate with more than 450 companies in 45 countries, driving innovation to ensure the public has clean, safe, reliable, affordable, and equitable access to electricity across the globe. Together, we are shaping the future of energy.

Programs:

Instrumentation and Control

Nuclear Power

© 2022 Electric Power Research Institute (EPRI), Inc. All rights reserved. Electric Power Research Institute, EPRI, and TOGETHER...SHAPING THE FUTURE OF ENERGY are registered marks of the Electric Power Research Institute, Inc. in the U.S. and worldwide.

3002023744

EPRI

3420 Hillview Avenue, Palo Alto, California 94304-1338 • PO Box 10412, Palo Alto, California 94303-0813 USA
800.313.3774 • 650.855.2121 • askepri@epri.com • www.epri.com