NTNU
Norwegian University of
Science and Technology

Multivariate Statistical Condition Monitoring

Håvard Oddan

Master of Science in Cybernetics and Robotics
Submission date: June 2017
Supervisor: Tor Engebret Onshus, ITK

Norwegian University of Science and Technology
Department of Engineering Cybernetics

ABSTRACT

Multivariate analysis methods have been studied for the purpose of improving condition monitoring of equipment in industry. This concerns fault detection, isolation and diagnosis for machinery in order to improve decision support capabilities for condition based maintenance, improved safety and reliability. Multivariate analysis techniques such as Principle Component Analysis and Partial Least Squares Regression have been investigated for this purpose, and applied to four different case studies.

Two case studies was based on data collected at Statoil's industrial processing facilities at Tjeldbergodden. One of these cases regarded a steam turbine compressor, where multivariate statistical methods where tested in order to evaluate their performance for fault detection and fault isolation. PCA and PLSR analysis was used and demonstrated a good performance for fault detection and fault isolation when data from a known fault condition was tested. The machine experience a non-stationary behaviour in the period of the analysis, which reduced the made use the multivariate methods more difficult and caused some false alarms.

The second case study regarded a steam turbine generator also from Tjeldbergodden. The performance of this machine was analyzed over a period of approximately two years. The results from the analysis was that a clear trend was detected from PCA analysis, which corresponded with a gradual loss of performance of the machine experienced in the same period. This case study demonstrated the use of multivariate analysis methods for discovering hidden information in relatively large data sets.

The third case regarded vibration analysis of a ball bearing. Vibration data from different working conditions of the bearing was used, including both healthy and faulty conditions. After signal conditioning, features from the vibration signal was extracted both from wavelet packet decomposition and time domain. The extracted features was then subject to the dimension reduction technique of PCA which was further used for fault detection and fault classification. The PCA analysis proved that faulty condition could be detected. SIMCA analysis was also used for classification. The classifier had good performance in classifying new samples of healthy condition and two distinct fault conditions.

The last analysis regarded acoustic signal monitoring of an air compressor, where wavelet packet decomposition was utilized for features extraction. PCA was further used in fault detection together with T^2 and SPE, based on training data and testing data for several fault conditions. SIMCA classification was used to train 8 different classes corresponding to the fault conditions and the healthy condition. Testing the classifier on new data samples revealed that the PCA based classification was not optimal, as nonlinear relations in the data could not be captured.

SAMMENDRAG

Tilstandsbasert vedlikehold har i de siste årene blitt viktigere i industrien for å redusere kostnadene forbundet med vedlikehold gjennom å overvåke den faktiske tilstanden til utstyret. Dermed kan levetiden på utstyr forlenges, vedlikeholdsoperasjoner kan planlegges på en optimal måte, og man oppnår en høyere pålitelighet og bedre sikkerhet. For å forbedre mulighetene for tilstandsbasert vedlikehold kan man benytte seg av mer intelligente tilstandsovervåkingssystemer. Gjennom en data-drevet tilnærming til tilstandsovervåking ved bruk av metoder fra multivariat data-analyse kan dette brukes til tidlig feildeteksjon og feildiagnose. I denne oppgaven er fire ulike studer gjennomført for å teste slike metoder.

Den første studien angikk en gasskompressor drevet av the dampturbin. Data fra rundt 55 målinger ble samlet inn, og metoder fra multivariat analyse ble brukt for beskrive de mange korrelerte variable i noen få latente variabler gjennom bruk av PCA og PLSR. Data fra normal drift ble brukt til å lage en modell av systemet, og ny data fra en kjent feilsituasjon ble brukt for å teste metodenes evne til å detektere, samt isolere feilen. Metodene klarte å detektere feilen i en tidlig fase, samt å isolere hvilken målinger som feilen materialiserte seg i.

Den andre studien ble utført på en dampturbin-generator. Driftsdata fra denne maskinen fra flere perioder i løpet av to år ble samlet inn å analysert. En tydelig trend ble i dataen ble tydelig gjennom PCA analyse, noe som kan ha en sammenheng med at denne maskinen har opplevd et tap av virkningsgrad over flere år. Bruk av multivariate metoder for å finne sammenhenger i store datasett ble her demonstrert.

En tredje studie angikk vibrasjonsanalyse av et kulelager. Vibrasjonsmålinger både fra normal tilstand og forskjellige tilstander ble brukt. Vibrasjonsmålingene ble behandlet med Wavelet Packet Decomposition for å ekstrahere informasjon fra tid-frekvensdomenet. PCA ble så brukt for å modellere normal tilstand til lageret, slik at feiltilstander kunne detekteres. I tillegg ble data fra ulike tilstander til lageret brukt for å trene en klassifiseringalgoritme baser på SIMCA, slik at nye målinger kunne klassifiseres i henhold til tilstanden lageret er i.

The siste studien som ble gjennomført omfattet akustisk overvåking av en luftkompressor. Her ble også Wavelet Packed Decomposition benyttet for ekstrahere informasjon fra tid-frekvensdomenet. På samme måte som for vibrasjonsanalysen for kulelageret, ble også PCA benyttet for å modellere normal tilstand slik at feiltil-

stander kunne oppdages. I tillegg ble 7 feiltilstander i tillegg til normal tilstand benyttet for SIMCA klassifikasjon. Når data fra ulike tilstander ble klassifisert, hadde SIMCA metoden noen vanskeligheter med å skille klassene.

CONCLUSION

This thesis describes the use of multivariate statistical methods for the purpose of condition monitoring. The thesis includes four case studies where these techniques were tested and evaluated. This includes two cases of analysis of large rotating machinery from Statoil's industrial complex at Tjelbergodden, a third study regards bearing vibration analysis and lastly an acoustic signal analysis of a reciprocating air compressor is covered.

The first case study regarded a steam turbine compressor for the purpose of testing multivariate condition monitoring techniques to achieve fault detection and fault isolation. Both Principal Component Analysis (PCA) and Partial Least Squares Regression was utilized for this purpose. These latent variable methods were able to reduce the dimension of the 55 measured variable, and the multivariate control charts Hotelling's T^2 and SPE were used for fault detection. Data from a known fault situation was used for testing, and showed that the methods were able to detect the faults at an early stage. The fault was also successfully isolated by the use of contribution plots. The data set used in this analysis did however have some limitation, as the machinery was not in a steady state of operation, reducing the modelling abilities for the latent variable methods and causing false alarms.

The second study regarded steam turbine generator where the its performance was analyzed over a period of approximately two years. The results from the analysis was that a clear trend was detected from PCA, which corresponded with a gradual loss of performance of the machine experienced in the same period. The analysis did however suffered heavily from missing data, reducing the quality of the analysis.

The third case regarded vibration analysis of a ball bearing. Vibration data from different working conditions of the bearing was used, including both healthy and faulty conditions. After signal conditioning, features from the vibration signal was extracted both from wavelet packet decomposition and time domain. The extracted features was then subject to the dimension reduction technique of PCA which was further used for fault detection and fault classification. The PCA analysis proved that faulty condition could be detected. SIMCA analysis was also used for classification. The classifier had good performance in classifying healthy condition and two distinct fault conditions.

The last analysis regarded acoustic signal monitoring of an air compressor, where wavelet packet decomposition was utilized for features extraction. PCA was further used in fault detection together with T^2 and SPE, based on training data and testing data for several fault conditions. SIMCA classification was used to train 8 different classes corresponding to the fault conditions and the healthy condition. Testing the classifier on new date revealed that PCA based classification was not optimal, as nonlinear relations in the data could not be captured.

PREFACE

This thesis concludes a Master's degree in Engineering Cybernetics from the Department of Engineering Cybernetics at NTNU.

The study was performed in collaboration with Statoil Tjeldbergodden. One of the objectives of this research is evaluate condition monitoring based on multivariate techniques for possible contributions to condition based maintenance at Tjeldbergodden.

The reader is assumed to have a basic understanding of linear algebra, but in-dept knowledge and experience with multivariate analysis should not be necessary.

Trondheim, June 2017
Håvard Oddan

ACKNOWLEDGMENTS

I would like to thank my supervisor Prof. Tor Onshus, for the guidance and support.

I would also like to thank Statoil Tjeldbergodden for the collaboration. In particular I would like to thank Ola Ness for the guidance and technical support, and Kari Faksvåg for facilitating the collaboration.

H. O.

CONTENTS

Part I

INTRODUCTION

1

INTRODUCTION

1.1 INTRODUCTION

1.1.1 *Background*

As modern industry companies are dependent on profit in order to be become and remain successful businesses, they are pressured to meet their production targets at the same time as they often are need to meet high safety requirements. In order to achieve this, the production equipment need to have high availability and reliability to minimize production downtime, and at the same time minimize maintenance related costs. Unexpected equipment failure can lead to large production losses, and unplanned maintenance can result in high additional costs. Due to these reasons, the concept of condition based maintenance (CBM) have become widespread in industry. This maintenance strategy can have great benefits compared to pure corrective or preventive maintenance in many situations, as it is based on the actual condition of the equipment, allowing maintenance to be performed in the most optimal manner.

Condition based maintenance relies on the ability to monitor the condition of the equipment. This is achieved through collection of information concerning the working condition of the equipment utilizing instrumentation to measure features such as vibration, temperature, pressure, acoustic emissions etc. Information from the condition monitoring can further be used by operational and maintenance personnel to take necessary actions, perform and plan maintenance operations. To further improve decision support capabilities for CBM, intelligent condition monitoring techniques can be used in order automate early detection of faults, and give diagnosis and prognosis of equipment health state. Such intelligent condition monitoring systems have in recent years been widely investigated in academia, and seem to be getting greater interest in industry. Such systems are therefore the main motivation for this thesis.

Data-driven methods for analyzing an ever-increasing amount of information are getting more important as we are in a time experiencing a large pressure for increased digitalization, with a revolution in industrial big data and internet of things. Industry companies are collecting a vast amount of data, such as from sensors at plants, which currently only parts of are made

use of, creating what is often known as an analytics gap. Being able to capture the meaning and hidden structures behind complex industrial quantitative big data would have large potentials for increasing the insight and understanding to help companies make better decisions.

1.1.2 *Objectives*

The objectives of the thesis is to investigate methods from multivariate analysis such as Principal Component Analysis and Partial Least Squares Regression to improve condition monitoring of industry machinery in terms of fault or anomaly detection, fault isolation and diagnosis.

The research concerns a investigation of the state of the art and related work, in addition implementing and evaluating multivariate condition monitoring methods for several different applications.

The methods are tested for two specific applications based on Statoil's process plant at Tjeldbergodden, where historical data is gathered from process equipment. The performance of the methods are then evaluated.

Additionally the multivariate analysis methods are tested for two other applications based on wavelet packet decomposition. One of these cases concerns vibration analysis of a bearing, and the other on acoustic signal analysis of a reciprocating air compressor.

1.1.3 *Report Structure*

This report first includes an introduction to intelligent condition monitoring in chapter 2, which covers the basics of condition based maintenance, different condition monitoring and data acquisition techniques, and different approaches to intelligent condition monitoring.

Chapter 3 is devoted to relevant background theory of multivariate analysis, including data pretreatment techniques, Principle Component Analysis (PCA) and Partial Least Squares Regression (PLSR).

Chapter 4 covers the data processing techniques of time-frequency decomposition including Short Time Fourier Transform and Wavelet Decomposition.

Chapter 5 involves an outline of the methods for multivariate statistical condition monitoring.

Chapters 6 and 7 regard two application cases for process equipment at Tjeldbergodden, involving a steam turbine compressor and a turbine generator unit, respectively.

Chapter 8 concerns a condition monitoring application based on vibration analysis of a bearing, while chapter 9 presents an application on acoustic signal analysis on an air compressor.

Chapter 10 provides a discussion of the methods used, the results, evaluation and further work.

1.1.4 *Software*

- PI Datalink was used for gathering historic process data from the database at Statoil Tjeldbergodden. This data was exported to Microsoft Excel files for further use.
- Matlab R2016b was used for preliminary data visualisation, conditioning, pretreatment and also for implementing algorithms for the respective applications along with presentation of resulting data and visualisation.
- The Unscrambler® X version 10.3 by CAMO Software. Used as main tool for multivariate data analysis. This includes model generation, analysis and visualisation.

2

INTRODUCTION TO INTELLIGENT CONDITION MONITORING

2.1 CONDITION MONITORING AND CONDITION BASED MAINTENANCE

Traditionally, condition monitoring methods have mostly been implemented in situations where there is of great importance to prevent catastrophic failure of critical machinery. In oil and gas industry such applications often include critical rotating machinery such as gas compressors, turbines or pumps. As the cost of instrumentation is tending to decrease, and information technology and analytical methods are improving, condition monitoring approaches are continuously expanding to a wider range of equipment.

As condition monitoring enables the health of equipment to be monitored so that deterioration is detected at an early stage, possible disasters can be avoided. Being able to diagnose, and optimally to preform a prognosis to predict remaining useful life, an optimal maintenance strategy is possible in the form of condition based maintenance, or sometimes known as predictive maintenance. Some of the main benefits of such maintenance strategy is listed below [36]:

Reduced repair time and costs

Compared to corrective maintenance, condition based maintenance enables maintenance activities to be planned in advance. This is advantageous as prior knowledge and planned activities often are performed more efficiently in less time and with better quality. Intelligent condition monitoring system also helps diagnosing fault conditions, reducing the time needed for manual troubleshooting and diagnosis.

Avoided revenue loss

Unplanned production downtime caused by sudden failure of critical equipment can lead to substantial loss of revenue. The downtime often tend be prolonged due to ordering of spare parts, as well as unplanned maintenance activities become less efficient. If instead a CBM approach is used, and equipment deterioration can be detected in and early phase, the maintenance activities can be planned

in convenient off-peak periods or in combination with other necessary maintenance activities or revision stops.

Maintenance cost savings

Reduced cost of spare parts and maintenance compared to preventive maintenance, as only the equipment that actually need to be repaired is fixed and the equipment lifetime can be increased. Failures induces by maintenance can also be reduced as the frequency of maintenance activities is lower compared to preventive maintenance. The size of the needed spare part inventory can also be reduced.

Improved safety

Condition monitoring enables increased equipment health verification and safety assurance. Early alarms also reduces the chance of severe machine damage which lowers the risk for personnel.

2.2 DATA ACQUISITION

Selecting the correct sensor to be able to gather accurate and sufficient information about the equipment health condition is of key importance for effective condition monitoring. Main considerations when choosing transducers is determining the principle fault conditions to be monitored, and defining the mechanical manifestation of faults. Most often a number of different transducers are used in combination to complement each other.

2.2.1 *Vibration Monitoring*

Vibration monitoring is used in a wide range of equipment monitoring applications, and is of special importance and is widely implemented for rotating machinery. Vibration monitoring enables detection of a problems in structural and rotating parts of the equipment. A particular benefit is that specific spectral components can be extracted from the vibration signal which reveals much information about certain components and fault conditions in the machine.

Different types of transducers are used for vibration monitoring, including proximity probes, velocity transducers and accelerometers. Proximity probes measure the motion of the rotating shaft directly in displacement, have the advantages of not being in direct contact with the shaft, and they have high low frequency gain. Velocity transducers measure the dynamic

motion and are effective for the speed range of many large rotating machines, have good sensitivity, but suffer from a narrow frequency response and is sensitive to damage from shock loadings. Accelerometers measure the dynamic response, an have good high frequency response, but poor low frequency response and sensitivity.

A range of features can be extracted from vibration signals including time domain features such as RPM, variance, absolute mean, peak, crest factor, skewness, shape factor and kurtosis, and so on. For frequency analysis of rotating machines, synchronous vibration spectra is important as many features about the machine condition may be revealed. For instance, the amplitude and phase for the first rotational order (1x), synchronous to the shaft rotation, can measure the state of rotor imbalance. Higher rotational orders (2x, 3x, 4x) can detect mechanical asymmetries such as rotor cracks [36].

2.2.2 *Thermal Monitoring*

Thermal monitoring is widely applied for condition monitoring as many faults are manifested in terms of certain thermal characteristics in temperature and heat transfer. It is particularly useful for monitoring bearings and gears in rotating machines often in combination with vibration monitoring.

Infrared thermal monitoring is also used for condition monitoring of machinery, equipment and processes, and one example of an application is to detect hot-spots caused by leaks of process fluids [36]

2.2.3 *Process Parameter Monitoring*

Process parameters used for process monitoring and control can be utilized for condition monitoring. Such parameters can include pressure, pressure drop, flow-rate, temperature etc. Deviation from the normal behavior of these parameters can be used to detect fault condition and process disturbances.

2.3 INTELLIGENT CONDITION MONITORING, DIAGNOSIS AND PROGNOSIS SYSTEMS

Figure 1 illustrates the main parts of intelligent condition monitoring, diagnosis and prognosis systems. Some of the main part of such system include sensor inputs, preprocessing, feature extraction, fault detection, fault isolation, fault classification and fault prediction.

2.3.1 *Feature identification*

Inputs to the system include a variety of information sources such as on-line sensor outputs from equipment instrumentation such as those described above, but also other sources from environmental information, observation and control information [36]. This information is further processed in order reduce noise, clean the data, handle missing values et cetera, such that the desired information can be extracted. Signal processing can be preformed in time domain, such as moving average filtering, or in frequency domain, such as high pass, low pass or band pass, or other techniques such as wavelet thresholding. When appropriate signal processing is performed, features are extracted from the signals. This can include time domain features such as RMS, mean value, kurtosis, crest factor, skewness et cetera. Frequency domain domain features such as spectral component extraction, spectral signature and spectral envelope are also in many cases useful. For rotated machinery, extracting spectral components is of particular high use as a lot of information of the machine condition is revealed from the spectral characteristics.

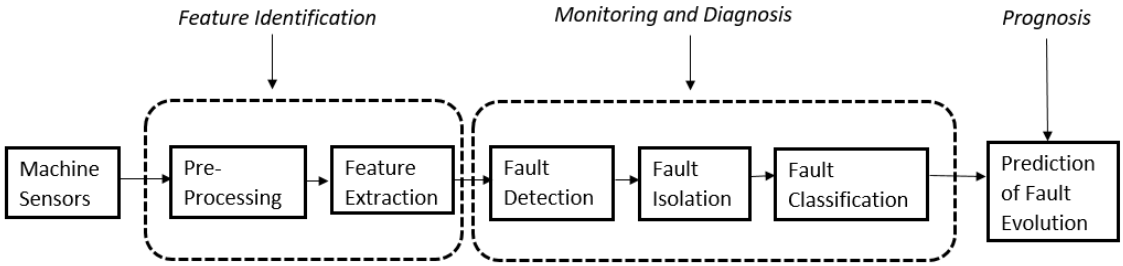

Figure 1: Flow chart overview of intelligent condition monitoring systems

2.3.2 *Fault Detection*

The features extracted is further monitored in a fault detection subsystem. The features are here compared to their healthy condition, and any deviations revealing equipment degradation can be detected. At this point a fault is detected in the form of a global deviation from normal condition, and the source of the fault is not ascertained.

2.3.3 *Fault Isolation*

The next subsystem includes fault isolation, which concerns identification and isolation of the fault that has revealed itself from the features. This can reveal which information sources, such as a component or subsystem, that is responsible for the deviations of the features. This does not necessarily imply that the causal relationship for the fault is found, only how it has influenced the features.

2.3.4 *Fault Classification*

The next part of the system concerns fault classification, which basically is where fault diagnosis is performed. The fault is here identified and classified in order to determine the cause of the fault and the equipment fault condition.

2.3.5 *Prognosis*

As fault diagnosis in essence only determines the current state and fault condition, a prognosis is needed to predict future condition of the equipment. Often one is interested in determining the remaining useful time (RUL) of the equipment in order plan and optimize maintenance.

Part II

BACKGROUND THEORY

3

MULTIVARIATE ANALYSIS

This chapter presents some main aspects concerning multivariate data analysis, including the theory and presentation of the two multivariate projection methods Principal Component Analysis and Partial Least Squares Regression. The theory presented below is mainly based on the results from [23], [27], [16].

3.1 LATENT VARIABLE METHODS

The multivariate analysis methods used in this study utilizes bilinear subspace models to analyze and model the structure in a data set. Bilinear modelling gives a generic description of data using loadings **P** and scores **T** in the following form [27]

$$\mathbf{X} = \mathbf{T}\mathbf{P}^{\mathbf{T}} + \mathbf{E} \tag{1}$$

where $\mathbf{X} \in \mathbb{R}^{m \times n}$, $\mathbf{P} \in \mathbb{R}^{n \times a}$, $\mathbf{T} \in \mathbb{R}^{m \times a}$, **X** is a data matrix with m rows representing samples, and n columns representing variables. $\mathbf{P} = [\mathbf{p_1}, ..., \mathbf{p_a}]$ is the loading matrix and the columns $\mathbf{p_i}$ are loadings. $\mathbf{T} = [\mathbf{t_1}, ..., \mathbf{t_a}]$ is the score matrix and the columns $\mathbf{t_i}$ are scores. E is the residual matrix and describes the unmodelled part of the data in **X**.

Latent variable methods are used to approximate a data set by finding linear combinations of the original variables and expressing new latent variables spanning a subspace in the original variable space. The methods assume that the data set is structured with correlated variables such that there exist latent variable describing the data in a lower dimension variable space. The set of loading vectors constitutes a basis for the LV-space, while the scores represent the coordinates of the projections of each sample or object onto the loading vectors [27].

3.2 PREPROCESSING

Several steps of preprocessing is often performed in order get the data for multivariate analysis on a suitable form and reducing the effect of unwanted noise and variation. A short summary of such methods are given below.

3.2.1 *Filtering*

Filtering techniques can be utilized in order to improve the signal-to-noise ratio and reduce the effects of outliers in the data.

One of the simplest types of smoothing filtering methods is the moving average filter. In this method, each reading x_{ik} of each variable $k = 1, 2, ..., K$ is replaced by a a weighted average of itself and its nearest neighbours from $k - D$ to $k + D$ [23]:

$$x_{ik} = \sum_{d=-D}^{+D} x_{i,k+d} u_d \tag{2}$$

Where u_d are the convolution weights.

3.2.2 *Centering*

Mean centering is used to remove the variable offset of the data. Each column is subtracted by its corresponding mean value so that all variables are centered around mean. This increases the ease of interpretation and numerical stability [16].

3.2.3 *Normalization*

In situations where the variables are scaled differently or have different units, some variables often tend to dominate the analysis. To avoid this effect such that all variables will have equal influence in the analysis, they can be normalized by dividing the variables with their respective variance in order to have the equal amount of variation.

Some of the pitfalls regarding normalization is that if calibration of nearly static variables with low variance, the normalization will amplify the noise and diminish the signal-to-noise ratio.

3.2.4 *Weighting*

Weighting can be performed on each of the variables in the data set in order do increase or decrease the influence of the corresponding variables in the model. This procedure is often approached in an empirical manner based on application specific knowledge and experience.

3.2.5 *Other Preprocessing Techniques*

A variety of other preprocessing techniques can be performed. Variable selection is a method where subsets of variables are selected for model calibration, and irrelevant variables or variables the diminishes the performance of the model is removed from the analysis.

Outlier handling concerns detecting and removing outliers from the data set that will otherwise greatly influence the bilinear models.

Nonlinear transformations may be utilized in order to linearize the original non-linear variables [27].

The reader is referred to [23] for more information regarding data proprocessing.

3.3 PRINCIPAL COMPONENT ANALYSIS

By diagonalizing the covariance matrix by extracting the cross-correlation between the variables in the data matrix, the PCA transforms the data matrix in a statistically optimal manner to a compressed and decorrelated form. In this way PCA achieves its strong ability to reduce the dimensionality of the data matrix while capturing the underlying covariation structure between the variables [10]. With highly correlated variables in a data set contaminated with noise, the first few principle components will capture the main relationship in the data, and the remaining components will mainly be composed of the noise. The mathematical representation with mean centered data matrix $\mathbf{X} \in \mathbb{R}^{m \times n}$,

$$\mathbf{X} = \mathbf{T}\mathbf{P}^{\mathbf{T}} + \mathbf{E} = \sum_{i=1}^{A} \mathbf{t}_i \mathbf{p}_i^{\mathbf{T}} + \mathbf{E} \tag{3}$$

where $\mathbf{P} \in \mathbb{R}^{n \times a}$, $\mathbf{T} \in \mathbb{R}^{m \times a}$. The data matrix may also be decomposed by singular value decomposition as [8],

$$\mathbf{X} = \mathbf{U}\tilde{}^{1/2}\mathbf{V}, \tag{4}$$

where Λ is a diagonal matrix of the eigenvalues $\mathbf{P}^{\mathbf{T}} = \mathbf{V}$ and $\mathbf{T} = \mathbf{U}\tilde{}^{1/2}$. The principle components are therefore ordered corresponding the eigenvalues. The elements corresponding to the the $A < n$ largest eigenvalues are extracted.

The computational algorithms for PCA are either based on singular value decomposition or the NIPALS (Nonlinear Iterative Partial Least Squares) algorithm. The reader is referred to [11] for a more detailed outline and pseudo code for the algorithms.

3.4 PARTIAL LEAST SQUARES REGRESSION

Partial Least Squares Regression (PLSR) is a regression technique that utilized latent structures in the data sets in order to build a regression model from explanatory variables $\mathbf{X}$ to response variables $\mathbf{Y}$. The latent structures are used such that the regression model most optimally connects the input and the output data.

PLSR model with mean centered data sets $\mathbf{X} \in \mathbb{R}^{m \times n}$ and $\mathbf{Y} \in \mathbb{R}^{m \times p}$

$$\mathbf{X} = \mathbf{Z_x P^T} + \mathbf{E_x} \tag{5}$$

$$\mathbf{Y} = \mathbf{Z_y Q^T} + \mathbf{E_y} \tag{6}$$

$$\mathbf{Z_x} = \mathbf{XW}(\mathbf{P^T W})^{-1} = \mathbf{XW}^* \tag{7}$$

where $\mathbf{Z}_x \in \mathbb{R}^{m \times a}$, $\mathbf{P} \in \mathbb{R}^{m \times a}$, $\mathbf{E}_x \in \mathbb{R}^{m \times n}$, $\mathbf{Z}_y \in \mathbb{R}^{m \times a}$, $\mathbf{Q} \in \mathbb{R}^{p \times a}$, $\mathbf{E}_y \in \mathbb{R}^{m \times p}$ and $\mathbf{W} \in \mathbb{R}^{n \times a}$ [27]. The reader is referred to [23] for derivations and further details regarding PLSR.

Partial Least Squares Regression is performed using the NIPALS algorithm [11].

3.5 VALIDATION

3.5.1 *Underfitting Versus Overfitting*

A challenge met when making a latent variable model is choosing model complexity, that is, choosing the number of latent variable that should be included. Choosing too high model complexity can result in overfitting, and thus creating a model with poor predictive performance and diminishing the signal-to-noise ratio. In this case the model will include noise and other unwanted phenomena. If on the other hand too few latent variables are included, we risk underfitting the model and loosing too much information and distorting the signal [23].

Different quantitative strategies for choosing model complexity exists and some of the are discussed in further sections. It should be mentioned that no method has shown to give optimal answers to this problem, thus this challenge is often based on a certain degree of empirical evaluation.

3.5.2 *Cross-validation*

Cross-validation is a method which uses the calibration data in order to determine the number of principal components that should be retained. In essence this method omits segments of the training data while performing PCA on the rest of the data sequentially by starting with only one PC, then two, then three and so on. This is performed on all data segments in turn so as all segments have been kept out once. The number of components are then chosen based on the PCA model complexity which shows the minimum prediction error [23].

3.5.3 *Independent Test Set*

The independent test set regards using a dedicated validation data set which is only used for validation. [12]

3.5.4 *The Explained Variance*

The explained variance is the variance that is explained by the model described by the signal. The variance not included in the model describes the noise in the data and is called residual variance. The explained variance increases as more and more principal components are added in the calibration model. Depending on the training data set at hand, this can be used to the determine how many principle components to retain in the model, as one for example could want to have about 95% explained variance by the model.

3.6 PLOT INTERPRETATIONS

Two of the most frequently used plots for graphical illustration of the results from multivariate analysis are presented below [12].

3.6.1 *The Score Slot*

The score plot shows the plot of each score corresponding the projected samples down on the principle components. Often the scores are plotted on a two or three dimensional space spanned by the corresponding two or three first two principle components, as these are the most describing of the data. The score plots can be seen as 'windows' into the PC-space, where each dot represent a sample or object. The score plot can

be used to visually interpret the calibration model, detecting groupings, outliers or trends in the data.

Figure 2 illustrates a two-dimensional score plot of PC1 and PC2. This plot is from the bearing vibration analysis where the calibration scores are from a healthy bearing condition. Samples for a bearing fault is projected on to this model, which are clearly grouped outside the region for normal condition, indicating a fault.

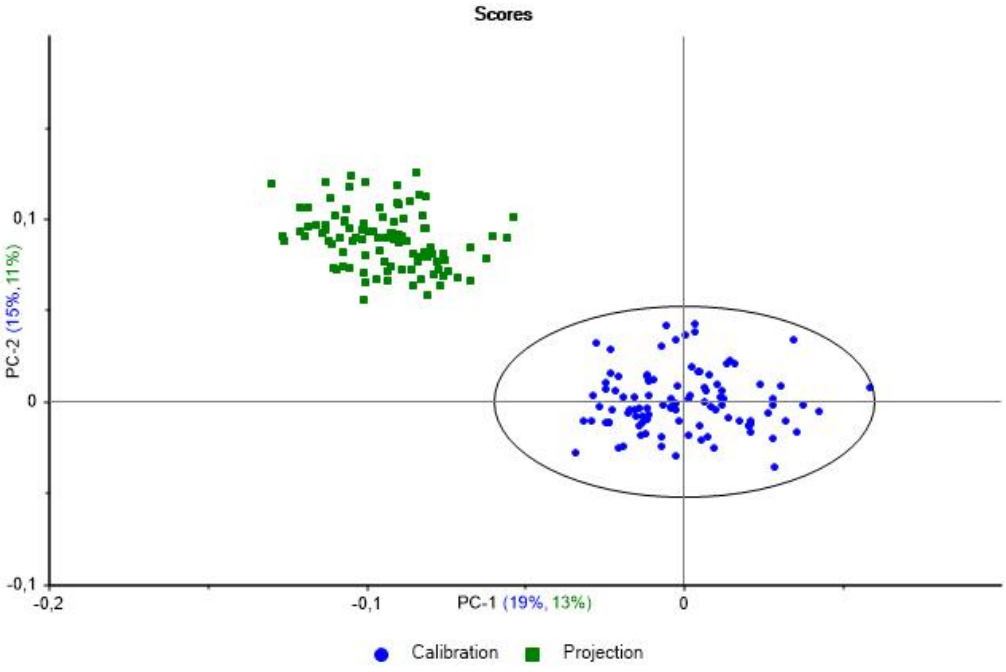

Figure 2: Score plot

3.6.2 *The Loadings Plot*

The loadings plot is used to illustrate how variables contribute to each PC and their relationship regarding to their contribution. In a correlation loadings plot, an absolute value of one indicates high variable contribution from the respective variable to the particular PC. A value of zero indicates no contribution. Clustered variables in the loadings plot indicate high correlation between them [16].

Figure 3 shows a correlation loadings plot for the bearing analysis. We see that PC1 have strong contributions from the *Variance*, *RMS* and *Abs mean* variables placed at the left in the plot. The strong clustering near the center shows that a large number of the variables are highly correlated and with limited influence on PC1 and PC2.

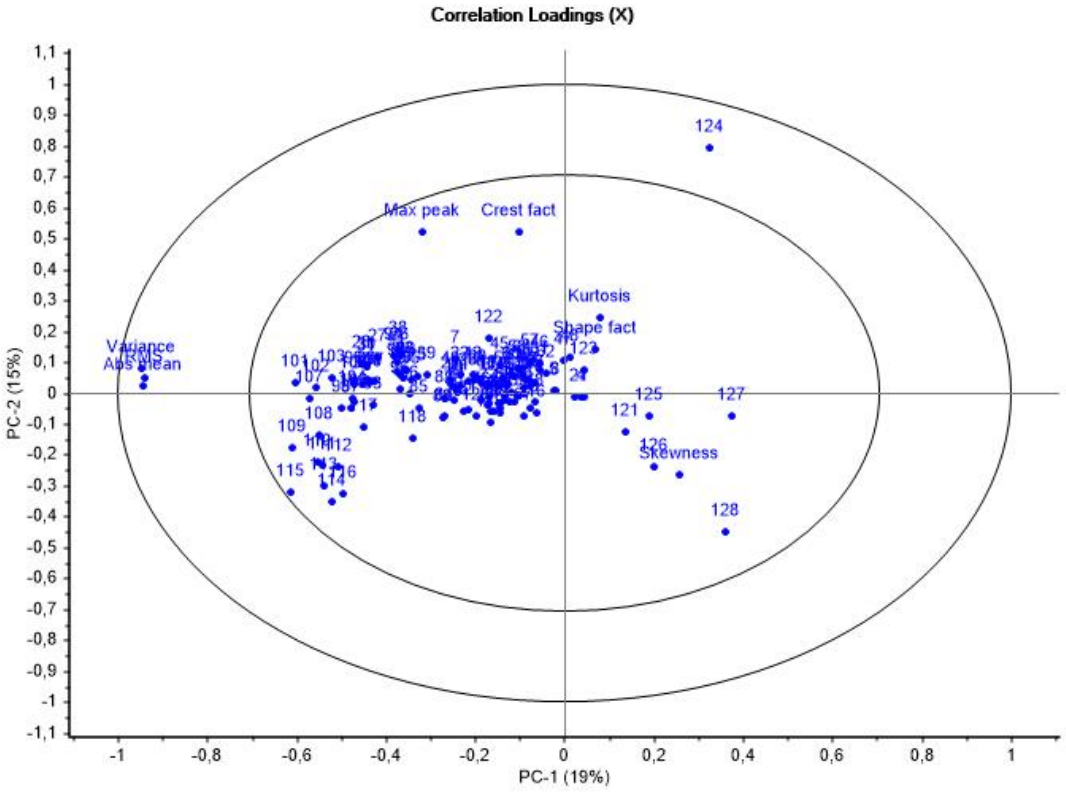

Figure 3: Loading plot

4

TIME-FREQUENCY ANALYSIS

Time-frequency analysis is an important subject in signal processing for a wide range of application when dealing when non-stationary signals [15]. Such techniques allows investigating both the temporal and spectral characteristics of a signal. These abilities make such techniques particularly useful for condition monitoring and fault diagnosis. Figure 4 illustrate a time-frequency decomposition in the form of a spectrogram for a vibration signal. The methods

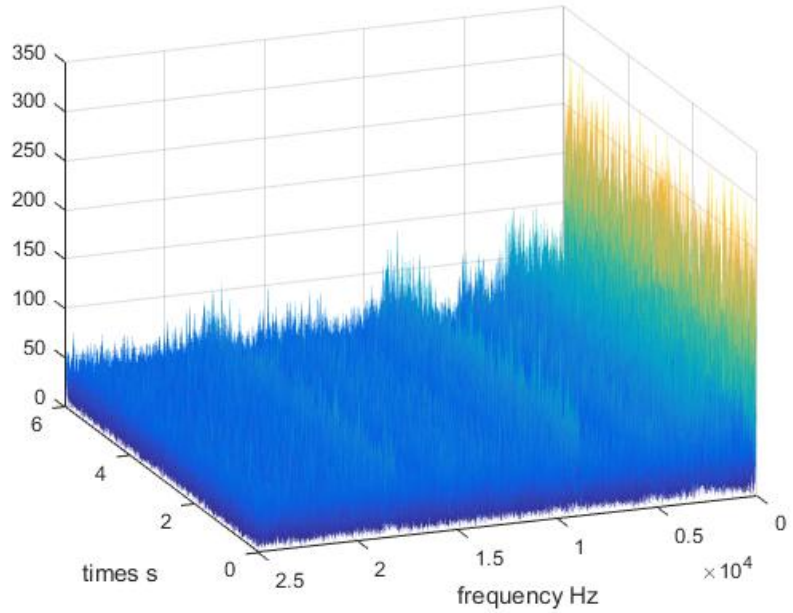

Figure 4: Spectrogram of bearing vibration

4.1 SHORT TIME FOURIER TRANSFORM

The traditional Fourier Transform transforms a signal between time domain and frequency domain. A time domain signal $X(t)$ is transformed into $X(f)$ in frequency domain in the following way,

$$X(f) = \int_{-\infty}^{\infty} X(t) \bullet e^{-2j\pi ft} dt \tag{8}$$

Discrete case,

$$\mathbf{X}_k = \sum_{n=0}^{N-1} x_n \cdot e^{\frac{-i2\pi kn}{N}} \tag{9}$$

The shortcoming of the traditional Fourier Transform is that the temporal information of a signal is lost when it is transformed into frequency domain. This is not a problem when dealing with stationary signals, but valuable temporal information is lost if the signal is non-stationary.

To be able to decompose the signal in both frequency and time simultaneously, a windowed Fourier Transform can be used. This is called Short Time Fourier Transform (STFT). Mathematically it is given as [15]

$$\mathbf{STFT}\{x(t)\}(\tau,\omega) \equiv X(\tau,\omega) = \int_{-\infty}^{\infty} x(t)w(t-\tau)e^{-j\omega t}dt \tag{10}$$

where $w(t-\tau)$ and is a window function and $x(t)w(t-\tau)$ is a windowed signal. The discrete case

$$\mathbf{STFT}\{x[n]\}(m,\omega) \equiv X(m,\omega) = \sum_{n=-\infty}^{\infty} x[n]\, w[n-m]\, e^{-j\omega n} \tag{11}$$

The spectrogram from STFT with squared magnitude is defined as

$$\mathrm{spectrogram}\{x(t)\}(\tau,\omega) \equiv |X(\tau,\omega)|^2 \tag{12}$$

The basic idea behind STFT is that it divides the signal up in segments, performing a Fourier Transform on each segment giving a sequence of temporal spectrums. The choice of window length is important in order to capture the desired signal characteristics and for the ability to suppress side blobs and spectral leakage [15]. One of the weaknesses of STFT is its fixed resolution, as shortening of the window length gives higher time resolution but poorer frequency resolution, and expanding the window length improves frequency resolution but degrades time resolution [24].

4.2 WAVELET TRANSFORM

An expansion of the Short Time Fourier Transform is the wavelet transform which have improved time-frequency localization features. Wavelet theory is relatively new, and wavelet transform

based technologies have in recent years been widely applied to a range of fields including signal and image processing, numerical computation, pattern recognition, speech analysis, fault diagnosis and others [30].

The goal of this section is not to provide a complete description of wavelet theory, but to present the basic ideas behind these technologies. The reader is referred to [22] for further details.

Continuous Wavelet Transform is defined by [24]:

$$G_s(w,t) = \int s\left(\frac{1}{\sqrt{a}}\right) \varphi\left(\frac{t-u}{a}\right) du \tag{13}$$

where a is a scale factor, u is the shift, $\varphi(t)$ is the mother wavelet and $G_s(w,t)$ is the wavelet transform of function $s(t)$. We see that the continuous wavelet transform correspond to the STFT with a shifted and scaled window function called mother wavelet.

4.2.1 *Discrete Wavelet Transform*

For the Discrete Wavelet Transform the signal is decomposed by convolution with high pass and low pass filter coefficients, called quadrature mirror filters [33].

$$y(n) = \sum_{l} f(n-l) \cdot g(l) \tag{14}$$

where $y(n)$ is the output from the convolving functions $f(n)$ and $g(n)$. The results from the high pass filters are called detail coefficients, and the results from the low pass filters are called approximation coefficients. A dyadic sampling is achieved as illustrated in figure 5.

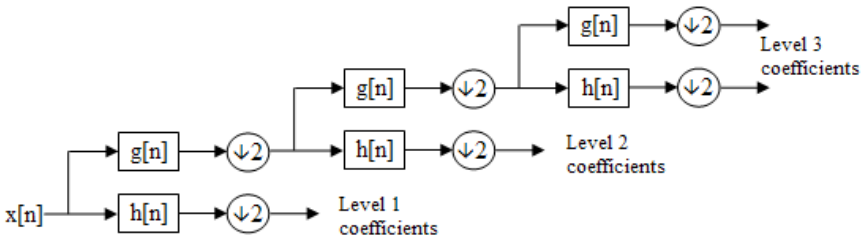

Figure 5: Discrete Wavelet Transform filter bank. (Figure from [1])

4.2.2 *Wavelet Packet Decomposition*

The wavelet packet decomposition is an extension of wavelet decomposition, where both approximation and detail coefficients are further decomposed providing a more precise frequency resolution. The wavelet packet decomposition makes a full binary tree as shown in figure 6 [37].

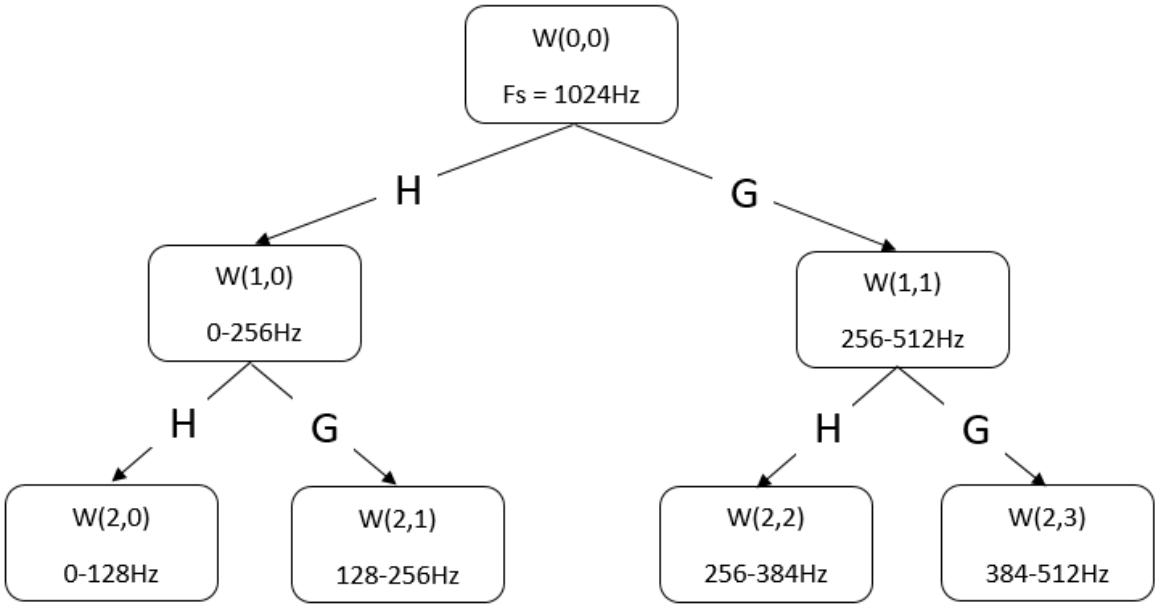

Figure 6: Wavelet Packed Tree

4.2.3 *Choosing Wavelet Basis Function*

The choice of wavelet basis function is important for meeting application specific needs. For condition monitoring, fault detection and diagnosis, orthogonal wavelet functions are preferred as they have the important property of being able to decompose and reconstruct signals efficiently. Wavelet families such as Haar, Coiflet and Daubechies can be used [33].

5

MULTIVARIATE STATISTICAL CONDITION MONITORING

Multivariate analysis have been successfully applied in industry both for multivariate statistical process control and quality control [20, 28]. The are large potential in utilization of multivariate statistical techniques for condition monitoring of machinery, equipment and industrial processes for fault detection and fault diagnosis. The data-driven based models such as Principle Component Analysis (PCA) and Partial Least Squares Regression (PLSR) models has proven to be effective for such application as deep knowledge of the processes and machines are not needed. Such quantitative latent variable methods are based on the actual data from the machinery and processes under monitoring and is then able to model and capture characteristics and conditions that are often not included in the first principle based models [17]. These models are not however causal models, but model only common-cause variation or correlation.

Latent variable methods are powerful as they are able to reduce the dimension of large data sets with many correlated variables in to a lower dimension of uncorrelated latent variables. These latent variable are linear combinations of the original variables, where the least significant variables are eliminated. By modelling a latent variable structure model from healthy historic data, these projection models can be used for online monitoring for the purpose of detecting faults and abnormal condition, and for fault diagnosis.

5.1 FAULT DETECTION

The first step in multivariate statistical condition monitoring is to develop an "in-control model" which models the normal situation from historic data using PCA or PLSR. The data from periods spanning the range for normal condition should be used for training of the in-control model, such that the mean and covariation structure for normal operating conditions are captured. The in-control model can then be applied on-line where new data is projected on the model. New data samples are subject to operations such as mean centering, scaling and transformation defined from the calibration data.

The projections of new data can be tested if they can reveal if any fault situation is apparent. Several multivariate control

charts can used for condition monitoring and fault detection, and the most important charts include the Hotelling's T^2 statistic and the Squared Prediction Error (SPE). These to indices are discussed in the further sections.

5.1.1 *Hotelling's T^2 Statistic*

Hotelling's T^2 statistic measures how close a new observation is to the model mean. This is analogous to the euclidean distance from the projected sample in the latent variable space, to the origin of the model. Hotelling's T^2 statistic for sample i [20]

$$T_i^2 = \sum_{j=1}^{A} \frac{\mathbf{t_j t_j^T}}{s_{t_j}^2} = \sum_{j=1}^{A} \frac{\mathbf{t_j t_j^T}}{\lambda_j} \tag{15}$$

The upper control limit (UCL) based on A first principle components can be defined as [20]

$$T_{A,UCL}^2 = \frac{(n^2-1)A}{n(n-A)} F_{\alpha(A,n-A)} \tag{16}$$

where $F_{\alpha(A,n-A)}$ is the upper $100\alpha\%$ critical point of the F distribution with $(A, n-A)$ degrees of freedom [20].

5.1.2 *Squared Prediction Error*

The Squared Prediction Error (SPE), often known as the Q-residual, is a measure of the goodness of fit for an observation to a model. SPE is the squared error of the projection of a sample onto the latent variable space [20]. In other words, the SPE measures the loss of information from projecting the sample in the latent variable space spanned in a lower dimension than the original.

$$SPE = \sum_{i=1}^{k} (x_{new,i} - \hat{x}_{new,i})^2 \tag{17}$$

The $100(1-\alpha)\%$ confidence interval upper control limit of SPE [20]

$$\theta_1 \left[\frac{z_\alpha \sqrt{2\theta_2 h_0^2}}{\theta_1} + \frac{\theta_2 h_0 (h_0 - 1)}{\theta_1^2} + 1 \right]^{\frac{1}{h_0}} \tag{18}$$

where z_α is the unit normal deviate corresponding to the upper $100(1-\alpha)$, α is the chance taken to incorrectly declare a fault because of the type I error [20],

$$\theta_i = \sum_{j=A+1}^{m} \lambda_j^i = Tr(\mathbf{E}^i) \tag{19}$$

for i = 1,2,3

$$h_0 = 1 - \frac{2\theta_1\theta_3}{3\theta_2^2} \tag{20}$$

An approximation SPE upper control limit at significance level α based on weighted Chi-squared distribution $\left[g\chi^2(h)\right]$

$$\frac{\nu}{2b}\chi^2_\alpha\left(\frac{2b^2}{\nu}\right) \tag{21}$$

where b is sample mean and ν is sample variance.

It should be noted that a good choice of calibration data is essential for the success of defining satisfying limits for Hotelling's T^2 and SPE control charts.

5.1.3 *Combined indices*

As the Hotelling's T^2 and SPE are two complementing indices, it is often preferable to combine them into a single index for a more practical use in condition monitoring. The Hotelling's T^2 and the Squared Prediction Error can be combined in the following manner [28]

$$\varphi = \frac{SPE(\mathbf{x})}{\delta_\alpha^2} + \frac{T^2(\mathbf{x})}{\chi^2_{l;\alpha}} \tag{22}$$

5.1.4 *The roles of SPE and T^2 in condition monitoring*

The indices T^2 and SPE measure different features for the samples projected to the latent variable space [28]. The SPE measures how well the new sample fits to the in-control model in regard to the modelled covariance structure, while T^2 measures the distance from the projected sample in the latent variable space to the model mean. This is illustrated in figure 7, where we have two original variables X1 and X2, and a PCA model with one principal component PC1. The SPE is the measure of how far samples in the original variable space are from PC1,

and thus how much it breaks the correlation structure. In condition monitoring, one could experience large T^2, but small SPE. In this case one should be careful with false alarms as it does not necessary have to be a fault, but rather a change in the operating condition. The SPE is often the most important index in condition monitoring as it indicates an abnormal situation, and has lower chances of type I and type II errors compared to T^2. The T^2 may also experience false alarms due to non-stationary processes, which makes the choice of alarm limits important.

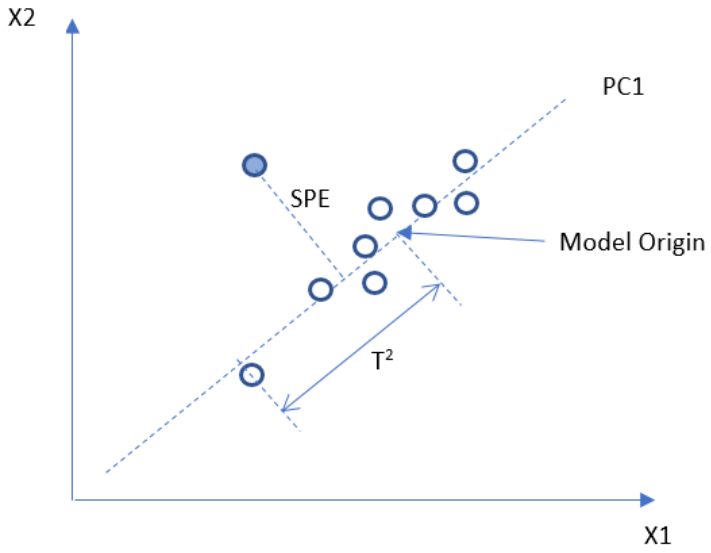

Figure 7: T^2 and SPE illustration

5.2 FAULT ISOLATION

Contribution plots, also known as diagnostic plots, are useful in the process of isolating faults that have been detected by metrics such as Hotelling's T^2 or SPE.

5.2.1 *Contributions to SPE*

When a situation with large deviation of SPE is detected, the contribution for each variable can be plotted in order to isolate the fault [20]. The SPE variable contribution is simply found by,

$$(x_{new,i} - \hat{x}_{new,i})^2 \tag{23}$$

where each contribution can be plotted in bar plot in order to find the variables which contribute the most to the SPE. This

variable can further be investigated in order to diagnose the fault.

5.2.2 *Contribution to Hotelling's* T^2

If experiencing an out of limits value of Hotelling's T^2, variable contributions can be found by first plotting the contributions from each of the scores $(t_i/s_{t_i})^2$. Variable contribution are then found for the scores with most contribution. The contribution from each variable to the score of component q is given by [20]

$$c_j = p_{q,j}(x_j - \bar{x}_j) \tag{24}$$

where c_j is contribution of the jth variable of the given sample, $p_{q,j}$ is the loading and $\bar{x}_j$ is the mean value of the variable. Variables with large contributions *and* with the same sign as the score should be further investigated.

Contribution plots are valuable tools for fault isolation and fault diagnosis, but care should be taken regarding their interpretation as it does not necessary reveal the cause of the fault, but only which variables that are related to it. In many cases where the variables are highly correlated, many of the variable will contribute to the deviation of T^2 or SPE which makes further diagnosis more difficult.

Alcala et al. [7] presented a contribution method for fault diagnosis called reconstruction-based contributions (RBC). This method shows improved performance compared to traditional contribution plots, and guarantees that the faulty variable has the most contribution.

5.3 PLSR

PLSR can be utilized for fault detection through its ability to predict output data Y from input data X. The PLSR analysis converts the input process data X into input latent variables, and output data Y into output latent variables. Through the regression analysis the relationship between the input and output variables is determined, and new process input data can be used to predict the output variables. The PLSR can thereby be used as a soft sensing technique where abnormal condition can be detected [17]. For instance, a PLSR model can be used to predict a particular process parameter from other process data, and if the predicted and the measured parameter deviates, this might indicate a fault.

5.4 FAULT CLASSIFICATION

In some cases classification of fault conditions can be used to automate the diagnosis. Classification is based on pattern recognition of extracted features in order to to identify the condition of the equipment. If a supervised learning approach is used, data from known fault conditions can be used to train a classifier. This can be used for classify future fault conditions.

5.4.1 *SIMCA Pattern Recognition Classification*

The Soft Independent Modelling of Class Analogies (SIMCA) is a classification technique based on PCA. It is a supervised learning approach where distinct features are extracted and modelled in unique corresponding PCA models. When all features of interest are learned, new data is tested for their belonging for each model, and hence new observations can be classified. Samples that pass a F-test based on the sample-to-model distance and sample leverage, are classified to the respective class analogy.

The reader is referred to [13, 14, 35] for more details about the SIMCA method.

5.5 OTHER RELEVANT USES OF MVA TECHNIQUES

5.5.1 *Data compression*

Often the storing capacity for historic process data is limited, in other cases the data transmission rate is limited when data is to be sent, for example from an offshore location to land. In such cases, efficient data compression techniques can be very beneficial. Multivariate analysis and latent variable methods are in this case useful as they function as dimension reduction techniques, utilizing the redundancy in correlated multivariate data to describe the information in the data in a compacted form.

5.5.2 *Visual presentation of hidden patterns in data*

Multivariate data analysis can be utilized for data mining and analysis of large collections of data. Latent variables can be modelled directly from the data itself without prior system knowledge, giving a pragmatic data compression that in combination with good graphical tools can give good predictive ability and good causal insight at the same time [23].

5.5.3 *Intelligent operator alarm systems*

Multivariate statistical process monitoring can be used to implement more intelligent alarm systems. In most cases process parameters are monitored independently in a univariate manner, with defined alarm limits. If many of the variables reach their alarm limit in a short period of time, the system can become uncontrollable as the operators are often unaware of the interdependency between variables [17].

Multivariate techniques handle this problem as they are able do find and model the correlations between variables, which allows for discovering underlying hidden patterns in data. The methods also reduces the dimensions of the variables such that the number of alarms can be reduced, making it easier for the operators to take the right actions.

Multivariate techniques can also give better alarms as they are able to detect deviations in the covariation structure between variables.

5.6 SELF-LEARNING AND ADAPTIVE TECHNIQUES

In some situations the conditions and processes are dynamic and time varying. If the system is very slowly time varying in the scale of months or years, it might suffice to manually retrain the in-control model to capture any new characteristics, or changes in covariance structure or model mean drifts. If the system have several distinct working conditions, such as different loads, startup, shutdown or normal, each situation can be modelled, and the system can automatically detect its actual mode of operation and use the corresponding model.

In other situations having a fixed in-control model might not be sufficient and can give rise to many false alarms. In such situations an adaptive approach is needed. Such systems continuously adapts its in-control model.

5.6.1 *Moving window PCA*

Moving window principal component analysis is an extension of PCA where the model adaptively updates the direction of the principal components, the model mean, and the covariation structure by moving the time-window on-line [18].

Such adaptive systems should however be used with care as choosing an incorrect moving window can lead to the system adapting to faults, rather than detecting them.

5.6.2 *On-The-Fly Processing*

Vitale & Martens et al. [34] have presented a promising proprietary On-The-Fly Processing (OTFP) technology for continuous processing of high-dimensional data streams. The basis for the technology is that reduced-rank bilinear subspace models are developed in order to summarize massive streams of multivariate responses, at the same time as evolving covariation patterns are captured in a self learning manner. Unlike traditional adaptive moving-window methods, past and recent points can be reconstructed and displayed simultaneously [34].

Part III

CASE STUDIES

6

STEAM TURBINE DRIVEN COMPRESSOR

6.1 INTRODUCTION

In the following section of the report is a case study with the use of multivariate statistical techniques for condition monitoring of a steam turbine compressor at Statoil's methanol production facilities at Tjeldbergodden.

Statoil Tjeldbergodden industrial facilities, located at Nordmøre in Norway, comprises of a gas receiving terminal, a methanol plant, an air separation unit and a LNG plant. Natural gas from the Heidrun field at Haltenbanken is received via the Haltenpipe pipeline. The methanol plant is the largest in Europe producing about 900,000 tonnes of methanol per year, while at the same time as being one of the most energy efficient methanol plants in the world [2].

The plant consists of a considerable number of process equipment units being critical for the production, including in particular large rotating machinery such as gas compressors and turbines. Unexpected failures of such equipment can cause unnecessary downtime with large revenue losses and high maintenance costs. Being able to detect equipment deterioration at an early stage through condition monitoring, actions can be taken before the systems reach an unrecoverable stage, saving the company from substantial losses in revenue.

The machinery under monitoring consist of a steam turbine side, which powers a gas compressor side compressing synthesis gas, or syngas, consisting mainly of hydrogen, carbon monoxide, and some carbon dioxide. This unit is of particular importance to monitor since it is critical for production, and a failure of this unit will result in production downtime. It is also a high cost machinery both in terms of capital and operational expenditures. This type of equipment is also highly technically complex and is often weakly understood both in regard to maintenance and operation. Obtaining replacement parts can also be difficult and can result in long delays, which again leads to extended periods of production downtime and revenue loss.

The main motivation for the study was to collect healthy training data in order to model the system under normal operational condition, for then to test condition monitoring abilities of this model from test data from known fault conditions. The

data-driven approach to this problem was motivated from the practicality and effectiveness of multivariate statistical methods compared to developing detailed physical models of complex processes, which can be difficult [38].

6.2 DATA ACQUISITION AND PRE-TREATMENT

In this study, a total of 55 different measurement points were available for the analysis, where historic data was available from a historic process database. Among the measurements included are condition monitoring instrumentation such as vibration and temperature measurements on each bearing as illustrated in figure 8 and described in table 1. The vibration measurements are based on proximity sensors, where the vibration signals are represented in the form of peak-to-peak displacement. Each bearing has vibration transducers both in the vertical and horizontal direction, and is complemented with bearing temperature sensors.

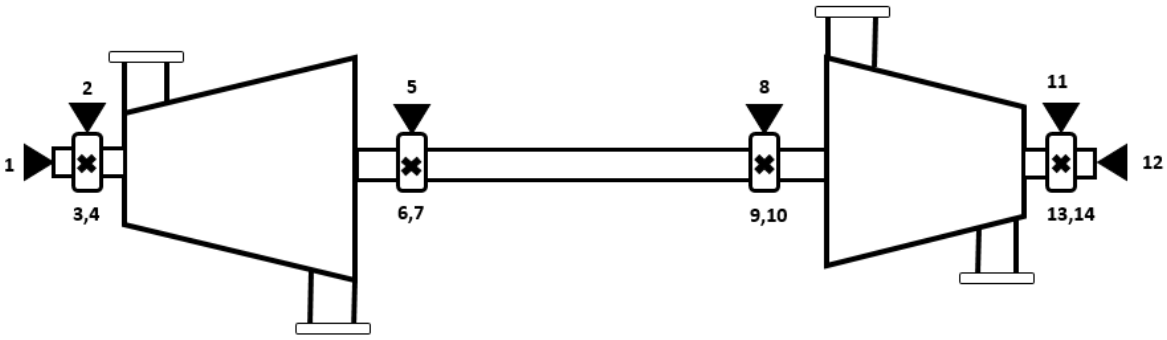

Figure 8: Principle illustration of measurement points for condition monitoring of bearings

In addition to the dedicated condition monitoring instrumentation, a variety of process parameters were included in the analysis. This includes parameters such as rotating speed, power, flow rates, pressure, temperature and so on.

The data was extracted with a sampling interval of 12 seconds and is interpolated.

In this case study, multivariate condition monitoring methods were applied on data from a specific event that occurred just after start-up of the machine after it had been off production due to maintenance service. After start-up and a stationary speed had been reached, the machine ran for approximately two days until it automatically tripped due to a gas leakage in a seal in the compressor. Some of the limitations from the data

Equipment	Measurement No.	Measurement Points	Measurement Type
Turbine	1	Free-end axial	Displacement [mm]
	2	Bearing, free-end vertical	Vibration [µm peak-to-peak]
	3	Bearing, free-end horizontal	Vibration [µm peak-to-peak]
	4	Bearing, free-end	Temperature [°C]
	5	Bearing, load-end vertical	Vibration [µm peak-to-peak]
	6	Bearing, load-end horizontal	Vibration [µm peak-to-peak]
	7	Bearing, load-end	Temperature [°C]
Compressor	8	Bearing, load-end vertical	Vibration [µm peak-to-peak]
	9	Bearing, load-end horizontal	Vibration [µm peak-to-peak]
	10	Bearing, load-end	Temperature [°C]
	11	Bearing, free-end vertical	Vibration [µm peak-to-peak]
	12	Free-end axial	Displacement [mm]
	13	Bearing, free-end horizontal	Vibration [µm peak-to-peak]
	14	Bearing, free-end	Temperature [°C]

Table 1: Measurement points table

from this specific event is that the machine had not reached a fully steady state of operation, as some slow dynamics had not fully settled, and there had been some minor changes in the setpoint of the rotational speed.

Prior to the further analysis, the data needed to be pre-treated. Firstly, the data was smoothed using a causal moving average filter to increase the signal-to-noise ratio and reduce the effect of outliers in the signal. A window size of 10 was chosen. Filtering the process data should however be done with care as is can be detrimental as it can suppress high frequency signal components and destroy the multivariate nature of the data [8].

Each column, representing variables of the data matrix, was subtracted by its mean value in order to center the data. Since the variables in the data set is a collection of different measurements with different units, performing an unscaled multivariate analysis would not prove to be sufficient as some variable will dominate the analysis completely, while other will have little contribution. To solve this, the data set was normalized by dividing each column with its corresponding standard deviation. After these steps, the data was mean centered which is important for having a meaningful PCA, and the normalized variables had equal influence on the model. Great care should be taken when normalizing each variable in the calibration set with their respective variance, as this can result in noise amplification of variables with low variance.

Another challenge is whether all variables should be included an the multivariate analysis, or if any of the variable should be weighted in a particular manner. In the exploratory MVA phase, different subsets of variables were included in the analysis in an iterative manner in order to find the selection with best per-

formance. Some variables tended to include more noise than others, and was excluded from the analysis.

The correlation matrix in figure 9 shows that the degree of correlation between the variables used in the analysis is considerable. This expected as much of the same process phenomena and physical laws affect several measurements in a certain manner such that they are correlated to each other.

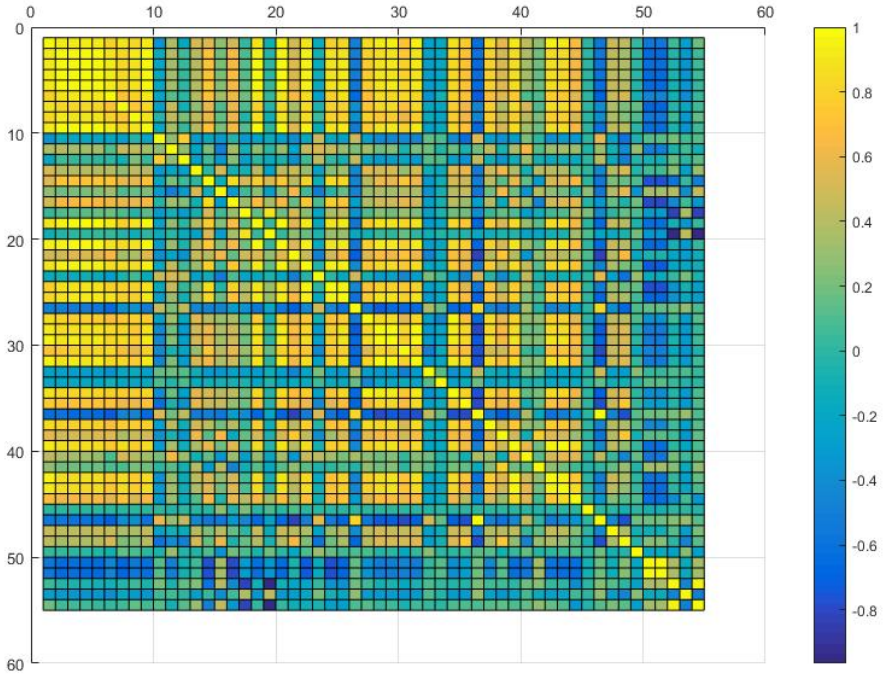

Figure 9: Correlation matrix of the 55 measurements monitored

6.3 PCA MODELLING OF NORMAL CONDITION

Several considerations must be made when making a calibration- or in-control model for use in condition monitoring. The data used to calibrate the model should most optimally span the entire range of variation for normal condition. Data from periods including any abnormal patterns should also be excluded from the calibration set.

In the exploratory analysis phase, several subsets of variables for different periods were explored in the calibration models in order to find the most optimal performance. Two different in-control PCA models are included in this paper. One of them includes all 54 variables with 5200 samples, making a 5200×54 calibration data matrix X. In the other model, one variable for a bearing temperature which showed an abnormal characteristic was removed, making a 2000×53 data matrix. The two

calibration data sets cover different periods of operation. The first model, including all 54 variables, was calibrated with a longer period of samples equivalent to about 20 hours, which included some dynamical variation of the machine. The second model was calibrated with samples from a period of about 7 hours, where the machine had a more stationary operation.

In condition monitoring, choosing model complexity is of high importance as an overfitted model will include the noise from the training data. As the intention is to detect deviations from the normal condition, having an overfitted model would imply that the observed data need to have the same noise characteristic as the training data in order to avoid false alarms. Underfitting the model by retaining too few principal component can on the other hand lead to distortions and give unsatisfying results.

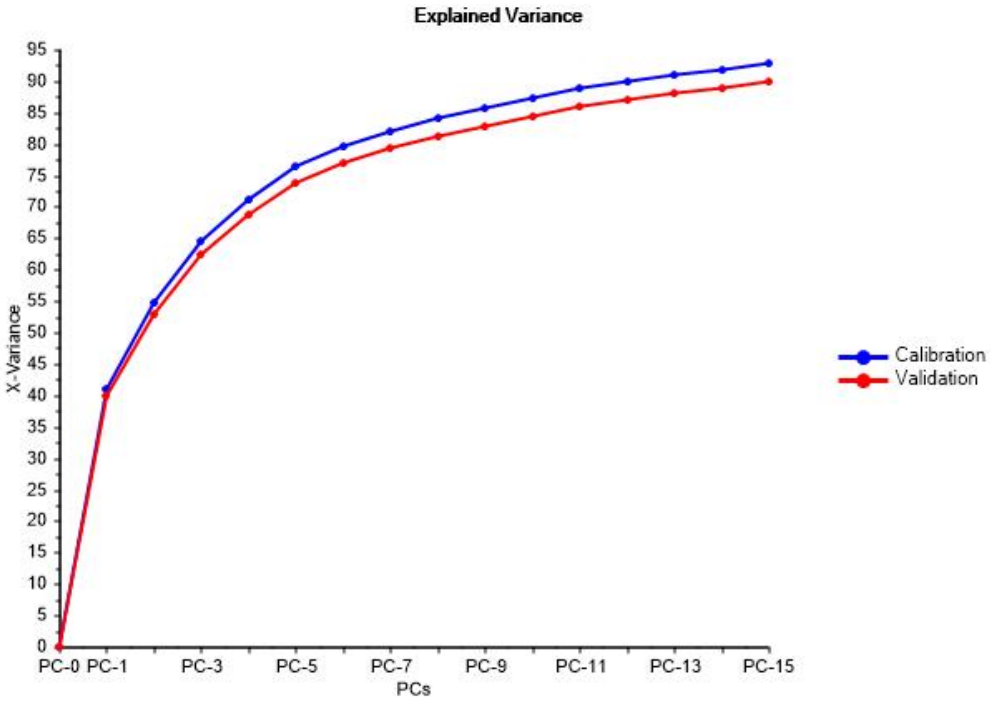

Figure 10: Explained variance of the first PCA in-control model with 54 variables

Figure 10 shows the explained variance of the training first PCA model as a function of the number of principal components added. We see that the 11 first components retain almost 90% of the explained variance with good validation, which might indicate that it is a good model complexity. The calibration models were validated using cross-validation divided on 20 segments with random selected samples in each validation segment. See Chapter 3 for a more detailed description of cross-validation. We also see that the PC1 explains about 40% of the

variance in the training data, revealing some of the considerable amount of structure in the data.

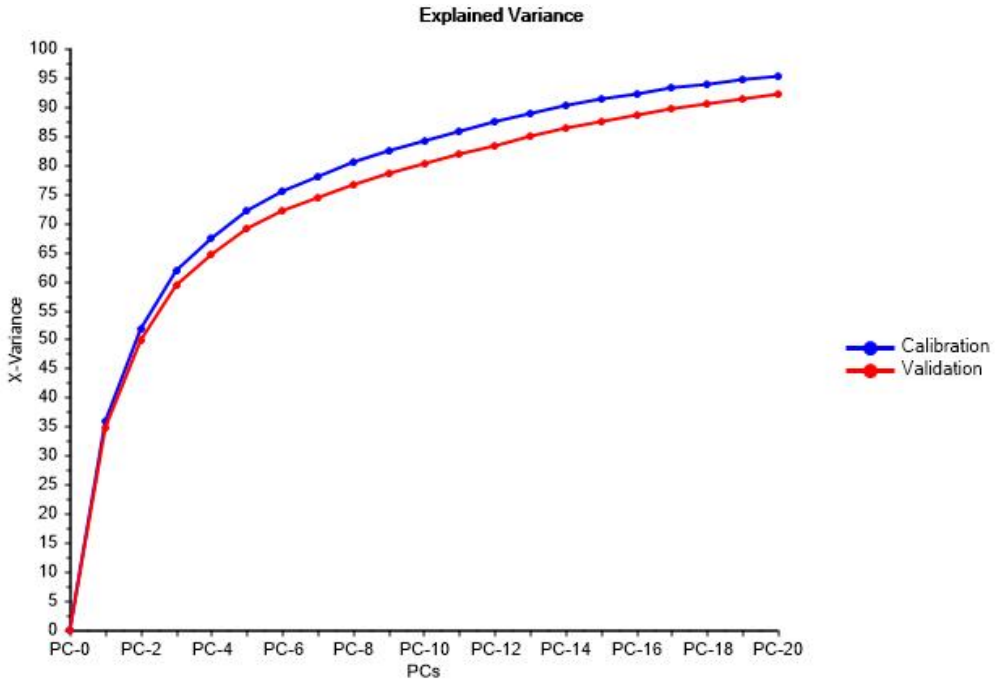

Figure 11: Explained variance of the second PCA in-control model with 53 variables

The explained variance for the second PCA calibration model with 53 variables are shown in figure 11 and reveals that the first principal components are less dominating and more components are needed in order to describe the same percentage of variance in the data. 16 components gave the best validating performance, but since the calibration data have less variation in this case, one should be careful not to overfit the model by adding too many noise dominated principal components.

6.4 FAULT DETECTION AND ISOLATION

For fault detection, new observations were projected on to the healthy in-control models, and based on multivariate control charts, this can detect deviation from normal condition indicating that a fault has occurred. The indices used for fault detection in this study are the Hotelling's T^2 statistic and the Squared Prediction Error (SPE), also known as Q-residual, as described in Chapter 5.

6.4.1 *First in-control model with 54 variables*

The first in-control model is projected with 5350 samples corresponding to about 18 hours of operation, from the end of the calibration data, until the time where the compressor seal leakage is clearly apparent in the differential pressure reading.

Figure 12 shows a score plot of projected samples onto PC-1 and PC-3 as a fault condition is developing. The plot shows that the projections remain in the 1% confidence limit for Hotelling's ellipse for PC-1 and PC-3 for the majority of time. At the last samples of the projection however, the samples start to move far beyond the limits for PC-3, indicating that a fault has occurred.

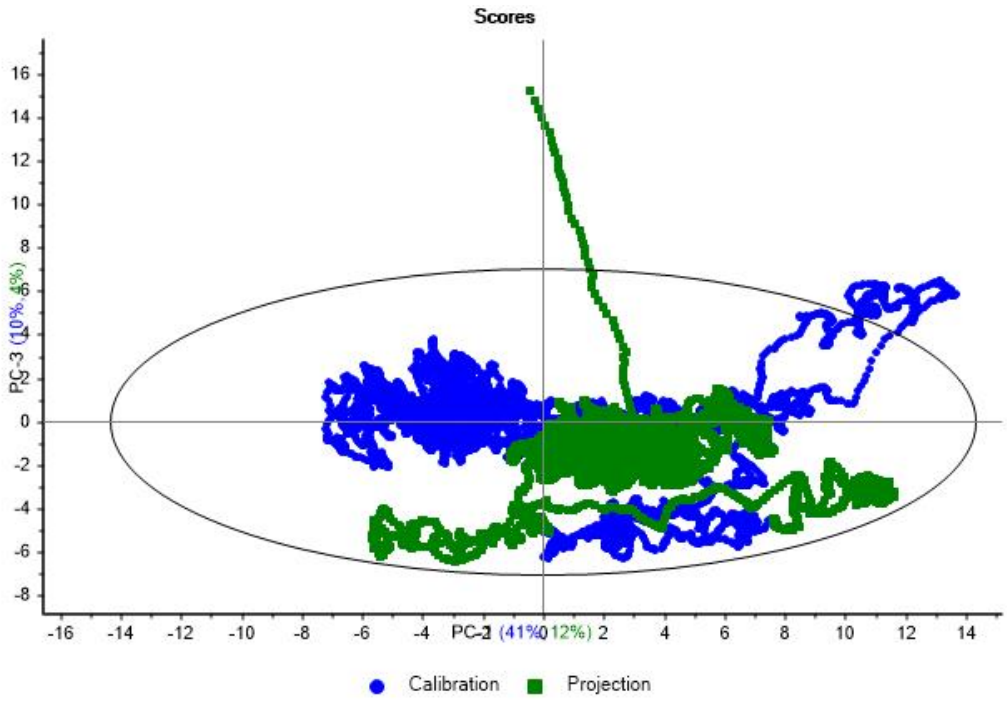

Figure 12: Score plot of projection for fault period on first PCA in-control model with 54 variables

Figure 13 shows the correlation loadings plot PC-1 versus PC-3 for the same model. It can be seen that there is a high degree of clustering of the variables which reveals that they are strongly correlated. In particular, variable number 1, representing the rotational speed of the machine is both correlated and negatively correlated with many of the variables.

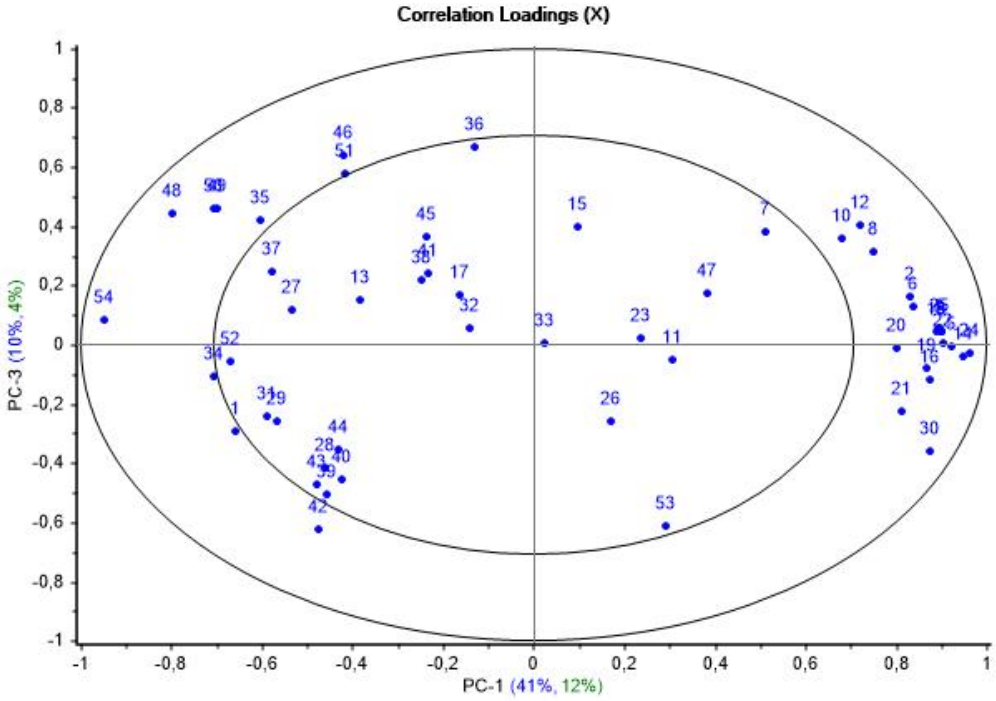

Figure 13: Loadings plot of projection for fault period on first PCA in-control model with 54 variables

Figure 14 shows the Hotelling's T^2 statistic for the same projection for 11 principle components. Around sample number 1200, T^2 starts to violate its 1% confidence limit considerably, showing an abnormal behaviour. This large variation in T^2 continues in the consecutive samples until the end of the projection period where the statistic increases beyond all bounds. The extreme violation of T^2 in the last samples is a clear detection of a fault that is appeared within the machine.

The SPE, or Q-resiual plot in figure 15 for the same projection shows an almost identical pattern compared the Hotelling's T^2 statistic. Already from around sample number 400, the SPE start to violate its 1% confidence limit, and near the end of the projection period, it increases to the extreme.

From the Hotelling's T^2 it is clear that the projected samples deviates far from the normal operating region for the normal condition within the model. The SPE plot reveals that many of the samples in the projection also heavily break the correlation structure from the in-control model.

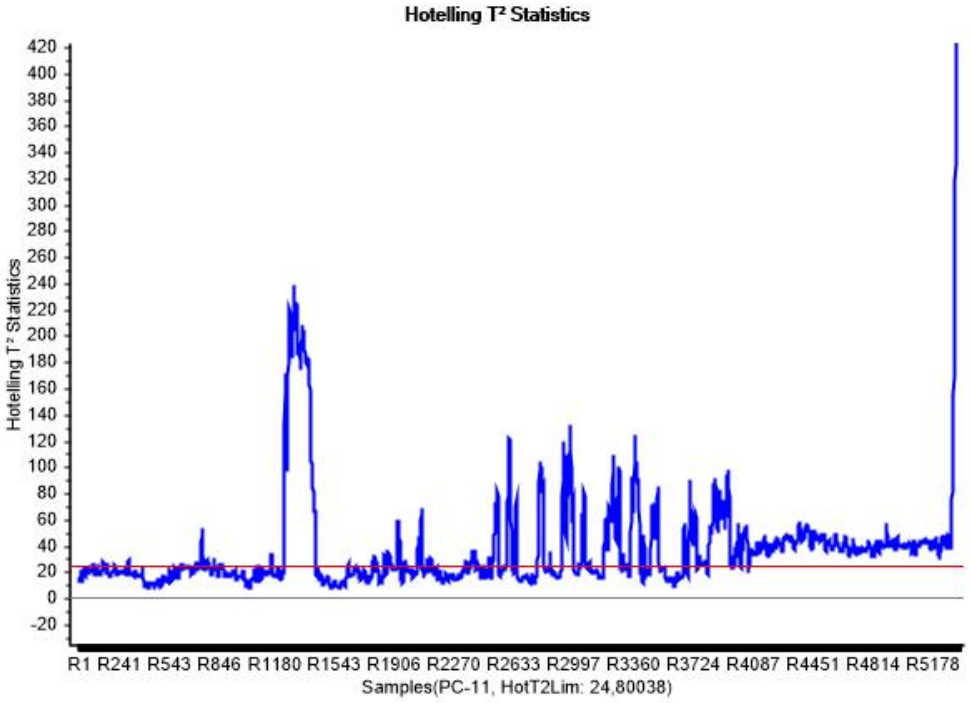

Figure 14: Hotelling's T^2 of projection for fault period on first PCA in-control model with 54 variables

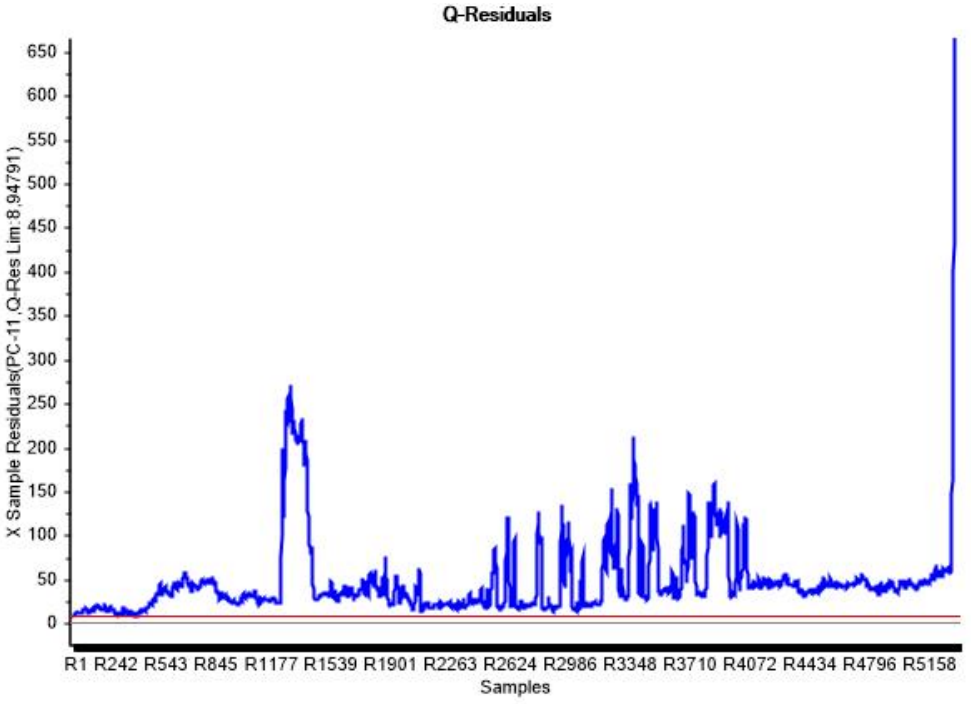

Figure 15: SPE, or Q-residuals, of projection for fault period on first PCA in-control model with 54 variables

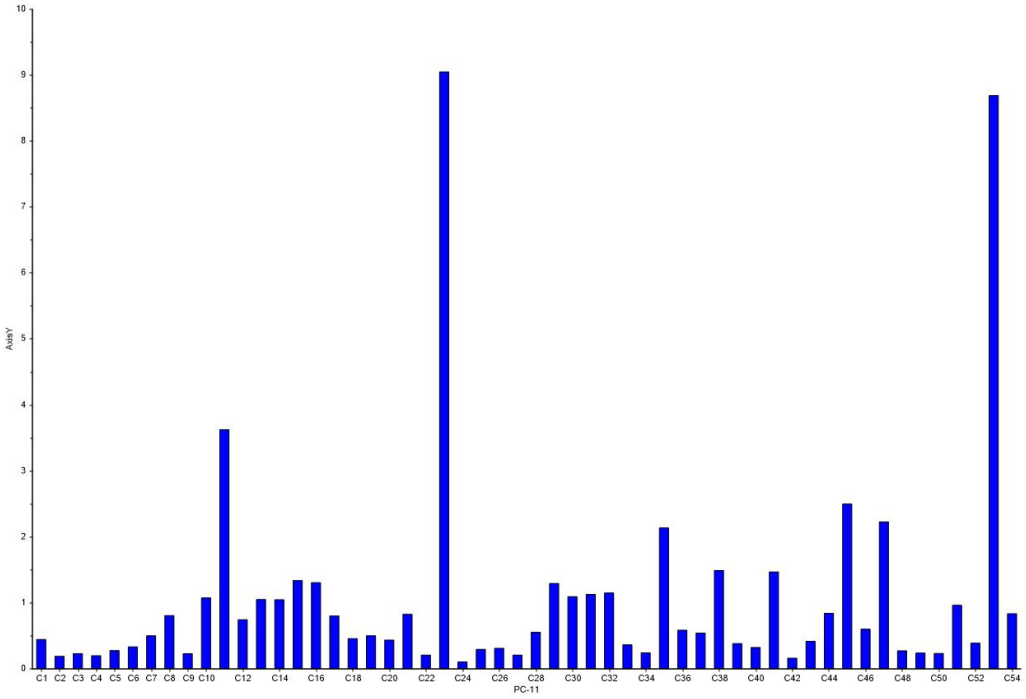

Figure 16: Variable contribution of residuals for projection of fault period on first PCA in-control model with 54 variables

After the abnormal situation have been detected from T^2 and SPE, it is of great interest to be able to isolate the fault in order further diagnose the fault condition. Fault isolation is helped with the use of variable contribution plots, also called diagnostics plots. Figure 16 shows the variable contribution for the residuals for the projection. It can clearly be detected that the largest contributions are from variable number 23, corresponded to a compressor bearing temperature, and variable number 53, corresponding to the differential pressure of inert gas in the compressor seal. A closer investigation of these variables reveal that the compressor bearing temperature at variable number 23 contributes to T^2 and SPE with the large spikes starting projection sample number 1200. The cause of abnormal behavior of this bearing temperature is however unclear, but the methods prove successfully in both detecting and isolating this fault.

The pressure difference reading in variable number 53 is the main contributor to the extreme violation of these indices in the last projection samples. This is caused by a drop in the difference pressure over the sealing due to the leakage. This fault is detected in a very early stage through large deviations in the scores, as detected by Hotelling's T, and heavy violation of the correlation model structure, indicated by SPE. The fault is detected about 1 hour and 40 minutes before the trip limit

for the pressure is reached. The fault is also detected before the univariate alarm limit is reached.

6.4.2 *Second in-control model with 53 variables*

A second calibration model was made where bearing temperature variable number 23 with its abnormal behavior was excluded from the analysis, as the cause of these deviations are unclear.

This new calibration model was projected with 2350 new samples corresponding to the period of about 8 hours, just after the calibration period and until the fault is clearly apparent from the pressure difference reading in variable 53.

Seen in figure 17 is the Hotelling's T^2 statistic of the projection including 16 principal components. The large spikes caused by the abnormal bearing temperature in the previous model are here not apparent. The T^2 statistic is however increasing beyond its 1% confidence limit from around sample number 700. The same phenomena is apparent in the SPE plot in figure 18, where the SPE is gradually increasing until the rapid increase occur at the end of the projection period.

Figure 19 show the variable contributions to the residuals. As expected, variable 53 corresponding to the seal difference pressure is dominating in contributing to the deviation, indicating it the cause of the fault condition.

It is however unclear if the gradual increase in Hotelling's T^2 and SPE is due to the developing fault condition or due to the limitations of the PCA model. The empirical linear PCA soft-modelling is not necessarily sufficient in modelling the dynamical and non-stationary nature of this data set. The calibration of the in-control model should optimally be performed on data from more stationary conditions of operation.

The methods are however successful in clearly detecting the fault caused by the seal leakage at an early stage.

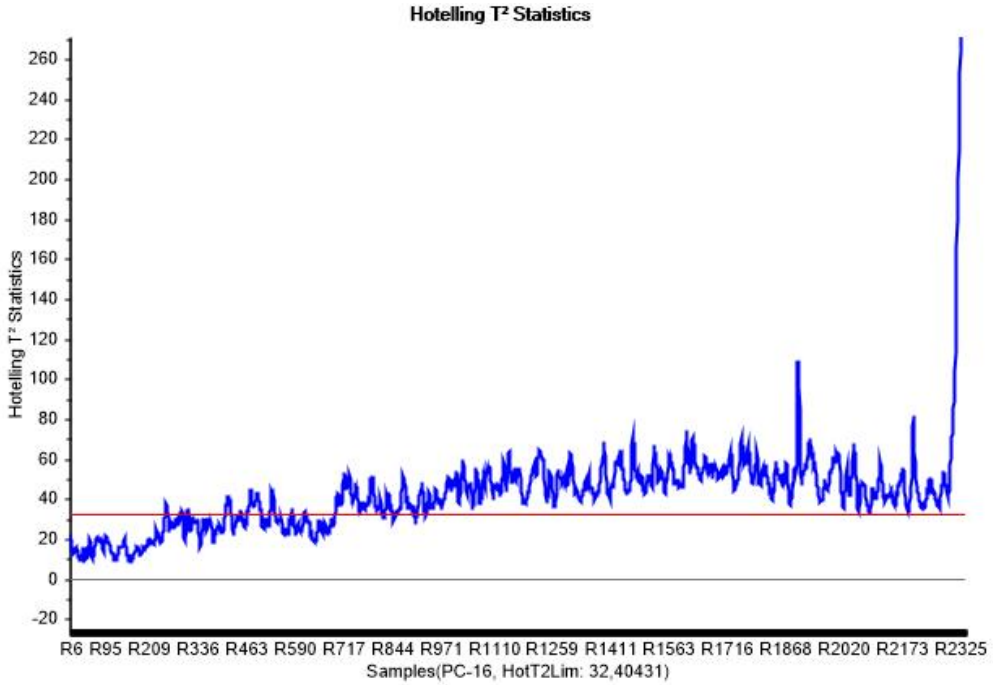

Figure 17: Hotelling's T^2 of projection of fault period on second PCA in-control model with 53 variables

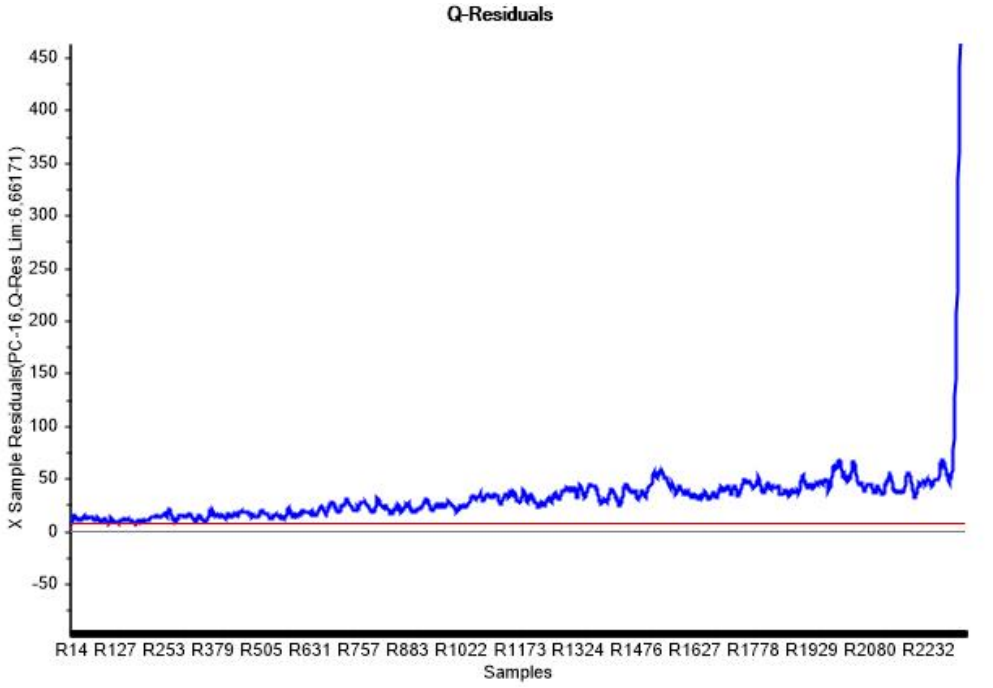

Figure 18: SPE, or Q-residuals, of projection of fault period on second PCA in-control model with 53 variables

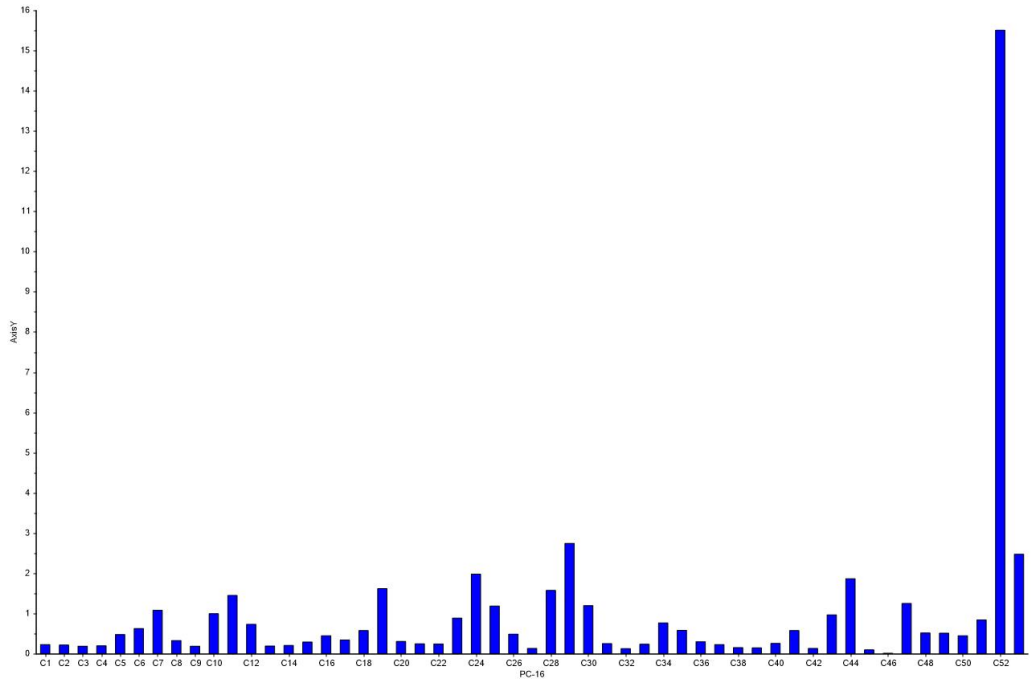

Figure 19: Variable contribution of residuals of projection of fault period on second PCA in-control model with 53 variables

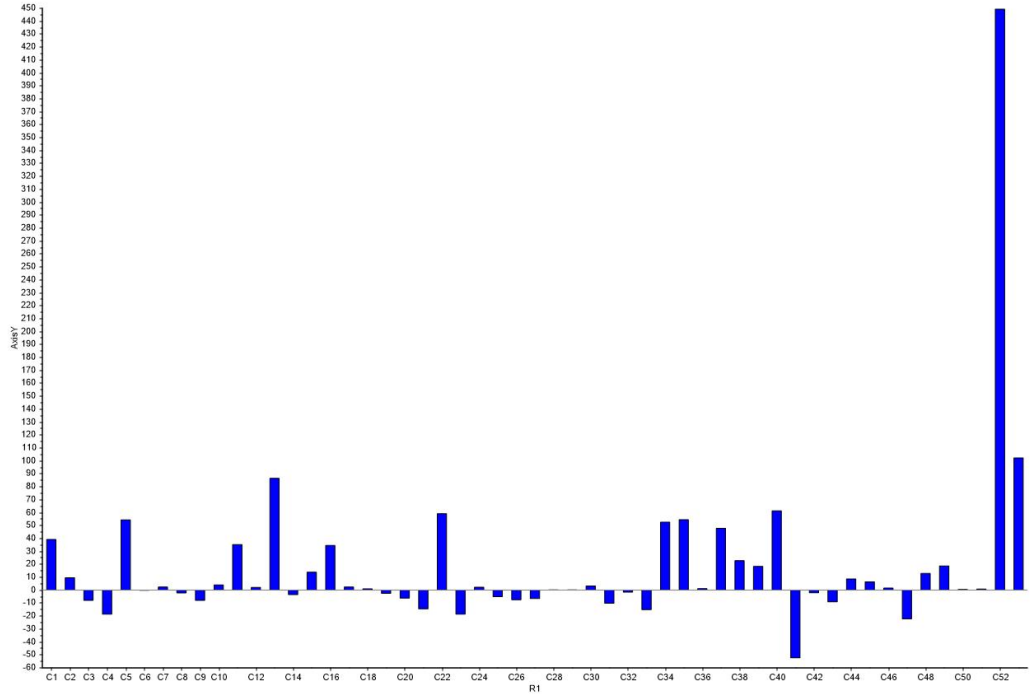

Figure 20: Variable contribution plot for Hotelling's T^2 for a sample in the fault situation

6.5 PLSR

In addition the PCA, which is based only on one data set X, PLS regression can be used to predict output variables Y from

input variables X. In this case two PLS models were made, one model where Y corresponded to the sealing pressure difference variable 53, and another case where Y corresponded to the bearing temperature of variable 23. Training data was used to calibrate the PLSR model used to predict Y from X. In each of the two cases, these predictions were then compared to their actual measurement data in order determine if any component in the process had changed.

Figure 21 shows the predicted pressure difference reading in the compressor seal for the fault period compared to the actual measurement. At around sample number 2250 the predicted and actual measurement start to deviate heavily. This is a clear indication that a fault has occurred which the in-control model did not take into account.

Figure 22 shows the predicted versus the actual bearing temperature from variable 53. From sample number 700, the predicted and actual bearing temperature starts to deviate indicating a faulty condition.

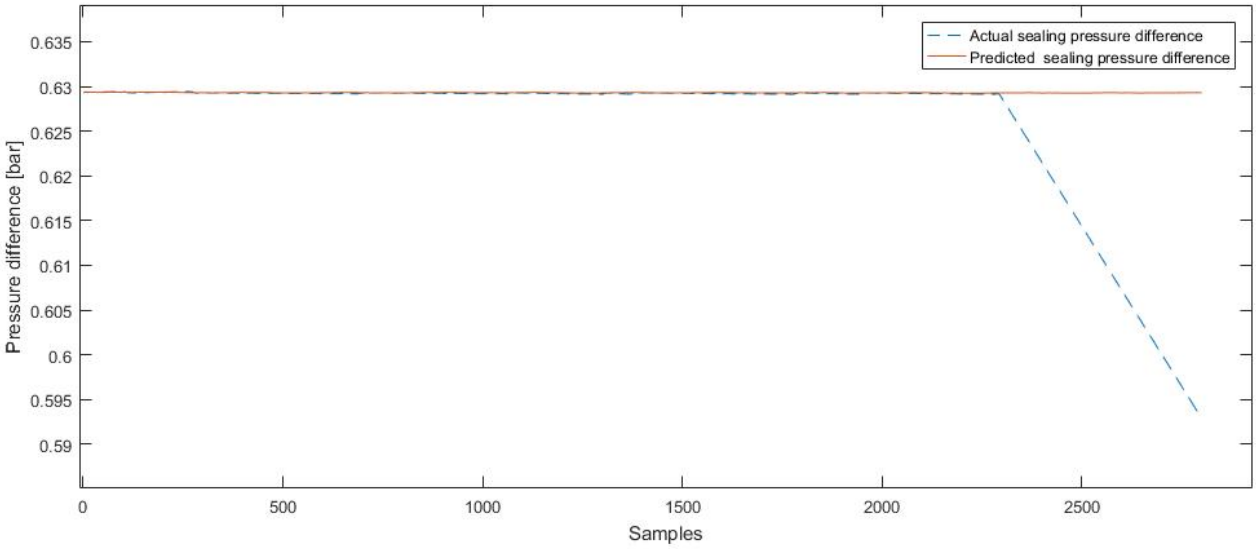

Figure 21: Loadings plot of PLSR model of differential sealing pressure

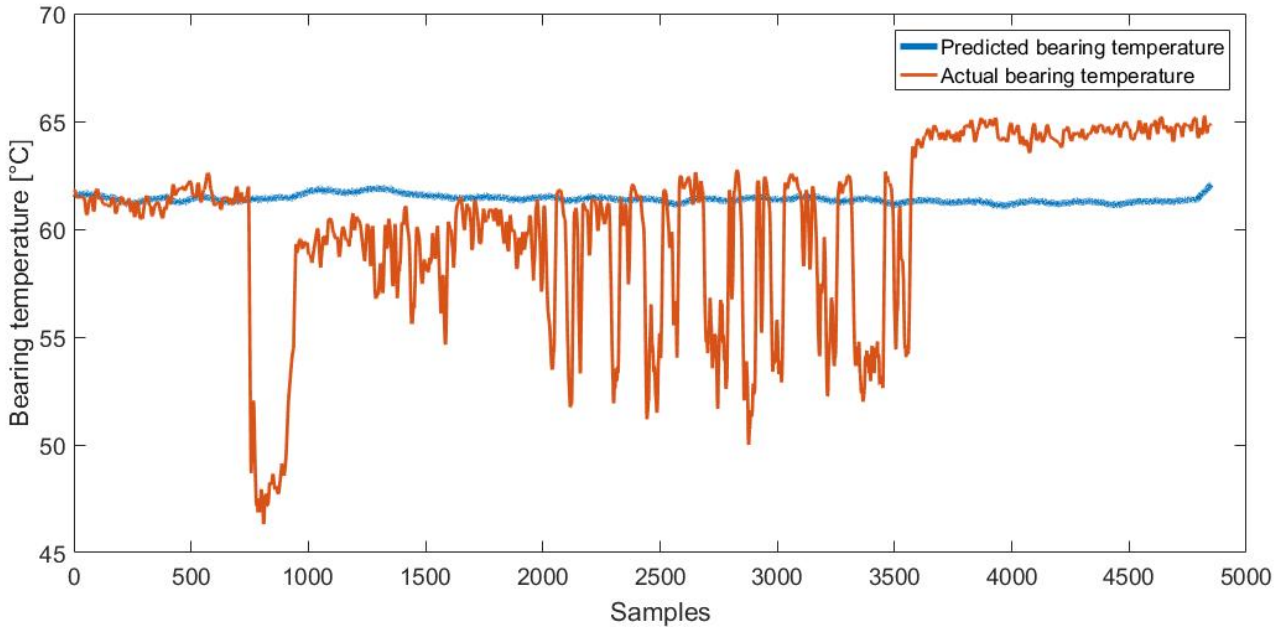

Figure 22: Predicted versus actual PLSR model for bearing temperature

7

STEAM TURBINE GENERATOR

7.1 INTRODUCTION

This section involves analysis of another large rotating machine at Tjeldbergodden, involving a steam driven turbine generator. This machine is also a type of equipment of high criticality, also being highly complex with large investment and operational cost. The benefits of intelligent condition monitoring systems would therefore be of high value to the operation and maintenance of this machine.

The machine had experienced a declining efficiency over several years, mainly due to internal steam leakage in the turbine. A major overhaul was performed where the rotor, inner-casing and labyrinth seals were replaced. The main motivation for this analysis was to explore any trends in the condition for machine over several years prior do this overhaul.

7.2 DATA ACQUISITION PRE-TREATMENT

Data from a historic process database collected from 6 periods during approximately 2 years, each period corresponding about 48 hours. These periods include June 2014, October 2014, February 2015, July 2015, December 2015 and May 2016. Considerable maintenance operation was performed in May 2016, and operational data after repair in August 2016 was also collected. The sampling was done a 1 minute intervals, and the data was interpolated.

Around 60 variables were available for the analysis, however, data for many of these variables were unavailable in the period of interest. Only 26 of these variables are therefor included in the analysis, consisting mainly of variables such as bearing temperatures, lubrication system parameters and other process parameters. A cleansing of the data was performed in order to remove any obvious outliers and faults due to instrumentation. Since many of the bad samples were removed from the analysis, the different periods consisted of different number of samples. Before analysis, the data was also mean centered and normalized in order to have mean centered PCA where all variables have equal contribution.

7.3 QUANTITATIVE DATA EXPLORATION USING PCA

Using PCA for data-driven analysis and empirical modelling is a promising approach as it gives the ability to learn different operating condition of the machinery without the need for any physical understanding of the process. PCA is a valuable technique for compressing a large amount of information in a compact and easily interpretative form. Score plots can give a window into the latent variable space to detect the underlying patterns in the data.

For this purpose all six periods during approximately two years was used to build a PCA model with 7 PCs. Figure 23 show a score plot of PC1 versus PC-3, which reveals a clear trend in the six consecutive periods numbered in chronological order. The data after the overhaul is also projected on this model for comparison.

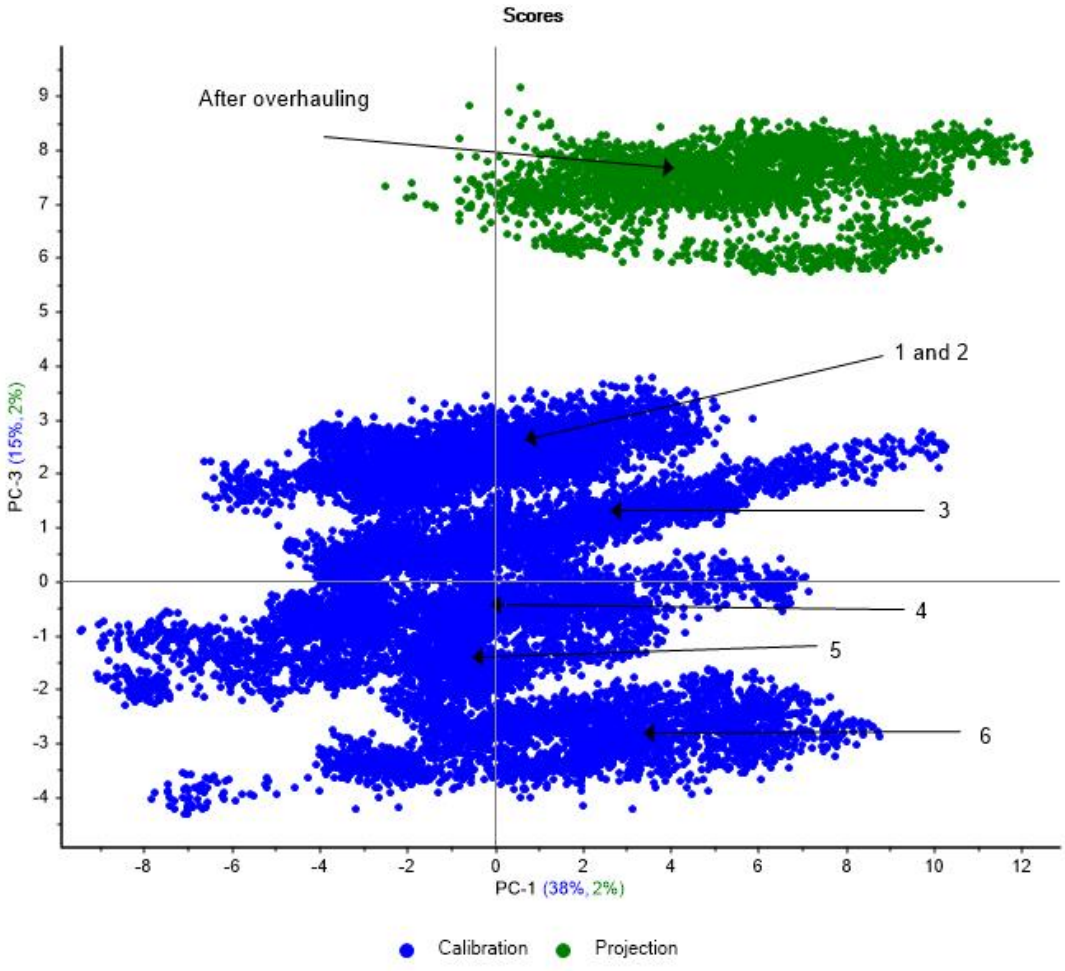

Figure 23: Score plot of PC-1 against PC-3

Figure 24, 25 and 26 shows score plots of PC-1, PC-2 and PC-3 respectively for the six periods prior do maintenance. We see that PC-1 and PC-2 shows little tendencies of any trends

over time, but describes the major dynamic variations in the data set. As previously indicated by the PC-1 versus PC-3 score plot in figure 23, PC-3 shows a clear trend with a continuously decreasing score in all 6 periods.

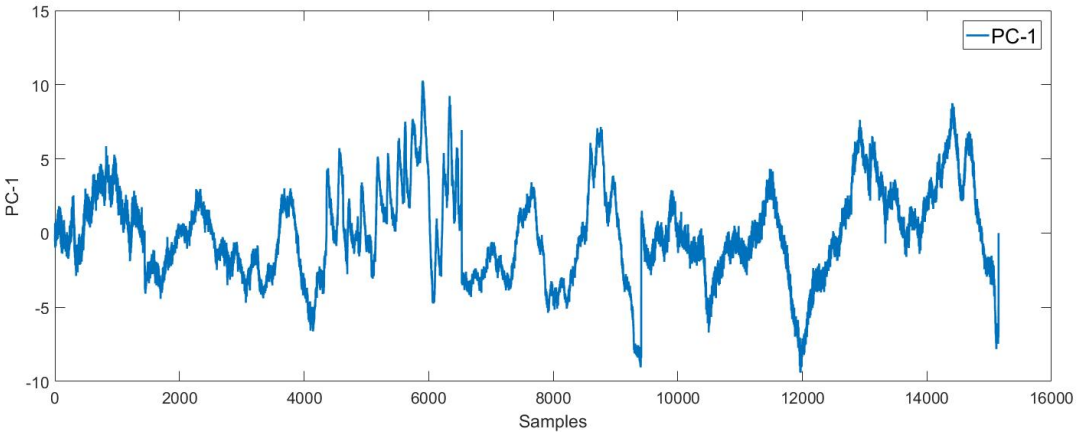

Figure 24: Score plot of PC-1 prior to overhauling

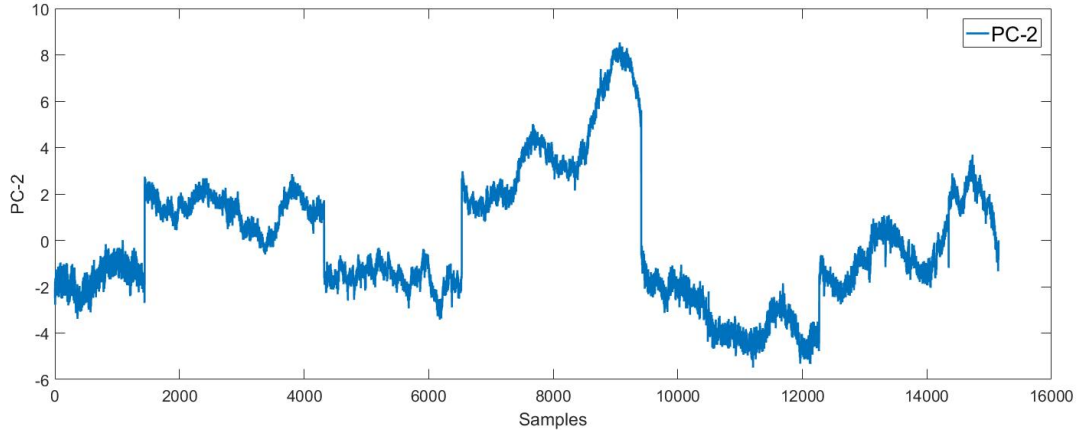

Figure 25: Score plot of PC-2 prior to overhauling

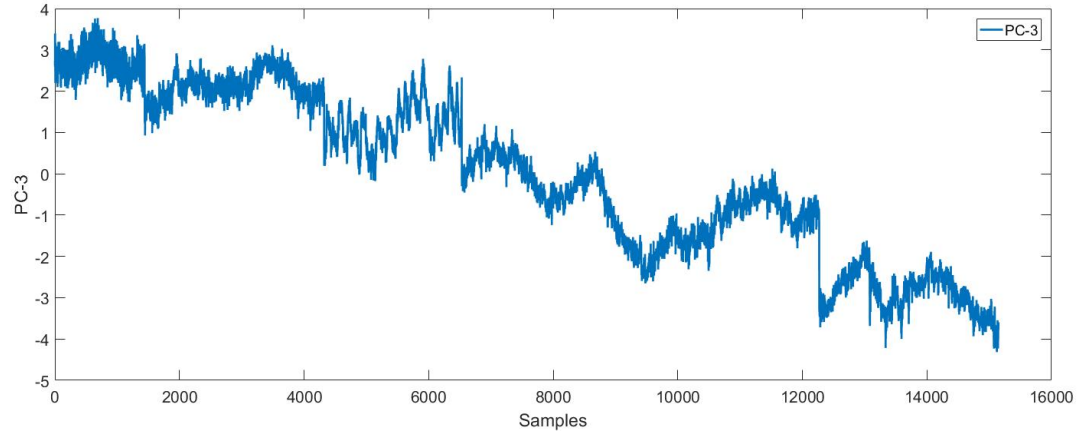

Figure 26: Score plot of PC-3 prior to overhauling

7.4 DEGRADATION TRENDING

Figure 27 shows a plot of the mean values of the scores for PC-3 of each period in the analysis. The plot illustrates the clear trend of the scores for PC-3. This trend may further be investigated if it has any causal relationship with the degradation in performance which has been experience over several years.

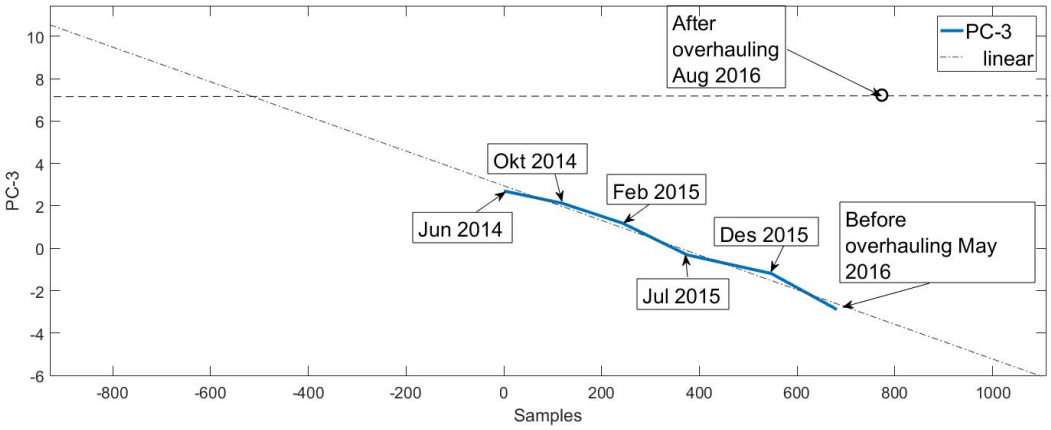

Figure 27: Trend for principal component 3

Being able to utilize this time series trend in order predict future condition of the machine would be of high value in regard to predictive maintenance. Time series analysis and modeling, such as autoregressive integrated moving average (ARIMA) model can be used for this purpose. ARIMA is a autoregressive moving average modelling technique based on the autocorrelation structure of time series data, which can be used forecast futures points in the series. Data-driven estimation of remaining useful life is however not be covered in this thesis, and the reader is referred to paper such [21] for more details.

The findings of this study is that the multivariate analysis methods applied on the turbine generator has shown that trends in the operating condition can be detected from otherwise hidden patterns in relatively large data matrices.

8

BEARING VIBRATION ANALYSIS

8.1 INTRODUCTION

This case study explores condition monitoring techniques for ball bearings based on vibration analysis. Bearings are essential components to support and locate rotating shafts in machines. The ball bearing studied here is a type of rolling-element bearing which operate with a rolling action as opposed to the sliding action of plain bearings. Rolling-element bearing health can deteriorate or fail due to causes such as improper load, operation or lubrication, and eventually all bearings will fail due to fatigue of the material [19]. Faults in such bearings starts as discontinuities in the bearing raceway or on a rotating element, and will grow and spread with time. Such faults give rise to changes in the vibration characteristics from the bearing, as every time the rolling-element passes a discontinuity, a pulse of vibration results. The rate of these pulses are dependent on the location of the discontinuity in the bearing, and such passing frequencies can be determined based on the geometry of the specific bearing. Trying to detect such passing frequencies from vibration frequency spectra can be difficult as they can be "buried" in noise. The pulses generated will however amplify the vibration signal in higher frequencies between 1 kHz and 20 kHz due to structural resonances [19].

Being able to monitor the condition of such bearings would be of high value for condition based maintenance strategies.

The main objective for this case study is to utilize Principal Component Analysis methods for fault diagnosis including fault detection, feature extraction and feature classification based on features both from time domain and time-frequency domain based on wavelet packed decomposition.

8.2 DATA ACQUISITION

Data is provided from The Society for Machinery Failure Prevention Technology, Vibration Institute, not-for-profit organization. Data Assembled and Prepared on behalf of MFPT by Dr Eric Bechhoefer, Chief Engineer, NRG Systems [3]. Data of bearing vibration have been collected for different conditions, including both normal and faulty conditions. A bearing test rig was used, where different faults were induced to the test bear-

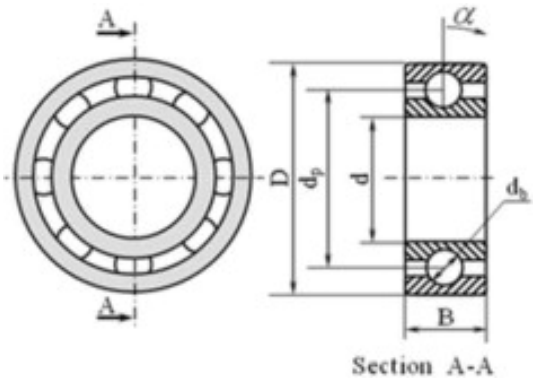

Figure 28: Geometry of a ball bearing. Figure from [29]

ings. Small grooves were made in the inner and outer race of the bearing in order to physically simulate the respective fault conditions.

The parts of the dataset from [3] included in this study is:

- Two baseline conditions for healthy normal condition, 270 lbs of load, input shaft rate of 25 Hz, sample rate of 97,656 Hz, for 6 seconds
- Three outer race fault conditions of varying severity: 270 lbs of load, input shaft rate of 25 Hz, sample rate of 97,656 Hz for 6 seconds
- One inner race fault condition reading: 200 lbs of load, input shaft rate of 25 Hz, sample rate of 48,828 Hz for 3 seconds

8.3 FILTERING

To increase the signal-to-noise ratio, a number of filtering techniques can be utilized, such as moving average filter, or high pass, low pass or band pass.

The filtering technique used in this study is called wavelet denoising, or wavelet thresholding, and is a nonlinear filtering technique that can show improved results as it better preserves sharp signal changes. The concept behind this method is that it utilizes the wavelet transform to concentrate the signal features in fewer large-magnitude wavelet coefficients. Small valued coefficients are thresholded or removed in order to remove the noise in the signal but still preserving the quality and important features. The signal can then be reconstructed using inverse wavelet transform [4]. The wavelet function used was the *Coiflot* of order 5 at level 4. Maximal overlap discrete wavelet transform was performed, and the thresholding was based on

soft universal threshold selection rule $\sqrt{2\ln(\cdot)}$. See the raw signal in figure 29 compared to the reconstructed denoised signal in figure 30.

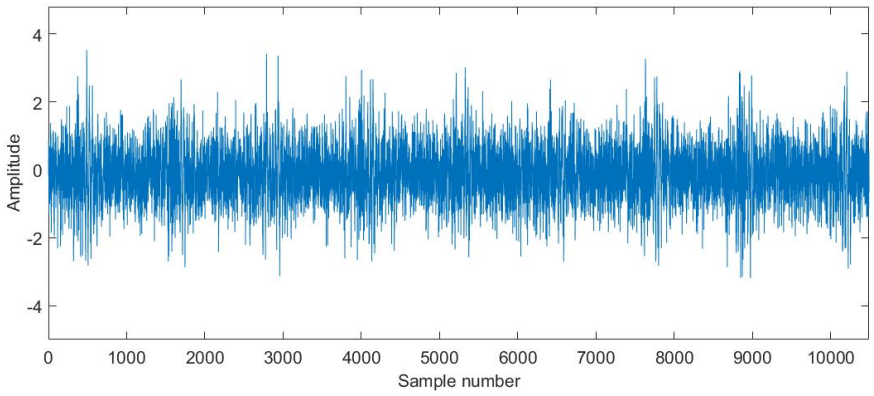

Figure 29: Raw vibration time series signal

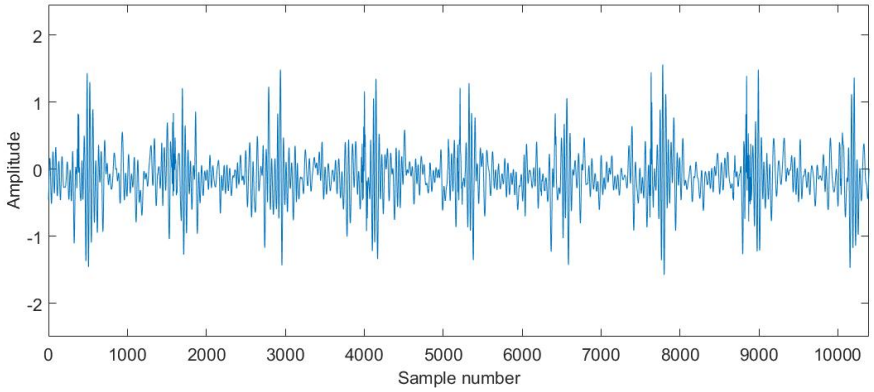

Figure 30: Vibration signal denoised using wavelet thresholding

8.4 FEATURE EXTRACTION

8.4.1 *Wavelet Packet Decomposition*

To be able to extract features from the frequency domain and at the same time achieve time-localization for real-time monitoring of non-stationary signals, we want to decompose the signal in both time and frequency. A time-frequency representation in the form of a spectrogram based in Short Time Fourier Transform (STFT) is illustrated figure 31.

As the STFT suffers from a weak time resolution and cross window contamination, a more optimal technology is the discrete wavelet transform. The stretching and translation through wavelet transform provides a good time domain resolution of high frequency parts of the signal is obtained, as well the lower

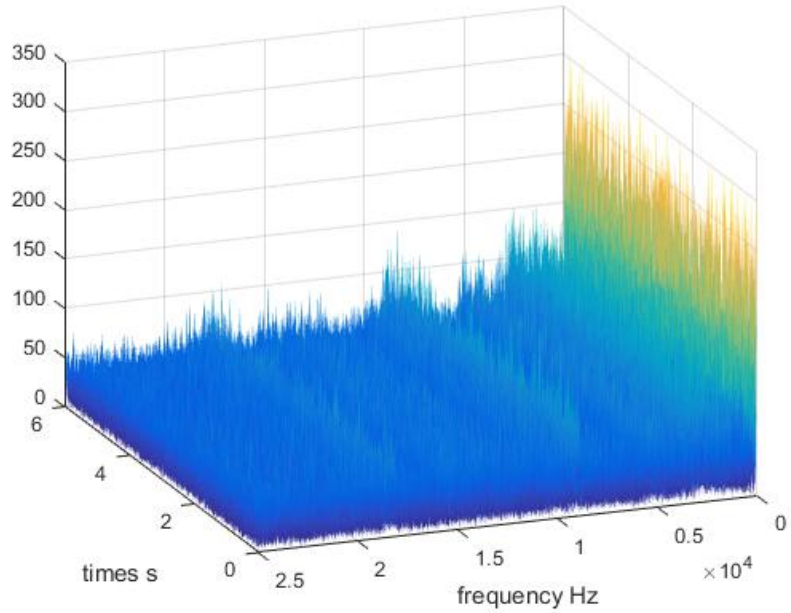

Figure 31: Spectrogram of raw signal based on Short Time Fourier Transform

frequency parts preserve good frequency domain resolution. Wavelet transform techniques are relatively new, and have in recent years expanded in use for applications such as signal and image analysis, pattern recognition, quantum mechanics, fault diagnosis and other fields [6, 30].

Wavelet packed decomposition is a type of wavelet transform where the signal is passed through more filters than DWT, allowing for a finer multilayer frequency band division, see section 4. This is therefore the technique applied in this study.

Several wavelet functions was tested, and the real *Coiflot* of order 5 was chosen for this analysis. This wavelet basis function have shown to be successful for bearing vibration analysis [15]. The decomposition was based on Shannon entropy.

The level of the decomposition was chosen in regard to the wanted level of scale or frequency band division that was to be utilized as features for the condition monitoring. A level of 7 was chosen, meaning that the wavelet packet tree consisted of 128 terminal nodes that represent the range of wavelet coefficients at the lowest scale divison. As the sampling frequency of the vibration was at 97.656Hz, the bandwidth of the spectrum was 48.828Hz and the width of the frequency bands of the terminal nodes was therefore approximately 381Hz. From this wavelet packet tree, a reconstruction of the wavelet packet power spectrum was computed with the same sampling frequency as the original signal.

As suggested by Shao et al. [30], the relative energy of each wavelet frequency band can be used in the feature vector. A summation of the power at each band over a window of $n = 5800$ samples formed the new samples of energy. Each feature vector or sample then consists of the energy of each band which was divided with the total sample energy, giving the relative energy. Energy at each band j:

$$E_j = \sum_{k=1}^{n} p_{k,j} \tag{25}$$

where $p_{k,j}$ is the band power at sample k and band number j.

$$T = [E_1, E_2, E_3, ..., E_{2^L}] \tag{26}$$

where L is the number of wavelet packet decomposition levels. The total energy is:

$$E = \sum_{k=1}^{L} E_k \tag{27}$$

Figure 32 shows the samples of the relative energy of the 128 bands for two different samples. The band number represent the frequency bands from 1 to 128, where 1 is the highest frequency band, and 128 is the lowest.

As concluded by [30], the energy located at different spectrum bands in the vibration signal reveals much information about the equipment condition. When faults appear, some parts of the spectrum will have relatively more energy than others in a way that is characteristic for that fault. This also enables faults to be classified if data for distinct fault conditions are available, as is the case in this study.

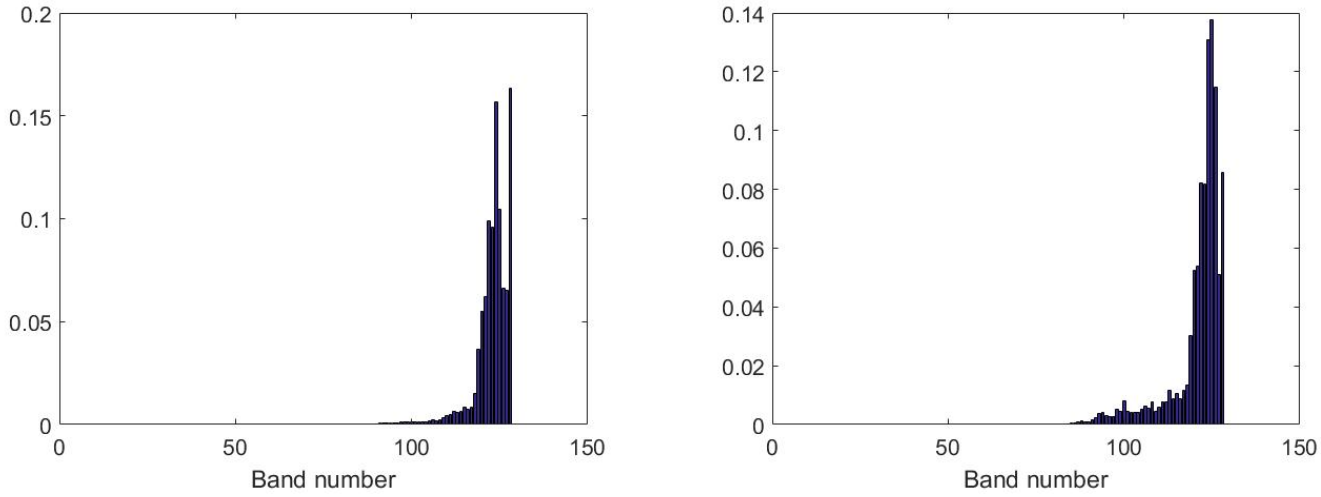

Figure 32: Two illustrative feature samples of relative energy at each frequency band.

8.4.2 *Time domain features*

A variety of time domain features may also be extracted for the purpose of condition monitoring. The time domain feature included in this study are listed below:

- **Absolute mean**. The absolute statistical mean of the time-series signal.
- **Maximum peak**. Maximum absolute peak value of a time-series signal.
- **Crest factor**. The crest factor is the ratio of peak to RMS values of the signal [33].
- **Kurtosis**. Kurtusis is a measure of the signal peakedness, and is defined as the signal's fourth order momentum about the mean [33].
- **RMS**. Root mean square of the signal.
- **Variance**
- **Shape factor**. The shape factor is defined as the ratio of the signal's RMS value to its absolute mean [33].
- **Skewness** The skewness of a signal is a measure of the signal asymmetry, and is defined as the third order moment about the mean [33].

The window size for the feature extraction in time domain was 5800 samples, corresponding to the same time division as the wavelet packed composition features.

8.5 FAULT DETECTION BASED ON IN-CONTROL PCA MODEL

The objectives for this section was to a create data-driven model from the normal bearing condition from the features, for then to investigate the fault detection techniques with testing data from known fault conditions. For most practical applications of these methods, data for the relevant fault conditions are not available or does not have the quality to be used for classification of future fault conditions. Being able to differentiate normal healthy condition from faulty conditions are in these cases important.

For further fault detection and classification, the features are subject to the dimension reduction technique of PCA (Principle Component Analysis) in order to compress the original feature space into lower dimension space containing new uncorrelated features. This method is widely used for feature extraction for condition monitoring [30, 33], and improves the ability for further fault detection and classification. In this case we had 128 features from the wavelet packet decomposition, and 8 time domain features, making a total of 136 features. In the further fault detection and classification, the methods were first tested with only the 128 wavelet packet features, and then with all 136 features.

Six vibration data series was included in this analysis. Two baseline conditions at a length of 6 seconds each, three different outer race fault conditions each with length 6 seconds, and one inner race condition at 3 seconds. The inner race condition was from a situation with a different load than the other conditions. This makes this specific data series less useful for testing for fault detection, as the PCA model for normal condition is expected to change as the load varies.

Each of the baseline and outer race fault conditions have a 100×136 data matrix with 100 samples and 136 features. Inner race condition have a 50×136 data matrix as the sampling period length is only 3 seconds.

8.5.1 *Fault Detection Using Only Wavelet Packet Features*

In this first part, only the 128 wavelet packet features were included in the generation of the in-control model and for the testing data. The in-control PCA model was calibrated with the baseline condition 1 with a 100×128 data matrix using singular value decomposition. Figures 33 shows the explained variance of the PCA in-control model as the number of retained PCs increase. 8 principal components was retained in the in-control model describing about 95% of the variance in the training data.

From the cross-validation we see that we have good validating performance all though the number of calibration samples is relatively low.

Figure 34 shows the correlation loadings of PC-1 by PC-2 for the in-control model. We see that the bands from the lowest frequencies have the most variation and dominates the influence on the model. Most of the bands of higher frequencies are highly clustered and correlated.

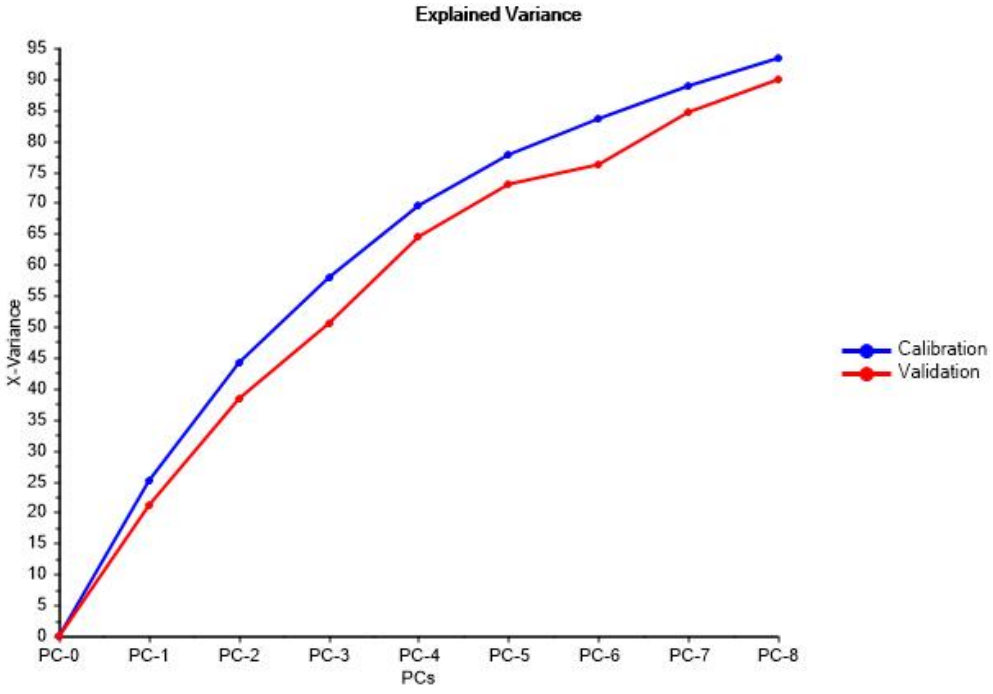

Figure 33: Explained variance of in-control model using only 128 wavelet packet features

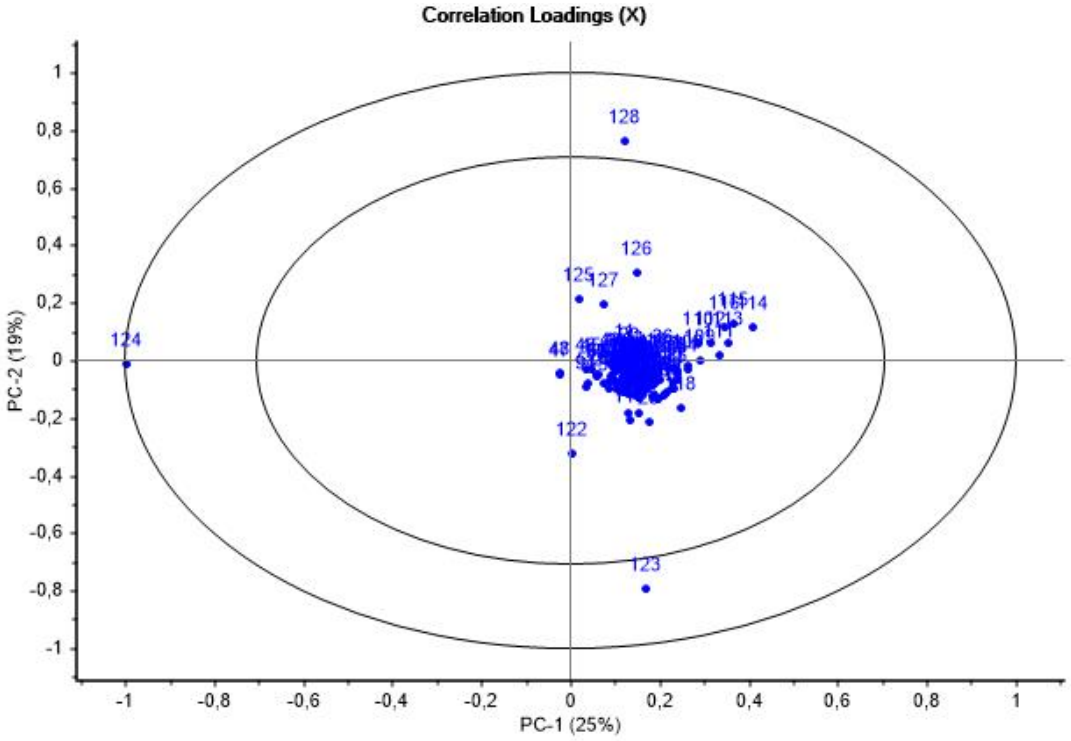

Figure 34: Correlation plot of in-control model using only 128 wavelet packet features

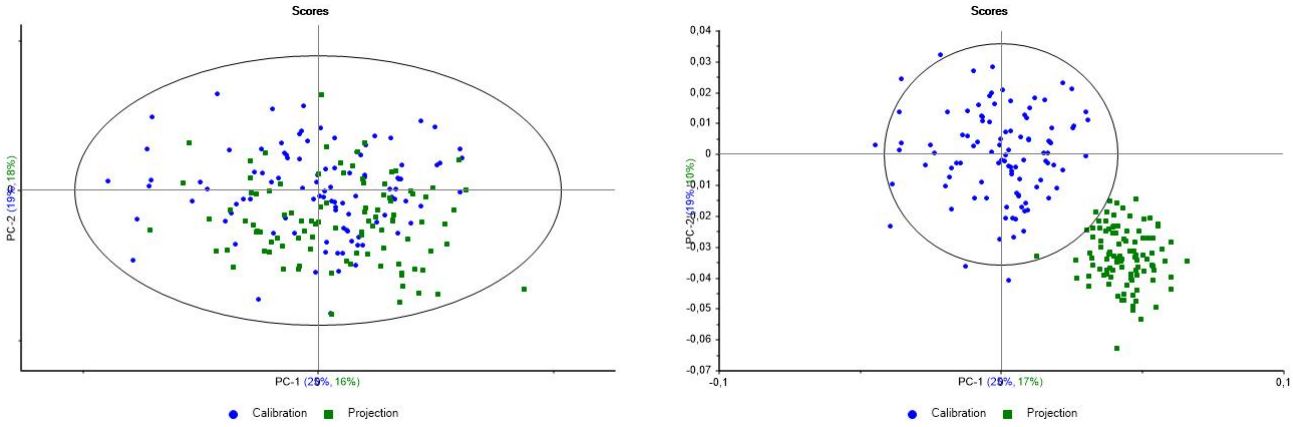

Figure 35: Score plot with projections of test data: Left: Baseline condition 2. Right: Outer race fault condition 2

After the in-control model was created, new testing data was projected on to the model in order to detect if the new samples deviated from normal condition. Figure 35 shows the PC-1 versus PC-2 score plots of the projections of the healthy baseline 2 testing data, and the faulty outer race condition 2 data. We see that the projections of the samples from baseline condition 2 is within the normal operating region of the in-control model, while the projections from the outer race fault condition lies outside this region, clearly detecting the presence of the fault.

Figures 36, 37, 38, 39 and 40 shows the Hotelling's T^2 statistic and the SPE or Q-residual for the projections of the test data from baseline condition 2, outer race fault condition 1-3, and the inner race fault condition, respectively. Each plot is compared to the 1% significance limit for T^2 and SPE. In figure 36 we see that Hotelling's T^2 for the baseline 2 condition is within the significance limit, except a few spikes. This might indicate that the operating conditions for baseline condition 2 was slightly different from baseline condition 1. Including more calibration samples in the in-control model calibration covering a wider range of operating conditions might improve this problem, thus lowering the number of false alarms. The SPE for the baseline 2 projection are well within the significance limit, indicating the the samples fits the model well.

The T^2 statistics in figure 37 for projections of the outer race fault condition 1 show that it is starting to violate the limit, but the deviation is small. For this condition, the SPE is still within the significance limit. Since the severity of this fault condition is the smallest, simulating a fault in an early stage, the small deviation in T^2 and SPE are expected. The fault is however still detected due to the relatively high number of samples breaking the T^2 limit.

For the remaining projections for the more severe outer race fault condition 2 and 3, as well as the inner race condition, we see that all significantly violates both Hotelling's T^2 statistics and SPE. In these cases the faults are clearly detected.

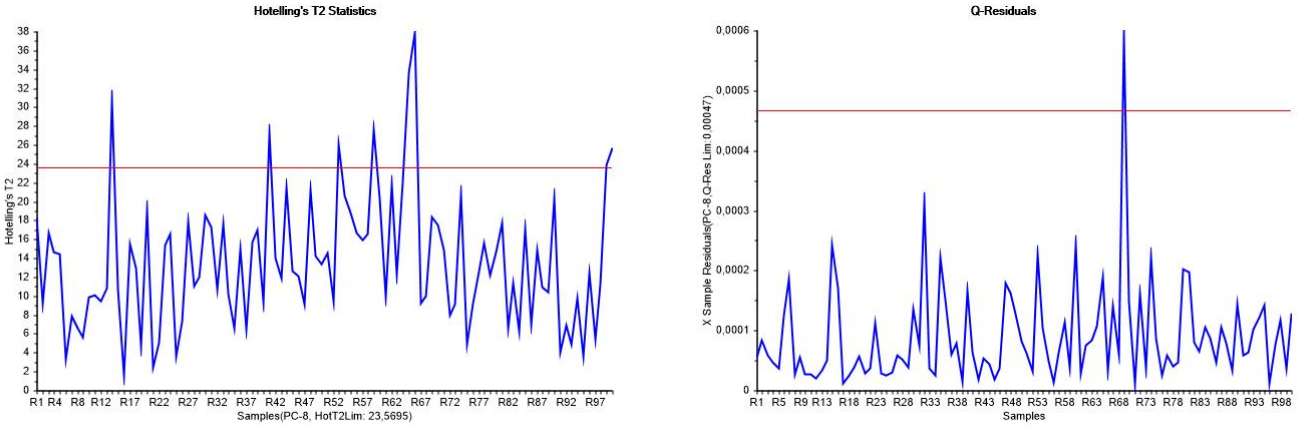

Figure 36: Baseline condition 2: Left: Hotelling's T^2. Right: SPE or Q-residual

.

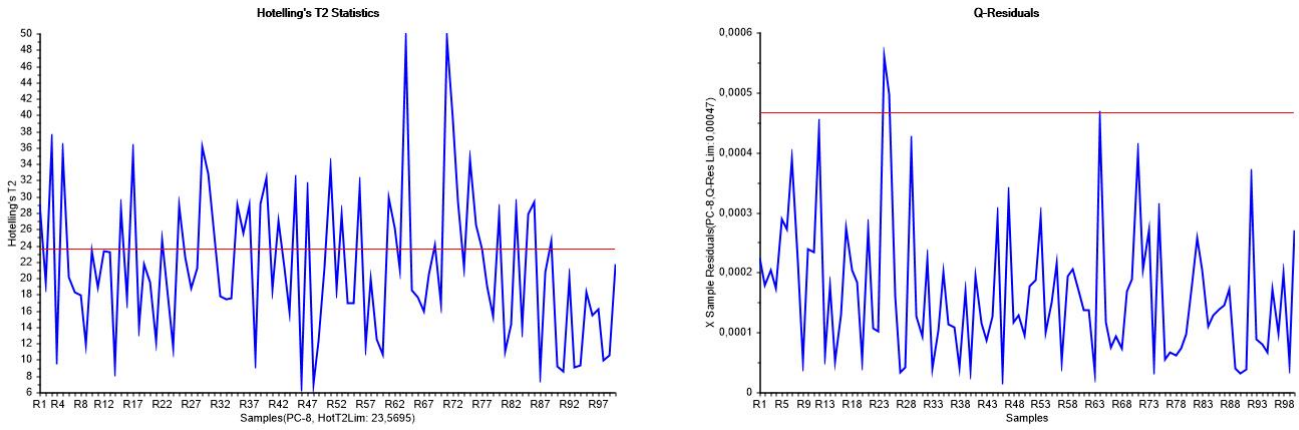

Figure 37: Outer race fault condition 1: Left: Hotelling's T^2. Right: SPE or Q-residual

.

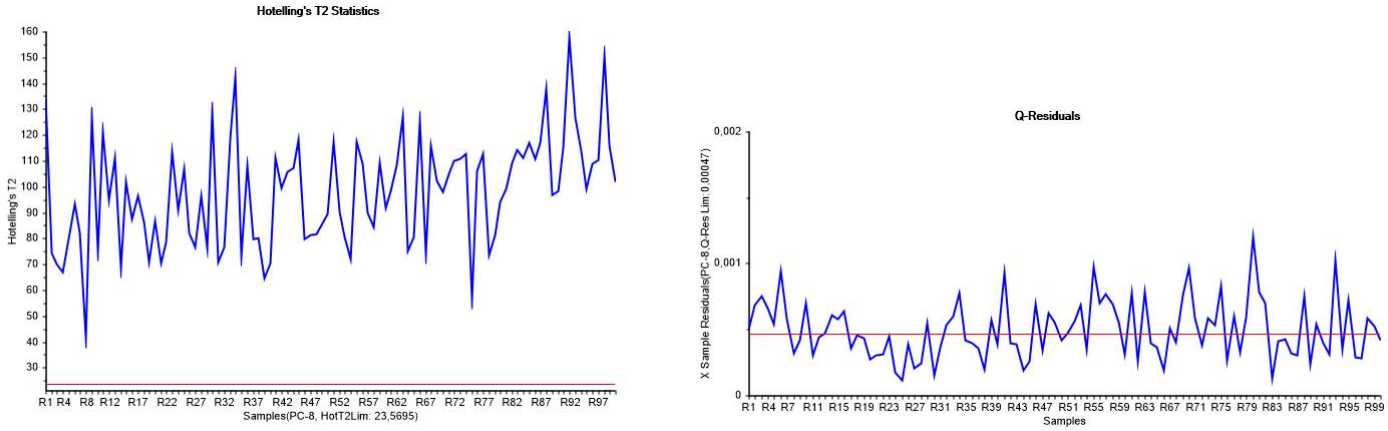

Figure 38: Outer race fault condition 2: Left: Hotelling's T^2. Right: SPE or Q-residual

.

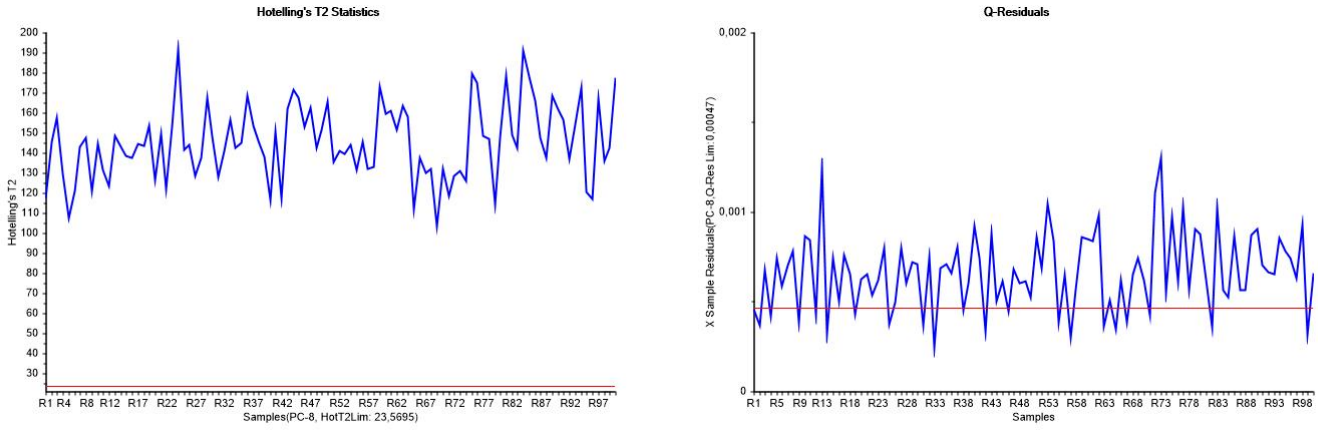

Figure 39: Outer race fault condition 3: Left: Hotelling's T^2. Right: SPE or Q-residual

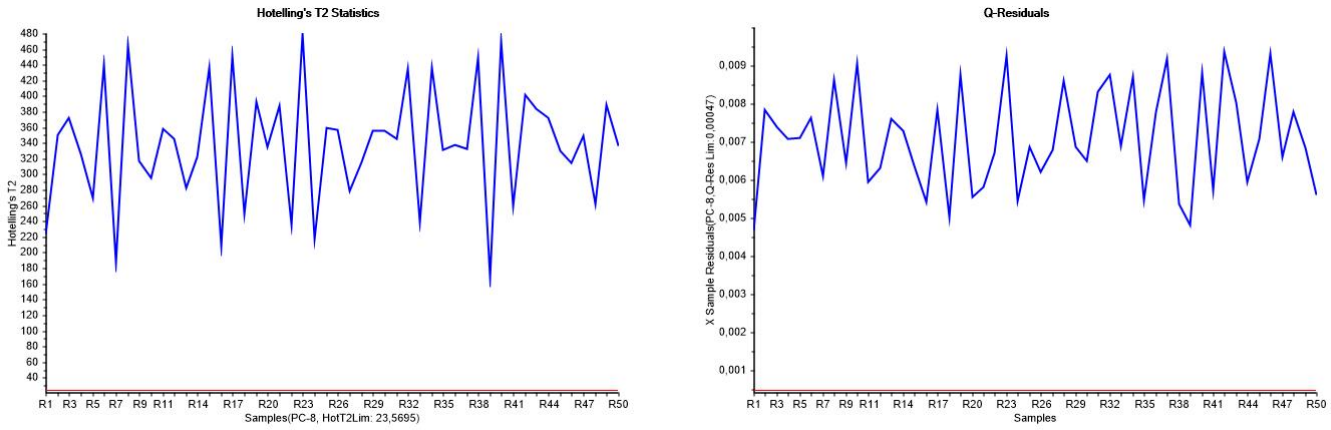

Figure 40: Inner race fault condition : Left: Hotelling's T^2. Right: SPE or Q-residual

.

8.5.2 *Fault Detection Using Both Time Domain and Wavelet Packet Features*

In this second part of the fault detection analysis, both the wavelet packet features and time domain features are included in the analysis. The data for all condition are the same as in the previous part, except that the time domain features are added to each data matrix.

As the time domain features are of varying magnitude and units, they are normalized in order to have equal contribution in the calibration of the in-control model. In order for the time domain features to have comparable contribution in the calibration as the wavelet packet features, the time domain features are down-scaled by a factor of 0.01.

Figure 41 shows the explained variance for the in-control model calibrated on the 100×136 baseline 1 training data. We see that the time features contribute to new types of variation, and in this case 11 principal components need to be retained in order to explain 95% of variance in the data. 11 PCs was retained in the in-control model, showing good performance by cross-validation.

The correlation loadings plot in figure 42 reveals that the time domain feature strongly contributes in the model. We also see that the variance, RMS and absolute mean are strongly correlated.

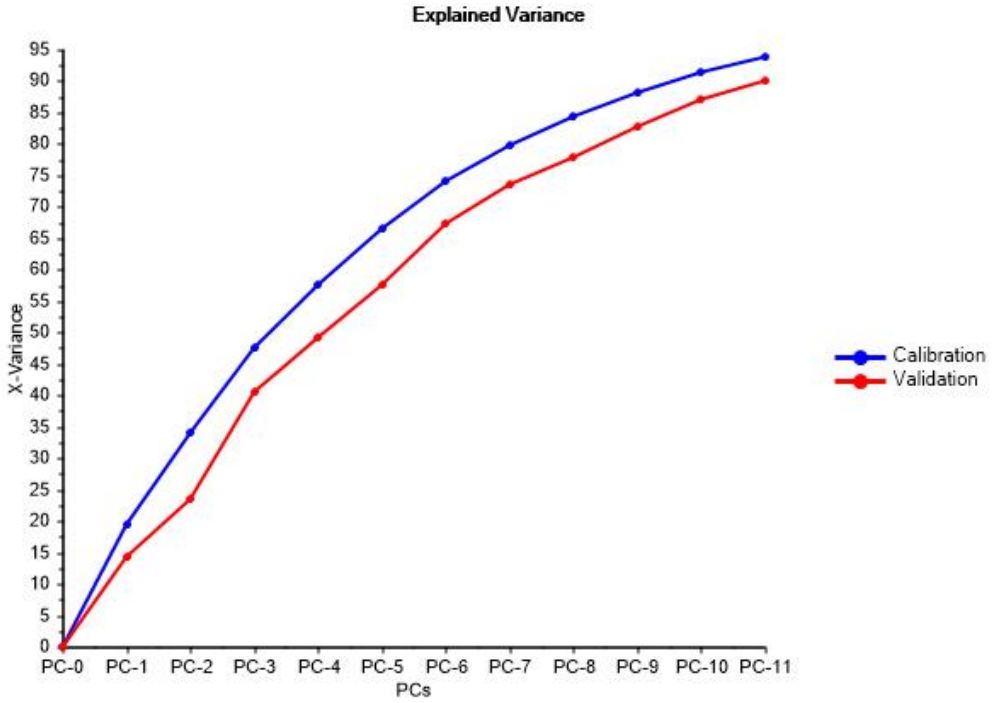

Figure 41: Explained variance of in-control model using all 136 features

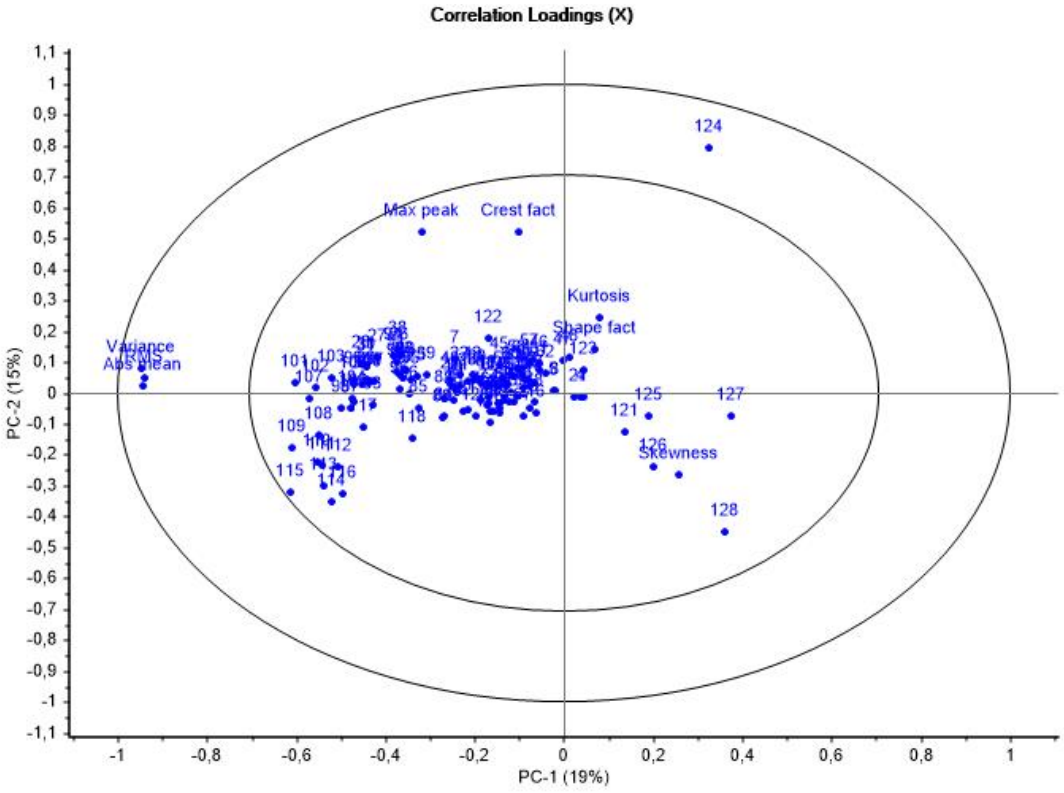

Figure 42: Correlation plot of in-control model using all 136 features

In the same manner as in the previous section, new testing data was projected on to the in-control model. Figure 43 shows

the score plot of projections for baseline condition 2 and outer race condition 2. We see that the samples from the baseline condition 2 are well within the normal operating region, while the outer race fault condition is clearly deviating. Adding the time domain features seem to improve the separability of this fault condition.

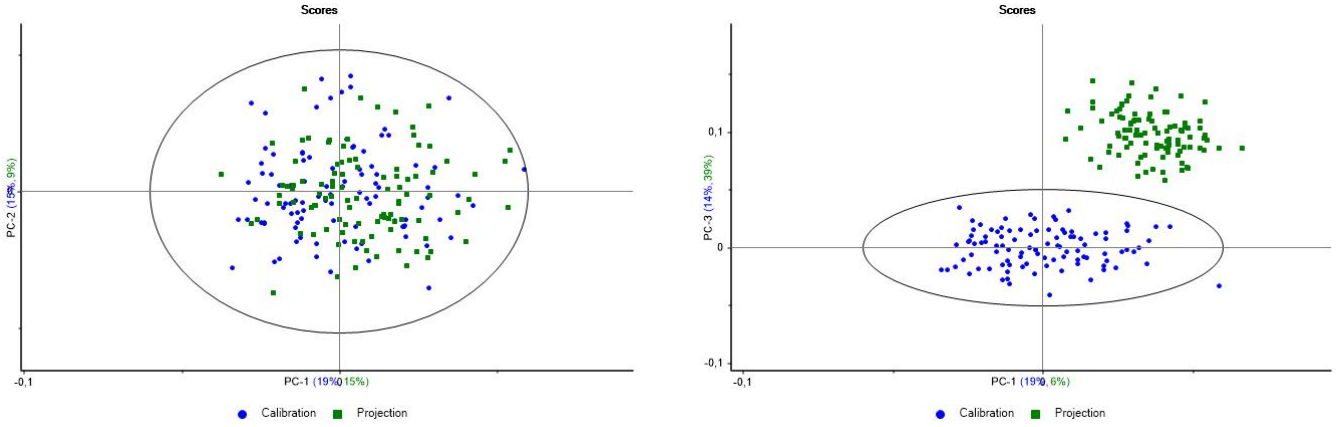

Figure 43: Score plot with projections of test data including all features: Left: Baseline condition 2. Right: Outer race fault condition 2

.

Hotelling's T^2 and SPE plots for all projections are shown in figures 44, 45, 46, 47 and 48. Similarly as was the case for the model with only wavelet packet feature, the projection of the baseline condition 2 samples are mainly within the limits for T^2, but some samples have larger deviations.

All of the projections for fault conditions are successfully detected as faulty by the Hotelling's T^2 statstics and SPE. We also see that adding the time domain features seem to improve the detection ability compared to using only wavelet packet features.

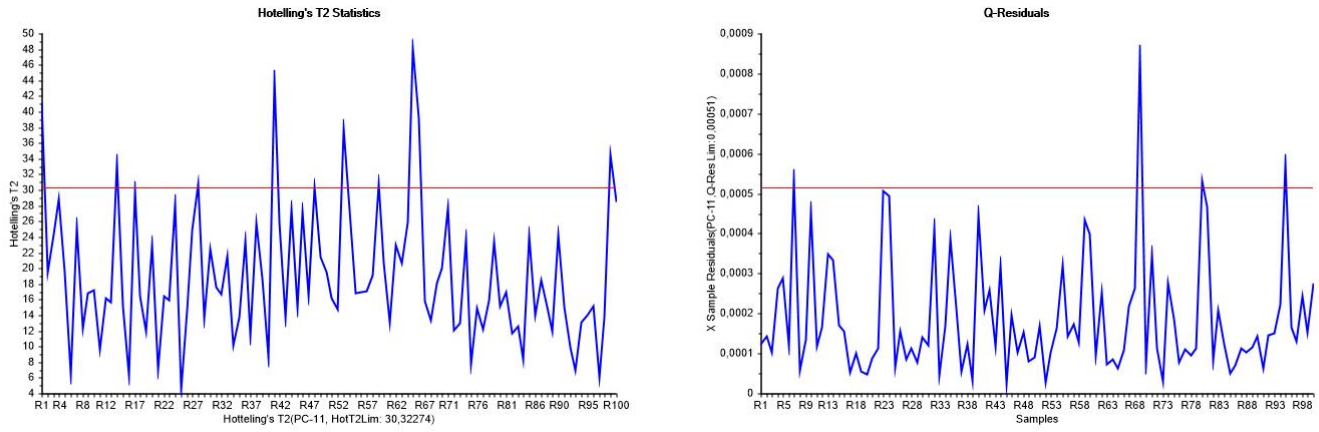

Figure 44: Baseline condition 2, all features: Left: Hotelling's T^2. Right: SPE or Q-residual

.

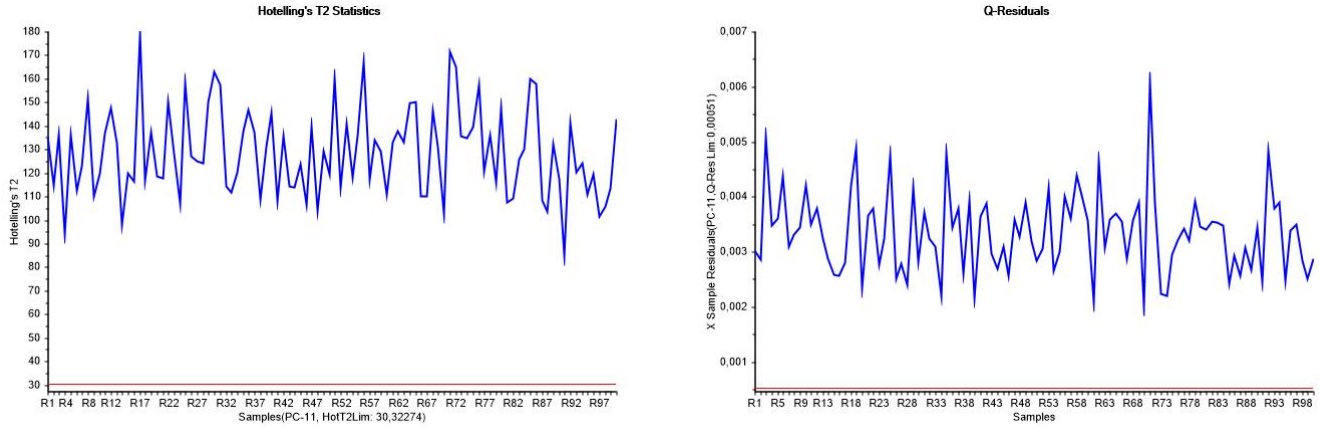

Figure 45: Outer race fault condition 1: Left: Hotelling's T^2. Right: SPE or Q-residual

.

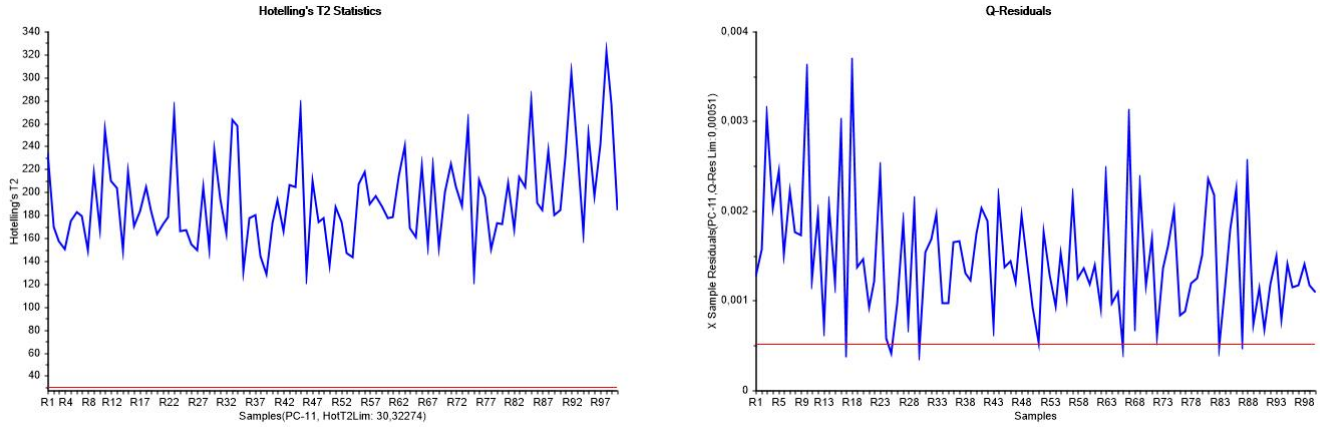

Figure 46: Outer race fault condition 2: Left: Hotelling's T^2. Right: SPE or Q-residual

.

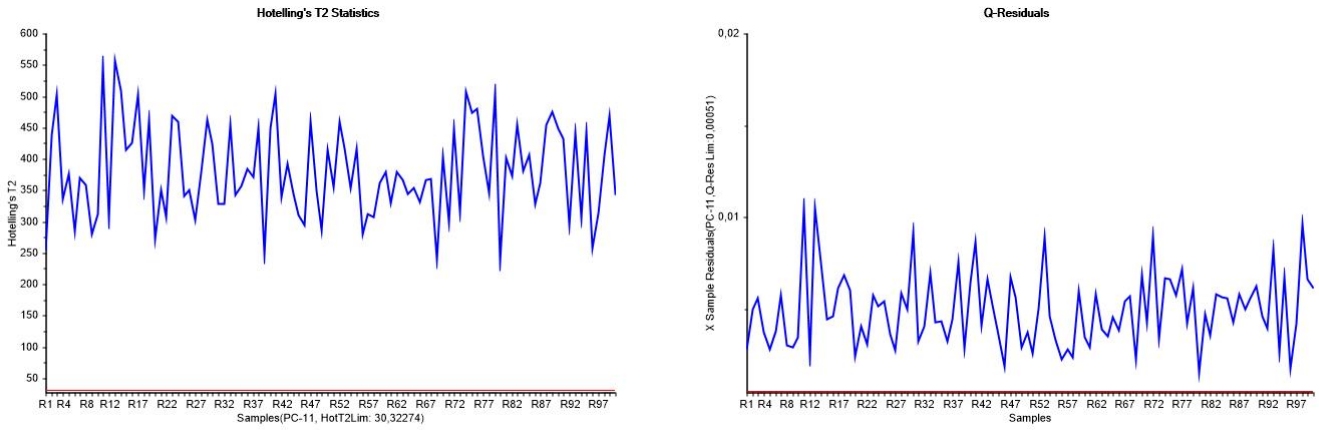

Figure 47: Outer race fault condition 3: Left: Hotelling's T^2. Right: SPE or Q-residual

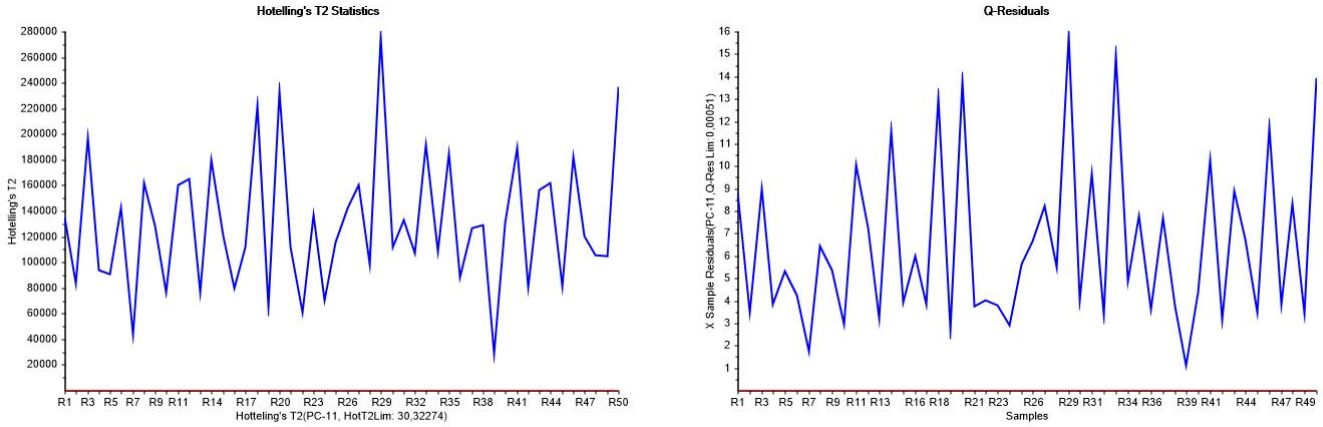

Figure 48: Inner race fault condition : Left: Hotelling's T^2. Right: SPE or Q-residual

.

8.6 SIMCA BASED PATTERN RECOGNITION AND CLASSIFICATION

When data from distinct fault condition are available, this can be used for supervised training for classification and diagnosis. SIMCA analysis is performed as a classification method using supervised training to diagnose faults from known fault conditions. The method involves generating PCA models for each condition, both healthy condition and the distinct fault conditions. In this case the different models included are healthy baseline condition 1 model, model for outer race fault condition 2, and a model for inner race fault.

For the calibration of the SIMCA class models, each of the data matrices for the three conditions were split into a training part and a testing part. The baseline condition and outer race

fault condition both had training matrices of size 50×128 for only WPD features and 50×136 for the case with all features included. The inner race fault class was calibrated with only the first 25 samples.

Table 2 shows the number of retained PCs for the class models using only wavelet packet (WPD) features. Figure 3 shows the explained variance and number of PCs for the SIMCA PCA models including all 136 features. The choice of model complexity was based on cross-validation and explained variance.

SIMCA Model	Number of PCs	Explained variance cal.	Explained variance val.	Number of cal. samples
Baseline 1	8	95%	87%	50
Outer race fault 2	9	95%	91%	50
Inner race fault	11	94%	76%	25

Table 2: SIMCA models, only 128 WPD features: model complexity and explained variance

SIMCA Model	Number of PCs	Explained variance cal.	Explained variance val.	Number of cal. samples
Baseline 1	10	94%	87%	50
Outer race fault 2	12	96%	91%	50
Inner race fault	4	93%	89%	25

Table 3: SIMCA models, with all 136 features: model complexity and explained variance

After the the SIMCA class models was trained with their respective calibration set in a supervised manner, new testing samples where projected on the the models for classification. The remaining 50 samples of the data sets for the baseline condition 1 and outer race condition 2 which was not included in the calibration are used as testing data. For the inner race fault condition the remaining 25 samples where classified. In addition, 100 samples of baseline condition 2 was also included in the classification.

The confusion table 4 shows the results of the the SIMCA classification for the case with only 128 wavelet packet features. Table 5 shows the same results for the case with all 136 features. The classification was based on the sample leverage, meaning the distance from the sample projection to the class mean, and the sample-to-model distance which measures the distance from a sample to a model. The result are based on a F-test with 5% confidence limit. The models including only wavelet packet features show an almost perfect classification of the test samples. The projections of the samples for baseline 2 condition, classified only 89 of the 100 samples to the baseline 1 class. The classification results for the models including all 136 features show a few type II errors for outer race 2 and inner

race test samples, but 94/100 test samples of baseline condition 2 is here classified to the the baseline 1 class.

		Predicted class membership			
		Baseline 1	Outer race 2	Inner race	Tot. no. samples
Actual Class	Baseline 1	50	0	0	50
	Outer race 2	0	50	0	50
	Inner race	0	0	25	25
	Baseline 2	89	0	0	100

Table 4: Confusion table for SIMCA classification with only 128 WPD features

		Predicted class membership			
		Baseline 1	Outer race 2	Inner race	Tot. no. samples
Actual Class	Baseline 1	50	0	0	50
	Outer race 2	0	48	0	50
	Inner race	0	0	22	25
	Baseline 2	94	0	0	100

Table 5: Confusion table for SIMCA classification with all 136 features

The results from the classification shows a good performance. The model-to-model distances shown in table 6 and 7 also shows that strong separation between the model, as all distances are much larger than 3. The classification methods should however be calibrated with larger calibration sets, and be tested with a larger number of representative samples. It is also uncertain if future faults will have a covariation structure that is captured in the trained fault models.

	Model-to-model distance		
	Baseline 1	Outer race fault 2	Inner race fault
Baseline 1	1	18.19	54.27
Outer race fault 2	–	1	99.24
Inner race fault	–	–	1

Table 6: WPD features: distances between models

	Model-to-model distance		
	Baseline 1	Outer race fault 2	Inner race fault
Baseline 1	1	22.20	147.02
Outer race fault 2	–	1	97.86
Inner race fault	–	–	1

Table 7: All features: distances between models

8.6.1 *Fault Isolation and Spectrum Analysis*

In most situations, data from known fault conditions are not available to be used in supervised learning based classification. Diagnosis of the fault condition can in these cases be based on known fault characteristics such as passing frequencies in ball bearings.

Different passing frequencies based on the geometry of the specific roller or ball bearing can be determined [9]:

Cage Pass Frequency:

$$CPF = \frac{f}{2}\left(1 - \frac{d}{e} \times \cos(\beta)\right) \tag{28}$$

Ball Pass Frequency Inner Race:

$$BPFI = \frac{N_e \times f}{2}\left(1 + \frac{d_B}{d_p}\cos(\alpha)\right) \tag{29}$$

Ball Pass Frequency Outer Race:

$$BPFO = \frac{N_e \times f}{2}\left(1 - \frac{d_b}{d_p}\cos(\alpha)\right) \tag{30}$$

Ball Fault Frequency:

$$BFF = \frac{d_p \times f}{d_b}\left(1 - \left(\frac{d_b}{d_p}\right)^2 \times \cos^2(\alpha)\right) \tag{31}$$

Where f is the driving frequency.

The test rig was equipped with a NICE bearing with the following parameters[3]:

- Roller diameter: $d_b = 0.235$
- Pitch diameter: $d_p = 1.245$
- Number of elements: $N_e = 8$
- Contact angle: $\alpha = 0$

The corresponding passing frequencies for the outer race fault in this study can be computed as BPFO $\approx$ 81Hz.

Zooming in on the time series vibration signal reveals the pulses at equal intervals caused by the discontinuities in the bearing raceways. Denoising of the signal by wavelet thresholding makes these pulsating characteristics more clear as seen in figure 49 and 50. Approximating the rate of the pulses in figure 50 gives a frequency at about 81Hz, which is corresponding to the passing frequency of a outer race fault of this bearing.

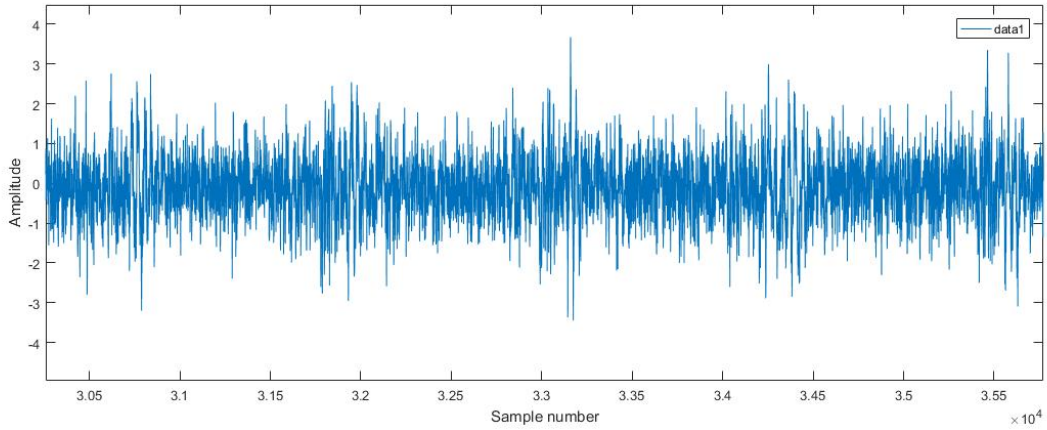

Figure 49: Outer race condition 2. Raw time-series signal.

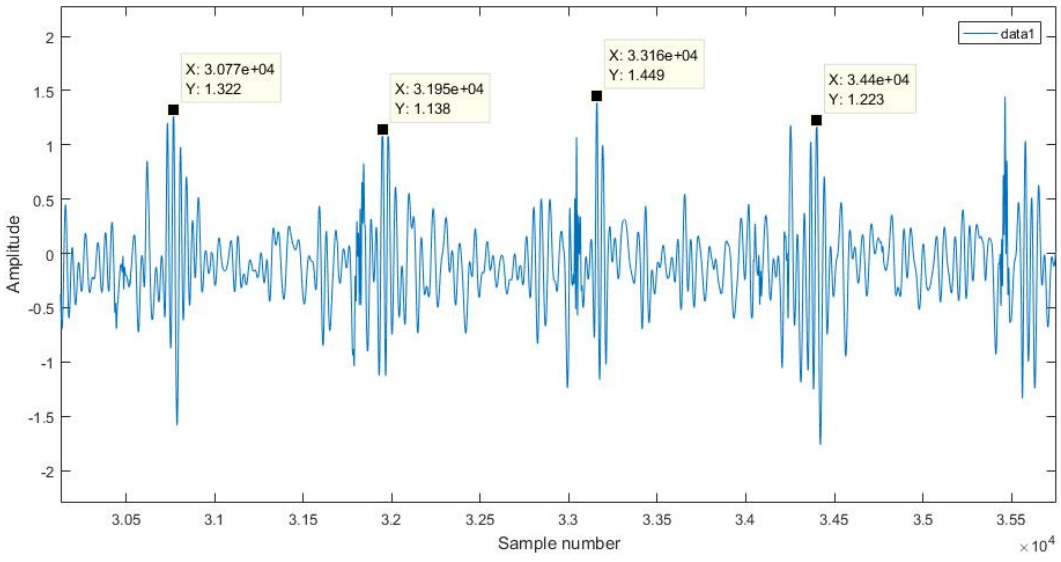

Figure 50: Outer race condition 2. Denoised time-series signal.

9

AIR COMPRESSOR ACOUSTIC SIGNAL ANALYSIS

9.1 INTRODUCTION

This case study concerns data-driven condition monitoring techniques for a reciprocating air compressor using acoustic signal analysis. Based on feature extraction from wavelet packet transforms, Principal Component Analysis was further utilized as a dimension reduction technique for fault detection and fault classification.

Some of the benefits of using acoustic signals for condition monitoring purposes is that they have a relatively low capital and installation cost. In contrast to vibration monitoring, acoustic measurements from microphones can be taken sufficiently far away from the machine to avoid conditions such as extreme temperatures [31]. Acoustic signals have shown to be effective for fault detection and fault identification of machinery both in time domain and frequency domain.

Verma et al. [33] performed a study on acoustic signal based condition monitoring of this air compressor setup, including data acquisition, sensor positioning analysis, data preprocessing, feature extraction, feature selection and classification. Verma et al. used the same data in their study, investigating a range of different feature extraction techniques, including Fast Fourier Transform, Wavelet Packet Transform, Discrete Cosine Transform, Short Time Fourier Transform and Wigner-Ville Distribution. The classification was based on one-against-one multiclass Support Vector Machine with radial basis function kernel, achieving good performance [33].

This case study uses the same data set in order to study the use of wavelet packet decomposition, PCA based fault detection and SIMCA fault classification.

9.2 DATA ACQUISITION

Specification for the reciprocating air compressor setup that was monitored is given below [33]:

- Air Pressure Range: 0-500 lb/m2, 0-35 Kg/cm2
- Induction Motor: 5HP, 415V, 5Am, 50 Hz, 1440rpm
- Pressure Switch: Type PR-15, Range 100-213 PSI

The reader is referred to Verma et al. [33] for more details regarding the air compressor test system setup, sensor location analysis and data acquisition.

The data set comprised of readings from 8 states which includes healthy state and 7 faulty states, namely [33]:

- **Healthy**. In the healthy state the air compressor is free from all faults.
- **Bearing fault**. Faulty state due to cracks in the bearing.
- **Flywheel fault**. Faulty state caused by wear in the flywheel.
- **Leakage Inlet Valve (LIV) fault**. Faulty state with leakage through the inlet valve due to damage.
- **Leakage Outlet Valve (LOV) fault**. Damaged outlet valve causing leakage.
- **Non-Return Valve (NRV) fault**. Damaged non-return valve causing leakage from the air tank, and hence increased load on the compressor.
- **Piston ring fault**. Loose piston ring in piston head causing air leakage.
- **Rider belt fault**. Faulty state caused by bad alignment of the rider belt with the pulley.

The data used in this thesis was collected from a freely distributed online repository provided by Indian Institute of Technology Kanpur, Verma et al [5], [33]. The data provided in their online repository was only available in the form of pre-processed acoustic signals. Only these conditioned signals are used in this thesis. Any other data from the online repository is not used.

The pre-processing performed by Verma et al. is descibed in detail in [33]. Below is a short summary the steps:

- **High and low pass filter** High pass filtering with cut-off frequency of 400Hz was performed in order to remove noise from external cooling fans. Additionally, high freqency noise was removed with a low pass filter with cut-off frequency 12kHz [33].
- **Clipping** The signal was cleaned by clipping up the signal, and removing the parts with large standard deviation.
- **Moving average smoothing** The effect of outliers was reduced by performing moving average smoothing.

- **Normalization** The signal was also normalized in order to scale the data within specified limits. See [33] for more details about these steps.

9.3 FEATURE EXTRACTION USING WAVELET PACKET TRANSFORM

A total of 225 recordings, each with a length of one second, were collected for all compressor states giving a total of 1800 recordings. The acoustic recordings were taken when the air compressor had a pressure range of 10 to 150 PSI [33].

To reduce the effect of the non-stationary operation of the air compressor, caused by the increasing load and pressure build-up in the air tank, only parts of these recordings were used in this study for time periods with near stationary conditions. 20 recordings, corresponding to a length 20 seconds, was collected for each state, where 10 recordings were used for training and the remaining 10 used for testing.

The feature extraction procedure was widely based on the same principle as for the bearing vibration analysis in chapter 8, except that time domain features were here not extracted. Wavelet packet decomposition was utilized for the acoustic signal in order to achieve a time-frequency decomposition.

Choosing the family and type of wavelet function is an important part of the wavelet analysis. Wavelet function families such as Haar, Coiflets and Daubechies have proven successful in acoustic analysis, and the Daubechies of order 4 (db4) is chosen for this case study as of its finite support size and low vanishing moment properties [26, 33]. The signal was decomposed in seven levels, meaning that the signal was divided in 128 frequency bands at the terminal nodes of the wavelet packet tree. The sampling rate for the acoustic signal was 50kHz, hence the frequency spectrum ranged from $0 - 25$kHz, and the width of each band was approximately 195Hz. The wavelet packet power spectrum was reconstructed from the terminal nodes with the original sampling frequency and 128 frequency bands.

The same procedure for extraction of spectral and temporal energy features as for the bearing analysis in chapter 8 was performed on the recordings. The wavelet packet power spectrum for each state was used to calculate the energy features based on equations 25, 26 and 27 from chapter 8. Each 10 second recording had 500.000 samples, and a window $n = 5000$ was chosen for the energy features at each of the 128 frequency bands. The resulting features comprised of a 100×128 training data matrix and a 100×128 testing data matrix for each of the 8 states. Figures 51, 52, 53 and 54 show the spectral energy

feature plots for all 8 air compressor states for 100 samples, corresponding to 10 seconds.

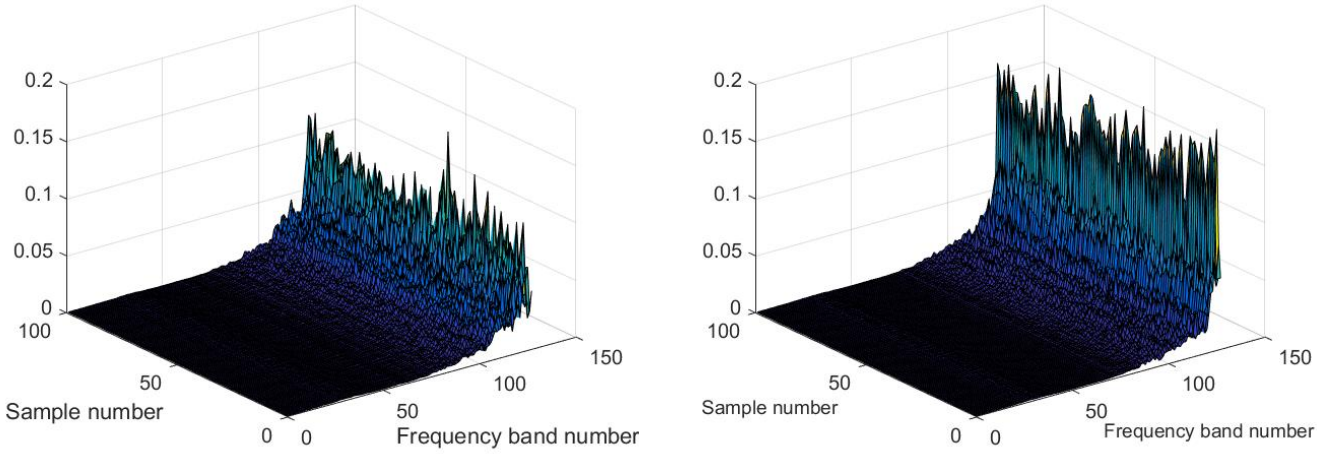

Figure 51: Wavelet packet energy spectrum: Left: Healthy. Right: Bearing fault

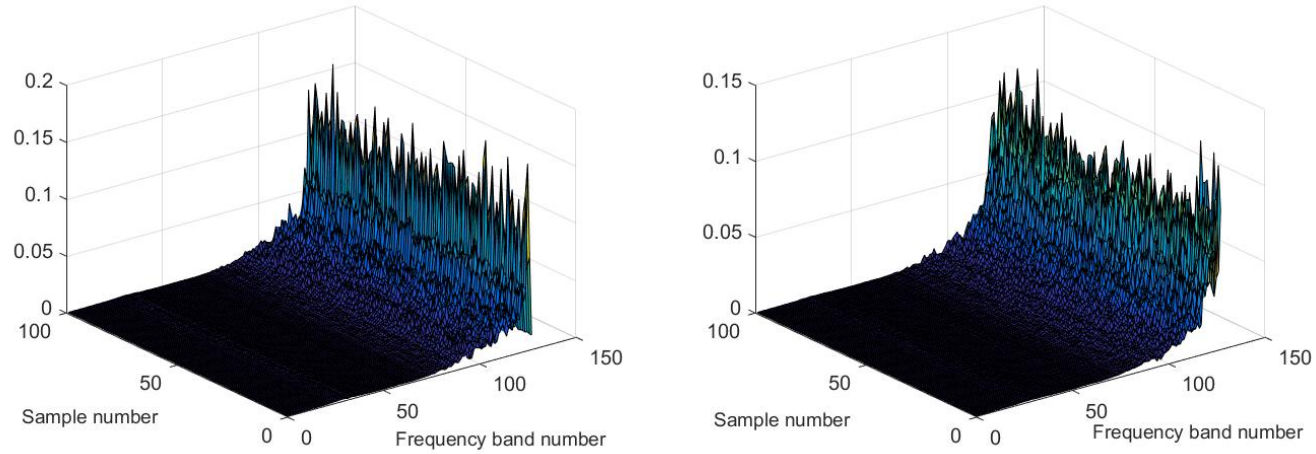

Figure 52: Wavelet packet energy spectrum: Left: Flywheel fault. Right: LIV fault

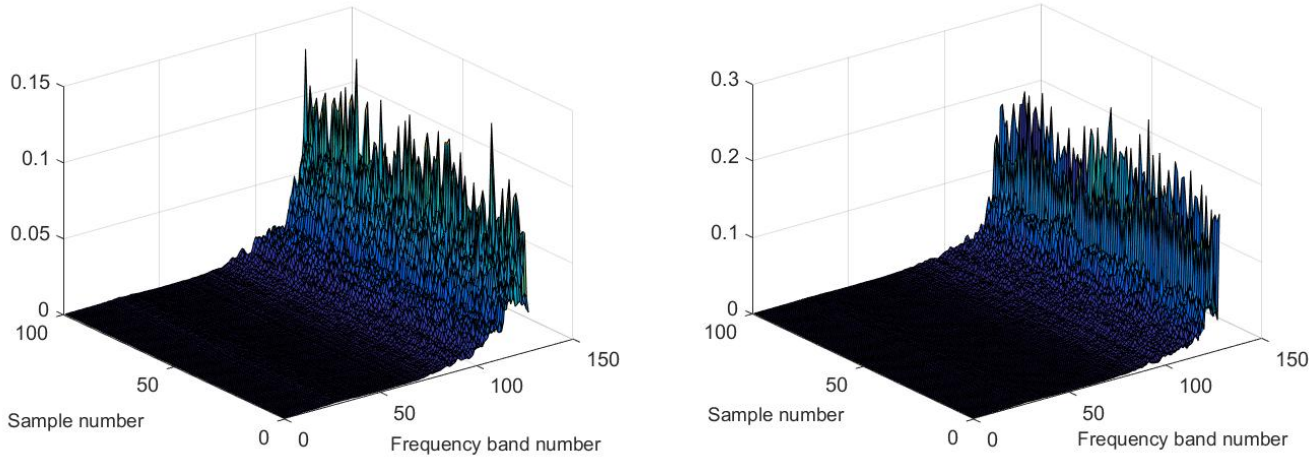

Figure 53: Wavelet packet energy spectrum: Left: LOV fault. Right: NRV fault

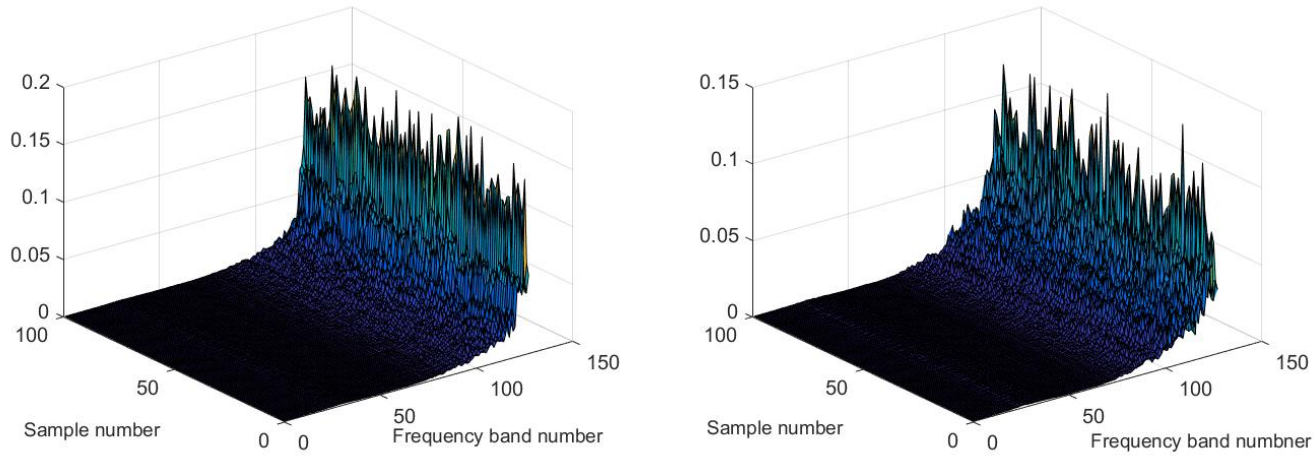

Figure 54: Wavelet packet energy spectrum: Left: Piston fault. Right: Rider belt fault

9.4 FAULT DETECTION BASED ON PCA IN-CONTROL MODEL

One of the most important aspects of condition monitoring systems is the ability to separate faulty conditions from normal conditions. To test this ability, PCA was utilized in order to reduce the dimension of the feature space and to generate an in-control model calibrated with data from the healthy state.

The in-control model was calibrated with the training data matrix for the healthy state with the size 100×128 using the SVD algorithm. From the explained variance plot in figure 55 we see that the first two components describe almost 50% of the variance in the data. This is expected since the most of the variation in the spectrum is located at the lowest frequency bands,

leaving the higher frequency bands with little contribution to the model. Based on cross-validation, 16 principal component was retained, explaining about 95% of the variance in the data.

The loadings plot in figure 56 shows each of the frequency bands, numbered from 1-128 where 128 is the lowest frequency, and how they influence PC1 and PC2. We see that the bands from the lower frequency range dominate the contribution to the model. Most of the bands in the high frequency range are clustered around zero, meaning that they have little influence.

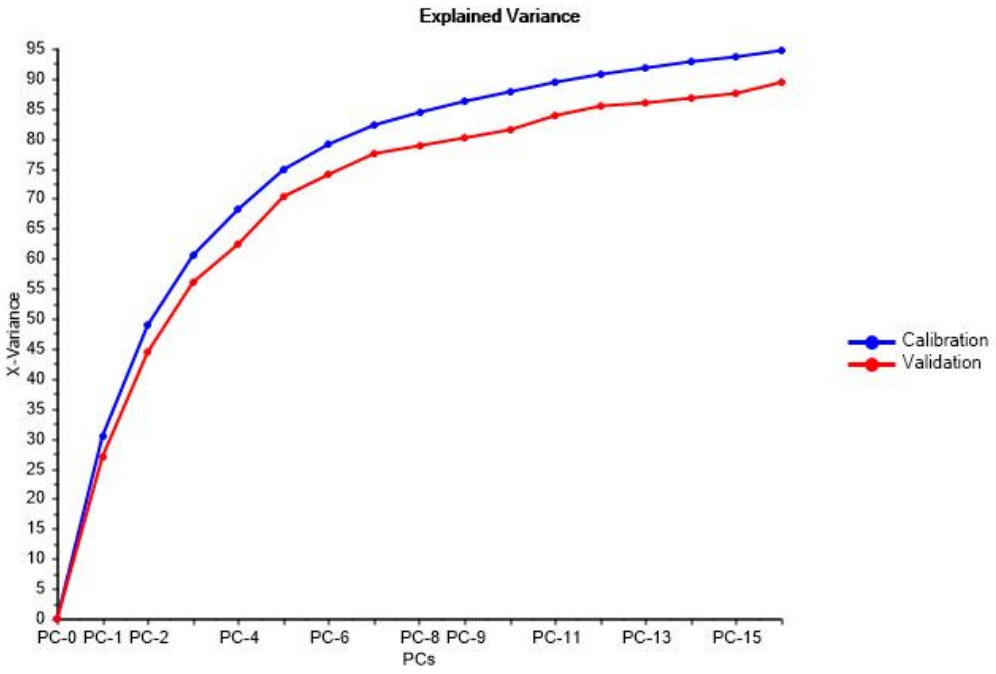

Figure 55: Explained variance of healthy calibration or in-control model

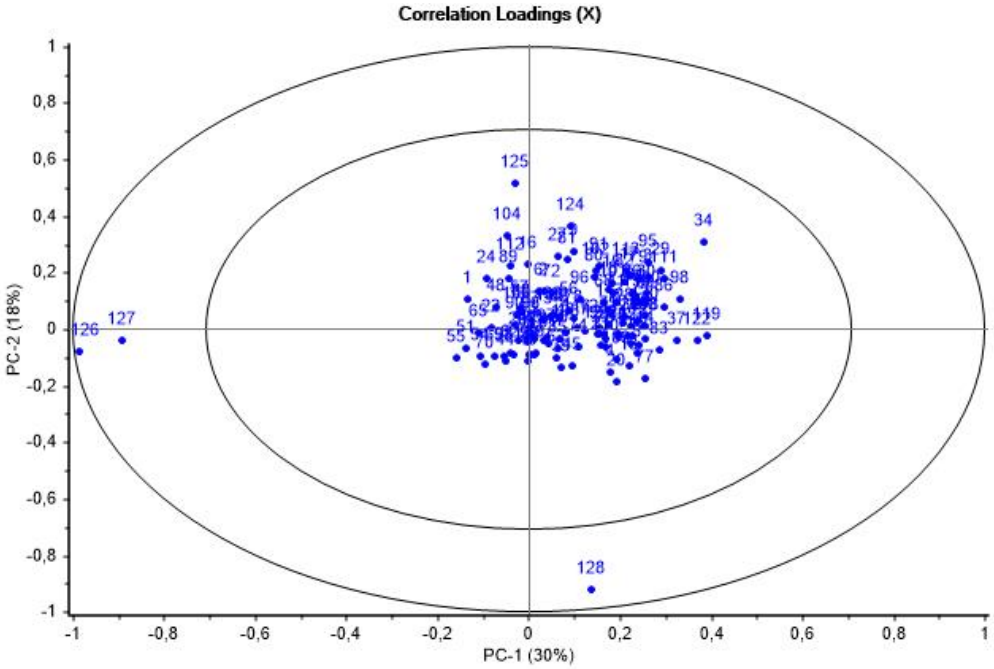

Figure 56: Correlation loadings plot i healthy in-control model

The in-control model was then projected with new testing data from all 8 states in order to investigate the ability to detect faults. The charts for Hotelling's T^2 and SPE, or Q-residual, was used for fault detection. Figures 57, 58, 59, 60, 61, 62, 63 and 64 shows the charts for T^2 and SPE for projections of testing data for each state of the air compressor with confidence limits of 1%. For the projection of healthy testing data we see that T^2 is within its limits, but the SPE chart shows some spikes that violates its confidence limit. This may be due to the non-stationary nature of the data or disturbances in the test data period. It is however clear that this projected features are form a healthy state.

All of the projections for the faulty states violate the limits for SPE, and most of the states also breaks the limits of T^2. The methods are therefore successful in detecting faulty conditions, and separating the faulty features from healthy ones.

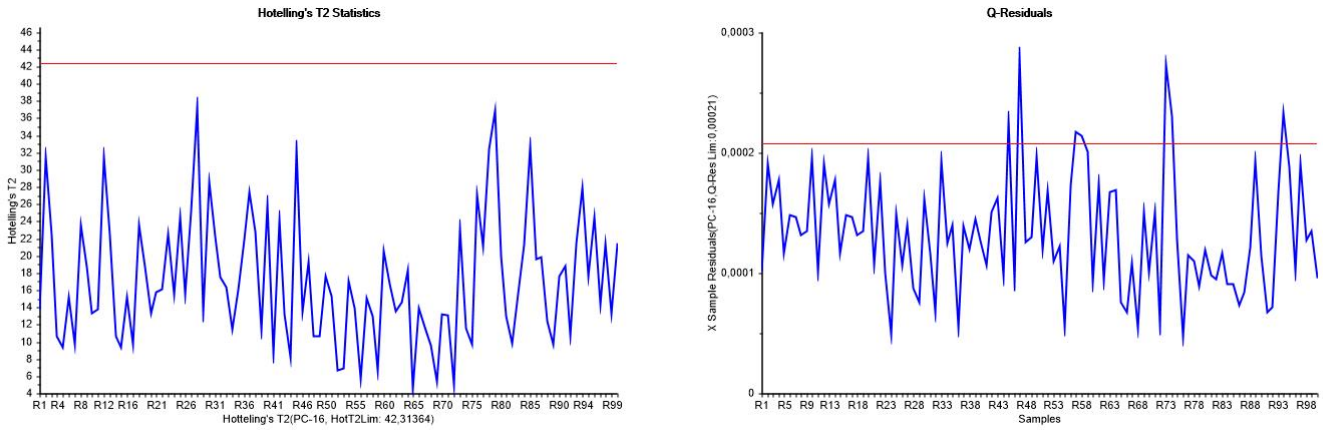

Figure 57: Projection of healthy testing data: Left: Hotelling's T^2. Right: SPE or Q-residual

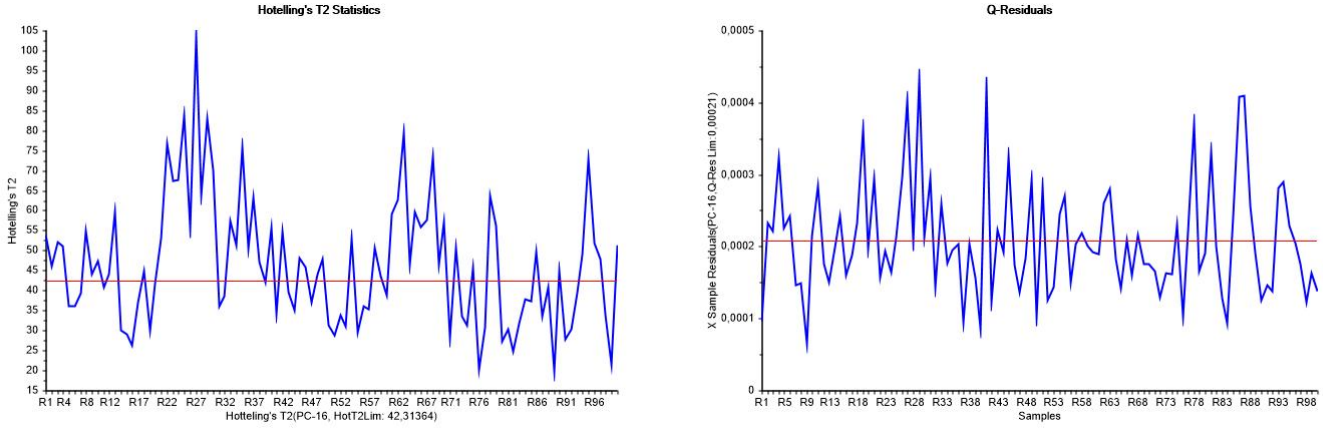

Figure 58: Projection of bearing fault testing data: Left: Hotelling's T^2. Right: SPE or Q-residualg

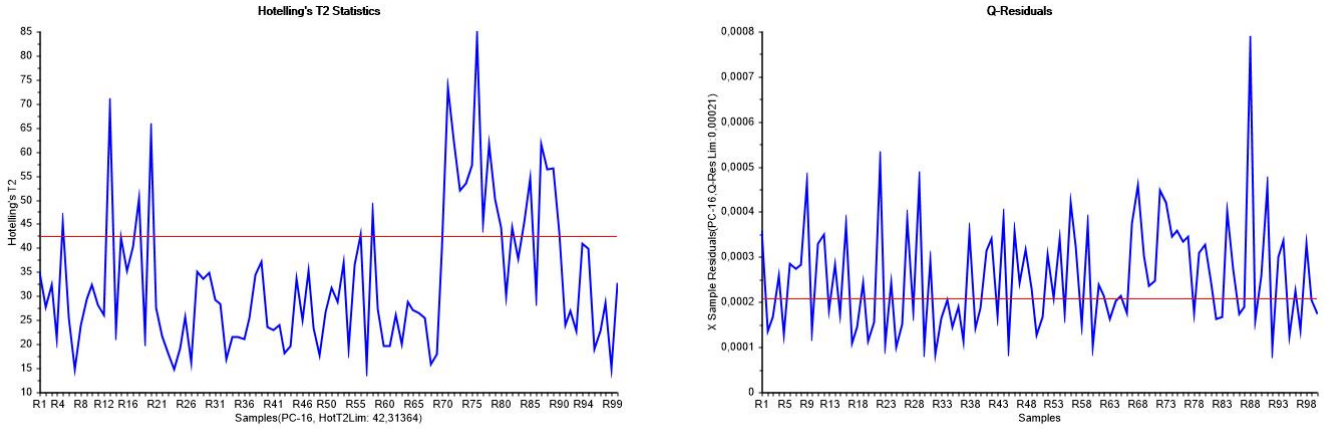

Figure 59: Projection of flywheel fault testing data: Left: Hotelling's T^2. Right: SPE or Q-residual

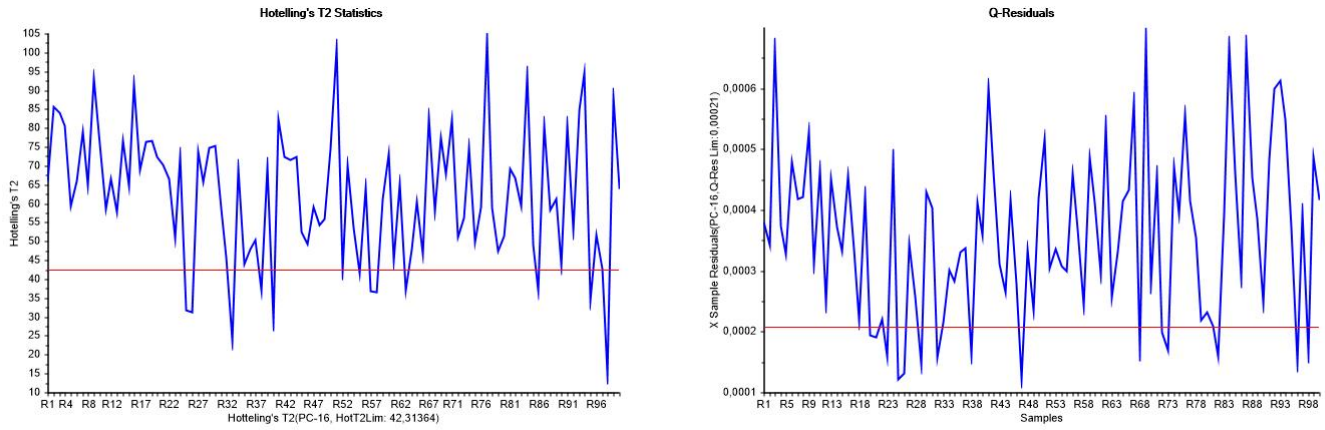

Figure 60: Projection of LIV fault testing data: Left: Hotelling's T^2. Right: SPE or Q-residual

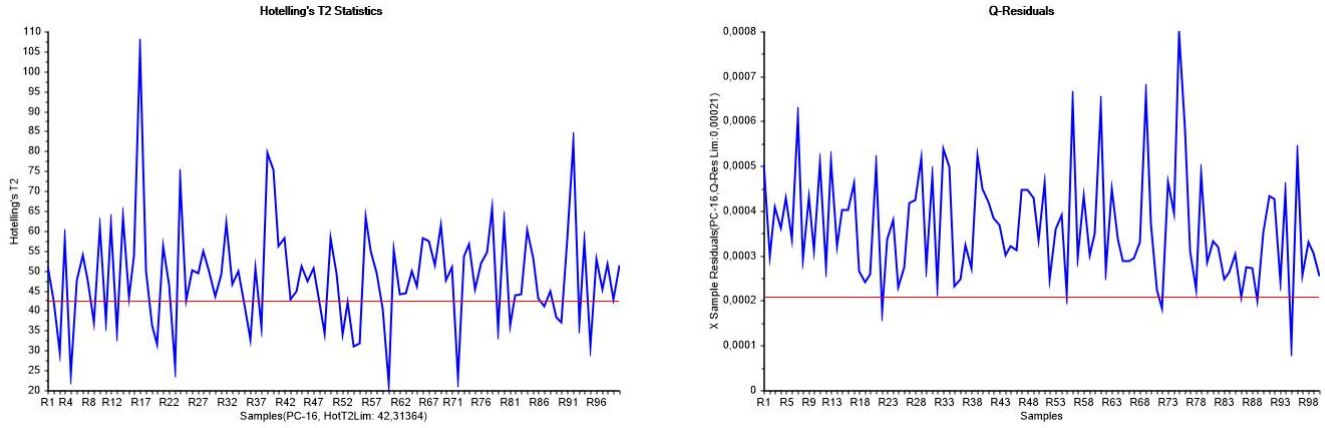

Figure 61: Projection of LOV fault testing data: Left: Hotelling's T^2. Right: SPE or Q-residual

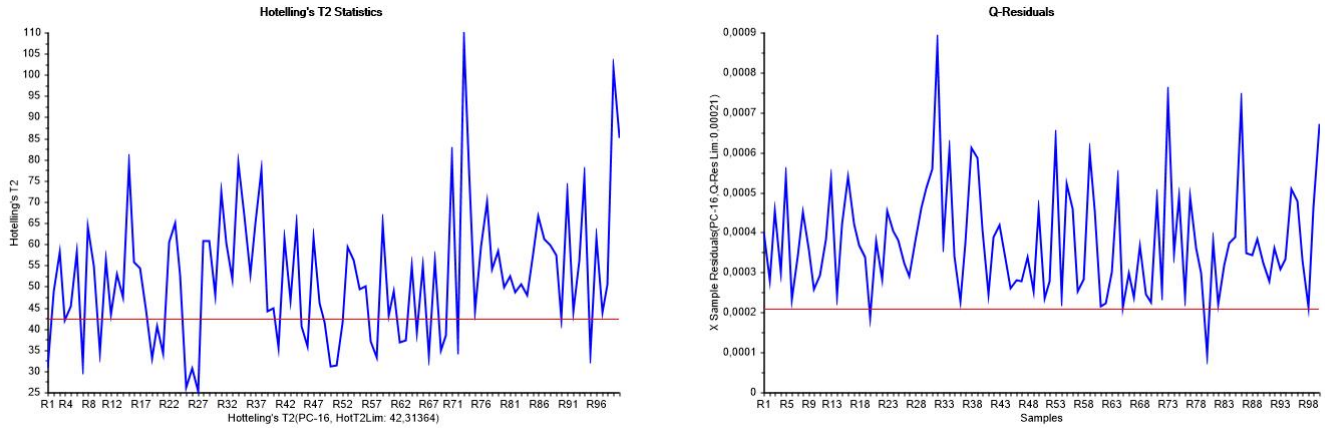

Figure 62: Projection of NRV fault testing data: Left: Hotelling's T^2. Right: SPE or Q-residual

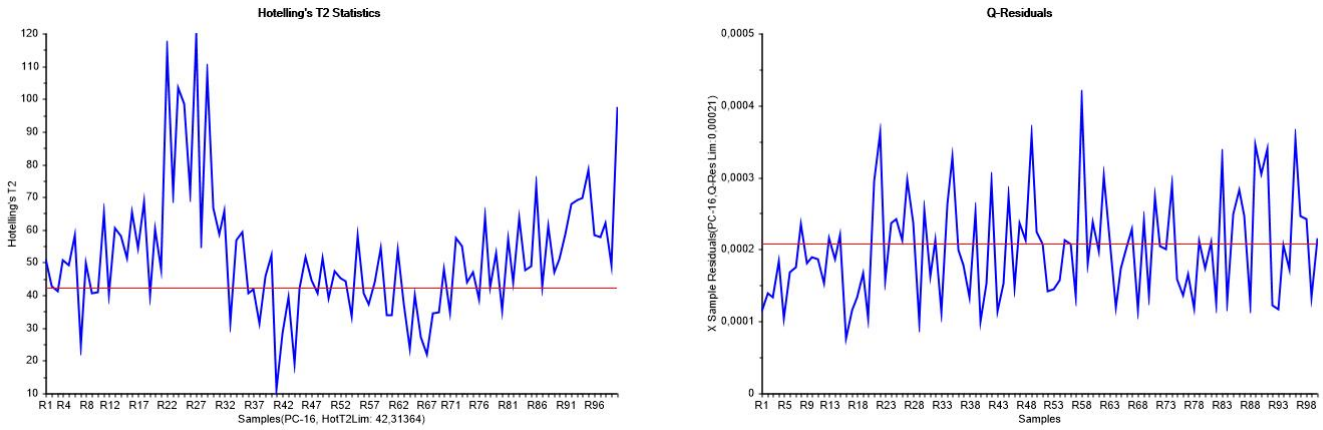

Figure 63: Projection of piston fault testing data: Left: Hotelling's T^2. Right: SPE or Q-residual

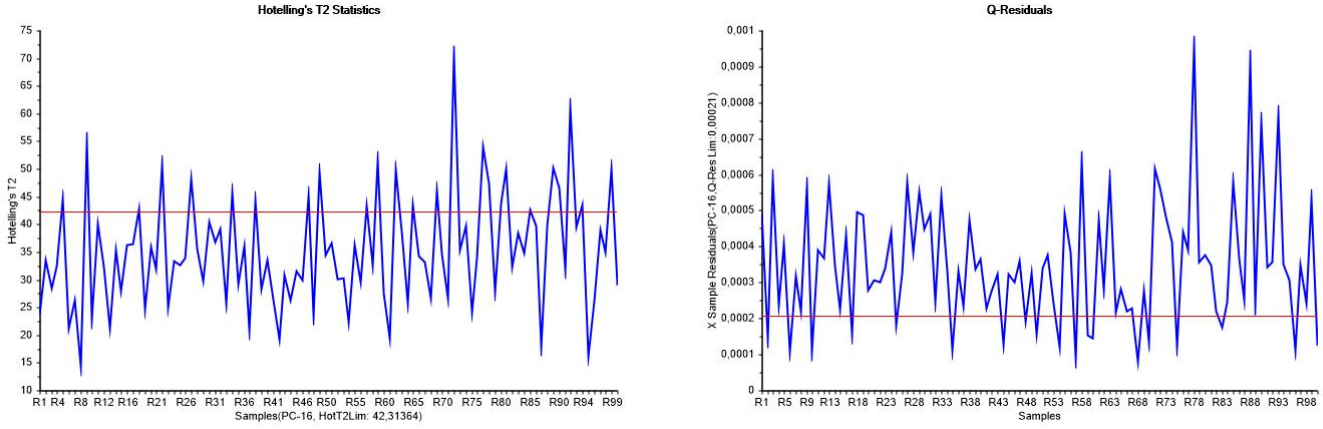

Figure 64: Projection of rider belt fault testing data: Left: Hotelling's T^2. Right: SPE or Q-residual

9.5 SIMCA PATTERN RECOGNITION AND CLASSIFICATION

This section regards classification of the states of the air compressor, utilizing pattern recognition of the extracted features using the SIMCA method. SIMCA based classification is a supervised learning approach where a PCA model corresponding to a class is trained with data from each of the states. Each PCA model will then capture the covariation structure for each state independently. The models may have different complexity. Table 8 shows the number of PCs retained and the explained variance for the PCA class models for each of the 8 states. The model complexity was chosen based on cross-validation. Each state was calibrated with training data of size 100x128, corresponding to 10 seconds.

Testing data from all states were then projected to all models. Each sample or feature vector was tested for belonging to

SIMCA Model	Number of PCs	Explained variance cal.	Explained variance val.
Healthy	16	95%	90%
Bearing	7	94%	93%
Flywheel	10	93%	90%
LIV	9	92%	90%
LOV	10	93%	90%
NRV	15	95%	92%
Rider belt	10	91%	87%
Piston	10	94%	92%

Table 8: SIMCA models: model complexity and explained variance

each models. This testing was based on two measures, namely sample-to-model distance (Si) and sample leverage (Hi). The sample-to-model distance is a measure of how far a sample lies from the modeled class, and is computed as the square root of the sample residual variance [13]. Sample leverage is a measure of the distance from the projection of a sample onto a model to the model mean. Small values of Si and Hi indicate that a sample belongs to the specific model.

The projected samples was subject to a F-test with a 5% confidence limit for Si, and a critical limit for Hi. The resulting memberships for the samples indicate that the SIMCA class models did not discriminate the class memberships very well, as many of the samples was positively classified to several classes. Table 9 shows the model-to-model distances for all classes. The distance between two models should be larger than 3 for the classes to be well distinguishable. We see that many of the distances are just above 3 which might explain why many of the samples where falsely classified to several classes.

	Model-to-model distances							
	Healthy	**Bearing**	**Flywheel**	**LIV**	**LOV**	**NRV**	**Rider belt**	**Piston**
Healthy	1	4.79	4.37	4.00	3.86	5.03	3.90	4.45
Bearing	-	1	3.15	8.71	7.57	9.94	5.28	4.75
Flywheel	-	-	1	5.95	6.59	6.03	3.44	3.41
LIV	-	-	-	1	2.57	3.05	3.49	3.34
LOV	-	-	-	-	1	3.10	3.75	3.13
NRV	-	-	-	-	-	1	3.88	4.18
Riderbelt	-	-	-	-	-	-	1	3.15
Piston	-	-	-	-	-	-	-	1

Table 9: Distances between models

A discriminating classification was then performed, where samples which had passed the F-test where classified to the class which had the smallest combination of Si and Hi in the following way: (Si/So)/(Critical F value) + Hi/(Critical Hi).

The results of the discriminating classification is shown in the confusion table 10. We see that this classification gives acceptable results for most of the classes. The test samples from the healthy state is 99% correctly classified. The classification method has some difficulty distinguishing the classes such as LIV fault and NRV fault, and bearing fault and flywheel fault.

		Predicted class membership							
		Healthy	**Bearing**	**Flywheel**	**LIV**	**LOV**	**NRV**	**Rider belt**	**Piston**
Actual class	**Healthy**	**99**	0	1	0	0	0	0	0
	Bearing	1	**86**	12	0	0	0	0	1
	Flywheel	1	6	**91**	0	0	0	2	0
	LIV	1	0	0	**70**	6	22	1	0
	LOV	0	0	0	0	**92**	2	2	0
	NRV	0	0	0	3	6	**88**	3	0
	Riderbelt	0	0	3	1	0	1	**92**	3
	Piston	1	3	0	0	8	0	4	**84**

Table 10: Confusion table for SIMCA classification with 5% significance limit with 100 test samples for each class

The poor class separation of the PCA based SIMCA classification approach might be due to the limitations of PCA as a linear analysis method, as high order nonlinear relationships of the data cannot be extracted [30]. Methods such as the improved PCA [32] or kernel PCA [30] might show improved results. Fuzzy clustering is also suggested techniques for class separation [14].

Part IV

DISCUSSION

10 DISCUSSION

10.1 DISCUSSION

From four different case studies, multivariate statistical techniques have shown to be effective methods to be used in condition monitoring applications where large sets of correlated data are analyzed.

In the case study for the steam turbine driven compressor in chapter 6, multivariate latent variable methods have proved their usefulness for effective dimension reduction of highly correlated process data. The study has illustrated how entire process units with large numbers of process parameters and condition monitoring instrumentation can be monitored much more effectively by analysis the covariance structure between the variables to find latent variables in a much lower dimension. The study has demonstrated how an in-control PCA model calibrated from historic process data from normal operating conditions, can be used for on-line monitoring of the condition of a steam turbine driven compressor. Projection of testing data from a period with a known fault condition has shown that the multivariate control charts Hotelling's T^2 and SPE are able to detect abnormal situations in the process at a early phase. Process variables that break their normal covariation structure are immediately detected before the their univariate alarm limits are reached. Through the use of contribution plots, the study has also shown how faults can be isolated after they have been detected. Process operators and engineers may have large benefits from the early fault detection and fault isolation.

The data set used for the study in chapter 6 had some limitations, as the data was from a period where the machine was not in a steady state operation. The non-stationary nature of the data made the modelling of a linear structure model difficult. This caused the residuals in the model to increase over time, causing false alarms. This problem could be avoided with a correct choice of calibration data from normal steady state operation. Miletic et al. [25] has covered an industrial perspective for implementation of multivariate statistic.

Chapter 7 covered the use of multivariate analysis for analyzing a long term trend in a steam turbine generator. By utilizing principal component analysis, a trend was discovered which seem to correlate with gradual degradation in performance that

has been experienced over several years. This shows how latent variable methods can give new insights in the hidden structures of multivariate data set.

The limitations of this study was that a vast number of the relevant process variables had missing values and could not be included in the analysis. This meant that a lot of the variation on the machine could not be analyzed, and hence potentially missing much of the relevant information.

The bearing analysis in chapter 8 showed that the use of wavelet packed decomposition in combination with principal component analysis was effective for fault detection and classification. Fault detection from the use Hotelling's T^2 and SPE has been effective, and all fault conditions were clearly detected. A problem met using these techniques was the occasional false alarms. Frequent false alarms can ruin the confidence in condition monitoring systems. Careful selection of the calibration set might reduce this problem, and filtering techniques for Hotelling's and SPE can be utilized.

The SIMCA classification shows a relatively good performance even if the linear PCA models does not capture higher order nonlinearity in the data. The performance of the classification should however be further tested to be able to tell if it can effectively classify future fault conditions.

Acoustic signal feature extraction and dimension reduction techniques for the purpose of fault detection and fault classification has successfully been analyzed from operational data from an air compressor in chapter 9. Using features from wavelet packet decomposition in combination with principal component analysis has been successful at detecting faulty from normal conditions. Using the SIMCA classification technique for supervised learning has also shown that different fault conditions can be classified. The SIMCA method had a relatively poor ability to effectively distinguish between different classes of fault states. This is due to the linear nature of the PCA models used in the SIMCA classification which are unable capture nonlinear relationships in the features. The data from the air compressor also shows a non-stationary nature as the load increases over time. This also make classification and false alarms more problematic.

10.2 FURTHER WORK

Suggestions for further work can be to investigate use of non-linear extensions of PCA such as kernel PCA, to improve the shortcoming of regular PCA when non-linear relationships are present in the data. KPCA could also improve fault classifica-

tion as it can better distinguish class memberships. Classification of multiple simultaneous fault condition could also be investigated.

Adaptive approaches to bilinear subspace models can also be powerful in many applications with time varying conditions.

Methods for avoiding false alarms are also important concerns that could be studied further. Fuzzy filtering is a technique that can be applied for this purpose [38].

Improving the fault isolation and diagnostic abilities of the traditional contribution plot by using reconstructions based contributions is also a promising technology.

There are also large potential for creating hybrid system through combining multivariate statistical methods with physical models. These techniques can compliment each other and improve the performance of condition monitoring systems.

Part V

APPENDIX

APPENDIX A

A.1 ACRONYMS

CBM Condition Based Maintenance

MVA Multivariate Analysis

PCA Principle Component Analysis

PLS Partial Least Squares

PLSR Partial Least Squares Regression

PLS-DA Partial Least Squares - Discriminant Analysis

SIMCA Soft Independent Modelling of Class Analogy

SPE Squared Prediction Error

MSPC Multivariate Statistical Process Control

NIPALS Nonlinear Iterative Partial Least Squares

STFT Short Time Fourier Transform

FFT Fast Fourier Transform

DWT Discrete Wavelet Transform

WPT Wavelet Packet Transform

WPD Wavelet Packet Decomposition

SVD Singular Value Decomposition

BIBLIOGRAPHY

[1] DWT Figure. 2017. URL: https://en.wikipedia.org/wiki/Discrete_wavelet_transform.

[2] Statoil ASA, Tjeldbergodden. 2017. URL: https://www.statoil.com/en/what-we-do/terminals-and-refineries/tjeldbergodden.html.

[3] The Society for Machinery Failure Pre- vention Technology, Vibration Institute, not-for-profit organisa- tion. Data Assembled and Prepared on behalf of MFPT by Dr Eric Bechhoefer, Chief Engineer, NRG Systems. 2017. URL: http://www.mfpt.org/FaultData/FaultData.htm.

[4] MathWorks. 2017. URL: https://se.mathworks.com/help/wavelet/ug/wavelet-denoising.html.

[5] Intelligent Data Engineering and Automation (IDEA) Laboratory, Indian Institute of Technology, Kanpur. 2017. URL: http://www.iitk.ac.in/iil/datasets/.

[6] F. Al-Badour, M. Sunar, and L. Cheded. "Vibration analysis of rotating machinery using time–frequency analysis and wavelet techniques." In: *Mechanical Systems and Signal Processing 25 (2011) 2083–2101* (2011).

[7] Carlos F. Alcala and Joe S. Qin. "Reconstruction-based contribution for process monitoring." In: *Automatica 45 (2009) 1593-1600* (2009).

[8] Bhavik R. Bakshi. "Multiscale PCA with Application to Multivariate Statistical Process Monitoring." In: *AIChE Journal, Vol. 44, No. 7* (1998). Dept. of Chemical Engineering, The Ohio State University, Columbus.

[9] Eric Bechhoefer. "A QUICK INTRODUCTION TO BEARING ENVELOPE ANALYSIS." In: *Green Power Monitoring Systems* (2012).

[10] Hocine Bendjama and Mohamad S. Boucherit. "WAVELETS AND PRINCIPAL COMPONENT ANALYSIS METHOD FOR VIBRATION MONITORING OF ROTATING MACHINERY." In: *JOURNAL OF THEORETICAL AND APPLIED MECHANICS 54, 2, pp. 659-670, Warsaw 2016* (2016).

[11] K. H. Esbensen and D. Guyot. *Multivariate Data Analysis - in practise*. Camo Process AS, 5 edition, 2002.

[12] K. H. Esbensen and D. Guyot. *Multivariate Data Analysis - in practise*. 5 edition. Camo Process AS, 2002.

[13] K. H. Esbensen and D. Guyot. *The Unscrambler Methods*. Camo Software AS, 2006.

[14] Jens C. Frisvad. *Classification 1 SIMCA*. 2010.

[15] Seyyed Reza Haqshenas. *Multiresolution-multivariate analysis of vibration signals: application in fault diagnosis of internal combustion engines*. 2012.

[16] Hanne Siri Amdahl Heglum. "Developing a Multivariate Prediction Model for Sleep Apnea from Polysomnography Data." In: (2016). NTNU.

[17] Honeywell. *Performance Monitoring of an Offshore Gas Compressor*. 2011.

[18] Manabu Kano, Shinji Hasebe, Iori Hashimoto, and Hiromu Ohno. "A new multivariate statistical process monitoring method using principal component analysis." In: *Computers and Chemical Engineering 25 (2001) 1103–1113* (2001).

[19] Bruel & Kjær. *Detecting faulty rolling-element bearings*. 2007.

[20] T. Kourti. "Application of latent variable methods to process control and multivariate statistical process control in industry." In: *International Journal of Adaptive Control and Signal Processing* (2005).

[21] Nabil Laayouj and Hicham Jamouli. "Prognosis of Degradation using Remaining Useful Life Estimation." In: *23rd Mediterranean Conference on Control and Automation* (2015).

[22] Daniel T.I. Lee and Akio Yamamoto. "Wavelet Analysis: Theory and Applications." In: *Hewlett-Packard Journal* (1997).

[23] H. Martens and T. Næs. *Multivariate Calibration*. Wiley, 1989.

[24] Neelam Mehala and Ratna Dahiya. "A Comparative Study of FFT, STFT and Wavelet Techniques for Induction Machine Fault Diagnostic Analysis." In: *COMPUTATIONAL INTELLIGENCE, MAN-MACHINE SYSTEMS and CYBERNETICS (CIMMACS '08)* (2008).

[25] Ivan Miletic, Shannon Quinn, Michael Dudzic, Vit Vaculik, and Marc Champagne. "An industrial perspective on implementing on-line applications of multivariate statistics." In: *Journal of Process Control December 2004, Pages 821–836* (2004).

[26] M. Misity, G. Oppenheim, and J. M. Poggi. *Wavelet Toolbox for Use with Matlab User's Guide*. Natick, MA, USA: MathWorks. 2007.

[27] Michael Nordmo. "Master's Thesis, Multivariate analysis." In: *NTNU* (2016).

[28] S. J. Qin. "Statistical process monitoring: basics and beyond." In: *J. Chemometrics Wiley InterScience* (2003).

[29] H. Saruhan, S. Sandemir, A. Çiçek, and I. Uygur. "Vibration Analysis of Rolling Element Bearings Defects." In: *Journal of Applied Research and Technology* (2015). Düzce University, Facult of Engineering, Düzce, Turkey.

[30] Renping Shao, Wentao Hu, Yayun Wang, and Xiankun Qi. "The fault feature extraction and classification of gear using principal component analysis and kernel principal component analysis based on the wavelet packet transform." In: *Measurement, Elsevier* (2014). School of Mechatronics, Northwestern Polytechnical University, Shaanxi, China.

[31] S.K. Singh. *Acoustics Based Condition Monitoring*. Dept. Of Mech. Engg., IIT Guwahati. 2008.

[32] Trendafilova, M.P. Cartmell, and W. Ostachowicz. "Vibration-based damage detection in an aircraft wing scaled model using principal component analysis and pattern recognition." In: *Journal of Sound and Vibration 313 (2008) 560–566* (2008).

[33] Nishchal K. Verma, Rahul Kumar Sevakula, Sonal Dixit, and Al Salour. "Intelligent Condition Based Monitoring Using Acoustic Signals for Air Compressors." In: *IEEE* (2016). Indian Institute of Technology Kanpur, U.P. 208016, India.

[34] Raffaele Vitale, Anna Zhyrova, João F. Fortuna, Onno E. de Noord, Alberto Ferrer, and Harald Martens. "On-The-Fly Processing of continuous high-dimensional data streams." In: *Chemometrics and Intelligent Laboratory Systems 161 (2017) 118–129* (2016).

[35] N.B. Vogt and H. Knutsen. "SIMCA pattern recognition classification of five infauna taxonomic groups using non-polar compounds analysed by high resolution gas chromatography." In: *MARINE ECOLOGY - PROGRESS SERIES l. 26: 145-156, 1985* (1985).

[36] Keshang Wang. *Intelligent Condition Monitoring and Diagnosis Systems*. Department of Production and Quality Engineering, Norwegian University of Science and Technology, Trondheim, Norway. IOS Press, 2003.

[37] Song Xing, Suting Chen, Zhanming Wei, and Jingming Xia. "Unifying Electrical Engineering and Electronics Engineering: Proceedings of the 2012 International Conference on Electrical and Electronics Engineering." In: Camo Software AS, 2013, pp. 1331–1332.

[38] S. M. Zanoli, G. Astolfi, and L. Barboni. "FDI of Process Faults based on PCA and Cluster Analysis." In: *Conference on Control and Fault Tolerant Systems* (2010).