Diagnosis of Corrosion Process in Nuclear Power Plant Secondary Piping Structures

M2CA-17-TN-VU__-0703-037

Koushik A. Manjunatha
Andrea L. Mack
Vivek Agarwal
David Koester
Douglas Adams

September 2019

INL
Idaho National Laboratory

Diagnosis of Corrosion Process in Nuclear Power Plant Secondary Piping Structures

Koushik A. Manjunatha, Andrea L. Mack, and Vivek Agarwal
Idaho National Laboratory

David Koester and Douglas Adams
Vanderbilt University

September 2019

Idaho National Laboratory
Idaho Falls, Idaho 83415
http://www.inl.gov

(This page intentionally left blank)

ABSTRACT

Corrosion inside secondary pipe structures in Nuclear Power Plants is difficult to detect without intrusion. Statistical and machine learning techniques are used to predict corrosion using pipe vibration data. 3D accelerometers were used to record pipe vibration data in three directions, using six sensors. Vibration data were processed using the Hilbert-Huang Transformation, which is known for its adaptable basis function to non-linear and non-stationary signals. The Hilbert-Huang Transformation using empirical mode decomposition generates multiple intrinsic mode functions of different amplitude and frequency levels from raw vibration data. From intrinsic mode functions features are extracted which include phase, energy, entropy, mean, standard deviation, skewness, and kurtosis. Logistic regression using only phase information correctly distinguished mass additions from baseline data in roughly 83% and 85% of the cases. Phase information varied substantially among different baseline runs and within trials of the same run. Random Forest resulted in >95% testing accuracy for both mass addition and mass removal data. Support Vector Machine had ~98% testing accuracy for mass removal and ~92% testing accuracy for mass additions when using Permutation Importance feature selection. When using Principle Component Analysis with Support Vector Machine, testing accuracy was ~68% testing accuracy for mass additions and ~85% testing accuracy for mass.

ACKNOWLEDGMENTS

(This page intentionally left blank)

CONTENTS

FIGURES

TABLES

(This page intentionally left blank)

ACRONYMS

EMD	empirical mode decomposition
HHT	Hilbert Huang Transform
IMF	intrinsic mode functions
ML	machine learning
NPP	nuclear power plant
PI	permutation Importance
PCA	principal component analysis
RF	random forest
SVM	support vector machine
VFD	variable frequency drive

(This page intentionally left blank)

Diagnosis of Corrosion Process in Nuclear Power Plant Secondary Piping Structures

1.   INTRODUCTION

This purpose of this work is to reduce the cost of preventive maintenance of secondary structures in nuclear power plants (NPPs) using state-of-the art sensor modalities, data fusion, and advanced data analytics. Specifically, this research aims to develop and validate three-dimensional sensor system capable of inspecting large piping components in secondary systems and providing a pathway to move from inspection based maintenance to condition-based maintenance. This is required because the majority of periodic piping inspections do not discover any degradation, while contributing significantly to operation and maintenance costs. This activity will enable reducing the cost of periodic piping inspections performed at NPPs. The success of this project will enable the following benefits:

- Improved capacity factor by reducing unplanned and planned outages

- Improved safety through fewer unexpected failures and repairs

- Optimized equipment operation and maintenance through early identification of faults

- Improved fault diagnostics through increased availability of data relating to faults and shared knowledge of fault behavior based on case studies and expertise

- Lifetime extension of existing NPPs by having an increased understanding of the current health of the plant components and remaining useful life estimation

- Minimizing human factors effects on non-destructive testing.

The sensor system with advanced analytics will have a larger area of coverage than the current nondestructive evaluation techniques and will aim to answer the question which current Aging Management Programs are facing (such as, where to inspect or identification of inspection locations). There is no currently available technology that would pinpoint which piping component needs to be inspected.

The rest of the report is organized as follows: Section 2 provides an overview of the piping testbed constructed at Vanderbilt University to mimic wall thinning due to corrosion at elbows and tee joints. Section 3 describes range of experimental testing performed and data collected under different operation conditions. Section 4 describes the Hilbert-Huang Transformation performed on experimental data to extract salient features. Section 5 presents data analysis performed and extraction of feature vectors to develop machine learning models for diagnosis. Section 6 presents initial results of machine learning techniques. Finally, summary and path forward is presented in Section 7.

2.   OVERVIEW OF THE PIPING TESTBED

The testbed was constructed with pipe sections fabricated out of 2-inch inner diameter schedule 80 carbon steel pipe and utilize a standard 4-bolt flange connections. A centrifugal pump (130 gpm at 0 head) is part of the testbed design to provide fluid flow through the piping system. The flow can be shut off or diverted using ball valves installed next to each fitting. The system included a 30 gallon tank with a heating element, a pump, and a network of pipes that fed back into the tank intake. The pump was driven by a variable frequency drive (VFD) motor controller, and manual ball valves were used to control flow patterns in the pipe network. Figure 1 gives an overview of the pipe network used in this research. The water temperature was measured with a resistance temperature detector probe in the tank, and the pipe surface temperature was measured with a Milwaukee handheld infrared thermometer.

The test bed was designed with two removable flanged fittings: a pipe elbow and pipe tee. The fittings were made removable so that studies of mass removal (representative of corrosion in the inner pipe walls) could be conducted, fittings could be replaced as needed, and so that the test bed provided flexibility. The pipe elbow fitting was the focus of this research. Six (6) PCB Piezotronics triaxial accelerometers used to collect vibration data (as shown in Figure 2) were mounted on one of the pipe elbow fittings. They are labeled from 101 to 106 to identify each accelerometer. The six triaxial accelerometers were attached approximately two centimeters apart along the bend of the pipe. Table 1 shows the model, serial number and sensitivity of the six triaxial accelerometers used for testing. These triaxial accelerometers were used to collect both experimental and operational modal analysis response data.
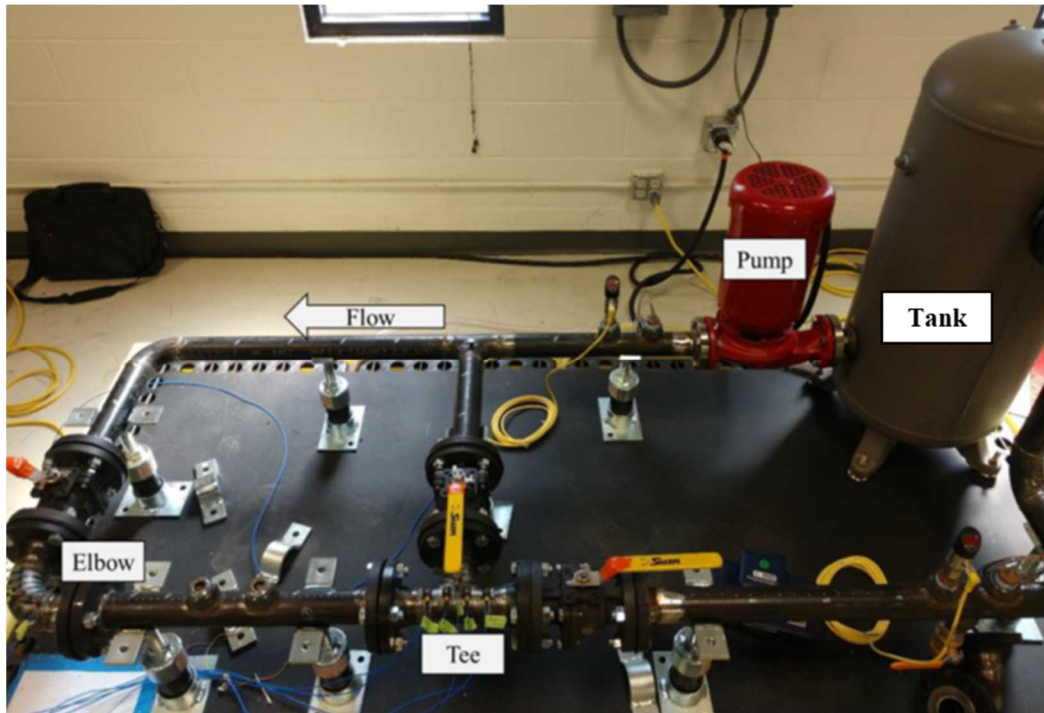
Figure 1. An overview image of the experimental pipe testbed.
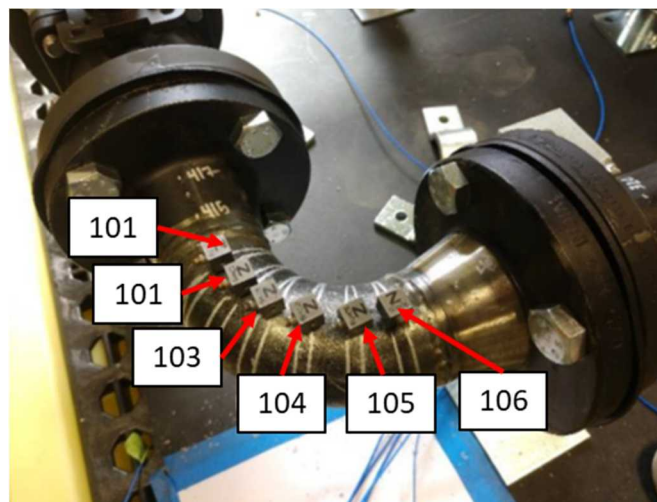
Figure 2. Pipe elbow shown with accelerometers.

Table 1. PCB Piezotronics triaxial accelerometers labeling and sensitivities.

Accelerometer label	Serial number	Sensitivity (mV/g)
101X	87143	100
101Y	87143	100
101Z	87143	100
102X	80778	100
102Y	80778	100
102Z	80778	100
103X	92855	100
103Y	92855	100
103Z	92855	100
104X	98930	100
104Y	98930	100
104Z	98930	100
105X	27044	100
105Y	27044	100
105Z	27044	100
106X	98745	100
106Y	98745	100
106Z	98745	100

A National Instruments NI cDAQ-9178 chassis with NI 9234 data acquisition cards shown in Figure 3 was used to collect data from the six triaxial accelerometers mounted on the pipe elbow.
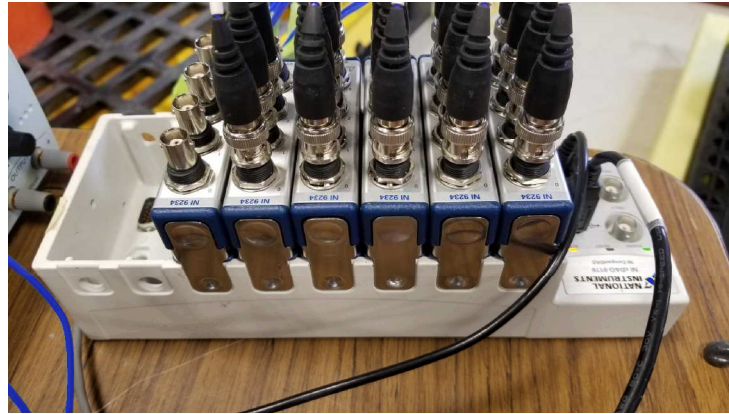
Figure 3. National Instruments NI cDAQ-9178 with NI 9234 data acquisition card.

3. EXPERIMENTAL TESTING AND DATA

A series of experiments were performed under different setting to understand the flow-loop and mimic corrosion in the elbow and in the tee. All the acceleration time history data collected for each experimental testing were collected by Vanderbilt University and shared with Idaho National Laboratory for model development using machine learning techniques. Prior to modify the condition of the experimental flow-loop, a series of baseline accelerometer data were collected. Figure 4 shows the standardized raw accelerometer signal for X, Y and Z directions of sensor 101.
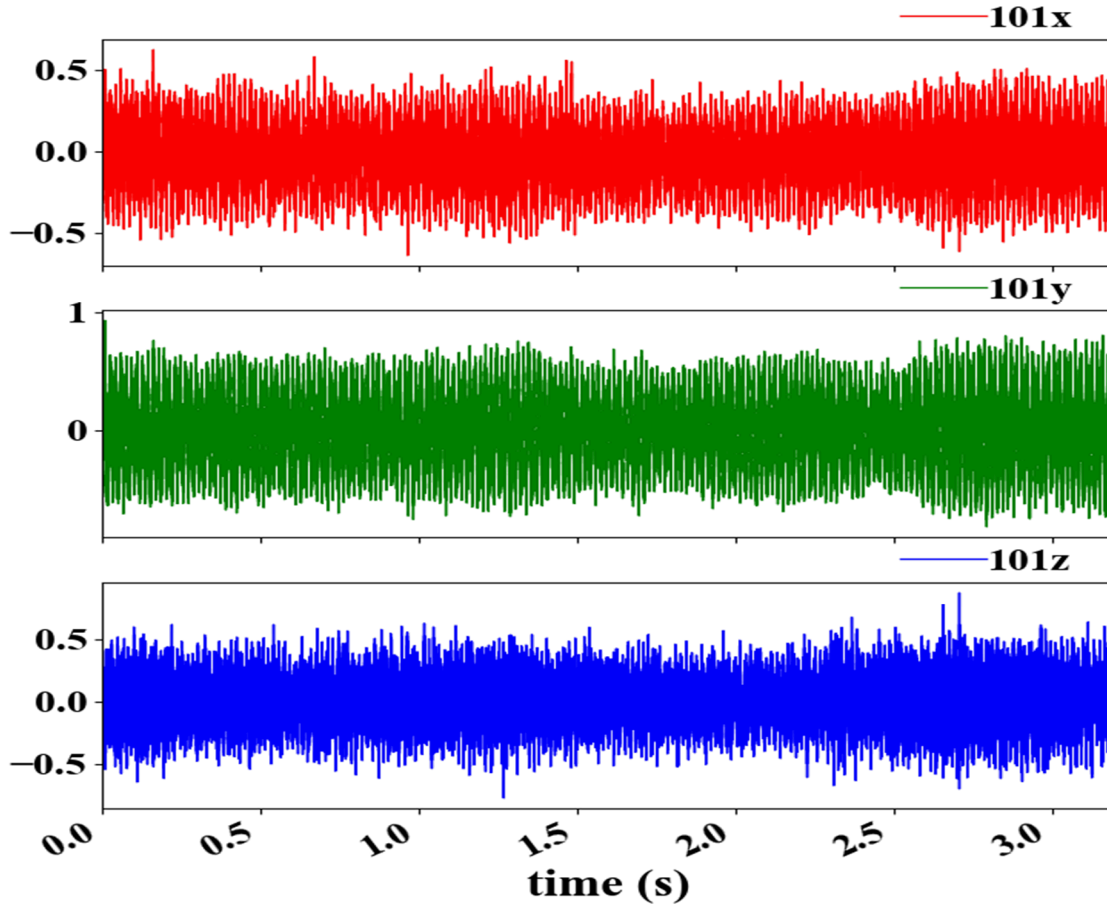
Figure 4. Raw accelerometer signal for X, Y, Z directions of accelerometer 101.

3.1   Addition of Mass at Multiple Pipe Elbow Locations

The operating data of the experimental loop was collected with masses added at multiple pipe elbow locations on June 11, 2019. The masses was placed under the accelerometer 101, between the accelerometers103 and 104, and under the accelerometer 106 as shown in Figure 5. The testbed water brought to 80 ☐ and held at temperature with heating unit. The pump was run at 60 Hz on VFD controller ('max flow'). Flowrate not directly measured, but same 60 Hz control input was used for all tests. Data was collected on all 6 triaxial accelerometers for 3.2 seconds at a sample rate of 51200 samples/second. Ten (10) independent samples of data were collected for each trial. All masses placed on bottom side of pipe elbow are shown in Figure 5. Data was collected with 2.9 grams, 6.4 grams, 11.4 grams, 13.8 grams, and 23.5 grams mounted to the elbow at the four locations described.

3.2   Addition of Mass at Multiple Pipe Tee Locations

The operating data of the experimental loop was collected with masses added at multiple pipe tee locations on June 13, 2019. The testbed was operated at 80 ☐ and head at temperature with heating element. The pump was run at 60 Hz on VFD controller ('max flow') in all trials. Data was collected for 3.2 seconds at 51200 samples/second. Figure 6 shows the location of the accelerometers mounted to the tee. Masses were placed at locations on the underside of the pipe tee as shown in Figure 7. Data was collected with 2.9g, 6.4 g, 11.4g, 13.8 g, and 23.5 g mounted to the tee at the four locations described.

6
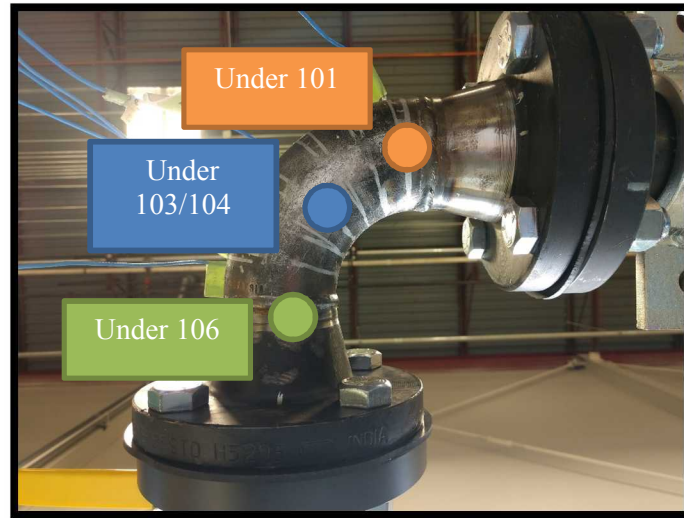
Figure 5. Underside of the elbow with locations of mass placement labeled. Note that 'under 103/104' refers to a mass placed between the planes of the accelerometers 103 and 104 (whereas in the case of accelerometers101 and 106, the mass was placed directly under their respective accelerometers).
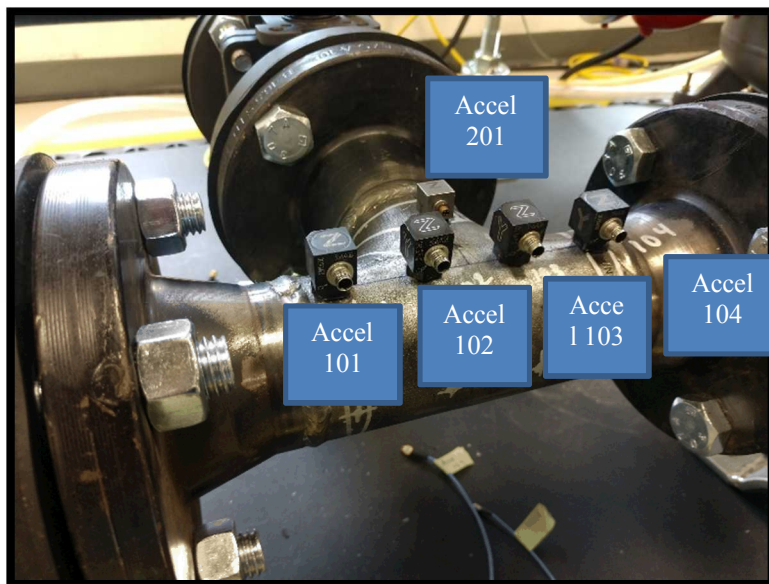
Figure 6. Accelerometer locations on the tee (accelerometer101 to the left, accelerometer 104 to the right, and accelerometer 201 at the back).
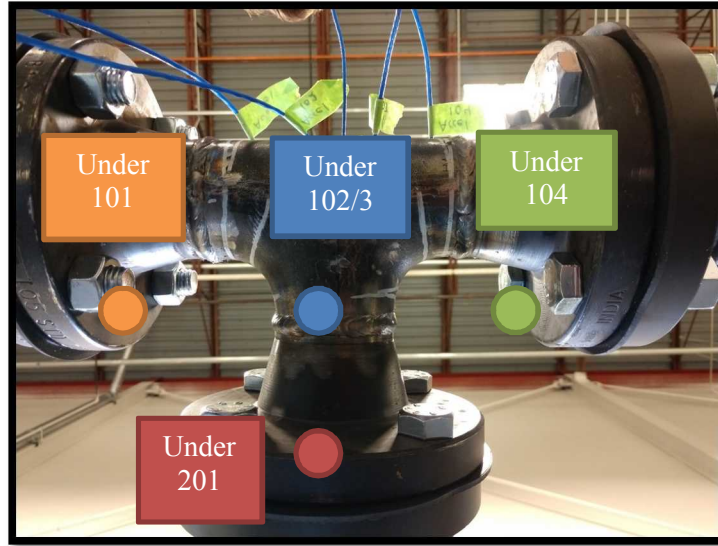
Figure 7. Underside of tee with locations of mass placement labeled. Note that 'Under 102/3' refers to a mass placed between the planes of accelerometers 102 and 103 (whereas accelerometers 101, 104, and 201 refers to the masses placed directly under respective accelerometers).

3.3   Variability in Operating Data along Pipe Elbow and Tee

The operating data of the experimental loop was collected with masses added at pipe elbow and tee locations from June 12, 2019 to June 21, 2019. The testbed was operated at 80 □ and head at temperature with heating element. The pump was run at 60 Hz on VFD controller ('max flow') in all trials. All of the data was collected at approximately 2:00 pm (+/- 1 hour) each day. The testbed was brought up to steady-state operating temperature with the thermostat set at 80 □. The temperature varied between 78 □ and 80 □ due to cycling of the heating element on the tank. Data was collected from elbow starting for 8 days. Data was collected for all flow configurations to include

- Elbow flow configuration: through elbow
- Tee flow configurations: through elbow, through both, and through tee.

3.4   Addition of Mass Roving on the Outside of Pipe Elbow

The operating data at multiple locations along the pipe elbow was collected on June 27, 2019. The testbed was operated at 80 □ and head at temperature with heating element. The pump was run at 60 Hz on VFD controller ('max flow') in all trials. Operating tests were conducted on the pipe elbow with masses placed on the 'outside' or 'outer channel' of the pipe elbow along the cross section as shown in Figure 8.

3.5   Removal of Mass from Pipe Elbow

The operating data as mass was removed from the pipe elbow was collected between June 28, 2018 and July 8, 2019. The testbed was operated at 80 □ and head at temperature with heating element. The pump was run at 60 Hz on VFD controller ('max flow') in all trials.

This section will cover the first mass-reduction experiments conducted on a pipe elbow. Each sub-header corresponds to a different mass reduction trial. In all of the trials, the pipe elbow was mounted with the same EPDM rubber gaskets in the test bed and operating data were collected at 80 □ equilibrium.

Pressure was not measured for this experiment, and it was not controlled. Table 2 shows the amount of mass removed for each trial.
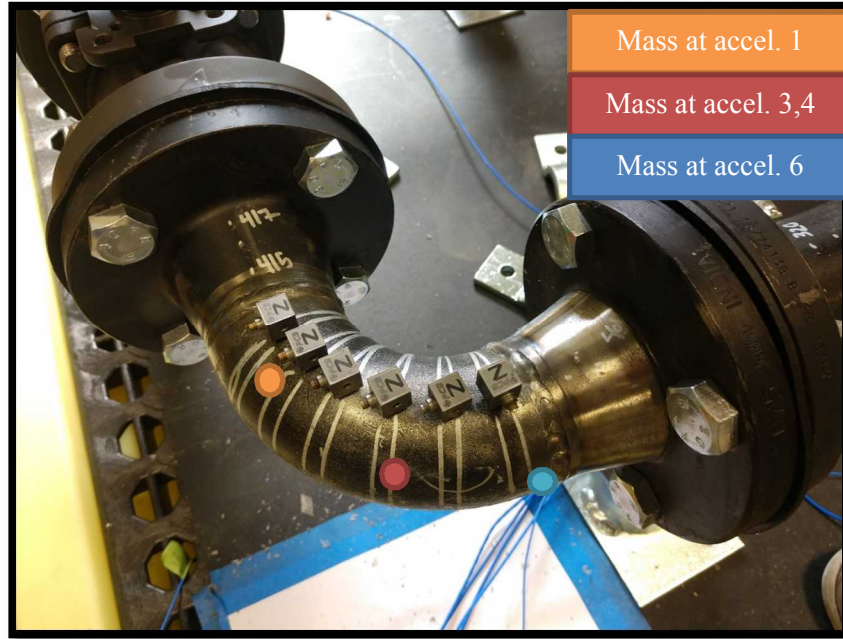
Figure 8. Pipe elbow with accelerometers and mass placement.

Table 2. Mass removed from the elbow for each trial.

Trial Number	Total Mass Removed	Percent of Initial Mass Removed
1	19 grams	0.3%
2	32 grams	0.5%
3	48 grams	0.75%
4	63 grams	1%
5	79 grams	1.25%
6	95 grams	1.5%

A 0.3% reduction in mass was used to simulate mass transfer from outside channel of pipe elbow (i.e. remove 20g evenly from one side of the elbow with a carbide rotary bit). Note that the baseline measurement from the mass roving experiment was used as the baseline measurement for the mass reduction trials on the elbow. The weight of the elbow before any mass removal was 6350g $\pm$ 1g. The mass of the elbow after first mass removal (0.3%) was 6330 $\pm$ 1g. Material was only removed inside of the flange weld to avoid the weld line. The entire outer surface was ground in final iteration of 0.3% material removal. Figure 9 shows where the mass was removed from the elbow.
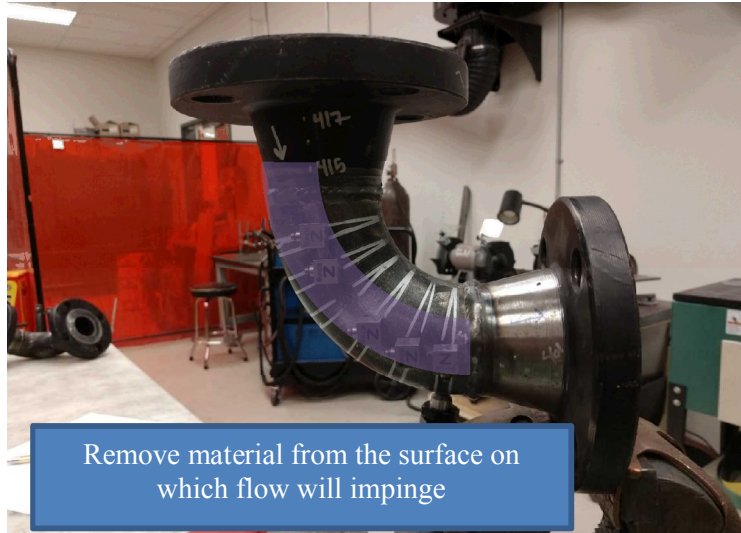
Figure 9. The mas was removed from the elbow is shown in the shaded area.

3.6 Mass Removal on Second Pipe Elbow

The operating data as mass was removed from the pipe elbow was collected between July 16, 2019 and July 18, 2019. The testbed was operated at 80 □ and head at temperature with heating element. The pump was run at 60 Hz on VFD controller ('max flow') in all trials. A similar process was used in the second iteration of the mass reduction trial on the second elbow, with the exception that the EPDM rubber gasket was replaced between installations of the pipe elbow. This change was adopted because the deterioration of the rubber gasket between installations and removals was deemed significant. It was determined that the variability introduced by using a fresh, new EPDM gasket on each installation would introduce less variability than using the same deteriorating gasket for each trial.

First temperature versus damage index test on second elbow – EPDM gasket

The operating data on the pipe elbow was collected on July 19, 2019. The water temperature was varied from 25 □ to 80 □ in steps of 5 □. The water pressure was 0 psi and pump was run at 60 Hz on VFD controller ('max flow') in all trials.

The second elbow with 1.50% of its mass removed was interrogated for the impact of operating water temperature on damage index. The elbow was left on the test bed after the second mass reduction trial (with EPDM gaskets). Temperature was recorded in 5 □ increments with the resistive temperature detector (RTD) temperature probe. Operating data collection was started within 20 seconds of the RTD probe reading the target temperature. The equilibrium operating temperature of the test bed without the heating element was found by operator experience to be approximately 68 □. Therefore, from 25 □ to 65 □ the test bed temperature was slightly increasing over the course of the data collection. Typically, the temperature had increased by either 0 or 5 □ by the time the temperature was checked again (the heating element and RTD probe were controlled by the same switch, so to check the temperature, the heating element was turned on as well. This heat input was not desired when a relatively constant operating temperature measurement was being taken, so the element and probe were left off during data collection). For 70 □ and above, the RTD probe and heating element were left on to maintain the desired operating temperature, and they automatically cycled the heating element to achieve the desired water temperature. Pressure was relieved to 0 psi (atmospheric) with a manual valve on the tank before every dataset was collected. 0psi was confirmed with an analog pressure gauge on the tank.

Second Temperature versus damage index test on second elbow – EPDM gasket

The operating data on the pipe elbow was collected on July 22, 2019. The water temperature was varied from 25 □ to 80 □ in steps of 5 □. The water pressure was 0 psi and pump was run at 60 Hz on VFD controller ('max flow') in all trials.

Pressure versus damage index test on second elbow – EPDM gasket

The previous temperature versus damage index test on second elbow was repeated again on July 23, 2019. The elbow was left on the rig and the same test was conducted. This was to confirm repeatability of the temperature vs damage index relationship. The operating data on the pipe elbow was collected at water temperature of 80℃ and pump was run at 60 Hz on VFD controller ('max flow') in all trials. The water pressure was varied from 8 psi to 0 psi in steps of 0.5 psi. The same damaged elbow with the same EPDM gasket was tested for the variation between measured damage index and measured operating pressure. The test bed, starting at approximately 80 □ fluid temperature and 0 psi, was operated up to 80 ℃ and allowed to pressurize to slightly above 8 psi. The tank was manually set to 8 psi, and then operating data was collected on the test bed. The tank pressure was reduced by 0.5 psi between data acquisition runs. The test bed was held at 80 □ with the RTD probe and heating element.

Pressure vs damage index test on second elbow – Aramid/Buna-N gasket

The operating data on the pipe elbow was collected on July 24, 2019. The operating data on the pipe elbow was collected at water temperature of 80 □ and pump was run at 60 Hz on VFD controller ('max flow') in all trials. The water pressure was varied from 8 psi to 0 psi in steps of 0.5 psi. The damage index vs pressure trial was repeated, but the EPDM gaskets were removed and replaced with stiffer Aramid/Buna-N gaskets. The goal was to observe the trend in damage index with respect to a change in pressure. If the relationship appeared to be relatively stable across boundary conditions, then the trend between damage index and pressure may be caused by some physical phenomena common between both boundary condition configurations. Similar to before, the test bed was heated and pressurized to its equilibrium operating conditions. Pressure was released from the tank in 0.5 psi increments and operating vibration data were collected between pressure releases. Note that the absolute value of the damage index measured was not of interest. Only the trend in damage index with operating parameter was of interest.

Temperature versus damage index test on second elbow – Aramid/Buna-N gasket

The operating data on the pipe elbow was collected on July 25, 2019. The operating data on the pipe elbow was collected at water temperature of 80 □ and pump was run at 60 Hz on VFD controller ('max flow') in all trials. The water pressure was varied from 8 psi to 0 psi in steps of 0.5 psi. The temperature versus damage index test was repeated with the stiffer Aramid/Buna-N gasket. The procedures used in this experiment were the same as described in the previous experiment using the Aramid/Buna-N gasket.

4. HILBERT-HUANG TRANSFORM

Vibration data were processed using the Hilbert-Huang transformation (HHT). The HHT is an empirical alternative to traditional signal processing methods such as Fourier transform and wavelet analysis. As an empirical method, the HHT allows the data to determine the basis functions, such that they do not have to be defined a priori. This flexibility is particularly beneficial as other signal processing methods are limited to stationary and/or linear signals, assumptions that are rarely met. As such, the HHT has emerged as a popular signal processing method for non-stationary, nonlinear data resulting in a time-frequency-energy representation [1]. The HHT consists of two steps: (1) Empirical Mode Decomposition (EMD) and (2) Hilbert transformation.

4.1 Empirical Mode Decomposition

EMD works by extracting the different intrinsic modes of oscillations, forming intrinsic mode functions (IMFs). IMFs are defined as having equal (or differing by one) extremums and zero-crossings and having a mean value of zero for the local maxima and minima at any point. Traditional signal

processing methods result in simple harmonics with constant amplitudes and frequencies. In contrast, IMFs represent simple intrinsic modes of oscillations with varying amplitudes and frequencies.

IMFs are found using a sifting process. As described in [1], the EMD process extracts local signal information by eliminating riding waves. The sifting process serves two purposes: "*to eliminate riding waves, and to make the wave profiles more symmetric. ... the first purpose must be achieved for the Hilbert transform to give a meaningful instantaneous frequency, the second purpose must also be achieved in case the neighboring wave amplitudes have too large a disparity*". Cubic splines are fit to connect local maximum and minimum values of the original signal as line "envelopes". The local means of the upper and lower envelopes (c) are found, and the observed signal ($x(t)$) is subtracted from the envelope mean to form the residual (r_1) for IMF (c_1)

$$r_1 = x(t) - c_1 \tag{1}$$

This sifting process continues to create the second IMF (c_2)

$$r_2 = r_1 - c_2 \tag{2}$$

The sifting process repeats until a stoppage criterion is met or the residual (r_n) is monotonic, not allowing further IMFs to be found. The stoppage criteria is subjective, but resulting IMFs have been found to be fairly robust to stoppage criteria chosen [1]. One stoppage criterion is to end the sifting process when the number of extrema and zero crossings differ by at most one in subsequent IMFs, after a predetermined S-number. The S-number sets a minimum number of IMFs before looking for stoppage.

The original signal can then be reconstructed by summing the IMFs and the residual

$$x(t) = \Sigma_k c_k + r_n \tag{3}$$

Once the signal has been decomposed into IMFs, the Hilbert transform is applied to extract the instantaneous phase and instantaneous frequency information for each IMF signal. Figure 10 and Figure 11 show the extracted IMFs for the baseline data. IMFs decrease in frequency by design. Not all IMFs contain useful information. Empirical studies suggest period increases with IMF. IMF energies from Gaussian white noise follow a normal distribution but deviate from normality at later IMFs [1]. Statistical significance of IMFs as a means to distinguish IMFs with signal information from noise may be explored in future work. IMF significance will aid interpretation in terms of the system parameters and dimension reduction may be explored in future work. Next the HHT process is explained.

4.2   Hilbert Transform

The Hilbert transform (H) extracts the imaginary portion of the complex IMF signal ($z(t)$). It is given as:

$$H[x(t)] = \frac{1}{\pi} PV \int_{-\infty}^{\infty} \frac{x(t)}{t-\tau} d\tau \tag{4}$$

As the convolution of $x(\tau)$ and $y(t) = \frac{1}{t-\tau}$, the Hilbert transform calculates the amount of overlap between x and y and scales it by π. Where the complex signal is:

$$z(t) = x(t) + iH[x(t)] = a(t)e^{i\theta(t)} \tag{5}$$

and instantaneous amplitude ($a(t)$), phase ($\theta(t)$), and instantaneous frequency ($\omega(t)$) subsequently can be found as:

$$a(t) = \sqrt{x(t)^2 + H[x(t)]^2} \tag{6}$$

$$\theta(t) = tan^{-1}(\frac{H[x(t)]}{x(t)}) \tag{7}$$

$$\omega(t) = \frac{d\theta}{dt} \tag{8}$$

The phase is interpreted as the number of rotations at a given time point. Phase information has been found to be very useful in detecting changes from normal operating conditions [1] and is described in more detail Section 5.

4.2.1   Instantaneous Phase

Constraining the phase function direction to clockwise or counterclockwise and restricting the phase to have a unique center at each time point will result in unique amplitude and phase combinations that satisfy *(5)* at each time point [1]. The total number of rotations at time t is the sum of the $\theta_k(t)$'s:

$$\theta(t) = \Sigma_k \theta_k(t) = \Sigma_k tan^{-1} \frac{H[c_k(t)]}{c_k(t)} \tag{9}$$

Unwrapping the total phase gives a cumulative phase function over time, such that the unwrapped phase is a monotonically increasing function. The derivative of the unwrapped phase is then used to calculate the instantaneous angular frequency ($f(t)$).

Phase information can be extracted using the raw vibration signal. However, when plotting the phase extracted from an original signal versus time, [1] shows how intrinsic modes of oscillation are missed. That is, the unwrapped phase displays further oscillations when plotted versus time. This suggests that phase extracted from the original signal may underestimate the number of rotations at unknown points in the time series

The absolute phase shift ($\phi_i(t)$) from the baseline (b) at each weight (w_i) at each time point allows differences in phase functions to be seen more clearly. The absolute phase shift is calculated as:

$$\phi_i(t) = |\theta_b(t) - \theta_{w_i}(t)| \tag{10}$$

Instantaneous phase, frequency, and amplitude are calculated from the IMFs and form the basis for all features used in data analysis. Other features are derived from those listed. Data analysis of the processed signal is described Section 5.

5.   DATA ANALYSIS

Both statistical and machine learning methods are considered for analyzing the vibration signal. Figure 12 depicts the flowchart of vibration data analysis to predict the pipe condition. Accelerometer data is normalized and fed into the HHT process. Features are extracted. Logistic regression is used to predict pipe condition using phase information. Traditional machine learning algorithms utilize additional features in predicting pipe condition.

5.1   Statistical Analysis

Data can be analyzed at each time point within each IMF, at each IMF, at each time point combined over IMFs, and at the file level. The statistical analysis used data at each time point, combined over IMF. Baseline data were collected on 9 days. Figure 13 displays the baseline data (yellow) relative to the data with mass added. Relative phase trends are similar between baseline data and mass added data.

Figure 14 facets the data in Figure 15 by pipe condition to make trends more apparent. Baseline data collected on June 13th have much different phase values than other days. Experimentalists should be consulted to try and explain this. After 1 second of data collection, differences in phase as a function of mass additions become apparent. A mass of 23.5g has more relative rotations than the baseline and is clearly distinguishable from mass additions of 2.9g and 13.8g after one second. Mass additions of 6.7g and 11.4g appear to increase in relative phase rate substantially around two seconds whereas other mass

additions do not. The pipe should vibrate at a lower frequency when more mass is added, which translates to fewer rotations than the baseline, and a larger absolute relative phase shift from the baseline. Therefore, we would expect data associated with more mass to have a higher relative phase at all time points. Figure 8 shows this is not always the case, indicating background variability and/or variables that are not accounted for.
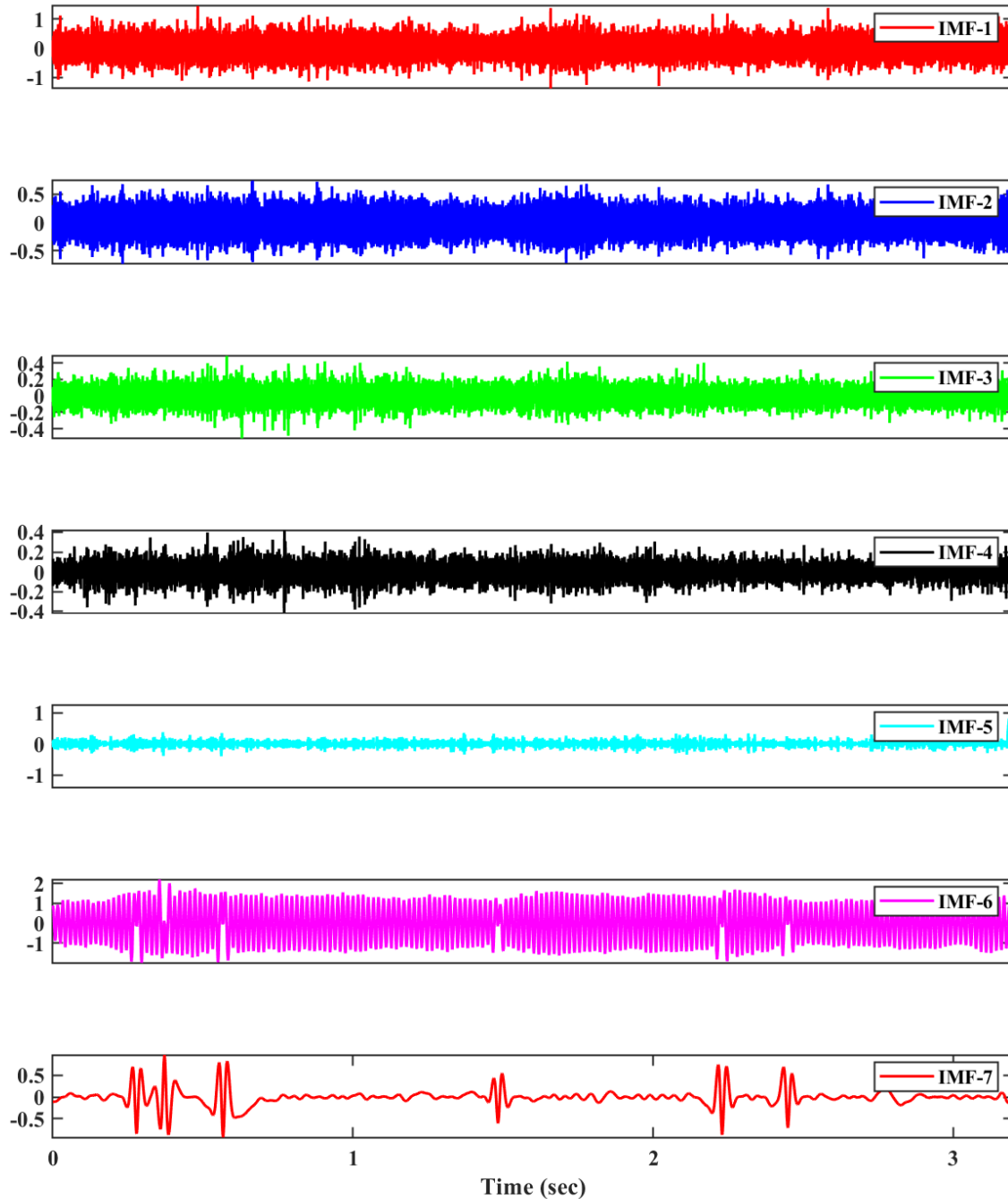

Figure 10. Extracted IMFs for baseline data (part 1).

14
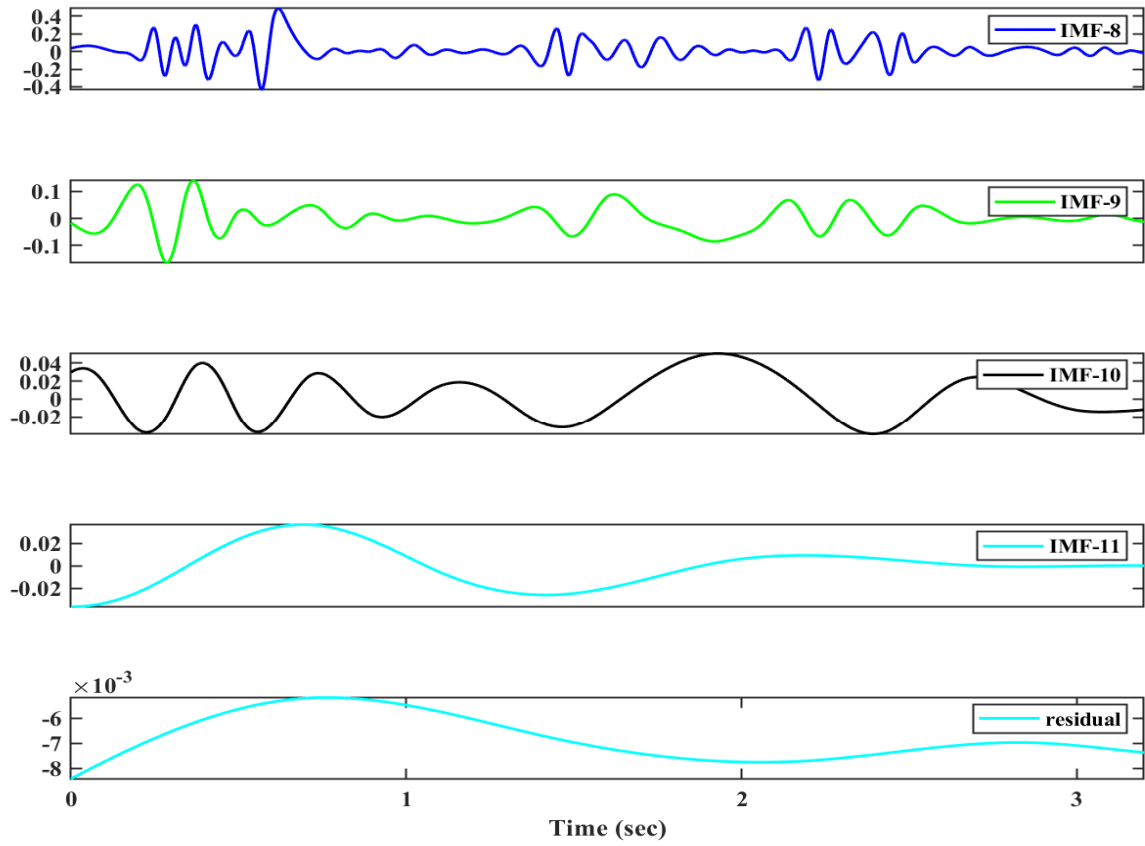
Figure 11. Extracted IMFs for baseline data (part 2).
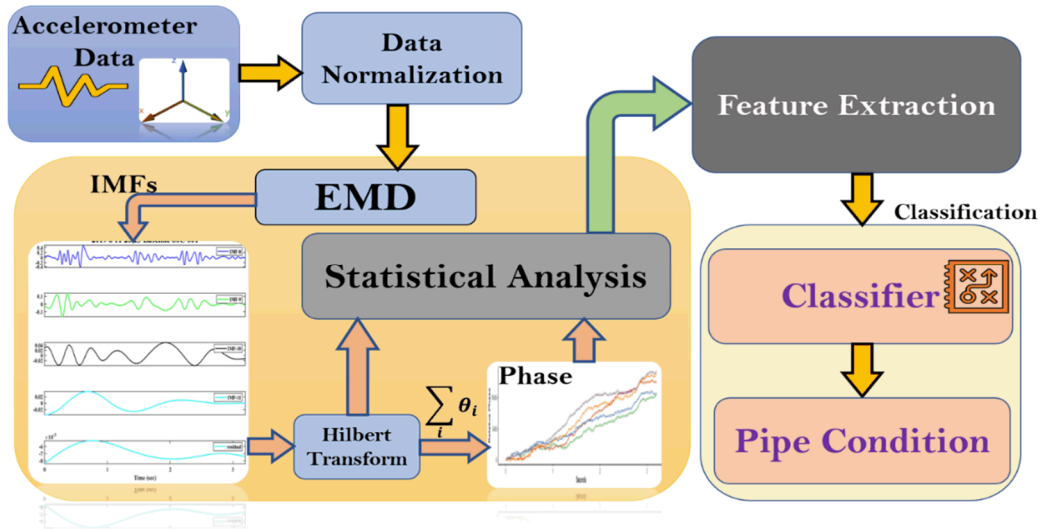
Figure 12: Data analysis flow diagram.
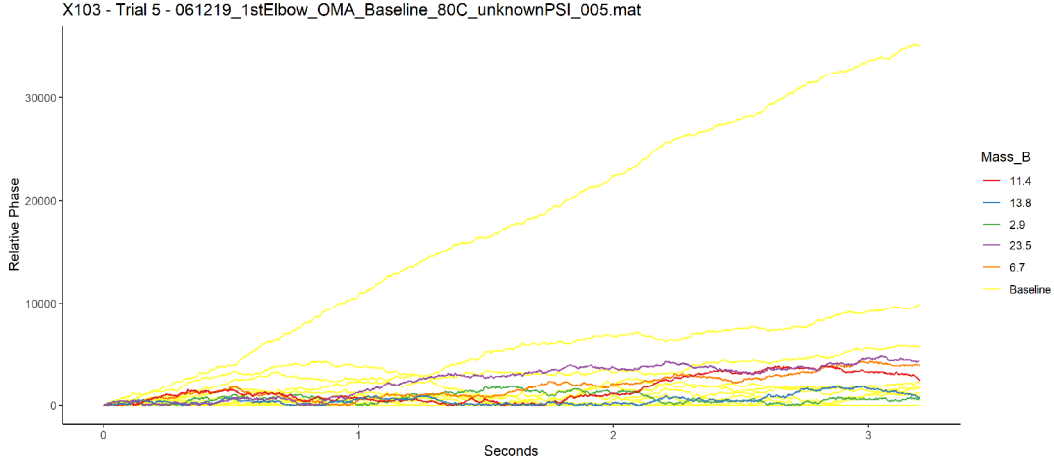
Figure 13: Trial 5 variability among baseline files and different mass additions.
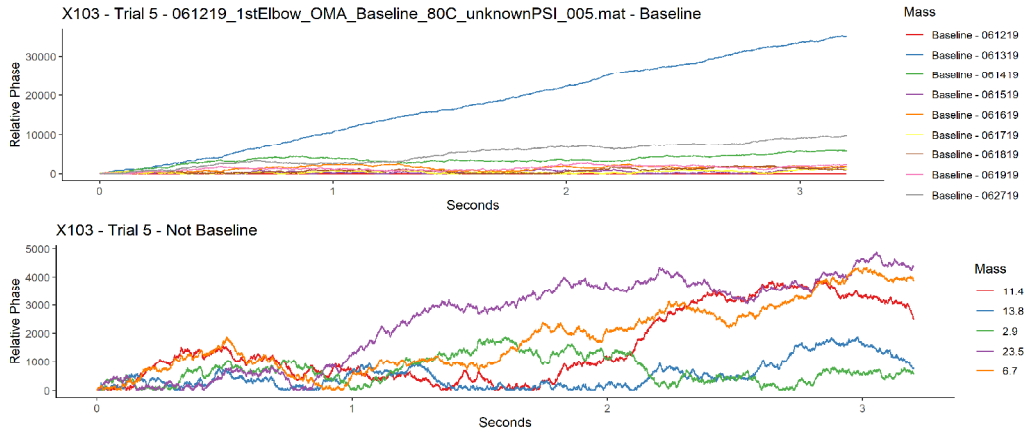
Figure 14: Trends in relative phase faceted by baseline/not baseline.

Data collected from sensors 103 and 104 in the x-direction for nine baseline runs and runs with 23.5g mass removal between locations 103 and 104 were used in this analysis. Baseline data were limited to that collected in the first elbow, as that is also the location of the mass variations. The baseline data recorded on the day (June 12, 2019) closest to 23.5g mass removal (June 11, 2019) were used. Each run had 10 separate trials. The analysis assumes the pipe system was reset between each trial.

Three statistical models were fit with increasing complexity. All models use the binary response distinguishing baseline data from 23.5g mass added data. Model A only uses data from Trial 5 and is given below:

$$Logit = log[\frac{\pi_B}{1-\pi_B}] = \beta_0 + \beta_1 * phase + \beta_2 * time + \beta_3 * phase * time \qquad (11)$$

where π_B = probability of baseline signal

β_0 = Logit (log-odds) of baseline signal at time 0, phase 0

β_1 = Rate of change in log-odds of baseline signal for each additional rotation at time 0

β_2 = Rate of change in log-odds of baseline signal for each additional second without rotation (phase=0)

16

β_3 = Each additional second, the rate of change in log-odds of baseline signal for each rotation

Note the parameters corresponding to 0 seconds and a phase of 0 are not meaningful. Model A assumes each observation is independent. It accounts for time in the mean estimate of the model but not in the covariance structure. Therefore, it is inherently mis-specified, and the covariance is under-estimated. Model A had 82.9% prediction accuracy. Figure 15 displays phase versus time facetted by baseline/23.5g. Color indicates prediction accuracy for Model A. The plot shows it is more difficult to distinguish from the baseline at earlier points in the time series. Red/green colors indicate accurate predictions for each pipe condition.

Model B includes data from all trials. It adds a covariance parameter (ρ) that models observations at lag (j) to be correlated as ρ^j within each trial, by using a generalized linear mixed model (GLMM) and a random effect for each trial. The benefit of the GLMM is the ability to correctly specify and estimate the correlation between observations. Additionally, when fitting a model with all trials the portion of trial-to-trial variability relative to total variability can be estimated. The fixed effects portion of Model B is the same as A. Model B did not converge. The covariance matrix is singular, likely because the lag-1 correlation for sensor 103, direction-X was 0.999, almost perfectly correlated. Finally, a GLM with all trials was fit with trial, phase, and time main effects and phase interaction terms with trial and time (Model C). Prediction accuracy for Model C was 85.2%. Figure 16 indicates correct predictions with green and incorrect predictions with red by trial.

Table 3 has the corresponding prediction accuracy by trial. Trials 1, 4, 5, and 9 had accuracy less than 90%. This suggests substantial trial-to-trial variability. That is, in certain trials it is easy to distinguish baseline from mass added and in others it is not. Figure 17 shows how variable the 23.5g mass added phase data are relative to the baseline, among the trials.

Other signal features should be included in the model or other environmental variables need to be controlled to more consistently get accurate predictions. Finally, the Table 4 compares prediction accuracy for the three models considered.
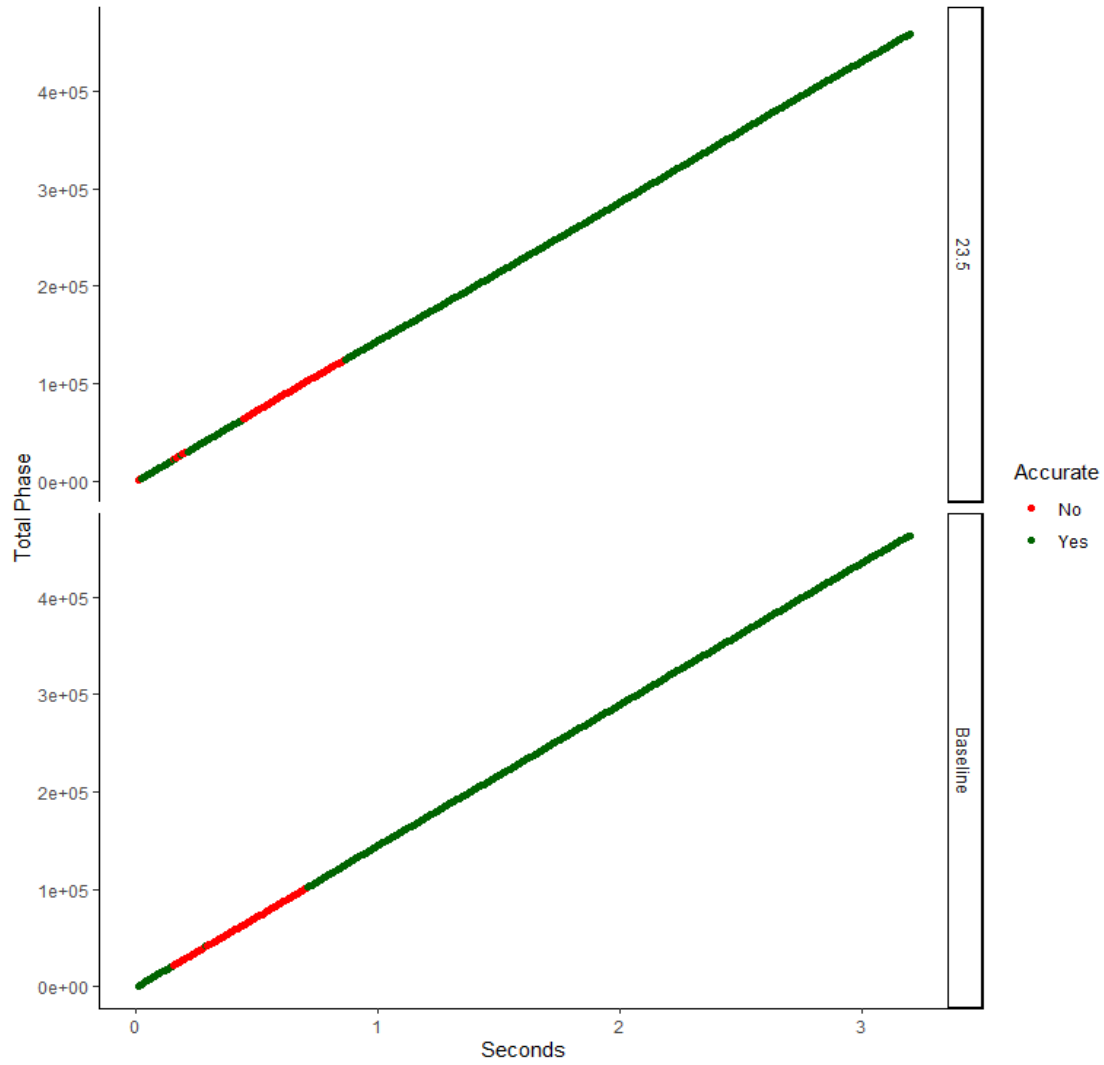
Figure 15: Trends in prediction accuracy from Model A.

Table 3. Model Prediction accuracy

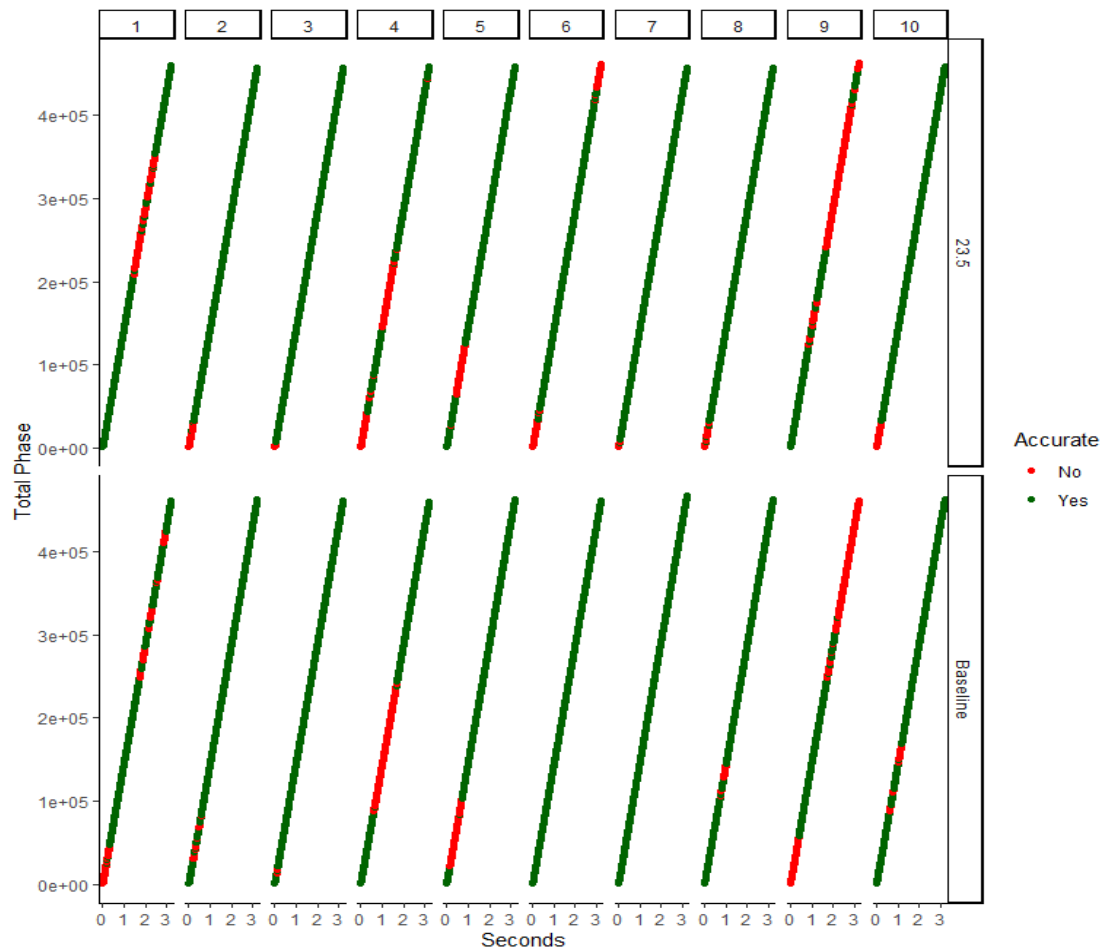| Trial	Prediction Accuracy	T rial	Prediction Accuracy
1	0.7595	6	0.9259
2	0.9389	7	0.9919
3	0.9849	8	0.9478
4	0.6854	9	0.5030
5	0.8437	10	0.9308

Figure 16: Trends in prediction accuracy for Model C.
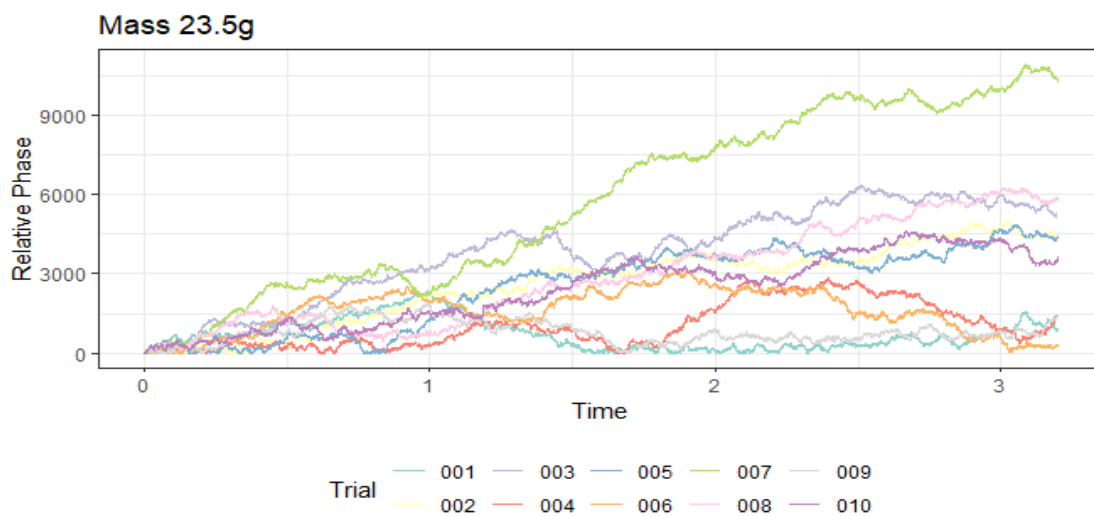
Figure 17: Variability in relative phase among trials.

Table 4. Prediction Accuracy

Model	Prediction Accuracy
Model A	82.9%
Model B	Did not converge
Model C	85.2%

While the statistical analysis only considered phase information at the time level, the Machine Learning approach used multiple features at various levels.

5.2   Feature Extraction

IMFs of each direction for every sensor are used to generate phase values and other statistical features. Summary measures of the phase values are further used to generate additional information such as slope. Based on the information collected the features are generated. Generated features are used in classical machine learning algorithms to predict condition of the pipe.

IMF level features include mean (μ), standard deviation (σ), entropy (EN), energy (E), skewness (SK), and kurtosis (KU) as shown in Table 5 [2]. Whereas phase features include maximum phase (MP), minimum phase (mP), mean phase (EP), maximum slope of the phase (MS), minimum slope of the phase (mS), mean slope of the phase (ES), standard deviation of slope of the phase (SDP). The slope/rate of the unwrapped total phase at each time point is calculated by:

$$\frac{d\theta_{tot}}{dt} = \theta_{tot,t+1} - \theta_{tot,t} \tag{12}$$

Table 5. Statistical features from IMFs.

Energy (E) $$\sum_{j=1}^{N}|X_j|^2$$ Entropy (EN) $$-\sum_{j=1}^{N}X_j^2\log(X_j)^2$$ Mean (μ) $$\frac{1}{N}\sum_{j=1}^{N}X_j$$	Standard Deviation (σ) $$\sqrt{\frac{1}{N}\sum_{j=1}^{N}(X_j-\mu)^2}$$ Skewness (SK) $$\frac{\sum_{j=1}^{N}(X_j-\mu)^3}{\sigma^3}$$ Kurtosis $$\frac{\sum_{j=1}^{N}(X_j-\mu)^4}{\sigma^4}$$

Using the aforementioned features for each IMF, a feature vector is constructed by combining features from X, Y and Z directions of each accelerometer. After feature extraction, finding dependencies within the features and with the class label is crucial to make more accurate predictions.

5.3   Feature Selection

The dependency of each feature on the final prediction is unknown. Not every feature from every IMF will contribute to the decision making. To what extent each feature influences the final prediction is also unknown. On the other hand, phenomenon of 'curse of dimensionality' should also be considered since, as the feature dimension increases the problem becomes computationally expensive and difficult to solve. Hence, according to Section 3.2, multiple features are generated, and using algorithms such as Principal Component Analysis (PCA) [3] , Permutation Importance (PI) [4] best features are selected. In addition, if advanced deep learning algorithms are used, then the algorithm can be trained with all the

features and the algorithms pick the most significant features to make predictions. In this work, the performance of feature selection is verified for both PCA and PI.

PCA is an unsupervised dimensionality reduction or feature elimination algorithm as it does not consider the output variables to reduce the dimension of the feature vectors. PCA combines input features in a specific way using a statistical procedure that uses orthogonal transformations to convert sets of possibly correlated variables into a set of values of linearly uncorrelated variables, called principal components. The first principal component has the highest possible variance, and each succeeding component indeed has the highest variance possible under the constraint that it is orthogonal to the preceding components. Hence, at the end, the least important features are dropped while still retaining the important information of all the variables. PCA also ensures that the new features are independent to each other.

To identify the importance/influence of each feature on the prediction, Permutation Importance (PI) is used after the model has been trained using ML algorithms. Given a trained model, test data set, and an evaluation metric, PI iteratively permutes (shuffles) a feature column (one at a time) and the performance is evaluated for a given model on the modified dataset. The performance score for each iteration is stored as $PFI = P_b - P_s$, where P_b is the base performance metric score and P_s is the performance score after shuffling. Higher the difference value implies the feature is more important. Advantage of this method is it considers the output variable while making feature selection but due to the iterative process it is time consuming. At the end, a subset of the original features is used in modelling and then for the predictions.

5.4   Machine Learning Models

Machine learning is a method of data analysis which automates an analytical model by learning from the data, identifying patterns and making decisions with minimum human intervention. Some of the machine learning algorithms used in the data analysis are discussed.

5.4.1   Random Forest

Decision tree is a supervised learning algorithm which splits the features into nodes of a tree. The root node of the tree is placed with the best feature and the leaf node is the decision variable (class label). Decision is made by following the tree from the root node along the tree to the leaf node. But simple decision tree suffers from high variance (overfitting) since they are extremely sensitive to data. Bagging reduces the overfitting by taking an average of several decision trees with bootstrap sampling. Bootstrap sampling draws samples from training data with replacement. Whereas Random Forests (RF) [5] are also bagged decision tree models but each decision tree has a random subset of features which eliminates any correlation between the trees and improves the performance. This means each model is trained with only a small subset of features (i.e. m out n features) as shown in Figure 18.
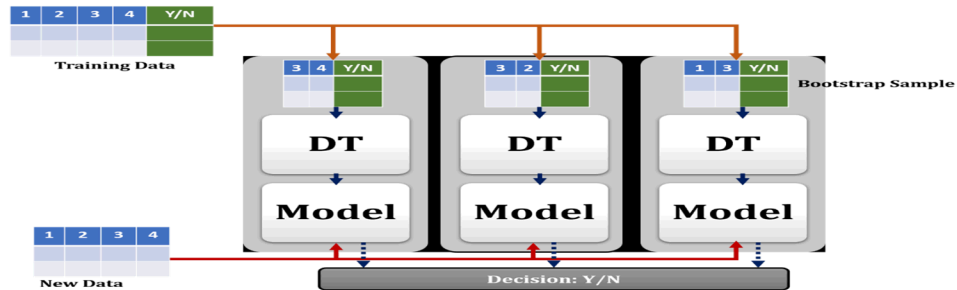

Figure 18: Random Forest (RF) Algorithm with bootstrapping.

It is important to note that the data does not need to be normalized or rescaled. In addition, RF are suitable for high dimensional data since it uses only subset of whole data parallelly, and it can also balance errors in unbalanced data sets.

5.4.2   Support Vector Machine

Support Vector Machine (SVM) [6] is a discriminative classifier that finds a hyperplane in an N-dimensional space (N- the number of features) that distinctly classifies the data points. To separate two classes there can be many possible hyperplanes but SVM finds a hyperplane of maximum margin having longest distance between data points of both classes. Maximum margin makes future data points to be classified with more confidence. Support vectors are data points closer to the hyperplane and influence the position and orientation of the hyperplane. Figure 19 shows the representation of hyperplane and its components in SVM.
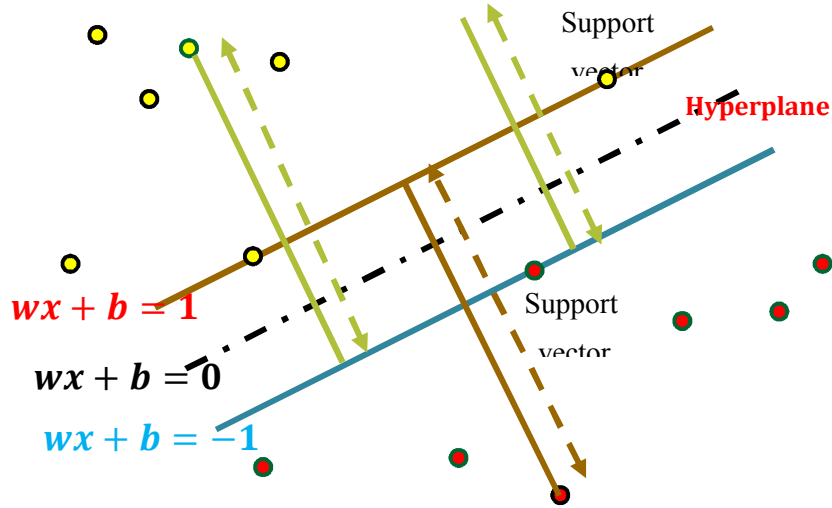

Figure 19: Representation of Hyperplane and its components in SVM.

For the input data points x with labels Y, the mathematical representation of hyperplane is given by

$$wx_i + b = 0 \qquad\qquad\qquad (13)$$

where w is the weight vector, b is the intercept, and $x_i \in x$. Similarly, the equation for the data points falling on either side of the hyperplane is given by

1 If $Y_i = +1 : wx_i + b \geq 1$

2. If $Y_i = -1 : wx_i + b \leq -1$ $\qquad\qquad\qquad (14)$

Using equations in (14), the expression for maximum margin is given by

$$\text{margin} \equiv \underset{x \in D}{argmin}(x) = \underset{x \in D}{argmin} \frac{|x \cdot w + b|}{\sqrt{\Sigma_{i=1}^{d} w_i^2}} \qquad\qquad (15)$$

The above expression is defined for linear classifier where data points are linearly separable. But the data points are not always linearly separable. In such case, the non-linear data points are made linearly separable by mapping to a higher dimension space using kernel function as $K(x_i, x_j) = \phi(x_i)\phi(x_j)$. Gaussian Radial Basis Function (RBF) kernel is the most widely used kernel function defined as

$$K = e^{\frac{\left\|x_i - x_j\right\|^2}{2\sigma^2}} \qquad\qquad\qquad (16)$$

where x_i and x_j are two datapoints and the parameter σ determines the width the of the Gaussian function. On the other hand, for a better classification SVM has tuning parameters such as Regularization and gamma. The regularization controls the SVM optimization to what extent the data points misclassification on training examples needs to be avoided. For a large value of regularization parameter,

the optimization chooses a smaller-margin hyperplane, if that hyperplane can classify all the training parameters correctly. Conversely, for a very small value of regularization parameter, the optimization tries to find a large margin hyperplane even if the hyperplane misclassifies more training examples. Whereas, the gamma parameter defines how far the influence of a single training example reaches. With low gamma value, training examples far from plausible separation plane are considered in calculating the hyperplane. Conversely, a high gamma value picks only training examples closer to the separation plane.

6.   MACHINE LEARNING RESULTS

For the two separated experiment set-ups such as 1. Mass addition and, 2. Mass removal, the machine learning algorithm is separately developed for classifying 1. Baseline v/s Mass addition, and 2. Baseline v/s Mass removal. The split ratio between training and testing examples is 80:20. The feature dimension for mass addition is 529 samples for Baseline, and 427 samples for Mass addition with 357 features. Whereas for mass removal is 529 samples for Baseline data, and 720 samples for Mass removal with 357 features. The feature selection is performed using PCA or PI. After the feature selection, the top 15 features are selected, and the model is retrained. In addition, K-fold cross-validation is also performed for performance analysis. Classification accuracy is considered as the main performance metric in both cases. The classification performance of the RF and SVM algorithms is discussed in the next sub sections.

6.1   Random Forest

Parameter settings such as number of decision trees, number of epochs, and number of features per decision tree are crucial to have maximum uncorrelation between decision trees and to avoid overfitting. The model parameters for RF are described in Table 6.

Table 6. Random Forest parameters.

Parameter	Value
Number of estimators (decision Trees)	10
Maximum features per estimator	60%
Number of epochs	500

Using parameters in Table 4, the RF model is separately trained for mass removal and mass addition classification against baseline data. The classification performance for both cases is shown in Table 7. The confusion matrix associated with classification of mass removal and mass addition is shown in Figure 20 (a. mass removal, b. mass addition). In addition, the performance result for K-fold cross-validation with 10 splits for mass removal and mass addition is shown in Figure 21 (a. mass removal, b. mass addition).

Table 7. RF Training and Testing accuracy: 1. mass removal v/s baseline 2. mass addition v/s baseline

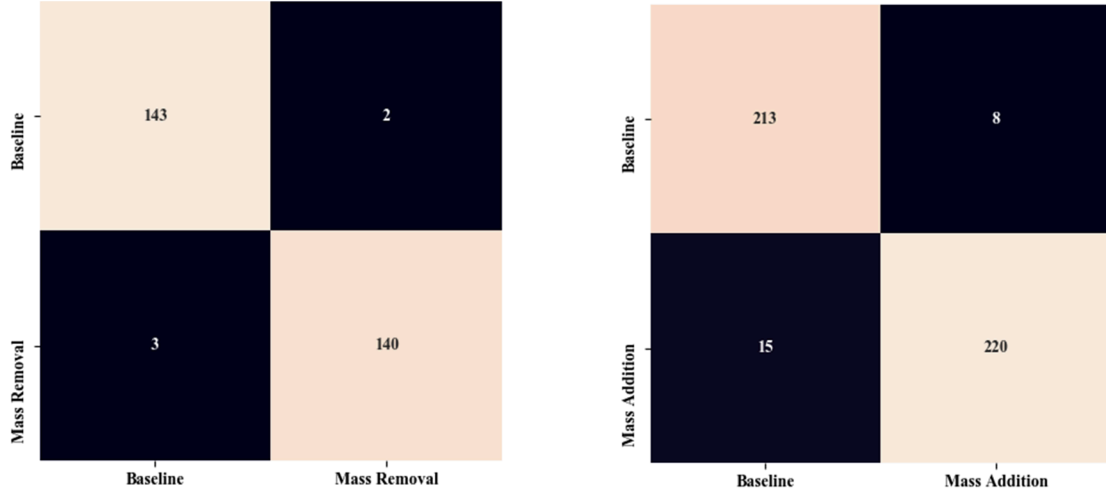| Case	Training accuracy	Testing accuracy
Mass Removal	99.6%	97.3%
Mass Addition	98.8 %	95.2%

Figure 20. RF Confusion matrix for mass removal v/s baseline (left) and mass addition v/s baseline (right).
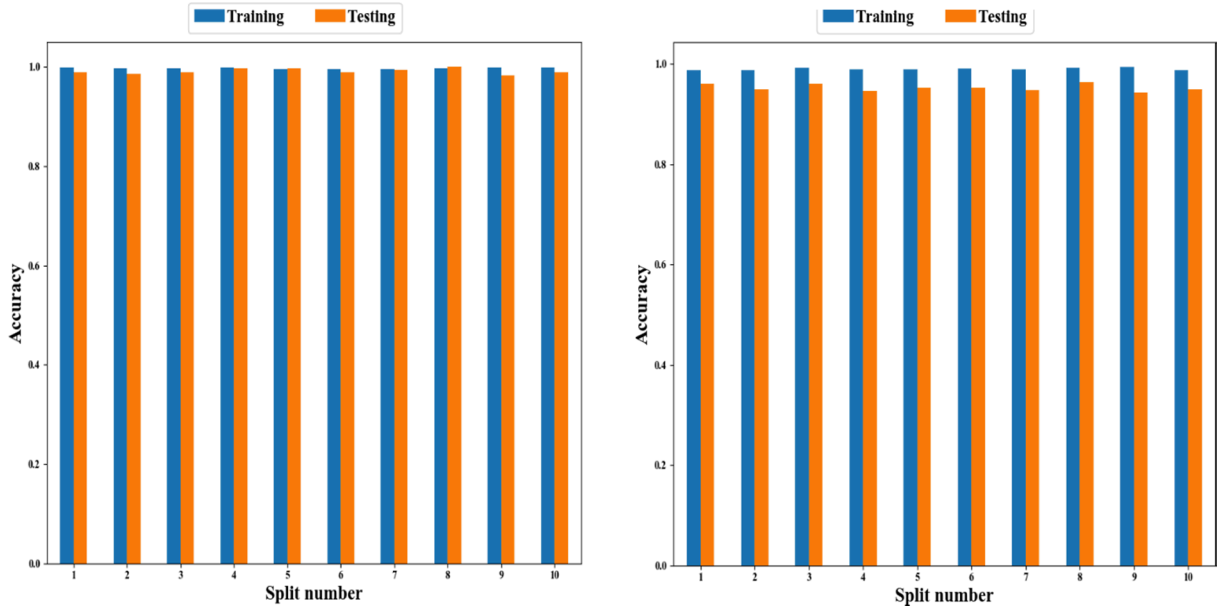

Figure 21. RF Training and Testing accuracy of K-fold cross-validation with 10 folds. left: mass removal, and right: mass addition.

6.2 Support Vector Machine

Using grid search, the gamma and Regularization parameter are selected. It is found that, the estimated parameters tend to overfit. Hence, the Regularization parameter is empirically decreased from the estimated value to compensate for overfitting. SVM is tested with PI and PCA based feature selection for both cases of classifying mass removal and mass addition. Like RF, SVM also trained with top 15 important features. Similarly, for PCA the feature dimension was reduced to 15. Performance results of SVM classification for mass removal and mass addition are shown in Table 8. The confusion matrix associated with classification of mass removal and mass addition is shown in Figure 22(a. mass removal, b. mass addition). In addition, the performance result for K-fold validation with 10 splits for mass

removal and mass addition is shown in Figure 23. It is noted that, the performance of SVM+PCA algorithm is poor compared to SVM+PI in both cases. In addition, like in RF, the algorithm is not very consistent in making good predictions in the case of mass addition.

Table 8. SVM Training and Testing accuracy: 1. mass removal v/s baseline 2. mass addition v/s baseline.

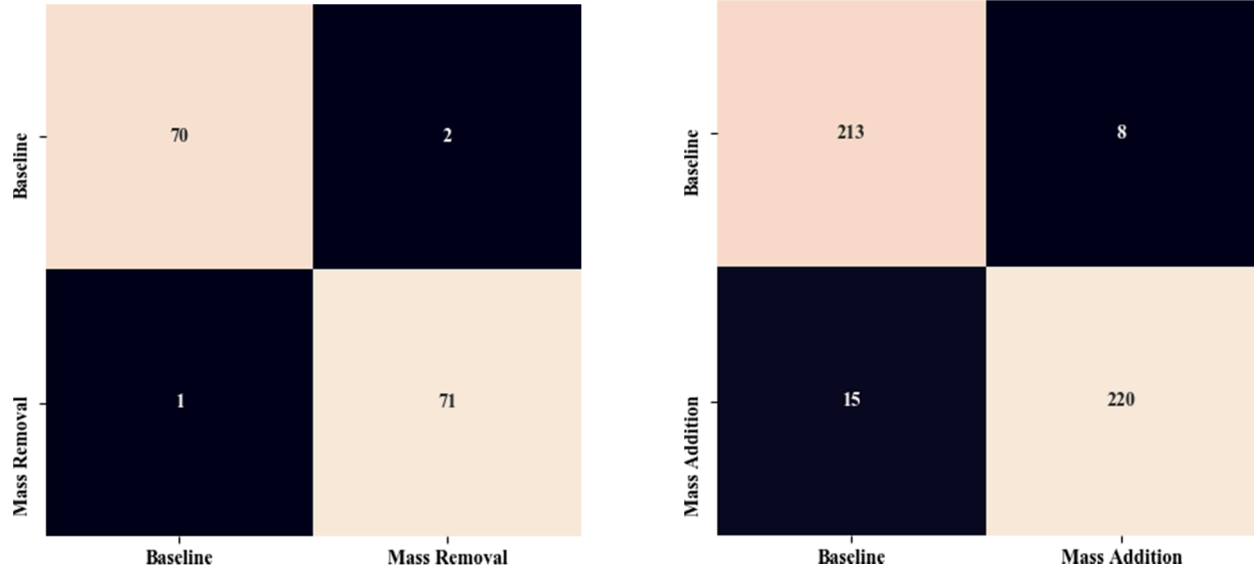| Case	Training accuracy	Testing accuracy
Mass Removal (SVM+ PI)	97.45	97.91
Mass Removal (SVM+PCA)	87.57	84.72
Mass Addition (SVM+PI)	92.8	91.57
Mass Addition (SVM+PCA)	70.3	68.14

Figure 22. SVM Confusion matrix for mass removal v/s baseline (left) and mass addition v/s baseline (right).

Overall, the PI based feature selection performs a better classification compared to PCA possibly due to the fact that the PI method considers both input features and class labels, whereas PCA does not consider class labels. Not considering class labels while reducing dimension, the features which are highly influential on class label might be dropped leading to reduced feature vectors which are less influential on class labels. As far as the different algorithms are considered both RF and SVM have almost a similar performance. SVM can be used as the final algorithm due to its simple implementation and control over hyperparameters tuning. In addition, for SVM, the regularization parameter is chosen less than the value predicted from the grid search to make sure that the performance gets generalized better.

Besides, the performance on classifying mass addition is not consistent, and in some cases, it showed overfitting. In this case, tuning hyperparameters need to be re-assessed to improve the performance without overfitting.
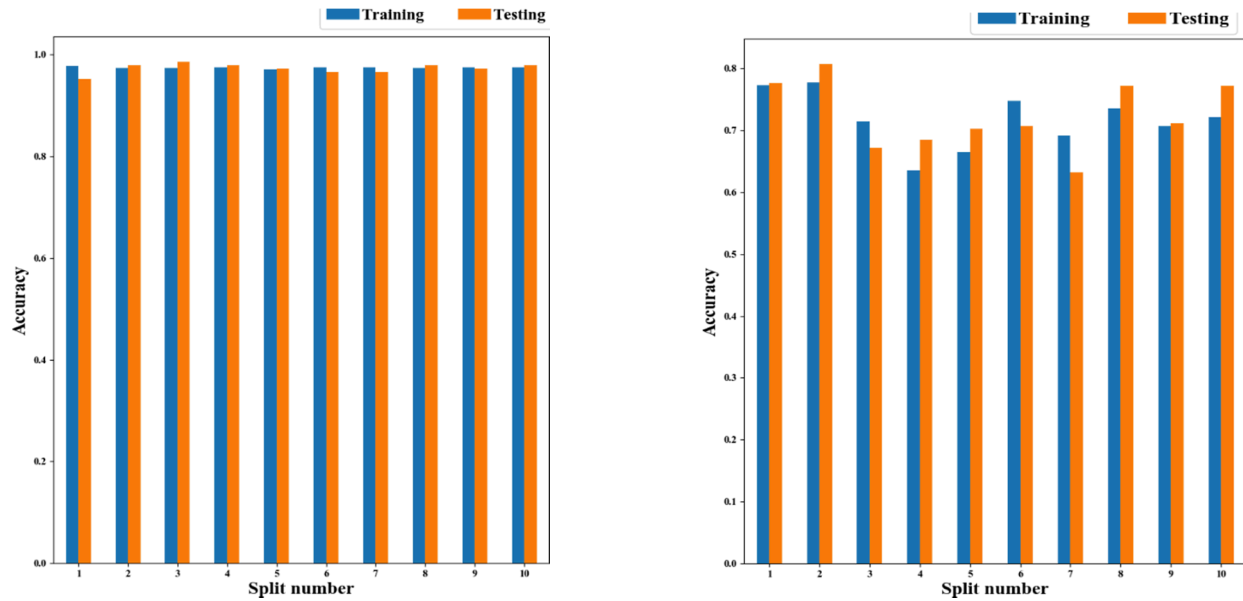
Figure 23. SVM Training and Testing accuracy of K-fold cross-validation with 10 folds. left: mass removal, and right: mass addition.

7. SUMMARY AND PATH FORWARD

Accelerometer data is converted to IMFs using the iterative EMD process, the phase features are extracted, and statistical tests are performed to make prediction on baseline and not baseline classes. The performance accuracy was found around 82 to 85%. Proceeding further, summary information from each IMFs is extracted along with total phase. Phase slope information is used to construct a feature vector. These feature vectors for X, Y, and Z directions of each sensor are extracted. Using the PI and PCA feature selection algorithms, the total dimension of the feature vector is reduced to 15 and trained with RF and SVM classifiers for the cases 1. Mass removal v/s baseline, and 2. Mass addition v/s baseline. Results showed that PI based feature selection outperforms PCA based feature selection. Classification performance of both algorithms for mass removal v/s baseline case has a better performance as opposed to mass addition v/s baseline case.

As a path forward, the work classifying mass removal v/s baseline will be continued to predict early stage of corrosion in secondary pipe structures. Statistical significance of IMFs, more complex statistical models, and integration into machine learning may provide more consistent, interpretable results. Level of corrosion will be considered in future work. Discretion and quantification of unexplained variability will be addressed.

8. REFERENCES

[1] Shen, H. &. (2014). *Hilbert-Huang transform and its applications*. Hackensack, NJ: World Scientific Publishing Co. Pte. Ltd.

[2] Lopez-Ramirez, M., Ledesma-Carrillo, L., Cabal-Yepez, E., Rodriguez-Donate, C., Miranda-Vidales, H., & Garcia-Perez, A. (2016). EMD-based feature extraction for power quality disturbance classification using moments. *Energies*, *9*(7), 565.

[3] Jolliffe, Ian. *Principal component analysis*. Springer Berlin Heidelberg, 2011.

[4] Altmann, A., Toloşi, L., Sander, O., & Lengauer, T. (2010). Permutation importance: a corrected feature importance measure. *Bioinformatics*, *26*(10), 1340-1347.

[5] Liaw, Andy, and Matthew Wiener. "Classification and regression by randomForest." *R news* 2, no. 3 (2002): 18-22.

[6] Suykens, Johan AK, and Joos Vandewalle. "Least squares support vector machine classifiers." *Neural processing letters* 9, no. 3 (1999): 293-300.

[7] Ignasiak, A., Gomolla, N., Kruczek, P., Wylomanska, A., & Zimro, R. (2017). "Long term vibration data analysis from wind turbine – statistical vs energy-based features." *JVE International Ltd. Vibroengineering PROCEDIA,* Vol. 13. Retrieved from *https://doi.org/10.21595/vp.2017.19081,* May 2019.