POWER PLANT VIBRATION MONITORING USING WAVELET FEATURE EXTRACTION AND FUNCTIONAL DESIGN OF EXPERIMENTS

A Dissertation presented to

the Faculty of the Graduate School

at the University of Missouri-Columbia

In Partial Fulfillment

of the Requirements for the Degree

Doctor of Philosophy, Industrial Engineering

---

by

BENJAMIN N. OGUEJIOFOR

Dr. Kangwon Seo, Dissertation Supervisor

MAY 2024

The undersigned appointed by the Dean of the Graduate School, have examined the dissertation entitled:

POWER PLANT VIBRATION MONITORING USING WAVELET FEATURE EXTRACTION AND FUNCTIONAL DESIGN OF EXPERIMENTS

presented by Benjamin N. Oguejiofor,

a candidate for the degree of Doctor of Philosophy,

and hereby certify that, in their opinion, it is worthy of acceptance.

_____

Dr. Kangwon Seo

_____

Dr. Suchi Rajendran

_____

Dr. Luis Occena

_____

Dr. Suhwon Lee

DEDICATION

This dissertation is dedicated to my late parents, Surveyor Joseph Nwakile Oguejiofor and Mrs. Christiana Nwakego Oguejiofor, for being an incredible source of inspiration.

ACKNOWLEDGEMENTS

I wish to express my immense gratitude and appreciation to my professor, mentor, and research advisor, Dr. Kangwon Seo. Your invaluable guidance, foresight, encouragement, passion for knowledge, and dedication greatly inspired me to push the boundaries of my intellectual capabilities to achieve academic excellence.

To Miriam my wife, sweetheart, words alone cannot express my profound gratitude and appreciation for your strong encouragement, advice, support, and the sacrifices you made that enabled me focus and dedicate time in completing this academic journey. I will also extend my appreciation to our children, Tobenna, Chidalu and Soluchi for their understanding and tolerance during my doctoral studies. To my late father-in-law, Engr. Anthony Belonwu, I appreciate all your fatherly advice, zeal, and encouragement especially in the pursuit of higher studies.

My doctoral dissertation committee members, Dr. Suchi Rajendran, Dr. Luis Occena, and Dr. Suhwon Lee; I am thankful for your time and dedication in reviewing my work and providing constructive feedback.

I would also like to thank Prof. Anthony Okafor (Missouri S&T), Curtis Wood (Ameren Missouri), and Dr. Asitha Edirisinghe, PhD (Discover Financial Services) for your words of encouragement. Lastly, the Graduate Fellowship and the Teaching Assistantship provided from the Industrial and Systems Engineering Department are gratefully acknowledged.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF ILLUSTRATIONS

ABSTRACT

In a nuclear plant power generation, analysis of vibration signal constitutes an integral part of predictive maintenance for rotating equipment such as pumps, motors, turbine generators, etc. Vibration signals are continuously monitored via sensors and thresholds for alarms maybe set up to identify equipment malfunction. Improved methods for decomposition and analysis of power plant vibration signals using wavelet feature extraction and functional design of experiment (FDOE) have not been sufficiently investigated for applicability in analyzing these signals for better detection of equipment faults. Chapter 1 introduces the general concepts and methods to be applied in our research study. In Chapter 2, we present the application of discrete wavelet transform (DWT) to decompose a reactor coolant pump vibration signal into frequency sub-bands and the generation of a number of features from some statistics. A principal component analysis (PCA) is used to reduce the large set of variables into a few principal components which can be applied in future monitoring of normal vibration signals. From the insights gained using PCA, Chapter 3 studies the linear discriminant analysis (LDA) to simulated vibration signals, to distinguish between normal and abnormal signals. In Chapter 4, we apply the functional principal component analysis (FPCA) in characterizing the vibration signals generated under several different levels of an environmental factor, a flow rate, associated with a condensate pump. An FDOE is applied with the target vibration curve and used to obtain an optimal flow rate. The obtained flow rate was found

comparable to the theoretical pump curve best efficiency point (BEP) and recommended for use for optimal pump performance and reliability. In Chapter 5, we perform an extensive review of literature on FDOE and its applications. A standard FDOE framework was shown and demonstrated the five basic steps to be applied when using JMP Pro 17. Chapter 6 provided overall conclusions and suggestions for future work.

Chapter 1

Introduction

In industrial environment, predictive maintenance (PdM) has proven to be the most efficient maintenance technique. PdM is based on the analysis of data collected through monitoring or inspections. The data are collected from equipment to determine the health status, define maintenance strategy, or detect faults (Villarroel et al. 2019). Vibration analysis can be an effective tool for diagnosing faults of looseness, eccentricity, imbalance, misalignment, defective bearings, damaged gears, etc. (Althubaiti et al. 2021, Tiboni et al. 2022). As a result, this technique has emerged as a powerful and well established PdM technique for rotating equipment (Ranjan et al. 2020). Vibration signals in rotating machinery are non-stationary which complicates their analysis due to changing time-frequency characteristics and could complicate accurate fault diagnosis. Overcoming these challenges requires advanced signal processing, feature extraction and fault diagnosis algorithms capable of handling nonlinear dynamics and extracting relevant information from complex vibration signals (Shaheryar et al. 2017).

Wavelet transform is a linear transformation in which a time signal is decomposed into wavelets, i.e. local functions of time endowed with a predetermined frequency content (Mohd Ghazali et al. 2021). By decomposing a non-stationary signal into its individual frequency components, wavelet transform can reveal time-varying features and identify transient events that may be missed by conventional Fourier transform-based methods. The advantages of using wavelet transform for vibration analysis in rotating machinery includes

the ability to detect local changes in vibration signals and improved time resolution (Lakis 2007).

The central idea of a principal component analysis (PCA) is to reduce the dimensionality of a dataset consisting of large number of interrelated variables while retaining the variation as much as possible present in the dataset (Jolliffe 2002). PCA simplifies the complexity in high dimensional data while retaining trends and patterns. It does this by transforming the data into fewer dimensions which acts as summaries for the features. PCA finds the principal components (vectors), which maximize the variance of the data along their direction. Each principal component explains a portion of the total variance in the data. The functional principal component analysis (FPCA) is similar to the PCA. The difference being that FPCA deals with functions whereas PCA deals with vectors. The principal components from the FPCA are also functions or curves and are called the functional principal components (FPCs). The linear discriminant analysis (LDA) also relies on linear transformation to maximize variance in a lower dimension. However, unlike the PCA, LDA finds the linear discriminants to maximize the variance between the different categories, while minimizing the variance within the class.

Functional data analysis (FDA) is about modelling data profiles with functions, which can be applied to a wide variety of shapes of profiles. FDA may apply B-splines, P-splines, Fourier bases, or wavelets as the basis functions in the smoothing step. The FPCA dimension reduction step simplifies the analysis by reducing profile data to a small number of scalar values called FPC scores. The FPC scores are the most efficient linear low dimensional decomposition of the observed profile-to-profile variation (Kenett and Gotwalt 2023). Functional data or curves can be found everywhere. It occurs whenever

we measure something over a continuum, like measuring temperature over time, pressure gradient across space, or radiation absorption across wavelength. Sometimes the shapes of the curves we measure are important to us, and we want to know how curve shape is affected by certain factors. This can be achieved using a functional design of experiments (FDOE) (Metusalem 2020). A combination of these techniques was used in this dissertation for analyzing vibration signals obtained from a power plant application.

Chapter 2

PCA-based Monitoring of Power Plant Vibration Signal by Discrete Wavelet Decomposition Features

2.1 Introduction

In the past few years, lots of investments have been made in continuous online monitoring of power plant equipment performance data. This enables timely monitoring of equipment operating condition to ensure improved system reliability. However, in a power plant application, a combination of wavelet analysis and principal component analysis (PCA) of vibration signals have not been adequately investigated for applicability in predictive maintenance. The investigation from this study could help understand how small changes in a continuously monitored vibration signal data can be better captured and used for predictive maintenance in detecting potential equipment fault, and thereby taking prompt actions to prevent catastrophic failure.

Wavelet analysis is the research hotspot in the field of mechanical fault diagnosis for many years due to its time-frequency resolution and excellent capability of detecting transients (Yan et al. 2014). Unlike Fourier transform, in which only the trigonometric functions are the basis for decomposing a signal, possible selections of wavelets for decomposition basis are numerous in wavelet analysis (Wang 1996). Discrete wavelet transform provides a time-scale information of a signal enabling extraction of features that vary in time. This property makes the wavelets an ideal tool for analyzing a signal of a transient or non-stationary nature (Nizwan et al. 2013).

The central idea of a PCA is to reduce the dimensionality of a dataset consisting of large number of interrelated variables while retaining as much as possible the variation present in the dataset (Jolliffe 2002). PCA simplifies the complexity in high dimensional data while retaining trends and patterns. It does this by transforming the data into fewer dimensions which acts as summaries for the features. PCA is an unsupervised learning method and reduces data by geometrically projecting them into lower dimensions called principal components (PCs), with the goal of finding the best summary of data using a limited number of PCs. The first principal component (PC) is chosen to minimize the total distance between the data and their projection onto the PC. By minimizing the distance, the variance of the projected points is also maximized. The second and subsequent PCs are selected similarly with the additional requirement that they should be uncorrelated with all previous PCs. For example, projection into PC2 is uncorrelated with projection onto PC1 and etc. The PCs are geometrically orthogonal. This requirement of no correlation means that the maximum number of PCs is either the number of samples or the number of features, whichever is smaller. The PC selection process has the effect of maximizing the correlation between the data and their projection and is equivalent to carrying out multiple linear regression on the projected data against each variable of the original data (Lever et al. 2017).

Wavelets decomposition methods, specifically the DWT have been used to analyze vibration signals in previous studies. However, the extraction of certain features from these decomposed signals have not been adequately investigated to determine which decomposed coefficients can be better utilized to further process the data. This study investigated the application of DWT decomposition of vibration signal data from a reactor

coolant pump used in a power plant application. Certain features were extracted from the decomposed coefficients. The PCA was used to determine from the features, which decomposed coefficients have more variability associated with it in a reduced PC variables from the large original PC variables.

2.2 Review of Previous Works

Wavelets decomposition and PCA methods have been investigated for analysing vibration signals in rotating machinery. Nizwan et al. (2013) applied discrete wavelet transform to decompose vibration signals from defective bearings. The Root Mean Square (RMS) for every decomposition level was calculated to detect the defect features in the vibration signals. Jayakumar and Thangavel (2015) transformed vibration and current signals from a motor ball bearing into power spectral density using wavelet transform coefficients. The steady state rotor frequency was used to introduce a new frequency pattern for fault diagnosis. Liu Zepeng et al. (2020) proposed an empirical wavelet thresholding to remove heavy noise and extract weak fault signals. This method was applied in diagnosing low-speed blade bearing faults. Bendjama et al. (2015) decomposed vibration signals to obtain multiple data series at different resolutions. The energy distribution was calculated using Parseval's theorem to select the optimal decomposition level for a possible diagnosis. Kumar et al. (2022) explored wavelet transform and empirical mode decomposition to decompose the vibration signals of a rotor-bearing system into a finite number of intrinsic mode functions to analyse fault characteristics. Patel and Giri (2016) applied principal component analysis (PCA) and self-organizing map (SOM) for clustering-based classification of bearing vibration data. Bendjama and

Boucherit (2016) proposed a combination of wavelet transform and PCA for fault frequency identification using Q-statistic for rotating machinery. Elsamanty et al. (2023) proposed a method that applied PCA, and back propagation neural network based on data fusion for vibration and electrical signatures. This method could be applied for diagnosing mechanical faults in rotating machines.

2.3 Power Plant Vibration Signal Data Application

Nuclear power plants rely on cooling systems to ensure safe and continuous operation of the nuclear reactor. During operation, the reactor coolant pump (RCP) function to circulate demineralized light water under pressure through the reactor vessel and loops. The water serving as coolant, moderator, and solvent for boric acid (used for chemical reactivity control), is heated as it passes through the reactor core. The water then flows through the steam generator where it transfers heat to the steam system and then returns to the reactor coolant pump to repeat the cycle. Therefore, the RCP provide forced primary coolant flow to remove and transfer the amount of heat generated in the reactor core.

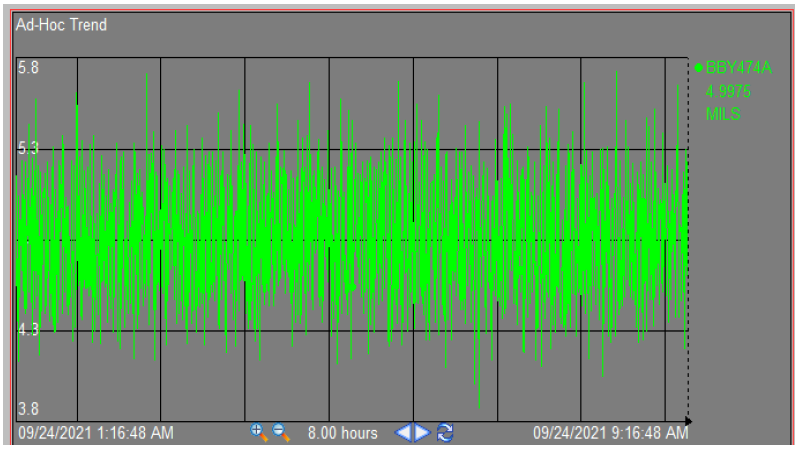

Figure 2.1: Snapshot of PI RCP vibration signal data

2.4 Wavelet Decomposition of Vibration Signals

The vibration signals obtained via PI was downloaded as a csv file and uploaded to python for wavelet decomposition. Using the array reshape function in python, a total of 2,750 vibration data obtained was broken down into 55 separate time windows of vibration signals containing 50 acquired vibration data for each.

The DWT of a function, $x(t)$ relative to real-valued wavelet, $\Psi(t)$ is described by:

$$W_{\Psi}(s,\tau) = \int_{-\infty}^{+\infty} x(t)\Psi_{s,\tau}(t)dt \tag{2.1}$$

$$\Psi_{s,\tau}(t) = \frac{1}{\sqrt{s}}\Psi\left(\frac{t-\tau}{s}\right) \tag{2.2}$$

where s and τ are called scale and translation parameters respectively. $W_{\Psi}(s,\tau)$ denotes the wavelet transform coefficients and Ψ is the fundamental mother wavelet (Haddadi et al. 2014). Because DWT uses scale and position values based on powers of two, the values of s and τ are: $s = 2^j, \tau = k \cdot 2^j$ and$(j,k) \in \mathbb{Z}^2$.

In a DWT, coefficients (weights) associated with the scaling function, called approximation coefficients(cA), capture low frequency information, and constitute the most important part of the signal. The coefficients associated with wavelet function, called detail coefficients (cD), capture high-frequency information. This down sampling process of the signal produces the two sequences cA and cD. The decomposition process can be iterated with successive approximations being decomposed in turn, so that one signal is broken down into many lower resolutions components as shown in Figure 2.2 below.

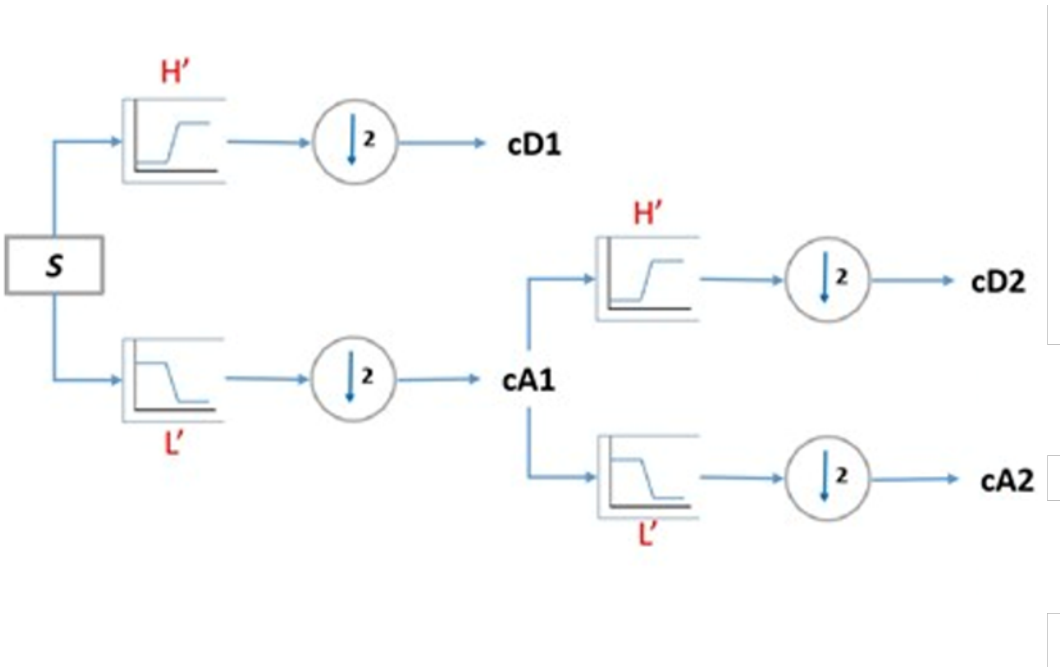

Figure 2.2: DWT decomposition process

Using the DWT decomposition process, each of the 55 vibration signals is decomposed into approximation and detailed coefficients with a possible maximum decomposition level of 2. The debauchee orthogonal wavelet (db6) was used for the DWT decomposition. Based on the level 2 decomposition, each of the vibration signal (S) was decomposed into approximation coefficients (cA2) and detailed coefficients (cD1 and cD2).

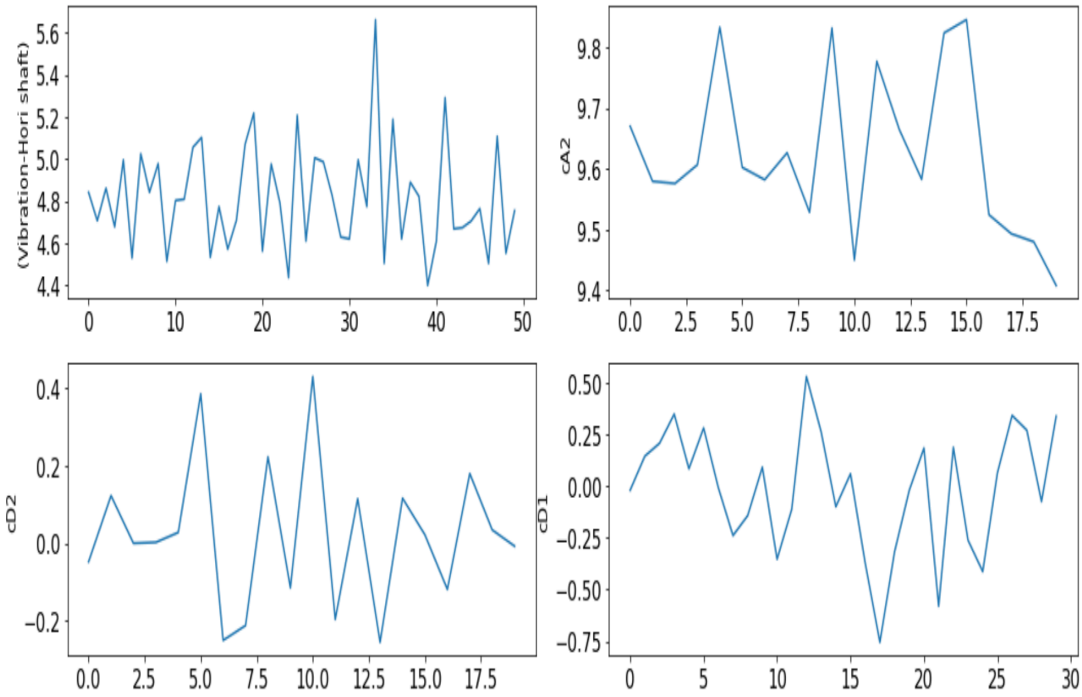

Figure 2.3: Python plots of a vibration signal and the corresponding decomposed approximation (cA2) and detailed (cD2 and cD1) coefficients.

2.5 Features Generation per Sub-Band of Vibration Signals

The DWT was used to split each of the 55 vibration signals into three different sub-bands consisting of the approximation coefficients (cA2) and the detailed coefficients (cD1) and (cD2). These frequency sub-bands were further investigated, and some features were generated. These features consist of some statistics such as the 5^{th}, 25^{th}, 75^{th} and 95^{th} percentiles, mean, median, standard deviation, variance, and root mean square (rms). Additionally, the zero-crossing rate, mean crossing rate and entropy of each signal were equally generated.

- 5^{th} percentile is the value associated with the location within each of the coefficients (cA2, cD1 and cD2) per each decomposed signal such that 5% of the data is below that value.
- 25^{th} percentile is the value associated with the location within each of the coefficients (cA2, cD1 and cD2) per each decomposed signal such that 25% of the data is below that value.
- 75^{th} percentile is the value associated with the location within each of the coefficients (cA2, cD1 and cD2) per each decomposed signal such that 75% of the data is below that value.
- 95^{th} percentile is the value associated with the location within each of the coefficients (cA2, cD1 and cD2) per each decomposed signal such that 95% of the data is below that value.
- Mean is a measure of central tendency used to determine the mid-point or central value for each of the coefficients (cA2, cD1 and cD2) per each decomposed signal.
- Median is also a measure of central tendency used to determine the middle value when the coefficients (cA2, cD1, and cD2) per each decomposed signal are ranked in order (ascending or descending).
- Variance is a measure of dispersion used to determine the spread of the recorded values for each of the coefficients (cA2, cD1 and cD2) per each decomposed signal.
- Standard deviation is the square root of the variance values.

These features were generated using the Python software and the final function returned a set of 12 features for any list values. Therefore, since one vibration signal was decomposed into 3 different sub-bands (cA2, cD1, and cD2) and following generation of the 12 features for each sub-band, a total of $12 \times 3 = 36$ features were generated per each vibration signal. These features were used as an input data for the PCA.

2.6 PCA-Based Feature Monitoring

Based on the 36 variables (features) obtained using the DWT of the vibration coefficients, a PCA was utilized to transform the large set of variables into smaller ones that would still contain information about the dataset. For the PCA application used in this study, we have a vector of random variables $\boldsymbol{y} = (y_1, y_2, \dots, y_{36})^T$, in a 36-dimensional space.

To obtain the principal components, let $(\lambda_1, \boldsymbol{e}_1), (\lambda_2, \boldsymbol{e}_2), \dots, (\lambda_{36}, \boldsymbol{e}_{36})$ represent pairs of eigenvalues and eigenvectors of $\boldsymbol{S} = Cov(\boldsymbol{y})$, the covariance matrix of $\boldsymbol{y}$, where $\lambda_1 > \lambda_2 > \cdots > \lambda_{36}$. $Z_1 = \boldsymbol{e}_1'\boldsymbol{y}$ has the maximum variance among all linear combinations of $\boldsymbol{y}$. $Z_2 = \boldsymbol{e}_2'\boldsymbol{y}$ has the maximum variance among all linear combinations which are orthogonal to Z_1. $Z_3 = \boldsymbol{e}_3'\boldsymbol{y}$ has the maximum variance among all linear combinations which are orthogonal to Z_1 and Z_2. …and so on.

The variance accounted for by the first principal component is:

$$Var(Z_1) = \boldsymbol{e}_1'\boldsymbol{S}\boldsymbol{e}_1 = \lambda_1\boldsymbol{e}_1'\boldsymbol{e}_1 = \lambda_1 \tag{3}$$

The first eigenvalue λ_1 is exactly the variance accounted for by the first principal component. The proportion of variance captured by the first principal component is

$\lambda_1/\sum_{i=1}^{36}\lambda_i$. The variance accounted for by the second principal component is λ_2. The proportion of variance computed by the first and second principal components is $(\lambda_1 + \lambda_2)/\sum_{i=1}^{36}\lambda_i$. The variance accounted for by the third principal component isλ_3. The proportion of variance computed by the first, second and third principal components is $(\lambda_1 + \lambda_2 + \lambda_3)/\sum_{i=1}^{36}\lambda_i$. The variance accounted for by the fourth principal component is λ_4. The proportion of variance computed by the first, second, third and fourth principal components is $(\lambda_1 + \lambda_2 + \lambda_3 + \lambda_4)/\sum_{i=1}^{36}\lambda_i$ and so on. If this fraction is substantially large, then we have reduced a 36-dimensional problem down to four dimensions with little loss of information. In this study, the correlation between the features (variables) was reviewed in sequence by the 36 features represented by$X0, X1, \ldots, X35$.

Table 2.1: Summary of components importance for the first seven components.

Components Importance							
	Comp. 1	Comp. 2	Comp. 3	Comp. 4	Comp. 5	Comp. 6	Comp. 7
Std. Dev.	3.328	2.788	2.033	1.252	0.944	0.283	0.161
Proportion of Variance	0.434	0.304	0.165	0.061	0.035	0.002	0.001
Cumulative Proportion	0.434	0.736	0.897	0.959	0.994	0.996	0.997

From the summary of importance of components as shown in Table 2.1, the standard deviations are given by $\sqrt{\lambda_1} > \sqrt{\lambda_2} > \cdots > \sqrt{\lambda_{36}}$ where the λ_i's are the eigenvalues of the covariance matrix. It is shown that the principal components (Comp.1 (Z_1), Comp. 2 (Z_2), Comp.3 (Z_3), and Comp. 4 (Z_4) account for approximately 96% of the

variability in the data, which means much of the variability within the original data can be explained by only four derived values.

The Biplot in Figure 2.4 shows the principal component (PC) loadings of the first four principal components (PCs) using arrows to indicate their direction. Because the plot is 2 dimensional, only the directions for principal components 1 and 2 (Comp. 1 and Comp. 2) can be easily seen. The biplot show that variability was observed in the variables (features) $X2, X13, X14, X25$ and$X26$.

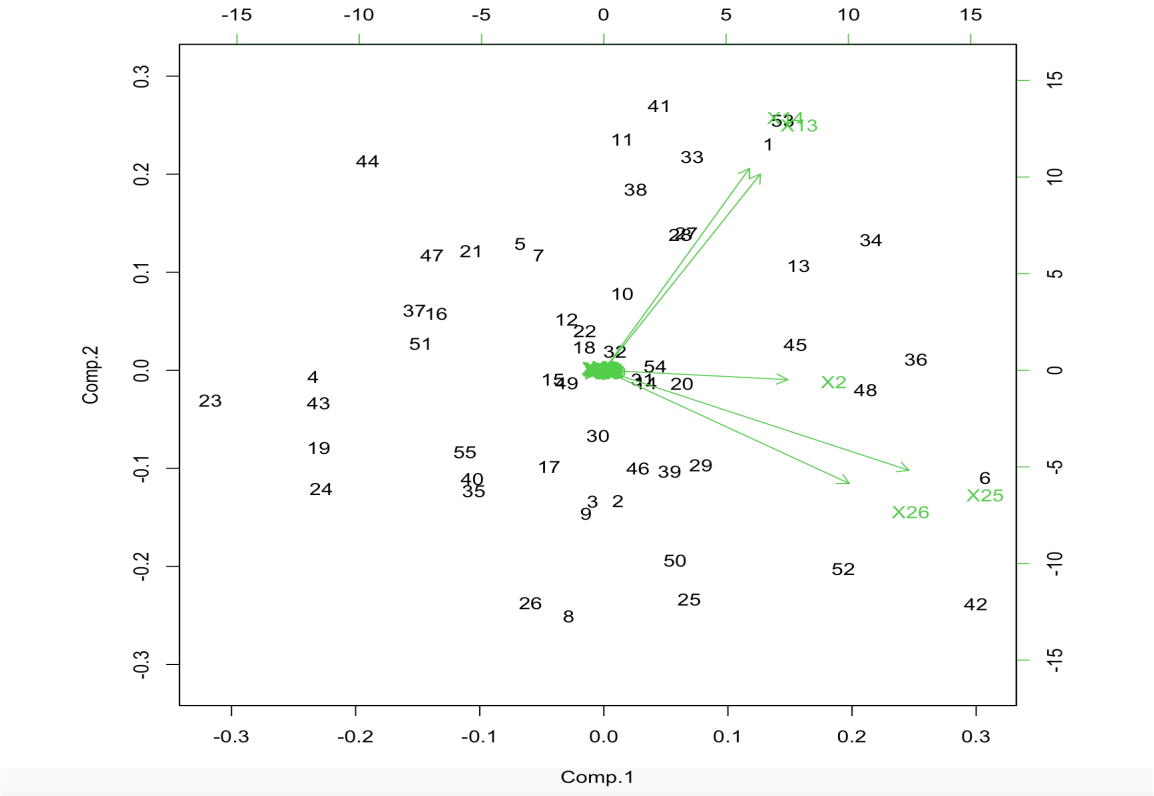

Figure 2.4: Biplot of important principal components

Additionally, because the PCs show that much variability was observed in $X2$, $X13$, $X14$, $X25$and $X26$ variables, a scatter plot of these features was plotted in Figure 2.5 to observe some correlation.

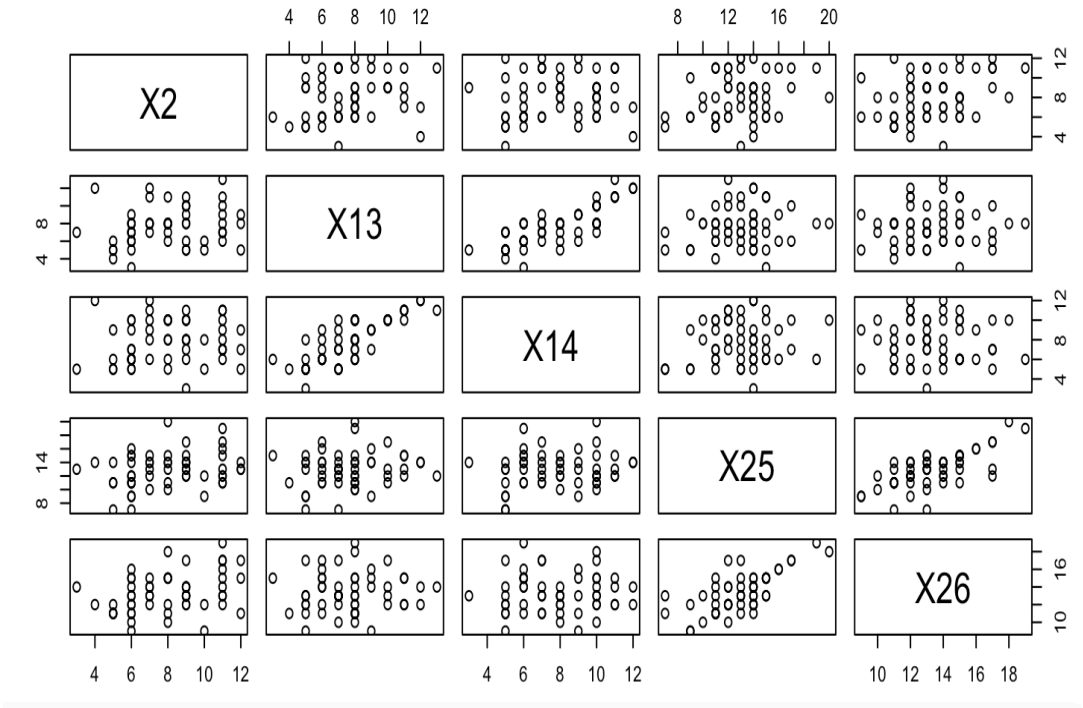

Figure 2.5: Scatter plot of important features

2.7 Results, Discussion and Conclusion

The following observations were made from the results:

- $X2$ is the feature corresponding to the 75th percentile of cA2 approximation coefficients.
- $X13$ and $X14$ are the features corresponding to the 25th and 75th percentiles of cD2 detailed coefficients.
- $X25$ and $X26$ are the features corresponding to the 25th and 75th percentiles of the cD1 detailed coefficients.

From the PCA results, it can be inferred that only 4 PCs explain most variability (about 96%) of the features. Therefore, we can monitor these small number of PCs variables instead of the original 36 PCs. These 4 PCs could provide the behavior of normal vibration signals and how to detect abnormal signals in the future. This is proposed for a future work.

Chapter 3

Simulation Study for Vibration Monitoring using Linear Discriminant Analysis, Discrete Wavelet Decomposition and Feature Extraction

3.1 Introduction

In Chapter 2, discrete wavelet transformation (DWT) was applied to a reactor coolant pump vibration signal. Certain statistical features were extracted from the decomposed wavelet coefficients. From the coefficients, principal component analysis (PCA) was used as a dimension reduction tool and determined that only 4 principal components (PCs) out of the original 36 variables explained the most variability (about 96%) existed in the feature data. It was proposed that these 4 PCs could provide the behavior of normal vibration signals and how to detect abnormal signals in the future. The challenge from Chapter 2 study was the lack of abnormal vibration data from the operating pump that could be contrast to the clustering of the normal vibration data.

In this study, simulation data are used to validate the above-mentioned procedure. Using simulation, we can obtain artificial vibration measurements with class labels; normal and abnormal. Instead of PCA, now we consider a different dimension reduction technique that can be used with the labeled data. Linear discrimination analysis (LDA) is an approach used in supervised machine learning to solve multi-class classification problems. LDA is based on Fischer's linear discriminant, a statistical method developed by Sir Ronald Fisher in 1936 (Chang et al. 2022). Fisher's method aims to identify a linear combination of

features that discriminates between two or more classes of labeled objects or events. LDA algorithms make predictions by using Bayes' theorem to calculate the probability of whether an input data instance will belong to a particular group or class. LDA works by identifying a linear combination of features that separates or characterizes two or more classes of objects or events, enables dimension reduction. LDA does this by projecting with more than two dimensions into one dimension so that it can be more easily classified (IBM November 2023). LDA distinguishes the known classes from each other by associating a unique label (or output value) with each class. The observations are then described as labeled observations. When there are two classes, i.e., $K = 2$, we need only one classifier, and when they are more than two classes, we need at least two and at most $K - 1$ classifiers to differentiate between the classes (Izenman 2008). LDA maximizes the distance between class means while minimizing the spread within each class. By projecting data points onto this discriminative axis, LDA reduces dimensionality and helps classifiers make more accurate predictions.

The motivation for the present study using LDA is based on the idea of finding a cluster of normal vibration data using some feature extraction and dimension reduction technique, then if some newly observed data points are plotted outside this cluster, we can detect that and use it to characterize abnormal data. From the insights gained from the previous study using PCA, this study simulated the normal vibration data and varying forms of abnormal vibration data using R software. DWT is used to decompose the signals with some statistical features extracted. LDA is then applied to distinguish between the classes for normal and abnormal vibration data with results presented.

3.2 Review of Previous Works

LDA have been investigated and applied for detecting rotating equipment faults. A lot of research has also been conducted on DWT feature extraction, dimension reduction and classification. The objective of this study is to validate if the application of LDA, DWT and feature extraction could distinguish normal and abnormal vibration data, especially when plotted in a low dimensional space.

Harmouche et al. (2014) applied LDA to improve on the classification of faults on ball bearings based on information from the spectral features. The experimental results show the LDA extracts from the spectral features new variables having more discriminatory power than the principal components from the PCA used in their previous study. The accuracy of the discrimination or class separability was determined using Bhattacharyya distance, a well-known measure of class separability. Bajaj et al. (2022) presented a Bayesian optimized discriminant analysis model to classify and monitor tool condition into three user defined classes. The study concluded the feasibility of discriminant analysis as a machine learning technique in the field of tool condition monitoring. Ahmed et al. (2021) transformed a centrifugal pump vibration signature using Walsh transform, and then applied a cosine linear discriminant analysis and K-nearest neighbor (KNN) to classify and detect fault signals. Tong et al. (2022) proposed a discriminant analysis using multi-view learning (DAML) and KNN to classify and diagnose bearing faults in rotating machinery. Shi et al. (2021) applied LDA based on adaptive divergence matrix (ALDA) to extract the status features (time and frequency domains) of rolling bearings. The ALDA algorithm was used to extract the status features of the bearings under different conditions. A support vector machine classifier was used to

verify the effectiveness of the proposed method. Jiang et al. (2023) proposed a fault diagnosis method based on a reference manifold (RM) learning. The vibration signals of a road header (equipment used for coal mining) were extracted and analyzed by time domain and wavelet packet energy. LDA was used to segment different characteristic parameters. Zamudio-Ramirez et al. (2022) proposed a diagnosis methodology based on the analysis of stray flux signals for detecting uniform wear in the gear teeth. The proposed methodology is based on the processing of the stray flux signals through feature calculation and extraction stages. The PCA and LDA techniques were used for dimension reduction of the estimated domain-based features. Ullah et al. (2024) introduced an innovative approach for fault diagnosis of a multistage centrifugal pump (MCP) using explanatory ratio (ER) LDA. To address the challenge of background noise and inference in vibration signals, the study utilized a fault-sensitive frequency band (FSFB) to extract some features. Chen et al. (2023) proposed a Gaussian assumptions-free interpretable LDA as learning model to physically locate informative frequency bands and fault characteristic frequencies for machine condition monitoring. Yan et al. (2022) theoretically investigated how LDA can locate informative frequency bands for machine health monitoring. Additionally, with the help of the spectral evolution mechanism of a bearing life cycle, a new proposition was made to theoretically illustrate that an optimal discriminant direction of LDA integrated with normal and abnormal spectral amplitudes can reveal informative frequency bands for machine health monitoring. Jaen-Cuellar et al. (2023) proposed a methodology based on statistical features and genetic algorithms to determine features for detecting gradual wear of gearbox. In addition, the methodology also makes use of LDA to generate a

bidimensional representation of the system conditions that are fed to a neural network for performing classification of the condition.

3.3 Simulation of Vibration Data

The data used in this study were simulated such that the created curves show similar shapes and visual characteristics with real vibration data that were observed in a reactor coolant pump of a nuclear power plant (Oguejiofor and Seo 2023). This study is an improvement from the previous work using PCA where abnormal vibration data could not be easily obtained. An example of the real vibration curve is shown on the left panel of Fig. 3.1. We first defined 128 equidistant discrete measurement points t_1, t_2, …, t_{128} in the interval [0, 100], and the signals were generated using the R software for each t_i according to the following model.

$$y_1(t_i) = A_1 \sin(\omega_1 t_i + \varphi_1) \tag{3.1}$$

$$y_2(t_i) = A_2 \cos(\omega_2 t_i + \varphi_2) \tag{3.2}$$

$$y(t_i) = y_1(t_i) + y_2(t_i) + \varepsilon(t_i) \tag{3.3}$$

where A_1 and A_2 are the amplitudes, ω_1 and ω_2 are the frequencies, φ_1 and φ_2 are the phases of the sinusoidal curves defined in Eq. 3.1 and Eq. 3.2, respectively. The sum of those two curves with some additional noise $\varepsilon(t_i)$ makes the final simulated vibration measurements. These parameters are randomly generated according to the following probability distributions.

$$A_1 \sim N(\mu_{A_1} = 5.3, \sigma_{A_1} = 0.3) \quad (3.4)$$

$$A_2 \sim N(\mu_{A_2} = 2.5, \sigma_{A_2} = 0.3) \quad (3.5)$$

$$\omega_1 \sim N(\mu_{\omega_1} = 0.35, \sigma_{\omega_1} = 0.02) \quad (3.6)$$

$$\omega_2 \sim N(\mu_{\omega_2} = 0.21, \sigma_{\omega_2} = 0.02) \quad (3.7)$$

$$\varphi_1 \sim N(\mu_{\varphi_1} = 0, \sigma_{\varphi_1} = 1) \quad (3.8)$$

$$\varphi_2 \sim N(\mu_{\varphi_2} = 0, \sigma_{\varphi_2} = 1) \quad (3.9)$$

$$\varepsilon(t_i) \sim ind.N(\mu_\varepsilon = 0, \sigma_\varepsilon = 0.25), i = 1,2, \ldots 128 \quad (3.10)$$

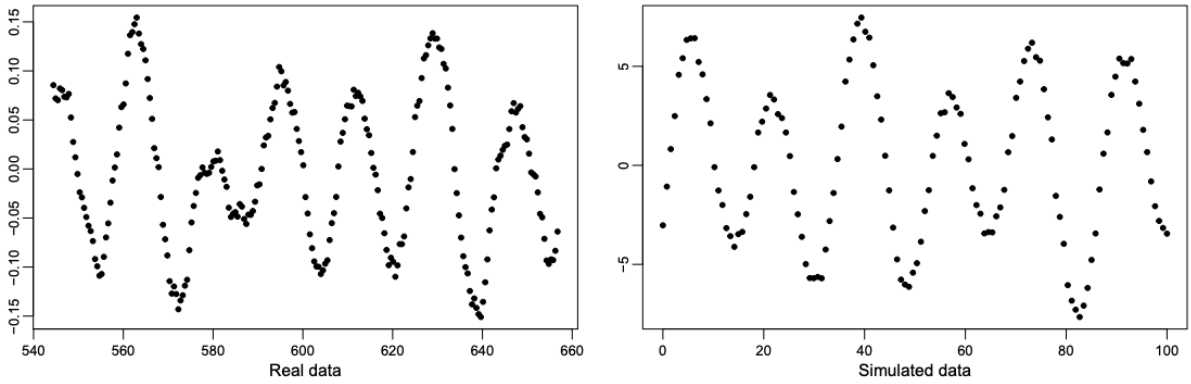

Figure 3.1: Examples of a real vibration curve (left) and a simulated vibration curve (right)

The right panel of Figure 3.1 depicts an example of simulated vibration curves, showing a similar shape with the real one in the left panel. The above-mentioned simulation data are treated as normal (N) vibration curves. For a classification study, data on the abnormal status are also required. We define three different classes of abnormality, i.e., deviations from the normal status, namely, high amplitude (HA), high frequency (HF), and high noise (HN). The HA curves were generated by replacing the parameters $\mu_{A_1} = 5.3$

with $\mu_{A_1} = 7.2$ and $\mu_{A_2} = 2.5$ with $\mu_{A_2} = 5.0$. The HF curves were generated by replacing $\mu_{\omega_1} = 0.35$ with $\mu_{\omega_1} = 0.7$ and $\mu_{\omega_2} = 0.21$ with $\mu_{\omega_2} = 0.5$. For the HN curves, $\sigma_\varepsilon = 0.25$ was replaced with $\sigma_\varepsilon = 1$. For each class of abnormality, all the other parameters except those replaced remained as the same as those of the normal data. The simulated dataset consists of 400 curves in total: 100 normal, 100 HA, 100 HF, and 100 HN curves. Figure 3.2 shows five randomly selected examples of normal and three classes of abnormal simulated vibration curves from the dataset.

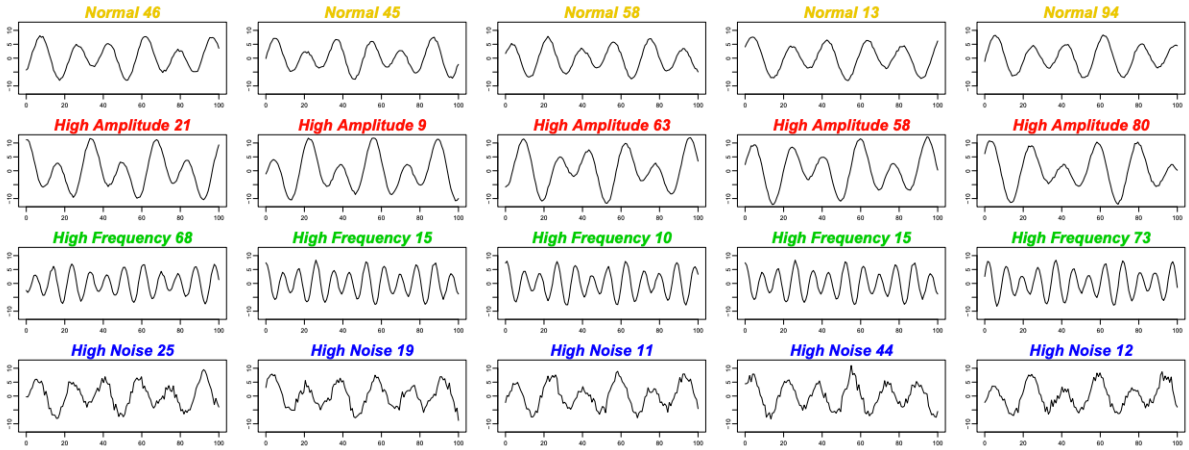

Figure 3.2: Examples of simulated vibration curves from each class.

3.4 Wavelet Decomposition of Simulated Signals

The process to decompose a signal into approximation and detail coefficients at different levels is called wavelet decomposition. The original signal, S passes through two complimentary filters, low pass (L) filter and high pass (H) filter, and is decomposed into two coefficients, cA and cD. cA is the lower frequency content often called the approximation, is the most important part of a signal, the identity of a signal. cD is the high-frequency content, usually called details, and imparts flavor or nuance. Similarly to

the approach in Chapter 2, DWT is applied to decompose the simulated signals. The decomposition process is iterated with successive approximations being decomposed in turn, such that one signal is broken down into many lower resolution components, called the wavelet decomposition tree. This down sampling involved a six-level process to decompose the signals into approximation (cA2) and detail coefficients (cD2, cD3, cD4, cD5, and cD6). The DWT decomposition process is shown in Figure 3.3.

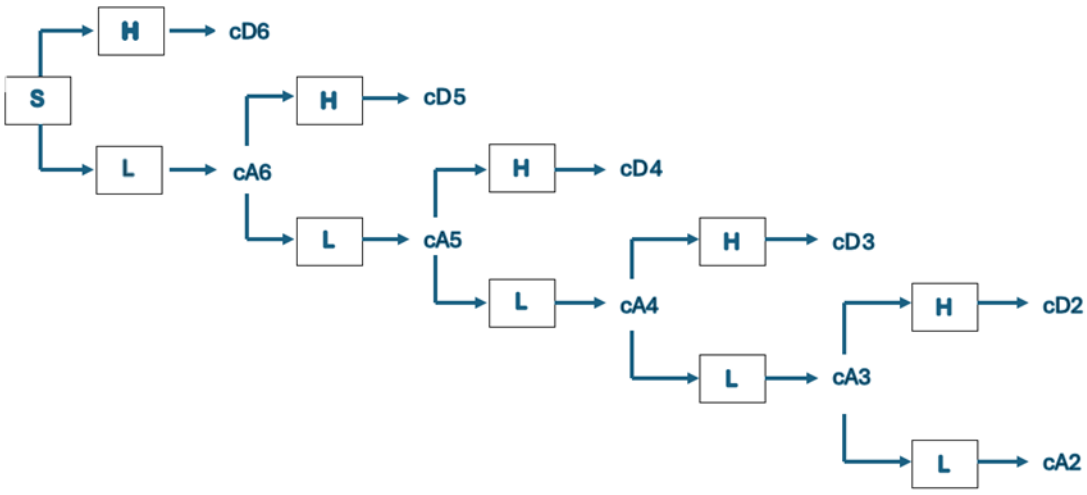

Figure 3.3: DWT decomposition process

3.5 Feature Extraction per sub-band of Simulated Signals

DWT was used to split each of the 100 vibration signals per class (i.e. normal, HA, HF, HN) into six different sub-bands (cD6, cD5, cD4, cD3, cD2, cA2) consisting of the detail and approximation coefficients. DWT was used to further investigate these frequency sub-bands and some statistical features were generated using the R software. A

total of 8 features were generated per each sub-band and consist of the 5th percentile, 25th percentile, 50th percentile, 75th percentile, 95th percentile, the mean, standard deviation (sd) and the zero-crossing rate (zcr). Therefore, since one vibration signal was decomposed into 6 different sub-bands, the final function returned a set of $8 \times 6 = 48$ features for any list values. These features were used as input for LDA.

3.6 LDA Classification with DWT and Feature Extraction

The purpose of the LDA following the DWT is to find the linear combinations of the variables (i.e. 48 features) that gives the best possible separation between the vibration categories (normal, HA, HF, HN) in the data set. Using the R software, three useful discriminant functions (LD1, LD2, and LD3) were used to separate these categories of vibration data.

Figure 3.4 is a snapshot of the linear discriminants' coefficients for the linear discriminant functions. The first discriminant function (LD1) is a linear combination of the variables and the coefficients of LD1:$(0.1198 \times d6\_05pct) + (1.4206 \times d6\_25pct) + \cdots + (-0.4512 \times c2\_zcr)$. The second discriminant function (LD2) is a linear combination of the variables and the coefficients and represented as LD2: $(-3.6696 \times d6\_05pct) + (-4.9242 \times d6\_25pct) + \cdots + (-0.3792 \times c2\_zcr)$.
Similarly, the third discriminant function (LD3) is a linear combination of the variables and the coefficients of LD3: $(-2.0771 \times d6\_05pct) + (1.2955 \times d6\_25pct) + \cdots + (-2.8082 \times c2\_zcr)$. In this illustration, $d6\_05pct$ is the 5th percentile of the 6th level detail coefficient; $d6\_25pct$ is the 25th percentile of the 6th level detail coefficient; and $c2\_zcr$ is the zero-crossing rate of the 2nd level approximation coefficient.

```
## Coefficients of linear discriminants:
##                    LD1           LD2           LD3
## d6_05pct   0.119852545 -3.669569393 -2.077098e-02
## d6_25pct   1.420665188 -4.924181546  1.295468e+00
## d6_50pct  -0.420328802  0.627927814 -3.323943e+00
## d6_75pct  -0.506922398  2.871701121  6.433704e-01
## d6_95pct  -0.092606583  1.573394346 -5.985168e-01
## d6_mean   -2.104715651  2.059233606  1.467447e+00
## d6_sd     -0.750777916  1.884561856 -1.277067e+00
## d6_zcr    -0.109494009  1.526084483  1.423562e-01
## d5_05pct  -0.195802553 -1.276730883  5.315772e-01
## d5_25pct  -1.014684166 -0.128256786  2.381193e-02
## d5_50pct  -0.064420649  0.063669595 -2.475690e-01
## d5_75pct   0.780433365  0.030121857 -5.694388e-01
## d5_95pct   0.503582302  1.437486119 -5.318095e-01
## d5_mean   -0.074951887  0.269828754  1.777649e-01
## d5_sd     -0.361271180 -1.743877003 -4.482054e-02
## d5_zcr     1.552438097  2.274280284 -2.563614e+00
       .             .             .             .
## d2_05pct  -0.009755845 -0.017409413  1.394147e-01
## d2_25pct   0.014675860 -0.009839710  2.660646e-02
## d2_50pct  -0.001496395  0.014236670  1.618433e-02
## d2_75pct  -0.018080410  0.018099458 -1.777285e-02
## d2_95pct   0.016519746 -0.008743870 -1.972984e-01
## d2_mean    0.001123900 -0.008197208  1.302626e-05
## d2_sd     -0.164966516 -0.054051048  1.734531e-01
## d2_zcr     1.456790436  1.519230254  1.366161e+00
## c2_05pct   0.244575285 -0.582950644 -1.708071e-01
## c2_25pct   0.035446674  0.063743582  1.324766e-01
## c2_50pct   0.237402960 -0.519974139 -2.594475e-01
## c2_75pct  -0.109158454  0.231099153 -3.833740e-02
## c2_95pct  -0.150823354  0.381926852  1.246417e-01
## c2_mean   -0.170302582  0.462451768  1.476356e-01
## c2_sd      0.258973050 -1.087941040 -2.997800e-01
## c2_zcr     0.451211952 -0.379154518 -2.808244e-01
```

Figure 3.4: Snapshot of linear discriminants coefficients for Chapter 3.6.

Table 3.1: Proportion of Trace for Chapter 3.6 LDA classification

Proportion of Trace		
LD1	**LD2**	**LD3**
0.6960	0.2661	0.0379

The proportion of trace is the percentage separation achieved by each discriminant function. Table 3.1 shows that LD1 has the highest separation with 69.60%, followed by LD2 with 26.61%, and LD3 with 3.79%.

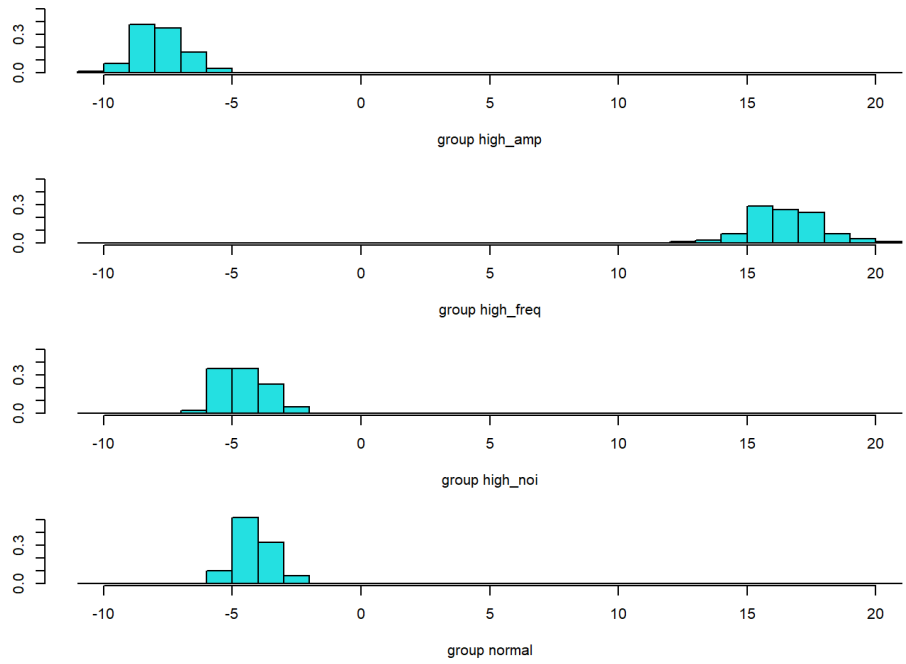

Figure 3.5: Stacked histogram of first discriminant function (LD1) for Chapter 3.6.

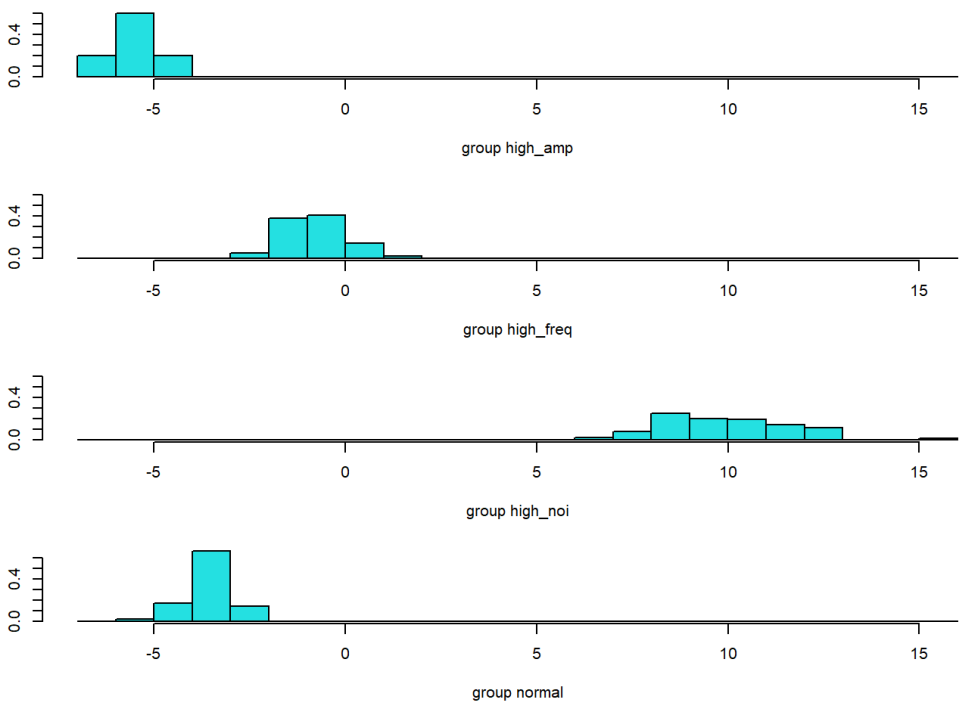

Figure 3.6: Stacked histogram of second discriminant function (LD2) for Chapter 3.6.

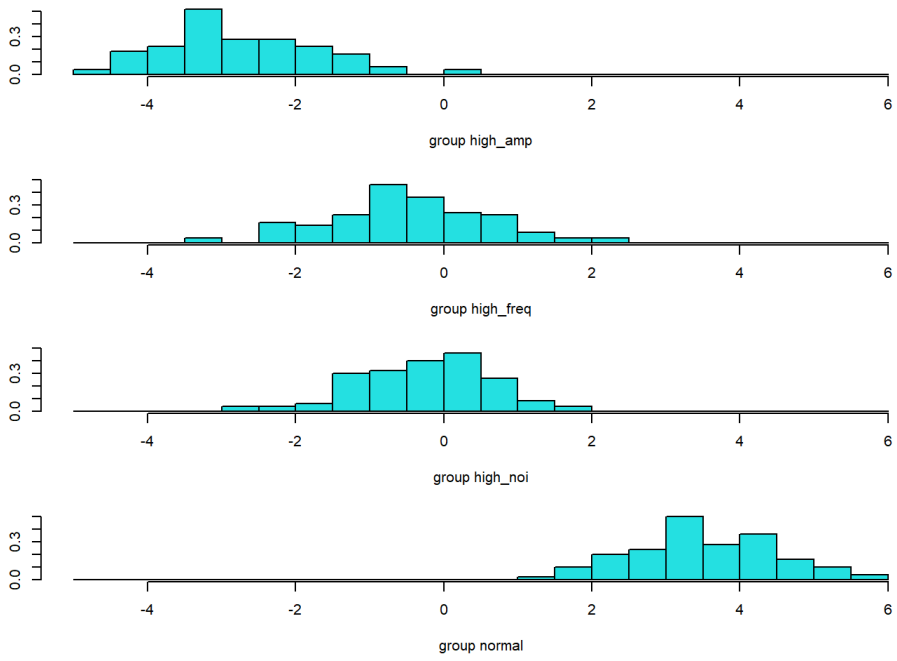

Figure 3.7: Stacked histogram of third discriminant function (LD3) for Chapter 3.6.

From the stacked histogram graph shown in Figures 3.5 for LD1, we can see the large separation between HF and other categories of vibration (HA, HN, and normal). On the contrary, there is some amount of overlapping between HA, HN, and normal. The LD2 histogram graph in Figure 3.6 shows some separation between HN and other categories of vibration (HA, HF and normal). We also observe some amount of overlapping between HA, HF, and normal. The LD3 histogram graph in Figure 3.7 shows a lot of overlap between the four vibration categories (HA, HF, HN, and normal). Table 3.1, proportion of trace, shows the percentage separation achieved by LD1 at 69.60%, followed by LD2 at 26.61%, and LD3 with the least separation at 3.79%. This explains why LD1 has the most separation, with LD2 having some separation, and LD3 with the least separation and lots of overlap as explained earlier.

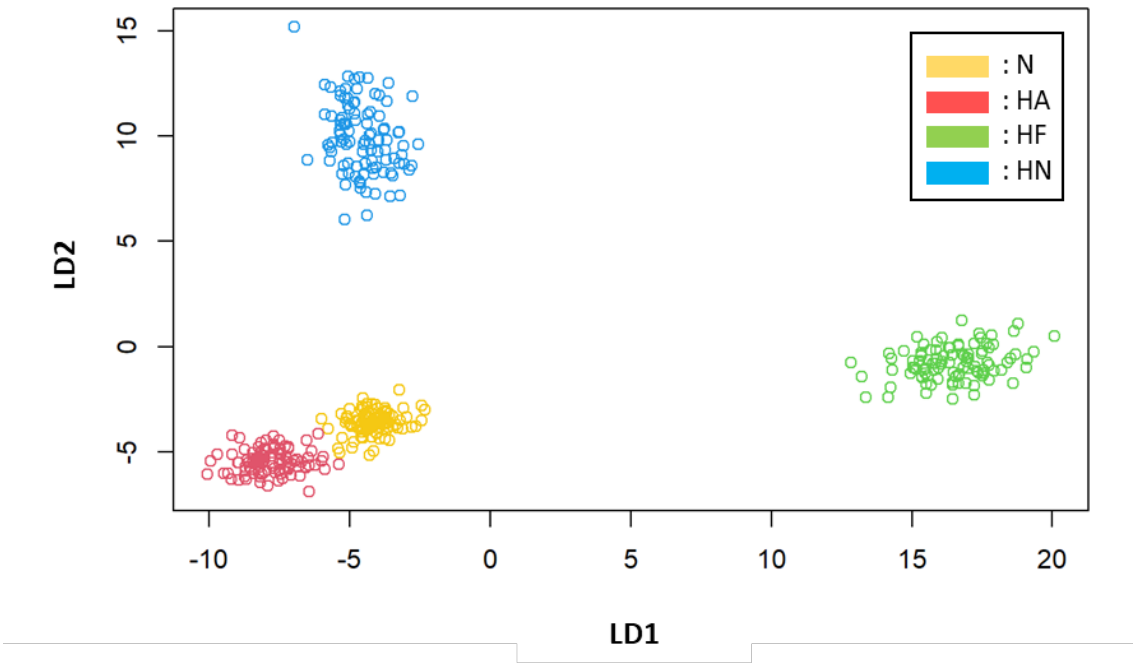

Figure 3.8: Scatterplot of LD1 and LD2 for Chapter 3.6.

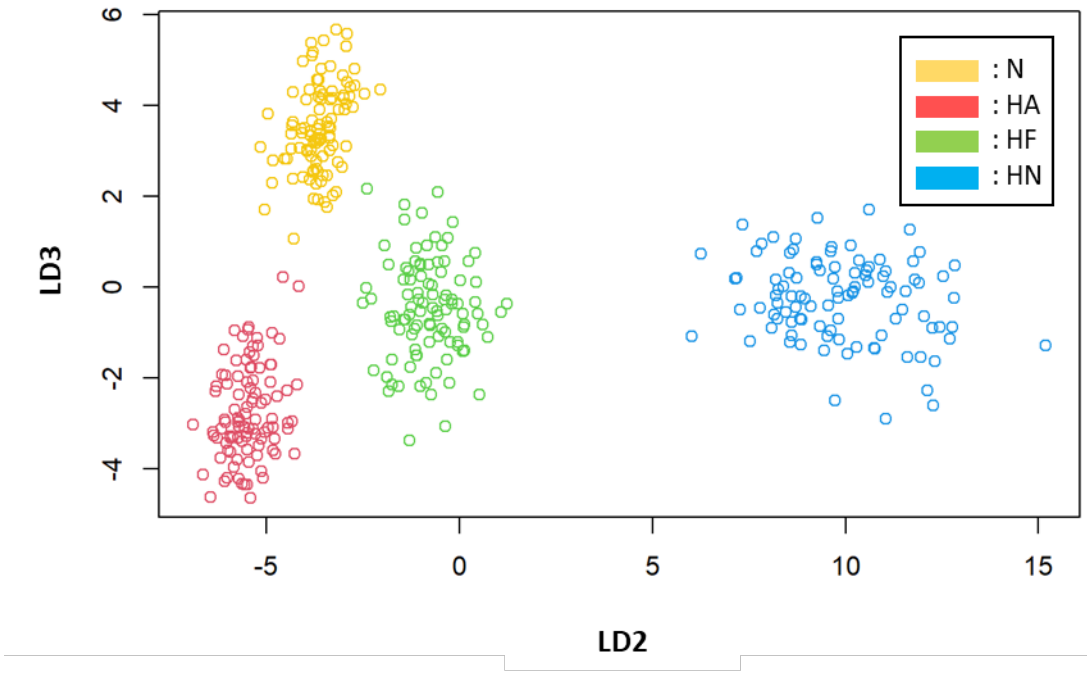

Figure 3.9: Scatterplot of LD2 and LD3 for Chapter 3.6.

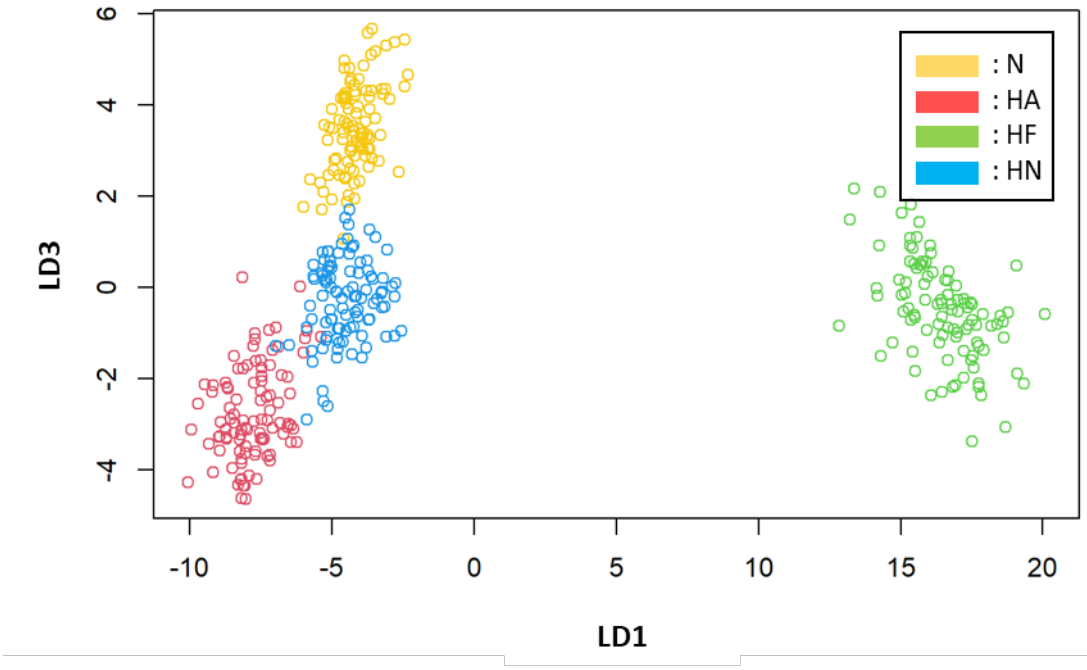

Figure 3.10: Scatterplot of LD1 and LD3 for Chapter 3.6.

In Figure 3.8, the scatterplot of the first two discriminant functions (LD1 and LD2) combined; we can see that the vibration categories are well separated. LD1 (x-axis) separates HF and other categories of vibration (HA, HN, and normal) distinctly well, but does not perfectly separate HA, HN, and normal. The LD2 (y-axis) achieves a good separation of HN and other categories of vibration (HA, HF and normal).

In Figure 3.9, the scatterplot of the second and third discriminant functions (LD2 and LD3) combined; we can see that the vibration categories are also well separated. LD2 (x-axis) separates HN and other categories of vibration (HA, HF, and normal) quite well, but does not perfectly separate HA, HF, and normal. The LD3 (y-axis) shows some overlap between the vibration categories.

In Figure 3.10, the scatterplot of the first and third discriminant functions (LD1 and LD3) combined, we can see that the vibration categories are not so well separated. LD1 (x-axis) separates HF and other categories of vibration (HA, HN, and normal) distinctly well, but does not perfectly separate HA, HN, and normal. The LD3 (y-axis) shows some overlap between the vibration categories.

From review of the three scatterplots, to achieve a good separation of the vibration categories, it will be good to use any of the two discriminant functions combined, but preferably the scatterplots of LD1 and LD2, or LD2 and LD3 as these two scatter plots displays the best separation between the vibration categories.

Table 3.2: Confusion Matrix and Statistics for Chapter 3.6.

```
Confusion Matrix and Statistics

          Reference
Prediction  high_amp high_freq high_noi normal
  high_amp       100         0        0      0
  high_freq        0       100        0      0
  high_noi         0         0      100      0
  normal           0         0        0    100

Overall Statistics

               Accuracy : 1
                 95% CI : (0.9908, 1)
    No Information Rate : 0.25
    P-Value [Acc > NIR] : < 2.2e-16
```

Table 3.2 shows the confusion matrix and some statistics. From the table, we can see a good prediction accuracy with diagonal values at 100%.

3.7 Comparison study and validation of LDA classification with DWT and feature extraction

In order to validate the results of LDA classification with DWT and feature extraction as shown in the Section 3.6, this research study also decided to perform additional LDA classification of the vibration signal categories using below methods:

- LDA classification with vibration signals only (i.e. without DWT and feature extraction).
- LDA classification with DWT coefficients (i.e. without feature extraction).
- LDA classification with features extracted from the original vibration signals (i.e. without DWT).

3.7.1 LDA Classification with Vibration Signals Only (i.e. without DWT and Feature Extraction)

In this comparative study, the LDA was applied to just the vibration signal categories (normal, HA, HF, HN) in the data set. In this application, the DWT and the feature extraction was not applied to the vibration signals prior to LDA classification.

Table 3.3: Proportion of Trace for Chapter 3.7.1 LDA classification.

Proportion of Trace		
LD1	**LD2**	**LD3**
0.7549	0.2016	0.0435

The proportion of trace is the percentage separation achieved by each discriminant function. Table 3.3 shows that LD1 has the highest separation with 75.49%, followed by LD2 with 20.16%, and LD3 with 4.35%.

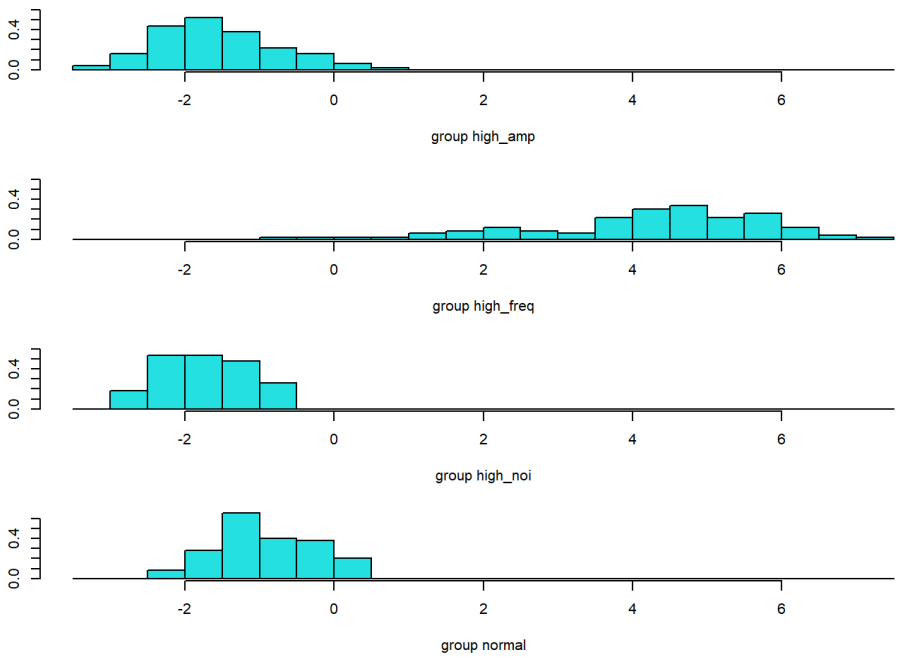

Figure 3.11: Stacked histogram of first discriminant function (LD1) for Chapter 3.7.1.

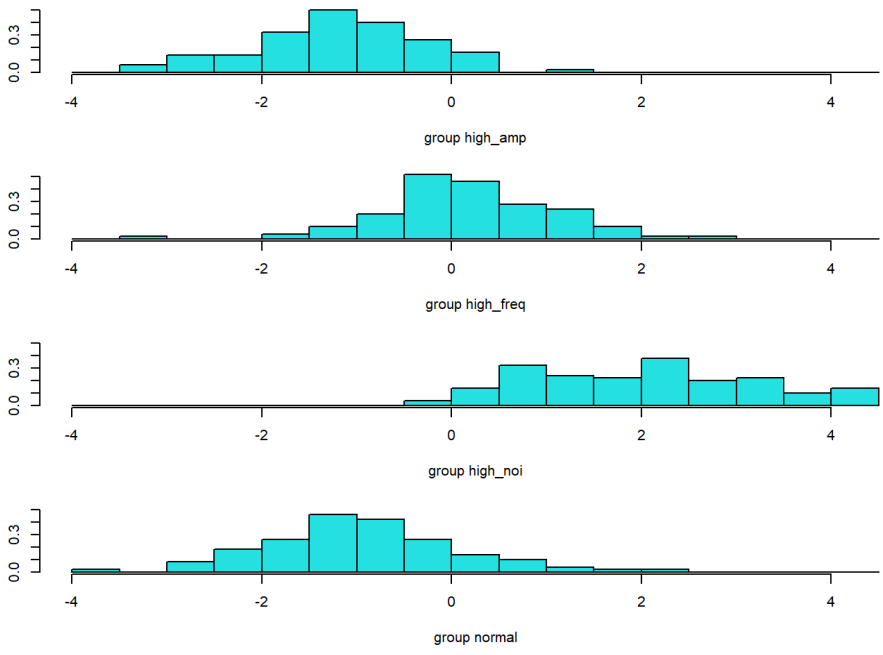

Figure 3.12: Stacked histogram of second discriminant function (LD2) for Chapter 3.7.1.

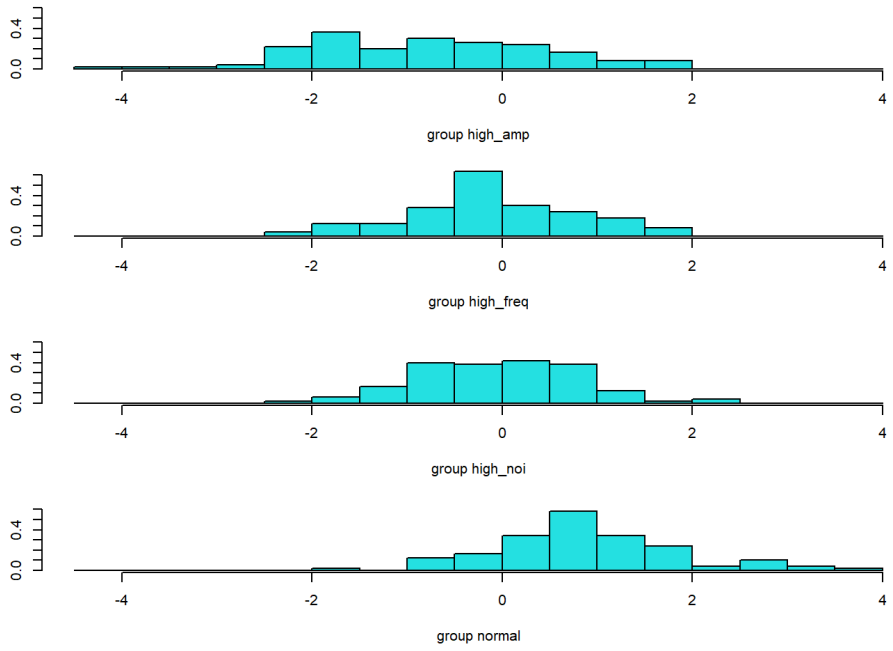

Figure 3.13: Stacked histogram of second discriminant function (LD3) for Chapter 3.7.1.

From the stacked histogram graph as shown in Figures 3.11 for LD1, we can see some separation between HF and other categories of vibration (HA, HN, and normal), but with some overlap. On the contrary, there is some amount of overlapping between HA, HN, and normal. The LD2 in Figure 3.12 and LD3 in Figure 3.13 would show the least amount of separation between the four vibration categories with significant amount of overlap.

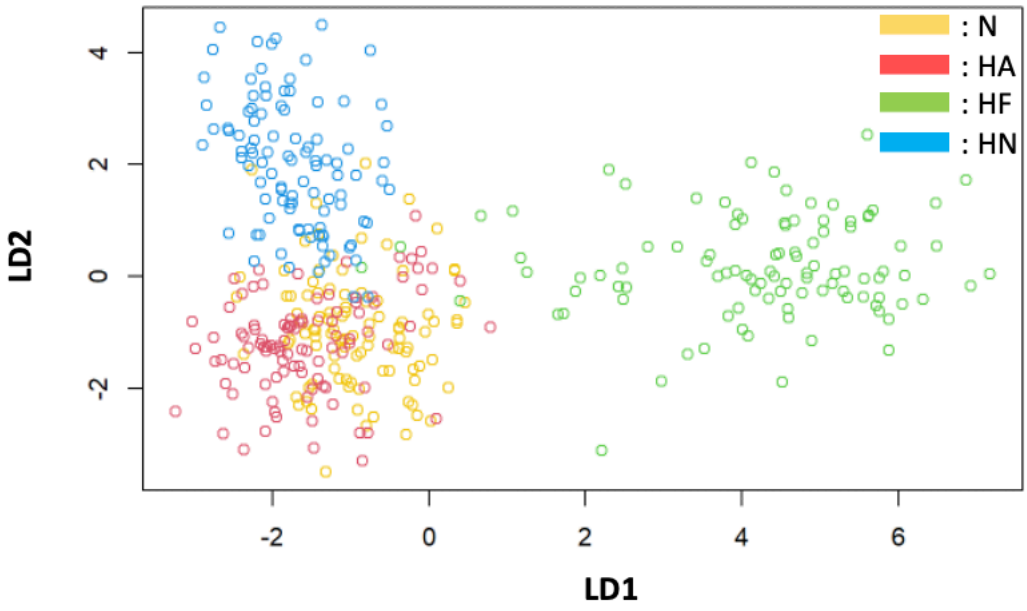

Figure 3.14: Scatterplot of LD1 and LD2 for Chapter 3.7.1.

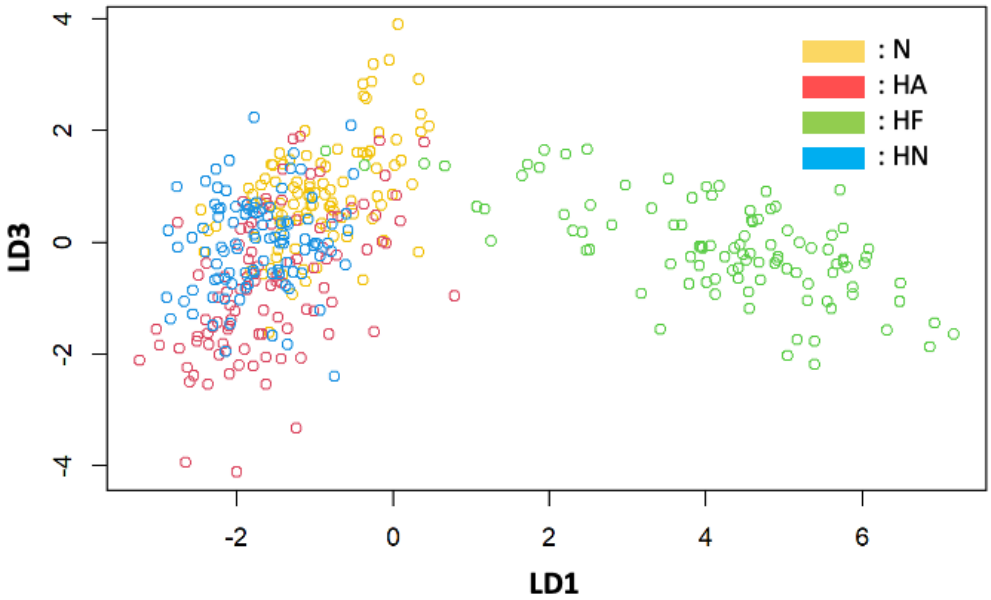

Figure 3.15: Scatterplot of LD1 and LD3 for Chapter 3.7.1.

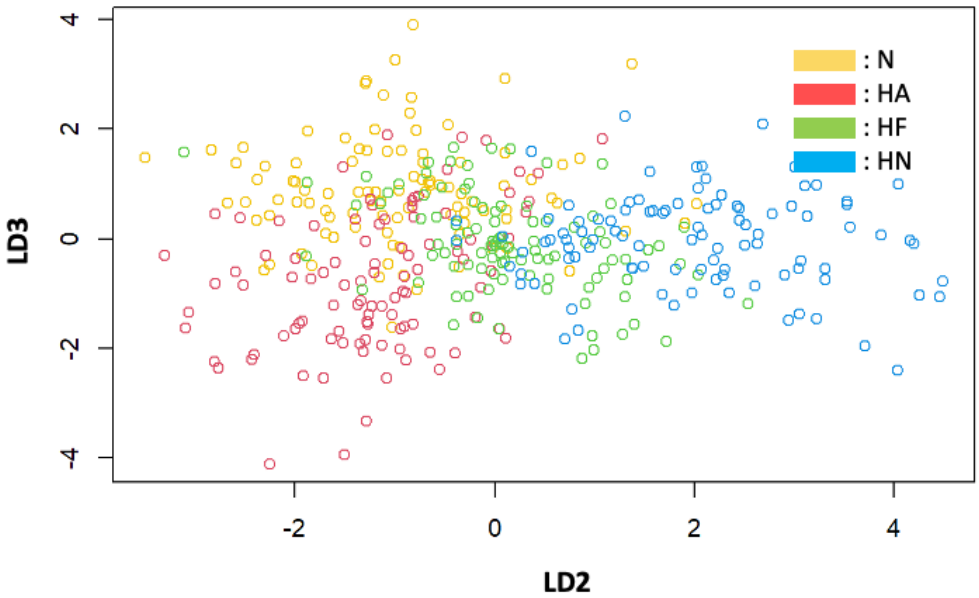

Figure 3.16: Scatterplot of LD2 and LD3 for Chapter 3.7.1.

In Figures 3.14 and 3.15, although the scatterplots separate HF and other vibration categories (HA, HN, and normal) to some extent, there is still a lot of scatter on the HF category. This would indicate variation still exists within the HF category. Figure 3.16 shows a lot of overlap between vibration categories and lots of scatter as well.

Table 3.4: Confusion Matrix and Statistics for Chapter 3.7.1.

```
Confusion Matrix and Statistics

          Reference
Prediction  high_amp high_freq high_noi normal
  high_amp        72         0        5     15
  high_freq        0        90        0      0
  high_noi         0         0       90      5
  normal          28        10        5     80

Overall Statistics

               Accuracy : 0.83
                 95% CI : (0.7895, 0.8655)
    No Information Rate : 0.25
    P-Value [Acc > NIR] : < 2.2e-16
```

Table 3.4 shows the confusion matrix and some statistics. From the table, the overall prediction accuracy is 83%.

3.7.2. LDA Classification with DWT Coefficients (i.e. without Feature Extraction)

In this comparative study, the LDA was applied to just the DWT coefficients of the vibration signal categories (normal, HA, HF, HN) in the data set. In this application, the feature extraction was not applied prior to LDA classification.

Table 3.5: Proportion of Trace for Chapter 3.7.2 LDA classification.

Proportion of Trace		
LD1	**LD2**	**LD3**
0.7549	0.2016	0.0435

The proportion of trace is the percentage separation achieved by each discriminant function. Table 3.5 shows that LD1 has the highest separation with 75.49%, followed by LD2 with 20.16%, and LD3 with 4.35%.

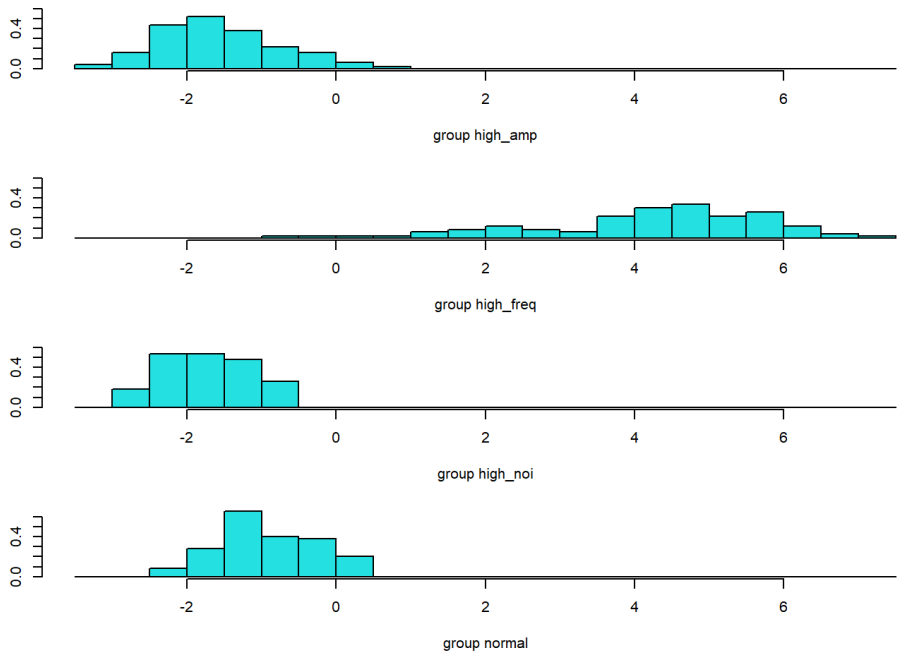

Figure 3.17: Stacked histogram of first discriminant function (LD1) for Chapter 3.7.2.

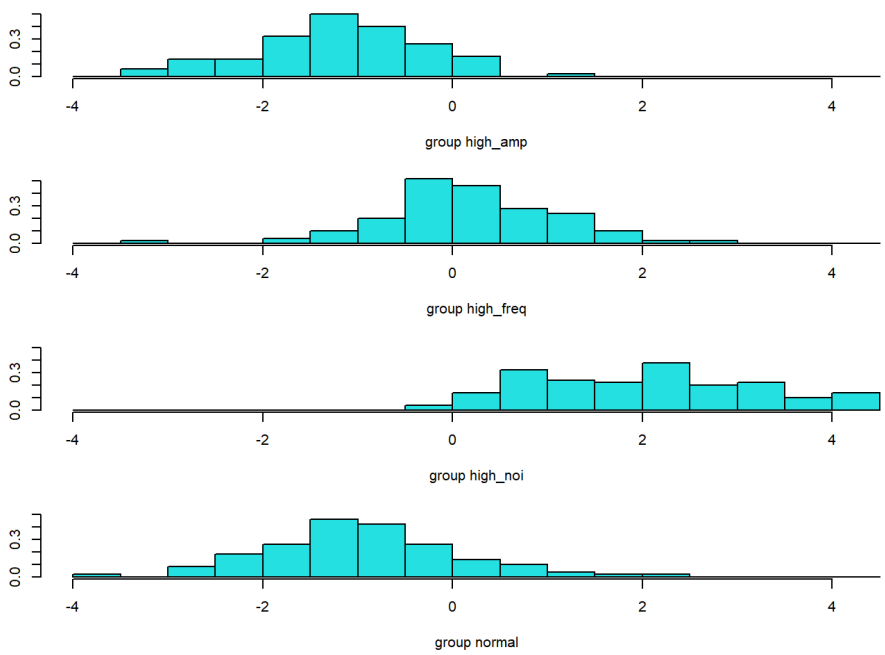

Figure 3.18: Stacked histogram of second discriminant function (LD2) for Chapter 3.7.2.

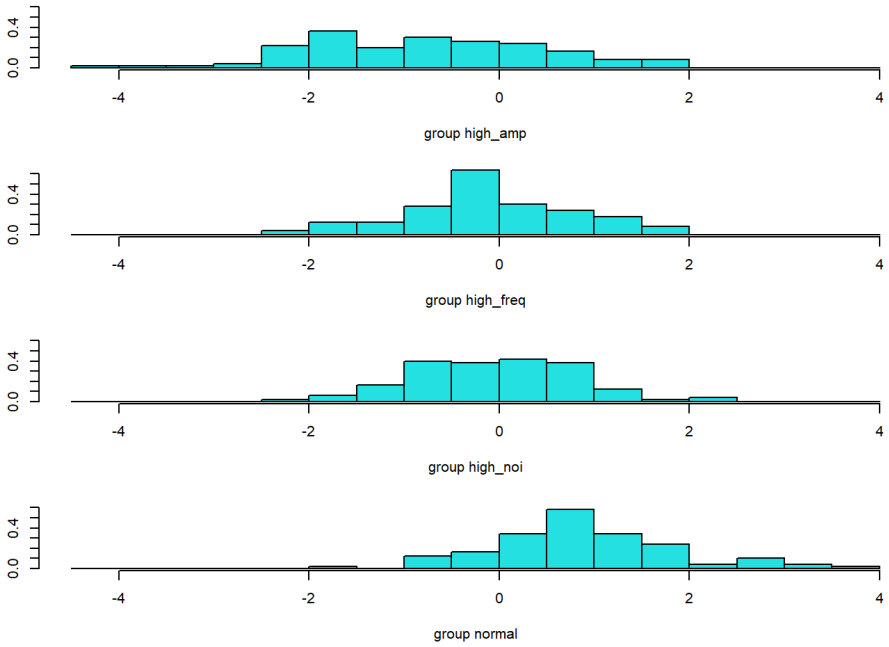

Figure 3.19: Stacked histogram of second discriminant function (LD3) for Chapter 3.7.2.

From the stacked histogram graph as shown in Figures 3.17 for LD1, we can see some separation between HF and other vibration categories (HA, HN, and normal), but with some overlap. On the contrary, there is some amount of overlapping between HA, HN, and normal. The LD2 in Figure 3.18 and LD3 in Figure 3.19 would show the least amount of separation between the four vibration categories with significant amount of overlap.

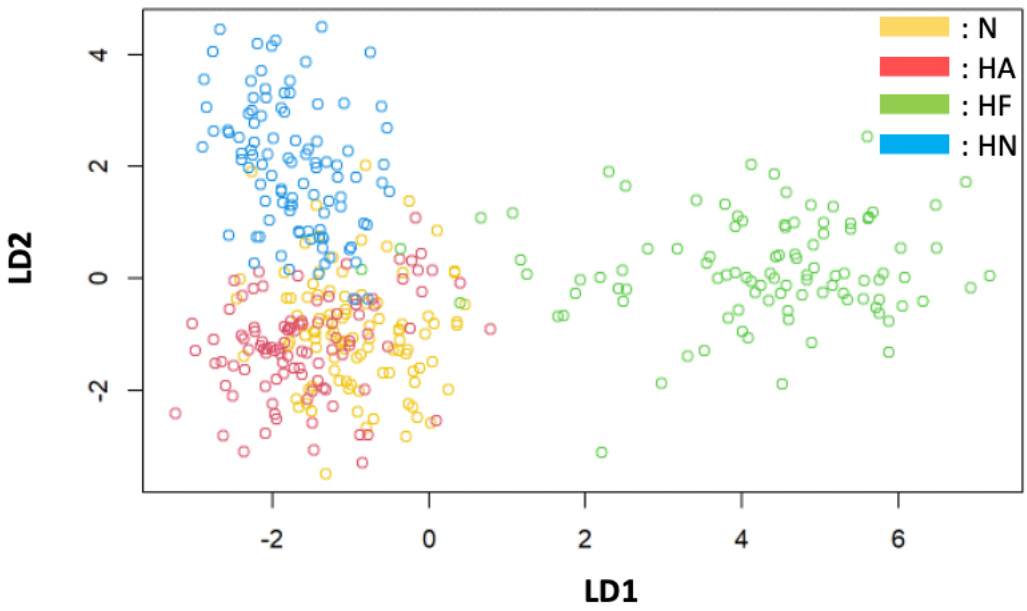

Figure 3.20: Scatterplot of LD1 and LD2 for Chapter 3.7.2.

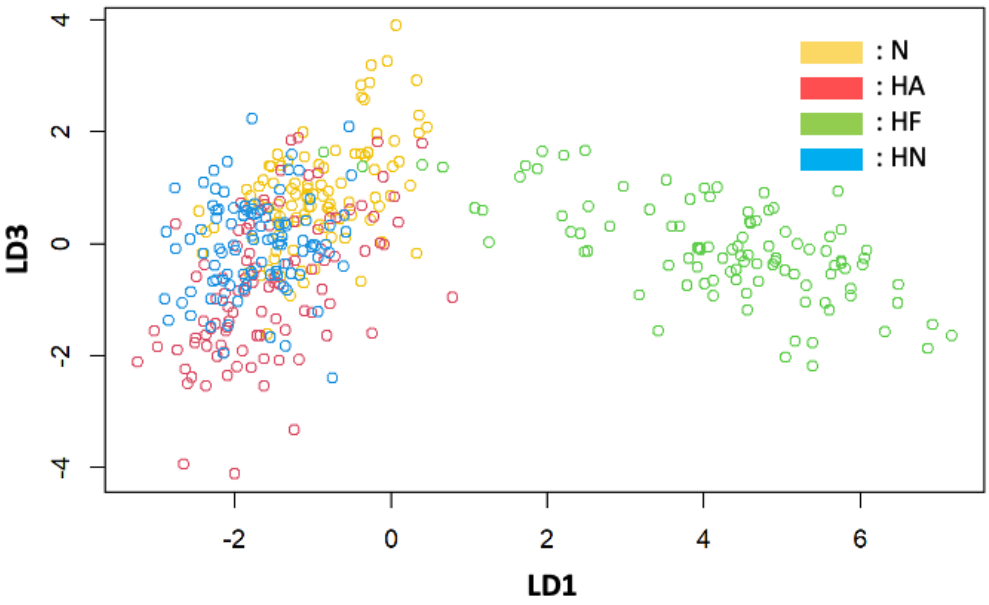

Figure 3.21: Scatterplot of LD1 and LD3 for Chapter 3.7.2.

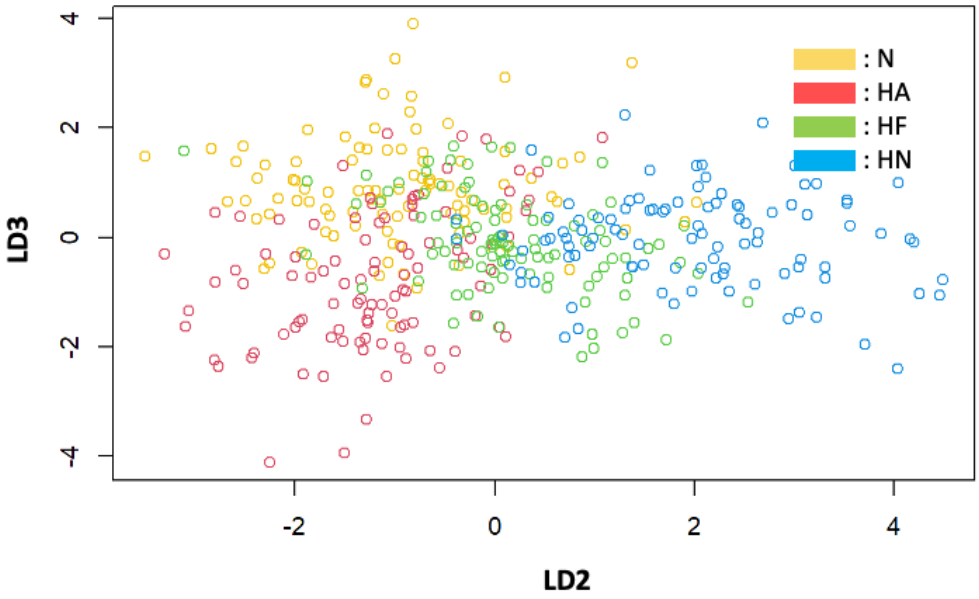

Figure 3.22: Scatterplot of LD2 and LD3 for Chapter 3.7.2.

In Figures 3.20 and 3.21, although the scatterplots separate HF and other categories of vibration (HA, HN, and normal) to some extent, there is still a lot of scatter on the HF

category. This would indicate variation still exists within the HF category. Figure 3.22 shows a lot of overlap between vibration categories and lots of scatter as well.

Table 3.6: Confusion Matrix and Statistics for Chapter 3.7.2.

```
Confusion Matrix and Statistics

          Reference
Prediction  high_amp high_freq high_noi normal
  high_amp        72         0        5     15
  high_freq        0        90        0      0
  high_noi         0         0       90      5
  normal          28        10        5     80

Overall Statistics

               Accuracy : 0.83
                 95% CI : (0.7895, 0.8655)
    No Information Rate : 0.25
    P-Value [Acc > NIR] : < 2.2e-16
```

Table 3.6 shows the confusion matrix and some statistics. From the table, the overall prediction accuracy is 83%.

Overall, an observation of the results of LDA classification with DWT coefficients found similar data was obtained when compared to Section 3.7.1 - LDA classification with vibration signals only. The DWT decomposes the vibration signals into some coefficients, and these coefficients would still retain the original characteristics of the original signals. This would explain the reason why the scatterplots and the histograms for Section 3.7.1 and this section would be similar.

3.7.3 LDA Classification with Features Extracted from the Original Vibration Signals (i.e. without DWT)

In this comparative study, the LDA was applied to the features extracted from the original vibration signals (normal, HA, HF, HN) in the data set. In this application, the DWT was not applied prior to extraction of features and LDA classification.

Table 3.7: Proportion of Trace for Chapter 3.7.3 LDA classification

Proportion of Trace		
LD1	**LD2**	**LD3**
0.7317	0.2629	0.0004

The proportion of trace is the percentage separation achieved by each discriminant function. Table 3.7 shows that LD1 has the highest separation with 73.17%, followed by LD2 with 26.29%, and LD3 with 0.4%.

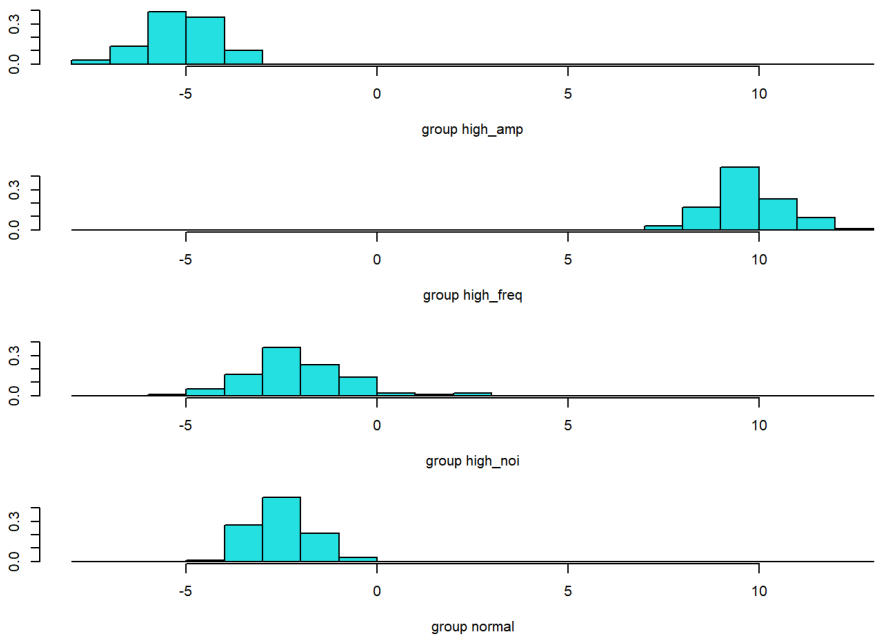

Figure 3.23: Stacked histogram of first discriminant function (LD1) for Chapter 3.7.3.

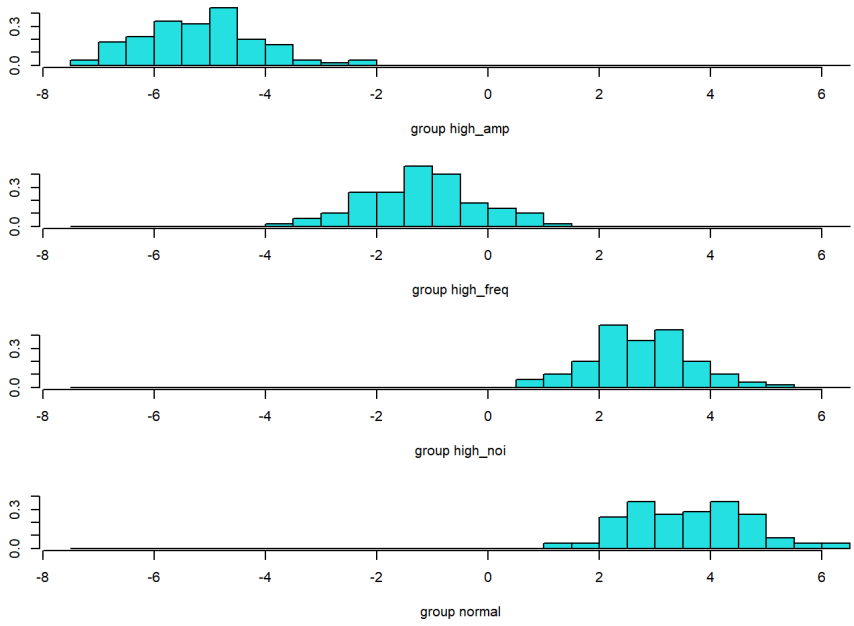

Figure 3.24: Stacked histogram of second discriminant function (LD2) for Chapter 3.7.3.

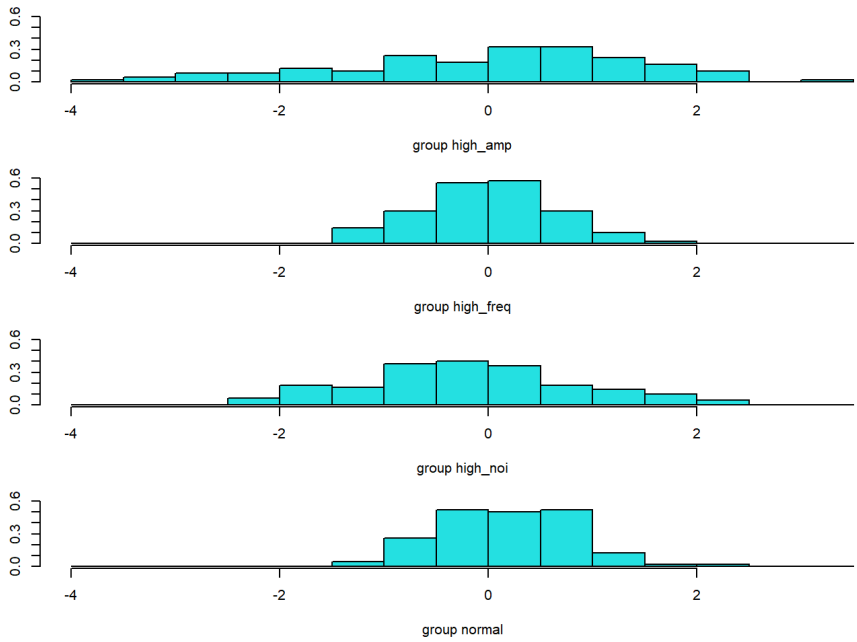

Figure 3.25: Stacked histogram of third discriminant function (LD3) for Chapter 3.7.3.

From the stacked histogram graph as shown in Figures 3.23 for LD1, we can see a good separation between HF and other vibration categories (HA, HN, and normal), but with no overlap. On the contrary, there is some amount of overlapping between HA, HN, and normal. The LD2 in Figure 3.24 show some overlap with no clear distinct separation between the vibration categories. The LD3 in Figure 3.25 would show the least amount of separation with significant amount of overlap between the four vibration categories.

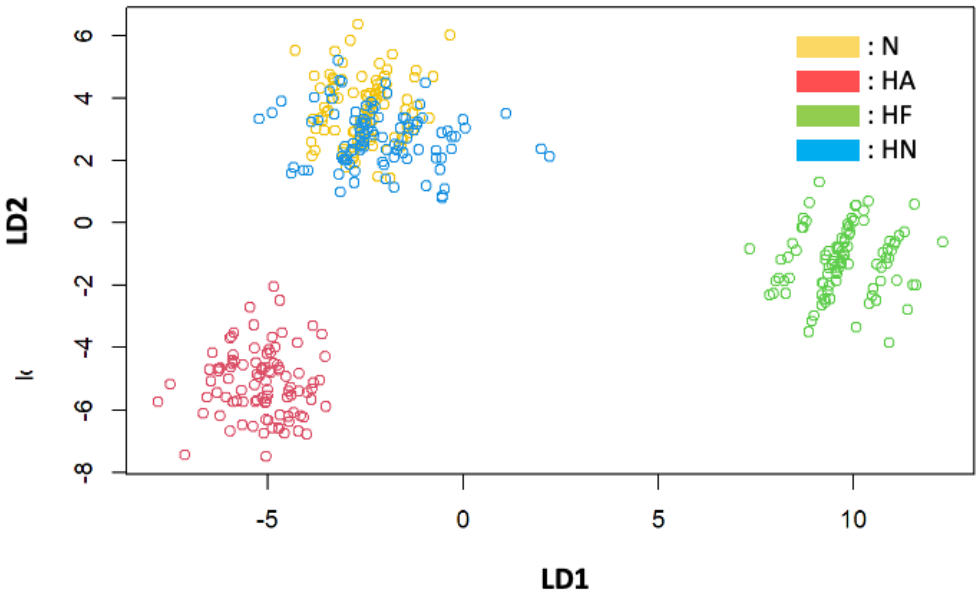

Figure 3.26: Scatterplot of LD1 and LD2 for Chapter 3.7.3.

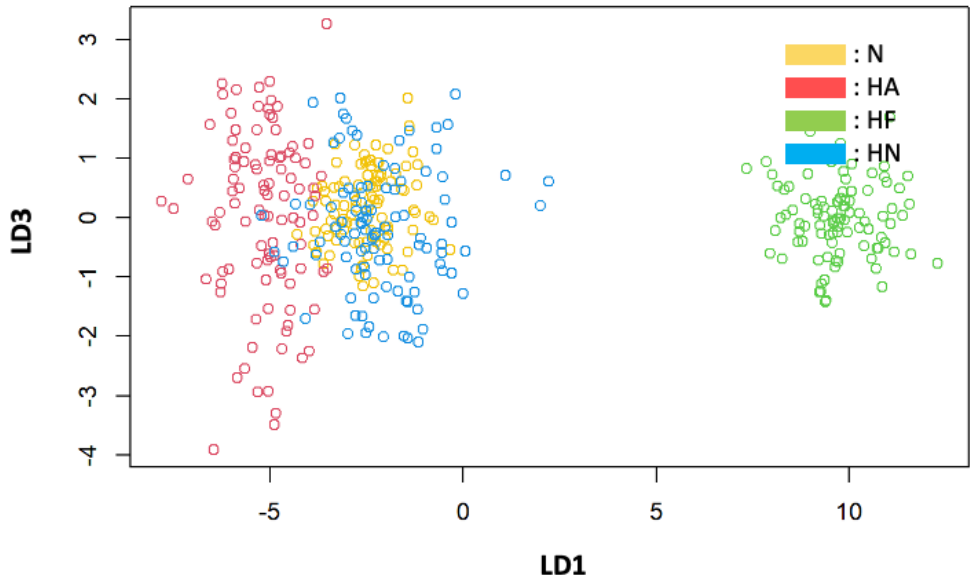

Figure 3.27: Scatterplot of LD1 and LD3 for Chapter 3.7.3.

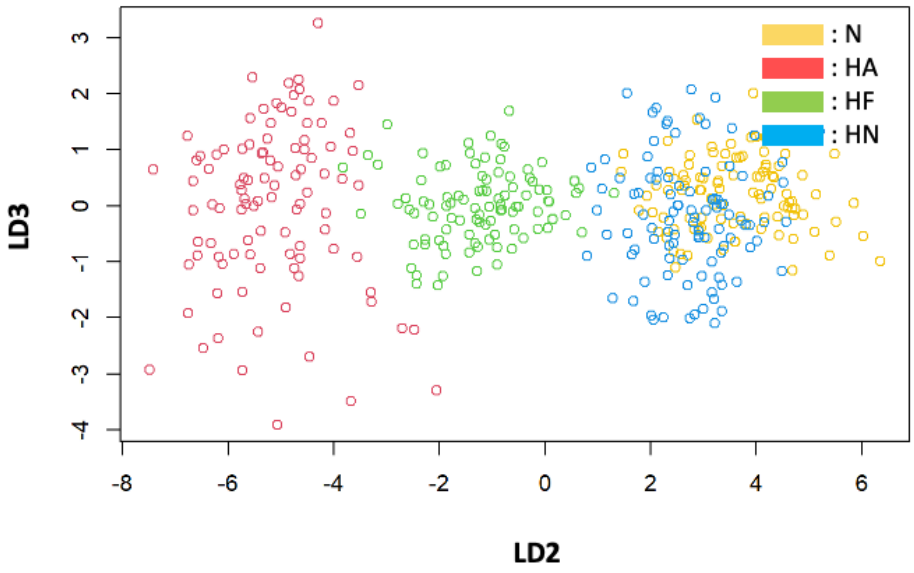

Figure 3.28: Scatterplot of LD2 and LD3 for Chapter 3.7.3.

In Figure 3.26 shows the scatterplot of the first two discriminant functions (LD1 and LD2) combined. It is observed the HF separates adequately from other vibration categories (HA, HN, and normal). The normal and HN categories are not separated with significant overlap observed when compared with HA that shows good separation. LD1 (x-axis) separates HF from other vibration categories with significant overlap observed between HA, HN, and normal. The LD2 (y-axis) show significant overlap between HN and normal, but HF and HA appear well separated with little amount of overlap.

Figure 3.27 shows the scatterplot of the discriminant functions (LD1 and LD3) combined. It is observed the HF separates adequately from other vibration categories (HA, HN, and normal), but significant overlap and some scatter was observed from the other vibration categories (HA, HN, and normal). On the LD1 (x-axis), a similar observation was observed with LD1 and LD3 combined. The LD3 (y-axis) show significant overlap

between all vibration categories. This is expected as LD3 only contributes 0.4% from the proportion of trace table.

Figure 3.28 shows the scatterplot of the discriminant functions (LD2 and LD3) combined. It is observed HF and HA separates well when compared to HN and normal vibration categories. A significant amount of scatter was also observed on HA. On the LD2 (x-axis), a similar observation was observed with LD2 and LD3 combined. The LD3 (y-axis) show significant overlap between all vibration categories. This is expected as LD3 only contributes 0.4% from the proportion of trace table.

3.8. Discussions and Conclusions

The results shown from this simulation study demonstrate the combined application of DWT decomposition, feature extraction, and LDA where it shows good results for vibration monitoring. The classification feature of LDA as applied in this study demonstrates how some abnormal vibration data such as HF, HN, and HA can be easily distinguished from the normal vibration data by observation of the linear discriminants scatter plots. The confusion matrix and statistics also show good prediction accuracy of the applied method when compared to other approaches that did not combine all three methods in the classification of the different vibration categories. This powerful technique can be applied for monitoring actual vibration signals of any rotating equipment.

The results from this study can be put into good practice in monitoring the vibration condition monitoring of rotating equipment used in power plants and other applications. The practical applicability will involve reviewing an equipment performance history and

obtaining the categories of vibration data (normal and various forms of abnormal data). The data can be trained using the LDA supervised learning as applied in this study for classification and distinguishing the vibration categories. One of the benefits of the applied method is the visual interpretation of vibration data from the linear discriminants scatter plots. This saves the analyst or predictive maintenance engineer precious time of navigating through complex analysis of vibration time waveform or frequency spectra which is typically the case for most power plants. The scatter plots will serve as a one-stop visualization of the rotating equipment condition. In this way, anomalies can be easily identified in detecting equipment degradation or malfunction. The necessary corrective actions can be planned in advance to fix equipment condition without resulting in a catastrophic failure. This would support the overall goal of predictive maintenance.

Chapter 4

Power Plant Vibration Monitoring and Optimization using Functional Design of Experiments.

4.1 Introduction

In a typical pressurized water reactor (PWR) nuclear power plant, the condensate pumps contribute the majority of the feedwater flow to the steam generators, which is converted into steam that spin the turbine generator to produce electricity. Adequate monitoring of the condensate pump feedwater flow is essential not only for the pump's reliability but also for efficient plant operations overall.

Vibration data in a time waveform acquired in a power plant application typically undergo a fast Fourier transformation (FFT) where the raw vibration signals in the time waveform are transformed to the frequency spectrum. The frequencies corresponding to the vibration amplitudes can be visualized and used to diagnose equipment faults. However, the time waveform could present some varying sharp peaks in the vibration amplitudes which may not be adequately modelled using the FFT. To better investigate this problem, a combination of wavelet transformation and functional principal component analysis (FPCA) is proposed and used in this paper to model a condensate pump time waveform vibration signal.

Discrete wavelet transform provides a time-scale information of a signal enabling extraction of features that vary in time. This property makes the wavelets an ideal tool for analyzing a signal of a transient or non-stationary nature (Haddadi et al. 2014).

Functional data comes in many forms, but their defining quality is that they consist of functions often, but not always, smooth curves (Ramsay and Silverman 2005). Functional data can be obtained from different fields and applications with examples including vibration signals, sensor streams, measurements taken over a range, spectral data, radar signatures, etc.

As explained in Chapter 2, principal component analysis (PCA) simplifies the complexity in high dimensional data while retaining trends and patterns. It does this by transforming the data into fewer dimensions which acts as summaries for the features. PCA finds the principal components (vectors), which maximize the variance of the data along their direction. Each principal component explains a portion of the total variance in the data. FPCA is similar to PCA; the difference is that FPCA deals with functions whereas PCA deals with vectors. The principal components from the FPCA are also functions or curves and are called the functional principal components.

In this study, a combination of the FPCA and FDOE is proposed and used to identify the effects of an environmental factor (pump flow rate) and optimize the shape of the vibration signal. Continuous vibration signals from the condensate pumps are obtained at different modes of nuclear power plant operations at varying flow rates. The FPCA is used to breakdown and reduce the dimensionality of these signals into a few empirical basis functions called functional principal components (FPCs). The FDOE is then applied to identify the effects of pump flow rates (environmental factor) and used to obtain the optimal shape of the target vibration curve.

4.2 Review of Previous Works

Data collected from measurements over time can be analysed in several ways. Examples of such continuous (functional) data include sensor streams, measurements taken over range, vibration signals, spectral data, tool wear, electrocardiograms (EKGs), trajectories of flights between cities, radar signatures, etc (Parker 2021).

The functional data analysis (FDA) is a non-parametric method that has been used to analyse functional data by which data profiles are modelled with functions.

FDA begins with two steps, an initial basis function smoothing step followed by a feature extraction dimension reduction step called FPCA. The basis function smoothing step allows FDA to be applied to a wide variety of shapes of the profiles. Let $x(t)$ denote a functional datum over time t, and $\beta_1(t), \beta_2(t), \dots, \beta_m(t)$ denote a standard set of functions, i.e., a basis. The function $x(t)$ of interest can be expanded in terms of the basis as

$$x(t) = \sum_{j=1}^{m} \xi_j \beta_j(t) \tag{4.1}$$

then the vector $\boldsymbol{\xi} = (\xi_1, \dots, \xi_m)$ determines the function's shape. In this study, the wavelet models are used as the basis functions.

FPCA dimension reduction step simplifies the analysis by reducing profile data consisting of potentially thousands of measurements per profile down to a small number (typically between 1 to 10) of scalar values called functional principal component (FPC) scores. Originally, the PCA is an unsupervised learning method and reduces data by geometrically projecting them into lower dimensions called principal components, with the goal of finding the best summary of data using a limited number of principal components.

They build an orthonormal basis, which means each data point is a linear combination of the principal components. Similarly, in a functional context, the FPC scores are the most efficient low dimensional linear decomposition of the observed profile-to-profile variation. This process results in a set of eigenvalues and (smooth) eigenfunctions extracted from the sample covariance matrix of the smoothed data, which can be thought of as empirical basis functions that reveal the shapes of curves on which most variability occurs. Typically, a few number of estimated eigenfunctions exhibit characterizations of the individual curve variation (Kenett and Gotwalt 2023).

He et al. (2014) performed an early literature review on various approaches and techniques used in handling dynamic responses in experimental design. The paper acknowledged that in some situations the response is a continuous variable and not a steady state constant. In such situations, therefore, it is realistic to consider the dynamic characteristics inherent in the actual response, instead of a single constant response value.

Kenett and Gotwalt (2023) used functional data analysis (FDA) and nonlinear regression (NLR) models in analyzing dissolution profiles of drug tablets where profiles of tablets under test are compared to a target reference dissolution curve. An information quality framework was used to evaluate the two approaches, FDA and NLR with a suggestion to combine both methods in order to enhance information quality.

Wang et al. (2016) presented an overview of FDA, starting with simple statistical notions such as mean and covariance functions, then converting some core techniques the most popular being the functional principal component analysis (FPCA). Additional core techniques such as linear regression, clustering and classification of functional data, and

other non-linear approaches such as feature time warping, manifold learning and empirical differential equations were discussed.

Spreafico et al. (2022) used a functional covariate cox model (FunCM) to study the association between time-varying processes. A time-to-event outcome is then proposed based on FunCM exploiting functional data analysis techniques to represent time-varying processes in terms of functional data.

Sozen and Oner (2022) applied the Fourier basis function and roughness penalty patch to transform the daily average temperature data from 18 cities in Turkey's black sea region into continuous function. This was used to evaluate changes in temperature functions.

Biswas et al. (2023) demonstrated the application of functional principal components and functional data clustering method to characterize the longitudinal patterns of vaginal microbiota.

Fidaleo (2020) combined FDA and design of experiments (DoE) to obtain the dynamic design space for the batch milling process of a hazelnut and cocoa based paste, using a face-centered central composite design with two functional responses (fineness and energy) and three factors (rotational speed, mass of balls, and ball diameter). The developed models were used to predict the functional responses as a function of the experimental factors and time.

4.2 Discrete Wavelet Transform (DWT)

In this study, wavelets are used to represent the discrete vibration measurements as a functional form. Herein, we briefly review the idea of discrete wavelet transform and

wavelet shrinkage which enables reducing noise included in the original measurements and thus obtaining smooth curves constructed with wavelet bases.

Suppose discrete sequence of data is given as $\boldsymbol{y} = (y_1, y_2, \dots, y_N)$ where $N = 2^J$, for some integer $J \geq 0$, so that the sequence is *dyadic*. The original measurements can be treated as the finest scale's signal with level $j = J$. For the case of Haar wavelet, by adding or subtracting non-overlapping consecutive pairs of measurements, $(y_1, y_2), (y_3, y_4), \dots, (y_{N-1}, y_N)$, coarser-scale coefficients are obtained with $j = J - 1$. By repeating this process, all the measurements are summarized at the coarsest level $j = 0$. Wavelet bases are generated by translations (shifting locations from left to right) and dilations (transition from a coarser scale to a finer scale by doubling the frequency) of a scaling function $\phi(x)$. The sequence $\boldsymbol{y}$ (or a function in general) can be represented by a linear combination of orthonormal basis of wavelets which are $\phi_{0,0}(x)$, the father wavelet at level $j = 0$ and location $k = 0$, and $\psi_{j,k}(x)$, the mother wavelets at levels $j = 0, \dots, J - 1$ and locations $k = 0, \dots, 2^j - 1$. The father wavelet represents the 'approximation (average)' or 'overall' level of signal and the mother wavelets represent the 'difference' or 'detail' of signal at the corresponding scale. It can be shown that, at a coarser level $j - 1$, the approximated signal (coefficients) can be decomposed to approximation and detail at a finer level j. Applying this in a cascading fashion, the function behind the original signal $\boldsymbol{y}$ can be represented by

$$f(x) = c_{0,0}\phi_{0,0}(x) + d_{0,0}\psi_{0,0}(x) + d_{1,0}\psi_{1,0}(x) + d_{1,1}\psi_{1,1}(x) + \cdots + d_{J-1,2^{J-1}-1}\psi_{J-1,2^{J-1}-1}(x) \tag{4.2}$$

that consists of the approximation at the coarsest scale and the details of all the scales and locations, where $c_{j,k}$ and $d_{j,k}$ are coefficients of basis functions. In general, the discrete

wavelet transform produces a vector of coefficients consisting of father wavelet coefficients at level j (not necessarily 0) and mother wavelet coefficients for finer levels $j+1, \ldots, J-1$. These coefficients can be interpreted as magnitudes of similarities between the given measurements and bases.

There exist many other wavelets besides Haar wavelet. For example, the *symmlet* family provides more smooth basis functions than Haar bases. For *symmlet-*N, the addition and subtraction operations used for Haar wavelets are replaced with linear combinations, more general operations of addition and subtraction, of N neighboring measurements. Figure 4.1 shows the coefficients used for the linear combinations for approximation (father wavelet) and detail (mother wavelet) for *symmlet*-4. From the original signal, eight measurements are combined to produce the next coarser-scale approximation and detail coefficients. From Figure 4.1, we can see the father wavelet is expected to produce the overall level of neighbouring measurements around $x = 2, 3,$ and 4, while the mother wavelet represents the difference between measurement at $x = 4$ and the neighbouring measurements. Repeating the process with the approximation coefficients, the next coarser-scale approximation and detail coefficients are produced, and this process is repeated. The outcome of DWT is often plotted by a bar plot representing magnitudes and signs of coefficients for each level.

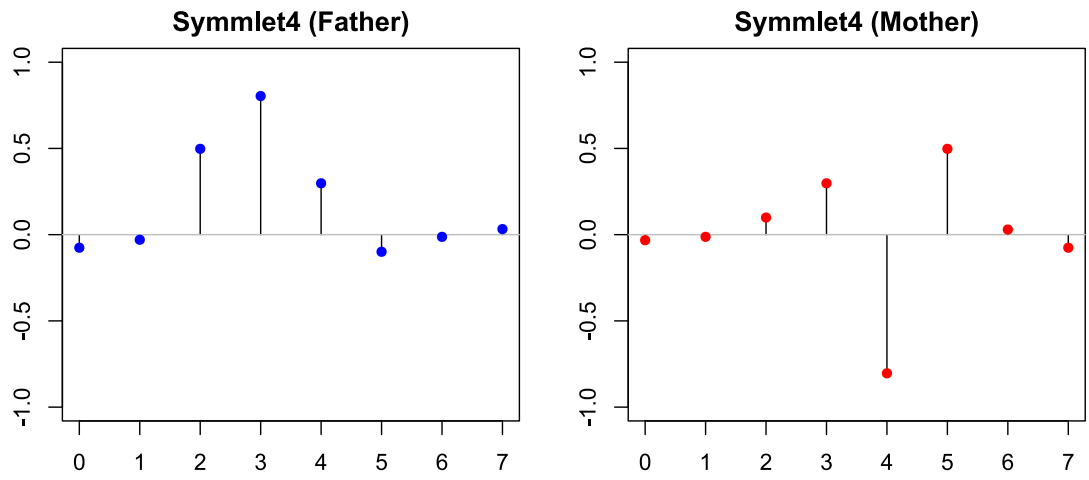

Figure 4.1: The (discretized) father and mother symmlet-4 wavelet bases used for linear combinations of signal.

The same outcome of DWT can be obtained by a matrix multiplication $\boldsymbol{d} = \boldsymbol{W}\boldsymbol{y}$ where W is the $N \times N$ orthonormal wavelet basis matrix evaluated at the N uniformly spaced observations. Figure 4.2 plots the columns of $\boldsymbol{W}$ with $N = 128$, created by the R package 'wavethresh', showing wavelet basis functions of *symmlet*-4 up to level 3. The inverse transform can be made $\boldsymbol{y} = \boldsymbol{W}^T\boldsymbol{d}$ to reconstruct the original signal from the DWT coefficients. In this process, the small coefficients are eliminated (become zeros) by applying a threshold, then we obtain a noise-reduced signal $\boldsymbol{y}^* = \boldsymbol{W}^T\boldsymbol{d}^*$. This is called a wavelet shrinkage. For more details of DWT and wavelet shrinkage, refer to (Nason, Wavelet method in Statistics with R), (Hastie et al. The elements of Statistical Learning), and references therein.

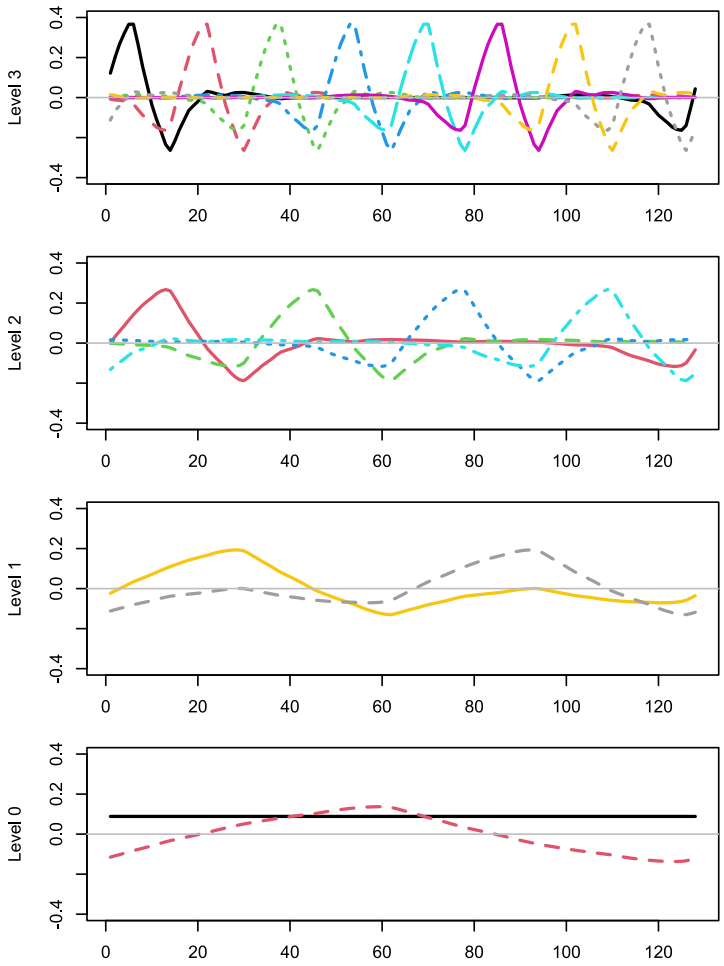

Figure 4.2: Selected columns of $\boldsymbol{W}$ with $N = 128$ for symmlet-4 plotted by connecting column values by lines. These can be interpreted as basis functions (up to level 3) for wavelet smoothing.

4.3 Functional Principal Component Analysis (FPCA)

FPCA is a method for investigating the dominant modes of variation in functional data. It is a widely used dimension reduction tool for functional data analysis. The FPCA aims to find the empirical basis functions which shape most variability based on different

vibration curves as discussed in the subsequent chapter of this paper. The empirical basis function estimated by the FPCA method are the eigenfunctions.

As described in Kenett and Gotwalt (2023), each curve in FPCA is considered an independent realization of a univariate stochastic process $X(t)$ with smooth mean function $E[X(t)] = \mu(t)$ and covariance function $Cov\{X(s), X(t)\} = \sigma(s,t)$. A spectral decomposition of the covariance function expresses $\sigma(\cdot, \cdot)$ as an orthogonal expansion in terms of its eigenvalues λ_j and associated eigenfunctions $V_j(t)$, so that

$$\sigma(s,t) = \sum_{j=1}^{\infty} \lambda_j V_j(s) V_j(t) \tag{4.3}$$

where eigenvalues quickly tend to zero and the first few eigenfunctions are slowly varying. The covariance function σ, is positive-definite and hence, the eigenvalues are nonnegative and can be ordered: $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_3 \geq 0$._The goal is to determine the primary components of functional variation in $\sigma(s,t)$, where the eigenvalues indicate the amount of total variance attributed to each component. A random curve can then be expressed as

$$X(t) = \mu(t) + \sum_{j=1}^{\infty} \xi_j V_j(t) \tag{4.4}$$

where the coefficient

$$\xi_j = \int \big(X(t) - \mu(t)\big) V_j(t) dt \tag{4.5}$$

is the inner product between the observed function, corrected by the mean function, and the j-th eigenfunction, called the j-th FPC score, which can be interpreted as a similarity measure of the observed function and the empirical basis. The eigenfunctions $V_j(t)$, called the shape functions, satisfy:

$$\int V_i(t)\, V_j(t) dt = \begin{cases} 1 & i = j \\ 0 & otherwise \end{cases} \tag{4.6}$$

and therefore, eigenfunctions are orthogonal to each other.

4.4 Power Plant Vibration Signal Data Acquisition

In a pressurized water reactor (PWR) nuclear power plant, the main feedwater system pressurizes and transports feedwater from the condensate system pumps, and heater drain pumps to the inlet of the steam generators. The condensate flow makes up of about 66% of total feed flow to the steam generators. The remaining 34% comes from the heater drain system. The condensate flow (from the condensate pumps) and heater drain flow (from the heater drain pumps) combine at the suction of the main feed pumps. The heated high-pressure water from the reactor core flows to the steam generator and transfers its thermal energy to the low-pressure water from the feed water system. The heat transfer from the high pressure (reactor core) system and low-pressure (feed water) system generates steam. This steam drives the turbine blades and spins an electric generator to produce electricity.

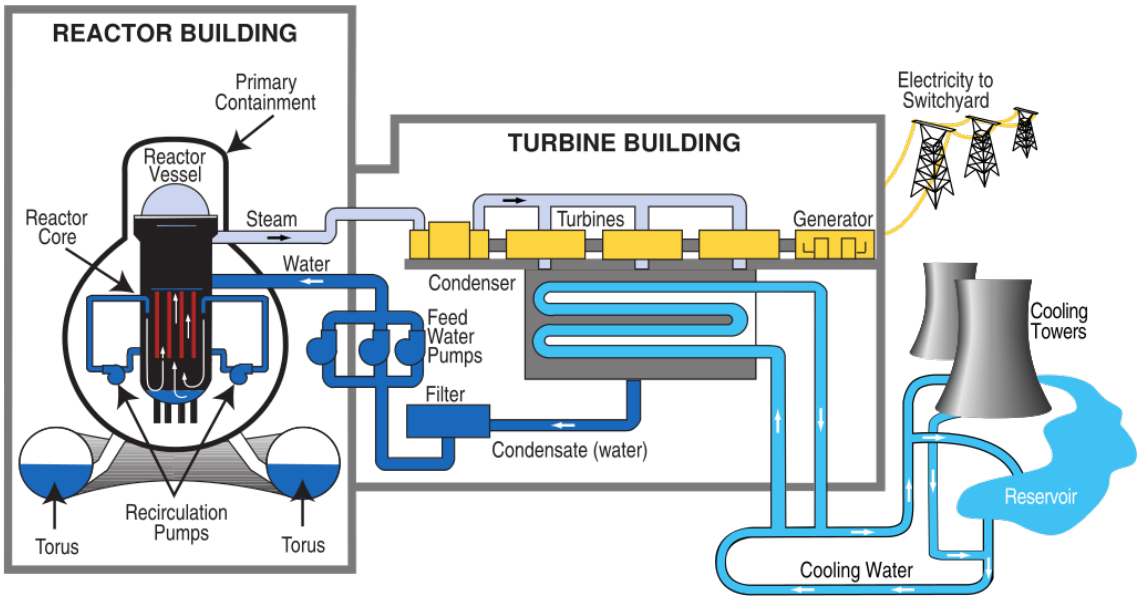

Figure 4.3: A sketch showing condensate and feedwater flows in a nuclear power plant. (Wikimedia 2013).

In this study, a sample of continuous vibration signal data in the time waveform obtained from a condensate pump of a PWR nuclear power plant is investigated. Vibration data was acquired using wireless vibration sensor. The sensor was positioned close to the pump outboard bearings in the direction parallel to pump flow as shown in Figure 4.4. This direction for vibration sensor placement was chosen from operating experience as this direction is the most sensitive to the pump vibration and typically captures the highest amplitudes of pump vibration when compared to other directions such as perpendicular or axial.

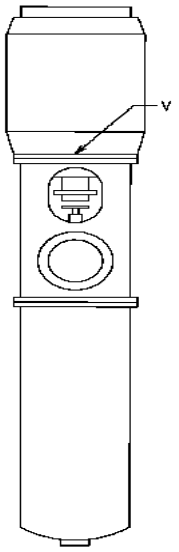

Figure 4.4: Position of vibration senor placement (v) on condensate pump outboard bearing.

In addition to the vibration data, the flow rates from the condensate pump during different operating cycles with various operating modes of the nuclear power plant was also recorded at six different levels with their corresponding motor current. The collected environmental factors and levels are shown in Table 4.1. Initially, these two environmental factors were considered as input variables for which the relationship with the vibration curve obtained at a given time interval was investigated. However, flow rate and motor current are highly correlated, and thus we only consider the flow rate as the independent variable.

Table 4.1: Condensate pump flow rates and corresponding motor current for vibration data acquisition.

	Flow rate (Klbm/hr)	Motor current (Amps)
1	3,641	370
2	3,374	350
3	3,632	368
4	4,435	395
5	3,635	369
6	1,033	269

As the vibration is continuously monitored, the functional instances should be prepared by extracting a sample with a given time window. To do this, each functional instance of vibration signal was obtained per flow rate that contained a total of 128 discretized measurement points and captured within the time window of 123 milliseconds (0.123 seconds). In order to account for replication, three vibration signals were obtained for each level of the flow rate as shown in Figure 4.5, resulting in a total of eighteen functional instances. All eighteen vibration signals possess a similar shape structure, starting from locations where a new cycle is made. These functional data were uploaded into JMP software row-wise in 128 columns along with corresponding flow rate for further analysis.

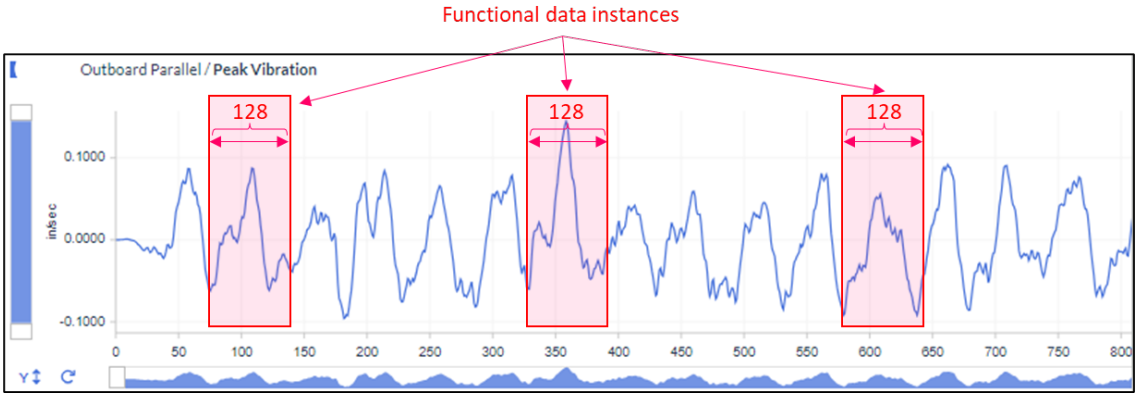

Figure 4.5: Snapshot example of the vibration signal (in/sec – peak) captured in the time waveform at a given time window.

4.5 Transformation of acquired data using FPCA

There is variation in a set of different curves from a data set. FDA utilizes the information obtained from the curves by first converting the discrete members into continuous functions. It then decomposes the variation in their shapes in ways that the functions can be used as responses in a DoE model or as predictors in a machine learning model. FDA allows efficient utilization of information obtained from a curve. This leads to more accurate results and equally saves time and resources.

There are five steps in the standard workflow of FDA, which is shown in Figure 4.6 (Gotwalt and Parker 2022).

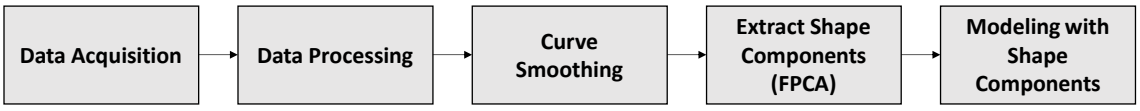

Figure 4.6: Standard workflow of FDA

- **Data Acquisition**: Design and run the experiment. Record and organize response data for each run.
- **Data Processing**: Removes outliers, transformations and aligning or warping functions for data that may need clean up.
- **Curve Smoothing**: Converts discrete data into continuous functions.
- **Extract Shape Components (FPCA)**: Identifies the shapes of functions in the data that characterizes the most variation in the response function. It reduces the functions down to a small number of score values that explain the most variation in the data.
- **Model Shape Components**: Applies the standard DoE analysis tools to these scores just like the regular scalar type DoE.

Using the functional data explorer (FDE) platform in JMP Pro 17, we fit many different wavelets and finds the best fitting one with respect to BIC (Bayesian information criterion), a model selection criterion. Then we identified the mean function and the characteristics shapes that explain variation in functions. The FPC model breaks the data into two kinds of variation. The FPC scores explain function-to-function variation while the eigenfunctions, also known as shape functions, explain the longitudinal variation. We fit the model using the FPC scores as responses. In this study, the extracted FPC scores characterize the variation between the different vibration signals, while the shape function characterizes all the time variation within vibration signals. These FPCs compress the vibration signal data information into the smallest number of scalar quantities possible. We can now treat these scores as regular DoE responses and find models that predict them using the DoE factors.

In this study, because the vibration signal data contains numerous peaks, the wavelet model was used at fitting the curve data. Wavelet modelling is much better at fitting curve data with many peaks when compared to the B-splines or P-splines or the Fourier Basis (Gotwalt and Parker 2022). The model selection results shown in Figure 4.7 suggest a Symlet 20 best fits the data out of 15 other wavelet bases based on any of three criteria, AIC, BIC, and GCV (generalized cross validation). The model with the lowest values of those criteria is the best for curve smoothing.

Model	Wavelet	AICc	BIC	GCV
1	Symlet 20	-23351	-19473.8	1.723366
2	Coiflet 3	-22957.7	-19420.2	2.461613
3	Coiflet 5	-23147.8	-19415.7	2.055773
4	Symlet 10	-23027.8	-19308.1	2.180373
5	Symlet 4	-22413	-19045.6	3.352219
6	Symlet 6	-22598.3	-19040.4	2.850251
7	Daubechies 6	-22562.9	-18935.2	2.799032
8	Daubechies 10	-22718.6	-18785.4	2.180926
9	Daubechies 20	-22752.4	-18688.2	1.92834
10	Coiflet 1	-21510.4	-18179.9	5.032272
11	Daubechies 2	-21489.5	-18159	5.078146
12	Biorthogonal 1.3	-21269.9	-17791.2	5.257088
13	Biorthogonal 4.4	-21019.1	-17681.2	6.210443
14	Haar	-18856.4	-15988.4	18.58562
15	Biorthogonal 2.6	-8912.33	-5255.55	1032.055

Figure 4.7: Wavelet model selection and best fit based on three criteria, AIC, BIC, and GCV.

Figure 4.8 below displays the outcomes of the wavelet model fits by Symlet 20 bases for 18 vibration signals.

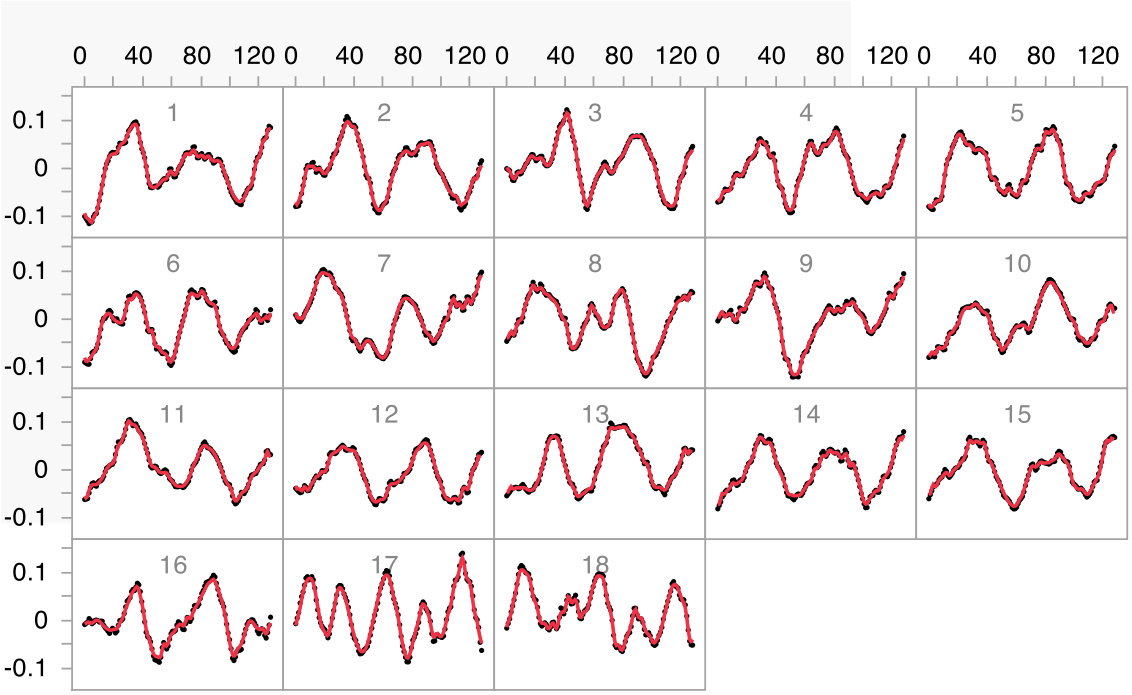

Figure 4.8: Curve smoothing based on Symlet 20 bases – the black dots are the original vibration measurements, and the red solid curves are fitted functions.

The FPCA is now used to investigate the dominant modes of variation of the vibration functional data. The FPCA breaks the data into FPC scores and eigenfunctions in a dimensional reduction technique that is closely analogous to classical PCA. Figure 4.9 shows the overall mean function and the individual shape functions, while Figure 4.10 shows the function summaries of the FPC scores for each functional instance.

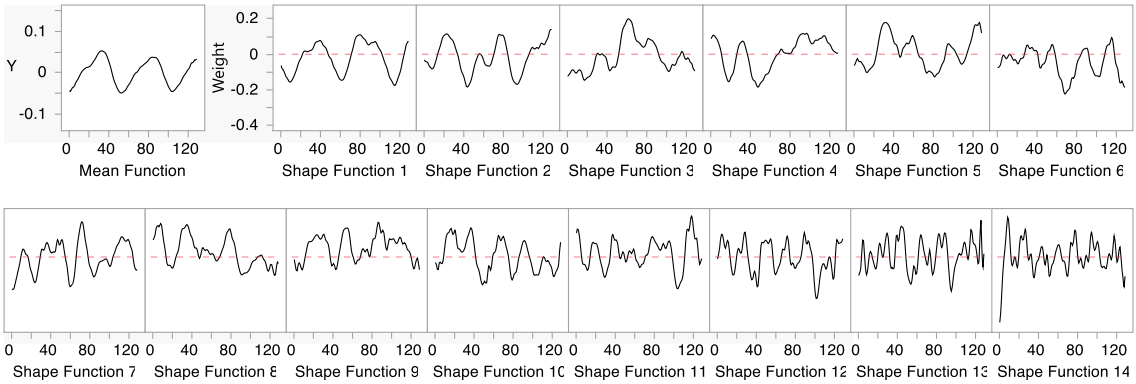

Figure 4.9: Mean function and individual shape functions

Function ID	Validation	FPC 1	FPC 2	FPC 3	FPC 4	FPC 5	FPC 6	FPC 7	FPC 8	FPC 9	FPC 10	FPC 11	FPC 12	FPC 13	FPC 14
1	Training	0.1371654	0.0987118	0.0990666	-0.106309	0.1814903	-0.051966	0.0576911	-0.077484	0.0382386	0.0729922	0.008742	-0.031477	-0.040413	-0.012134
2	Training	0.2447902	-0.223666	-0.088981	-0.041736	-0.041992	0.0255988	0.0820984	0.0952179	-0.007815	0.028998	-0.081138	-0.021072	-0.025379	0.0317874
3	Training	0.169901	-0.328649	-0.141312	-0.101722	-0.020667	-0.112926	-0.07941	0.0037307	-0.005703	-0.015325	0.0112522	-0.045669	0.0157983	-0.036505
4	Training	0.0555744	0.1241216	0.2345542	-0.010879	-0.047501	-0.07578	-0.089107	0.0272837	-0.053957	0.0555233	-0.069914	0.0108798	0.0210584	-0.001278
5	Training	0.1655813	0.1028193	0.0070001	-0.104419	-0.106618	0.0515438	-0.038543	-0.048057	0.0934218	0.0668989	0.0140949	0.0318565	0.02908	0.0084903
6	Training	0.1232033	0.0743194	0.0303309	0.0080672	-0.114522	0.0810924	0.1747677	-0.031904	-0.037662	-0.0029	0.0215686	-0.048331	0.0330937	-0.018307
7	Training	-0.189457	0.3127516	-0.297751	0.0407536	-0.105527	-0.057526	0.0045553	-0.006712	0.0676669	-0.01991	-0.037252	0.0172614	-0.02334	-0.011534
8	Training	-0.209179	0.3439485	0.0333257	-0.193088	7.6123e-6	0.0375359	-0.050937	0.0798837	-0.075168	-0.038666	0.0275679	-0.037738	-0.014898	-0.002268
9	Training	0.0066114	0.0806112	-0.227332	0.2334748	0.0896699	-0.038035	-0.045328	0.0013268	-0.042178	0.0543965	0.0403085	-0.037018	0.0250557	0.0276292
10	Training	0.1545254	-0.01461	0.1251636	0.0555045	-0.034791	0.0201918	-0.068484	-0.110312	0.0258118	-0.095701	-0.011595	-0.047771	-0.011376	0.0358631
11	Training	0.0841055	-0.046545	-0.018764	-0.115424	0.1400266	0.0911969	-0.022295	0.0968094	0.0942047	-0.037509	0.0228066	0.0095478	0.0245012	0.0060598
12	Training	0.1684135	-0.111639	-0.047655	-0.006948	0.0072151	0.0931322	-0.056944	-0.059637	-0.044713	-0.019184	-0.033863	0.0391782	-0.007632	-0.022736
13	Training	0.1237368	0.0074091	0.1942559	0.1432211	-0.028845	-0.161495	0.0651893	0.0648888	0.049293	-0.052422	0.0296028	0.0220846	-0.00076	-0.005548
14	Training	0.0989557	0.0690711	0.0424466	0.0170647	0.0788895	-0.010667	0.0357023	-0.010178	-0.05475	-0.01372	0.0025365	0.0632683	0.0147748	0.0100645
15	Training	0.1004944	-0.013228	-0.120174	0.0278452	0.088	0.0066781	0.0453914	-0.029451	-0.053947	-0.041039	-0.00516	0.0527137	-0.004843	-0.0067
16	Training	0.0505719	-0.126158	0.0822893	0.1668419	-0.098348	0.0958577	-0.055576	0.045844	-0.008223	0.0422828	0.0748935	0.0163006	-0.043972	-0.010165
17	Training	-0.675272	-0.100863	0.1021118	0.1684821	0.072759	0.0705136	0.0056873	0.0014877	0.0397546	0.0031817	-0.055986	-0.022069	0.011115	-0.019194
18	Training	-0.609721	-0.248405	-0.008575	-0.180731	-0.059247	-0.064945	0.0355415	-0.042739	-0.024276	0.0121035	0.0415342	0.0280552	-0.001863	0.0264748
Target	Validation	0.0031311	0.0041485	-0.000223	-0.000437	0.0021939	-0.000757	0.0006712	0.0012405	-0.000894	0.0032647	0.000536	0.0010767	0.0016262	0.0016384

Figure 4.10: Function summaries of the FPC scores for each functional instance.

The FPC scores combined with the shape functions model each functional instance. That is, a functional instance = Overall mean + (FPC 1 Score × Shape Function 1) + (FPC 2 Score × Shape Function 2)+ ⋯ + (FPC 14 Score × Shape Function 14). The eigenvalue plotted in Figure 4.11 shows how much the individual shape functions explain the shape variation. We can see that 14 FPCs (shape functions) explain 99.7% of the shape variation. The first FPC (FPC 1) contributes the largest percentage (41.3%) of the variation and does the second FPC (FPC 2) the second largest (18.9%), so on. Figure 4.12 shows the model selection criterion with different number of FPCs where 14 FPCs produce the best BIC value.

FPC	Eigenvalue	20 40 60 80	Percent	Cumulative
1	0.06811		41.3%	41.3%
2	0.03115		18.9%	60.2%
3	0.01910		11.6%	71.8%
4	0.01499		9.09%	80.8%
5	0.00795		4.82%	85.7%
6	0.00584		3.54%	89.2%
7	0.00479		2.9%	92.1%
8	0.00352		2.13%	94.2%
9	0.00287		1.74%	96%
10	0.00214		1.3%	97.3%
11	0.00172		1.04%	98.3%
12	0.00134		0.81%	99.1%
13	0.00055		0.34%	99.5%
14	0.00041		0.25%	99.7%

Figure 4.11: Eigenvalue and the proportion of variability explained by each FPC.

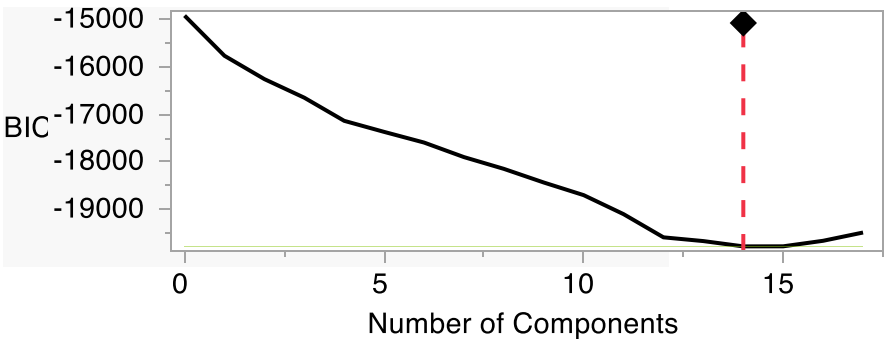

Figure 4.12: Model selection outcome with different number of FPCs

The FPCA diagnostic plot in Figure 4.13 shows the actual by predicted plot for every vibration measurement from all 18 functional instances. The plot compares the actual measurements with those reconstructed by 14 FPCs. The plot shows a good fit as the points are close to the line with narrow deviations.

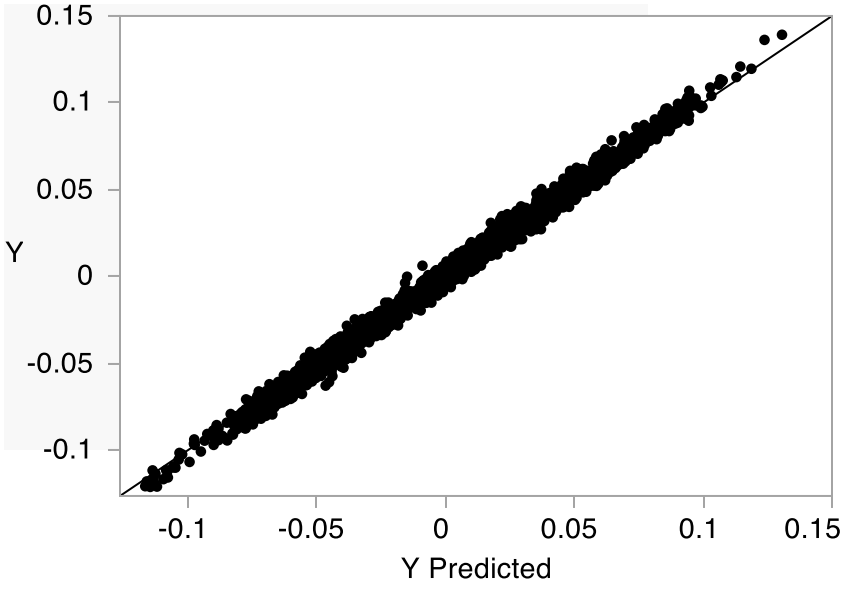

Figure 4.13: FPCA diagnostic – Actual vs. Predicted plot.

The score plot is a visual representation showing the projection of the FPC scores for each vibration function into two-dimensional plane. Figure 4.14 shows the score plot if we choose FPC 1 and FPC 2 as two variables. It is observed that vibration function IDs #7 and #8 are clustered together and away from the rest of other vibration functions. Also, vibration function IDs #17 and #18 are clustered together and away from the rest of other vibration functions. The reason for that can be observed from Figure 4.8, where it is shown the mentioned function IDs have multiple peaks for the same number of vibration measurements within the fitted functions. Vibration function IDs #7 and #8 have similar number of multiple peaks and are clustered together from the rest. Vibration function IDs #17 and #18 have similar but most prevalent number of multiple peaks and are also clustered together from the rest.

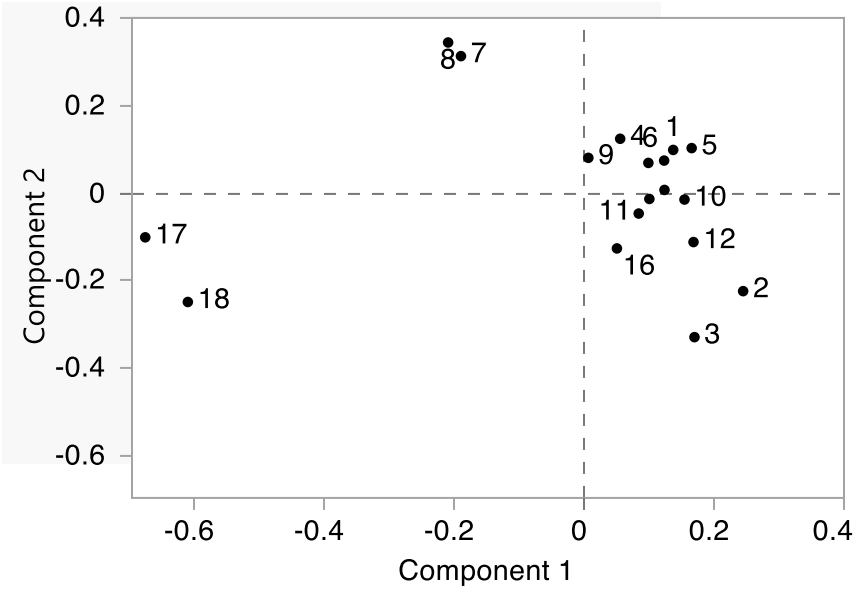

Figure 4.14: Score plot of FPC 1 vs. FPC 2 scores for 18 functions

The FPC Profiler is a visual representation showing the values of the FPC scores associated with a particular vibration Function ID and the direction the FPC scores moved when compared to the mean response function. Figure 4.15 shows the FPC Profiler for the example for vibration function ID #3.

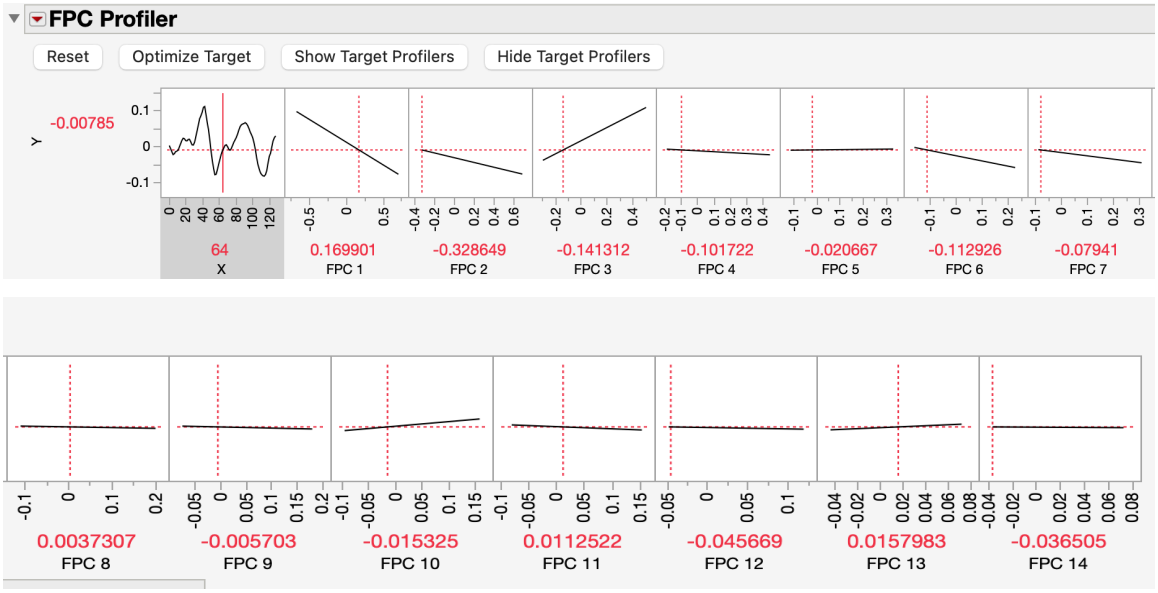

Figure 4.15: FPC Profiler for vibration Function ID #3

Figure 4.16 shows the mean vibration response function showing the FPC scores as zero (0) values.

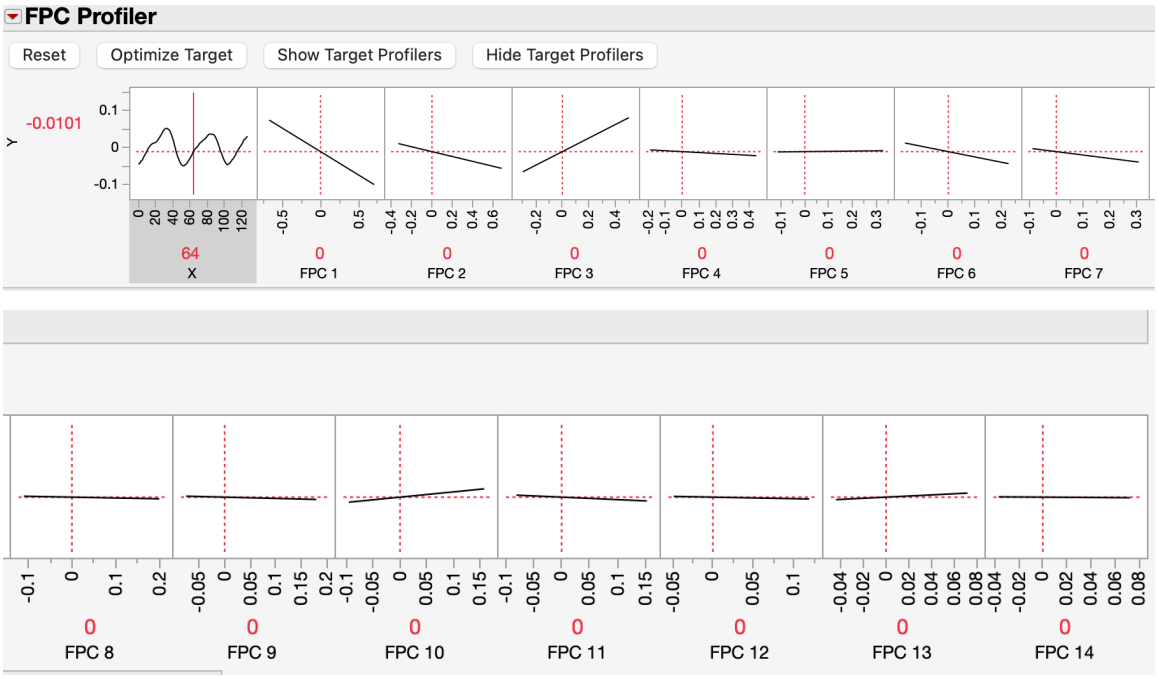

Figure 4.16: Mean vibration response function showing the FPC scores as zero values.

4.6 Optimization of environmental factor using FDOE

The goal of functional design of experiments (FDOE) is to identify a relationship between controllable factors and the functional response (the curve). We fit the FPC scores as functions of the DOE factors. Using (FPC score) × (eigenfunctions) as the intermediate formula, the (modelled FPC score) × (eigenfunctions) is used as the final prediction formula.

In this study, a target function was added in the analysis. This target function represents the ideal vibration curve. In this research, the mean function of 18 curves was

used as the target function because it provides smooth and noise reduced shape of a vibration curve. The shape of the target vibration curve or the mean function as shown in Figure 4.9 appears like a sinusoidal waveform in shaped with reduced amplitudes. In vibration analysis, a sinusoidal time waveform typically corresponds to the operating speed of the equipment when transformed in the Fast Fourier Transform (FFT). If the vibration amplitudes are not significantly high, then it represents normal equipment (Eshelman 1999). Therefore, the shape and amplitude of the target vibration curve obtained in the FDOE analysis is acceptable and reasonable for use in this research application.

The JMP software was used to find the factor settings that will closely match the target vibration function or curve. This was obtained by using the FDOE Profiler – Optimization and Desirability – Maximize Desirability. From the FDOE analysis results, the optimized flow rate that will match the target vibration curve corresponds to a target flow rate value of 3308.5 Klbm/hr as shown in Figure 4.17.

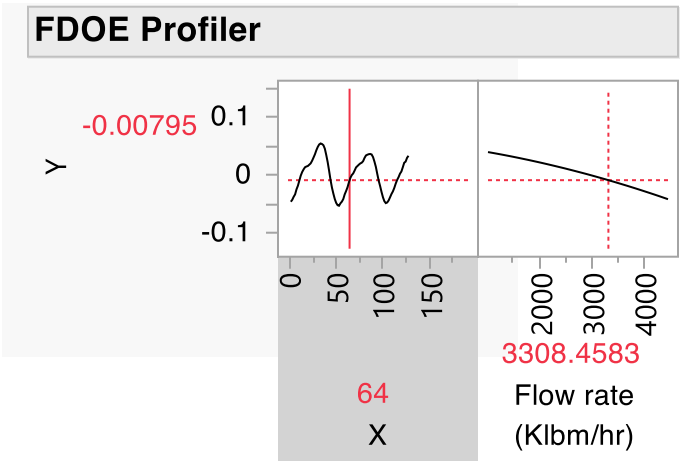

Figure 4.17: Optimized flow rate (3308.5 Klbm/hr) corresponding to the target vibration curve.

Figure 4.18 shows an example of shape of the vibration curve when the flow rate value is reduced lower than the target flow rate value (2870 Klbm/hr). Figure 4.19 shows an example of shape of the vibration curve when the flow rate is increased higher than the target flow rate value (3891 Klbm/hr). In Figure 4.18 example, the reduced flow rate (2870 Klbm/hr) when compared with the target or optimized flow rate (3308.5 Klb/hr), the observed the shape of the vibration curve appear clipped or truncated and no longer sinusoidal (non-harmonic) in shape. However, in Figure 4.19, with flow rate increased (3891 Klbm/hr), the observed shape of the vibration curve still appears sinusoidal in shape. The Interpretation of these observations.is discussed under the 'Results and Discussions' section of this chapter.

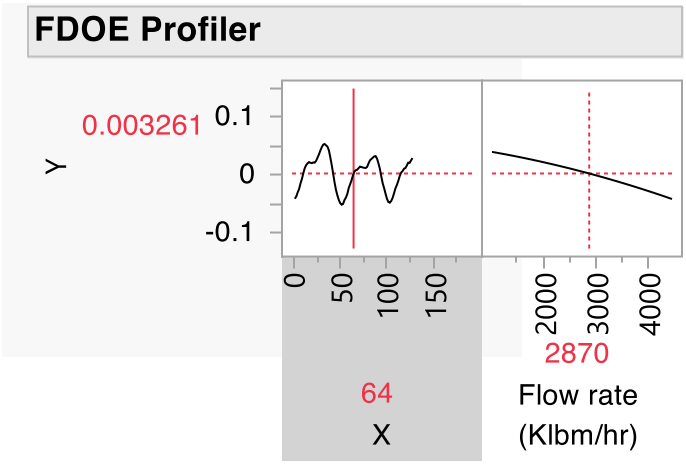

Figure 4.18: Reduced flow rate (2870 Klbm/hr) vibration curve example

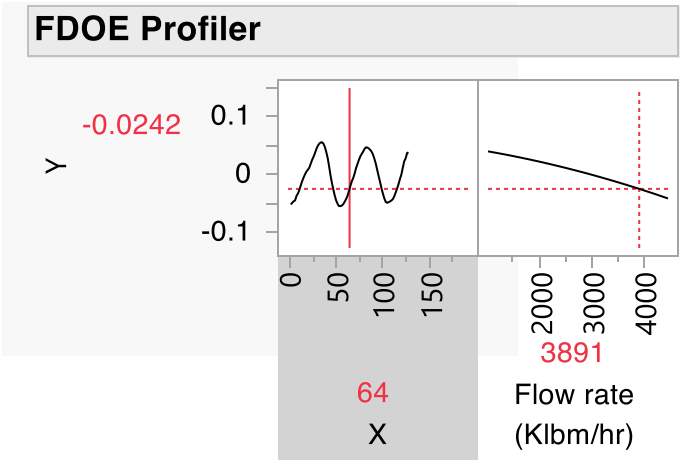

Figure 4.19: Increased flow rate (3891 Klbm/hr) vibration curve example

4.7 FDOE analysis for regression plots

For the FPC 1, the FDOE regression plot in Figure 4.20 shows an increasing upward slope as you go from 1,000 Klbm/hr flow rate to 4,500 Klbm/hr flow rate, with the corresponding distribution of the vibration functional data. It shows a generalized R-square value of 0.5. This shows the FDOE model can explain about 50% of the variability of the outcome data using FPC 1.

Similarly, For the FPC 10, the FDOE regression plot in Figure 4.21 shows a curvilinear relationship as you go from 1,000 Klbm/hr flow rate to 4,500 Klbm/hr flow rate, with the corresponding distribution of the vibration functional data. It shows a generalized R-square value of 0.32. This shows the FDOE model can explain about 32% of the variability of the outcome data using FPC 10.

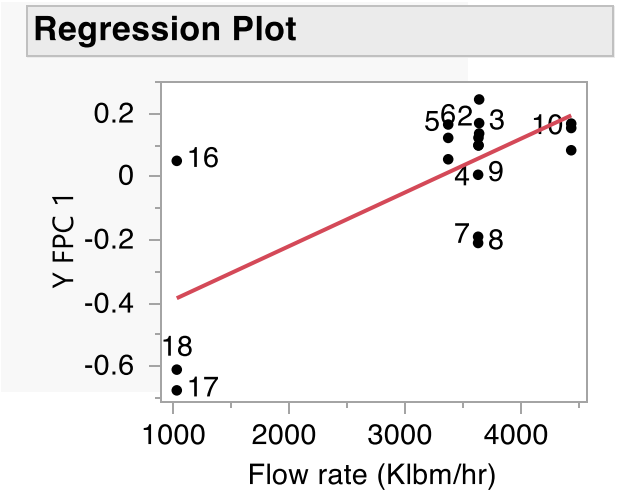

Figure 4.20: Generalized regression plot for FPC 1.

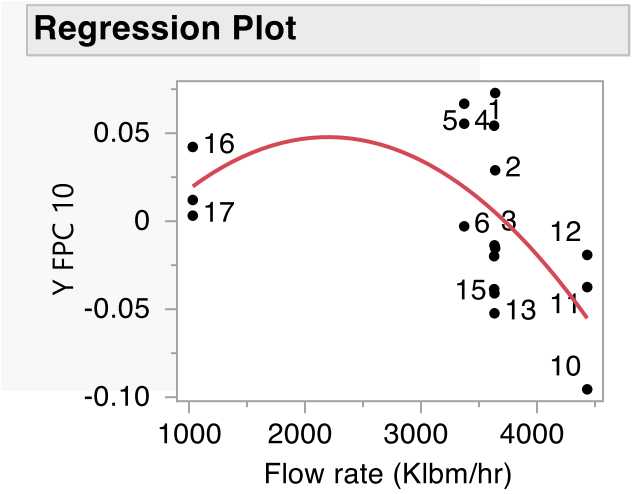

Figure 4.21: Generalized regression plot for FPC 10.

4.8 Results and discussions

The original pump curve for the condensate pump used in this study was reviewed. From the pump curve as shown in Figure 4.22, the pump duty point is typically the rated flow and head at which the pump is designed to operate. The duty point flow rate is (3598.5 Klbm/hr (7,200 gpm). The optimized flow rate from Figure 4.17 which corresponds to the

target vibration curve is 3308.5 Klbm/hr (6619.6 gpm). A pump's best efficiency point (BEP) is the best operating point where the possibility of energy waste, unwanted forces, vibration, and noise are at minimum. Therefore, the BEP is closely related to the highest theoretical reliability and lowest possible risk or maintenance for the pump. A pump running at its BEP is considered optimal in performance; however, this is typically not the case for practical pump applications.

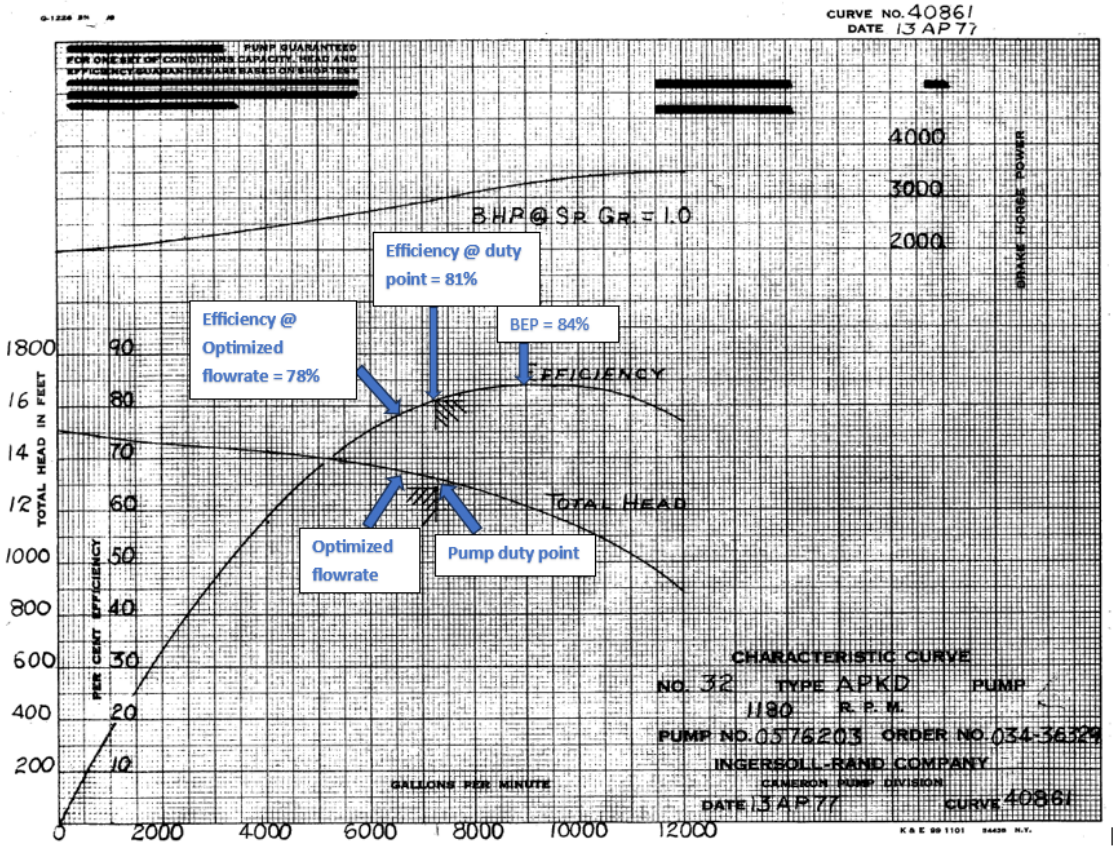

Figure 4.22: Original pump curve showing optimized and duty point flow rates and efficiencies comparison with BEP (Ingersoll-Rand 1977).

From Figure 4.17, the optimized flow rate of 3308.5 Klbm/hr (6619.6 gpm) corresponds to the target vibration curve. The target vibration curve appears similar to a sinusoidal wave form in shape, and typically indicates a pump operating at its operating speed (one frequency). This optimized flow rate obtained from the FDOE indicates a pump efficiency of 78% when compared to duty point and BEP efficiencies of 81% and 84% respectively. In terms of pump performance, the optimized flow rate of 6619.6 gpm is acceptable for practical application. The pump efficiency from the optimized flow rate is within 3% and 6% when compared with the duty point and BEP efficiencies respectively. The results show that with the optimized flow rate, the pump can operate within the bounds of optimal performance where vibration, energy waste and unwanted forces are minimized.

From Figure 4.18, it was observed as pump flow rate is reduced at 2870 Klbm/hr (5742.3 gpm) from the recommended optimized flow rate value, the vibration curve appears clipped or truncated, thus losing its sinusoidal waveform appearance. In vibration analysis, a truncated waveform typically would indicate a generated or second higher order (multiple frequencies). This may indicate some type of equipment fault.

From Figure 4.19, it was observed as pump flow rate is increased at 3891 Klbm/hr (7785 gpm) from the recommended optimized flow rate value, the vibration curve still appears sinusoidal in shape. This shows the flow rate although above the duty point flow rate, 3598.5 Klbm/hr (7,200 gpm) is still below pump BEP. The higher flow rate although still acceptable with vibration performance is operating at a higher flow rate than pump design. This would not be recommended for long term pump operations.

Some of the short comings from the regression plots as shown in Figures 4.20 and 4.21 could be attributed to the single factor flow rate that was used in the research study.

The data was obtained from an operating power plant. Motor current could not be used as an input factor because of the strong correlation this factor has with pump flow rate. This is because higher pump flow rates tend to draw more motor current and lower flow rates will draw lesser motor current.

4.9 Conclusions

The following conclusions were reached from this study:

1. Wavelets was applied to smoothen the functional vibration data. It was demonstrated that wavelets performed well in modelling the various peaks in the time waveform vibration data and converting these discrete data into continuous functions.
2. A FPCA was applied to successfully characterize the variation in the vibration response function.
3. A FDOE analysis was applied to the target vibration curve and used to obtain an optimal flow rate value of 3308.5 Klbm/hr (6619.6 gpm).
4. The optimized flow rate pump efficiency (78%) falls within the bounds of BEP efficiency for optimal pump performance.
5. The shape of the target vibration curve appears sinusoidal and reflects a pump's operating speed with minimal harmonics.
6. Operating the pump at lower pump flow rates 2870 Klbm/hr (6619.6 gpm) when compared to the optimal flow rate reflects a truncated shape in vibration curve. This would indicate some type of harmonics and potential equipment issues and

therefore flow rates below 2870 Klbm/hr is not recommended for the condensate pump operations as applied in this study.

Chapter 5

Applications of Functional Design of Experiments: A Review

5.1 Introduction

Functional design of experiments (FDOE) applies to experiments where the response variable takes the form of a curve instead of a single value. In FDOE, we are interested in knowing how the factors of interest might affect the features of a response curve we are interested in. Also, based on the response curve shape, we may be interested in how we can optimize the response variables (curves) by manipulating the factors, and be able to model the relationship between the two. Examples of practical applications include: Fidaleo (2020) applying various rotating speed and properties of a ball mill (factors) to determine fineness and energy (functional responses); (Oguejiofor and Seo) applying different pump flow rates (factor) to determine vibrations (functional response) in a power plant; (Metusalem 2020) investigating different material properties (factors) and the effect on the load-deflection curve (functional response) and etc. FDOE also find applications in various disciplines such as engineering during optimizing process designs, process parameters and optimization; pharmaceuticals and healthcare during drug formulation optimization and clinical trial design; agriculture during crop yield optimization and precision agriculture; and social sciences and psychology during behavioral studies and survey optimization.

Functional data occur when we obtain or measure data over time. In many applications and disciplines, it is important to consider the entire curve as an output instance of a system, instead of an aggregated statistic, e.g., the average, from a curve. In both the traditional design of experiments (DOE) and functional design of experiments (FDOE), a systematic approach is used to study the relationship between input variables (factors) and the output variables (responses). In FDOE, the response is in a functional form or a curve. The shape of curve we measure is important because we want to understand how the curve shape is affected by certain factors. Advancements in research and data acquisition have made functional data analysis (FDA) very attractive and the application of FDOE could play a key role in understanding and analyzing such data.

This study first presents a general framework for FDOE and outlined the various steps to be followed during its application to functional data using the approach utilized in JMP software. FDOE is an emerging area in statistical experimental design and analysis, with lack of scholarly practical applications found in literature. FDOE could play an important role in future research involving analyzing and characterization of functional data. Despite these clear benefits, adequate understanding and recognition of the key features and importance of FDOE have been limited to date. This review paper will reveal the importance of FDOE and stimulate its application in various fields of study.

5.2 General FDOE Framework

The below standard workflow shows five steps used in FDOE. This be applied as a general framework using JMP Pro 17 software and shown in Figure 5.1 below.

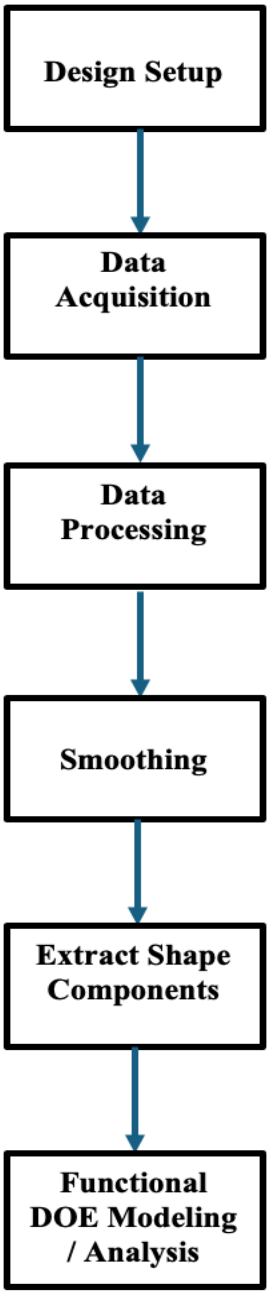

Figure 5.1: General FDOE framework

Functional data is different from the traditional experimental design because you are dealing with a data set or experimental run where you do not have a singular value for the response. Instead, you have a form of continuous data from measurements that have

been recorded over time. The goal of the FDOE is to relate the shape of this functional data to other variables and find the best settings of your input factor which will match the target curve you are expecting.

5.2.1 Design Setup

The first step in FDOE is setting up your experimental design. JMP Pro 17 DOE platforms could help design, evaluate, and analyze experiments. Most of the platforms focus on constructing designs while other platforms support the design effort. Any of the DOE custom design platforms such as random block, factorial, split plot, etc. can be used to construct the experimental design. The definitive screening design from JMP Pro 17 has also been found to support functional response data input. The custom design platforms would help produce a table from your chosen experimental design to record functional data. Alternatively, the experimental design style chosen can be constructed in an excel or csv file. The uploaded functional data in JMP can be arranged in a data table using the Functional Data Explorer (FDE). This can be accomplished row-wise using 'row as functions', or column-wise using 'column as functions', or also in a 'stacked data format' where two or more columns are stacked into a new single column.

5.2.2 Data Acquisition

After the experimental design set up, the next step is to run your experiment by acquiring the functional data. The functional data should be organized based on your design set up and be ready for using the FDE. Functional data is collected through various means. In most FDOE applications, data is collected over a period using sensors. Sensors gather and generate information based on the physical conditions surrounding it. Sensor

data collection usually involve the following: a processor to convert physical signals into digital data, a communication capability to transmit data to people or machines, and a power source. Some sensor examples and types of sensor data processing techniques below offer insights into the diversity of their applications:

Accelerometers*:* can sense the tilt or movement in a particular direction such as vibration or impact.

Photosensors*:* can detect the presence of visible light, infrared transmission, or ultraviolet energy.

Temperature sensors**:** measure the amount of heat generated from an area of object such as temperature.

Infrared sensors**:** light-based sensors used for proximity and detection.

Gyroscope*:* capture the speed and rotation around an axis of an object.

Following data acquisition and results documented in the FDE data table, the functional data will then be processed for further analysis.

5.2.3. Data Processing

A significant problem with most functional data analyses is that of misaligned curves. Without adjustment, even an analysis as simple as estimation of the mean will fail (James 2017). Unprocessed data may contain human, machine or instrument errors. It might also be in a different format, some entries might be suspect (outliers), out of date, poorly structured, large in volume and highly complex. To avoid such issues, data processing plays an important role in filtering data to ensure high quality prior to further

processing. Data processing will involve data cleaning to address irrelevant data, duplicate data, noisy data, missing values, etc. Part of the pre-processing will involve data transformation to remove outliers and unacceptable format. Additionally, depending on data type, a data reduction may be necessary for data with too many dimensions or categories.

The FDE platform in JMP Pro 17 contains various data processing options to address some of the issues identified above for unprocessed data. The FDE platform data processing can be used to clean up data such as removing observations with zero values, filtering certain x or y values that fall within certain intervals, transforming the data through standardization, square root, logarithms or exponential. A dynamic time warping (DTW) can also be used as a function alignment technique that finds an optimal warping to align two or more functions together. For spectral data, some options available in JMP Pro17 include:

- *Standard Normal Variate (SNV) Method:* Standardizes each function with their individual means and variances. This is an alternative to standardize which uses global mean and variance.
- *Multiplicative Scatter Correction (MSC):* Applies a simple linear regression to fit each individual function (level of the ID variable) where the response is output values for the function and the regressor is the output values for the mean function. The MSC corrects light scatter for each function to the light scatter of a reference point.
- Savitzky-Golay Filter: This method fits local polynomials to several collections of points across the domain. The polynomials are fit using least squares and the

number of points in each fit is determined by the bandwidth. When you select this option, several fits are made for polynomials of order 0, 1, and 2 and bandwidths up to 10. The best fitting models for each function are selected based on the Akaike information Criterion (AIC). The order of the polynomial and the bandwidth can be different for each function.

- Baseline Correction: Subtracts a baseline function from each individual function. A baseline correction is used when there is a known trend, or baseline, that you want to remove. For example, this could be due to an artifact of how the data is measured. Usually, the information is in the peaks of the data, so these regions are not included in the baseline model. (Gotwalt and Parker 2022, JMP 2023).

5.2.4 Curve Smoothing

Smoothing techniques play an important role in FDA as they provide insight in the functional behavior of stochastic process. In some cases, a summary statistic of this function, such as the average response or the maximum response will be sufficient for analyzing the system under study. However, there are many other occasions when the experimenter desires to know the entire response curve to understand the dynamics of the system under study (Pan and Saleh 2019). Also, in statistics and image processing, to smooth a data set is to create an approximation function that attempts to capture important patterns in the data, while leaving out noise or other fine scale structures. Smoothing is the first step in any FDA, and its purpose is to convert raw discrete points into a smoothly varying function. This emphasizes patterns in the data by minimizing short-term deviations due to observational errors, such as measurement errors or inherent system noise (Ullah and Finch

2013). Various models exist in JMP Pro 17 Functional Data Explorer (FDE) application to fit functional data or a response curve. Such models include B-Splines, P-Splines, Fourier Bases and Wavelets. The smoothing steps convert discrete measurements into continuous functions. FDE applies the chosen smoothing method and finds the best fitting curves. The spline models use the number of knots, the degree of the polynomial with the lowest Bayesian Information Criterion (BIC) to find the spline with the best fit. The Fourier bases is based on number of Fourier bars and uses the BIC. The wavelet model applies different wavelet types and will find the wavelet with the lowest BIC as the best fit. A brief background discussion on the smoothing models found in JMP Pro 17 is explained below.

B-Splines: A piecewise polynomial curve and the knots are the points where the pieces meet. It is typically used for non-periodic data. B-splines are a powerful tool commonly used to model smooth functions. They can be used in settings where we wish to fit a curve to data without making strong functional form assumptions. These curves are called B-spline functions. These functions are defined in terms of a set of basis functions. A common approach towards B-spline smoothing is to construct a large number of knots as the smoothing parameter to reduce the effective degrees of freedom and increase smoothness in the overall function estimate (Ullah and Finch 2013). B-splines have found several applications in modeling of functional data as shown in below review of selected literature.

Houhou et al. (2021) used B-splines basis function to investigate Raman spectral data in the functional framework. Kramer et al. (2008) applied a combination of B-spline basis function and partial least squares (PLS) to benefit a noisy functional data. Kermarrec

et al. (2020) approximated the distance between two-point clouds using B-spline surfaces. The study was able to determine which distance is better suited in case of heteroscedastic and correlated terrestrial laser scanner (TLS) observations for local deformation analysis. Wang et al. (2022) utilized tensor network B-spline (TNBS) to model the nonlinear autoregressive exogenous (NARX) system to alleviate the computation and store burden for processing high-dimensional systems.

Shahmohammadi et al. (2020) applied the iso-geometric B3-spline and Langrangian interpolation along the longitudinal and transverse directions to analyze the vibration of shell panels made of functionally graded material (FGM). Hoek et al. (2021) presented a method that utilized B-splines to generate cooperative trajectories to retain the versatility of an autonomous vehicle.

P-Splines: P-splines or penalized splines are regression splines fit by least squares with a roughness penalty. P-splines have much in common with smoothing splines, but the type of penalty used with a p-spline is somewhat more general than for a smoothing spline. Also, the number of locations of the knots of a P-spline is not fixed as with a smoothing spline. Generally, the knots of a P-spline are at a fixed quantities of the independent variable and the only tuning parameters to choose are the number of knots and the penalty parameter (Ruppert 2022). P-splines have found several applications in modeling of functional data as shown in below review of selected literature.

Aguilera-Morillo et al. (2020) proposed a penalized version of functional linear regression (FLM) by introducing a P-spline penalty in the least square fitting criterion. Ruppert (2022) used two algorithms, the myopic and full search, to study the effects of the

number of knots on the performance of P-splines. Alejandra-Hernandez et al. (2023) illustrated the use of P-spline models to estimate derivatives in the context of longitudinal data. The study also proposed a new penalty acting on the population and the subject specific levels to address under-smoothing and boundary problems in derivative estimation. Gressani et al. (2022) used a P-spline approximation baseline hazard with a logistic regression model to specify the conditional survival function of susceptible subjects. Maturana-Russell and Meyer (2021) proposed a Bayesian approach to estimating the spectral density of a stationary time series using a prior based on a mixture of P-spline distributions. Xu et al. (2022) applied a P-spline method to fit the functional data related to the original time series data. The study used the obtained smooth-fitting data as input of the clustering algorithm to enhance the ability to process data.

Fourier Basis: The Fourier basis function is a method to smooth out data varying over a continuum and exhibiting a cyclical trend. Fourier smoothing works by taking a dataset and decomposing it into series of sine and cosine waves using a technique called Fourier transform. The high frequency components, which represent noise in the dataset are then removed, resulting in a smoother signal.

The main advantage of Fourier analysis is that very little information is lost from the signal during the transformation. The Fourier transform maintains information on amplitude, harmonics, and phase and uses all parts of the waveform to translate the signal into the frequency domain. Preservation of phase information by the Fourier transformation means that the signal can be transformed back into the time domain. The major disadvantage of the Fourier transformation is the inherent compromise that exists between

frequency and time resolution (Parsons et al. November 2000). The Fourier basis have found several applications in modeling of functional data especially for vibration data.

Wavelets: A wavelet modeling capability was newly added to JMP Pro 17. Wavelet modeling have been found much better at fitting curve data with varying peaks than the splines or Fourier methods (Worley 2022). Wavelets have found several applications in modeling of functional data as shown in below review of selected literature.

Michis (2017) applied a wavelet smoothing method to improve conditional forecasts generated from linear regression response models. The study found that when compared to a moving average and band-pass filter, the wavelet smoothing method provided the best results when applied on highly volatile marketing time series. Lu et al. (2020) proposed a hybrid model based on wavelet threshold denoising (WTD) and recursive feature elimination with cross-validation (RFECV) to improve measurements in quantitative analysis of coal properties using laser-induced breakdown spectroscopy (LIBS). Volvach et al. (2024) applied wavelet decomposition method and local smoothing to filter high frequencies and eliminate the influence of outliers in studying the spectral composition of local average annual surface air temperature with solar activity. Vanus et al. (2023) employed a smoothing procedure based on wavelet transformation to reduce the inaccuracies in a predicted signals data. Zhao et al. (2024) utilized wavelet transform to identify outliers in correcting abnormal velocities and direction angles in a three-step vehicle trajectory reconstruction. The study also applied a Savitzky-Golay filter to smoothen the trajectories. As a result of problems from poor real-time predictability of oil operation status and difficulty in offline interpreting the played back data, (Feng et al.

2020) proposed a wavelet-based Kalman smoothing method for processing uncertain oil well-testing data.

5.2.5. Extract Shape Components

Following the smoothing step, FDE automatically performs a Functional Principal Component Analysis (FPCA). This is identified using a Bayesian Information Criterion (BIC) based on the chosen smoothing model. The FPCA shows how much percentage of variation of the data is explained by a single mean function and an identified number of individual shape functions. The first shape function typically explains the most variation in the data, followed by the second shape function and so on. The Actual vs Predicted plot can be used to determine the dimensionality we would want to work with. In general, because we want our predictions to be as close as possible to the actual data, it is good if the Actual vs Predicted plot follows the 45 degrees line. The FPCA model breaks the data into two kinds of variation. The shape function characterizes the time variation between the response functions while the extracted Functional Principal Component (FPC) scores characterizes the variations between the different functions under investigation (examples could be batch to batch variation, vibration response at different flow rates). The shape functions can be used to understand the curve-to-curve variation in the data around the mean. They indicate the shape features of the data curves. The basis function model can be rewritten as a linear combination of shape functions. The coefficients of the shape functions are the functional principal component scores. These functional principal components are new scalar variables that contain the shape information in the data (Michelson 2024). The FPC will compress the response function information into the

smallest number of scalar quantities possible. The scalar scores are treated as regular DOE functions and this process is called FDOE analysis. JMP Pro 17 will perform this modeling and connect resulting DOE models for the shape scores to the shape functions (Gotwalt and Parker 2022).

5.2.6. Functional DOE Modeling / Analysis

The FPC scores help capture the shape of the curve for each functional data response. JMP PRO 17 FDOE Analysis Functional Data Explorer can be used to build a model that relates the experimental factors to these shape scores. JMP will use a generalized regression platform to build a model to relate the experimental factors to the shape scores. The FDOE Profiler helps visualize how changes in our experimental factors affect the shape of the response curve by these shape scores. Basically, by manually adjusting the individual factors, the shape of the response curve that serves best for the experimental application can be determined. Alternatively, if a target function or curve is included in the initial analysis, the FDOE Profiler can find the factor settings that will closely match the target curve.

5.3. Practical Applicability of FDOE using JMP

FDOE is an emerging field in experimental design with scarcity of publications found following review of literature. The insight gained from this study is that JMP Functional Data Explorer with FDOE Analysis option has made functional data analysis much easier. Researchers can plan their functional experiments in advance using the

FDOE framework as outlined in this paper. The conclusions from conducting such experiments can show how each factor or factor settings can influence the shape of the functional response curve. Following a review of literature, one publication was found demonstrating practical applicability of FDOE using JMP. The other research study on FDOE is part of this PhD dissertation and have been submitted to a journal for review.

Fidaleo (2020) used a combination of Functional Data Analysis (FDA) and DOE to obtain a dynamic design space for the batch milling process of a hazelnut and cocoa based paste in a stirred ball mill. The designs setup involved three factors (shaft speed, ball diameter, overall ball mass) at three levels (low, center, high), with two response variables, Energy and Fineness. Data acquisition was obtained through calculation and direct measurement (micrometer). A natural logarithm transformation was used as part of the data processing. B-splines were used for smoothing. The FPCA scores were modeled as a function of the experimental factors through response surface models. This was used to build a final model for the functional responses including one principal component for energy and two for fineness. The developed models were able to predict with good accuracy the functional responses as a function of the experimental factors.

Oguejiofor and Seo (2024) applied the FPCA in characterizing the vibration signals generated under different levels of flow rates, an environmental factor, of the condensate pump. The designs setup involved one factor (pump flow rate) at six different levels, with one response variable, pump vibration. Data acquisition was obtained by sensors attached to the pump. No processing was required for the data, and wavelet was used for smoothing. A FDOE was used in combination with a target vibration curve to obtain an optimal flow rate. The obtained flow rate was found comparable to the theoretical pump curve best

efficiency point (BEP) and recommended for optimal pump performance and reliability. Table 5.1 below shows the JMP FDOE framework methodology comparison for the two studies.

Table 5.1: JMP FDOE framework methodology comparison

	Fidaleo	**Oguejiofor & Seo**
Design Setup	Three factors at three levels with two response variables. Face-centered central composite design. 8 total runs.	One factor at six levels with three replications. One response variable. 18 total runs.
Data Acquisition	By calculation and direct measurement	By sensors
Data Processing	Natural logarithm transformation	Not required
Smoothing	B-splines	Wavelet
Extract Shape Components	FPCA	FPCA
Functional DOE Modeling / Analysis	Dynamic design space for the batch milling process	Target vibration curve with optimal pump flow rate.

5.4. Conclusion

FDOE can play a key role in future research in analyzing and characterizing functional data. This review demonstrated a larger context of applicability of FDOE not just for academic work but also its usefulness especially in industrial application. It is hoped that this review would help create awareness and encourage other researchers advance their work as related to the use of FDOE with its numerous applications.

Chapter 6

Conclusions and Future Work

This Chapter provides concluding remarks on this dissertation and future work. In Chapter 2, we investigated and applied DWT decomposition of vibration signals from a reactor coolant pump. Certain statistical features were extracted from the decomposed coefficients. The PCA was used to determine which coefficients had more variability from the reduced PC variables. In Chapter 3, we applied the insights gained from the study using PCA to simulate the normal vibration data and three categories of abnormal vibration data. DWT was used to decompose the signals with some statistical features extraction. We applied the LDA in distinguishing between the classes for normal and abnormal vibration data. In Chapter 4, we applied the FPCA to breakdown and reduce the dimensionality of the vibration signals into few empirical basis functions called FPCs. We also applied the FDOE to identify the effects of pump flow rates (environmental factor) in determining the optimal shape of the vibration curve. In Chapter 5, we performed a comprehensive review of FDOE and its applications. A standard workflow was shown as a general framework for FDOE using JMP Pro 17.

The LDA classification combined with DWT and statistical feature extraction showed good results from our study in distinguishing various categories of vibration data. For future work, we propose comparing our methodology with other popular classification techniques used for supervised learning problems such as K-nearest neighbour (KNN), support vector machines (SVM), naïve bayes, decision trees, etc. This can help determine

which method works best for practical applicability especially for power plant vibration monitoring of rotating equipment.

BIBLIOGRAPHY

Aguilera-Morillo, M., et al. (2020). "Variable selection with P-splines in functional linear regression: Application in graft-verus-host disease." Biometrical Journal **62**(7): 1670-1686.

Ahmed, Z., et al. (2021). "A Novel Framework for Centrifugal Pump Fault Diagnosis by Selecting Fault Characteristic Coefficients of Walsh Transform and Cosine Linear Discriminant Analysis." IEEE Acess **9**: 150128 - 150141.

Alejandra-Hernandez, M., et al. (2023). "Derivative curve estimation in longitudinal studies using P-splines." Statistical Modelling **23**(5-6).

Althubaiti, A., et al. (2021). "Fault diagnosis and health management of bearings in rotating equipment based on vibration analysis – a review." Journal of Vibroengineering **24**(1): 46-74.

Bajaj, N. S., et al. (2022). "A Bayesian Optimized Discriminant Analysis Model for Condition Monitoring of Face Milling Cutter Using Vibration Datasets." ASME Journal of Nondestructive EValuation, DIagnosis and Prognosis of Engineering Systems **5**(2).

Bendjama, H. and M. S. Boucherit (2016). "Wavelets and principal component analysis method for vibration monitoring of rotating machinery." Journal of Theoretical and Applied Mechanics.

Bendjama, H., et al. (2015). Selection of Wavelet Decomposition Levels for Vibration Monitoring of Rotating Machinery. . The Ninth International Conference on Advanced Engineering Computing and Applications in Sciences, Nice, France.

Biswas, R., et al. (2023). "Functional data analysis to characterize disease patterns in frequent longitudinal data: application to bacterial vaginal microbiota patterns using weekly Nugent scores and identification of pattern-specific risk factors." BMC Med Res Methodol **23**(1): 251.

Chang, F., et al. (2022). "Assisted Selection of Biomarkers by Linear Discriminant Analysis Effect Size (LEfSe) in Microbiome Data." Journal of Visual Experiments (JoVE).

Chen, Y., et al. (2023). "Gaussian assumptions-free interpretable linear discriminant analysis for locating informative frequency bands for machine condition monitoring." Mechanical Systems and Signal Processing **199**.

Elsamanty, M., et al. (2023). "Principal component analysis approach for detecting faults in rotary machines based on vibrational and electrical fused data." Mechanical Systems and Signal Processing **200**.

Eshelman, R. (1999). Basic Machinery Vibrations: An Introduction to Machine Testing, Analysis and Monitoring, VIPress.

Feng, X., et al. (2020). "Wavelet-Based Kalman Smoothing Method for Uncertain Parameters Processing: Applications in Oil Well-Testing Data Denoising and Prediction." Sensors **20**(19).

Fidaleo, M. (2020). "Functional Data Analysis and Design of Experiments as Efficient Tools to Determine the Dynamical Design Space of Food and Biotechnological Batch Processes." Food and Bioprocess Technology **13**(6): 1035-1047.

Gotwalt, C. and R. Parker (2022). "Developer Tutorial: Modeling Spectral Data using JMP PRO 17." Retrieved 2nd August 2023, from https://community.jmp.com/t5/Mastering-JMP/Developer-Tutorial-Modeling-Spectral-Data-Using-JMP-Pro-17/ta-p/579098.

Gressani, O., et al. (2022). "Laplacian P-splines for Bayesian inference in the mixture cure model." Statistics in Medicine **41**(14): 2602-2626.

Haddadi, R., et al. (2014). Discrete Wavelet Transform Based Algorithm for Recognition of QRS Complexes. International Conference on Multimedia Computing and Systems (ICMCS), IEEE**:** 5.

Harmouche, J., et al. (2014). Linear Discrminant Analysis for the Discrimination of Faults in Bearing Balls by using Spectral Features. 2014 International Conference on Green Energy ICGE. Sfax, Tunisia, IEEE.

He, Z., et al. (2014). "A Review of Analysis of Dynamic Response in Design of Experiments." Quality and Reliability Engineering International **31**(4): 535-542.

Hoek, R. v., et al. (2021). Cooperative Driving of Automated Vehicles using B-Splines for Trajectory Planning. IEEE Transactions on Intelligent Vehicles, IEEE.

Houhou, R., et al. (2021). "Comparison of functional and discrete data analysis regimes for Raman spectra." Anal Bioanal Chem **413**(22): 5633-5644.

IBM (November 2023). "What is Linear Discriminant Analysis (LDA)?". Retrieved March 13, 2024, from https://www.ibm.com/topics/linear-discriminant-analysis.

Ingersoll-Rand (1977). Instruction Manual for Condensate Pumps, M-007-00029, Rev. 013**:** 103 of 118.

Izenman, A. J. (2008). Modern Multivariate Statistical Techniques- Regression, Classification, and Manifold Learning.
, Springer (New York).

Jaen-Cuellar, A. Y., et al. (2023). "Gear Wear Detection Based on Statistic Features and Heuristic Scheme by Using Data Fusion of Current and Vibration Signals." Energies **16**(2): 20.

James, G. M. (2017). "Curve Alignment by Moments." Annals of Applied Statistics **1**(2): 480-501.

Jayakumar, K. and S. Thangavel (2015). "Industrial drive fault diagnosis through vibration analysis using wavelet transform." Journal of Vibration and Control **23**(12).

Jiang, H., et al. (2023). "Vibration Signal Analysis of Roadheader Based on Referential Manifold Learning." Shock and Vibration **2023**: 1-11.

JMP (2023). "JMP Pro Data Processing Options." Retrieved February 29, 2024.

Jolliffe, I. T. (2002). Principal Component Analysis, Springer New York, NY.

Kenett, R. S. and C. Gotwalt (2023). "Functional data analysis and nonlinear regression models: an information quality perspective." Quality Engineering **35**(3): 480-492.

Kermarrec, G., et al. (2020). "Deformation Analysis Using B-Spline Surface with Correlated Terrestial Laser Scanner Observations- A Bridge Under Load." Remote Sensing **12**(5): 829.

Kramer, N., et al. (2008). "Penalized Partial Least Squares (PLS) with Applications to B-Spline Transformation and Functional Data." Chemometrics and Intelligent Laboratory Systems **94**(1): 60-69.

Kumar, A., et al. (2022). Vibration signal analysis of a rotor-bearing system through wavelet transform and empirical mode decomposition. International Conference on Materials Science and Engineering (ICMSE 2022).

Lakis, A. A. (2007). Rotating Machinery Fault Diagnosis using Time-Frequency Methods. 7th WSEAS International Conference on Electric Power Systems, High Voltages, Electric Machines. Venice, Italy**:** 139-144.

Lever, J., et al. (2017). "Principal Component Analysis." Nature Methods **14**: 641-647.

Liu Zepeng, et al. (2020). "Vibration analysis for large-scale wind turbine blade bearing fault detection with an empirical wavelet thresholding method." Renewable Energy **146**: 99-110.

Lu, P., et al. (2020). "Accuracy improvement of quantitative LIBS analysis of coal properties using a hybrid model based on a wavelet threshold de-noising and feature selection method." Appl Opt **59**(22): 6443-6451.

Maturana-Russell, P. and R. Meyer (2021). "Bayesian spectral density estimation using P-splines with quantile-based knot replacement." Computational Statsitics **36**: 2055-2077.

Metusalem, R. (2020). "A Primer on Functional Design of Experiments." October 10, 2023, from https://community.jmp.com/t5/JMP-Blog/A-primer-on-functional-design-of-experiments/ba-p/267566.

Michelson, D. (2024). "Sometimes special data need special modeling tools. JMP® Pro Functional Data Explorer to the rescue!". from https://community.jmp.com/t5/JMPer-Cable/Sometimes-special-data-need-special-modeling-tools-JMP-Pro/ba-p/729341.

Michis, A. A. (2017). "A wavelet smoothing method to improve conditional sales forecasting." Journal of the Operational Research Society **66**(5): 832-844.

Mohd Ghazali, M. H., et al. (2021). "Vibration Analysis for Machine Monitoring and Diagnosis: A Systematic Review." Shock and Vibration **2021**: 1-25.

Nizwan, C. K. E., et al. (2013). A Wavelet Decomposition Analysis of Vibration Signal for Bearing Fault Detection. 2nd International Conference on Mechanical Engineering Research (ICMER 2013). Pahang, Malaysia**:** 9.

Oguejiofor, B. N. and K. Seo (2024) Power plant vibration monitoring and optimization using functional design of experiments, University of Missouri-Columbia.

Oguejiofor, B. N. and K. Seo (2023). PCA-based Monitoring of Power Plant Vibration Signal by Discrete Wavelet Decomposition Features. 2023 Annual Reliability and Maintainability Symposium (RAMS), IEEE Xplore.

Pan, R. and M. Saleh (2019). "On designing experiments for a dynamic response modeled by regression splines." Applied Stochastic Models in Business and Industry **36**(2): 251-267.

Parker, R. (2021). "Analyzing Functional Data with Direct Functional PCA in Functional Data Explorer." Retrieved 30 July 2023, from https://community.jmp.com/t5/Discovery-Summit-Americas-2021/Analyzing-Functional-Data-with-Direct-Functional-PCA-in/ta-p/398753.

Parsons, S., et al. (November 2000). "Advantages and Disadvantages of Techniques for Transforming and Analyzing Chiropteran Echolocation Cells." Journal of Mammalogy **81**(4): 927-938.

Patel, R. and V. K. Giri (2016). "Analysis and Interpretation of Bearing Vibration Data Using Principal Component Analysis and Self-Organizing Map." Int J Advanced Design and Manufacturing Technology **1**: 111-117.

Ramsay, J. and B. W. Silverman (2005). Functional Data Analysis, Springer New York.

Ranjan, R., et al. (2020). "Fault Diagnosis of Journal Bearing in a Hydropower Plant using Wear Debris, Vibration and Temperature Analysis: A Case Study." Journal of Process Mechanical Engineering **234**(3): 235-242.

Ruppert, D. (2022). "Selecting the number of knots for penalized splines." Journal of Computational and Graphical Statistics **11**(4): 735-757.

Shaheryar, A., et al. (2017). "Robust Feature Extraction on Vibration Data under Deep-Learning Framework: An Application for Fault Identification in Rotary Machines." International Journal of Computer Applications **167**(4).

Shahmohammadi, M. A., et al. (2020). "Free vibration analysis of sandwich FGM shells using isogeometric B-spline finite strip method." Steel and Composite Structures **34**(3): 361-376.

Shi, M., et al. (2021). "Feature extraction method of rolling bearing based on adaptive divergence matrix linear discriminant analysis." Measurement Science and Technology **32**(7): 1-10.

Sozen, C. and Y. Oner (2022). "The investigation of temperature data in Turkey's Black Sea Region using functional data analysis." J Appl Stat **49**(9): 2403-2415.

Spreafico, M., et al. (2022). "Modelling time-varying covariates effect on survival via functional data analysis: application to the MRC BO06 trial in osteosarcoma." Statistical Methods & Applications **32**(1): 271-298.

Tiboni, M., et al. (2022). "A Review on Vibration-Based Condition Monitoring of Rotating Machinery." Applied Sciences **12**(3).

Tong, Z., et al. (2022). "Bearing Fault Diagnosis Based on Discriminant Analysis Using Multi-View Learning." Mathematics **10**(20).

Ullah, S., et al. (2024). "Fault Diagnosis of a Multistage Centrifugal Pump Using Explanatory Ratio Linear Discriminant Analysis." Sensors **24**(6).

Ullah, S. and C. Finch (2013). "Applications of Functional Data Analysis: A Systematic Review." BMC Medical Research Methodology **13**(43): 12.

Vanus, J., et al. (2023). "A innovative wavelet transformation method optimization in the noise-canceling application within intelligent building occupancy detection monitoring." Heliyon **9**(5): e16114.

Villarroel, A., et al. (2019). "Development of a Low-Cost Vibration Measurement System for Industrial Applications." Machines **7**(1).

Volvach, A., et al. (2024). "Wavelets in the analysis of local time series of the Earth's surface air." Heliyon **10**(1): e23237.

Wang, J.-L., et al. (2016). "Functional Data Analysis." Annual Review of Statistics and Its Application **3**(1): 257-295.

Wang, W. J. (1996). "Wavelet Transform in Vibration Analysis for Mechanical Fault Diagnosis." Shock and Vibration **3**: 17-26.

Wang, Y., et al. (2022). "Modeling nonlinear systems using the tensor network B-spline and the multi-innovation identification theory." International Journal of Robust and Nonlinear Control **32**(13): 7304-7318.

Wikimedia (2013). "BWR Nuclear Power Plant Diagram." 2023, from https://commons.wikimedia.org/wiki/File:BWR_nuclear_power_plant_diagram.svg.

Worley, B. (2022). "Analyzing Spectroscopic Data - What's New in JMP Pro 17." https://community.jmp.com/t5/JMP-Blog/Analyzing-Spectroscopic-Data-What-s-New-in-JMP-Pro-17/ba-p/550771 2024.

Xu, J., et al. (2022). "A Fast Weighted Fuzzy C-Medoids Clustering for Time Series Data Based on P-Splines." Sensors **22**(16).

Yan, R., et al. (2014). "Wavelets for fault diagnosis of rotary machines: A review with applications." Signal Processing **96**: 1-15.

Yan, T., et al. (2022). "Investigation on optimal discriminant directions of linear discriminant analysis for locating informative frequency bands for machine health monitoring." Mechanical Systems and Signal Processing **180**.

Zamudio-Ramirez, I., et al. (2022). Detection of Uniform Gearbox Wear in Induction Motors Based on the Analysis of Stray Flux Signals Through Statistical Time-Domain Features and Dimensionality Reduction Techniques. IEEE Transactions on Industry Applications IEEE Trans. on Ind. Applicat. Industry Applications, IEEE.

Zhao, J., et al. (2024). "Vehicle trajectory recontruction for intersections: an integrated wavelet transform and savisty-golay filter approach." Transportmetrica A: Transport Science **20**(2).

VITA

Benjamin Oguejiofor was born in Nigeria. He works as a Consulting Engineer for Ameren Missouri at the Callaway Nuclear Plant. He is currently completing his PhD degree in Industrial Engineering at the University of Missouri, Columbia under the advisement of Dr. Kangwon Seo. He is a recipient of the Multicultural Engineering Program and the Dean's Excellence Graduate Fellowships from the University of Missouri for his doctoral program. He obtained a Master of Science (M.S) degree in Mechanical Engineering from the Missouri University of Science and Technology, Rolla. He also obtained a Bachelor of Engineering (B.Eng.) degree in Mechanical Engineering from the University of Nigeria, Nsukka. His research interests include vibration analysis, predictive maintenance, functional design of experiments, machine learning and reliability engineering.