Diagnostic Features from Aircraft Propulsion Bearings in Accelerated Aging Experiments

Brian Dykas
(former) Research Aerospace Engineer
CCDC Army Research Laboratory
APG, MD, USA

Adrian A. Hood
Research Mechanical Engineer
CCDC Army Research Laboratory
APG, MD, USA

Nenad G. Nenadic
Research Faculty
Rochester Institute of Technology
Rochester, NY, USA

Edward Zhu
Research Mechanical Engineer
CCDC Army Research Laboratory
APG, MD, USA

ABSTRACT

Over the course of a long-duration aging of helicopter drivetrain bearings to examine the consumption of grease life, both vibration and acoustic emission sensing was used to monitor the bearing response as the grease life was consumed through this aging. Acoustic emission is evaluated against vibration in terms of signal trends over the course of the experiment. Common signal metrics are calculated to yield condition indicators, and machine learning techniques are applied to the vibration and acoustic emission data. For the 862 hour duration test run equivalent to over 6700 hours on wing, features of these signals trend with increased degree of aging. Autoencoders were used to enrich existing set of traditional condition indicators and principle component analysis was effectively used for feature fusion. This measured trending shows promise for future onboard Health and Usage Monitoring Systems which may adopt new sensing and data analysis modalities to trend the condition of mechanical systems.

INTRODUCTION

Vibration monitoring has become ubiquitous on modern helicopter mechanical systems following several decades of study into the dynamics and signal processing of vibration from components containing damage including spalls resulting from rolling contact fatigue (Refs. 1–4). Despite enduring interest in prognostic methods for developing accurate estimates of remaining useful life of components, achievement of this goal has remained elusive for some failure modes. The consumption of grease life in bearings has been studied and grease life models factor in to the determination of required inspections and other airworthiness and maintenance considerations. However, no studies in the open literature show clear trends of diagnostic features in vibration signals over long periods of time associated with the gradual consumption of grease life in otherwise healthy bearings. This paper advances this capability by examining next generation diagnostic techniques including (1) acoustic emission sensing and (2) application of machine learning techniques to improve the diagnostic capabilities in Future Vertical Lift aircraft.

Conventional vibration measurements on wing have not been demonstrated to give a reliable quantitative measure of the grease condition, although corrosion and fatigue damage may be more readily measured. However, the application of deep learning techniques to data recorded over long periods holds promise to better classify the grease age than conventional mechanical diagnostics relying on a small number of engineered signal features. Furthermore, features of acoustic emission signals have demonstrated potentially higher sensitivity to lubrication condition due to increase AE generation through asperity contact, than can be measured in structural vibration signals (Refs. 5, 6).

Acoustic emission techniques have been well established for structural components and are increasingly being considered for dynamic mechanical systems (Refs. 7, 8). Signal characteristics for the latter are quite different and features have not been well developed for dynamic components operating in noisy systems. In this work, acoustic emission signals are recorded near a test bearing operating in a complex mechanical assembly over a period of several hundred hours. Signal

Presented at the Vertical Flight Societys 75th Annual Forum and Technology Display, Philadelphia, PA, USA, May 13-16, 2019.

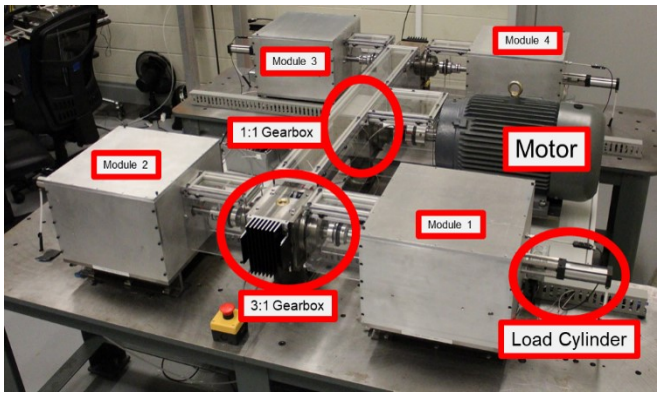

Fig. 1: Bearing Test Rig showing Four Test Stations

snapshots are recorded from both an acoustic emission sensor and a conventional accelerometer as the bearing and lubricant are subjected to an accelerated aging environment. Model order reduction techniques are applied to both the acoustic emission signal as well as the acceleration signal to assess and compare their diagnostic sensitivity and accuracy.

Autoencoders are unsupervised artificial neural network used to learn efficient representations of the input data by attempting to recreate the input. In doing so, a much lower representation is expressed at the inner layers. Autoencoder topologies, trained on data that corresponds to normal operation, have been demonstrated as effective anomaly detectors (Ref. 9). The recent revolution in representation learning furthered the usefulness of this machine learning. An early successful demonstration of autoencoder neural network as an anomaly detector using vibration data provided is given in (Ref. 10). In the meantime, there has been a revolution in training deep neural networks, which facilitated training of deeper, more expressive neural network models and demonstrated its performance on many important machine learning tasks (Ref. 11). The significant advantage of deep learning approach in HUMS is that it can be trained on abundantly available normal data (Ref. 12), whereas classical machine learning model require a more *statistically large data sets* (Ref. 13), with significant number of instances of failure which are expensive to collect.

METHODS

Experimental Test Stand and Specimens

These experiments were conducted in conjunction with a test program aimed at aging grease lubricated bearings from a US Army utility helicopter for 862 hours at its nominal operating speed of 4115 rpm, axial load of 25 lbf, and an elevated temperature of 250 $^{\circ}F$ to accelerate the consumption of grease life. A total of eight test bearings were aged in this manner, distributed across four test chambers with each pair of bearings loaded against one other to ensure equal loading. A single electric motor drove all test chambers through a series of gearboxes and shafts, as shown in Figure 1

The test bearing is a deep groove ball bearing with an inner diameter of 1.97 in. (50 mm) and an outer diameter of 3.54 in.

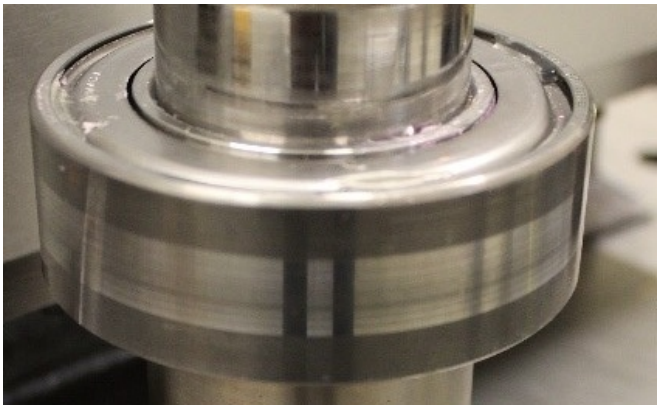

Fig. 2: Test Bearings Mounted on Shaft with Inner Race Spacer

(90 mm), ball diameter of 0.5 in (12.7 mm) and 10 balls. The DN is 288,000 which is considered a moderate speed. The test bearing is shown in Figure 2. It is grease lubricated with Chevron SRI grease, containing a polyurea thickener. The bearing has a two piece, stamped and riveted cage. For this testing, the bearings were degreased and a reduced charge of 3.5 cc grease was reintroduced in order to reduce the amount of base oil available to lubricate the bearings. This reduced fill compares with a nominal charge of approximately 30% of the free volume in the bearing or 17 cc, so the grease volume is reduced by roughly 80% for these experiments (Ref. 14). Operation for long periods at elevated temperatures results in a slow consumption of the lubricant and increased asperity contact between bearing components, which has been shown to result in increased acoustic emission activity (Ref. 15).

In order to estimate the life acceleration factor introduced by elevated temperature operation in the laboratory, two models are employed. The first adopts tabular data from the SKF catalog (Ref. 16) and the second is an empirical model given by Booser and Khonsari (Refs. 17, 18). Using an appropriate empirical factor, the model for grease life given by Booser and Khonsari is given in Equation 1

$$\log \bar{L} = -2.40 - \frac{2540}{273+T_c} - S_g \log 2 \\ -.26\frac{DN}{(DN)_L} - .18\frac{DNW}{C^2} \quad (1)$$

where:

$\bar{L}$	Life (Hours)
T_c	Temperature (^{o}C)
S_g	Half life subtraction factor = .3 (unitless)
N	Shaft Speed (RPM)
$\frac{DN}{(DN)_L}$	Speed penalty (unitless)
D	Averge Bearing Diameter (mm)
W	Radial load (lbs.)
C	Specific dynamic capacity (lbs.)

The acceleration factor is given as $AF = \frac{\bar{L}_{70}}{\bar{L}}$ where $\bar{L}_{70}$ is the life at the operating temperature of 70°. The bearings are exposed to a range of temperatures on wing under operation, but

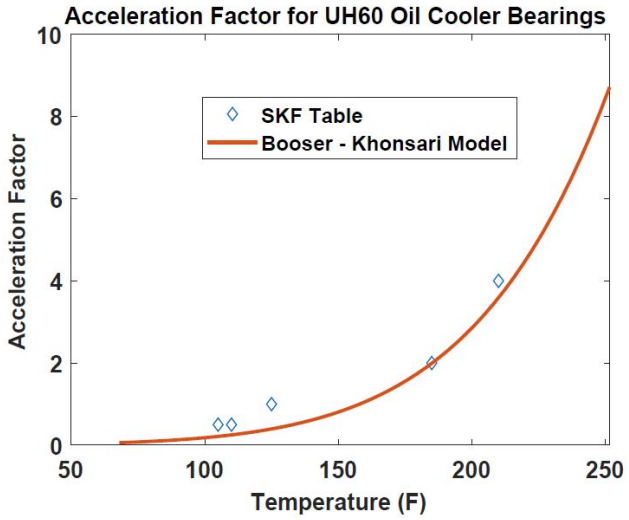

Fig. 3: Dependence of Grease Life Acceleration Factor on Temperature

speeds and loads are reasonably constant. The life reduction factor is shown graphically versus temperature in Figure 3. These models suggest that the acceleration factor for the 250 $^{\circ}F$ outer race temperature is on the order of 8, and the 842 hours of operation in these experiments would translate into approximately 6,700 hours of operation on wing. The reduction in grease charge compounds this acceleration factor by reducing the available reservoir of mobile oil.

Data snapshots were sampled from an accelerometer, acoustic emission sensor and a tachometer sensor at two-minute intervals over the entire experiment at 1 MHz. Figure 4 shows the date plotted against run-time. The breaks in the line represent times for which the stand was operating but high speed data was not collected. The test was divided into four periods: (1) break-in, 0-40 hours; (2) Baseline, 40-307 hours; (3) Aged 1, 307-591 hours; (4) Aged 2, 591-862 hours. During break-in the bearing experiences initial wear and grease redistribution. The duration of this process is inexact but is given as 8-10 hours by a major bearing manufacturer (Ref. 16) and 50 hours in the Army Aeronautical Design Standard 79E (Ref. 19). For the purposes of data analysis here, an intermediate value of 40 hours is chosen, with the remaining periods being approximately 275 hours each. The Baseline period represents an early period in the bearing/grease life, while Aged 1 and Aged 2 represent progressively more aging of the bearings although these conditions are relative to the initial condition. The bearing did not fail prior to the pre-determined test duration so the final data points do not necessarily represent the end of grease life.

Only a single bearing was instrumented for this experiment due to limited instrumentation and data acquisition equipment on hand although future test campaigns can increase this number of instrumented bearings to establish repeatability of the diagnostic results across different test specimens. The bearings were labeled such that the bearing closest to the load cylinder was designated as 1 and the bearing closest to the input shaft was designated 2. The instrumented bearing B_{22} was

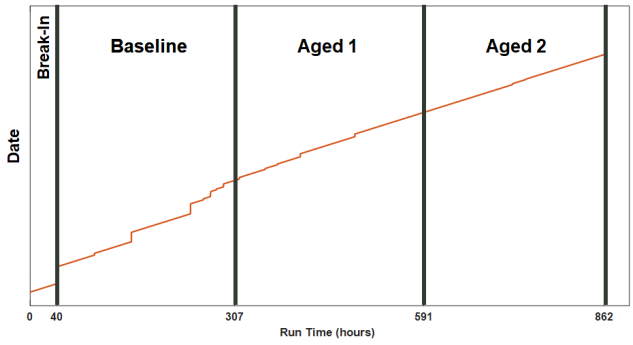

Fig. 4: Data Summary showing Test Phases over Time

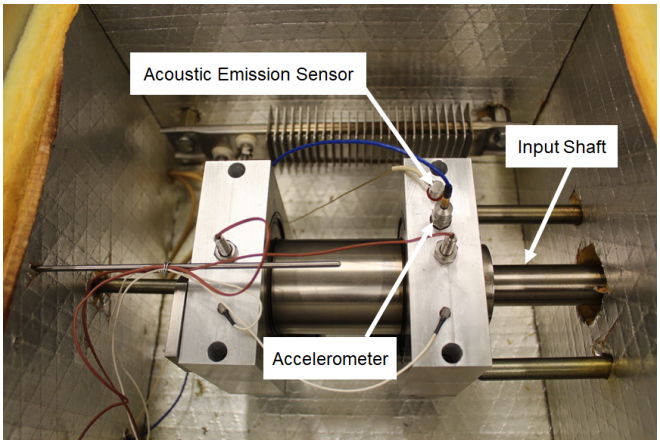

Fig. 5: Bearing test chamber with instrumentation on the right-side bearing

in test chamber #2 on the input shaft side. The instrumented bearing is shown installed in the test chamber in Figure 5.

The two sensors used are a PCB 3224A1 accelerometer and Mistras Nano30 AE Sensor. Both sensors have temperature ranges up to 350 $^{\circ}F$. The accelerometer has a frequency range of 20 kHz frequency and the frequency range of the AE sensor is 125-750 kHz. The acoustic emission sensor was adhered to the bearing block using a silicone adhesive and the stud mounted accelerometer was mounted to a nut attached to the bearing block using a high temperature two part epoxy.

SIGNAL PROCESSING AND CONDITION INDICATORS

Common signal processing techniques were first applied to acoustic emission and vibration data to generate Condition Indicators (CIs) including statistical characteristics, spectral methods and characteristics of the envelope spectrum including bearing fault frequencies. A total of 38 condition indicators were calculated for acoustic emission and 35 condition indicators for the vibration signal. Most were similarly defined with some targeting different bands of the frequency spectrum. A full list of the condition indicators considered is included in Table 4 in the Appendix.

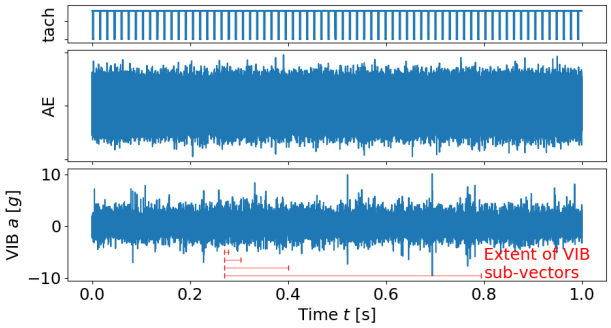

Fig. 6: Contents of a typical data file: tachometer (top), acoustic emission, vibration (bottom)

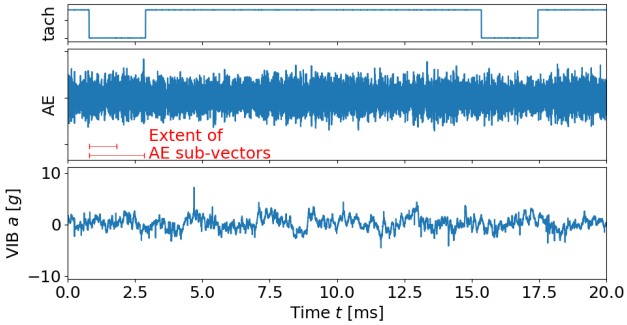

Fig. 7: AE and vibration signals over one revolution of the shaft.

Pre-processing

Figure 6 shows the input waveforms of a sample file: tachometer, acoustic emission, and vibration, with one million samples over the course of one second of operation. Figure 7 zooms into time and illustrates data variation on the order of one revolution of the shaft.

The frequency-domain information of acoustic emission and vibration is shown in the same normalized graph of Figure 8. The two normalized spectra, shown in the same plot, had very little overlap. The acoustic emission signal shows much higher frequency content associated with the range of the sensor and amplifier filter, in the range of 100-450 kHz. This measurement of Rayleigh stress waves at the surface shows continuous mode emissions resulting from the large amount of generation in dynamic mechanical systems.

Over the course of the experiment, a time varying DC component of AE was noted. This non-physical artifact in the data acquisition implementation drifted over time occasionally dropping back to lower values as shown in Figure 9. Because the variation in DC component was an artifact of instrumentation/data-acquisition channel and not a real, physical characteristic of the signal, and it was removed from the data files prior to analysis. As a consequence, the calculated root mean square of the signal is mathematically identical to the variance.

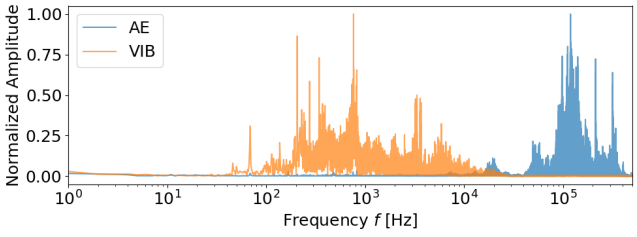

Fig. 8: Acoustic emission and vibration spectra

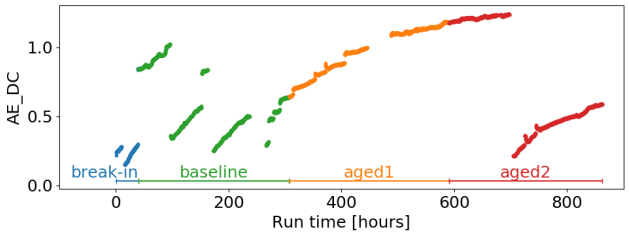

Fig. 9: DC Component of AE Signal, removed for analysis

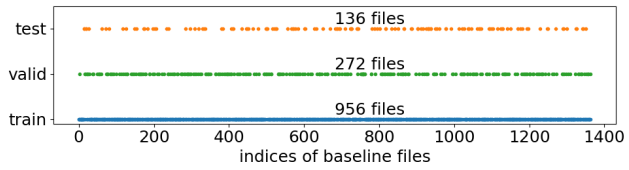

Fig. 10: Three data sets are uniformly distributed across the baseline

Three pre-processing steps were performed per 1 second data record 1) Because the means, or DC component of acceleration and acoustic emission do not carry relevant information on health but can be affected by the instrumentation drift, the means were subtracted from the signal. 2) the samples before the first falling edge and after the last falling edge of the tachometer signal were removed 3) Data associated with files where the RPM $<$ 2000 RPM was excluded from training the models. 4) All data was normalized using min-max. Machine learning algorithms do not perform well when the input numerical attributes have very different scales. This scaling is used to compress the range of the data from 0 to 1. (Ref. 20)

The models were trained, validated, and tested on the baseline data segment and later evaluated on the remaining segments. Before the modeling started, 15% of the baseline data was randomly selected from the baseline data set and set aside for testing. The remaining 85% was used for building models (training and validation). Of the 85% of the data used for modeling, 80% was used for training and 20 for validation, with all three dataset being uniformly distributed across the baseline dataset, as illustrated in Figure 10.

MACHINE LEARNING MODELS

The autoencoder implementation was based on *fully-connected, dense* layers. In this type of layers each of the inputs is connected to each of the output through the matrix of weights as conceptually depicted in Figure 11. Because the input neurons are full-connected to the output neurons, the output vector of the i^{th} layer $\bar{\phi}^{(i)}$ is obtained as a nonlinear

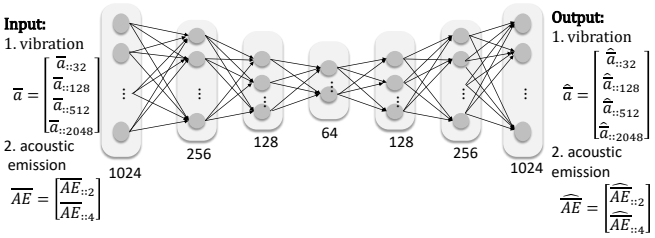

Fig. 11: Autoencoder Topology Applied to Test Data

transformation of an affine transformation of the output vector of the preceding layer $\bar{\phi}^{(i-1)}$:

$$\bar{\phi}^{(i)} = \mathrm{f}(\bar{\bar{W}}^{(i-1)}\bar{\phi}^{(i-1)} + \bar{b}^{(i-1)}) \tag{2}$$

where $\bar{\bar{W}}^{(i-1)}$ and $\bar{b}^{(i-1)}$, and f are the associated matrix of weights, offset vector, and the nonlinear transformation, respectively. Traditionally, the activation function f() of choice in neural network was a sigmoid, defined as

$$\mathrm{f}(z) = \sigma(z) = \frac{1}{1+\mathrm{e}^{-z}}. \tag{3}$$

Modern implementation prefer simpler activation functions, such as the rectifier function $\mathrm{f}(z) = \mathrm{ReLU}(z) = max(0,z)$. Our implementation started with $\mathrm{ReLU}(z)$ activation function, but obtained better performance with $\mathrm{ELU}(z)$ defined as

$$\mathrm{f}(z) = \mathrm{ELU}(z) = \begin{cases} z, & z > 0 \\ e^z - 1, & z < 0 \end{cases} \tag{4}$$

The number of layers and the number of neurons per layer are hyperparameters that are varied. For example, the topology in Figure 11 depicts the starting topology for acceleration data which first encodes a 1024-component input vector to a 64-component encoding vector via two layers with 256- and 128-outputs respectively and then decodes it back to 1024 components via 128- and 256-output output layers.

Sampling Concept

The number of parameters needed to train an autoencoder is dependent on the number of inputs, hidden layers, and neurons per layer. An input size of 1024 features was chosen due to the author's experience in training neural networks of this size. The challenge then is to determine ways to compress the 1 million data points per record into the pre-determined input size. Three methods were investigated: 1) A collection of decimated samples of the data and 2) Representations of the spectrogram images 3) Calculated condition indicators.

Input Type 1: Signal Decimation To efficiently capture multiple frequency ranges of the input data, the input vector of length n consists of multi-scale sub-vectors of length L obtained by decimating the signals sampled at 1 MHz.

In the proposed scheme the input vector $\bar{s}$ consisted of sub-vectors at different time scales $\bar{s_{:k}}$. For example

$$\bar{s} = \left[\bar{s}^T_{:2^0}\bar{s}^T_{:2^1}\cdots\bar{s}^T_{:2^M}\right]^T \tag{5}$$

where $\bar{s}_{:1}$ denotes a sub-vector of consecutive samples of the input data, $\bar{s}_{:2}$ denotes a sub-vector that takes every other point of the input data, and $\bar{s}_{:2^M}$ is a sub-vector decimated by 2^M where $M = n/L$

For acoustic emission, where frequency of interest is in the $100 \leq f \leq 300$ kHz range (see Figure 8), the initial configuration employs $L = 512$ element long sub-vectors, starting by sampling every other sample and ending by sampling every 4th sample.

For vibration data, where most energy was below 30 kHz (see Figure 8), four decimation levels were used: 32, 128, 512, 2048, as shown in Figure 11.

Each sub-vector probed into a specific time scale starts at the same random location within the data record. Up to 1000 input vectors of this construction were randomly drawn from a single file. The files contained one million points per signal and the length of the vector per signal was 1024.

Referring again to Figures 6 and 7, it also compares the data extraction lengths and the tach signal to get a sense of the fraction of the rotation that is being captured. For vibration, all except the 1/212th sample captures at least one shaft cycle of dynamics. The AE sensor's samples corresponds to rotation on the order of the tachometer target painted on the shaft. During training, 1000 random samples are taken from each dataset resulting in 3,879,400 training examples so it is felt that dynamics associated with each shaft angle is accounted for in the training. Evaluation employed the same process, except fewer samples - typically about 40 - were drawn per file and a statistical function was applied to the distribution of predictions vectors.

Input Type 2: Spectrograms A 2D Convolutional Neural Network (CNN) autoencoder was used to assess the performance of using spectrogram images as input. These images serves as a form of data compression as 22,499 points represented by pixel values are used instead of the 1 million sample points. CNNs emerged from the study of the brains visual cortex in which many neurons react only to a visual stimuli located in a limited region of the visual field. This motivated an architecture in which the neurons are not fully connected. These networks determine optimal filter parameters to extract features from the image and then feeds those features into an an autoencoder. For each convolution layer, a filter kernel is convolved with the image to produce a reduced order representation.

The AE signal spectrograms were generated by applying the STFT, with a Hann window of length 300 and overlap of 100 points, on randomly sampled 30k point sequences from the 22,871 data records. Ten of such sequences were sampled from each record, resulting in a total of 228,710 images of size 149 x 151. The vibration signals were processed in a largely similar way. However, due to the signal energy residing in a lower frequency range, each data record was decimated by

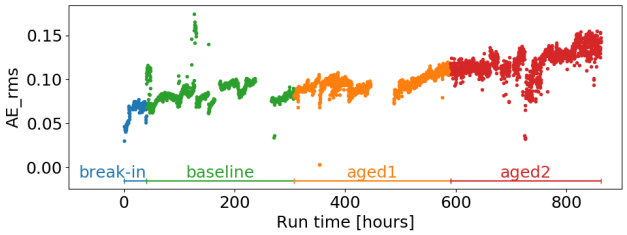

Fig. 12: Trend of AE Root Mean Square throughout Test

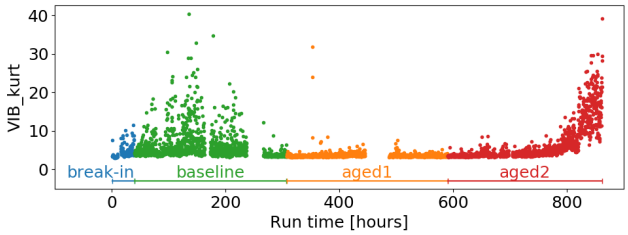

Fig. 13: Trend of Acceleration Kurtosis throughout Test

3, keeping the first 30k points. The spectrograms were then generated using the downsampled signals to create a total of 22,871 images of size 149 x 151.

Input Type 3: CIs Analyzed Using an Autoencoder

As mentioned earlier, autoencoder is a data reduction strategy that reduces redundancies. For this reason, the calculated CIs were also processed by an autoencoder and the codings analyzed for comparison.

RESULTS

Vibration and acoustic emission data were recorded for the duration of the 862 hour experiment, with only occasional interruptions in the acquired data due to computer shutdowns or operator interruption for routine intervention and data handling. A total of 22,871 vibration and acoustic emission snapshots were sampled at 1 MHz. A subset of 4,564 data records, representing every 10 minutes, was used for the current analysis of input types 1 and 3. Input type 2 used the full data-set.

Representative Condition Indicators

One of the easiest and most common current approaches to reporting acoustic emission for mechanical systems with continuous mode emissions is the root means square of the signal. Examining the rms of each snapshot of acoustic emission, Figure 12 shows a clear upward trend throughout the experiment with some early scatter during a period of break-in of the grease within the bearing. This represents an improvement over vibration which often trends up and down, and is influenced heavily by other aspects of the dynamic system (Figure 13). The gaps are due to times when the rig was operating but the high speed data acquisition system was not recording. Vibration RMS did not show a clear trend. However, vibration kurtosis suggested potentially different degradation mechanism compared to the one suggested by AE RMS. While AE RMS trend is monotonic, almost linear, vibration kurtosis suggest onset of degredation at the end of Aged 2. In particular, while some values of vibration kurtosis were large even during the baseline portion of the test the trend of the second part of Aged 2 are different.

There are 38 CIs computed for the AE sensor and 35 computed for the accelerometer. A Principle Component Analysis (PCA) is conducted on the CIs to avoid comparing every CI to

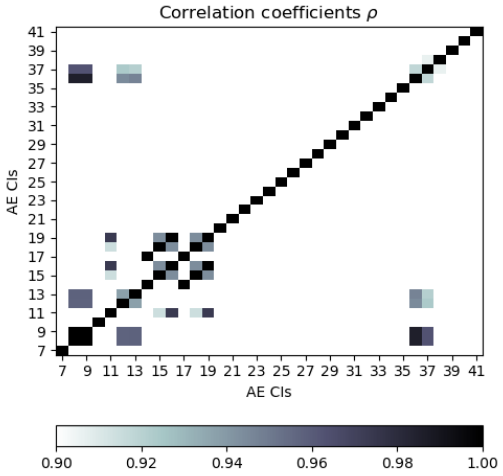

Fig. 14: AE: CI Cross-Correlation Coefficients > .9

those determined by the autoencoder. Since many of the signals have a high degree of correlation, the number of CIs are down-selected prior to the PCA. This is done by producing a cross-correlation matrix and isolating the CI combinations with cross correlations greater that .9 Plots of the cross correlation coefficients are give in Figures 14 and 15. The index values along the x and y axis correspond to Table 4 in the Appendix.

For AE, the following CIs were used: 8, 10, 11:12, 14, 20:35,

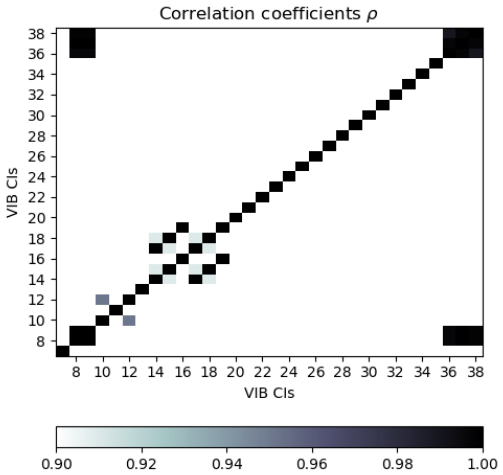

Fig. 15: Vibration: CI Cross-Correlation Coefficients > .9

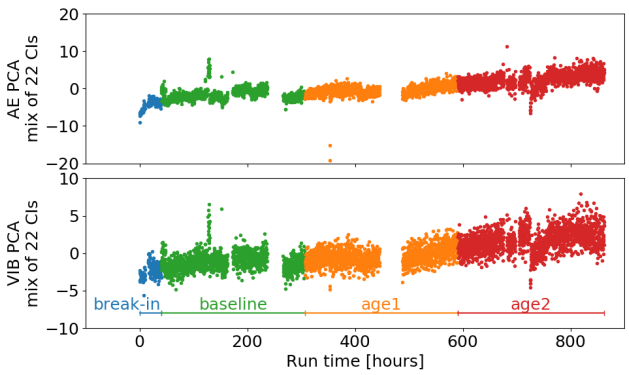

Fig. 16: CIs Projected onto PC1 vs Runtime

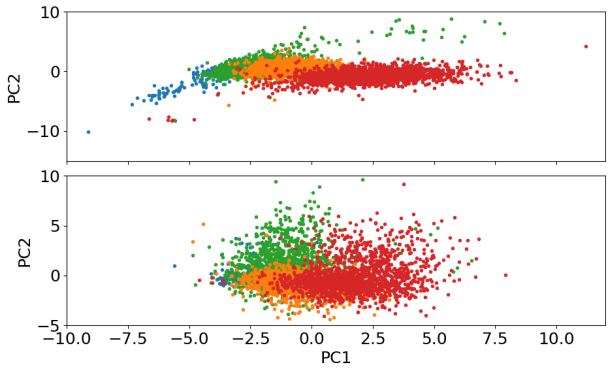

Fig. 17: CIs Projected onto PC1 and PC2 AE:Top Vib:Bottom

39, 40. For acceleration, the following CIs were used: 8, 10, 11, 13:14, 16, 20:35.

For each data record, the corresponding subset of CIs are weighted and summed into a single value. The weights are obtained from the values in the 1st eigenvector obtained from the PCA. The results are shown in Figure 16 where each values is plotted vs run-time.

The two plots show similar information. Both trend upward with apparent activity at around 175 hours and 750 hours.

Another visualization is to project the select CI values onto the first 2 principle axes. This is shown in Figure 17. Distinct clusters associated with age are apparent for the acoustic emission signal but there is considerable overlap. The vibration signal does not have as much spread and has a higher degree of overlap than the AE signal.

Image Type 1: Decimated Signals

Two separate encoders were trained: one based on the acoustic emission signal and the other based on the vibration signal. The results are discussed in turn.

A TensorFlow/Keras (Refs. 21,22) implementation is depicted in the model summary given in Table 1. It shows that seven layers (including the input) were used, progressing from 1024 to 64 and back to 1024 outputs requiring 608,064 parameters trained. Consistent dropout rate of 10% was applied to each layer. Dropout is method for preventing overfitting, based on

Table 1: The summary of autoencoder model by layers used for AE and vibration.

Layer	Number of neurons	Number of params	Activation
fully-connected 0 (input)	1024	0	None
fully-connected 1	256	262400	ELU
fully-connected 2	128	32896	ELU
fully-connected 3	64	8256	ELU
fully-connected 4	128	8320	ELU
fully-connected 5	256	33024	ELU
fully-connected 6 (output)	1024	263168	Linear
Total params		608,064	

Dropout: 10% for all layers

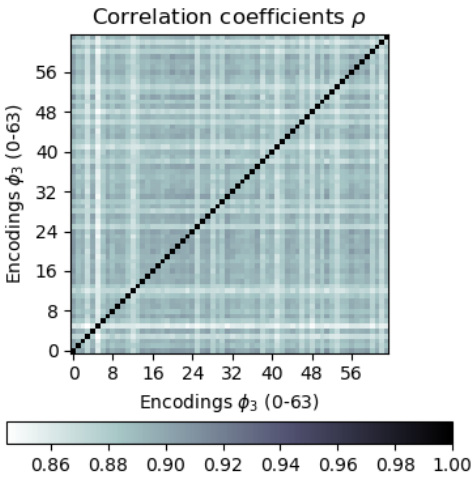

Fig. 18: Cross-correlation matrix of AE encodings

excluding a subset of nodes randomly during the training process (Ref. 23). 100 epochs were used to get a loss of 0.02. The model for AE was similar but, due to the limited number of decimations available, started with 512 and compressed to 8 and back to 512. There were 140,616 parameters trained to a loss of 0.02.

The AE encodings were strongly correlated as shown in Figure 18, which plots the image of the correlation coefficient matrix computed on all sixty four encodings. The black diagonal corresponds the self-correlation coefficient equal to unity ($\rho = 1$). The color bar on the bottom of the plot shows that most coefficients are about 85%.

A few representative encodings, viz. 0, 6, 26, 34, and 64, are shown in Figure 19. The overall trend resembles that of AE RMS, shown in Figure 12.

The same training process was applied to vibration signals. While the AE input vector consisted of two sub-vectors of lengths $L = 512$, the vibration input vector consisted of four sub-vectors of lengths $L = 256$. Although the experimentation varied the widths of the layers independently for vibration and acoustic emission, the configuration of Table 1 was found the yield the smallest error for both AE and vibration. It is important to note that this was preliminary modeling effort and that more accurate topologies of neural networks will likely be found in the future work. The cross-correlations of the encodings of the vibration-based autoencoder are shown

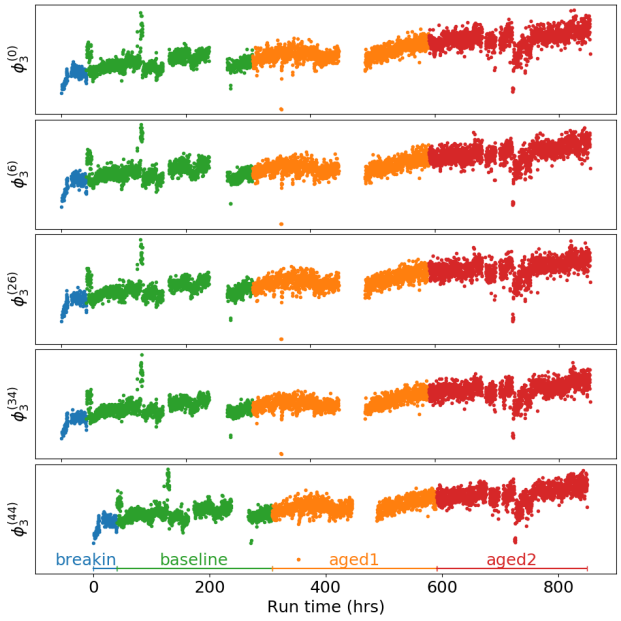

Fig. 19: Encodings: Acoustic Emission

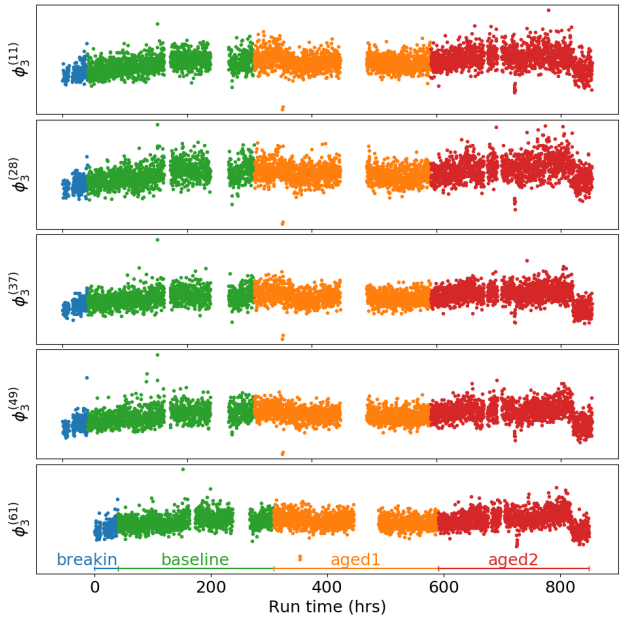

Fig. 21: Encodings: vibration

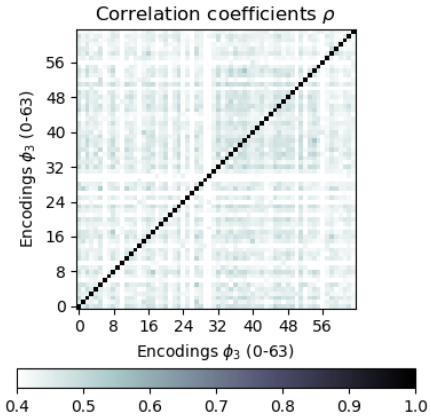

Fig. 20: Cross-correlation matrix of VIB encodings

in Figure 20.

Figure 21 shows a few representative encodings of the vibration-based autoencoder. These encodings did not indicate a notable trend.

Input Type 2: Spectrograms

Three examples of generated spectrograms (top row) as well as the corresponding autoencoder reconstructions (bottom row) are shown in Figures 22 and 23. In these figures, the x and y axes are time and log-frequency respectively with low frequencies at the top of the image.

The CNN stage produced a feature vector of size 5,776. More details about the CNN can be found in (Ref. 24). This feature vector is followed by a fully connected layer of size 512 and finally the encoding layer of 128 points.

Figures 24 and 25 show the results from the 2D CNN autoencoder analysis of the spectrograms. The codings were not analyzed vs time as was the case in Figure 19 and 21 but instead, a PCA was performed and the values were compressed to a 2D plot along the first two principle coordinates. The colors indicate the age of the bearings as indicated from each file's time stamp.

The AE result in Figure 24 also forms clusters and there is a definite trend with grease age. In addition, there is less overlap than that obtained using PCA on the calculated CIs. The acceleration result in Figure 25 also have trending clusters with age but appears to also rotate within this PC1, PC2 plane.

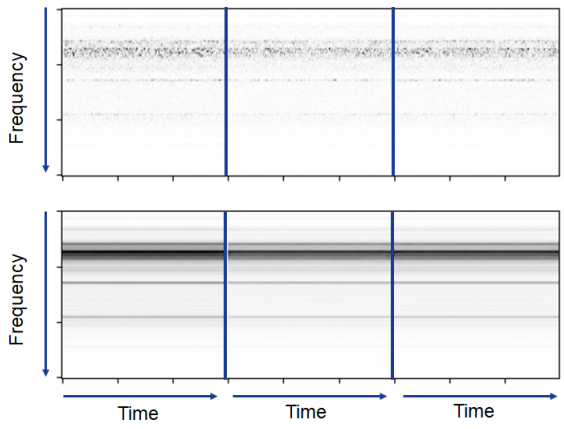

Fig. 22: Example spectrogram of AE signal

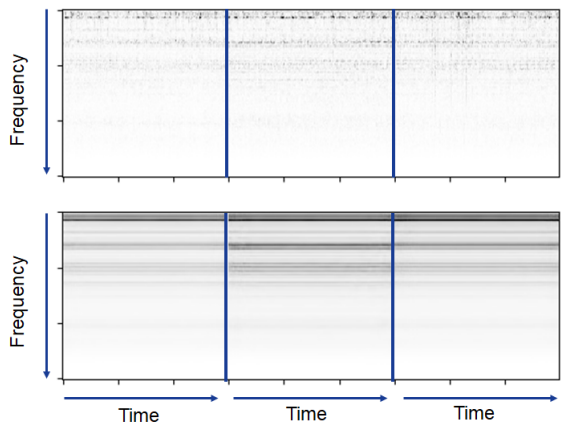

Fig. 23: Example spectrogram of acceleration signal

Input Type 3: Codings of the Condition Indicators

The computed CIs were also processed through the autoencoder to compress the information and compare with the other results. The configuration of the autoencoder is that of Figure 11, with the weights summarized in Table 2.

The set of CIs for each data record was processed through the trained autoencoder and the encodings from center layer are all plotted vs run-time.

Figures 26 and 27 show the results using all of the CIs.

These signals combined represent a reduced order representation of the data. The scatter plot of Figure 27 shows low correlations coefficients for each encoding indicating independence. Encodings 0 and 5 show a slight trend with age

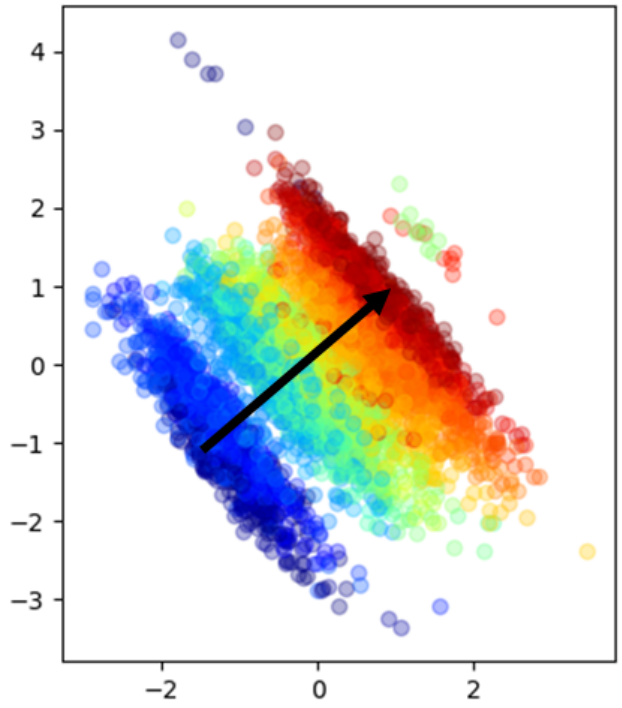

Fig. 24: PCA Plot: 2D CNN on Acoustic Emission

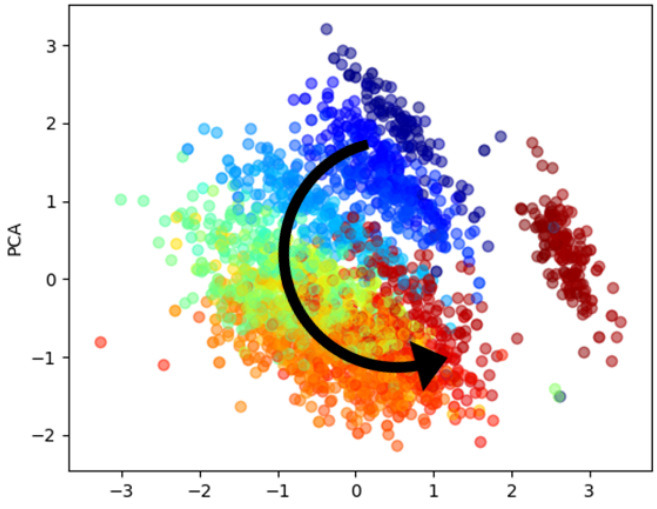

Fig. 25: PCA Plot: 2D CNN on Vibration

Table 2: The summary of autoencoder model by layers used to encode AE-based CIs.

Layer	Number of neurons	Number of params	Activation
fully-connected 0 (input)	37	0	None
fully-connected 1	32	1216	ELU
fully-connected 2	16	528	ELU
fully-connected 3	8	136	ELU
fully-connected 4	16	144	ELU
fully-connected 5	32	544	ELU
fully-connected 6 (output)	37	1221	Linear
Total params		3,789	

Dropout: 10% for all layers

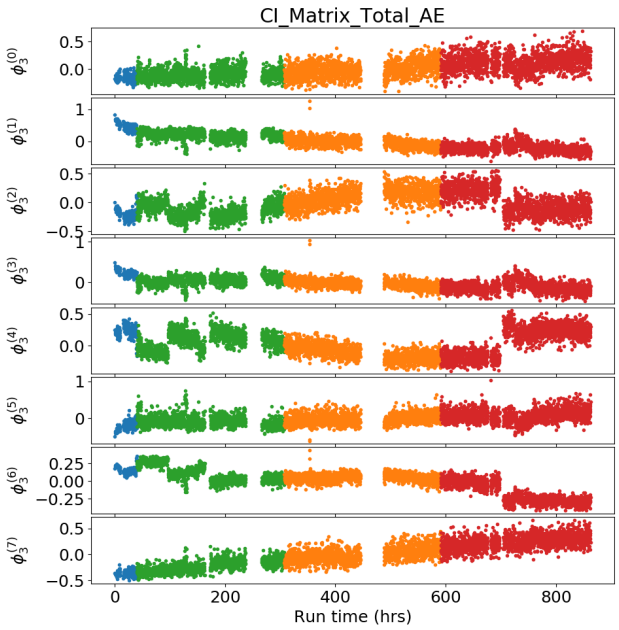

Fig. 26: Autoencoder encodings based upon AE-based CIs

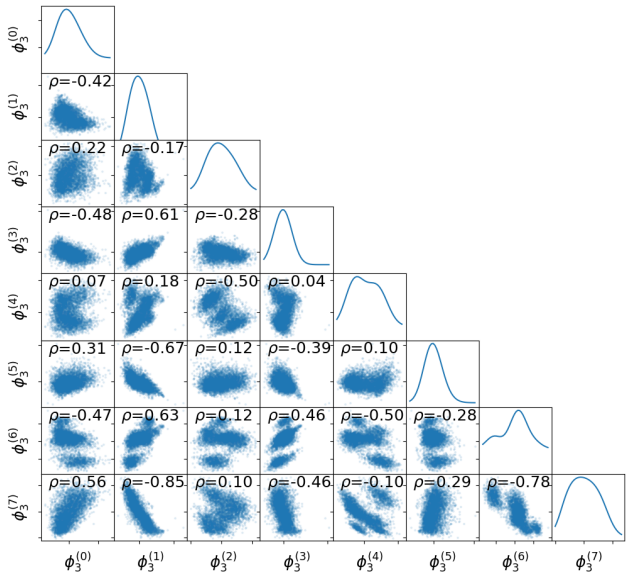

Fig. 27: Scatter Plot: autoencoder encodings based upon AE-based CIs

Table 3: The summary of autoencoder model by layers used to encode vibration-based CIs.

Layer	Number of neurons	Number of params	Activation
fully-connected 0 (input)	33	0	None
fully-connected 1	32	1088	ELU
fully-connected 2	16	528	ELU
fully-connected 3	8	136	ELU
fully-connected 4	16	144	ELU
fully-connected 5	32	544	ELU
fully-connected 6 (output)	33	1089	Linear
Total params		3,529	

Dropout: 10% for all layers

after 500 hours. Encodings 1 and 7 show monotonic variations with age. Encodings 2, 4 and 6 show breaks around 750 hours.

Similar analysis was conducted based upon vibration CIs, with Table 3 summarizing the information on the autoencoder network and Figure 28 and Figure 29 showing the encodings and their scatter plots.

Like the case for the AE signal, the encodings are distinct and have low correlation coefficients. Many of the encodings seem to indicate activity around 750 hours. Encoding 2,3, and 5 have sharp increases in amplitude and coding 1,4, and 5 seem to begin an incline at that rum-time. Encoding 7 appears to be monotonic with age.

CONCLUSIONS

This experiment investigated novel approaches to diagnosing the condition of grease lubricated bearings over hundreds of hours of test time has established feasibility of improved diagnostics for the lubrication condition that can be implemented in future aircraft. It leveraged previous feasibility work performed at the authors laboratories (Ref. 5) and the volume of data sampled gives additional insight to the variability in these signals.

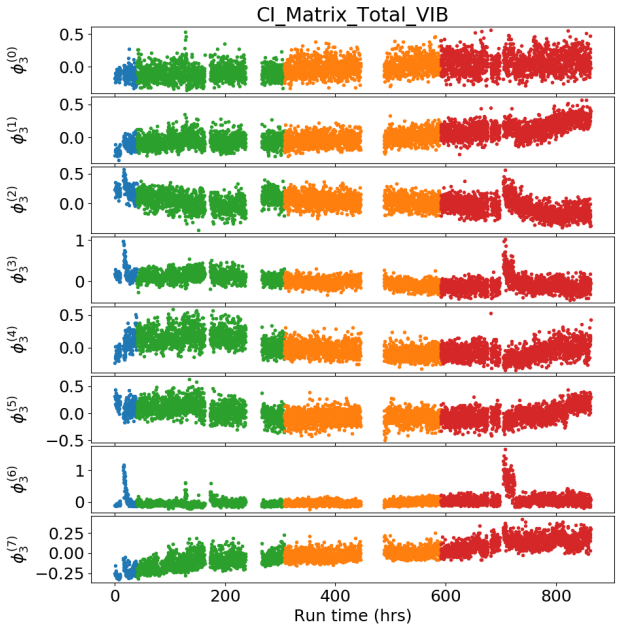

Fig. 28: Autoencoder codings based upon vibration-based CIs

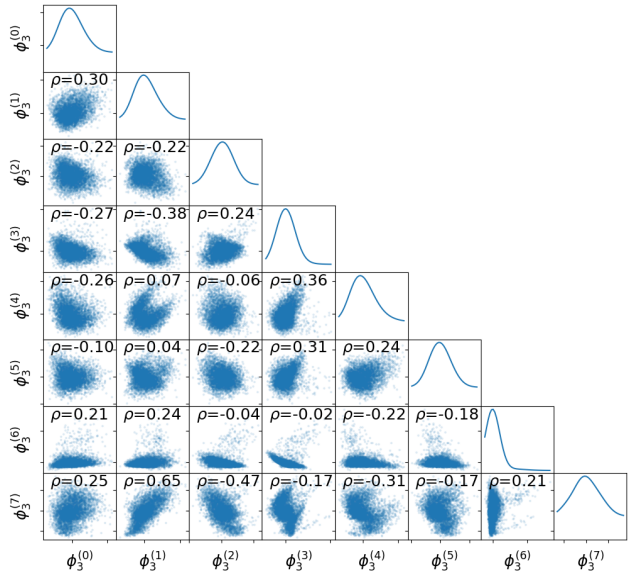

Fig. 29: Scatter Plot: autoencoder encodings based upon vibration-based CIs

The autoencoders based on the raw data showed a developing trend, but no new insights, beyond those already available from CIs. On the other hand, the CI-based autoencoders suggested more than one type of plausible degradation. In addition, the 2D CNNs were examined and they demonstrated the ability to clearly separate the 4 subsets of data into in the PC1-PC2 space. For those based on AE, the progression was linear. For vibration, the separation appeared to rotate.

The work has found that: 1) extended duration experiments are required to produce evidence of aging. 2) acoustic emission signals in mechanical components contain diagnostic features that trend over time as grease life (lubricant) is con-

sumed. 3) autoencoders enrich the existing set of engineered CIs and provide improvements over individual CIs and those fused using PCA. 4) the principle component space provides a powerful mechanism for viewing and organizing features from unsupervised learning.

Future work must next reproduce these results across multiple specimens to establish reproducibility, extend the duration of the tests to more severe grease conditions to establish feature thresholds and end of life trends, and examine computationally efficient methods of embedding the diagnostics onboard aircraft. In addition, the reported on work on autoencoders only scratched the surface of the possibilities for building useful models. For example, similar models based upon 1D CNN layers (instead of fully-connected layers) will be explored in the future and compared with these models. In addition, as more data becomes available, with more discrete instances of ground truth, the performances of different models will be reevaluated.

Author contact: Brian Dykas brian.dykas@gmail.com, Adrian Hood adrian.a.hood.civ@mail.mil, Nenad Nenadic nxnasp@rit.edu

ACKNOWLEDGMENTS

The authors wish to acknowledge the support of the NSRDEC CCDC Aviation & Missile Center and AMRDEC CCDC Data & Analysis Center for allowing these opportunistic measurements to be made during bearing aging experiments supporting the Army helicopter fleet.

APPENDIX

Table 4: Description of Calculated Condition Indicators

Index #	CI Name	CI Description
7	DC	DC component
8	Std	Standard Deviation
9	rms	Root Mean Square
10	CF	Crest Factor
11	kurt	Kurtosis
12 , 13	min,max	Minimumm and max.value
14,15,16	[3,4,5]xRMS	Count of signal threshold crossings @ [3x,4x,5x] rms
17,18,19	[3,4,5]xRMSrate	Count rate of signal threshold crossings @ [3x,4x,5x] rms
20	EnvBSFmax	Peak of envelope spectrum at BSF
21	EnvBPFOmax	Peak of envelope spectrum at BPFO
22	EnvBPFImax	Peak of envelope spectrum at BPFI
23	EnvFTFmax	Peak of envelope spectrum at FTF
24	EnvBSFmax3	Peak of envelope spectrum at BSF (max.of 3 nearest frequencies)
25	EnvBPFOmax3	Peak of envelope spectrum at BPFO (max.of 3 nearest frequencies)
26	EnvBPFImax3	Peak of envelope spectrum at BPFI (max.of 3 nearest frequencies)
27	EnvFTFmax3	Peak of envelope spectrum at FTF (max.of 3 nearest frequencies)
28	FFT_BSFmax	Peak of spectrum at BSF
29	FFT_BPFOmax	Peak of spectrum at BPFO
30	FFT_BPFImax	Peak of spectrum at BPFI
31	FFT_FTFmax	Peak of spectrum at FTF
32	FFT_BSFmax3	Peak of spectrum at BSF (max. of 3 nearest frequencies)
33	FFT_BPFOmax3	Peak of envelope at BPFO (max.of 3 nearest frequencies)
34	FFT_BPFImax3	Peak of spectrum at BPFI (max.of 3 nearest frequencies)
35	FFT_FTFmax3	Peak of envelope at FTF (max.of 3 nearest frequencies)
36	RMS_100_150	RMS of signal bandpassed from 100-150kHz (V + 20 dB)
37	RMS_150_200	(AE) RMS - bandpassed from 150-200kHz (V + 20 dB)
38	RMS_200_250	(AE) RMS - bandpassed from 200-250kHz (V + 20 dB)
39	Data_250_300	(AE) RMS - bandpassed from 250-300kHz (V + 20 dB)
40	Data_300_350	(AE) RMS - bandpassed from 300-350kHz (V + 20 dB)
41	Data_350_450	(AE) RMS - bandpassed from 350-500kHz (V + 20 dB)
35	Data_11_15	(VIB) RMS - bandpassed from 11-15kHz
37	Data_15_20	(VIB) RMS - bandpassed from 15-20kHz
38	Data_17_23	(VIB) RMS - bandpassed from 17-23kHz

The bearing fault frequencies and their harmonics are given in Table 5

Table 5: Bearing Fault Frequencies (Hz.)

Harmonic	CFF	BSF	BPFO	BPFI
1	28.1	182.7	280.7	405.1
2	56.1	365.5	561.4	810.2
3	84.2	548.3	842.1	1215.4
4	112.3	731.2	112.8	1620.5

Cage Fault Freq. (CFF), Ball Spin Freq. (BSF), Ball Pass Freq.-Outer (BPFO), Ball Pass Freq. - Inner (BPFI)

REFERENCES

[1]Balderston, H. L., "The detection of incipient failure in bearings," *Materials Evaluation*, Vol. 27(6), 1969, pp. 121–128.

[2]Halme, J. and Andersosson, P., "Rolling contact fatigue and wear fundamentals for rolling bearing diagnostics-state of the art," *Proceedings of the Institution of Mechanical Engineers, Part J: Journal of Engineering Tribology*, Vol. 224(4), 2010, pp. 377–393.

[3]Howard, I., "A Review of Rolling Element Bearing Vibration Detection, Diagnosis and Prognosis," *Defence Science and Technology Organisation*, 1994.

[4]Randall, R. and Antoni, J., "Rolling element bearing diagnosticsA tutorial," *Mechanical Systems and Signal Processing*, Vol. 25(2), 2011, pp. 485–520.

[5]Dykas, B., Hood, A., and Becker, A., "Acoustic Emission in Grease-Lubricated Helicopter Drivetrain Bearings," *Proceedings of the 10th Australian DST Group International Conference on Health and Usage Monitoring Systems (HUMS 2017), Melbourne, VIC Australia*, Feb 26-Mar.

[6]Miettinen, Anderson, P., and Wikstr, V., "Analysis of grease lubrication of a ball bearing using acoustic emission measurement," *Proceedings of the Institution of Mechanical Engineers, Part J: Journal of Engineering Tribology*, Vol. 215, 2001, pp. 535–544.

[7]Mba, D., "Development of Acoustic Emission Technology for Condition Monitoring and Diagnosis of Rotating Machines: Bearings, Pumps, Gearboxes, Engines, and Rotating Structures," *The Shock and Vibration digest*, Vol. 38, 03 2006, pp. 3–16.
doi: 10.1177/0583102405059054

[8]Sikorska, J. Z. and Mba, D., "Challenges and obstacles in the application of acoustic emission to process machinery," *Proceedings of the Institution of Mechanical Engineers, Part E: Journal of Process Mechanical Engineering*, Vol. 222, (1), 2008, pp. 1–19.
doi: 10.1243/09544089JPME111

[9]Eklund, N., "Anomaly Detection Tutorial," Proceedings of the annual conference of the prognostics and health management society, 2018.

[10]Japkowicz, N., Myers, C., Gluck, M., *et al.*, "A novelty detection approach to classification," IJCAI, Vol. 1, 1995.

[11]LeCun, Y., Bengio, Y., and Hinton, G., "Deep learning," *Nature*, Vol. 521, (7553), 2015, pp. 436–444.

[12]Yan, W. and Yu, L., "On accurate and reliable anomaly detection for gas turbine combustors: A deep learning approach," Proceedings of the annual conference of the prognostics and health management society, 2015.

[13]Bishop, C. M., "Model-based machine learning," *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, Vol. 371, (1984), 2013, pp. 20120222.

[14]Lugt, P., "A Review on Grease Lubrication in Rolling Bearings," *Tribology Transactions - TRIBOL TRANS*, Vol. 52, 06 2009, pp. 470–480.
doi: 10.1080/10402000802687940

[15]Cockerill, A., Clarke, P., A., R., T., Bradshaw, Cole, P., and Holford, K. M., "Determination of Rolling Element Bearing Condition via Acoustic Emission." *Proceedings of the Institution of Mechanical Engineers, Part J: Journal of Engineering Tribology*, Vol. 230(11), 2016, pp. 1377–1388.

[16]Group, S., *SKF Group. Super-precision bearings*, SKF Group - pub BU/P1 13383/1 EN, 2014.

[17]Booser, E., "Grease Life Forecast for Ball Bearings," *Lubrication Engineering*, Vol. 30(11), pp. 536–541.

[18]Booser, E. and Khonsari, M., "Grease Life in Ball Bearings: The Effect of Temperatures Limiting Factors Define Grease Life in Four Distinct Temperature Zones," *Tribology and Grease Technology*, September 2010.

[19]Directorate, U. A. A. E., *ADS-79E Condition-Based Maintenance System for US Army Aircraft*, US ARMY, 2016.

[20]Geron, A., *Hands on Machine Learning with Scikit-Learn and TensorFlow*, O'Reilly Media Inc., 2017.

[21]Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., Corrado, G. S., Davis, A., Dean, J., Devin, M., *et al.*, "TensorFlow: Large-scale machine learning on heterogeneous systems, 2015," *Software available from tensorflow. org*, 2015.

[22]Chollet, F. *et al.*, "Keras: Deep learning library for theano and tensorflow," *URL: https://keras. io/k*, 2015.

[23]Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., and Salakhutdinov, R., "Dropout: A simple way to prevent neural networks from overfitting," *The Journal of Machine Learning Research*, Vol. 15, (1), 2014, pp. 1929–1958.

[24]Zhu, E., Mulugeta, H., Hood, A., and Dykas, B., "Learning Condition Indicators via Deep Variational Autoencoding Fault Detection in Rolling Element Bearings," *The 12th International Workshop on Structural Health Monitoring September 10-12, 2019, Stanford, California, USA*, 2019.