8-2014

Prognostic-based Life Extension Methodology with Application to Power Generation Systems

Seyed Ahmad Niknam
University of Tennessee - Knoxville, sniknam@vols.utk.edu

To the Graduate Council:

I am submitting herewith a dissertation written by Seyed Ahmad Niknam entitled "Prognostic-based Life Extension Methodology with Application to Power Generation Systems." I have examined the final electronic copy of this dissertation for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy, with a major in Industrial Engineering.

John E. Kobza, J. Wesley Hines, Major Professor

We have read this dissertation and recommend its acceptance:

Belle R. Upadhyaya, Peter K. Liaw, Mingzhuo Zin

Accepted for the Council:

Carolyn R. Hodges

Vice Provost and Dean of the Graduate School

(Original signatures are on file with official student records.)

Prognostic-based Life Extension Methodology with Application to Power Generation Systems

A Dissertation Presented for the
Doctor of Philosophy
Degree
The University of Tennessee, Knoxville

Seyed Ahmad Niknam
August 2014

DEDICATION

I dedicate this work and give special thanks to my best friend ***MOJDE*** *a.k.a. the wife.*

ACKNOWLEDGEMENTS

I express my gratitude and deep appreciation to my advisors Prof. Hines and Prof. Kobza who were more than generous with their expertise and precious time. They have consistently helped me by sharing their knowledge and experiences kindly and without reproach.

I must express my grateful thanks to Prof. Liaw for his inspiration on doing fruitful research which is accompanied by his strong support. He always gives students the courage to face critical challenges.

For his helpful advice and support, I am especially grateful to Prof. Upadhyaya. His willingness to provide feedback made my teaching time an enjoyable experience. Appreciation is extended to Prof. Jin for his great helps and management in the reliability program and also for serving on my committee.

I am grateful to many people who assisted, advised, and supported my research over the years. In particular, I wish to thank faculties, staff, and graduate students with the Department of Industrial and Systems Engineering.

ABSTRACT

Practicable life extension of engineering systems would be a remarkable application of prognostics. This research proposes a framework for prognostic-base life extension. This research investigates the use of prognostic data to mobilize the potential residual life. The obstacles in performing life extension include: lack of knowledge, lack of tools, lack of data, and lack of time.

This research primarily considers using the acoustic emission (AE) technology for quick-response diagnostic. To be specific, an important feature of AE data was statistically modeled to provide quick, robust and intuitive diagnostic capability. The proposed model was successful to detect the out of control situation when the data of faulty bearing was applied. This research also highlights the importance of self-healing materials.

One main component of the proposed life extension framework is the trend analysis module. This module analyzes the pattern of the time-ordered degradation measures. The trend analysis is helpful not only for early fault detection but also to track the improvement in the degradation rate. This research considered trend analysis methods for the prognostic parameters, degradation waveform and multivariate data. In this respect, graphical methods was found appropriate for trend detection of signal features. Hilbert Huang Transform was applied to analyze the trends in waveforms. For multivariate data, it was realized that PCA is able to indicate the trends in the data if accompanied by proper data processing. In addition, two algorithms are introduced to address non-monotonic trends. It seems, both algorithms have the potential to treat the non-monotonicity in degradation data.

Although considerable research has been devoted to developing prognostics algorithms, rather less attention has been paid to post-prognostic issues such as maintenance decision making. A multi-objective optimization model is presented for a power generation unit. This model proves the ability of prognostic models to balance between power generation and life extension. In this research, the confronting objective functions were defined as maximizing profit and maximizing service life. The decision variables include the shaft speed and duration of maintenance actions. The results of the optimization models showed clearly that maximizing the service life requires lower shaft speed and longer maintenance time.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1 INTRODUCTION

Based on the estimated potential capacity of wind power, there has been a heightened interest in the United States to boost the share of wind turbines in electricity supply to over 20% by 2030 [1]. A critical aspect in the evaluation of economic viability of wind turbines is the difference between the market value of energy and the energy generation costs. From the turbine design and manufacturing point of view, this is a challenging problem because of the varying operating conditions, e.g. wind speed and unsteady aerodynamic loads. In effect, the operations and maintenance (O&M) costs are a major contributor to the energy generation costs. The O&M costs of offshore wind farms contribute about 25–30% to the energy generation costs [2]. It is to be noted that O&M costs vary based on location (offshore or onshore), weather conditions, turbine size and age, accessibility to drive train components, and availability of personnel, cranes and lifting equipment. Moreover, there are always concerns about unexpected equipment failure which can result in large maintenance costs e.g. intricate corrective actions as well as the downtime cost. Generally speaking, O&M costs are a major concern in all types of power plants particularly in the aging nuclear and fossil power plants which exceed the nominal design life. For nuclear power plants, O&M costs are estimated to be 60–70% of the overall generating cost [3].

It should be emphasized that the maintenance strategy is the main contributor to the O&M costs. As indicated by Lu et al. [4], if the maintenance strategy does not work appropriately, the cost of replacement related to a failed bearing in wind turbines would jump from $5000 to $250,000. Maintenance strategies can assure low maintenance costs based upon corrective maintenance or low operating costs through time-based maintenance. In essence, O&M costs can be minimized through improved reliability of the turbine's components, condition-based maintenance (CBM), and optimal management of maintenance requirements. This is often the case with all nuclear and fossil-fueled power plants as well.

The CBM of wind-turbines has received much attention during the past few decades [4-7]. Practically, CBM can restore the balance between the maintenance costs and operations costs by minimizing unplanned downtime and unnecessary maintenance. A CBM system consists of three major units: (1) data acquisition, (2) data processing, and (3) maintenance

decision-making. Of special concerns are system-level and subsystem-level fault diagnosis and prognosis, since modern wind turbines are equipped with CBM systems. Diagnostics and prognostics are two important aspects of a comprehensive CBM program, which provide information for maintenance decision-making with respect to safety, operations planning and service parts logistics. Nowadays, prognostic-related activities are mainly embedded within the domain of Prognostics and Health Management (PHM) systems. Figure 1.1 shows a typical PHM system.

Figure 1-1 A typical PHM system

A key component missing from current PHM programs is the trade-off between the loss of power generation and O&M costs. The cost associated with the loss of power generation (i.e., plant outage) is a significant portion of the total downtime cost. In effect, the main purpose of PHM program is to reduce both loss of power generation and O&M costs. However, uncertain operational, meteorological, and logistical situations along with diverse technical health states of turbine components requires PHM programs to balance between O&M costs on one side and the loss of power generation and turbine durability on the other side. To provide a platform for this multiple criteria decision making (MCDM), it seems essential to have adequate information with respect to failure-causing faults and the remaining operational life of the faulty sub-systems and components. In other words,

decision-making and planning for O&M need to take diagnostic and prognostic systems into account. The primary objective of this research is to develop a prognostic-based life-extending methodology in order to improve the current PHM practices in power generation.

1.1 Problem Statement

Although considerable research has been devoted to developing prognostics algorithms, rather less attention has been paid to post-prognostic issues such as maintenance decision making. In particular, an effective prognostic system may play a significant role in improving the durability of engineering systems. It is to be noted that applying prognosis for extending the operable life of power generating systems has been noticeably lacking in published literature.

In this study, it is assumed that the first priority of the decision maker (DM) is to mobilize the potential residual life (i.e., extend the life of the system) to improve the return on investment. The obstacles in performing life extension include: lack of knowledge (e.g. about aging mechanisms, life-limiting factors, premature or impending failures and relevant collateral damage), lack of tools such as trust worthy prognosis algorithm, lack of data (e.g. an adequate RUL estimation), and lack of time. In essence, to overcome the above-mentioned obstacles, a life extension program requires numerous elements including: diagnostic, prognostic, maintenance and spare parts management, knowledge about operations and environmental conditions (before and after life extension), decision support system, and a methodology to analyze the efficiency of a life extension program. The major purpose of this research is to introduce a credible framework for prognostic-based life extension and demonstrate its applicability by elaborating certain elements of this framework. Figure 1-2 describes the framework for prognostic-based life extension.

The first and foremost element of an efficient life-extending methodology is effective and robust diagnosis, which detects and identifies the impending faults. A degraded component or system may affect other parts of the system. Also, competing failure modes need to be taken into account. This research primarily considers using acoustic emission (AE) technology for quick-response diagnostic. To be specific, an important feature of AE data will be statistically modeled to provide quick, robust and intuitive diagnostic capability.

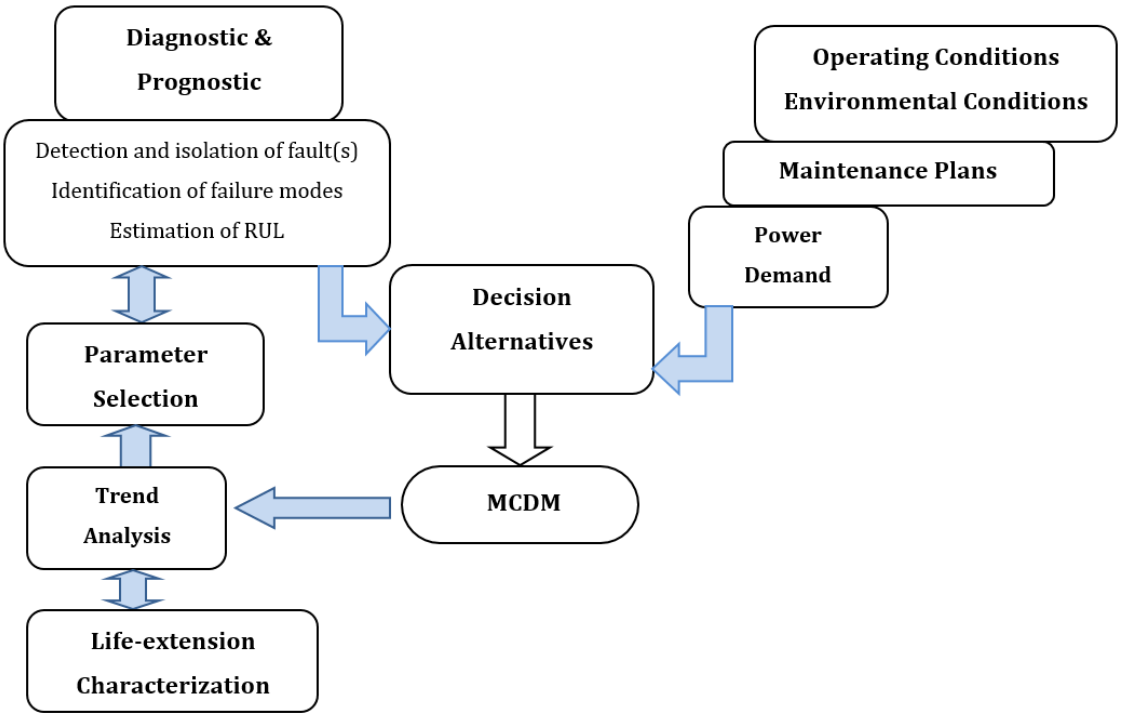

Figure 1-2 Framework for prognostic-based life extension

Next, real-time estimates of remaining useful life (RUL) need to be used to characterize and quantify the life extension practices. Useful life can be defined as the ability for power generation to fully or partially meet demands. RUL is the time that a component or system is able to operate with the same predefined specifications without the need for major repair or maintenance or without a significant rate of minor faults. In this study, the primary focus is on individual-based prognostic using sensed or inferred degradation measures. It is not unusual to combine degradation measures into a single prognostic parameter. In essence, RUL is subject to uncertainties and this creates the need for detailed characterization of prognostic parameters and selection of the appropriate parameters.

One main component of the proposed framework is the trend analysis module. This module analyzes the pattern of the time-ordered degradation measures. The trend analysis is helpful not only for early fault detection but also to track the improvement in the degradation rate. This research considers trend analysis methods for the prognostic

parameters, degradation waveform and multivariate data. In addition, algorithms to address non-monotonic trends will be considered.

All the activities in the diagnostic and prognostic modules need to be followed by fruitful decision making. Maintenance-decision making (MDM) is by nature a complex process due to the followings:

1. Presence of various decision alternatives, and therefore, the need for compromising,
2. Multiple criteria which usually conflict and make the judgment of alternatives a complex endeavor,
3. Uncertainty associated with the available data i.e. RUL estimates.

The major step in decision making is to determine alternatives which are defined as possible courses of action [8]. Possible actions may include fault accommodation (i.e., modifying control rules or using redundancy), altering operation conditions or tactical control, and maintenance practices. Such actions need to restore the strength distribution or diminish the degradation, e.g., by reducing the stress. The alternatives may include urgent repair, working with modified/manipulated operating conditions, or working with no changes in operating conditions.

Thus, an optimization procedure is required to clarify the parameters of these alternatives. For instance, the optimal shaft speed must be determined that allows the generator to meet demands and also diminishes the effects of component failure at the same time. In addition, it is essential to update RUL estimates after applying each alternative in order to characterize the life extension or the need for repeating the decision making. Moreover, in the case of life extension one can expect explicit change in the trend of component degradation or aging. The trend analysis module tests the trends of prognostics parameters (monotonic vs. non-monotonic) in order to ensure an appropriate parameter selection.

Another concern in developing the alternatives is to find the optimal time to perform the actions. In effect, performance of maintenance at inconvenient times results in substantial time and economics losses. In this respect, logistics costs associated with maintenance can justify the option of run to failure. For instance, the cost and availability of vessels with cranes used in off-shore wind turbine maintenance may justify run to failure or even out of

schedule maintenance for multiple turbines. For this, maintenance plans need to be considered as inputs to the decision support system along with operations and environmental conditions. On the other hand, the plan for electricity generation needs to be taken into account. In effect, loss of power generation should be considered for the period of shut down and the period of inefficient performance.

1.2 Original Contributions

It is important to emphasize that life extension is case–dependent. In addition, it has a wide range of issues and options to be discussed. For these reasons, this research does not cover all the aspects of a prognostic-based life extension program. Hence, the discussion focused on the presence of non-monotonic degradation during life extension, trend analysis of degradation data and a decision support system that use prognostic information for life extension. Accordingly, the original contributions are as follows:

- Development of a novel diagnostic method based on AE energy, Zero-inflated Poisson (ZIP) regression and control charts. The contribution can be further expanded by applying pattern recognition methods to identify faults and the limiting failure modes.
- Performing self-healing experiments using AE.
- Enhancement of parameter selection procedures through the techniques of trend analysis. Since the life extension activities may result in non-monotonic trends, there is a need for a trend analysis module to test the trends and modify the degradation paths.
- Development of a prognostic-based decision support system using multi-objective optimization methods. The goal is to extend turbine durability. The contribution can be further expanded by taking superior fatigue life models into account and estimating the optimal time for life extension activities.

1.3 Organization of the Document

Chapter 2 provides a detailed literature survey. Because of the wide scope of this research, the literature review was divided into several subsections in order to provide better insight. It is to be noted that, when necessary, some literature review is presented in other chapters to support a particular subject.

Chapter 3 covers certain examples of nonmonotonic degradations. This discussion is followed by trend analysis in Chapter 4. Next, the interactive decision support system using prognostic information is presented in Chapter 5. Concluding remarks and recommendations for future work are presented in Chapter 6.

2 LITERATURE SURVEY

2.1 Diagnostics

Diagnosis systems deal with fault detection, isolation and identification of the nature of faults. Diagnostic methods can be classified into two major groups: data-driven and model-based. The former consists of pattern recognition, statistical approaches and artificial intelligence (AI) methods. Pattern recognition methods include power spectrum graphs, wavelet phase graphs, phase spectrum graphs, cepstrum graphs, spectrograms, and wavelet scalograms. Examples of statistical approaches are regression methods, cluster analysis, support vector machines, and hidden Markov models. AI approaches include artificial neural networks, evolutionary algorithms (e.g. Generic Algorithms), fuzzy logic systems, neural-fuzzy, and fuzzy-neural systems.

The application of the above-mentioned approaches is a function of the quantity and quality of data as well as the data analysis techniques. This may cause some obstacles for developing generic diagnostic systems in complex mutli-component systems such as wind turbine. Diagnostic data can be classified into three groups:

- Value type data is collected for single value variables like temperature, pressure, and humidity,
- Waveform type is time series data which is collected over a specific time epoch and for specific variables (e.g. vibration data, acoustic data), and
- Multi-dimension type such as image data.

In this respect, signal processing is the process of data analysis and feature extraction in waveform and multi-dimensional data which can be categorized into three types: time-domain analysis, frequency-domain analysis (e.g., fast Fourier transform, signal average method, cepstrum analysis, and high frequency resonance technique), and time-frequency analysis. The most common time-frequency methods include wavelet transforms, Hilbert-Huang transforms (HHT), and short-time Fourier transforms). These methods can handle non-stationary and nonlinear signals. In essence, traditional signal analysis relies on the limiting assumption that signals are stationary. However, in the real physical world, the

signal features change over time. The research on advanced machine diagnostics has tended to focus on extracting information from non-stationary signals.

Wavelet analysis is a powerful method to examine non-stationary signals, detect the transient signal, measure the time evolution of the frequency transitions, and segregate the amplitude and phase component [9]. In wavelet analysis, the signal is decomposed into different detail signals that correspond to different frequency bandwidths.

HHT is applied in this research for the trend analysis of degradation waveforms. The Hilbert-Huang transform (HHT) [10] is a relatively new time-frequency technique. The HHT method is able to analyze large signals. The empirical mode decomposition (EMD) is the first major operation of the HHT. The details of EMD process can be found in [11, 12]. The EMD relies on the local characteristic time scales of a signal and can decompose the signal into a collection of successive intrinsic mode functions (IMFs). In other words, EMD intends to accurately reveal the signal characteristics. An IMF is a function that reveals a simple oscillatory mode embedded in the signal [12]. An IMF may suffer from mode mixing problem. After finding the IMFs, the Hilbert transform (HT) is applied to produce a full time-frequency-energy distribution of the signal. The HT process provides instantaneous frequency and amplitude information for each IMF.

The fault diagnosis of bearings is elaborately discussed in the next section. In addition, the review particularly considers the applications of acoustic emission (AE) in fault diagnosis of rotating components. Bearing and gearing systems have some common failure mechanisms such as pitting, and therefore they have similar diagnostic techniques. Moreover, AE has been extensively studied for gearing systems rather than bearings. For these reasons, fault diagnosis of gear systems is also presented in the literature review.

2.1.1 Fault diagnosis in bearing and gearing systems

Rolling Element Bearings (REB) are one of the primary causes of breakdown in rotary systems. Examples for catastrophic bearing failure include automatic processing machines and helicopters [13]. The major failure modes in bearing include spalling, pitting, flaking, brinelling, fluting, seizure, etc. Fatigue spalling is the major cause of diminished life for bearing races or rolling elements. Spalling occurs mostly in the rolling direction in the region of maximum shear stresses. Failures in bearings occur because of friction, wear, damage to

housing due to turning of the outer race, cage fatigue or broken cage rivets. In effect, relative motion between the balls would be the root cause of the cage failure. The relative motion happens because of angular misalignment between the inner and outer bearing races or a decrease in diametrical clearance.

In years past, the L10 bearing life prediction [14] was used as the reliable method to denote bearing life. L10 Life is also introduced as "basic rating life" in bearing catalogs. The bearing life in this method is the time or number of stress cycles that 90 percent of a population survives. Based on this principle, the relation between bearing life and the dynamic load capacity (C_D) with 90 percent probability of survival for carrying one million inner-race revolutions has been used for many years. In this equation, p is the load life exponent and B_L is the equivalent bearing load [14]

$$L_{10} = [C_D / B_L]^P \qquad \text{Eq. 2-1}$$

ISO 281 specifies the standard way to calculate the rated life of a bearing. In addition, ISO 16281 [15] provides the methods for calculating the fatigue life for universally loaded single-row REB. The presented calculation methods can be derived for more complex designs based upon the given geometry references. The standard considers lubrication, contamination, fatigue load limit of material, internal load distribution on the bearing, tilting, misalignment, and operation clearance. Dynamic effects such as centripetal and gyroscopic forces were not taken into account. In reality, dynamic effects would be significant in high speed applications.

Based on Hertz stress-life relation, life in a ball bearing is a function of cubic power of load. In roller bearings, life is a function of forth power of load. Zaretsky et al [16] discussed various life theories for REBs. The authors believed that Lundberg-Palmgren theory is the best life prediction method. This theory is based on orthogonal shear stress. A few researchers have modeled bearing crack propagation based on Paris's formula [17]. This formula relates the rate of defect growth to the instantaneous defect area D as follows:

$$dD/D_t = C_0 (D)^n \qquad \text{Eq. 2-2}$$

Here, C_0 and n are material constants. Accordingly, under constant operating conditions, a deterministic crack propagation model in bearings relates the instantaneous defect area

with material constants that have no correlation with the defect size. Nevertheless, these models lack the exact shape of crack geometry, and therefore the stress intensity factor was roughly determined. In essence, bearing fatigue processes are highly stochastic. Numerous factors contribute to the fatigue crack growth including stress states, lubrication, material properties, and other environmental and operating conditions. In this respect, the propagation mode of fatigue spall is a subject deserving of attention. The investigation conducted by Hoeprich [18] was used as the industry standard for unacceptable fatigue spall damage size which is 0.01 in^2 (6.25 mm^2).

Gearing systems are one of the acknowledged examples of the weakest-link-in-the-chain components in engineering systems such as automotive, helicopters, mining equipment, and Mega-Watt-level wind turbines. Gearing systems in wind turbines transmit force from the blades to the generator by means of a group of solid bodies i.e., gears, shafts and bearings. The most used gearing systems in wind turbines are planetary gearing systems, also known as epicyclic gearing. The major advantages of planetary gearing systems include high power density, compact design and various gearing configurations. The planetary gearing systems in wind turbines should be able to handle the torque generated by heavy and long blades and increases in speed of the main shaft by the scale of 50 to 300. With an operational period of 5 years, gearbox replacement in wind turbines may lead to 1-6 months of production loss and a cost up to 10 percent of the original construction cost [19]. Moreover, a higher frequency of failures has been reported in bigger size wind turbines.

One approach to improve the reliability and availability of gearboxes is to reduce the number of moving components. For instance, recent developments in wind turbine design advocate the use of direct drive generators (e.g. permanent magnet gearless synchronous generators) which avoids the need for an additional gearbox and leads to significant reduction in mechanical wear.

Major gearbox failures include shaft misalignment, local and distributed faults of gears and local faults in rolling element bearings [19, 20]. For instance, the movement of machine chassis in wind turbines results in misalignment of the output shaft of the gearbox, which finally leads to bearing failure. Moreover, wind gust loads (which cause cyclic stresses) and manufacturing inconsistencies may cause minor misalignment. Highly reliable gearboxes

such as the Geared Turbofan Engine (GTF) take the advantage of magnetic bearings to minimize wear, self-align the transmission shafts and operate lubrication-free.

Gear faults include cracks, tooth bending, tooth breakage, and tooth surface deformation such as scuffing, pitting and spalling. Pitting is a fatigue effect caused by micro-cracks due to surface and subsurface stresses. Pitting is generally initiated within the hardened layer and near the pitch surface or on the dedendum flank where the oil film is thin [21-24]. The micro-cracks propagate and consequently lead to removal of tiny particles and formation of cavities on tooth flanks. Aslantas and Tasgetiren [21] provided a numerical model of pitting formation life for spur gears using the finite element method and linear elastic fracture mechanics analyses. This work demonstrated that pitting failures might have irregular morphology, but rolling direction has no influence on the shape of pitting. In general, heat treatment prolongs the surface fatigue failures of gears. For more information on gear wear and damage mechanisms, the readers are referred to ISO 10825 [23].

The CM programs for bearing and gearboxes are mainly based on vibration analysis, acoustic emission (AE), and oil debris analysis. AE is a promising CM technique with early fault detection ability. AE is broadly reviewed in the next section. Oil debris monitoring deals with detecting the number and size of oil particles that used to be a piece of the contact surfaces. Early stage of pitting can be identified via oil debris monitoring. This technique is not able to provide helpful information for fault localization.

Vibration analysis is a well-established monitoring technique for all rotary systems. It is pertinent to mention that monitoring of bearings and gearboxes can be directly solved by root cause analysis of the factors influencing the vibration [20]. In effect, vibration analysis techniques based on the bearing or gear case vibration signatures reveal a range of faults. It was noted that the stiffness of the operating gear system is the main influencing parameter in terms of vibration level [25]. Furthermore, it should be noted that varying loads affect the vibration signal modulation. Condition indicators for gearbox condition monitoring include [26]:

- Root mean squared (RMS) value for the velocity: this parameter deals with the energy content of the signal. RMS is a simple and robust descriptor of the overall condition.

Although it is not sensitive to incipient tooth failure, RMS would increase with progress in tooth failure (e.g., pitting) and wear out.

- Delta RMS is associated with differences between consequent RMS values. This parameter observes the trend of the vibration signal. Gear faults impose rapid increases in Delta RMS,
- Crest factor is the ratio of peak value of the signal over the RMS. This parameter may reveal the damage in gearing systems at an early stage.
- Kurtosis is the 4^{th} centralized moment which deals with the shape of amplitude distribution,
- Energy ratio is the ratio between the energy of the regular meshing and the energy of the signal. This parameter is useful in indicating heavy wear (i.e., where more than one tooth is damaged),
- Sideband index, and
- Peak value.

Bartelmus and Zimroz [27] introduced a simple regression-based diagnostic feature relying on the linear relationship between the sum of the meshing component amplitudes and the instantaneous input speed as the operating conditions indicator. The gearing system under study is a multistage gearbox used in a bucket wheel excavator. This gearing system examines non-stationary operating conditions. The general and cyclic variations in operating conditions come from varying properties of ground and the manual control. After 20,000 hours of operation, the gearbox suffers from over limit radial backlash in all rolling elements bearings and scuffing of almost all the gears. The author considered this stage of life as a "bad condition" which is also more susceptible to load than the gearbox in good condition.

The linear relation described above depends on load susceptibility. Hence, the new feature is the difference of the slope of the linear relationship in different cases, i.e. the good and the bad condition. In effect, the linear relation would be applicable in a limited range of operating conditions. Moreover, changes in the conditions of tooth surface invalidate the assumption of linear relationship. It is of note that increased backlash is the major reason of degradation in tooth surface conditions.

It may be worth stating that gear dynamic models have been investigated with the hope to provide superior analysis of vibration generation mechanisms and dynamic behavior of faulty gears. The dynamic models appearing in the literature can be divided into three groups; (1) the models that consider dynamic factors to determine root stress formulas, (2) models that merely considered tooth stiffness as the storing element, and (3) recent dynamic models of gears which take into account shaft and bearings. For diagnostics purposes, Howard et al. [28] simulated gear dynamics of a single stage reduction gearbox focusing on the friction between the gear teeth through a multiple-degree-of-freedom model. The match between the simulation and the experimental results was promising.

2.1.2 Acoustic Emission

Acoustic Emission (AE) is a powerful passive method for real-time investigation of materials behavior under mechanical and/or thermal stress. AE is a natural phenomenon in a material deforming under stress, in which localized physical changes within the material result in rapid release of strain energy that leads to emission of transient elastic waves [29]. In fact, AE is a well-established diagnosis method for static structures. In recent decades, AE applications in health monitoring of structures and materials have been extended to other areas such as rotating machineries and cutting tools. The ISO 22096:2007 standard provides the general principles of AE application.

The major advantage of AE technology is the sensitivity to surface and subsurface micro-damage. For this, AE can ensure superiority over vibration-based monitoring systems with respect to discovering crack propagation and early fault detection, and therefore, it offers good potential for prognostic capabilities [24, 29-31]. Price et al. [32] confirmed the detectability of AE in severe sliding and pitting situations by observing changes in AE energy. The authors identified changes in the frequency patterns of AE signals prior to fault appearance and emphasized the superiority of AE in wear detection over vibration monitoring or wear debris analysis. In addition, Tan et al. [25, 33, 34] verified the supremacy of the AE technique in detecting and monitoring pitting in comparison with vibration analysis or spectrometric oil analysis techniques. These studies also emphasized the usefulness of AE for prognostics in rotating machinery considering the near linear relationship between AE and pit progression. Badi et al. [35] simulated scuffing and pitting

defects on gear tooth. The authors used crest factor and kurtosis for both AE and vibration analysis. While both monitoring techniques were able to detect scuffing, the pitting was detected only through the AE technique.

The potential sources of AE in bearings are subsurface cracks, rubbing between damaged mating surfaces, excessive temperatures, and lubrication malfunction. It is worth noting that increases in AE parameters may happen prior to bearing failure particularly in dry runs. It is also the case when the operating conditions such as temperature or speed change. For these reasons, selecting the most appropriate failure threshold is a challenging task.

Al-Balushi and Samanta [36] introduced energy-based features extracted from AE signatures. This index was defined as the square of the ratio of the RMS value for a section of the signal to the RMS value of the entire signal. The proposed index was successful in detecting and locating the broken and pitting teeth. In general, the proposed technique worked better than the traditional kurtosis and crest factor methods. In addition, Al-Balushi and Samanta [37] employed wavelet transforms to decompose AE signals to low frequency and high frequency components. The output of wavelet transforms was used as input to an ANN based diagnostic approach.

The literature reports an increase in AE parameters (e.g. RMS, amplitude and energy) due to a increase in gear defect size [22, 24, 30, 38]. Singh et al. [39] concluded that pitting of gears can be detected earlier by AE at the crack initiation and growth stages, although this conclusion would not be acceptable for unloaded gears at extremely high speeds. By contrast, vibration monitoring reveals the fault at a later stage of the crack growth process.

Yao et al. [9] believed that an efficient fault diagnosis requires a match between the time-frequency structure of the wavelet transform and the transient components of the signal. In view of this, the authors proposed a new adaptive Morlet wavelet filter for more effective diagnosis of incipient gear crack faults. Kurtosis maximization was applied to optimize the wavelet parameters. In practice, tooth crack or breakage generates periodic impulses (0.044 second in this case) which produce resonance features in the vibration signal. However, the periodic impulses generated by tooth crack are usually hidden in the signals. The use of Morlet wavelets was proven to be effective in extracting such impulses. In general, to

appropriately decompose these features, one should apply wavelet transforms with superior time resolution at high-frequencies.

Fault detection in split torque gearboxes (STG) using acoustic emission and vibration sensors was studied by He and Eric [40]. STG is a reliable gearbox because of fewer gears and bearings and a reduced number of speed reduction stages. This work found a decrease in signal to noise ratio and accordingly diminished fault signal because of a large number of synchronous components meshing or rolling in close proximity. The following AE parameters were used in this work for fault detection: ring-down count, duration, peak amplitude, rise time, rise time slope, RMS, and kurtosis. A Hilbert-Huang transform based algorithm was used to extract gear fault features from the vibration signals. It was concluded that AE offers 100% accuracy in damage classification. In addition, AE signals do not require complicated algorithm to generate gear fault features and accurate AE signal would be in the form of high counts and low rise time.

Raad et al., [41] used visual comparison and confirmed that spalling is the source for the AE bursts. Sentoku [38] claimed that increasing in pitting would result in friction which consequently increased the AE amplitude. Tandon and Mata [24] experimentally demonstrated the detectability of AE for defect size (pit diameter) around 500 micrometer or greater in diameter. The detectability of vibration monitoring was recognized for a defect size more than 1000 micrometer.

Eftekharnejad and Mba [30] considered the fault detection of helical gears. It is pertinent to mention that the AE signal has two types of time domain waveforms; burst and continuous. This study shows that transient AE bursts come into sight on the continuous type AE waveform of helical gears exposing the defects in helical gears. While the accelerometer was located on the bearing pedestal, the authors believe in the superiority of AE over vibration analysis in identifying the seeded defects on helical gears. In the case of rolling element bearings, significant changes in vibration can be observed when the remaining operational life is very short [29, 42].

Moreover, insensitivity to structural resonance and mechanical background noise gives the AE technology an extra advantage over the vibration-based monitoring systems. Depending on the data acquisition system, the smallest scale of AE inspection would be the

small dislocations within the elastic stress field. In fact, there is an acoustical reaction to permanent deformation processes such as twinning, slip and micro-crack formation. Nonetheless, because of the wide frequency range (10 kHz to 10MHz) AE is applicable not only for crack and corrosion detection but also for identifying certain processes like friction, solidification and phase transformation.

It is recognized that the interaction of component's elements in relative motion is the primary source of AE in rotating machinery. This interaction can be in the form of impacting, cyclic fatigue, friction, turbulence, material loss, leakage, etc [29]. Thus, in rotating machineries, the number and arrangement of component interfaces, and loading conditions can influence the propagation of the AE signal.

Furthermore, AE is not geometry-sensitive but it is material sensitive. In effect, AE signals depend on material microstructure, non-homogeneities and deformation mode. Principally, ductile materials show low emissivity in comparison with brittle material [43]. Suman et al. [44] believed that AE in metals can be generated due to dislocation movements in plastic deformation or the initiation and growth of cracks. Miyachika et al. [45] considered bending fatigue failure for two types of gears. The authors postulated that defect detection through AE would be easier in the case of case-hardened gears in comparison with normalized gears. Marked increases of AE parameters were adequate to characterize both crack initiation and crack growth for case-hardened gears.

The major drawbacks of AE-based monitoring systems include the distance of AE sensors to the source and the weakness of signals in force or intensity, i.e., attenuation [22, 25, 33]. In practice, severe attenuation and reflections may happen due to sensor positioning and complexity of machines. For this, an attenuation test (i.e. Nielsen source test) is recommended in order to display different interfaces [22]. Singh et al. [46] realized that attenuation at each individual interfaces contributes to the overall attenuation across the gearing system. According to [22], attenuation would be greater for lighter loadsing conditions.

The positioning of AE sensors in gearing systems and bearings is a subject deserving of attention. While it seems the most convenient place would be the gearbox casing [22, 24], several researcher [25, 33, 38, 45] attempted to place the sensor on the side of the gear

namely close to the gear teeth. For such experiments, a slip ring was applied to transfer the monitoring data. Toutountzakis et al. [5] critically reviewed the developments in application of AE for spur gear systems. The authors believed that gear defect detection with AE is accompanied by technical difficulties that are mainly centered on understanding the parameters that influence the AE signal. These parameters include design factors (e.g., meshing mechanisms and asperity contact), operational factors (e.g., speed, torque and specific film thickness), and degradation factors (e.g., removed wear particles or debris interfering with gear mating). It is to be noted that, for spur gears, a gear meshing cycle is a combination of rolling and sliding. The rolling portion of gear meshing principally occurs on the pitch line. Asperity contact occurs during the sliding.

Tan et al [33] explained that high amplitude AE transient bursts are generated at the gear mesh frequency because of the rolling contact on the pitch line of the spur gear mesh. It is interesting to note that in the time domain of AE signals, it is possible to calculate the gear mesh frequency by inversing the periodic time between two succeeding AE bursts [22]. However, the continuous waveform is produced because of the sliding contact. In principle, sliding, rolling or a combination of both will occur during the gear mesh, which is known as asperity contacts. It is notable that, in the gear industry, rolling contact fatigue cracks is one of the most challenging problems. The authors in [33] believed that under the conditions of constant temperature the load has no significant influence on AE RMS levels. On the contrary, speed would significantly influence AE levels since speed is associated with the sliding speed of the meshing gears. In brief, the source of AE signals is the elastic deformation of the material at asperity contacts [25].

Several studies consider the important effects of lubricant on AE activity [22, 25, 47]. In effect, lubricant is crucial in order to maintain the mechanical integrity of gears and bearings [47]. The thickness of oil film directly affects the rate of wear and asperity deformation and consequently AE activity is influenced. Therefore, one application of AE technology in gearboxes would be monitoring the lubricating conditions. The thickness of the oil film depends on oil temperature, surface roughness, load and speed of the meshing gears [47]. Hamza and Mba [47] realized that the oil film thickness in meshing gears with rolling and sliding (spur gear) has more effects on AE activity in comparison with meshing gears with

pure rolling (helical gear). Oil temperature influences AE activity in the case of fixed speed and load conditions particularly for higher rotational speeds [22, 25]. Oil temperature was introduced as the main reason for unsatisfactory seeded defect identification in [25] since increasing loads and speeds increased the oil temperature.

To this end, it is possible to say that the development of AE applications for bearings and gears is merely dependent on experimental findings using seeded defects. In most cases, researchers are hardly aware of mounds or protrusions, which may exist at the boundaries of the seeded defect. Protrusions generate AE activity and may lead to growth in vibration [30]. Hence, it is worthwhile to consider more detailed experiments based on natural pitting or wear.

The most common measurable characteristics of the AE signal include peak amplitude, energy, duration, counts, and rise time [24, 43, 45]. Figure 2.1 illustrates these measurable characteristics. Peak amplitude is a function of AE source scale and velocity [48]. The mentioned parameters can be used to provide more signal features such as RMS, AE cumulative event count, counts to peak, rise time slope, crest factor and Kurtosis [45]. Moreover, the distribution of events versus the above-mentioned parameters can be used as signal features [24].

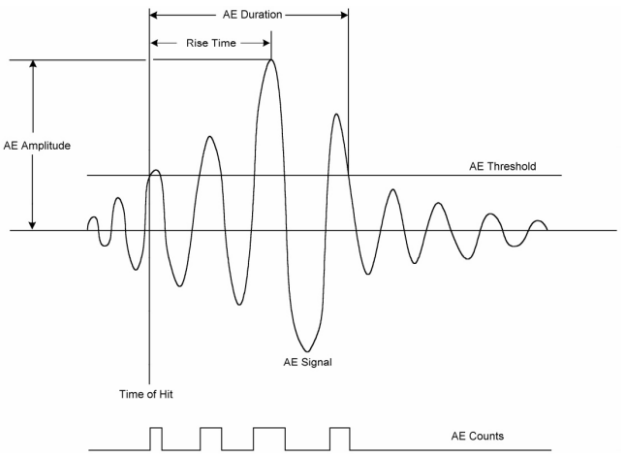

Figure 2-1 Measurable characteristics of AE signal [43]

AE count is defined as the number of threshold crossings of the AE signal. In the literature, this parameter is also known as ring down count (RDC) or threshold crossing count. In effect, AE counts imply the existence of a transient event (e.g. rolling action of meshing gears). One major drawback of this parameter is the dependence on the threshold level, signal frequency and the amplitude of the AE pulses. It is to be noted that the average counts is a measure of AE intensity, i.e., the size of the emission signals detected [43]. Furthermore, AE counts divided by duration gives the average frequency of the signal. It is therefore difficult, if not impossible, to identify the origins of defects in bearings and gears using AE counts.

Traditionally, RDC has been widely used in the literature as the condition indicator. Miyachika et al. [45] postulates that AE cumulative event count increases with crack growth in the case-hardened gears. Tandon and Mata [24]claimed that AE counts showed better results than other AE parameters in gear defect detection. Tandon and Nakra [49] considered AE counts for condition monitoring of radially loaded ball bearing. The authors observed a direct relationship between AE count and speed in the case of outer race defect. The results showed that AE count is a practical indicator for defects less than 250μm in diameter. However, Morhain and Mba [50] declared that AE count is able to detect large defects up to 15 mm in lengths and 1 mm in width. The authors in [50] calculated the AE counts at several different threshold values proportional to the maximum background amplitude (MBA) in order to determine the optimal threshold values. At the lowest speed and load, a threshold level at or above 30 percent of MBA is suggested. This study also emphasized the sensitivity of AE counts to the level and grade of lubricant within the bearing.

AE parameter analysis for tool condition monitoring has received attention mainly for real-time applications. AE count rate was introduced as a reliable parameter for monitoring tool wear during turning, although AE signals highly depend on process parameters [51, 52]. In [53], a direct relationship between tool wear during drilling operations and AE count was observed. Carpinteri et al. [54] applied AE technology to monitor concrete and masonry buildings. The cumulative number of counts was found to be informative for analyzing the evolution of cracks and determining the released strain energy. The authors observed that the maximum counting of AE corresponds to the maximum velocity of crack propagation.

The parameter of interest in the study is energy which is a 2-byte parameter derived from the integral of the rectified voltage signal over the duration of the AE hit [43]. This parameter covers the significance of count, duration and peak amplitude. It is to be noted that peak amplitude is a function of AE source scale [48]. The unit of energy is 10 micro-volt-seconds per count. The resolution of energy is 1 count ranging from 1 to 65535. It is evident that the energy is a discrete variable. For this, there may be a high frequency of zero energy in AE signals.

2.2 Prognostics

This section provides an overview of the current methods in prognostic modeling. A number of different definitions of prognostics have appeared in the literature [55-64]. The key phrases of the various definitions of prognostics are listed as follows:

- Estimation of time to failure,
- Estimation of risk of unacceptable behavior or accomplishing the mission,
- Estimation of remaining useful/operational life,
- Predicting future states or conditions,
- Predicting failure progression, and
- Predicting fault before it occurs.

Considering all the definitions, it is clear that prognostics require knowledge about the system under study, engineering models, technology for data collection and data processing, validated methodologies and standardized procedures.

ISO 13381 [55], provides general guidelines for a broader picture of prognostics. This standard believes that prognosis is an integral phase in the complete process of condition monitoring (CM). According to ISO 13381, prognosis is a case dependent process, and therefore, it is not reasonable to specify certain approaches or methodologies in prognosis standards. The standard covers the list of required data and the specific objectives for each group of data. Since prognosis is mainly based on data, determining the degree of certainty of the prognosis process, i.e., the confidence level, is an essential task. There is also an emphasis on the importance of failure modes data. Such data must include existing failure

modes, future failure modes, alarm limits and trip set point. In addition, symptoms of existing failure modes, i.e., influence factors and failure mode initiation criteria, need to be carefully considered. More importantly, the models of failure mode behavior should be used to provide more prognosis data. In addition, time to failure needs to be estimated using intuitive and/or empirical methods.

This standard divides a prognosis process into four phases: pre-processing, existing failure mode prognosis, future failure-mode prognosis, and post-action prognosis. Pre-processing deals with identifying and modeling the failure modes. The second and third phases are associated with the severity of existing and future failure modes as well as estimation of time to failure. The post-action phase mainly considers how to prevent the initiation of future failure modes.

Sheppard et al. [56] attempted to outline IEEE standards for prognostics and health management (PHM). The authors believed that standards might reduce cost and design time in PHM systems. In this study, the term Time to Failure was preferred over the popular term Remaining Useful Life (RUL). This is because the term RUL would be deceptive at the system level. The necessary elements of a PHM system were mentioned as follows: current state estimation, future state prediction and impact assessment, which should be followed by mitigating action. This study emphasized the significance of information management in PHM, which involves the need for more improvements in information exchange, information processing and combining information from multiple models.

There are extensive reviews of prognostics modeling in publications by , Jardine et al. [57], Lee et al. [58], Heng et al. [60], Farrar et al. [63, 64], Sikorska et al. [65], Si et al. [66], Hines and Usynin [67] and Dragomir et al. [68].

Jardine and his colleagues [57] provided an extensive review of machinery diagnostics and prognostics for condition-based maintenance up to 2005. The authors believed that prognostics approaches are even superior to diagnostics for maintenance decision support in a CBM program. To have effective prognostics models, this work underlined the importance of understanding failure mechanisms besides the knowledge on the fault propagation.

Lee et al. [58] provides an overview of the methods and information flow infrastructure that actualize the development of e-maintenance based on intelligent prognostics. Heng et al. [60] reviewed the prognostic of rotating machineries. Limitations and merits of current prognostics models have been indentified in this study. For efficient prognostics modeling, the authors suggested integrating condition monitoring and reliability data, utilizing incomplete measures of degradation, and considering failure interactions.

Sikorska et al. [65] discussed business issues that are central to appropriate model selection. This topic emerged because prognostic models are not similar in terms of satisfying the business needs and the level of risk they impose. On the other hand, businesses are also required to meet the requirements of prognostic models, e.g., data availability. The authors supported a staged approach in applying prognostics models. The first stage was defined as predicting the RUL along with the confidence limits for all the identified failure modes. From a practical point of view, the extent to which a system is broken down for RUL prediction is considerable. The subsequent stages need to consider the failure modes interactions and the effects of maintenance actions. This work classified the prognostics models based upon practical implementation issues of remaining life prediction.

Bond et al. [62] emphasized the significance of monitoring and managing materials degradation in nuclear power plants. In this way, prognostic models need to understand the phenomena of stressor-material interactions and quantify the rate of material degradation. The authors advocated the staged prognostics approach. In this way, the best prognostic model at the early stage of a component's service life would be a reliability model since there is no adequate degradation information. For active systems, monitoring of the stressor, e.g., temperature, fluid cavitations, mechanical vibration, is more beneficial than monitoring the subsequent effects of degradation and aging (i.e., degradation caused by numerous factors). However, this creates the need for physics-based models to relate the rate of aging or degradation to the stressors.

Hines and Usynin [67] reviewed various empirical methods for process and equipment prognostics. Bond and Meyer [69] emphasized the significance of online continuous monitoring to ensure optimal life management and operability of critical systems in nuclear power plant. The authors criticized the current mitigation policies in which mitigation

actions are being applied after discovering the degradations. The current online monitoring in nuclear power plants is limited to measuring environmental parameters and detecting the degradation in the physical aspects of active components, such as pumps and valves. This work highlighted the importance of monitoring and managing passive systems, such as pressure vessels and concretes, for the purpose of long-term operation (i.e., 80 years).

Greitzer et al. [70, 71] presented a high-level framework for prognostics and life extension to facilitate the communication of logistics requirements. These works focused on developing statistical methods for characterizing the degradation in mechanical systems with an emphasis on real-time and onboard prognostics.

2.2.1 Outputs and outcomes of prognostic models

In general, prognostic outputs are dependent upon historical data and/or the output of diagnostics units. Historical data consists of reliability data, maintenance data and the old condition monitoring data. To be useful for prognostics, a diagnosis system should be able to provide detectable short-term or long-term trends of degradation. However, one can expect diagnostic systems to provide posterior event analysis depending on the nature of failures. For example, Tan et al. [72] experimentally compared the diagnostic and prognostic capabilities of spectrometric oil analysis (SOM), acoustic emission (AE) and vibration analysis for monitoring of spur gears subject to natural pitting. It was concluded that the linear relationship between AE signals and the pitting rates would lead to better prognostics capability. The AE technique, in this case, offers much earlier diagnosis than vibration analysis. Even SOM has better diagnostics capabilities than vibration analysis at higher torque conditions.

The common prognostics systems outputs include RUL and the estimated time to failure (TTF). Furthermore, prognostic models might be used to predict the chance that a system operates without a major failure up to a specific time. However, in all cases, the output should be accompanied by confidence limits. Figure 2.3 shows different cases of TTF prediction. It is evident that more data would result in a more accurate model as shown in Figure 2.4. Here, the accuracy is defined as the closeness between the predicted values and the actual values. In principle, precision is a measure of the narrowness of the interval that covers the remaining life [68].

Moreover, to be effective, the outputs should be calculated for all the ongoing failure modes. In addition, censored data can be helpful in improving the accuracy of prognostic models as recommended in [59]. Successful implementation of prognostics promises reductions in life-cycle costs, downtime cost, number of failures, mean time to repair, maintenance labor costs and hazardous conditions. It is important to mention that appropriate action must be taken based upon the prognostics information. In practice, earlier detection of failures through prognostics gives more time for mitigation. This creates the need for prognostics-enabled scheduling of maintenance and spare part management.

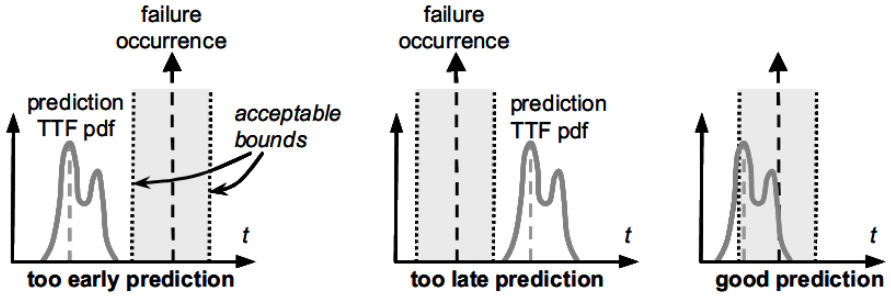

Figure 2-2 TTF predictions [68]

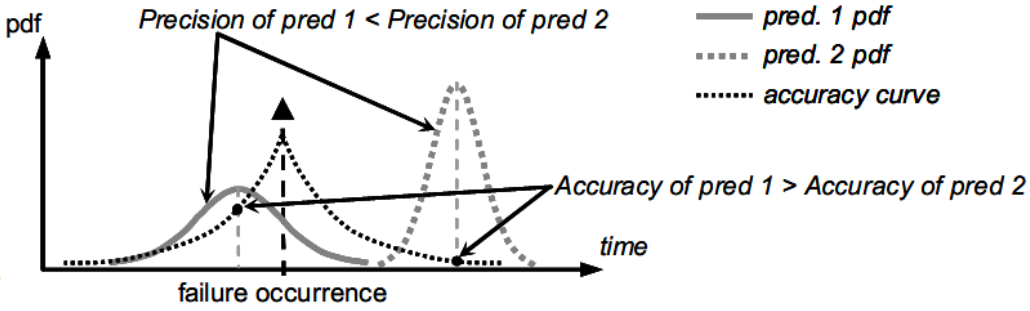

Figure 2-3 Accuracy of TTF prediction [68]

2.2.2 Prognostic approaches

This section classifies the existing prognostics approaches into model-based and data-driven approaches. This classification is based on the source of information or inputs, which is directly related to the cost and accuracy.

2.2.2.1 Model-based approaches

Such approaches require physical models, i.e., mathematical representation of the mechanistic knowledge, and theory of the degradation processes, such as crack propagation and spall growth. Model-based approaches are normally case dependent. If the physical models remain consistent across systems, they may provide the most accurate prognostics. However, in reality, the stochastic nature of systems increases the complexity of the physical models.

To name a few examples, Li et al. [73] introduced a defect propagation model via failure mechanism modeling for bearing prognostics. Qiu et al. [74] proposed a prognostic model for a bearing system by linking the damage mechanics to the parameters of vibration analysis. Damage mechanics were used to theoretically relate failure time to the stiffness variation. The authors assumed the bearing system was a single-degree-of-freedom vibratory system and related the vibration amplitude and natural frequency to the system stiffness. By using the vibration measurements, on-line prediction of bearing failure time can be calculated. Amiable and his coworkers [75] provided a comparison of the numerical models for estimating the lifetime in a thermal fatigue experiment. The studied models include closed-form, finite element solutions, boundary element solutions and fatigue criteria.

Ray and Tangirala [76] provided real-time damage rate of mechanical systems through a nonlinear stochastic model of fatigue crack dynamics. As mentioned earlier, the primary sources of AE in REB include the extension of cracks and contact between the rollers and the defected area, i.e., impact mechanism. Due to the fact that it is not possible to directly measure the bearing's defect area, Li et al. [13] developed an adaptive prognostic scheme to tune their deterministic defect propagation model - without interrupting the system - using

the estimation of defect size based on vibration and AE signatures. The following deterministic model was used in this study:

$$t_0 = \frac{C_0}{-n+1} D_0^{n+1} \quad \text{Eq. 2-3}$$

This model is based on Paris' formula where C_0 and n represent material properties, D_0 is the smallest detectable defect area and t_0 is the time when D_0 occurs. The authors successfully reduced the prediction errors by applying a non-linear recursive least square algorithm.

2.2.2.2 Data-driven approaches

These approaches are highly influenced by the quantity and quality of data. They are able to transform high-dimensional and noisy data into lower dimensional information. There are several types of data-driven approaches:

I. Traditional times to failure (reliability-based) approaches are also known as event data-based or experience-based prognostics. By categorizing the empirical models based on the type of information, the reliability-based approaches are introduced as Type I prognostics in some literature [67, 77, 78]. Such approaches use priori distribution of failure times, and therefore, it is assumed that operating conditions will remain the same in the future. Weibull analysis is the most common parametric reliability model because a it can handle increasing, decreasing and constant failure rates. The Weibull model with two parameters (β as the shape parameter and θ as characteristic life) is represented below:

$$\lambda(t) = \frac{\beta}{\theta}\left(\frac{t}{\theta}\right)^{\beta-1} \quad \text{Eq. 2-4}$$

The common practice in reliability-based approaches to estimate RUL is the use of Mean Residual Life (MRL) based upon parametric or nonparametric distributions. The MRL at time t with survival function S is defined as

$$MRL(t) = E(X - t \mid X > t) = \frac{\int_t^{\infty} S(u)du}{S(u)} \quad \text{Eq. 2-5}$$

Here, X corresponds to a non-negative random variable. For identical units, reliability approaches provide an estimate for the entire population, and therefore, one cannot expect accurate prediction for individual units. The application of experience-based approaches is limited to non-critical items with low failure rates or where it is not possible to obtain sensory data.

II. Statistical approaches can be categorized into two main groups. The first group uses direct CM data. This group include the following approaches:

a. Regression-based models are popular models that have no probabilistic orientation. They are mainly used for the trend extrapolation. Simple regression models do not provide probability density function of RUL, which is important for risk analysis and maintenance decision-making. In this category, the General Path Models (GPM) is a very popular non-linear empirical model. The model was developed based on the famous article by Lu and Meeker [79] in which the authors considered reliability prediction based on degradation measurements where the time-to-failure data is not sufficient to estimate failure distributions. Degradation measurements (i.e. prognostic parameters) show the degradation paths to the end of life and it can be assumed that there is a unique degradation path for each individual component as shown in Figure 2.5. The degradation path *y* for unit *i* at time *t* can be expressed as:

$$y_i = \eta(t, \varphi, \Theta_i) + \varepsilon \qquad \text{Eq. 2-6}$$

where ε shows the error, φ is a vector that represents fixed effects and Θ_i is a vector that represents individual effects for the i^{th} component [79].

End of life is usually indicated by passing over a pre-defined critical threshold. To estimate RUL of an individual component through GPM, the fitted model needs to be extrapolated to the failure threshold. It implies that the model parameters are constantly used for all RUL estimation regardless of the actual trend of degradation measures. Hence, it is possible to apply a Bayesian updating technique for a GPM model in order to consider the prior information. Bayesian updating is based on combining the prior information

and new observations to update the model parameter predictions. This would result in more accurate model fitting. Coble [78] applied Bayesian updating technique for GPM models.

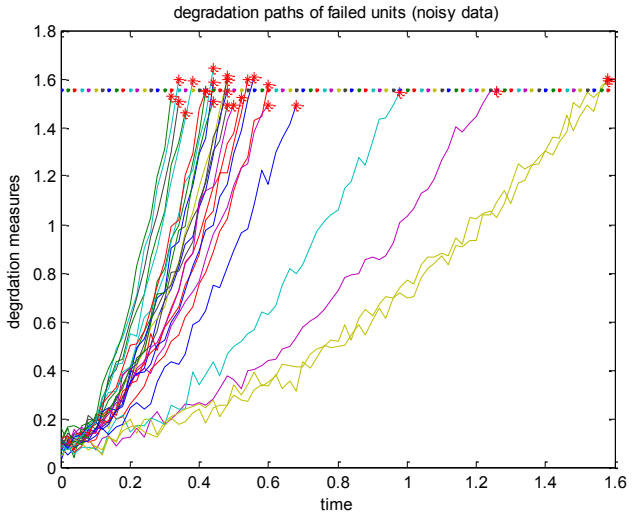

Figure 2-4 Degradation measurements

Another popular method in this group is the auto-regressive moving average (ARMA) which can be used in real-time applications to provide short-term predictions of RUL. Advanced ARMA can be used for non-stationary data and does not require historical failure data.

b. Brownian motion with drift (Weiner processes) is a useful technique where the degradation processes vary bi-directionally over time and the noise follows a normal distribution. Liao and Elsayed [80] applied a simple Brownian motion process to estimate the lifetime for the light intensity of scanners' LED lamps.

c. Gamma processes are used when a degradation trend takes place gradually in a series of tiny positive increments. Thus, the increment for each time interval follows a Gamma distribution. Henc, a Gamma variable is calculated as the sum of Gamma distributed increments. This technique has a

straightforward mathematical background, but it is limited to monotonic processes.

d. Markovian-based models that are normally used for discrete state systems.

The second group works with indirect CM data:

a. Stochastic models include aggregate reliability functions, and conditional probability methods such as RUL probability density functions, static Bayesian networks and dynamic Bayesian networks such as Kalman filters and particle filtering. Bayesian networks are useful when the prior knowledge is reliable. Static Bayesian networks can deal with incomplete data and can be used when multivariate training data is available. Dynamic Bayesian networks are helpful for modeling time-series data. This method is applicable when the multivariate posterior distribution is available.

b. Covariate-based hazard models: The factors that cause the degradation process are known as covariates. For instance, temperature is a covariate for a large number of mechanical and electrical components. The covariates normally have a stochastic nature and can be used to indicate lifetime. Covariate-based hazard models incorporate covariates in prognostic modeling. The Proportional Hazard (PH) model is a widely accepted semi-parametric model for stress-based prognostic where each operating condition is associated with a specific degradation level. The general form of a PH model is given by:

$$\lambda(t \mid \underline{z}(t)) = \lambda_0(t)\psi(\underline{z}(t); \underline{\beta}) \qquad \text{Eq. 2-7}$$

where $\lambda_0(t)$ is called baseline hazard rate, and $\psi(\underline{z}(t);\underline{\beta})$ is a function of covariates z with coefficient β. Both parametric and non-parametric forms of the baseline hazard rate can be utilized in practice. Weibull PH models are the most used parametric models. PHM should not be applied when covariates are not available or failure cannot be separated into individual failure modes. In addition, PHM is not a good method for non-stationary covariates.

Morkov Chain (MC) methods can be also used for stress-based prognostics. However, in MC processes, the transition probabilities do not vary over time and the future state does not depend on the past state, which may be considered as a drawback for some cases.

c. Hidden Markov models (HMM) and hidden semi-Markov Models (HSMM) do not have the limiting assumption of the exponential distribution.

III. Artificial Intelligence techniques include some non-linear, nonparametric techniques such as:

a. Artificial neural network (ANN) can be used to provide direct RUL prediction or to estimate some parameters for other models. ANNs are the best choice when other methods are not applicable. The development process of ANNs is time-consuming and requires large amount of data, although it is not recommended to apply ANN for complex data or when minimal data is available for training. Moreover, confidence limits cannot be provided. Gebraeel et al. [81] used ANNs to project the degradation signals to find the best exponential fit. Huang et al. [82] developed a life prediction model for ball bearings based on a self-organizing map and Multi-Layer Perceptron neural networks trained with a Back-Propagation algorithm. Herzog et al. [83] tried to enhance the predictive capability of ANNs in prognostics applications.

b. Expert systems and fuzzy systems are based on a set of rules to detect the similarity between the observations and the previously defined failures. Fuzzy methods can deal with noisy and incomplete data.

2.3 Life Extension

Power generation systems are an excellent example of capital-intensive engineering systems that are expected to operate safely and efficiently beyond their nominal design life. The Recent downturn in the world economy has restricted access to investment capital.

Moreover, return-on-investment (ROI) is a major concern with certain power generation systems including wind turbines. Accordingly, limited investment capitals and the need for ROI are the motivations for life extension in engineering systems. From investors point of view, the 2000 MW of installed wind energy capacity in 1990 still needs to remain operable [84]. On the other hand, useful design life is limited for highly stressed components and sub-systems. Table 2.2 shows typical reliability field data for two types of WTs inferred from a population of over 300 WTs with average age up to 15 years [85].

Table 2-1 Failure rate of 300kW & 1MW WT (Failure per turbine per year)

Assembly	300kW	1MW
Generator	0.059	0.126
Brake	0.029	0.056
Hydraulics	0.039	0.096
Yaw system	0.079	0.152
Sensors	0.037	0.151
Pitch system	0.034	0.237
Blade	0.078	0.308
Gearbox	0.079	0.255
Shaft/bearings	0.002	0.046

After investigating the reliability data of over 6,000 wind turbines of all sizes, Spinato and his co-workers concluded that smaller WTs are more reliable than larger WTs [86]. Not surprisingly, there has been substantial recent growth in utilizing large WTs. It is also intriguing to note that O&M costs contribute at least 10-15% of the total power generation income would increase over the 20 years of WT operating life [7]. To this end, considering all the addressed issues about WT reliability, the attempts to develop life-extending methodologies can be easily justified.

Reinertsen [87] provided a discussion on the published literature with respect to life extension practices. The author highlighted the lack of research on determining the residual

life of repairable systems. He concluded that there are no consistent and good methods for extending residual life up to the time of his publication.

In a series of research studies, Ray et al. attempted to connect the dynamics of material degradation with the current active control technologies to come up with a damage-mitigating control system (i.e., life extending control system) [76, 88-106]. The major goal of these studies is to achieve an optimized trade-off between structural durability and dynamic performance under the assumption that a small manipulation in the dynamic performance would result in significant improvement in the service life of critical components and prevent the mechanical structures from overstraining. Consequently, enhanced safety, productivity, availability and reliability are expected.

Ray and Lorenzo [89] provided an example about overheating of steam generator tubes when the control system reduces the feed water flow or increases the fuel flow. A small reduction in plant performance through damage-mitigating control would extend the creep-fatigue life of tubes. One integral part of a damage-mitigating control system is the damage prediction system, which applies the available sensory and operational information. However, it should be emphasized that an analytical model of the material properties (i.e., fatigue model) can be more useful in predicting the impending failures. For instance, sensors may not be able to provide valuable information for high cycle fatigue of a critical plant component.

To perform damage-mitigating control, the major challenge is to characterize the fatigue damage model and make it compatible with state-space representation of the controlled process (in a continuous-time setting). In effect, a fatigue damage model is based on experimental data. In such models, damage rate or other damage properties should be applied as inequality constraints. Ray and Newman [90] introduced a stochastic model of fatigue damage dynamics with an extended Kalman filter for estimating fatigue damage accumulation. In this model, the information on damage accumulation is a function of time and can be used for on-line life prediction.

To this end, it is evident that prognostic systems are still in a research and development phase. Reliable and physically meaningful prognostic systems require long-term and high-quality data streams and ongoing diagnostics. In this way, AE is a good candidate to be used

in PHM systems. This creates the need for more appropriate AE-based diagnostic techniques. The interest in long-term online monitoring creates the need for reliable long-term operation of monitoring systems as well as fast and precise prognostic approaches that can deal with multiple failure modes and novel events. The prognostics output is affected by degradation rate, novel events that change the rate, influencing factors (e.g., noise, type of model) and the sensor performance. Despite all the concerns in fruitful applications of prognostic models, it seems PHM has the potential to provide applicable life-extending methodologies and more appropriate maintenance decision making.

The next chapter provides examples of non-monotonic degradation. The experiments that resulted in non-monotonic degradations are discussed. It is followed by trend analysis non-monotonic degradation measures in Chapter 4.

3 EXAMPLES OF NON-MONOTONIC DEGRADATION

This chapter attempts to cover certain examples of non-monotonic degradation. One objective of this chapter is to provide a profound insight into the sources of non-monotonic trends. In this way, two degradation experiments along with the relevant diagnosis techniques are discussed. Generally speaking, little attention has been paid to the analysis of non-monotonic degradation.

Bae and Kvam [107] studied the degradation of light displays, such as plasma display panels (PDP) and vacuum fluorescent displays (VFD). The critical characteristic of light display quality is luminosity (or brightness). PDP and VFD have similar degradation characteristic mainly caused by manufacturing impurities. In particular, the burn-in characteristics are of interest when there is an unstable increase in emitted light before the continuous drop. Figure 3-1 illustrates the non-monotonic pattern of the VFD's luminosity degradation.

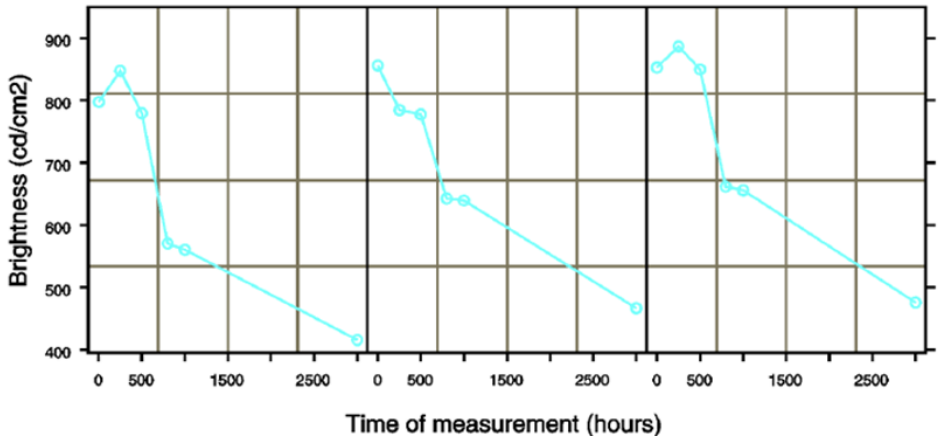

Figure 3-1 VFD's luminosity degradation [107]

Fouling is a common problem with heat exchangers which leads to degradation in the heat transfer efficiency [108]. Figure below shows the overall thermal resistance of a heat exchanger before and after cleaning (i.e. removing unwanted materials). Here, cleaning can

be considered as an imperfect maintenance that changes the degradation level as shown in Figure 3-2.

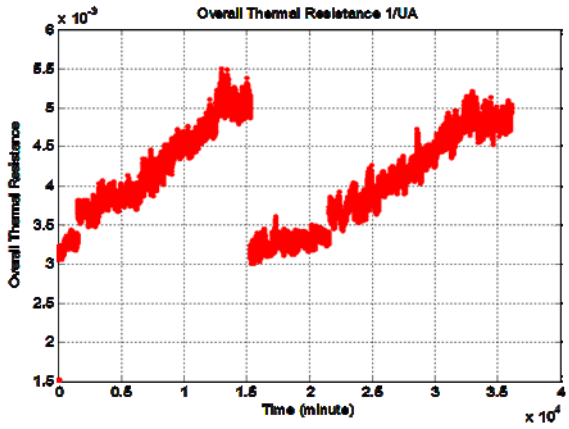

Figure 3-2 Degrdation in heat exchangers before and after cleaning

Gebraeel and Pan [109] proposed a stochastic model to handle time-varying operating conditions in real-time. In their experiments, load and speed were considered as the factors that influence the bearing's lifetime. Thus, the authors applied different load and speed settings. Figure 3-3 shows the degradation signal for two different rotational speeds. This plot clearly shows the reduction in the amplitude of the degradation signal as a result of the reduction in speed.

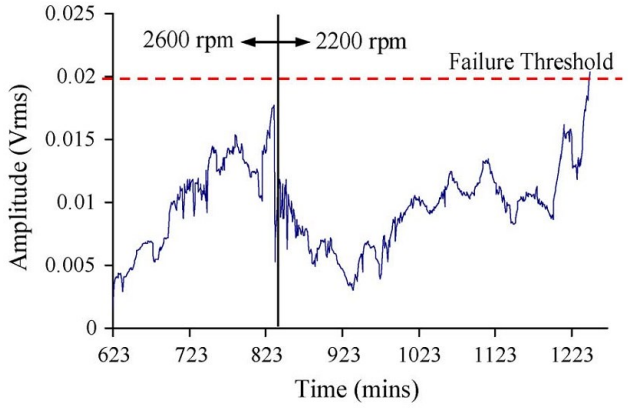

Figure 3-3 Speed change led to reduction in the amplitude of the degradation signal [109]

The remainder of this chapter provides discussion on the experiments that resulted in non-monotonic degradation patterns. The first experiment, presented in Section 3.1, deals with applying different types of unbalance in a rotary system. The second experiment, presented in Section 3.2, deals with the prospect of crack closure in materials subject to fatigue.

3.1 Unbalance in Rotary Systems

The lack of balance causes excessive vibration in rotary systems. Vibrations impose centrifugal force and oscillatory force. Unbalance occurs primarily due to uneven material deposit on the rotor about its rotating centerline. High vibration amplitude at the rotating speed is the primary indicator of this fault. Other sources of unbalance include:

- Imperfect manufacturing (e.g. eccentricity, clearance tolerances and cracks, loose parts in hollow places, crack, etc.)
- Operational changes (e.g. maintenance, corrosion and erosion, wear at the journal or bearing, thermal or gravitational distortion)

Bearings are a significant contributor to unbalance in rotary system because of internal clearance and run-out. In mechanical systems that utilize rolling element bearings, the rotational axis is not the same as the shaft axis. Even the locking mechanism (e.g. taper and setscrew) of bearing would affect the balance.

In effect, the existence of heavy spot in rotors leads to shaft bending and cyclical forces on bearings. Both centrifugal and oscillatory forces affect the bearing life. The former would be the cause of fatigue. The magnitude of centrifugal force can be obtained as

$$F_c = m\, r\, \omega^2 \qquad \text{Eq. 3-1}$$

where m represents mass, ω represents the angular speed and r is the radius from the center of rotation. It is important to note that centrifugal forces appear on stationary parts of the system. The vibration measurements for a point on the stationary parts show the sum of all oscillatory forces transmitted to that point.

There are various types of unbalance based upon the relation between the center of gravity (CG) and the heavy spot:

- Static unbalance happens when the CG and the heavy spot are in the same plane and the principal mass axis is parallel to the shaft axis.
- Quasi-Static Unbalance happens in systems such as motor-pulley where the shaft axis and the principal mass axis do not intersect on the CG.
- Couple unbalance occurs due to existence of two equal heavy spots which are 180 degrees apart (Figure 3-4). In this case, the shaft axis intersects the principal mass at the CG. Couple unbalance imposes radial force on bearings.
- Dynamic unbalance happens when there is no relation between the principal mass axis and the shaft axis. Dynamic unbalance is the combination of static and couple unbalance, and therefore, it is the most common type of unbalance.

Figure 3-4 Couple Unbalance

Balancing is a well-known operation in the category of filed service and repair. It is important to note that perfect balance cannot be achieved or even measured. Regardless of the above-mentioned unbalance types, two weights in two separate planes would correct the unbalance faults. In essence, while the objective is to reduce vibration, the balancing treats the cause rather than the symptom. Balancing may involve the followings: correction of mass distribution, creating centrifugal force, changing orientation of parts and adding/removing mass from non-rotating part [110]. These actions would provide mass symmetry, change the center of gravity, and affect bending moments. In practice, single-plane balancing contribute to about 60-70 of all balancing. The major assumption for single-plane balancing is the absence of couple unbalance.

The cyclical forces on bearing measured through vibration analysis are used for balancing. Since the sensors are normally placed on the stationary parts, only oscillatory

forces can be measured. Thus, there is no guarantee that the unbalance situation becomes better for the rotating parts. More importantly, balancing cannot treat other system faults that appear at the frequency of the shaft speed. In particular, in the case of couple unbalance, balancing would diminish the effects of unbalance on bearings but impose more bending load on the shaft. In general, the vibration induced by unbalance is affected by several factors including: mass, shaft speed, and stiffness of the supports. In essence, stiffer supports result in more force on bearings. It seems, however, further investigations are needed in order to elaborately diagnose unbalance faults in rotary systems.

In this research, several types of unbalance were investigated to provide a platform for the proposed diagnosis method and also to illustrate the possibility of non-monotonic degradation signals due to imperfect balancing.

3.1.1 Experimental set-up and procedure

A multi-purpose test rig was designed and developed to simulate the drive train components of wind turbines. Figure 3-5 shows the test rig. A WindMax 2 kW wind turbine generator was used in this setup. The test rig's motor provided a rotational speed in the range of 10- 1760 revolutions per minute (RPM). The control drive along with a LabVIEW-based program allowed the motor to provide variable speed according to a pre-defined speed pattern. Other main components of the test rig include a single phase planetary gearbox, support bearings, and two test bearings of different size.

The test bearing used was a SKF self-aligning ball bearing type 1205 ETN9 (Figure 3-6). The characteristics of the test bearing are as follows: internal (bore) diameter 25 mm, external diameter 52 mm, 26 rollers, and diameter of roller 7.4 mm. To make a faulty bearing, a groove with the width of 0.5 mm and the maximum height of 1.41 mm was made by using electrical discharge machine on the outer race of a bearing (Figure 3-6). Prior to the main tests, two bearings were run for 20 hours at the speed of 1700 RPM i.e. over 2 million cycles. These bearings will be called "used bearings" in the relevant analysis. It is important to note that all the bearings were used in dry condition. One reason for this was to ensure that the used bearings will not remain as good as new bearings after 2 million cycles.

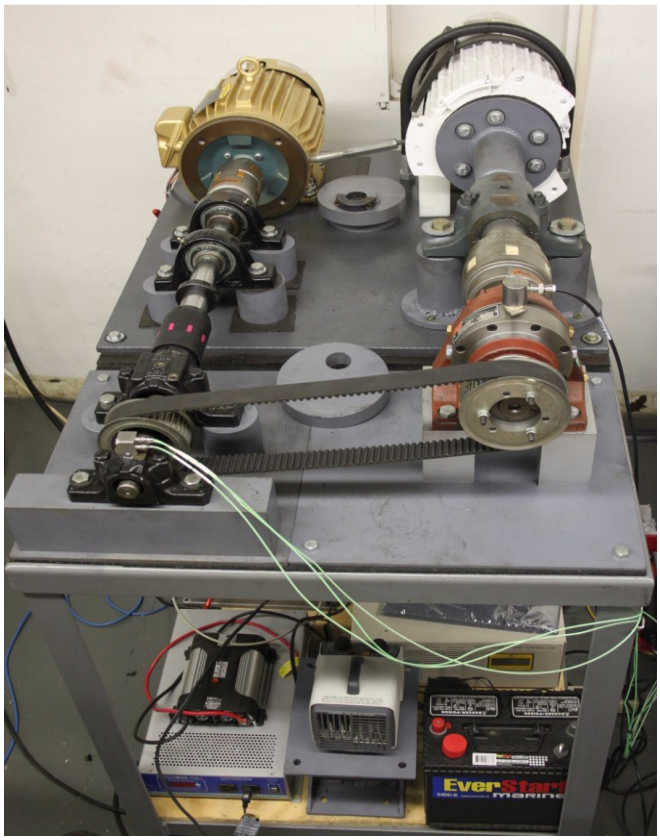

Figure 3-5 Test Rig

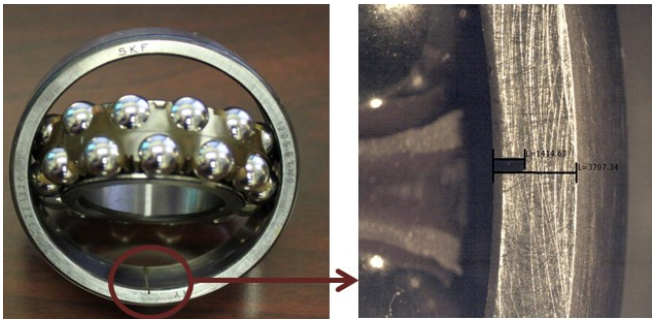

Figure 3-6 SKF 1205 ETN9 with a seeded fault

To impose the unbalance force, two uniform disks with the diameter of 6 inches and the thickness of 1 inch were used. To be able to attach the disks to the shaft and also change the angle of the disks, a shaft collar with a thickness of 0.5 inch was welded to each disk (Figure 3-7). An off-center hole with the diameter of 1.375 inch was then made for each disk as

shown in Figure 3-8. The mass of the disk with and without the hole were approximately 8 pounds and 7.6 pounds respectively. The new center of gravity moved off-center about 0.078 inches. The disks could be placed between the test bearings as shown in Figure 3-7. Consequently, several different sets of experiments could be performed by changing the number and orientation of the disks. In this study, the following loading configurations were used: no disk, one disk, two disks with an angle of zero, two disks with an angle of 120, and two disks with an angle of 180 namely couple unbalance. The magnitudes of the balancing masses for each configuration were found graphically.

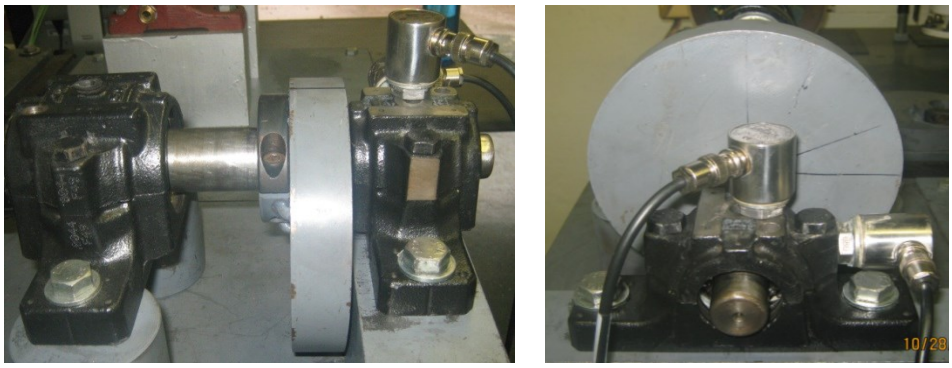

Figure 3-7 Sensor placement and the unbalance disk

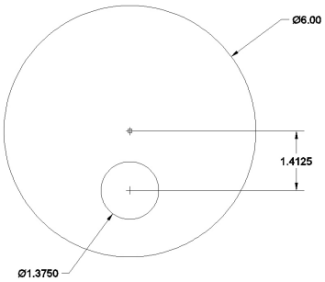

Figure 3-8 Disk with off-center hole (units in inches)

A 2-channel PCI-2 based AE system was utilized for data acquisition. The PAC R15I-AST acoustic sensor was used. This sensor has an operating frequency range of 80-200 kHz. The sensor housing contains a filter and an integral preamplifier of 40 dB. With an improved 18 bit analogue to digital conversion scheme, PCI-2 board provides a sampling rate of up to 40

MHz and a dynamic range of more than 85 dB. This data acquisition system is able to record up to 65535 counts per hit. AE signals were sampled at a high rate of 5 MHz. The upper and lower limit of filter was set to 10 KHz and 2 MHz respectively. The sensors were placed on the bearing housing as shown in Figure 3-7. Due to the interfaces that exist between the source and the side sensor, the sensor that was placed on the top of the housing provided more useful data. In Figures 3-9 the left graph shows the energy collected from the top sensor and the graph on the right shows the energy collected from the side sensor. As shown in this figure, data from the side sensor is not really helpful. By experimenting various position of the seeded fault, it was realized that placing the fault exactly below the top sensor provides more useful AE signals.

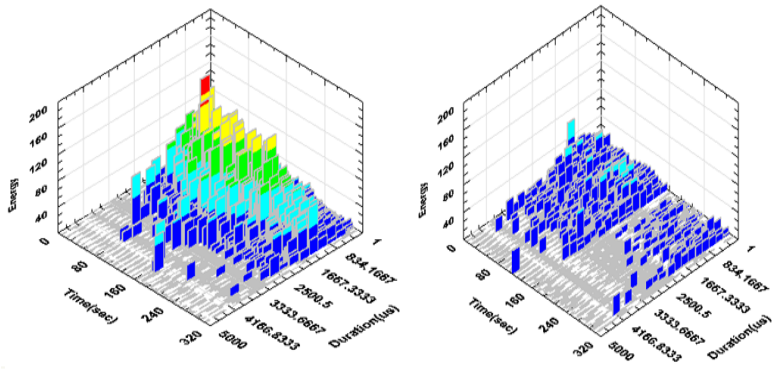

Figure 3-9 PAC-energy from two sensors (Left) the top sensor, (Right) the side sensor - Fault was under the top sensor

Moreover, the movement of the bearing inside the housing was negligible. The results of variable speed tests indicate the sensitivity of AE count to the shaft speed as shown in Figure 3-10 where the speed was gradually changed from 0 to 300 RPM and dropped back to 0.

In this study, three major variables were considered: (1) bearing type, (2) shaft speed, and (3) unbalance. Table 1 shows different categories of the variables. As shown in the table, three speed levels, three bearing types, and 5 loading patterns were designed for the experiments.

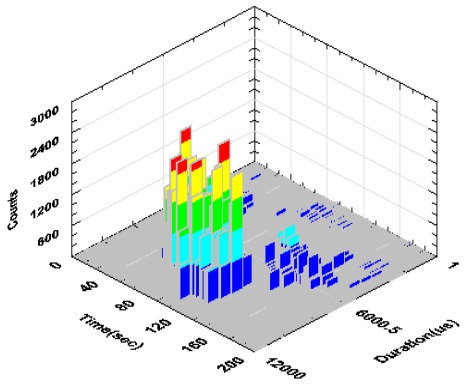

Figure 3-10 Dependence of AE parameters to the shaft speed

Table 3-1. Test Data Description

	Settings	Category
Unbalance	No Disk	1
	1 Disk	2
	2 Disk w 0°	3
	2 Disk w 120°	4
	4 Disk w 180°	5
Bearing	New	1
	Used	2
	Faulty	3
Shaft Speed (RPM)	150	1
	300	2
	450	3

3.1.2 Unbalance types and degradation trend

This section provides the graphical results for different cases of unbalance experiment monitored by AE technique. Keeping in mind that there is no perfect balancing, the objective of providing these figures is to show the need for more appropriate diagnostic techniques that could distinguish between various types of unbalance. Twelve bearings were used in these experiments. It is important to mention that the results unbalance experiments for some bearings are not similar to what presented in the following figures. The reasons for this dissimilarity include the moving structure of the test rig due to unbalance load and also using magnet mounting rather than stud mounting in the vibration monitoring. However, the

results for the unbalance tests prove the difference among various unbalance types for all the bearings tested. In addition, degradation of bearing in low speeds deserves more attention as sometimes unexpected degradation measures were collected in lower speeds. Figure 3-11 shows the difference among the unbalance types for the shaft speed of 300 RPM (there are four types of unbalance - categories 2-5 in Table 3-1).

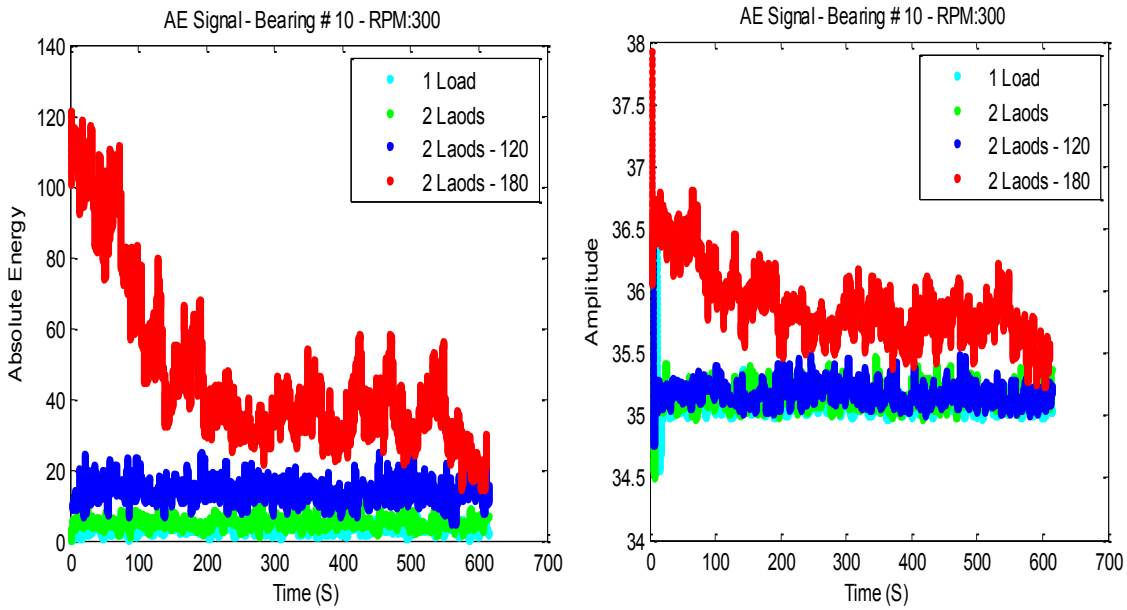

Figure 3-11 AE signals including absolute energy in Joules (left) and amplitude in dB (right) for various unbalance types with same speed (300 RPM)

Fogure 3-12 illustrates the effect of the shaft speed on the amplitude of AE signals in the case of couple unbalance. In this respect, the fluctuation of the AE amplitude is a subject that deserves more attention. As expected, the amplitude is higher for the shaft speed of 450 RPM. However, unbalance at lower speed (e.g. 150 RPM) may exhibit surprising results that mainly refer to the nature of REBs (i.e. the interactions beween the elements of bearings). Figure 3-13 shows the unexpected AE signals for two unbalance cases at 150 and 300 RPM. Finally, the cumulative plot for the energy of all the different types of unbalance and various shaft speeds are depicted in Figure 3-14. This plot clearly shows the change in trends of degradation when the load and shaft speed are changing.

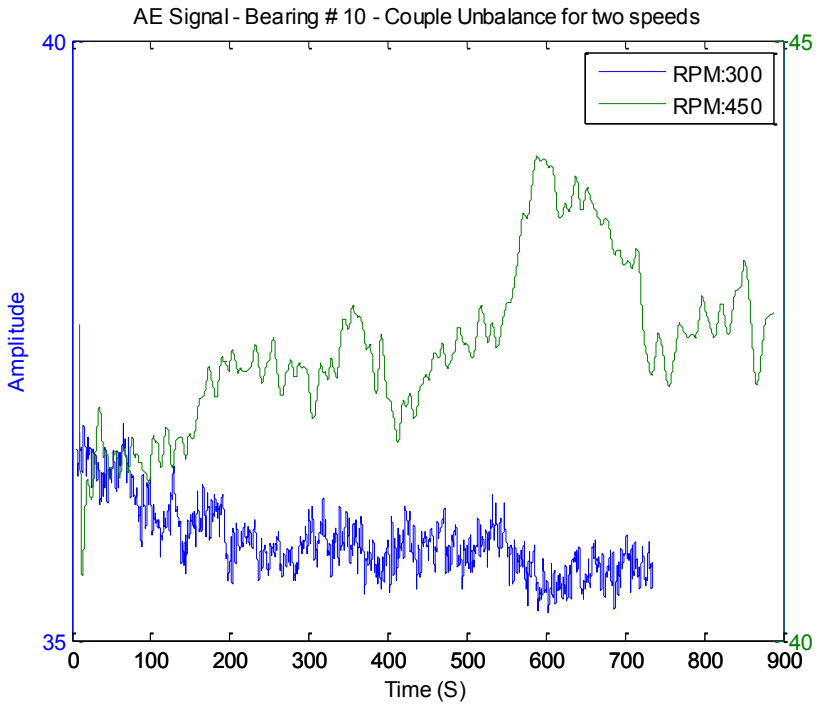

Figure 3-12 Amplitude (in dB) of the AE signal for couple unbalance at two speeds

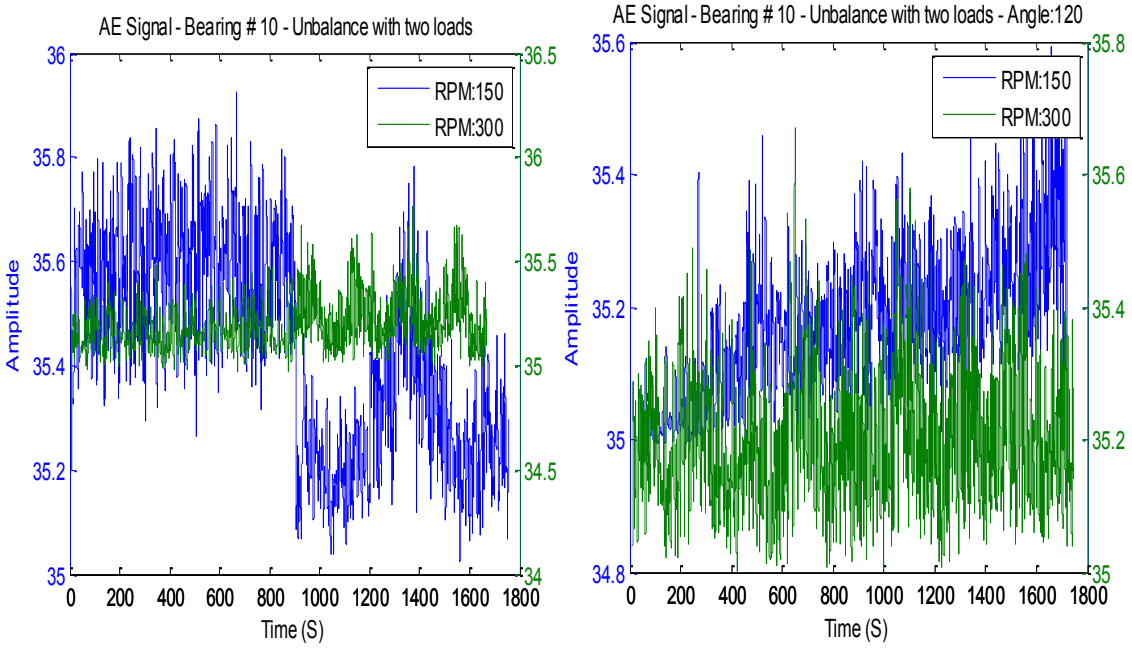

Figure 3-13 Surprising behavior of bearings in lower speed (Amplitude in dB)

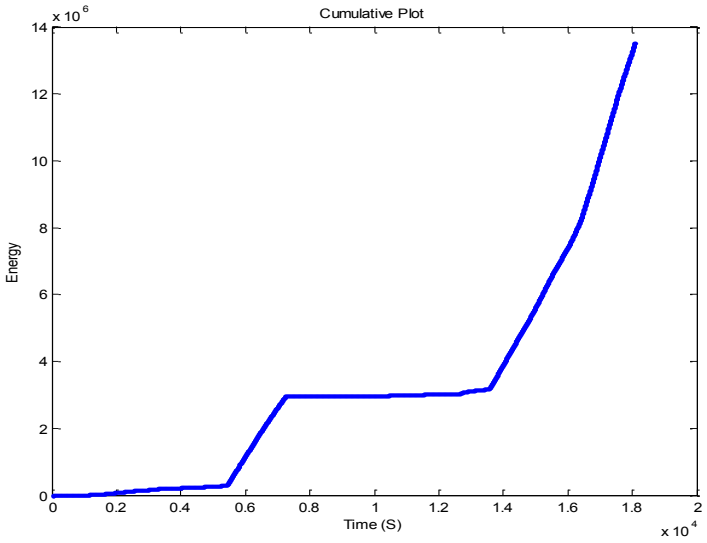

Figure 3-14 Cumulative plot for all different types of unbalance and various speeds (150, 300, and 450 RPM); the unit of energy is 10 μv-sec/count

3.1.3 Count data modeling and categorical data analysis

This section focuses on applying regression methods to model AE count data. In the literature, linear regression models have been primarily used to correlate the AE signatures and the physical features of interest. Traditional regression (non-Bayesian) methods have been widely used to model count data in both natural and social sciences. Considering the proven usefulness of AE counts for various AE applications, it seems an appropriate modeling of this parameter will lead to fruitful AE-based diagnosis.

In effect, the response variable in such models is a nonnegative integer. The most common regression-based count data model is Poisson generalized linear models (GLM) which is an extension of ordinary least squares regression and agrees with distributions from the exponential family. The Poisson GLM has the form

$$y \sim Poisson(\lambda)$$

The log link is specified by

$$\log \lambda = \beta_0 + \sum_i \beta_i x_i \qquad \text{Eq. 3-2}$$

The logistic model where p is the probability of success is given by

$$y \sim Binomial\ (p,n)$$

The logit link is given by

$$\text{logit}\ p = \log\frac{p}{1-p} = \beta_0 + \sum_i \beta_i x_i \qquad \text{Eq. 3-3}$$

The major drawback of Poisson regression is the restrictive assumption of equality between the variance and the mean i.e. equidispersion [111-113]. On the other hand, overdispersion occurs when the variance is greater than the mean. The overdispersion can be handled through negative binomial GLM. Generalized linear mixed models (GLMM) add an error term to GLM which is not part of our discussion.

Categorical data analysis based on GLM was performed for on AE data to provide a meaningful interpretation of the test variables' effect on AE features. The data set for the categorical data analysis consists of data from six bearings: one faulty bearing, two used bearings and three new bearings. The cumulative energy (CE) was selected as the response variable. The objective of this categorical analysis is to reveal the effects of explanatory variables (i.e. bearing, speed and unbalance) on the response variable. Due to the time limit, it was not possible to perform all the loading patterns for the speed level of 150 and 450 RPM. Although, this can be a drawback for the categorical analysis, the results are satisfactory.

The GENMOD procedure in SAS was used to conduct the statistical tests. This procedure fits a GLM to the data by maximum likelihood estimation. Through an iterative fitting process, the parameters of the model are numerically estimated in this procedure. Based on the asymptotic normality of maximum likelihood estimators, standard errors and p-values of the estimated parameters are computed. The parameter with the smallest p-value will be the most significant parameter in that category. The GENMOD procedure provides a number of probability distributions and link functions such as log-link function.

The categorical data analysis (CDA) shows that speed has the most significant effect among the explanatory variables. The bearing type has the least significant effect with the p-value of 0.6295 (Table 3-2). Figure 3-10 supports the SAS results that speed would positively affect the number of count and the energy of the signal.

Next, the interaction between explanatory variables (speed, unbalance and bearing) is added to the model. This 3-way interaction is significant (p-value = 0.1041). This model shows that speed and unbalance have significant effects on the CE level (Table 3-3).

Table 3-2 Categorical data analysis

Parameter	Estimate	Standard Error	Pr > Chi Sq.
intercept	0.1278	0.3054	0.6757
bearing	-0.0475	0.0985	0.6295
speed	0.0011	0.0007	0.1254
unbalance	0.0336	0.0493	0.4962

Table 3-3. Categorical data analysis with 3-way interaction

Parameter	Estimate	Standard Error	Pr > Chi Sq.
intercept	-0.6358	0.5584	0.2549
bearing	0.1680	0.1617	0.2987
speed	0.0022	0.0010	0.0229
unbalance	0.1717	0.0971	0.0769
Interaction	-0.0002	0.0002	0.1041

A new model was built based on the interactions between every two explanatory variables. The most significant interaction is the bearing-unbalance. In this case, unbalance type is the most significant variable. The effect of categories of each variable on the AE generation was then investigated. One category was selected as the baseline and the significance of the other categories were analyzed. In the case of shaft speed, 450 RPM was selected as the baseline. As expected, the results indicate that this speed level is more significant than the other two levels (i.e. 150 and 300 RPM). The bearing type 3 (i.e. faulty bearing) was the baseline to analyze bearing types. As shown in the Table 3-4, type 1 (new bearing) and 2 (used bearing) are not significant compared to the faulty bearing.

For the unbalance, type 5 was the baseline. As shown in the Table 3-5, type 1 and 2 are not significant at all compared to the couple unbalance i.e. type 5. Interestingly, the SAS

output results indicate that couple unbalance is the most significant type of unbalance in comparison with type 3 and 4. This result is similar to the graphical balancing calculation that shows couple unbalance needs the highest amount of balancing mass. Since it was realized that shaft speed has the most significant effect on CE, the next set of statistical tests were conducted with the speed level of 300 RPM. Having more data was the reason for selecting this speed level. The results of this data analysis indicated that unbalance is more significant in the absence of speed effect. Similar to the case with various speeds, adding the interactions of these variables lead to the results that all the inputs are significant.

Table 3-4 CDA - Bearing 3 is used as the baseline

Parameter	Type	Estimate	Standard Error	Pr > Chi Sq.
intercept		0.1088	0.3034	0.7199
bearing	1	0.0101	0.199	0.9593
bearing	2	-0.2117	0.2195	0.3347
speed		0.0011	0.0007	0.1256
unbalance		0.0338	0.0493	0.4928

Table 3-5 CDA – Unbalance type 5 is the baseline

Parameter	Type	Estimate	Standard Error	Pr > Chi Sq.
intercept		0.2736	0.4105	0.505
bearing		-0.0421	0.0986	0.6696
speed		0.0009	0.0009	0.3432
unbalance	1	0.0808	0.214	0.7058
unbalance	2	-0.2115	0.2814	0.4521
unbalance	3	-0.4037	0.2917	0.1664
unbalance	4	0.3487	0.251	0.1648

In practice, depending on the preset reference threshold value, there may be a high frequency of zero counts in the AE signals, which indicates excess zeroes. In this case, a simple Poisson regression would not satisfactorily fit the data. Zero-inflated Poisson (ZIP)

model, introduced by Lambert [112], fittingly handles overdispersion and excess zeroes. Principally, ZIP model considers the data set as a mixture of a process that generates only zeroes and a process that generates counts from a Poisson or a negative binomial model. In other words, ZIP models calculate the probability (p) of having observations of 0. Therefore, *1-p* would be the probability of having non-negative integers. Therefore, for a vector of responses $\boldsymbol{Y} = (Y_1, \dots, Y_N)^T$, the count response can be written as

$$Y_i \sim 0 \qquad \text{with probability } p_i\,, \qquad 1 \le i \le N$$
$$Y_i \sim Poisson(\lambda_i) \quad \text{with probability } 1 - p_i \text{ ; and } \lambda_i \ge 1$$

Thus,

$$Y_i = 0 \quad \text{with probability } p_i + (1 - p_i)e^{-\lambda_i} \qquad \text{Eq. 3-4}$$
$$Y_i = k \quad \text{with probability } (1 - p_i)e^{-\lambda_i}(\lambda_i)^k/k! \qquad \text{Eq. 3-5}$$

Consider the following regression models for

$$\lambda = [\lambda_1, \lambda_2, \dots, \lambda_n]' \text{ and } \mathrm{p} = [p_1, p_2, \dots, p_n]' :$$

$$\ln(\lambda) = \mathbb{B}\beta, \qquad logit(\mathrm{p}) = \ln\left(\frac{\mathrm{p}}{1-\mathrm{p}}\right) = \mathbb{G}\gamma \qquad \text{Eq. 3-6}$$

where $\mathbb{B}$ and $\mathbb{G}$ are covariate matrices (design matrices), β and γ are regression coefficients. Accordingly, the log-likelihood function is

$$L(\beta, \gamma|\mathrm{y}) = \sum_{y_i=0}^{n} \ln(e^{\mathbb{G}_i\gamma} + \exp\left(-\left(e^{\mathbb{B}_i\beta}\right)\right) - \sum_{i=1}^{n} \ln(1 + e^{\mathbb{G}_i\gamma})$$
$$+ \sum_{y_i>0}^{n}(y_i\, \mathbb{B}_i\beta - e^{\mathbb{B}_i\beta}) - \ln(y_i!) \qquad \text{Eq. 3-7}$$

By using $\ln(\lambda) = \mathbb{B}\beta$ and $logit(\mathrm{p}) = \ln\left(\frac{\mathrm{p}}{1-\mathrm{p}}\right) = \mathbb{G}\gamma$ we conclude

$$\lambda_i = \exp(\mathbb{B}_i\beta), \text{ and } p_i = \frac{\exp(\mathbb{G}_i\gamma)}{1+\exp(\mathbb{G}_i\gamma)}$$

Hence, using the delta method, the variances of $\hat{p}$ and $\hat{\lambda}$ are

$$cov(\hat{\lambda}_i, \hat{p}_i) = [[\frac{\partial\lambda_i}{\partial\beta}]'[\frac{\partial p_i}{\partial\gamma}]']cov(\hat{\gamma}, \hat{\beta})[[\frac{\partial\lambda_i}{\partial\beta}]'\,[\frac{\partial p_i}{\partial\gamma}]']'$$

$$\hat{\sigma}_{0(\hat{p}_i)} = \sqrt{var(\hat{p}_i)},\ \hat{\sigma}_{0(\hat{\lambda}_i)} = \sqrt{var(\hat{\lambda}_i)}$$

The independent variables in this study are the shaft speed and unbalance. The amount of mass required for balancing was used as the unbalance values in the ZIP model. The graphical balancing techniques were utilized to calculate the amount of the balancing mass. The procedure of balancing considers both the balance of forces and couples. It is to be noted that couple unbalance (type 5) needs the highest amount of mass for balancing.

The COUNTREG procedure in SAS was used to develop the ZIP models. This procedure performs nonlinear optimization. Two iterative minimization method were applied; (1) the quasi-Newton method and (2) the Newton-Raphson method. The ZIP models provide the probability of obtaining zeroes (p) and the parameter for the Poisson model (λ). The probability of obtaining zero for the faulty bearing is almost zero for all the tests. This implies the existence of strong burst signals in all the AE hits. To provide a fair comparison, Figure 3-15 shows the λ for different unbalance types at the shaft speed of 300 RPM.

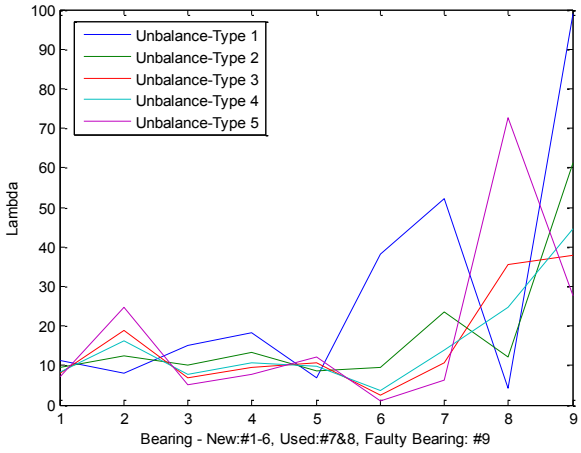

Figure 3-15. Lambda from the ZIP model for different unbalance type at 300 RPM

In this graph, bearing 9 represents the faulty bearing; bearing 7 and 8 represent the used bearings. It is clear that the highest values of λ belong to either used or the faulty bearings. It implies that these bearing generate stronger burst signals in comparison with all the new bearings used in this study. It is interesting to note that the highest λ (= 429) was obtained for the couple unbalance of the faulty bearing.

Control charts can be used to detect anomalies in the trend of the above-mentioned parameters (i.e. p and λ). In essence, control charts can be a supportive tool for fault detection and diagnosis particularly when an accurate process model does not exist [114]. For instance, Leger and his coworkers [115] successfully linked Cumulative Sum (CUSUM) control charts and radial basis function neural network as a fault detection and diagnosis strategy. The CUSUM control chart is responsible for fault detection and generation of the fault signature pattern, which in this case is the current process means. The neural network performs the pattern recognition.

The statistical changes (i.e. variability) fall into two categories: (1) natural variations which deals with regular and inherent variations, (2) unnatural variations [116]. The application of control charts in fault detection is based on the assumption that charts display the special causes of unnatural variations through a number of detectable patterns. General format of control charts consist of a centerline, which represents the expected value of the plotted data, the upper control limit and the lower control limit indicating the limits of variation. A process is in statistical control if the plotted statistic lay within the control limits. Control limits are generally set with respect to the standard deviations of the plotted statistic. The collected data is usually formed in time-ordered subgroups. Then, the values of interest need to be extracted from each subgroup. The minimum of 25 subgroups are recommended to initiate univariate control charts [116]. It is to be noted that data that comes from non-homogenous processes is a drawback for successful implementation of control charts. Traditional Shewhart charts can be divided into variables charts and attribute charts. The former deal with continuous data and the latter is associated with discrete data. CUSUM, first introduced in 1954, and moving average charts are special kinds of variables charts. In case of discrete data, c and u charts are suitable for Poisson-based data. The most common chart for data with binomial distributions are the np and p charts.

In this research, CUSUM chart was applied to diagnose the anomalous changes in the parameters of the ZIP model [117-119]. This chart detects the deviation of the process mean through cumulative sums of the shift between sample averages from a target value. CUSUM charts are sensitive to small and moderate changes in the process mean. Such changes, e.g. one-sigma shift in the mean, are hardly detectable by Shewhart-charts. The run chart of

CUSUM displays the successive differences between the sample average and the target i.e. process mean. Recent schemes of CUSUM takes the advantage of likelihood ratio that comes from the probability distribution of sample statistics. In addition, multivariate CUSUM (MCUSUM), and Hotelling's T^2 are the most useful types of multivariate control charts because of sensitivity to small and moderates variations. Moreover, nonparametric control charts have been recently developed to provide distribution-free control charts. Nonparametric control charts works better than their parametric counterparts in case of heavy-tailed distributions or where the true distribution is quite different from normality [120-122]. Nonparametric univariate control charts include Shewhart-type, and CUSUM-type control charts using statistical schemes such as ranks, signed ranks, and hodges-lehman statistics.

The two-sided tabular CUSUM for n samples chart is defined by

$$S^{\pm}(n)\begin{cases} 0, & if\ n = 0 \\ \max\{0, S^{\pm}(n-1) \pm (x_n - \hat{\mu}_0) - k\hat{\sigma}_x\}, & if\ \ n = 1,2,\dots \end{cases} \quad \text{Eq. 3-8}$$

where $x_n = \lambda_n$ (or p_n) is the nth observation, $\hat{\mu}_0$ is the target (i.e estimate of in-control mean) and $\hat{\sigma}_x$ represents the known or estimated standard deviation of the sample mean. Here, $S^{+}(n)$ and $S^{-}(n)$ give the accumulation on high side and low side respectively and k is the threshold for accumulation which is also called allowable slack. This parameter is the minimum difference between the target and sample average. The V-Mask is the usual visual procedure used to determine an out-of-control process as shown in Figure 3-16. Thus, we have the decision interval as H = $h\sigma_x$ = $\tan(\theta)\ d\ \sigma_x$ as and h is called the decision parameter.

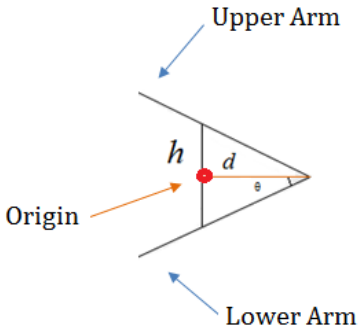

Figure 3-16 Sample V-mask for a Two-Sided CUSUM Chart

The origin point of the V-Mask is the most recent plotted point (i.e. cumulative sum). The origin will be relocated by collecting new data. The V-mask arms slide backward. The process is out-of-control if one or more of the previously plotted points crosses the arms of the V-mask. If the points cross the lower arm it is possible to conclude that the process mean is increasing. Figure 3-17 show the two-sided CUSUM charts of λ and p for the case when only the new bearings were used. Here, as none of the plotted points cross the arms of the V-mask we conclude that the process is in control.

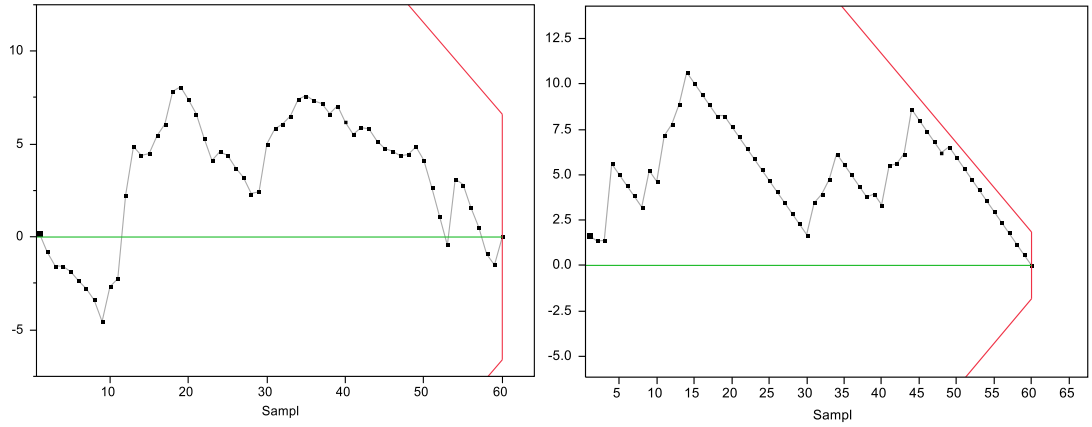

Figure 3-17. (Left) CUSUM charts of λ for new bearings – (Right) CUSUM charts of p for new bearings

Figure 3-18 show the CUSUM charts of λ and p for the case when used bearings were also taken into account. The chart for λ clearly shows that the bearings experienced certain deviation and they are not as good as new bearings. It is important to note that the points in the CUSUM charts of λ crossed the lower arm. It implies the increase in the mean of the energy in the used bearing comparing to the new bearing. Figure 3-19 show the CUSUM charts of λ and p for the case when all types of bearings were taken into account. These charts clearly depict the out of control condition since faulty bearing was taken into account. Interestingly, the points in the CUSUM charts of λ crossed the lower arm which means the increase in the mean of the energy due to existence of the fault in the system.

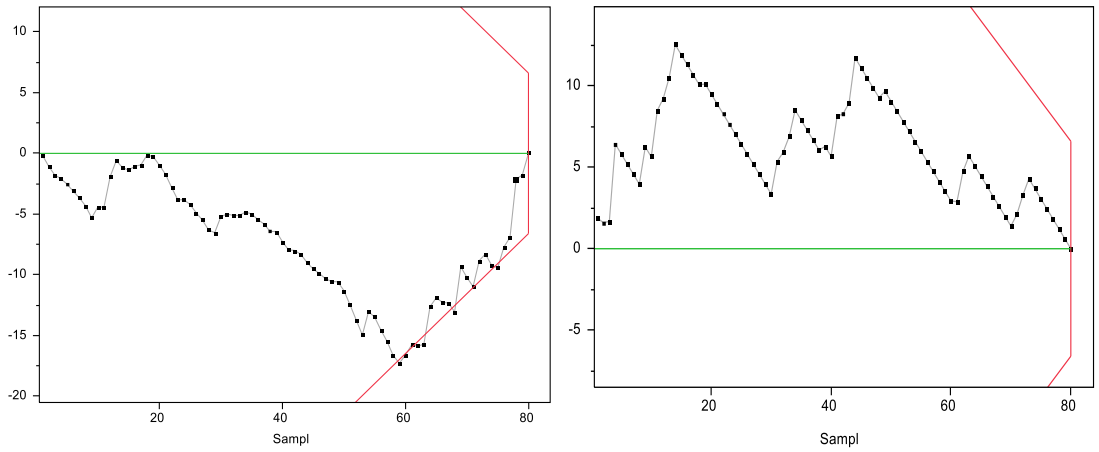

Figure 3-18. (Left) CUSUM charts of λ for new and used bearings – (Right) CUSUM charts of *p* for new and used bearings

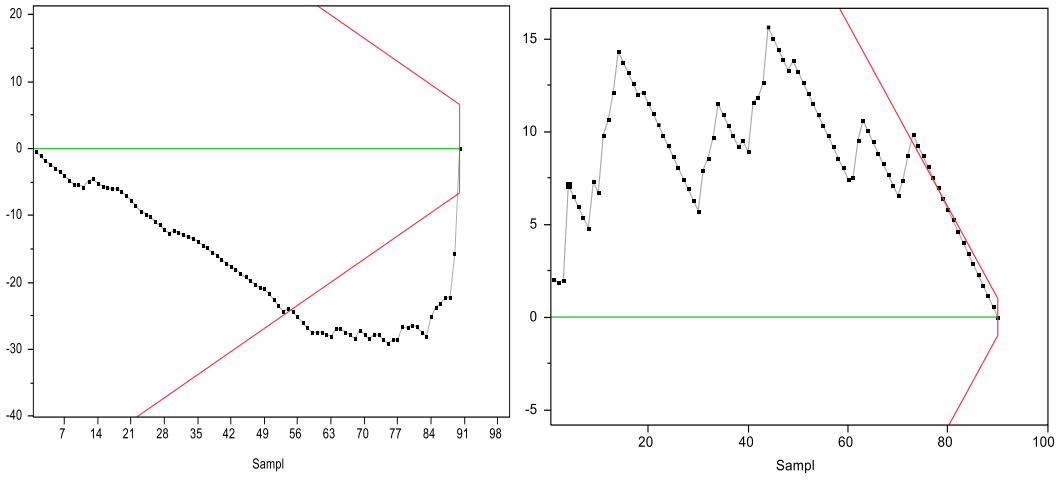

Figure 3-19. (Left) CUSUM charts of λ for all bearings - (Right) CUSUM charts of *p* for all bearings

3.2 Fatigue Crack Closure

Self-healing material has become a favorite topic for investigation in material science and engineering. Self-healing is defined as the ability of a material to repair damage [123, 124]. Self-healing can be classified into two groups: autonomous and non-autonomous. The latter requires some form of external intervention. For instance, certain healing processes require heating to activate the repair process. In essence, an effective healing process restores thermodynamic force by unsettling the thermodynamic equilibrium.

Many recent studies in this area have focused on nonmetallic materials i.e. polymers, concretes and ceramics. Considerable success has been reported for polymers because of their large rates of diffusion and plasticity [123]. In ceramic materials small cracks can be repaired via oxidative reactions. Despite the countless practical prominence, less attention has been paid to self-healing alloys and composites mainly due to intrinsic features such as strong bonds and low diffusion rates.

The study of natural materials and tissues has become an important aspect of self-healing materials. In this respect, the new fields of biomimetic and bio-inspired materials have emerged. Biomimetics intends to study and mimic biological systems for engineering applications such as self-healing, self-lubricating, and self-cleaning. Most biological systems can implement partial or complete self-repair. The ability of self-repair in living organisms implies the existence of a complex hierarchical structures which is very difficult to be imitated in artificial self-healing materials. However, self-healing materials require mechanisms such as artificial vascular system or healing agent to partially heal damages such as voids and cracks. A healing agent is designed to fill the cavities in material through phase transition or chemical reaction. Thermosetting polymers such as epoxy can be used as healing agent. Epoxy can be stored in the matrix along with a catalyst in the form of macro-capsules and act when the cavity is growing.

There are several strategies to outfit the metallic engineering materials with self-healing properties. Figure 3-20 shows the most common strategies:

- Encapsulation of a healing agent which can be an alloy with a low-melting point
- Damage prevention by diffusion of the atoms and forming precipitates to halt the growth of crack or void

- Reinforcement of an alloy matrix with embedding shape-memory alloys

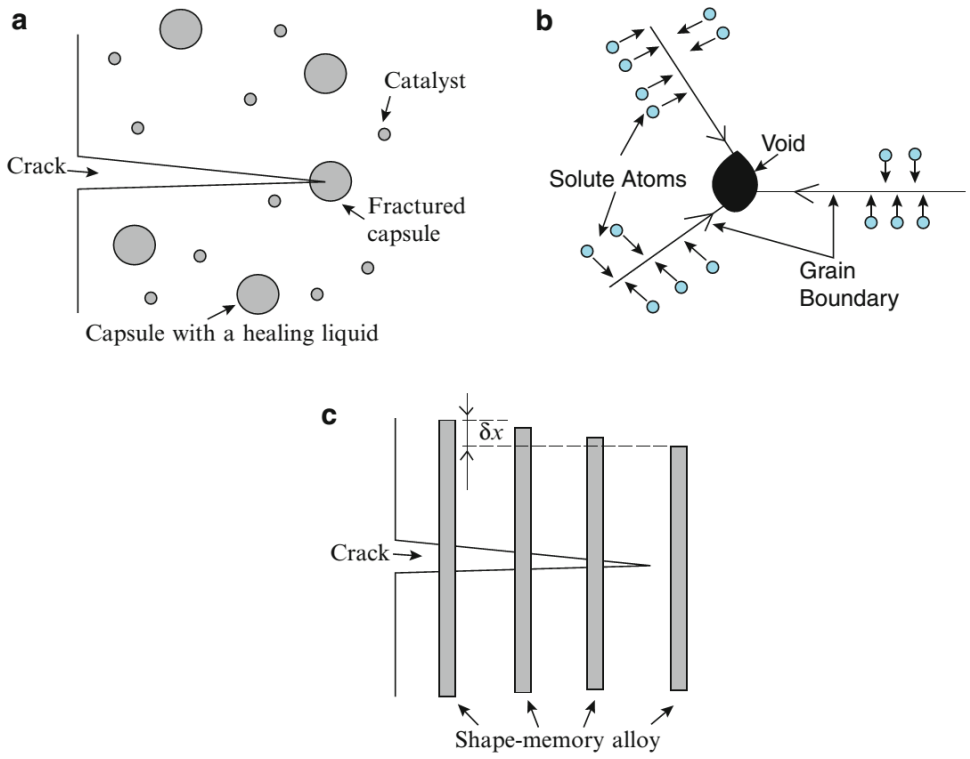

Figure 3-20 Three types of self-healing in metals; (a) healing agent, (b) precipitation in an over-saturated alloy, (c) shape-memory alloy micro-wire [123]

Fatigue crack closure (FCC) is the self-healing mechanism that was considered and experimented in this research. FCC, first pioneered by Elber [125], is an important phenomenon in evaluating the effective driving force for crack growth. Elber identified that plastically deformed material left in the wake of a propagating crack would result in partial or complete crack closure for a portion of the applied loading range [125]. It implies that the nature of the crack-face behind the crack tip is a major contributor to the crack-growth rate. The initial discovery of FCC was pursued by attempts for further elaborations using numerical and analytical methods such as finite element analysis and modern tomography methods [126]. However, there are few skeptical studies that advocate reviewing the role of FCC in the modeling of fatigue crack growth [127, 128].

FCC is not similar in different materials. It is still a challenging topic in predicting fatigue lives. There are several contributing factors to the phenomenon of FCC including: loading pattern [129], micro-structure of material, environment, geometry, crack length [130] and so on [129, 131]. For instance, Newman and Elber [132] realized that because of decrease in strain amplitude at the crack tip, shafts would show signs of lower crack growth rate in salt water. Furthermore, Liu and Wu [131] discovered different fatigue closure behavior in different geometries. They also emphasized the significant effects of stress ratio on FCC. In practice, these contributors would harness the existence, intensity and the dominant mechanism of crack closure. The crack surfaces behind the crack tip can contact due to a variety of mechanisms including: plasticity-induced crack closure (PICC), oxide-induced crack closure (OICC), roughness-induced crack closure, and fluid-induced crack closure [133]. PICC, introduced by Elber, is the dominant factor in the most FCC cases. To successfully investigate this mechanism, the size of plastically deformed area should be estimated. OICC happens when crack faces meet as a result of oxide debris or other corrosion-based reasons. In practice, it is possible to verify the dominant FCC mechanism by using experimental evidence along with analytical or numerical models.

Precise assessment and quantification of FCC would clarify ambiguities on materials behavior in terms of crack propagation. In this way, in situ investigation of the crack-tip deformation, plastic deformation near a fatigue-crack tip, and FCC micro-mechanisms are of primary importance. Equally important is the techniques to calculate and validate the effective-stress intensity-factor range (ΔK_{eff}) based on the below relationships

$$Fully\ open\ crack: K > K_{cl} \qquad \Delta K_{eff} = K_{max} - K_{cl}$$

The prerequisite for the determination of ΔKeff is the precise measurement of the crack closure load. Inconsistent results in load measurements are influenced by the measurement location and the data interpretation method.

The techniques to quantify FCC can be divided into three groups: experimental, numerical and analytical models [134]. In essence, analytical models provide quick approximation. On other hand, experimental and numerical models may provide better results. However, they are typically time-consuming and costly. Experimental methods

include [135-138]: crack opening displacement (COD) gauge, strain gauge, interferometry, nonlinear ultrasonic, direct current potential drop, synchrotron X-ray micro-tomography, neutron diffraction, polychromatic X-ray micro-diffraction and acoustic emission (AE). The ability of AE to reliably measure the crack closure level was appreciated in several works [135, 139]. Chang and his coworkers [139] believe that AE signals do not depend on the length of crack. This would be one of the significant advantages of AE signals in measuring FCC. In this research, AE was used to study fatigue crack growth behavior subjected to a single tensile load. It is to be noted that FCC elucidates the crack growth retardation subject to high overload. The single tensile overload is of interest since higher stress level can cause the contact of crack surface behind the crack tip. In essence, cycling loading (with periodic overload) causes more crack closure and consequently it may reduce the rate of fatigue fracture.

A set of experiments was conducted to study FCC subject to a tensile load using AE. The experiments intend to analyze the growth of crack for single-edge-notched samples of HAYNES® HR-120™ according to ASTM E647 (Figure 3-21 & 3-22). This material is a heat-resistant alloy that provides excellent strength at elevated temperature. The length of crack for this test is 20 millimeters. When the crack reached the middle point of the intended crack length (i.e. 10 millimeters) the test should be stopped and a single tensile load (STL) needs to be applied. The magnitude of the STL is equal to 150% of the maximum load applied for crack propagation.

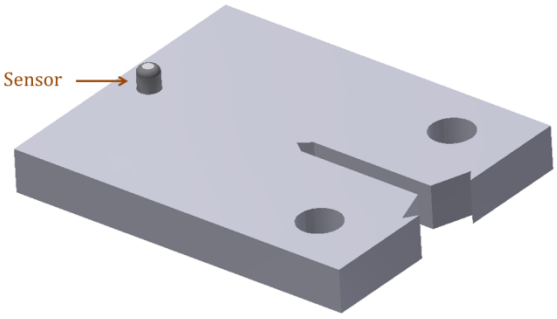

Figure 3-21 Sample with wire cut perpendicular to the loading direction and the sensor location

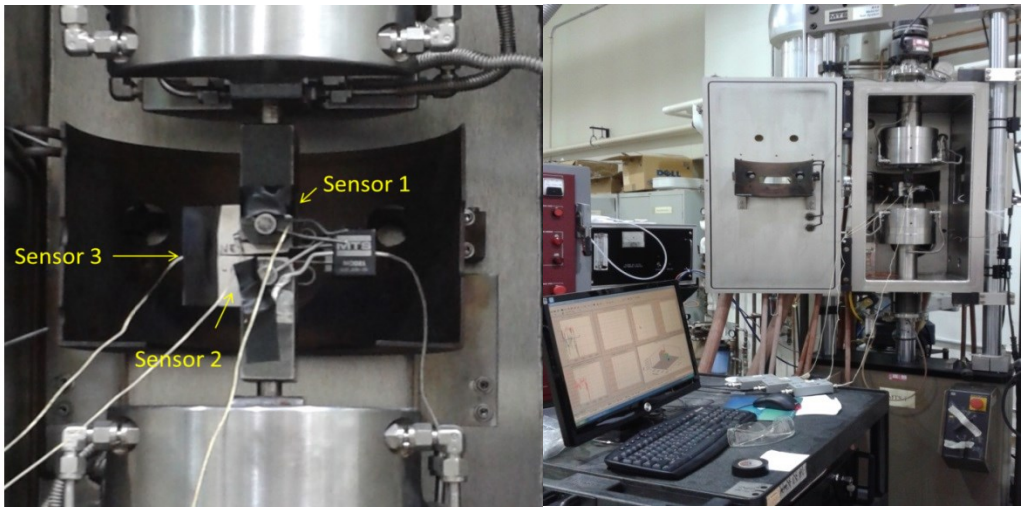

Figure 3-22 FCC Experiment

The experiments start with 2-3 mm pre-crack to provide the early stages of crack growth for appropriate settings. Fatigue tests were carried out in ambient laboratory air using a servo-hydraulic material testing systems. AE measurements were performed using AE equipment (Digital Signal Process (DiSP) system) manufactured by the Physical Acoustic Corporation. The AE signal was amplified by preamplifiers (40 dB) and passed through the band-pass filter from 20 kHz to 2 MHz. The threshold set as 40 dB so the signals exceeding this level were recorded and used for further analysis. Table 3-6 provides the duration of the tensile loads and the number of cycles to reach the pre-determined length of crack (i.e. 20 millimeters)

Table 3-6 Details of FCC experiments

	Sample 1	Sample 2	Sample 3	Sample 4	Sample 5
Last cycle	223860	238270	225481	239565	211502
Duration of the tensile load (sec.)	0	30	300	450	300

Figure 3-23 shows the crack length versus cycle for all the samples. Sample 1 is clearly different from other samples that experienced STL and consequently FCC. Figure 3-24 shows

the rate of crack growth versus the effective stress intensity factor range. The period of retardation is obvious in this graph.

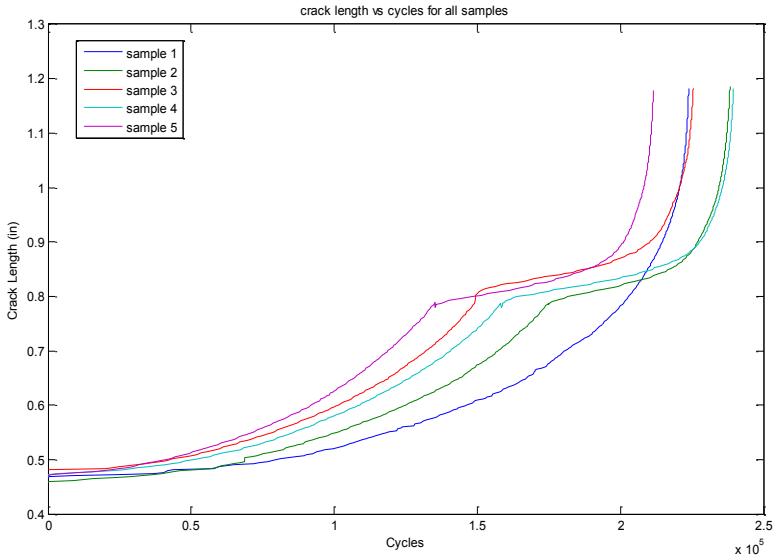

Figure 3-23 Crack length vs. Cycle

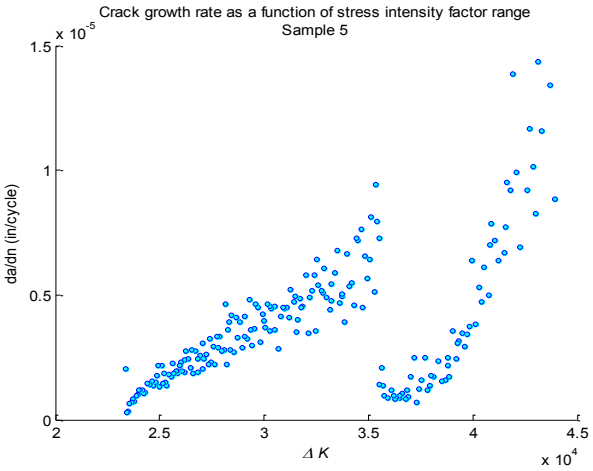

Figure 3-24 Effective fatigue crack growth curve (ΔKeff with unit of stress x length$^{0.5}$)

Figure 3-25 provides the AE features for a sample that experienced STL. The red and green points in these plots are related to the start and the end of STL respectively. The period of retardation can be clearly seen right after applying the STL. Roughly speaking, the majority

of research works on fatigue life analysis are based on assumptions that exclude FCC from consideration. In essence, FCC is still a topic for basic research and no significant and profitable application has been identified for this phenomenon due to two reasons: (1) complexity of numerical crack growth calculations, (2) understanding exact micro-mechanism of FCC is still a challenging research topic mainly due to experimental difficulties.

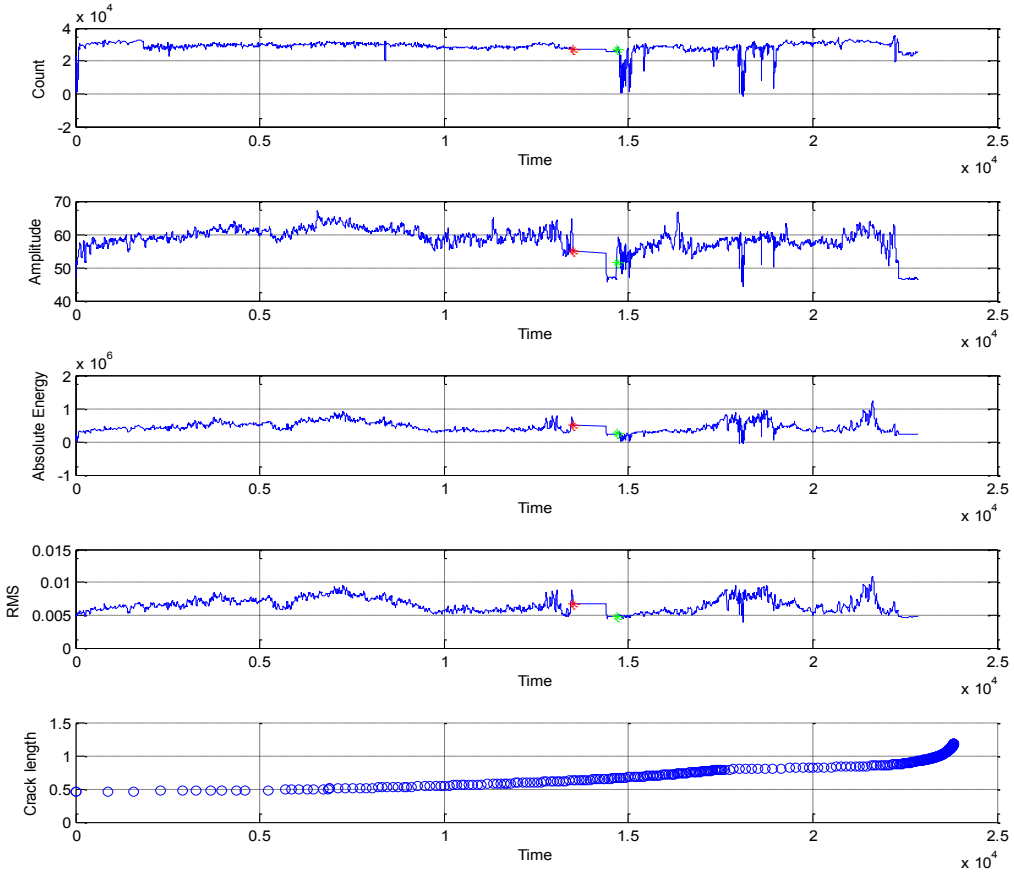

Figure 3-25 AE Signal for the FCC experiment (the time unit is second)

This chapter was dedicated to prove the existence of non-monotonic degradation due to imperfect maintenance and self-healing. The next step would be the analysis of non-monotonic degradation measures for the use in prognostic models. Chapter four presents various graphical and analytical trend analysis methods for analyzing degradation data.

4 TREND ANALYSIS

4.1 Background

As mentioned previously, the focus of this research is on empirical degradation-based prognostic which takes into account the measured or inferred conditions of a specific unit under study. This individual-based life estimation requires a population of the degradation paths for a specific fault mode. Success in accurate RUL estimation is primarily a matter of appropriate prognostic parameter selection. In this way, combining several degradation measures is a common practice in prognostics and it has the potential benefits of developing vigorous parameters. Coble and Hines introduced three parameter features to characterize suitable prognostic parameters [77, 78] : monotonicity, prognosability, and trendability. These metrics are useful for comparing the parameters in individual-based prognosis models. The optimal parameter can be selected through any optimization technique, such as Genetic Algorithms. According to Coble [78], the definitions and the relevant equations of the above-mentioned parameters are as follows:

1. *"Monotonicity: Underlying positive or negative trend of the parameter. It is given by the average difference of the fraction of positive and negative derivatives for each path."*

$$\text{Monotonocity} = \left(\left| \frac{\#\ \text{pos}\ \frac{\text{d}}{\text{dx}}}{\text{n}-1} - \frac{\#\ \text{neg}\ \frac{\text{d}}{\text{dx}}}{\text{n}-1} \right| \right) \qquad \text{Eq. 4-1}$$

2. *"Prognosability: A measure of the variance in the critical failure value of a population of systems."*

$$\text{Prognosability} = \exp\left(\frac{\text{std(failure values)}}{\text{mean}(|\text{failure values} - \text{starting values}|)} \right) \qquad \text{Eq. 4-2}$$

3. *"Trendability: Indicates the degree to which the parameters of a population of systems have the same underlying shape and can be described by the same functional form."*

$$\text{Trendability} = \min\left(\left| \text{Corr Coef}_{ij} \right| \right) \qquad \text{Eq. 4-3}$$

The focus of this research is on monotonicity which deals with subjacent positive or negative trend of parameters based on the assumption that systems do not experience any form of self-healing or human intervention (i.e. maintenance). In essence, this research deals with repairable systems in which self-healing or human interventions are desirable. A system with prognostic capabilities can lead to improved and more efficient maintenance. Thus, one can expect non-monotonic prognostic parameter upon post-prognostic maintenance and/or self-healing.

Degradation data can be considered as data series with a certain trend plus stochastic variations ($\varepsilon(t)$).

$$D(t) = f(t) + \varepsilon(t) \qquad \text{Eq. 4-4}$$

The stochastic variation (i.e. noise) is an inherent feature of degradation data. There are numerous methods for noise reduction and smoothing to minimize the stochastic variations. However, this study intends to analyze the main trend of the data regardless of the existence of the stochastic variation. In the following sections, the GPM will be used for the analysis of degradation data. It is important to remember that the GPM is based upon several significant assumptions:

- The model needs to assume a functional form of the degradation paths. The function can be derived from physical models.
- Degradation data are available for a population of identical components.
- A critical failure threshold can be assumed for the population of identical components
- Degradation measures collected under similar conditions.
- Finally, the degradation paths are monotonic.

For degradation data, trend can be defined as the pattern of the time-ordered degradation measures. Trend analysis is a prevalent tool in different fields of science (e.g. hydrology and environmental science) for analyzing time series of random variables to detect a significant change over time. A monotonic trend suggests unidirectional and consistent change in the mean level of degradation measures. It must be emphasized that the assumption of a monotonic trend limits the number of available functional forms (usually to

a linear, exponential, or logarithmic function). Moreover, using polynomial functions does not retain the monotony property. Table 4-1 shows the typical patterns and associated slopes in degradation signals [140].

Table 4-1 Typical patterns in degradation signals

	Concave Upward	Concave Downward	Linear
Monotonic Increase	$[\partial x] = +$ $[\partial\, \partial x] = +$	$[\partial x] = +$ $[\partial\, \partial x] = -$	$[\partial x] = +$ $[\partial\, \partial x] = 0$
Monotonic Decrease	$[\partial x] = -$ $[\partial\, \partial x] = +$	$[\partial x] = -$ $[\partial\, \partial x] = -$	$[\partial x] = -$ $[\partial\, \partial x] = 0$

A fruitful trend analysis requires proper sampling frequency and continuous recording of degradation measures. It also requires preprocessing of data in terms of selection and calculation of the quantities that describe the time-series for any specific task. Trend detection is a subset of trend analysis and it refers to identifying significant decreases or increases in the magnitude of a reference variable in time-series data [140]. Efficient trend detection is critical in early discovery of an impending failure. Furthermore, in the case of self-healing or human interventions, trend changes can be considered as the sign of improvement in the degradation rate. Melek and his colleagues [140] investigated various architectures for trend detection. They compared the performance of these techniques on physiological data. The authors focused on fuzzy approaches mainly because of the aptitude to identify underlying trends in the presence of noise and high fluctuation.

Trend analysis has been investigated for aging properties, such as failure rate and mean residual life (MRL) [141, 142]. Lim and Park [141] developed certain tests to distinguish increasing or decreasing MRL form constant MRL. The authors considered the situations where monotonicity is not an appropriate assumption in modeling aging properties.

Researchers also have worked on the trend (i.e. pattern) of failure process in repairable systems [143-147]. This type of trend analysis is mainly helpful for selecting the most suitable model. For example, the existence of trend in the failure times implies that the time between failures are not identically distributed, and therefore, renewal process models would not be appropriate for this data. Obviously, the primary focus is on the trend exhibited by the time between failures data to discover whether alterations in the pattern of failures are statistically significant. It is pertinent to mention that censored data is a major concern in trend analysis of failure data.

Both graphical and analytical methods have been used for the trend analysis of failure data. The analytical methods are mainly statistical hypothesis tests such as the Mann test, Laplace test, Anderson-Darling test, and the military handbook test. Normally, in these tests, the null hypothesis is formulated as no trend in the data, namely a non-homogenous Poisson process (NHPP) and the goal is to detect the increasing or decreasing trend in recurrent failures.

It is of note that increasing trend would be an indication of degradation over time and it happens when the time between failures tends to get shorter. Decreasing trend would be the indication of improvement over time where the time between failures tend to increase. Figure 4-1 shows the decreasing and increasing trends in failure times.

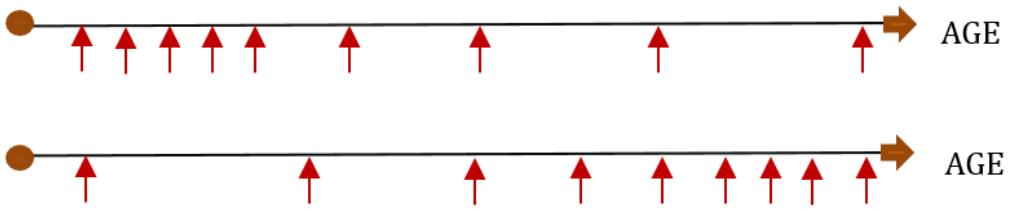

Figure 4-1 Decreasing and increasing trends in failure times

In this regard, a non-monotonic trend denotes the change of trends with time or the existence of cyclic failure times. For instance, the well-known bath tub curve can be considered as one form of non-monotonic trends in which time between failures tend to increase in the beginning, and after a period of stability, time between failures become

shorter i.e. indicate deterioration [145]. It is to be mentioned that the trend analysis of the failure times is a decent candidate for measuring prognosability.

The remainder of this chapter is dedicated to introduce several graphical and analytical trend analysis methods for analyzing degradation data. These methods have been utilized for time-series data in various fields of science. Methods to analyze and handle non-monotonic degradation are practically nonexistent in the literature.

4.2 Statistical trend analysis

Statistical trend analysis refers to a set of parametric and nonparametric statistical tests to detect the existence of a trend in time series. Hypothesis tests in statistical trend analysis usually contemplate the null hypothesis as "no deterministic trend". Each test has a statistical quantity to perceive the existence of a deterministic trend. The significance level indicates the probability of wrongly rejecting the null. The power of a test refers to the probability of rejecting the null when it is false. It is important to note that failure to reject the alternative means that there is no enough evidence to conclude the existence of a trend in data.

It is to be noted that a system under stress might experience step trend (e.g., due to shock) which appears as an abrupt shift in the degradation measures at a specific point in time. There are a number of parametric and nonparametric tests to detect and evaluate step changes (e.g. Kruskal-Wallis test, sign test, paired t-test).

The parametric tests (e.g. t-test) are mainly based upon the linear regression coefficients for normally distributed and homosedastic random variables [148]. Thus, the parametric tests are limited to linear trends. For this reason, parametric tests are not considered in this research. On the other hand, nonparametric tests are not highly affected by outliers and large data gaps. Mann-Kendal (MK) is the most used non-parametric trend test which is based on the relative ranking of data (i.e. a rank-based test). The MK test computes difference between the sequential data$(x_i - x_j)$. Then, it assigns an integer values of 1, 0, or -1 to the computed differences:

$$\text{sign}(x_i - x_j) \begin{cases} -1 & \text{for } (x_i - x_j) < 0 \\ 0 & \text{for } (x_i - x_j) = 0 \\ 1 & \text{for } (x_i - x_j) > 0 \end{cases}$$

If n represents the length of the data series, the test statistic, S, is computed using the sum of the integers as:

$$S = \sum_{i=2}^{n} \sum_{j=1}^{i-1} \text{sign}(x_i - x_j) \quad \text{Eq. 4-5}$$

Thus the MK test is related to the sample size, trend magnitude and coefficient of variation. Generally speaking, statistical trend analysis methods can be used only as a preliminary trend check, and therefore, it is not possible to justify further discussion on these methods.

4.3 Graphical methods

Graphical tests do not usually need complicated calculations. Thus, they are simple to perform and are powerful in detecting the strong trends. On the other hand, they might not be very useful for data with slight trends, and therefore they need to be combined with analytical tests. In addition, graphical tests are based upon subjective interpretation which is prone to error. There are a number of graphical methods to be applied in trend analysis. Using control charts was even mentioned in the literature as a graphical test. This section covers two graphical tools that were found suitable for degradation data: 1. cumulative plots, 2. temporal shape analysis.

Also, two sets of data were applied in this chapter. The first set was made through simulation of bearing degradation. The fault signal of bearings were modeled through the combinations of the following parts [149, 150]: repetitious impulses, load, bearing induced vibration, and machinery induced vibration (see section 5.3.3 for detailed explanations). The simulation used the ball pass frequency outer race as the frequency component of a bearing fault. This frequency is a linear function of the shaft speed. Figure 4-2 shows six different cases of degradation measures simulated for rolling element bearings with different shaft speeds. The last three cases provide the typical non-monotonic degradations after reducing the shaft speed. As shown in Figure 4-2, the simulation results are similar to the experiment results produced by Gebraeel & Pan [109] (see Figure 3-3).

The parameter features explained in section 4.1 (i.e. monotonicity, prognosability, and trendability) are calculated for the above cases using the codes provided in [78]. Not

surprisingly, the monotonicity for the first three are close to one. Table 4-2 provides the metrics for the non-monotonic case (i.e. case 4-6) for a population of 50 samples.

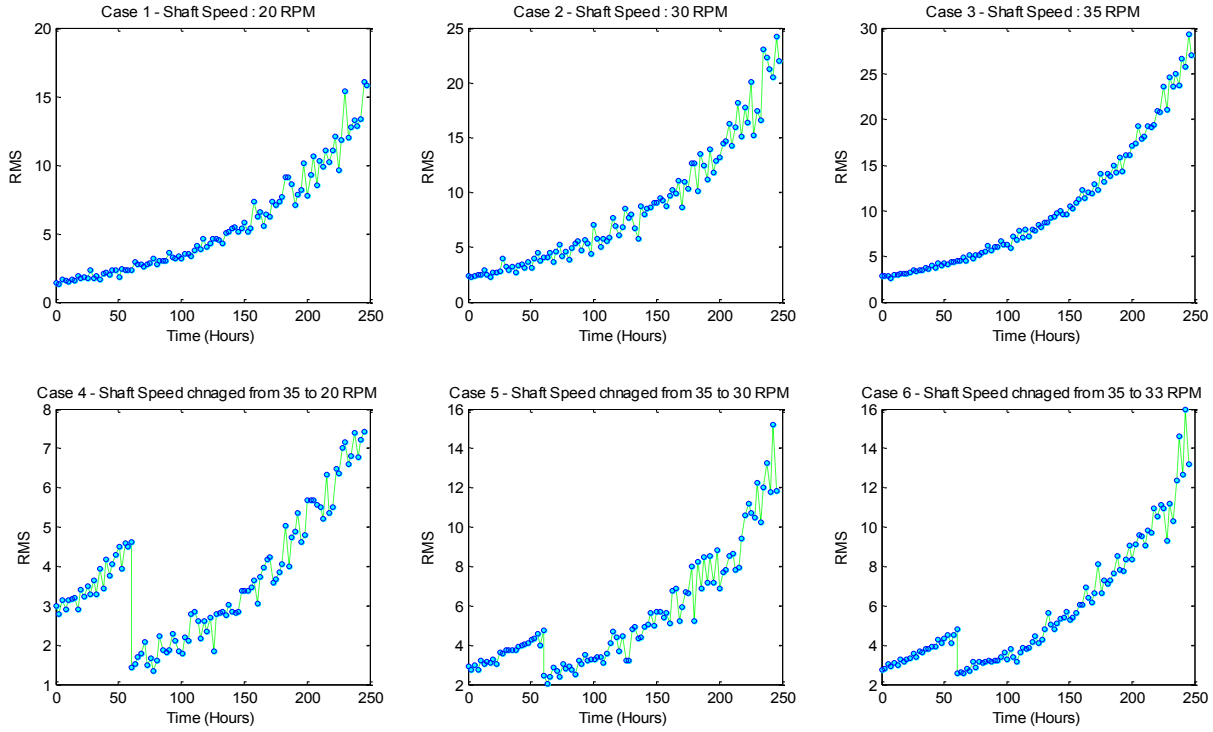

Figure 4-2 Different cases of degradation measures simulated for REBs

Table 4-2 The metrics for the non-monotonic degradations (case 4-6)

	Case 4 - Speed change from 35 to 20	Case 5 - Speed change from 35 to 30	Case 6 - Speed change from 35 to 33
Monotonicity	0.333	0.75	0.85
Prognosability	0.856	0.855	0.951
Trendability	0.916	0.9	0.942

For certain systems, damage is considered to be cumulative [78]. Thus, any change in slope of an increasing degradation can be a sign of improvement in durability. Figure 4-3 shows the cumulative plots of the simulated degradation data for 50 bearings with constant shaft speeds i.e. a monotonic degradation trend. Unlike the original plot of RMS, the

cumulative plot does not consider the slight trends or fluctuations in the data set. Figure 4-4 clarifies the change in trends where the shaft speed changes.

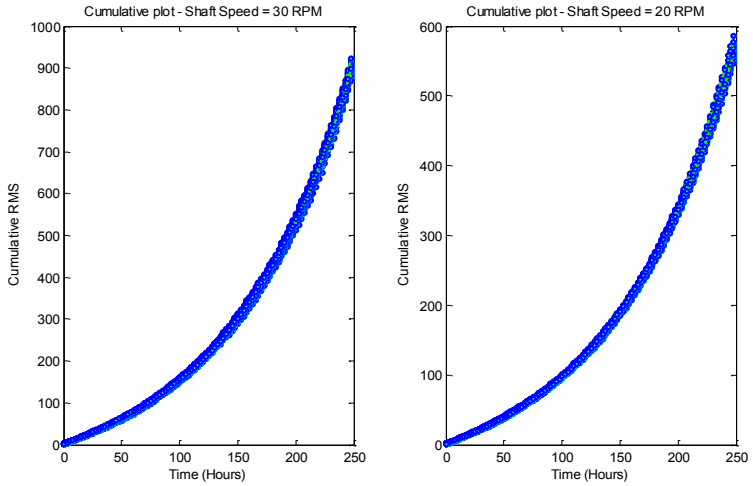

Figure 4-3 Cumulative plot for monotonic degradations

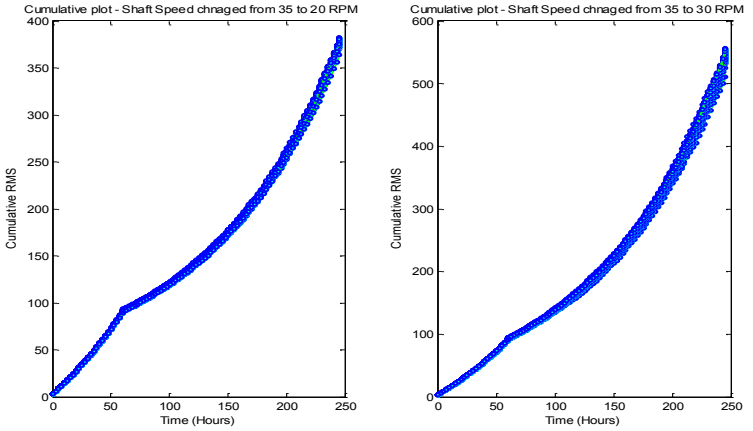

Figure 4-4 Cumulative plot for non-monotonic degradations

The cumulative plots clearly detect the strong trend in the degradation data due to the change in shaft speeds. Cumulative plots were also produced for the AE signal features. It is important to note that only energy-related features showed the change in trend during the retardation period.

The second graphical method known as temporal shape analysis is a qualitative analysis proposed by Konstantinov and Yoshida [151]. This analysis starts by approximating the variable *v* by a polynomial with the order of *m*. Then, the values of the polynomial are evaluated at *v*. At the next step, the first and the second derivatives of the polynomial are calculated.

The sign of difference between every data points and the preceding in the first and the second derivate should be extracted. The combination of the extracted signs forms the qualitative shape of the variable. Figure 4-5 displays the qualitative shape for the six cases discussed in Figure 4-2. It is obvious that the first derivatives (i.e. the first line of the qualitative shape) shows the change in trend and the second derivatives provide insight about the shape of the variable. It is also possible to show the qualitative shape by +1 and -1. In this way, no difference between the samples (i.e. zero value) can be taken into account. Figure 4-6 shows the qualitative shape using +1 and -1 for the polynomial with the order of four.

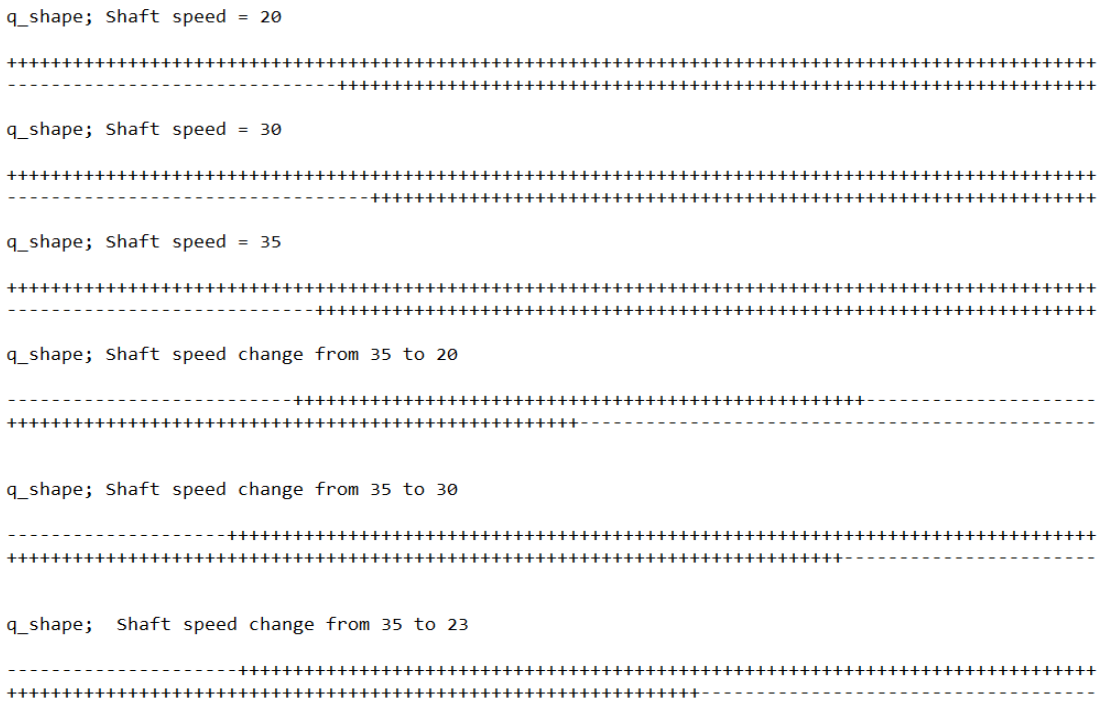

Figure 4-5 Qualitative shape for the six cases discussed in Figure 4-2

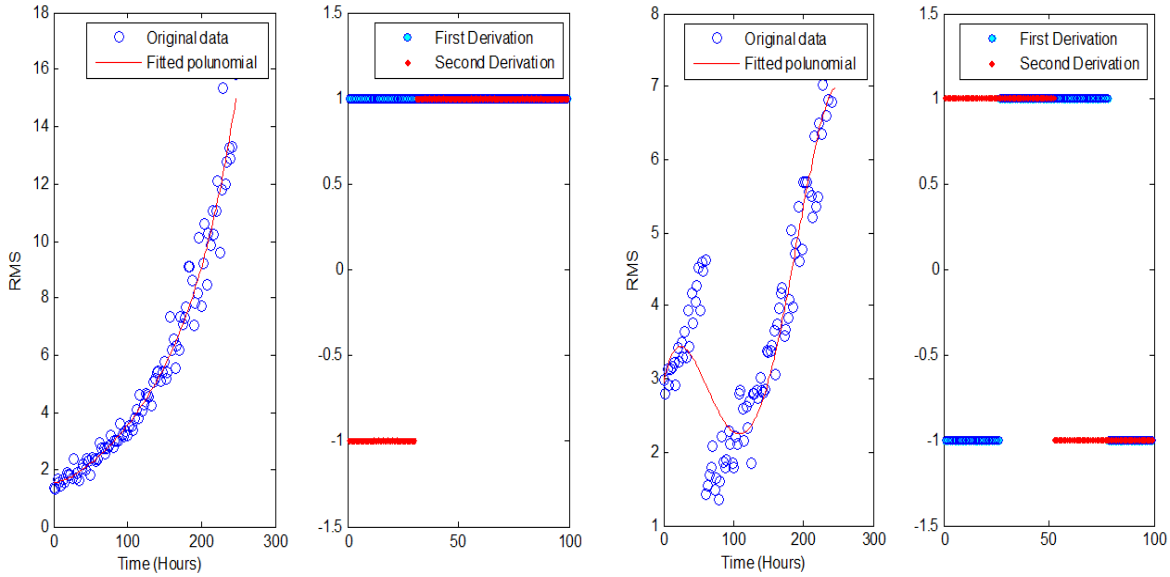

Figure 4-6 Qualitative shapes with +1 and -1; (left) Speed shaft equals 20, (Right) Speed shaft changed from 35 to 20

Next, the monotonicity can be calculated using the qualitative shape data and the functional form obtained by the temporal shape method. Table 4-3 provides the results for monotonicity with various polynomial order. In this table, qualitative shape refers to direct use of number of positive or negative derivation according equation 4-1. The values related to the polynomial fit depend on the order. It was realized that with order of 4 or more we can obtain the monotonicity of 1 for cases 5 and 6. In essence, high order polynomial take the noise into account and might not be reliable.

Table 4-3 Monotonicy values using the temporal shape method

		Case 4: Speed change from 35 to 20		Case 5: Speed change from 35 to 30		Case 6: Speed change from 35 to 33	
		Qualitative shape	Polynomial	Qualitative shape	Polynomial	Qualitative shape	Polynomial
Monotonicity	Order of 3	0.971	0.33	0.874	0.506	0.92	0.746
	Order of 5	0.338	0.33	0.513	1	0.558	1
	Order of 10	0.105	0.3733	0.171	1	0.236	1

4.4 Trend analysis using Hilbert-Huang Transform

The analysis of waveforms can be performed in the time domain or the time-frequency domain prior to the signal feature extraction. Time–frequency analysis has gained increasing attention in the field of trend analysis. These techniques analyze non-stationary signals (i.e. time series in trend analysis) in time and frequency domains simultaneously. Wavelet analysis is the most popular time–frequency analysis for detection of meaningful trends [140, 152] or finding the turning points [153]. In actual practice, wavelet-based method are more appropriate for real-time implementation because of better computational efficiency. This section investigates the use of Hilbert-Huang Transform (HHT) which was introduced in Chapter 3.

The Hilbert Transform (HT) within HHT delivers the instantaneous amplitude and frequency and information for each signal component i.e. IMF. In essence, the amplitude of the signal waveform can be used as a useful feature to display a non-monotonic trend. Nonetheless, amplitude in degradation signal might not be a good indicator of the incipient fault, particularly in the absence of certain physical quantities such as impacting. Figure 4-7 shows a typical waveform which is a combination of small waveform collected periodically.

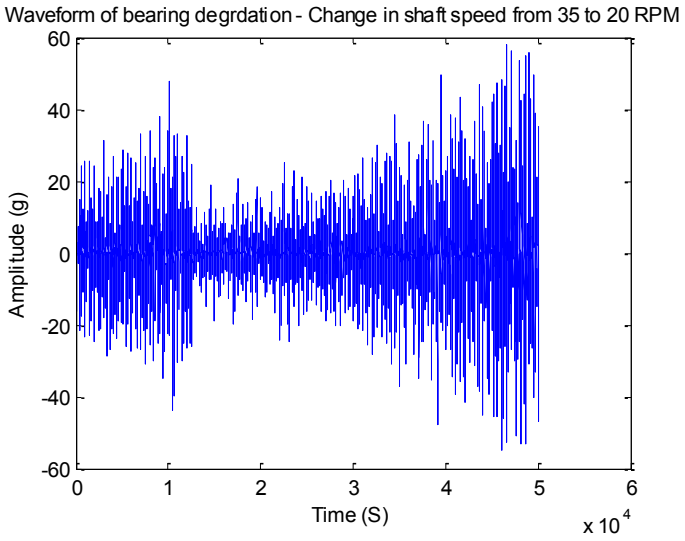

Figure 4-7 Typical waveform

The change in the amplitude of the signal represents the change in the shaft speed. However, through analysis is required if there is slight trend change or when the comparison of several degradation signal is required. In essence, HHT can decompose the signal to several components and provide the instantaneous amplitude. Thus, instead of the whole waveform, the important components of the signal can be analyzed and the component associated with noise can be removed [154, 155]. Furthermore, more useful signal features can be utilized for the instantaneous amplitudes.

Figure 4-8 shows the first four components of the wave shown in Figure 4-7. Figures 4-9 and 4-10 shows the Hilbert spectrum i.e. time-frequency analysis of the signal and the instantaneous amplitude for each components respectively.

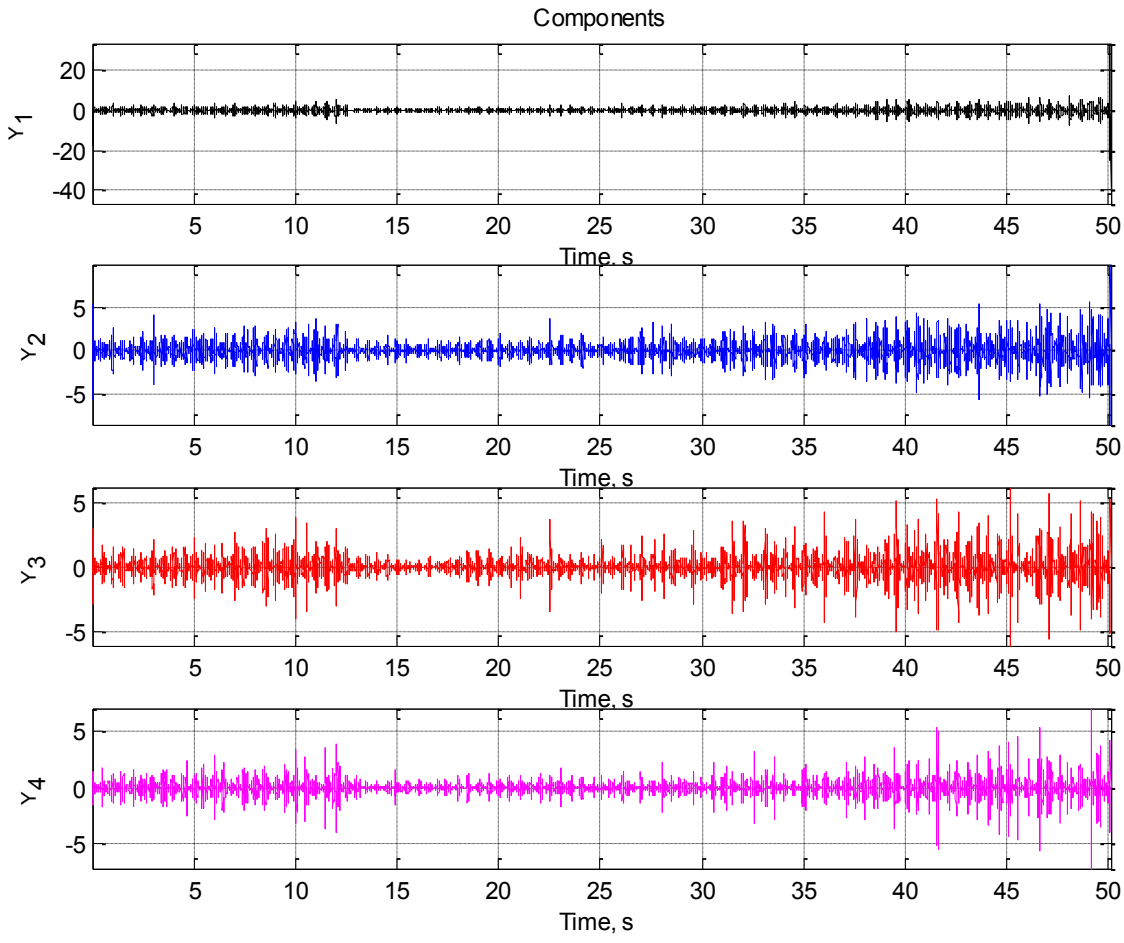

Figure 4-8 Components of the waveform

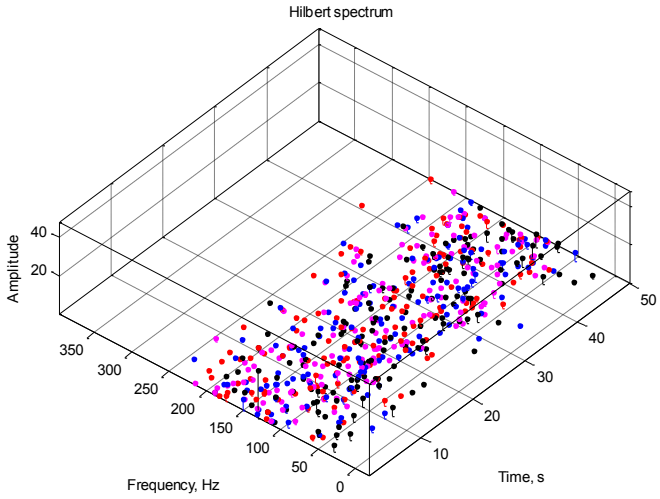

Figure 4-9 Time-frequency analysis

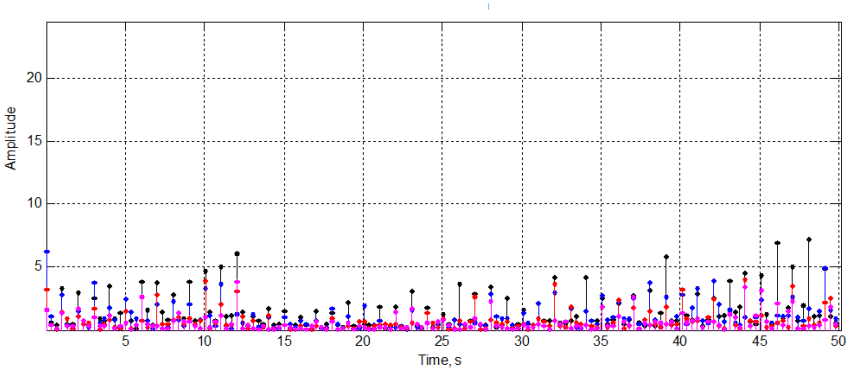

Figure 4-10 Instantaneous amplitude for each component

In order to further improve this analysis, the concept of crest factor is utilized. For a waveform, crest factor is the ratio of the peak amplitude to the RMS value. This dimensionless feature is an indication of the significance of the peaks. The square of the crest factor is known as the peak-to-average power ratio (PAPR). Figure 4-11 illustrates the crest factor calculated by using the instantaneous amplitude of the first signal component for the three non-monotonic cases shown in Figure 4-2. It is evident that the crest factors in the above figure is not helpful for both detecting the trend change and the comparison of different cases of the speed change. For this, the calculation of crest factor was modified. In the new form, instead of using the RMS for the corresponding waveform, the maximum RMS

up to the time of calculation was used. This can insure that the instantaneous amplitudes are compared against the waveform with the maximum energy. The trend changes using the modified crest factor is illustrated in Figure 4-12.

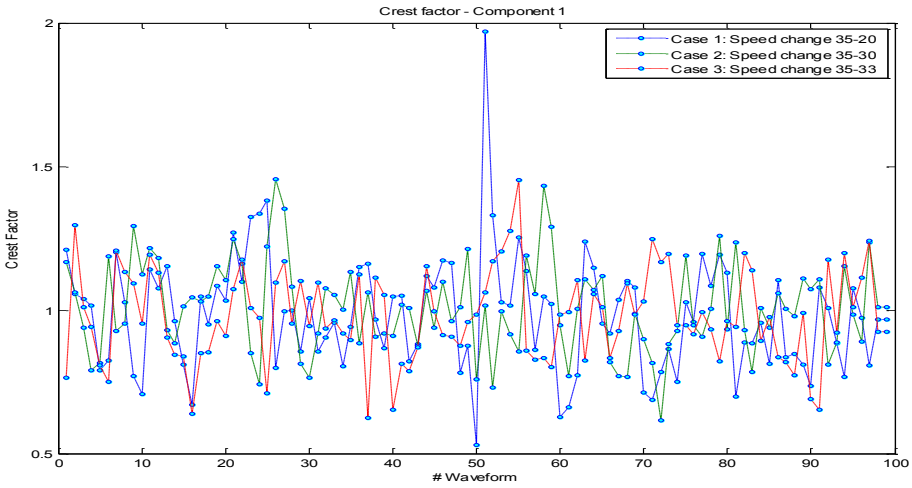

Figure 4-11 Crest factor

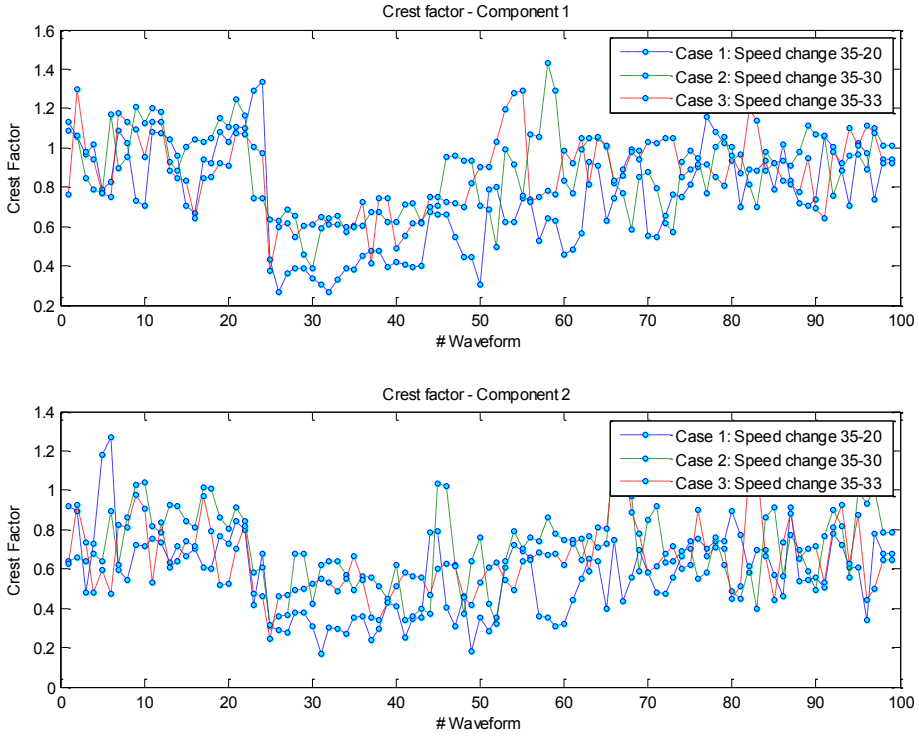

Figure 4-12 Modified Crest factor (Peak/max RMS)

In order to further improve the calculation of the crest factor, the ratio of the maximum RMS over the RMS for each waveform was used. Accordingly, the trend changes can be properly displayed as shown in Figure 4-13. This form of crest factor provides better comparison among the presented cases for the first two signal components.

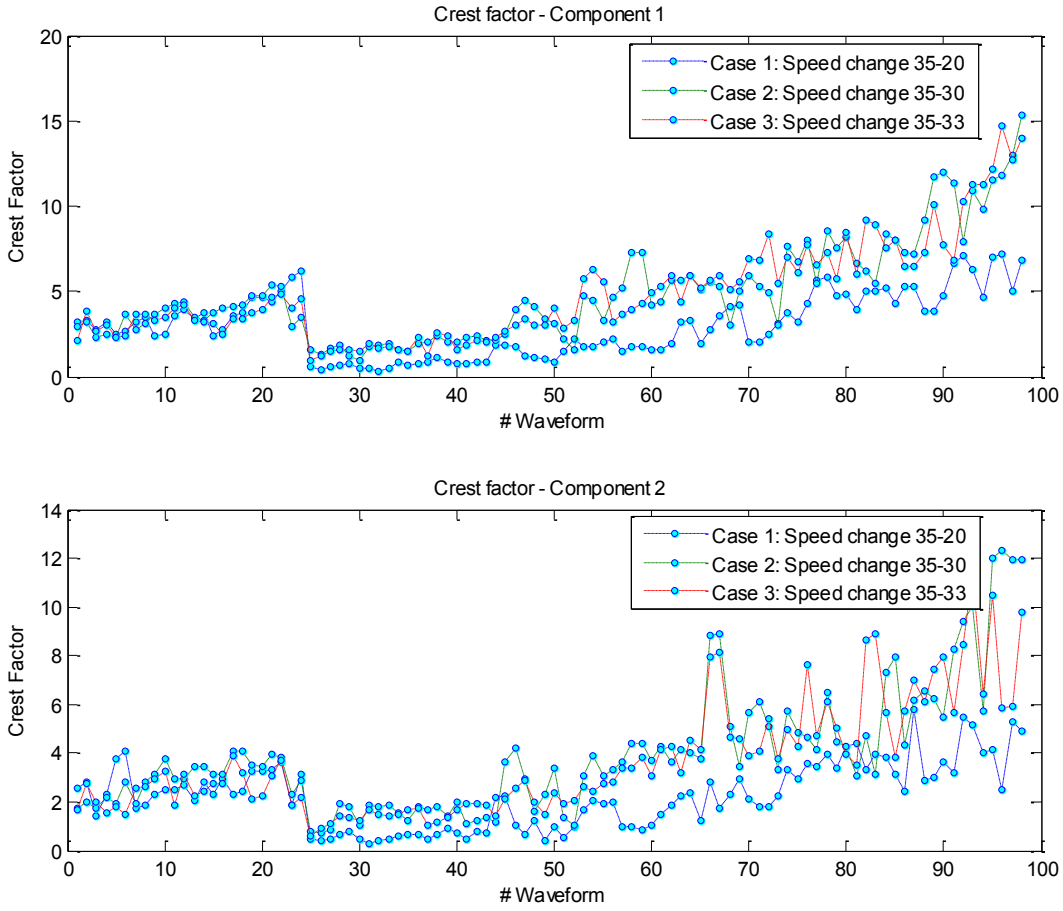

Figure 4-13 Modified Crest factor (Peak/ (max RMS / Waveform's RMS))

4.5 Multivariate trend analysis

For some data sets, including AE data, it might be of interests to analyze the whole complex multivariate data. There are many approaches to investigate multivariate data sets. Among those methods, Independent Component Analysis (ICA) and Principal component analysis (PCA) seem to be good candidates for trend analysis of multivariate degradation signals.

4.5.1 Principal components analysis

PCA is a widely used linear statistical technique for reducing the dimensionality of complex multivariate data. In PCA, the goal is to analyze covariance structure and reduce the complexity of the data by projecting onto a lower dimensional subspace while retaining the variability. Through the PCA, the data will be represented by the products of scores (i.e. mutually orthogonal data) and principal component loadings (i.e. transposed linear transformation matrix) plus the matrix of residuals [156]. PCA has been utilized in the literature to analyze the features and the waveform of AE signal [156-158].

In this analysis, the AE data set related to one FCC experiment was utilized. This data set contains over 27000 observations and 11 response variables (i.e. signal features) as follows:

- X1: PAC-Energy
- X2: Average frequency (over the entire AE hit = AE count / Duration)
- X3: Initiation frequency, i.e. rise-time based frequency
- X4: Count
- X5: Amplitude; maximum AE signal excursion
- X6: Absolute signal level (ASL) is the averaged amplitude of the AE signal
- X7: Reverberation frequency (ring down frequency = (AE count - count to peak)/ (duration – rise-time)
- X8: Absolute energy is the true energy measure of AE hit.
- X9: Rise-time is the time from the start of AE hit to the time of the peak amplitude
- X10: Root Mean Squared (RMS)
- X11: Signal strength; similar to energy but it is calculated over the entire AE

Figure 4-14 shows the typical raw data of the first variable (Energy). This figure clearly shows the need for smoothing the data especially due to high level of background noise. Smoothing algorithms extricate the data from the periodic components and retain the long term trends. The Savitzky–Golay smoothing filter takes care of high frequency information

that should be retained. This filtering method increases the signal-to-noise ratio by performing local polynomial regression. The Butterworth filter is utilized to remove high frequency noise in the signal. Figure 4-15 shows the results of applying four methods of smoothing. The areas shown in the red ellipses represents the period of retardation in the signal i.e. crack closure. Butterworth filtering was chosen for further analysis on the data.

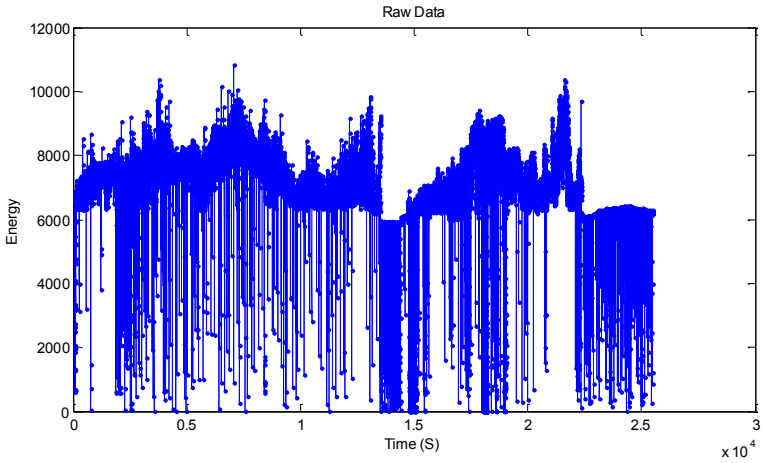

Figure 4-14 Typical AE raw data for the FCC experiment

Using scatter plots and correlation map the highly correlated variables were identified. As shown in Figure 4.16, the correlation map shows the pseudocolor map of the correlation matrix for the data set. It is obvious that Energy (variable 1) is highly correlated with count (variable 4) and absolute energy (variable 11). All these variables are related to the energy of the signal. Also, RMS (variable 10) is correlated with ASL (variable 6).

In essence, correlated variables increase the contribution of their related PC. Consequently, the highly correlated variables were eliminated for the next stages of analysis. Thus, data with 7 variables was used for the next phase. PCA was performed for the smoothed AE data that was free of highly correlated variables. Figure 4-17 shows the scatter plot and the bi-plot for all the vectors. The direction and length of the vectors in the bi-plot represent the contribution to the first two principal components. For example, the first PC on the horizontal axis has positive coefficients for all the variables except variable 3. Figure

4-18 provides the scree plot of the data. With the first four PCs it is possible to represent over 90% of the variability of the data.

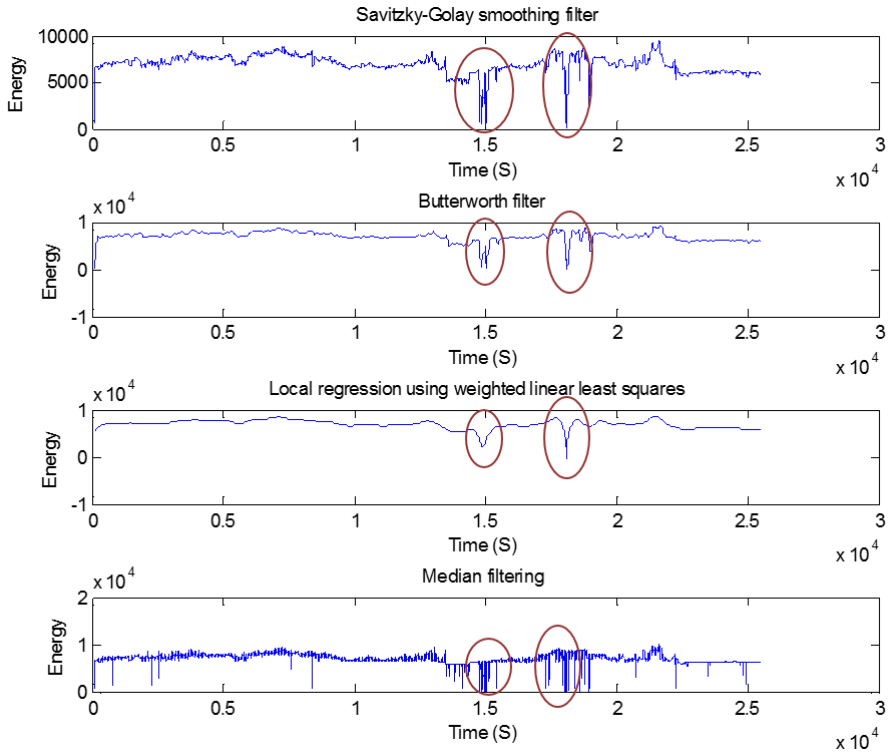

Figure 4-15 Results of applying various methods of smoothing

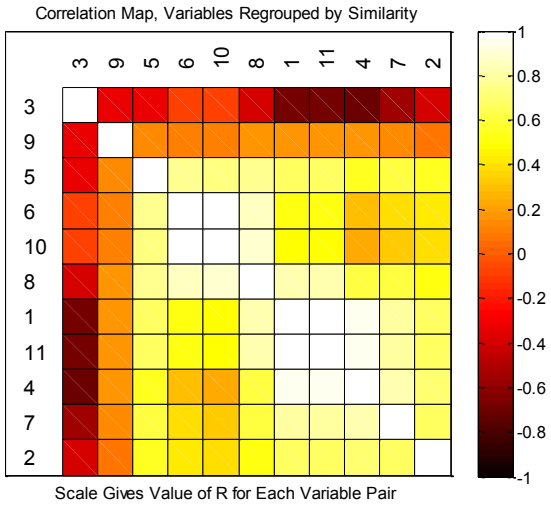

Figure 4-16 Correlation maps for the data set filtered by Savitzky-Golay

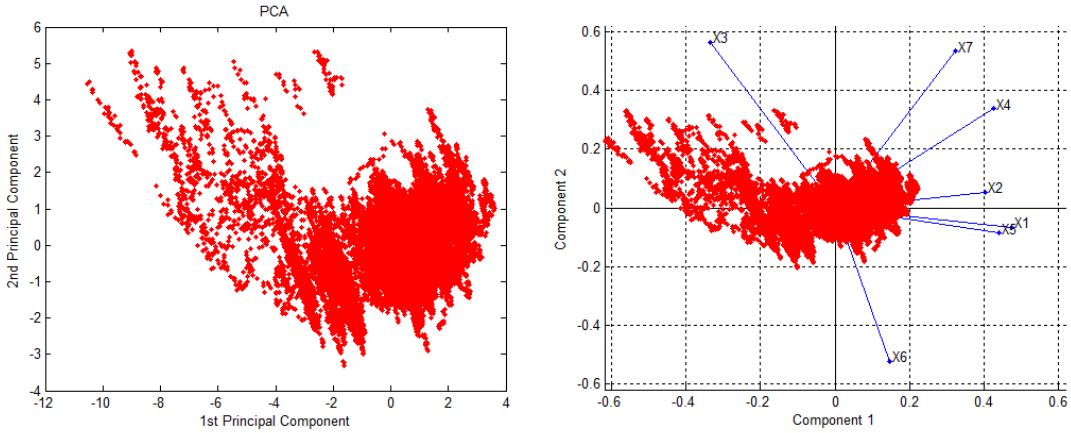

Figure 4-17 The first two PCs

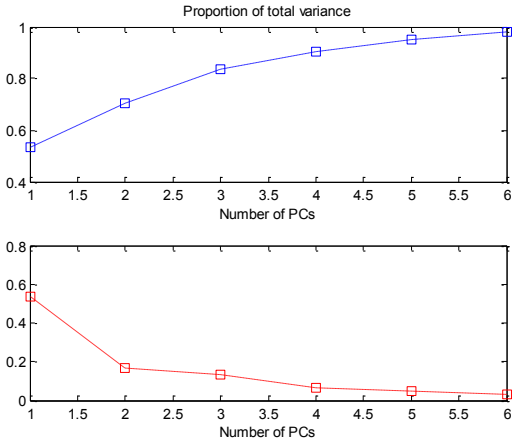

Figure 4-18 The scree plot of the AE data with 7 variables

Similar analysis was performed for the portion of the data that contain the retardation as shown in Figure 4.19. Interestingly, similar to the bi-plot for the whole data (see Figure 4-17, variable 3 cause the variation on the negative portion comparing other variables (Figure 4-20). As clearly shown in figure 4-19 all the retardation period for the signal features have some negative values except for variable three that goes upward. To this end, by eliminating variable 3, better results are expected. Furthermore, it is possible to look at the portion of the data that has no retardation. The first 14000 observations used for the next analysis as displayed in Figure 4-21. Finally, by removing variable 3 and comparing the presented bi-plots, it is possible to say that the shift in the relationship between PCs can represent the

retardation period. Thus, it can be concluded that PCA can show the trends in the data (Figure 4-22).

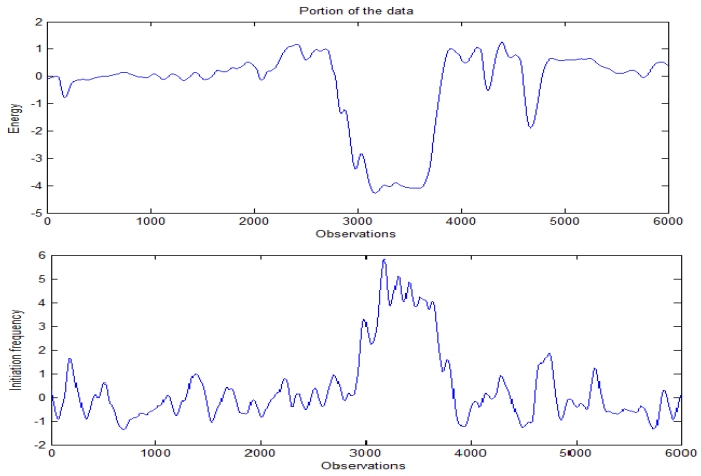

Figure 4-19 Portion of the data that show the retardation

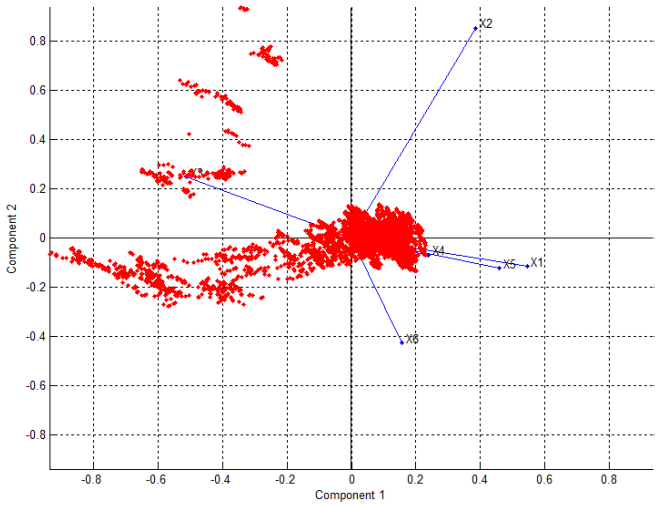

Figure 4-20 The first two PCs (Retardation portion)

There are always two important questions regarding PCA: 1. How many principal components should be retained for the analysis, 2. Is there any useful information in the remainder of PCs (i.e. PC4, PC5, and PC6) that was not selected. For the answer of the second question the information complexity criteria was used [159]. The results show that one PC

would be enough to cover the remainder of information that not covered in the selected PCs. It seems variable 6 may contain significant information. It is clear from Figure 4-22 that variable 6 has different direction and it should contain some information that is not appeared in the first three principal components. Variable 6 is the RMS which is basically different from other features of the signal variables.

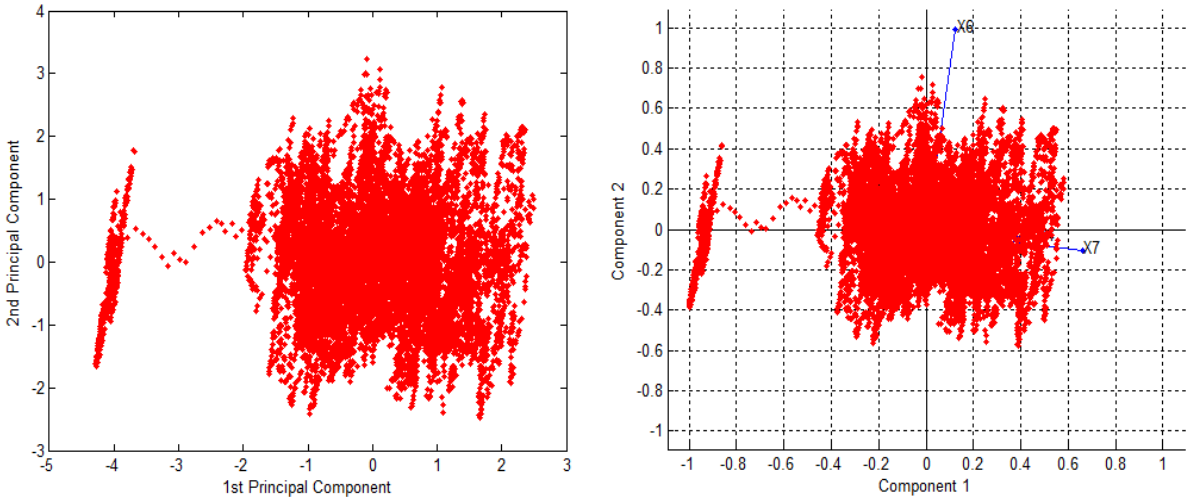

Figure 4-21 The first two PCs (portion of the data without retardation)

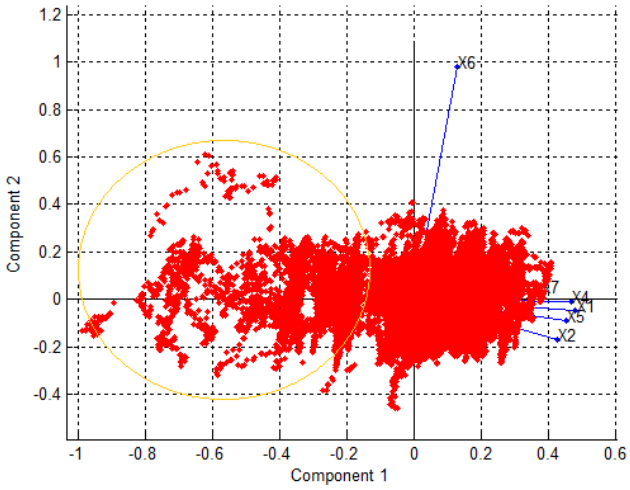

Figure 4-22 The first two PCs–Orange circle represents the retardation period

Considering the new scree plot for the modified data, we can conclude that three principal components are appropriate to represent that data even with the retardation period. Figure 4-23 shows the selected PCs. Also, it is possible to use the PCs to reconstruct the original samples as shown in Figure 4-24.

In brief, it was realized that PCA can show the trends in the data providing its perfect performance. Preliminary analysis of the data including smoothing, filtering, elimination of the correlated variables were very helpful in proper implementation of PCA.

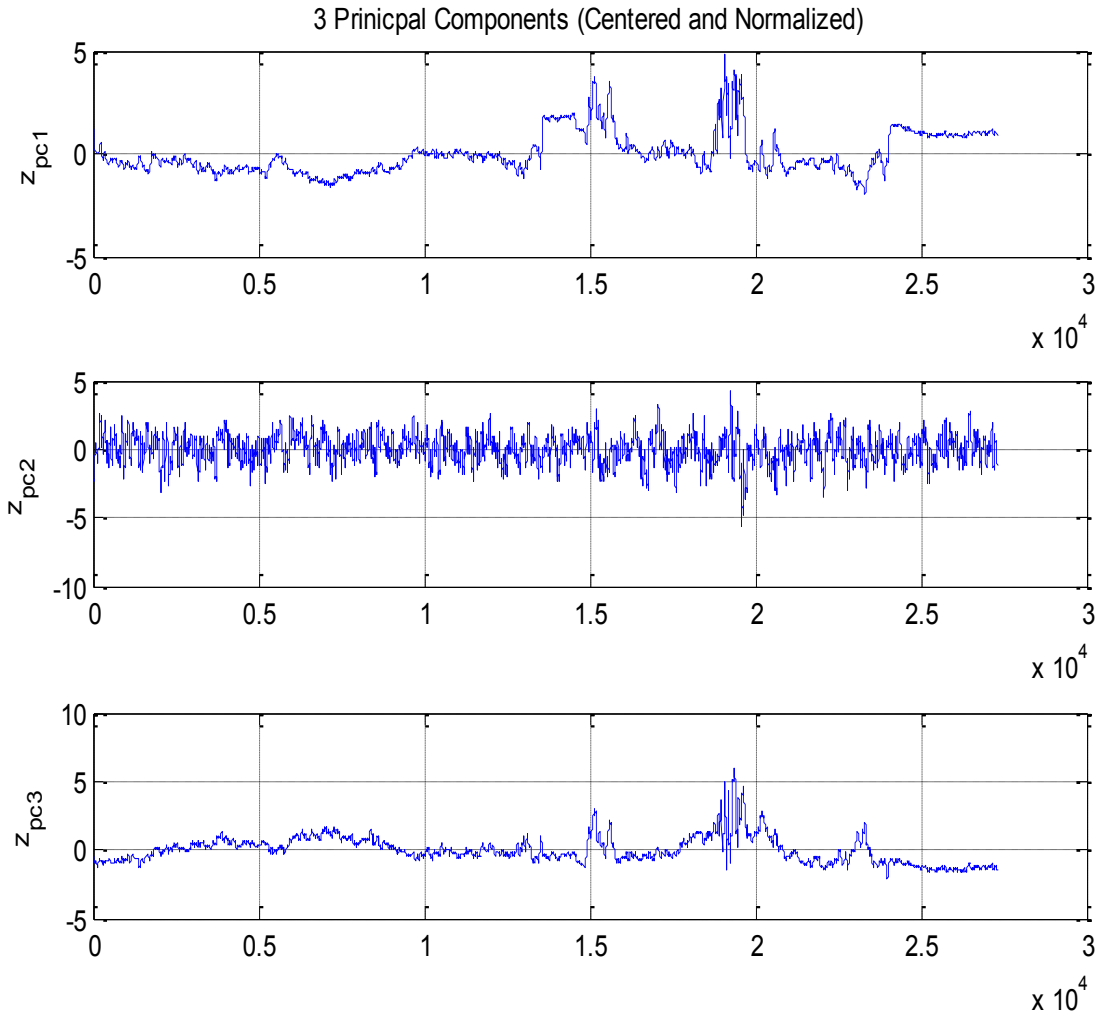

Figure 4-23 Selected PCs

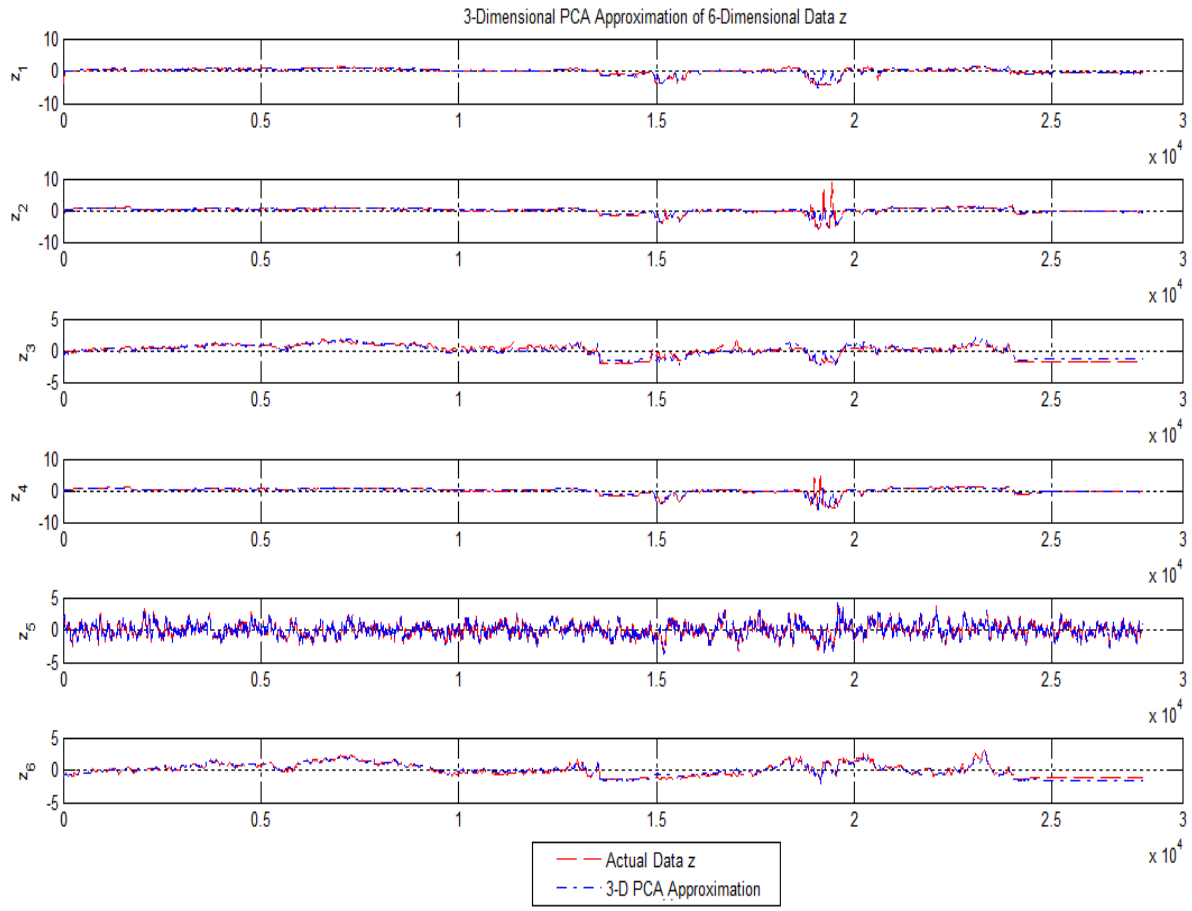

Figure 4-24 Reconstruction of the original samples

4.5.2 Independent Component Analysis (ICA)

ICA is a popular signal processing method to decompose a multivariate signal into uncorrelated and maximally independent components [160]. This method is also helpful in selecting the proper variables for further analysis. This method was applied on the AE data for the FCC test and it was not able to appropriately handle the retardation period. Figure 25 shows the cases when we select to have one and two independent components. Figure 26 which displays the decomposition of the data based on ICA, we can conclude that this method was not able to appropriately handle the retardation period in the AE signals.

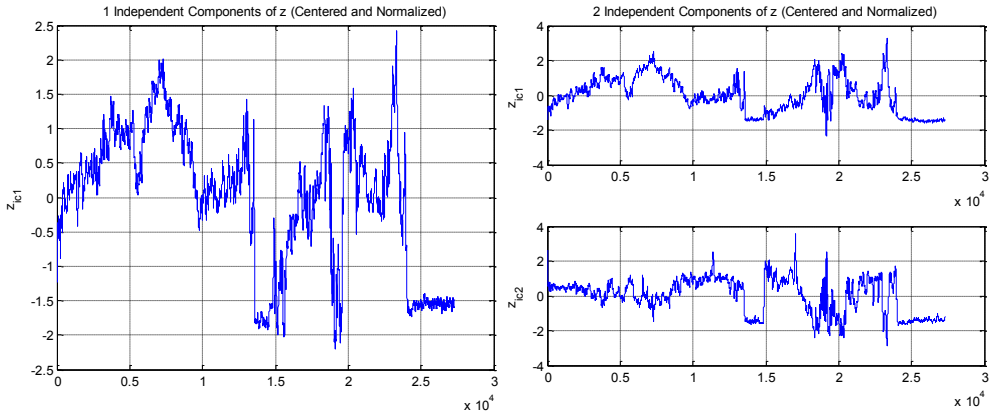

Figure 4-25 ICA components

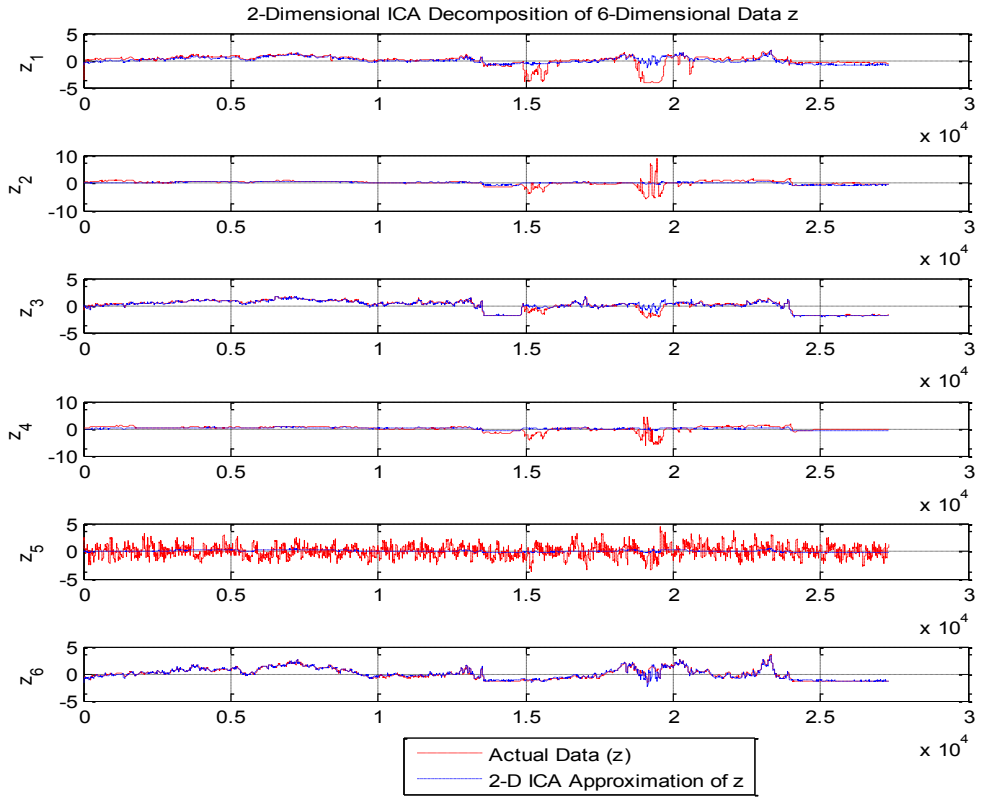

Figure 4-26 ICA decomposition of the data

The next two sections introduce two methods to treat non-monotonic degradation signals. Of course, the usefulness of these methods is case-dependent.

4.6 Average Conditional Displacement Algorithm

This section considers the application of an algorithm known as Average Conditional Displacement (ACD) for automatic estimation of monotonic trends. ACD describes a trend by piecewise linear curves. This approximation algorithm was first introduced by Vamoş [161]. It is to be noted that ACD is based upon the signal values interval not the time interval. In this algorithm, the slope of each estimated linear segment is proportional to the average of the time series values in the corresponding interval [161, 162].

The advantage of ACD algorithm is twofold: 1) no need for initial assumptions such as functional form of the trend, 2) the development of an automatic algorithm. The accuracy of ACD is comparable with well-tested methods such as polynomial fitting and moving average particularly for the signals with stationary noise.

The ACD works well in estimating the monotonic trend for time series data with arbitrary stationary noise. More importantly, ACD reveals one of the possible monotonic components of a non-monotonic trend. Suppose that a time series $\{x_n\}$, with $1 \leq n \leq N$, can be generated by a discrete stochastic process $\{Zn\}$ and the values of the trend f(t):

$$X_n = f_n + Z_n \qquad \text{Eq. 4-6}$$

Considering δt as the sampling interval, f_n represents the value f(t) at the moment

$$t_n = (n-1)\,\delta t. \qquad \text{Eq. 4-7}$$

Instead of the unknown initial and final values (i.e f_1 and f_N) , the ACD algorithm uses the extreme values of the time series. In addition, we assume that Z_n does not depend on f_n. Furthermore, we assume that $\{Z_n\}$ is stationary and $\{X_n\}$ is a non-stationary process. The probability distributions of $\{Z_n\}$ and $\{X_n\}$ are denoted by $p_z(z)$ and $p_x(x,n)$ respectively. Therefore, we have $p_x(x,n) = p_z(x- f_n)$. The theoretical background of the ACD is elaborately explained in [161]. Here, only the numerical quantities of the ACD are presented.

Figure 26 shows the one-step variation of the time series $\{x_n\}$ for the interval $(\xi_j, \xi_{j+1}]$. There are J disjoint intervals that embrace all the values of x_n. Therefore, N_j is the number of the values of x_n in the interval $I_j = (\xi_j, \xi_{j+1}]$, for j = 1,2,…, J. An increase in N results in improved accuracy of the trend estimation.

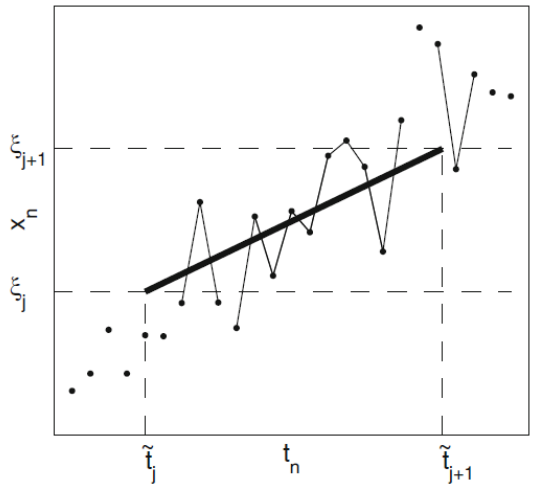

Figure 4-27 The one-step variation of the time series in which the thick straight line represents the ACD approximation [162]

However, the ratio between the noise fluctuation and the amplitudes of the trend variation is the major contributor to the accuracy of the ACD algorithm. In essence, when homogeneous intervals are used, the difference between the values of N_j should not exceed a unit at the most. The one-step variation of the time series can be defined as

$$\delta x_n = x_{n+1} - x_n \qquad \text{Eq. 4-8}$$

The sample average of δx_n (i.e. the average variation of $\{x_n\}$ within I_j) is computed as follows:

$$\hat{g}_j = \frac{1}{2N_j}\left(\sum_{x_n \in I_j} \delta x_n + \sum_{x_{n+1} \in I_j} \delta x_n\right) \qquad \text{Eq. 4-9}$$

It is to be noted that the interval I_j should contain the initial or final values. Moreover, the values of $\hat{g}_j$ should have similar sign in order to be used in the numerical approximation of the monotonic trend. Otherwise, the monotonic trend cannot be determined. If the values of $\hat{g}_j$ have different signs, repeated central moving average (RCMA) can be utilized to smooth the fluctuation of the time series. RCMA provides a gradual smoothing of the time series based upon two parameters: (1) length of the averaging window, (2) the number of averaging.

In Figure 26 the thin straight segments denote the pieces of $\{x_n\}$ that enter into the computation of the sample average i.e. $\hat{g}_j$. The thick continuous line denotes the ACD approximation of the monotonic variation of $\{x_n\}$ for the interval $(\xi_j, \xi_{j+1}]$. Thus, the points $(\tilde{t}_j, \xi_j)$ and $(\tilde{t}_{j+1}, \xi_{j+1})$ demarcate the j-th straight segment which has the slope of $\hat{g}_j$. Hence

$$\tilde{t}_{j+1} = \tilde{t}_j + \frac{\xi_{j+1} - \xi_j}{\hat{g}_j} \quad \text{Eq. 4-10}$$

Therefore, for $t \in (\tilde{t}_j, \tilde{t}_{j+1})$ the estimated monotonic trend (i.e. piecewise linear curve) is

$$\tilde{f}(t) = \xi_j + (t - \tilde{t}_j)\hat{g}_j \quad \text{Eq. 4-11}$$

Furthermore, $\hat{g}_j$ are negative for decreasing trend. Thus the points $(\tilde{t}_j, \xi_{J-j+2})$ and $(\tilde{t}_{j+1}, \xi_{J-j+1})$ demarcate the j-th straight segment which has the slope of $\hat{g}_{J-j+2}$. Hence

$$\tilde{t}_{j+1} = \tilde{t}_j + \frac{\xi_{J-j+1} - \xi_{J-j+2}}{\hat{g}_{J-j+2}} \quad \text{Eq. 4-12}$$

Thus, we have

$$\tilde{f}(t) = \xi_j + (t - \tilde{t}_j)\hat{g}_{J-j+2} \quad \text{Eq. 4-13}$$

The automatic ACD algorithm requires three things [162]:

1. Distribution of the time series values into disjoint intervals I_j. The primary purpose of the ACD is to specify the trend shape as accurate as possible. For this, the distribution of the time series values into disjoint intervals is inversely proportional to the noise standard deviation. It implies that small number of intervals needs to be used if the noise fluctuation is high.
2. Determining the parameter that controls the speed on noise damping in the averaging process. This parameter is the maximum value of the semi-length of the averaging window. The noise fluctuation can be strongly damped by large averaging window. However, it can lead to distorted trend.

3. Determining the threshold to stop the iteration of the ACD algorithm when the noise damping reached an acceptable level. For this, we seek the maximum value of the ratio between the noise variance and the final residual variance.

Figure 4-28 and Table 4-4 show the application of the ACD for treating the non-monotonicity in the degradation signals.

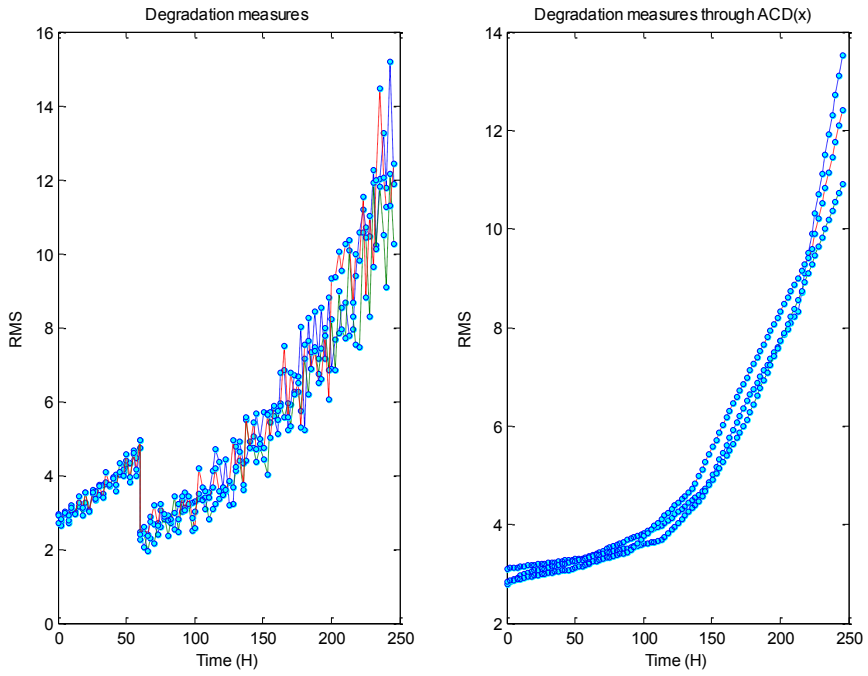

Figure 4-28 Applying ACD for degradation signals

Table 4-4 Applying ACD for degradation signals

	Case 4 - Speed change from 35 to 20	Case 5 - Speed change from 35 to 30	Case 6 - Speed change from 35 to 33
Monotonicity	1	1	1
Prognosability	0.866	0.914	0.948
Trendability	0.877	0.967	0.983

4.7 Trend-Based Segmentation

Trend based segmentation methods (TBSM) are able to locate the turning points in the time series data. Segmentation methods are popular in financial time series data to locate a set of trading points [153, 163, 164]. The most common segmentation methods include piecewise linear representation (PLR), Fourier transform, and wavelets. This study performs PLR. The details of PLR along with a pseudocode is provided in [153, 163]. Figure 4-29 shows three stages of segmentation process which was enough to group the data set into 4 segments in the third stage.

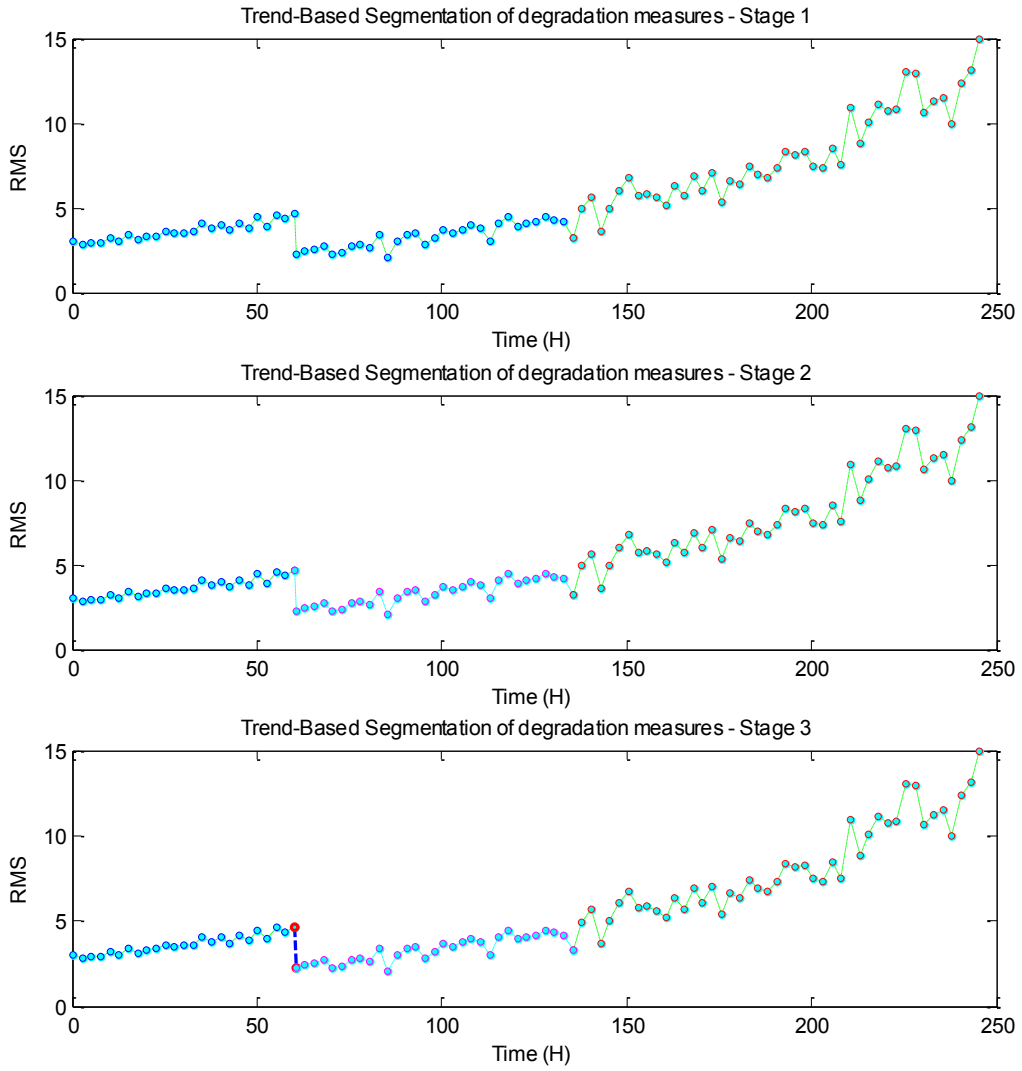

Figure 4-29 Three stages of segmentation process

The first segment represents the degradation with high speed and the second segment represents the reduction in speed (i.e. change in operating condition). It is to be noted that the Bayesian GPM did not provide any results for a set of non-monotonic data similar to what is shown in Figure 4-29. Therefore, the data was segmented based on the PLR method. This time only the segments after the change in speed were taken into account (i.e. segment 3 & 4) and the prognostic model worked properly. Next, Bayesian GPM was applied for a set of non-monotonic data similar to what is shown in Figure 4-29. The prognostic model did not provide any results. Hence, in the next step the data was segmented based on PLR method. This time only the segments after the change in speed was taken into account, and the prognostic model worked properly.

In chapters 3 and 4 the existence and analysis of non-monotonic trend were discussed. The next chapter is dedicated to fact that we can extend the life engineering systems by manipulating the operating conditions such as shaft speed. In this respect, it is required to apply the trend analysis methods presented in the current chapter. However, Chapter 5 presents the an example of post-prognostic decision making with application to a simple power generation system.

5 INTERACTIVE DECISION SUPPORT SYSTEM

5.1 Introduction to Decision Making

The procedure of decision making falls into one of the following categories:

- Intuitive analysis of information which results in a decision that cannot be justified.
- Programmed decision making based upon past analysis and experience.
- Analytical approaches that consider possible alternatives and their consequences

This chapter talks about a Decision Support System (DSS) to gain insight into the analytical approaches with respect to operation and maintenance decision making (OMDM) based on prognostic information. A Decision Support System (DSS) may be defined as a computerized information system to support decision-making activities. Hence, using condition monitoring data and degradation models, an intelligent DSS should be able to transfers the monitored machine conditions and prognostic information into maintenance schedules [165].

As mentioned earlier, one objective of this dissertation is extending the application of prognostics to OMDM. It is pertinent to mention that the knowledge about the remaining life would offer more operation and maintenance alternatives. The estimation of RUL, also called prognostic distance, gives the DM lead-time and flexibility to perform maintenance actions at any time until the failure [166].

To be specific, OMDM of a typical wind-based power generating systems is considered in this chapter. Naturally, OMDM is a complex process due to the followings:

- Presence of various decision alternatives and decision criteria, and therefore, the need for compromising.
- Multiple objectives which usually have conflict and make the judgment of alternatives a complex endeavor.
- Uncertainty associated with the outcomes and information e.g. future cost.

Thus, one can look forward to a form of multiple criteria decision making (MCDM) problem to provide rational and consistent solutions for operation and maintenance problems. The general model of MCDM can be represented as follows:

$$\text{Maximize } [C_1(x), C_2(x), \ldots, C_k(x)],\ x \in X$$

Here, *X* represents the set of available alternatives, *x* is an alternative and C_i is the i^{th} evaluation criterion. Alternatives are defined as the possible courses of action. It is evident that one needs to develop all the possible alternatives in order to have a fruitful decision making. The descriptors of alternatives are called attributes which include performance parameters and inherent characteristics. Criteria are the rules or standards for evaluating the alternatives. Needless to say, the selected criteria should be clearly understood by decision maker (DM). A directly measurable criterion is called true criterion. When it is difficult to directly measure a criterion, it can be substituted by one or more surrogate criteria. It is to be noted that, in some situations, several MCDM methods need to be utilized to validate the results. In this respect, there are some methods to aggregate the results of various MCDM methods.

5.1.1 Multiple attribute decision making

Decision making problems can be divided into two classes depending on the availability of alternatives. The first class is known as selection problems or multiple attribute decision making (MADM) that focuses on selecting the best alternative among a finite set of predetermined alternatives by evaluation against a set of attributes. The general model of MADM is:

$$X = [A_1, A_2, \ldots, A_m]; \text{ set of alternatives}$$

$$C = [C_1, C_2, \ldots, C_n], \text{ set of criteria}$$

x_{mn} : performance of the n^{th} criteria on the m^{th} alternative

$$Z = \begin{bmatrix} x_{11} & \cdots & x_{1n} \\ x_{21} & \cdots & x_{2n} \\ x_{m1} & \cdots & x_{mn} \end{bmatrix}$$

The MADM model is usually represented in pay-off tables as shown below. It is clear that this type of decision analysis starts with alternative formulation and criteria selection. For instance, the decision criteria in power and energy systems fall into different categories: technical, economic, social, and environmental. Efficiency, safety and reliability are the most

used technical criteria. Economic criteria are broader and may include investment cost, operation and maintenance cost, expected lifetime or service life, product cost, and payback period. The methods of criteria selection are discussed in [167, 168].

Table 5-1 Pay-off Table

	C_1	C_2	…	C_m
A_1	x_{11}	x_{12}	…	x_{1n}
A_2	x_{21}	x_{22}	…	x_{2n}
…	…	…	…	…
A_n	X_{n1}	x_{m2}	…	x_{mn}

Often, in the next step, the relative importance of criteria i.e. criteria scores needs to be assigned. The evaluation scores may come from ordinal, interval or ratio scales. In addition, normalization is required when different units of evaluation measures are applied. The relative importance of criteria involves the application of weighting methods. In general, weighting methods falls into two main categories: equal weight and rank-ordered weighting. It is obvious that the former ignores the relative importance. The latter can be classified into three categories: subjective, objective and combination weighting methods [167]. Weights determination is followed by determination of preference orders of criteria by applying MADM methods. There are a number of publications that review such methods [167]. It should be emphasized that weighting and prioritization methods can also be used in multi-objective optimization procedures.

Literature survey shows that analytical hierarchy process (AHP) is by far the most popular selection and prioritizing method. The objective stands at the top of the hierarchy and criteria and sub-criteria are decomposed into levels and sub-levels of the hierarchy. Thus decision alternatives appear at the bottom of the hierarchy. The ratio-scaled importance of alternatives is calculated through pair-wise comparison of evaluation criteria. It is possible to mix qualitative and quantitative criteria through AHP.

5.1.2 Multiple objective decision making

Multiple objective decision making (MODM) is another class of MCDM problems. In practice, it is not possible to predetermine the alternatives in many decision situations. Therefore, to specify the alternatives, decision makers (DM) attempt to concentrate on explicit mathematical relationships between decision variables. Hence, DM needs to formulate a set of objective functions to be optimized subject to a set of constraints. In other words, MODM deals with developing the best alternative from a large number of alternatives.

In certain cases, it is required to develop a subset of these alternatives and select the best solution through MADM methods. To name a few, multiple objective linear programming (MOLP), ε-constraint method and parametric variation of weights method can be used to generate the sub-set of alternatives. MOLP identifies non-dominated extreme points where decision variables have linear relationships.

The significant challenge in MODM is centered on defining suitable decision variables and the mathematical relationships between them. MODM is also known as multiple criteria mathematical programming (MCMP) or vector optimization problem. Objectives are defined to make improvements based on decision maker (DM) preferences. The thresholds or desired targets of objectives are called goals. One approach to deal with MODM problems is to deform it to a single objective problem by either combining all the objective functions or moving all but one objective to the constrain set [169]. In both cases, however, a single solution is returned. An MODM problem with *m* objectives is formulated as follows:

$$\text{Maximize } [f_1(x), f_2(x), \ldots, f_m(x)]$$

Subject to:

$$g_i(x) \leq 0,\ j=1,2,\ldots, n$$

where

$f_i(x)$ = i^{th} objective function, i=1,2,…, m

$g_j(x)$ = j^{th} constraint function, j=1,2,…, n

and $\mathbf{x} = [\, x_i \mid i=1,2,\ldots, k\,]$ is a k-dimensional decision variable vector in the solution space X (restricted by $g_i(x)$)

The details with respect to the general formulation of a multi-objective constrained problem are presented in [169, 170]. Here we seek a vector **x*** that maximize the set of

objective functions. From a decision making standpoint, the conflicting criteria would not lead to an optimal solution. For this, the aim is to uncover the best compromise solution that satisfies the DM. Thus, a feasible solution that meets or exceeds DM's goals is called satisfying solution. From this perspective, there can be a feasible solution (i.e. alternative) x_1 that is relatively equivalent to x_2 with respect to all criteria but it dominates x_2 solutions in at least one criterion. Moreover, there can be a feasible solution (x_1) that dominates all the other feasible solutions. Therefore, $x_1 \in X$ is a non-dominated solution if for all $x \in X$:

$$C_i(x_1) \geq C_i(x),\ i = 1,2,...,k;\ \text{and}\ C_j(x_1) > C_j(x),\ j \neq i.$$

In the MCDM literature, the terms "Pareto optimal set", and "efficient solution" refer to a set of all feasible non-dominated solutions. Pareto-optimal front refers to the corresponding objective function values of a given Pareto optimal set. Having a non-dominated solution, it is required to degrade the performance of at least one objective in order to improve the performance of another objective. An increase in the number of objective would result in increase in the size of Pareto-optimal set. Thus, it is intuitive to investigate a set of solutions rather than the whole Pareto optimal set. Moreover, it is desirable to apply algorithms with ability to find the Pareto optimal set with reasonable amount of iterations. In fact, classical optimization methods, such as weighted sum methods and goal programming, are not efficient in this respect. This fact emerges the development and application of evolutionary algorithms.

There is an increasing trend of using MCDM in power generation systems and also for PHM. Kusiak and Zheng [171] proposed a bi-objective power optimization model for a single power generator unit in which the power factor and the power output are improved by adjusting set points of blade pitch angle and generator torque. The dynamic process of a wind turbine is the prerequisite for optimizing the power produced and its quality. For this, the author applied the actual wind farm data to derive the required dynamic process models. Through simulation experiments, Munteanu et al. [172] studied the trade-off between the efficiency of energy conversion and input variability. Furuta et al. [173] considered improving safety level as one of the objectives in their bridge maintenance planning. Safety level, also known as durability level depends on the current safety level with respect to the

initial safety level. Orcesi and Frangopol [174] believed that short duration of monitoring programs can be a source for the error associated with the decision. Thus, the authors attempted to capture the uncertainties inherent to the structural degradation and decision processes in their probabilistic methodology.

5.1.3 Evolutionary optimization

Evolutionary optimization procedures refer to heuristic and meta-heuristic methods that are mainly used to solve combinatorial optimization problems. A number of evolutionary methods have been developed since the 1970s including: Genetic algorithms [169, 170, 174, 175], simulated annealing, particle swarms, differential evolution, tabu search, constraint propagation, artificial bee colonies, artificial immune systems, artificial neural networks, and Monte Carlo based-method. In essence, there is no distinctive algorithm with the ability to solve all the problems perfectly. However, in practice, multi-objective genetic algorithm (MOGA) is the most popular evolutionary method.

Genetic algorithms (GA) refer to a set of probabilistic search algorithm based upon natural selection and genetic operations. GA underlines combining information from good parents and provides approximate optimal solution. In principle, GA considers information as a series of binary string. The binary string format of a solution is known as chromosome. The chromosomes represent individuals' characteristics and they should have the same length. The set of solution is called population (i.e. a series of binary strings codes). In essence, to minimize the risk of being trapped in local maxima it is appropriate to use a population of solutions. Genetic operators such as crossover and mutation are utilized to obtain new population (i.e. the new generation of individuals). Crossover operation implies combining information from good parents and mutation means making changes in the code string of the new individuals. The outputs of GA include individuals who better fitted to their environment. Hence, a mathematical criteria i.e. fitness function is required to rank solutions.

In this study, the optimal solutions sets of the multi-objective optimization problem are determined using non-dominated sorting genetic algorithm-2 (NSGA-II) developed by deb

and his co-workers [176]. NSGA-II is based on the repetitive procedure of classifying the Pareto solutions to *n* Front. There are many derivatives for NSGA-II in the literature.

5.2 Using prognostics in maintenance decision making

The process of using prognostic information for life extension and safety improvement has recently become a favorite topic for investigation in the field of PHM. It seems significant amount of research needs to be undertaken to develop prognostic decision making programs with the ability to deal with distributed and interconnected systems in which various decision-making formulations are required for different subsystems. In this way, particular attention should be paid to the significance of information fusion and uncertainty managements in prognostic decision making. Moreover, it is worthwhile to consider the situations where no feasible solutions are found or where the Pareto optimal set does not help in achieving the performance objectives.

Balaban and Alonso [177] provides a review of the articles with respect to using prognostic information for fault-tolerant control and automated contingency management. The authors also outlined the key requirements for determining an aeronautic system's actions using prognostic information. This work applied probabilistic optimization methods and left the evolutionary algorithms for future works. Iyer and his coworkers [178] proposed a framework for human-centered decision support system (DSS) based upon prognostic indications. This system consists of a multi-objective decision support module in which optimal set of actions are selected via generating the corresponding Pareto Frontier (i.e. non-dominated alternatives).

In addition, there has been a tendency to calculate the Return on Investment (ROI) associated with the opportunities created by PHM. In the most recent publication, a methodology for determining ROI was proposed in [179] that considered uncertainties in both PHM performance and the associated costs e.g. unscheduled maintenance costs. Haddad et al. [180] believed that systems with prognostic capabilities have the potential to reduce life-cycle cost and optimize availability. The authors proposed an economic approach based on real options theory in order to quantify maintenance options. In this research, the decision maker has a set of options after a prognostic indication. This will give the decision

maker a flexibility to take actions in the period of the remaining operational life. The real options theory was utilized to monetarily manage the flexibility induced by prognostic considering the optimal time of maintenance action. In real options theory, the uncertainty parameter may be signified by a measure of the total value of an asset over the lifetime.

A central issue in human interventions is repair activities which can be divided into two classes: minor and major repairs. Minor repairs reduce the rate of aging and require short period of downtime e.g. routine services, oil change and alignments. On the other hand, major repairs have direct influence of the age of system. Normally major repairs special equipment and have longer process e.g. part replacement [87]. A prognostic-based life extension program should be able to plan and manage repair activities. Furthermore, the end-of-life threshold of repaired system would not be the same as the new one. There might be a need for a model to adjust the threshold after repair.

5.3 Problem formulation and methodology

This section attempts to elaborate the inherent features of the decision problem. The objective is to develop an interactive and computer aided techniques for aiding DM with application to a wind-powered electrical generators. In this way, the focus will be on formulating realistic models with conflicting and non-commensurable objectives to investigate a set of solutions. The primary function of a wind turbine is to convert wind kinetic energy into electrical energy within a definite speed range (i.e. cut-in and cut-out wind speed). In practice, wind turbines dynamics diminishes the power variations that come from the wind behavior. The various configurations and specific power dynamics of wind turbines will not be considered in this study. However, it is to be noted that almost all the wind turbines consist of the following components: aerodynamic rotor, transmission system, generator, controller, reactive power compensation, step-up transformer, and network [181]. It is evident that dynamic operation of the wind turbines is a function of the components performance and their interactions. Figure 5-1 shows the interaction between the components of a wind turbine. In this Figure, ω is the rotational speed, T represents the torque, U is the voltage, I refers to the current and f is the frequency. The aerodynamic rotor together with the mechanical transmission system is known as the aeroelastic part.

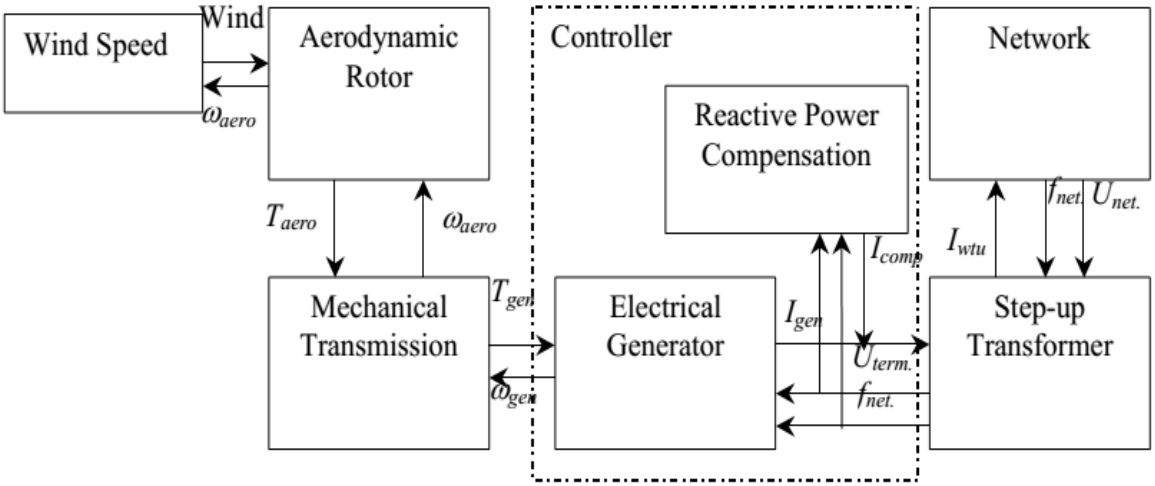

Figure 5-1 Interaction between the components of a wind turbine [181]

The wind energy captured by the wind turbine rotor can be expressed as:

$$P_r = 0.5\,\rho\,\pi\,R^2 C_p(\lambda, \beta_{pitch}) W^3 \qquad \text{Eq. 5-1}$$

where ρ is the air density, R is the rotor radius, W is wind speed before passing the rotor (also known as effective wind speed), and $C_p(\lambda, \beta)$ is the aerodynamic power coefficient and it is a normalized function of the blade pitch angle and tip-speed ratio ($\omega_{rotor} * R$), where ω_{rotor} represents the rotational speed of rotor then we have

$$\lambda = \frac{\omega_{rotor}}{W} R \qquad \text{Eq. 5-2}$$

The pitch control system controls the mechanical torque of wind turbine when the rotor speed exceeds the rated speed [182]. It is to be noted that the conventional wind turbines have a stall-regulated rotor with fixed pitch angle. The power coefficient can be defined as

$$C_p(\lambda, \beta) = \frac{P_{aerodynamic}}{P_{available}} \qquad \text{Eq. 5-3}$$

Hence, the aerodynamic power on the main shaft is (per unit):

$$P_{aerodynamic} = \frac{0.5\,\rho\,\pi\,R^2 C_p(\lambda,\beta_{pitch}) W^3}{P_{baserotor}} \qquad \text{Eq. 5-4}$$

It is to be noted that $A = \pi\,R^2$ is the area that swept by the blades, and it is known as the swept area or capture area. Here the $P_{baserotor}$ is the rated power of the rotor. Thus, the aerodynamic torque in per unit can be formulated as:

$$T_{aerodynamic} = \frac{P_{aerodynamic}}{\omega_{rotor}\, P_{baserotor}} = \frac{\rho\, \pi\, R^2 C_p(\lambda,\beta_{pitch}) W^3}{2\, \omega_{rotor}\, P_{baserotor}} \quad \text{Eq. 5-5}$$

The aerodynamic torque can be simplified as:

$$T_{aerodynamic} = 0.5\, \rho\, A\, \frac{C_p(\lambda,\beta_{pitch})}{\lambda} W^2 \quad \text{Eq. 5-6}$$

Hence, for a given wind speed, λ and β_{pitch} (i.e. the power coefficient) control the mechanical torque. The power coefficient can be approximated by [183]:

$$C_p(\lambda, \beta_{pitch}) = (0.44 - 0.0167\beta) \sin\left(\frac{\pi(\lambda-3)}{15-0.3\beta}\right) - 0.00184(\lambda-3)\beta \quad \text{Eq. 5-7}$$

The dynamic model of wind turbine can be described by the first order differential equation [172]:

$$\dot{\omega} = \frac{1}{J}(T_{aero} - C_d\omega - T_{genrator}) \quad \text{Eq. 5-8}$$

Here, $T_{genrator}$ is the generator torque and C_d is the damping.

5.3.1 Assumptions

The optimization model is formulated under the following assumptions:

- For the sake of simplicity, the focus of this section is mainly on a variable-speed power generator unit (PGU) which consists of a three-bladed rotor and a generator. The capacity of the PGU can be defined as:

$$C_{PGU} = \frac{\text{Amount of the generated electricity}}{\text{Amount of the generated electricity in full power run}} \quad \text{Eq. 5-9}$$

- The functional failures of the PGU include: (1) complete loss of energy generation capability; (2) over speeding; (3) partial loss of energy conversion capability.
- In fixed speed wind turbine, there is only one tip-speed ratio that can provide the maximum efficiency. However, the modern variable speed wind turbines intend to ideally provide maximum efficiency across a range of wind speeds. For this, the rotor speed can manipulated by a controller that takes into account the difference between generator torque and the aerodynamic torque. In this research, it is assumed that controlling the rotor speed is possible as one of the main decision variables.

- MODM methods provide interim results due to the existence of uncertainty.
- The monitoring and control systems ensure the ability of the system to fail well. In other words there is no failure with catastrophic consequences.
- Because of real-time applications, the computational times are negligible.
- In essence, the overall efficiency of the wind turbine is influenced by the losses of its various components. The bearing loss is mainly due to mechanical friction loss, and therefore, it depends on several parameters including the rotational speed and rotor weight. The bearing loss is usually very small and can be ignored.
- The alternatives for decision making include: (1) Stop the operation if degradation level is high and maintenance resources are available; (2) Let the operation continues until a planned time/age or until the availability of resources; (3) Continue operation with modified operating parameters and/or imperfect maintenance to extend life. This alternative depends on the goal for extension (e.g. meeting the demand), and the reduction in the degradation rate.
- We are allowed to stop the operation for imperfect maintenance practices that extend the life i.e. decrease the degradation/failure rate. The main condition to perform imperfect maintenance is

$$COST_{maintenance} <<< COST_{failure}$$

- Imperfect preventive maintenance should be routinely done. The routine imperfect PM can be with two outcomes: no change in the degradation/failure rate or decrease in the degradation/failure rate.
- The operational and environmental conditions have great impact on the rate of the degradation process.
- Wind speed will be utilized as the random input of the system. One can expect disturbances in the system inputs due to wind speed and load variations. The output of aerodynamic rotor is a function of the wind speed and on the rotor speed.
- There is no replacement (i.e. no renewal process).
- There is sufficient time to estimate and apply RUL (remaining useful life).

- The time horizon for life extension is from the time of fault detection to the time of failure i.e. end of life. Upon fault detection, imperfect corrective maintenance can be performed with the objective of life extension.
- It is possible to manipulate the variables for life extension
- Life extension is a kind of availability improvement. However, the objectives of a cost-effective life-extending program include: (1) increase in durability by reducing degradation with no significant reduction in performance; (2) performance enhancement. In this way, it is critical to the degradation or failure rate to the dynamic performance parameters.

5.3.2 Bearings

As discussed earlier in Chapter 2, roller bearings are critical parts of any machine that transmits power and motion. In essence, bearing has direct influence on the functionality of the machine. There are a number of formulas that can simplify the calculation of bearing life. In this study bearing life is expressed as:

$$L = \left(\frac{C}{P}\right)^n \qquad \text{Eq. 5-10}$$

where L is the expected lifetime of a bearing in millions of race revolution, C is the bearing load rating (i.e. dynamic load capacity), P is the equivalent load acting on the bearing, n is an exponent dependent on the type of bearing. In theory, C is the load to have a life of 1 million race revolutions with 90% reliability. This formula takes into account the applied load on a bearing. It is obvious that bearing lifetime is directly dependent the shaft speed and the applied load. There are two important modification factors that can be added to the bearing lifetime formula as expressed below:

$$L = a_f \, a_m \left(\frac{C}{P}\right)^n \qquad \text{Eq. 5-11}$$

where a_m is the modification factor that deals with manufacturing and material variations, and a_f is the modification factor of the life that represents the severity or size of defect. P can be calculated using equations 5-1 to 5-6.

To calculate a_f we need a model that measure and update the defect size using the degradation data. A deterministic model for bearing fatigue life is proposed by Li et al. [13, 184]. This crack propagation model is based on Paris's formula

$$\frac{da}{dN} = C_0(\Delta K)^n \quad \text{Eq. 5-12}$$

and relates the rate of defect growth to the instantaneous defect area *D* as follows:

$$\dot{D} = \frac{dD}{dt} = C_0(D)^n \quad \text{Eq. 5-13}$$

Here, C_0 and n are material constants. Accordingly, under constant operating condition, a deterministic crack propagation model in bearing relates the instantaneous defect area with material constants that have no correlation with the defect size. The bearing defect size at time can be estimated through the tome domain integration of the above equation:

$$\ln(D) = \alpha + \beta \ln(t + t_0) \quad \text{Eq. 5-14}$$

where t is the bearing running time, t_0 shows the time associated with the smallest defect area D_0.

$$t_0 = (\frac{C_0}{1-n})(D)^{n+1}$$

$$\alpha = \frac{1}{1-n} ln(\frac{C_0}{1-n})$$

$$\beta = \frac{1}{1-n}$$

$$\theta(t) = [\alpha\ \beta\ t_0]'$$

A recursive least square algorithm (RLS) can be used to update the vector of unknown parameters in the defect propagation process. The RLS is an adaptive estimation algorithm and can be expressed by the following equations [13, 73]:

$$e(t) = Y(t) - \hat{Y}(t, \hat{\theta}(t-1))$$

$$\Psi(t) = \frac{\hat{Y}(t, \hat{\theta}(t-1))}{d\theta} \Big| \theta = \hat{\theta}(t-1)$$

$$P(t) = \lambda^{-1}\left(P(t-1) - \frac{P(t-1)\Psi(t)\Psi(t)^T P(t-1)}{\lambda + \Psi(t)^T P(t-1)\Psi(t)}\right)$$

$$\hat{\theta}(t) = \hat{\theta}(t-1) - P(t)\Psi(t)e(t)$$

where *P(t)* is covariance matrix, and $\hat{Y}(t, \hat{\theta}(t-1))$ is the estimated value of $Y(t)$, and therefor, $e(t)$ represents the prediction error. Under varying operating condition the above equation can be modified to

$$\frac{dD}{dt} = C_0(D)^n \, e^{Z(t)}$$

where $e^{Z(t)}$ represents the amount of uncertainty in material properties or environmental factors. $Z(t)$ is a random variable which can be expressed by Gaussian-Markov Process with the following equation:

$$\dot{Z} = -\zeta \, Z(t) + w_z(t)$$

$w_z(t)$ is a Gaussian noise with zero mean.

5.3.3 Bearing Degradation Modeling

One requirement of the decision support system is the availability of degradation data to produce the relevant prognostic information. In fact, there is a need for changing the operating condition in the decision making process. This creates the need for a large database of degradation measures for each specific operations setting. For this reason, in this research, bearing degradation measures where simulated. The fault signal of bearings can be modeled through the combinations of the following parts [149]:

1. Repetitious impulses: A single defect in bearings generates a series of impulses of equal amplitude. The frequency of this faulty signal is equal to one the above mentioned frequency components. The impulses signals is normally represented by:

$$d(t) = d_0 \sum_{n=-\infty}^{n=\infty} \delta\left(t - \frac{n}{BPFO}\right)$$

Here, $\delta(t)$ denotes the Dirac function, and d_0 is the constant impulse amplitude. BPFO (Ball pass frequency outer race) represents one the major the major frequency components associated with a faulty bearing. For ball bearings the BPFO is calculated by:

$$BPFO = \frac{N_B}{2} f_s(1 - \frac{D_B \cos(\theta)}{D_P})$$

where D_B is the ball diameter, D_P represents the bearing pitch diameter, f_s is the shaft speed, and N_B is the number of balls. BPFO is the most common fault in bearings and it is obvious from the above formula that BPFO is a linear function of the shaft speed.

2. Load distribution: the formula below represents the non-uniform radial load on the around the circumference of a bearing

$$q(t) = q_0[1 - (1/2\varepsilon)(1 - \cos(\theta))]^n$$

3. Resonant vibration: $z(t) = \sum_{n=-\infty}^{n=\infty} \cos\left(2\pi f_r(t - \frac{n}{BPFO})\right)$
4. Exponential decay: $e(t) = \sum_{n=-\infty}^{n=\infty} e^{-\alpha(t-\frac{n}{BPFO})}$
5. Gaussian noise *n(t)*: a basic noise model to simulate the effects of random processes.

Finally, the signal model is denoted as

$$w(t) = [d(t)q(t)z(t)] * e(t) + n(t)$$

Based on the above simulation the waveform of the typical faulty bearing signal is shown in Figure 5-2.

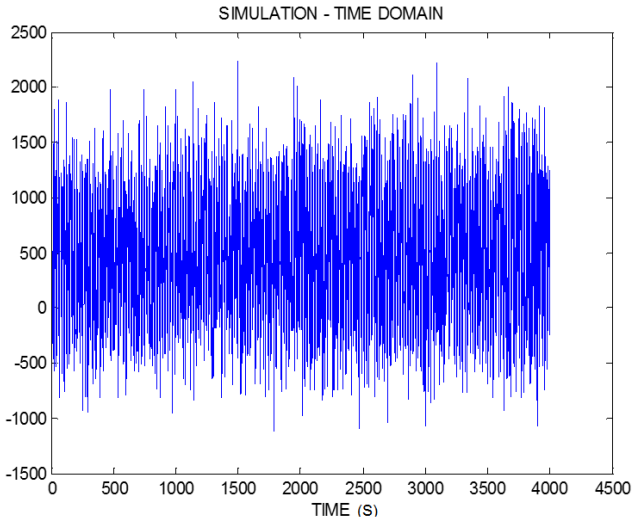

Figure 5-2 Simulation of bearing degrdation

However, this simulation approach is computationally intensive. Thus, the second approach was utilized which approximate the elements of the degradation signal as listed below [150]:

1. Repetitious impulses: $X_f(t) \approx A_0 \sum_{n=1}^{N_f} cos\left(2\pi n f_l - \emptyset_f^n\right)$
2. Load: $X_q(t) \approx \sum_{n=1}^{N_q} A_q^k cos\left(2\pi n f_q t + \emptyset_q^n\right)$
3. Bearing Induced Vibration: $X_{bs}(t) \approx \sum_{n=1}^{N_{bs}} A_{bs}^k e^{-\alpha n t} cos\left(2\pi f_q t + \emptyset_{bs}^n\right)$
4. Machinery Induced Vibration: $X_{bs}(t) \approx \sum_{n=1}^{N_{bs}} A_{bs}^k e^{-\alpha n t} cos\left(2\pi f_q t + \emptyset_{bs}^n\right)$
5. Noise*: n(t)*

Finally, the fault signal is represented by

$$x(t) = X_f(t).X_q(t).X_{bs}(t) + X_s(t) + n(t)$$

Figure 5-3 shows the RMS of simulated degradation data for 25 bearings. The effects of shaft frequency on degradation measures are displayed in Figure 5-4. In addition, the effects of change in the speed shaft is illustrated in Figure 5-5. This Figure is similar to the experimental results presented by Gebraeel and Pan [109](see Figure 3-3).

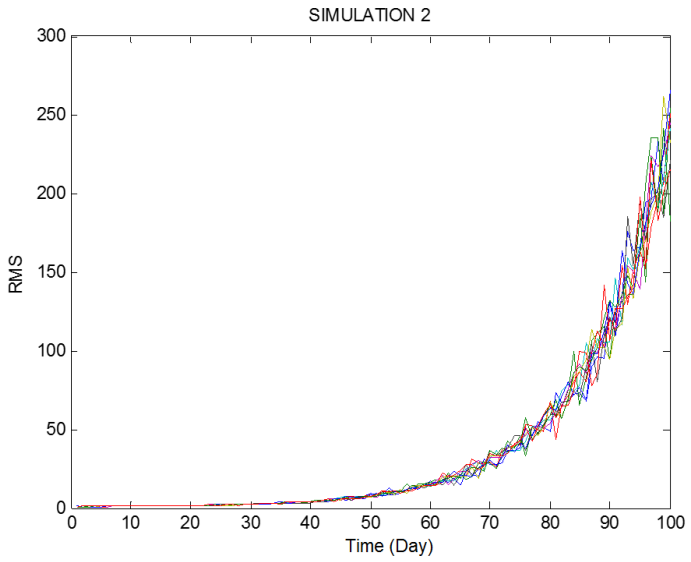

Figure 5-3 Simulated degradation data for 25 bearings

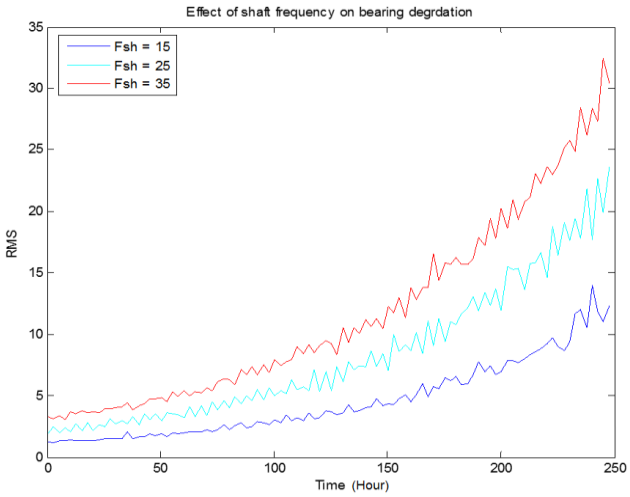

Figure 5-4 Degradation pattern for three different shaft speeds

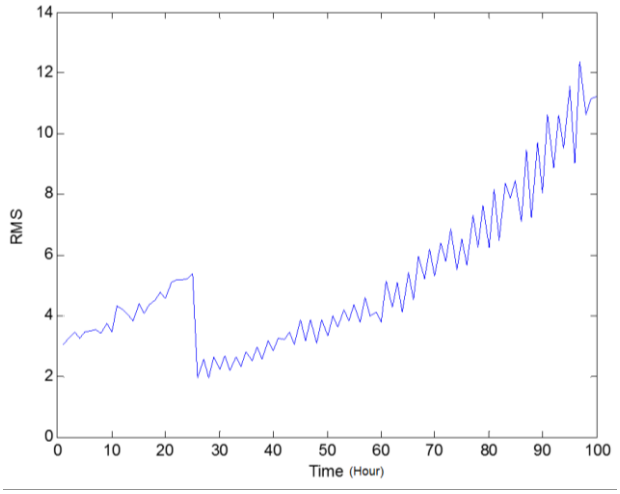

Figure 5-5 Effects of change in the speed shaft

5.3.4 Wind model

Wind speed variation is the major cause of fluctuating power in the conventional wind turbines [181, 185]. A wind model is divided into deterministic and stochastic parts as shown in Figure 5-6. The former denotes the wind speed profile which is assumed to be constant in short periods of time (e.g., 10 minutes).

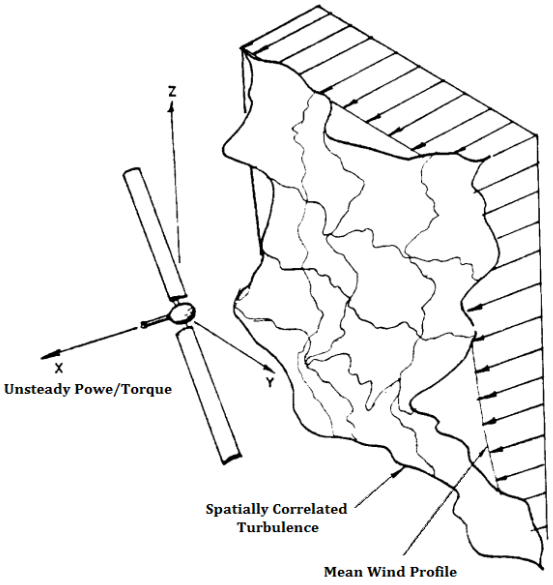

Figure 5-6 Illustration of wind speed variation on the rotor area [185]

The stochastic part represents the spatial turbulence on the rotor area. Through knowledge of turbulence fluctuations is essential for the design of pitch control systems. The effects of spatial turbulence on power variation become even more significant in large diameter rotors. The wind turbine module provided by the Princeton Satellite Systems was used to simulate wind data [186]. The codes developed based on the model provided by Rosas in [181]. Figure 5-7 shows the typical stochastic wind data. It should be noted that the wind speed has been also modelled by three parameters Weibull distribution [187].

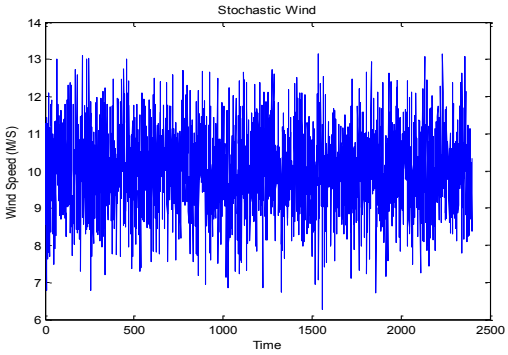

Figure 5-7 Typical stochastic wind data

5.3.5 Decision variables & Objective function

Using the average wind speed for each day is a common practice in industry. For this, the unit of time in this optimization model is a single day. It is important to mention that the optimization model considers the period of time between the fault detection and the end of life (EOL) i.e. the time that system needs to be shut down. The EOL is estimated through the prognostic algorithms. Thus, the decision making for life extension can be performed on any day in this period. The optimization model needs to have the revenue from power generation from the decision making day (DMD) to the EOL. It is assumed that wind speed, and other important parameters are available for each day in the period of DMD-EOL (i.e. DMD to DMD+RUL). Hence, the optimization model attempts to consider the decision variables to provide optimal values for service life and the profit from generating power in the period of DMD-EOL. It is to be noted that because of change in the variables, the EOL needs to be estimated for each set of variables. The data for the prognostic algorithm can be provided by the simulation methods explained in Section 5.3.3.

The bi-objective model in this study is defined as follows:

$$\textbf{max } \boldsymbol{Profit = Revenue - COST_{O\&M}}$$

$$\textbf{max } \boldsymbol{Service\ Life}$$

$$\boldsymbol{Revenue_i = RUL_i * WT_{PR} * \frac{C_f}{100} * C_{EH}, \quad DMD \leq i \leq EOL}$$

$$\boldsymbol{L = a_{LE}\, a_f\, a_m\, (\frac{C}{P})^n}$$

The decision variables are as follows:

- Rotor Speed [e.g. 15-35 RPM]
- Pitch Angle [e.g. 20-40]
- Time (day) of maintenance [e.g. between day 55 and day 85]
- Duration of maintenance [e.g. 1-10] (i.e. determines the effects of maintenance on life extension

Net profit can be simply defined as the difference between the revenue generated from electricity sale and the operation and maintenance (O&M) costs. The cost of maintenance is a combination of several different elements including: cost of material, cost of labor, cost of access, cost of production loss during maintenance, and cost of prognostic and health management. Each element of the maintenance cost can be subdivided into several units. For instance, some details such as cost of transportation, cost of loading/unloading, and cost of a crane might be added to the cost analysis. Furthermore, there are some variables in the cost analysis that may come from predefined distributions. For example, cost of labor and total cost of production loss during downtime are function of total working hour which can be approximated by the expected time to repair. This cost element can be further elaborated if one wants to consider the cost of applying different maintenance types.

The penalty of not satisfying the demand is another element of operations cost. However, demand for electricity can be also used as a constraint. The figure blow shows the pattern of daily demand based on the data provided by National Grid [188]. The data is scaled to provide better insight.

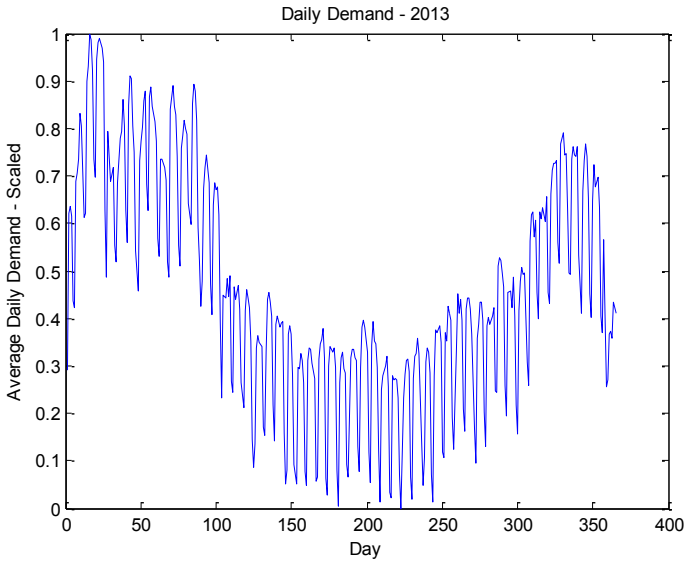

Figure 5-8 Pattern of daily demand

5.3.6 Results and Discussion

This section summarizes the results of applying prognostic information for decision making with application to a wind-powered PGU. This section applies the models developed in the previous sections particularly the optimization model from section 5.3.5. Figure 5-9 provide the typical pattern of stochastic wind data and the associated power generated by the PGU's rotor. In Figure 5-10, the effects of modification factor on the bearing life is illustrated.

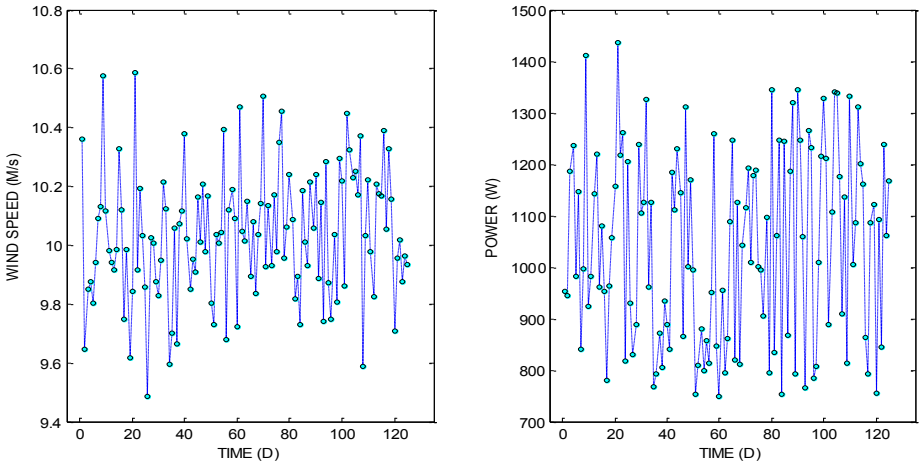

Figure 5-9 Stochastic wind the associated generated power

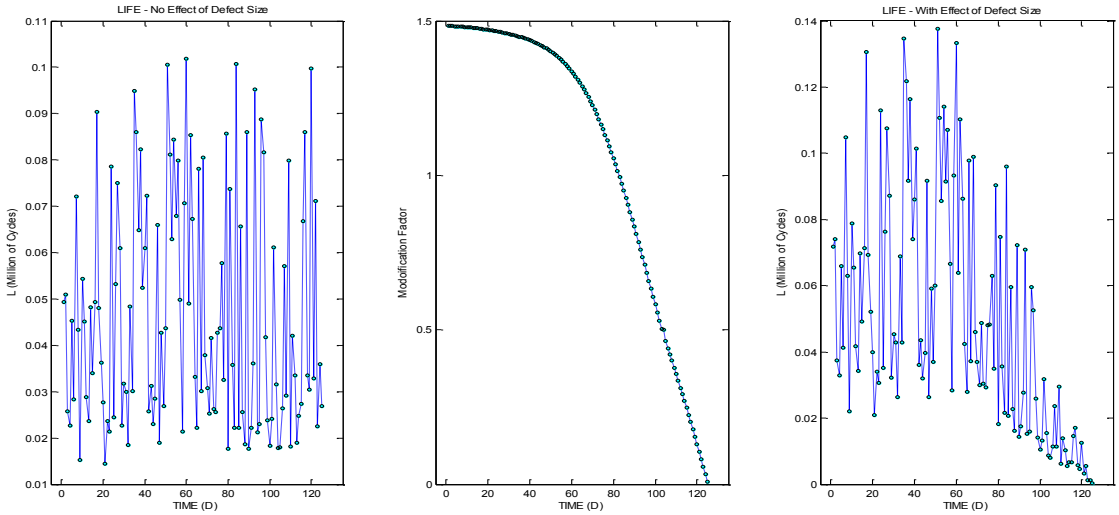

Figure 5-10 Effects of modification factor on bearing life

Although the power applied on the bearings is a function of the stochastic wind speed, considering the defect size can ensure the right trend for the life of bearing. Similarly, the estimated revenue can be obtained as shown in Figure 5-11. Finally, the result of NSGA optimization is depicted in Figure 5-12.

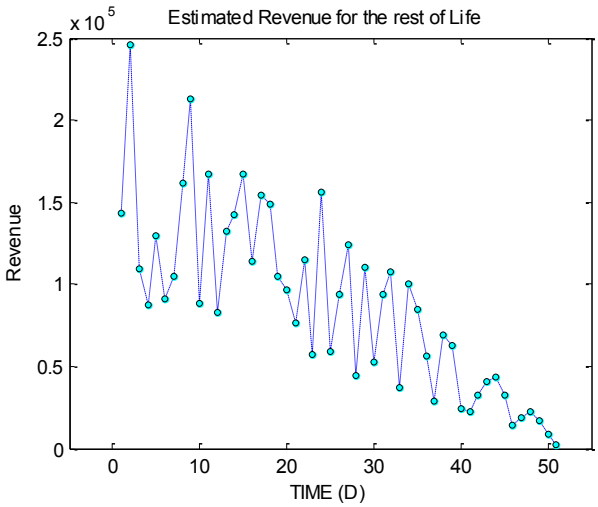

Figure 5-11 Estimated revenue

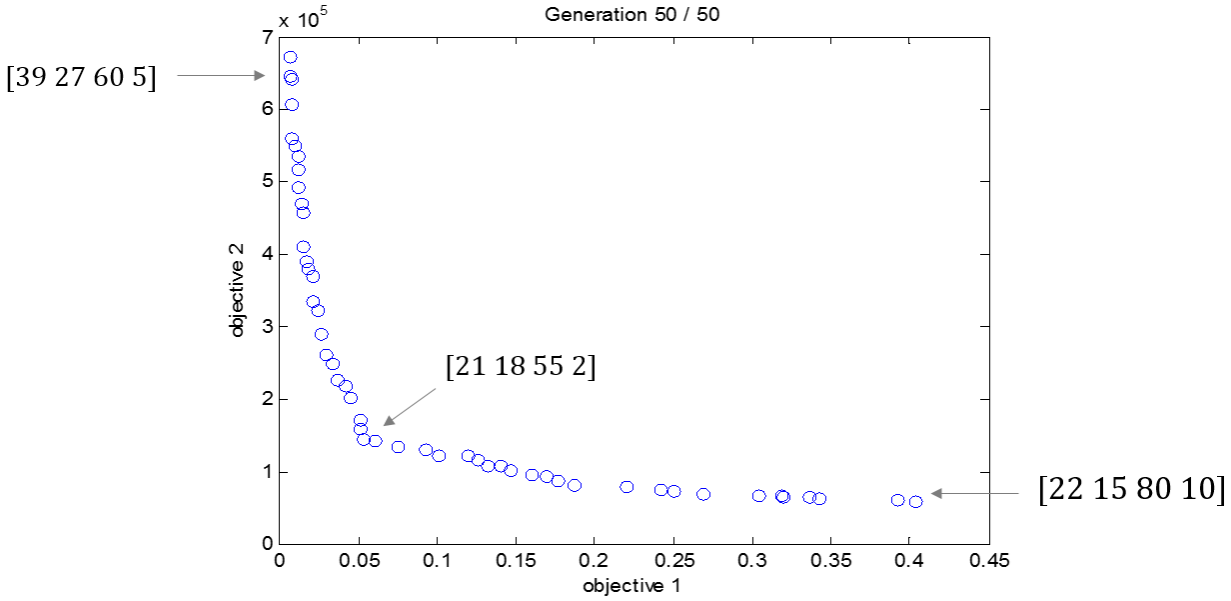

Figure 5-12 NSGA optimization results

Figure 5-12 shows the confronting nature of objectives. Obviously, in order to increase the profit (i.e. objective 2) we need to accept less bearing life which is represented by objective one. The numbers in the bracket represent pitch angle, shaft speed, the suggested time (day) of maintenance, and the proposed duration of maintenance. Accordingly, by increasing the shaft speed the profit increases with the expectation of shorter life for bearings. In addition, higher pitch angle and less maintenance are required for higher profit. Interestingly, the maximum life extension needs the minimum speed and the highest value of the maintenance days close to the EOL.

5.3.7 Uncertainty Considerations

Uncertainty management is an integral component in all aspects of PHM. Uncertainties in PHM arise from imperfections in predictability of prognostic models. The failure in predictability is a function of several factors including type and quality of data, inherent assumptions in the model, level of computations, and understanding of underlying physical processes. The uncertainty estimation and quantification for PHM applications has been extensively considered in the literature [189-199]. Baraldi et al. [190] classified the sources of uncertainties in PHM into three categories: 1. randomness related to future degradations, 2. errors in modeling, and 3. inaccuracies in degradation data. All these sources are significant when it is required to use the outputs of a PHM system for maintenance decision making.

In essence, a large uncertainty bound is a risk in the application of RUL estimates in decision making. In practice, uncertainty would be higher at the early stage of life. On the other hand, as we get close to the EOL, the uncertainty is low but the risk for unanticipated failure in high. Therefore, there is a need to identify decision making horizon. Figure 5-13 shows the examples of degradation measures where two degradation path were selected for uncertainty analysis. Figure 5-14 clearly shows the need for a decision making horizon. In this figure the RUL is estimated for the selected degradation paths (i.e. unit 1 and unit 2). Accordingly, by getting close to the EOL the uncertainty decreases namely the actual RUL and the estimated RUL get closer.

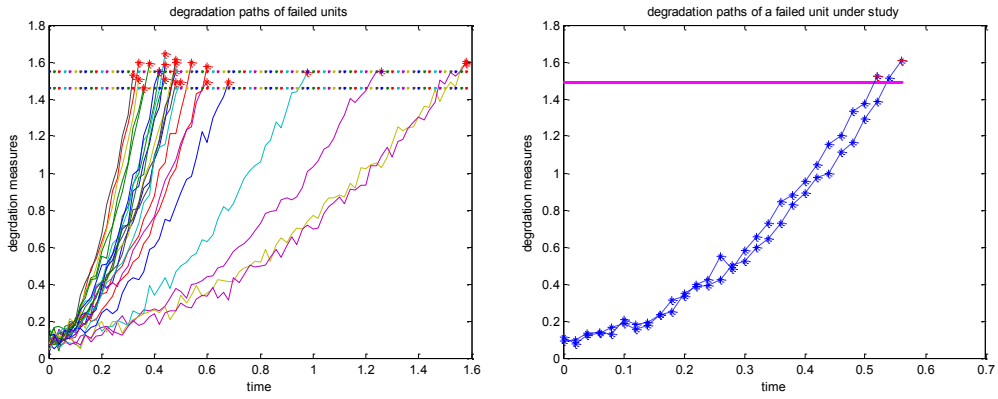

Figure 5-13 Examples of degradation measures

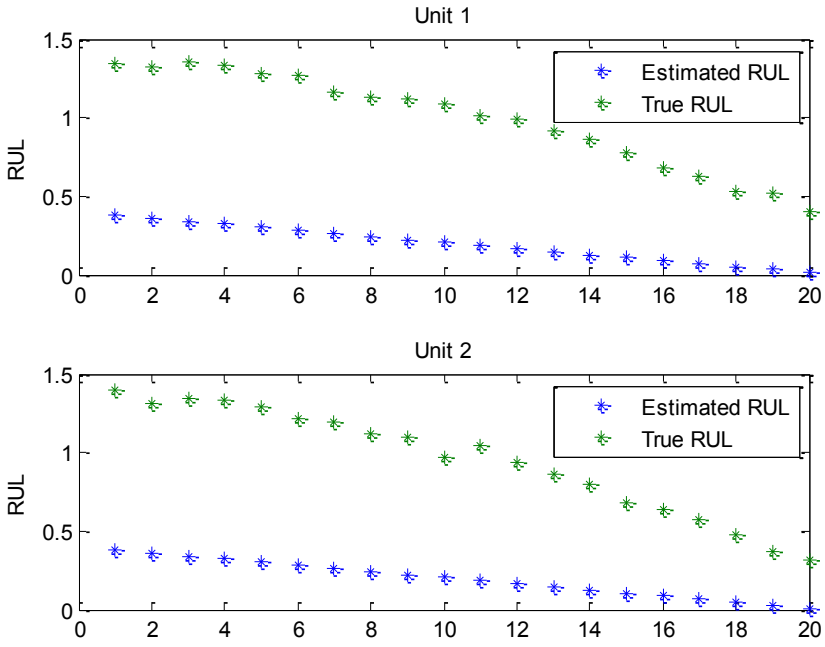

Figure 5-14 Estimation of remaining life as getting close to the EOL

6 CONCLUSIONS AND FUTURE WORKS

6.1 Conclusions

Applying prognostics for extending the operable life of engineering systems has been noticeably lacking in published literature. This research investigates the use of prognostic data to mobilize the potential residual life. The obstacles in performing life extension include: lack of knowledge (e.g. about aging mechanisms or life-limiting factors), lack of tools such as trust worthy prognosis algorithm, lack of data (e.g. an acceptable RUL estimation), and lack of time. The first and foremost step in an efficient life-extending methodology is effective and robust diagnosis, which detects and identifies the impending faults. In this way, the concerns in successful application of proactive on-line monitoring include sensor calibration, quantification of uncertainty, sensor advancement, better understanding of physical aspects, wireless data transferring, enhanced data interrogation, data fusion for multichannel monitoring data, design of monitoring experiments, recognition of failure modes with similar patterns, monitoring small components in noisy areas, and controlling simultaneous multiple failure.

In Chapter 3 the possibility of non-monotonic degradation measures was proven experimentally (unbalance and FCC experiments). This research primarily considers using acoustic emission (AE) technology for quick-response diagnostic. To be specific, an important feature of AE data is statistically modeled to provide quick, robust and intuitive diagnostic capability. The proposed ZIP model was successful to detect the out of control situation when the data of faulty bearing was applied. This diagnostic techniques can be further improved by considering pattern recognition methods to detect the type of fault in the system. Moreover, it was proved through categorical data analysis that couple unbalance is the most significant types of unbalance. This research also highlights the importance of self-healing materials. It seems promising research on self-healing materials would affect the current methods of PHM in near future.

One main component of the proposed life extension framework is the trend analysis module. This module analyzes the pattern of the time-ordered degradation measures. In many cases trend analysis is an essential part of PHM modeling for both ensuring

monotonicity and discovering the improvements in the degradation signal after maintenance or self-healing. Trend analysis for prognostic parameters, degradation waveform and multivariate data are considered in this study. It was concluded that graphical methods are appropriate for trend detection of signal features. Particularly, cumulative plots are useful for certain features, e.g. energy-related features. Analyzing the waveform might be a better option before performing the feature extraction phase. HHT is very useful in noise removal and extracting the main components of the signal. Those components can be used for trend analysis and feature extraction. In this respect, the modified crest factor worked appropriately in analyzing and comparing the waveforms. For multivariate data, it was realized that PCA is able to indicate the trends in the data. Preliminary analysis of the data including smoothing, filtering, elimination the correlated variables were very helpful in proper implementation of PCA. In addition, two algorithms are introduced to deal with a non-monotonic trend: ACD and TBSM. Obviously, the usefulness of these algorithms is case-dependent. However, it seems, both algorithms have the potential to treat the non-monotonicity in degradation data.

Although considerable research has been devoted to developing prognostics algorithms, rather less attention has been paid to post-prognostic issues such as maintenance decision making. All the activities in the diagnostic and prognostic modules need to be followed by fruitful decision making. A multi-objective optimization model is presented for a power generation unit. This model proves the ability of prognostic models to balance between the power generation and life extension.

The major step in decision making is to determine alternatives which are defined as possible courses of action. Thus, an optimization procedure is required to clarify the parameters of these alternatives. Another concern in developing the alternatives is to find the optimal time to perform the actions. In effect, performance of maintenance at inconvenient times results in substantial time and economics losses. Nevertheless, despite the confronting nature of objectives, the proposed optimization model provides acceptable results. It was discovered that by increasing the shaft speed the profit increases with the expectation of shorter life for bearings. Interestingly, the maximum life extension needs the minimum speed and the highest value of the maintenance days close to the EOL.

6.2 Recommendations for future works

The followings are the recommended future works in various aspects of prognostic-based life extension:

1. System-level frameworks are in demand to be applied for a variety of mechanical systems. Current methods are suited for component RUL estimation. Moreover, challenging research is needed in order to handle the situations with overlapping failure mechanism. Moreover, integration of the qualitative methods such as failure mode and effects analysis (FMEA) or fault tree analysis (FTA) into the quantitative diagnosis analysis is desirable.
2. Part of the future work with respect to the FCC experiments is providing in depth discussion on the possible difference in the characteristic signal in various stage of life. Preliminary results show that there are different fundamental frequencies in the first two signal components for various stages of crack propagation. Further analysis are required to link the signal features with various stages of crack propagation particularly for the period of retardation.
3. In essence, uncertainty continues to be a major concern for the fruitful application of prognostic models. In this respect, assumptions and simplifications can be significant sources of uncertainty in prognostic models. Monotonic fault progression is an important assumption used in a number of prognostic systems. This assumption can be violated through human intervention or self-healing. Hence, non-monotonic degradation will appear in the modeling process which not only increases the uncertainty but also may cause model failure in some situations. Trend analysis techniques are able to provide a platform for uncertainty management prior to applying the degradation data in prognostic models. In addition, a large uncertainty bound is a risk in the application of RUL estimates in decision making. In practice, uncertainty would be higher at the early stage of life. The uncertainty seems to be lower close to the EOL but the risk for unanticipated failure in high. Therefore, there is a need to identify the optimal decision making horizon.
4. By obtaining an acceptable level of diagnosis and prognosis in the system, it is possible to take full advantage of maintenance activities and ensure better decision making. Many

researchers have focused on the maximizing the power capture from the wind through the optimization of wind turbine control parameters such as set points of blade pitch angle, and generator torque. This research intends to simultaneously maximize the unit life and power generation through multi-objective optimization by applying prognostic information. The followings list the future works for the proposed multi-objective optimization model:

- It is always difficult for the DM to pick one best solution from a large set of alternatives. Therefore, more constraint can be added to the MCDM problem.
- For a superior decision making, the amount of infeasibility and the number of violated constraints need to be taken into consideration.
- The safety level can be considered as an objective function for the critical components. In particular, it is pertinent to mention that waiting to do the maintenance close to the end of life increases the risk of unanticipated system failure and intimidates the system safety.
- This decision making problem can be extended to a level that DM selects a portion of a solution and ask the system to provide exclusive feasible Pareto frontier [178]. Furthermore, sensitivity analysis needs to done mainly to improve the acceptability of the overall decision making model.
- To enhance the models performance it is essential to consider dependent competing risk (e.g. the combination of degradation wear and random shocking).
- Power quality can be considered as a constraint when the operating parameters need to be manipulated. There are several metrics to measure power quality of wind turbine such as power factor, reactive power, and harmonic distortion. Power Factor (PF): is defined as the ratio of real power over apparent power. It provides a measure for the efficiency of the PGU. The goal is to preserve a power factor of 1.

REFERENCES

1. DOE, *20% Wind Energy by 2030 Increasing Wind Energy's Contribution to U.S. Electricity Supply*. 2008.
2. Xiang, J. and S. Watson, *Assessment of Condition Monitoring Techniques for Offshore Wind Farms.* Solar Energy Engineering, 2008. **130**(3): p. 031004-1.
3. Bond, L.J. and R.M. Meyer. *Online Monitoring to Enable Improved Diagnostics, Prognostics and Maintenance*. in *ICI2011*. 2011. Daejeon,Korea.
4. Lu, B., et al., *A review of recent advances in wind turbine condition monitoring and fault diagnosis*, in *Power Electronics and Machines in Wind Applications*. 2009: Lincoln, NE p. 1-7.
5. Hameed, Z., et al., *Condition monitoring and fault detection of wind turbines and related algorithms: A review.* Renewable and Sustainable Energy Reviews, 2009. **13**(1): p. 1-39.
6. Hameed, Z., S.H. Ahn, and Y.M. Cho, *Practical aspects of a condition monitoring system for a wind turbine with emphasis on its design, system architecture, testing and installation.* Renewable Energy, 2010. **35**(5): p. 879-894.
7. Kusiak, A. and W. Li, *The prediction and diagnosis of wind turbine faults.* Renewable Energy, 2011. **36**(1): p. 16-23.
8. Ravindran, A.R., *Operations Research and Management Science Handbook*. 2007, Boca Raton, FL: CRC Press.
9. Yao, X., et al., *Wind Turbine Gearbox Fault Diagnosis Using Adaptive Morlet Wavelet Spectrum.* 2009: p. 580-583.
10. Huang, N.E., et al., *The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis.* Proceedings of the Royal Society of London, 1998. **454**: p. 903-995.
11. Yu, D., J. Cheng, and Y. Yang, *Application of EMD method and Hilbert spectrum to the fault diagnosis of roller bearings.* Mechanical Systems and Signal Processing, 2005. **19**: p. 259-270.
12. Lei, Y. and M.J. Zuo, *Fault diagnosis of rotating machinery using an improved HHT based on EEMD and sensitive IMFs.* Measurement Science and Technology, 2009. **20**(12): p. 1-12.

13. Li, Y., et al., *Dynamic prognostic prediction of defect propagation on rolling element bearings.* Tribology Transactions, 1999. **42**(2): p. 385-392.
14. Vlcek, B.L., R.C. Hendricks, and E.V. Zaretsky, *Determination of Rolling-Element Fatigue Life From Computer Generated Bearing Tests*. 2003, NASA: Cleveland, Ohio.
15. ISO, *16281: Rolling bearings - Methods for calculating the modified reference rating life for universally loaded bearings*. 2008.
16. Zaretsky, E.V., J.V. Poplawski, and S.M. Peters, *Comparison of Life Theories for Rolling-Element Bearings.* Tribology Transactions, 1996. **39**(2): p. 237-248.
17. Salehizadeh, H. and N. Saka, *Crack Propagation in Rolling Line Contacts.* Journal of Tribology, 1992. **114**(4): p. 690-697.
18. Hoeprich, M.R., *Rolling Element Bearing Fatigue Damage Propagation.* Journal of Tribology, 1992. **114**(2): p. 328-333.
19. Ragheb, A. and M. Ragheb. *Wind turbine gearbox technologies*. in *International Nuclear and Renewable Energy Conference*. 2010. Amman, Jordan.
20. Bartelmus, W. and R. Zimroz, *Vibration condition monitoring of planetary gearbox under varying external load.* Mechanical Systems and Signal Processing, 2009. **23**: p. 246-257.
21. Aslantas, K. and S.Tasgetiren, *A study of spur gear pitting formation and life prediction.* Wear, 2004. **257**(11): p. 1167-1175.
22. Toutountzakis, T., C.K. Tan, and D. Mba, *Application of acoustic emission to seeded gear fault detection.* NDT&E International, 2005. **38**: p. 27-36.
23. ISO, *Gears - Wear and damage to gear teeth - Terminology*. 1995.
24. Tandon, N. and S. Mata, *Detection of Defects in Gears by Acoustic Emission Measurements.* Journal of Acoustic Emission, 1999. **17**(1-2): p. 23-27.
25. Tan, C.K. and D. Mba, *Identification of the acoustic emission source during a comparative study on diagnosis of a spur gearbox.* Tribology International, 2005. **38**: p. 469-480.
26. Večeř, P., M. Kreidl, and R. Šmíd, *Condition Indicators for Gearbox Condition Monitoring Systems.* Acta PolytechnicaVol. 45 No. 6/2005, 2005. **45**(6): p. 35-43.

27. Bartelmus, W. and R. Zimroz, *A new feature for monitoring the condition of gearboxes in non-stationary operating conditions.* Mechanical Systems and Signal Processing, 2009. **23**: p. 1528-1534.
28. Howards, I., S. Jia, and J. Wang, *The Dynamic Modelling of a Spur Gear in Mesh Including Friction and a Crack.* Mechanical Systems and Signal Processing, 2001. **15**(5): p. 831-853.
29. Mba, D. and R.B.K.N. Rao, *Development of Acoustic Emission Technology for Condition Monitoring and Diagnosis of Rotating Machines; Bearings, Pumps, Gearboxes, Engines and Rotating Structures.* The Shock and Vibration Digest, 2006. **38**(1): p. 3-16.
30. Eftekharnejad, B. and D. Mba, *Seeded fault detection on helical gears with acoustic emission.* Applied Acoustics, 2009. **70**: p. 547-555.
31. Tana, C.K., P. Irvinga, and D. Mba, *A comparative experimental study on the diagnostic and prognostic capabilities of acoustics emission, vibration and spectrometric oil analysis for spur gears.* Mechanical Systems and Signal Processing, 2007. **21**: p. 208-233.
32. Price, E.D., A.W. Lees, and M.I. Friswell, *Detection of severe sliding and pitting fatigue wear regimes through the use of broadband acoustic emissio.* Journal of Engineering Tribology, 2005. **219**(2): p. 85-98.
33. Tan, C.K. and D. Mba, *A Correlation Between Acoustic Emission Activity and Asperity Contact During Meshing of Spur Gears Under Partial Elastohydrodynamic Lubrication.* Tribology Letters, 2005. **20**(1): p. 63-67.
34. Tan, C.K. and D. Mba, *Limitation of Acoustic Emission for Identifying Seeded Defects in Gearboxes.* Journal of Nondestructive Evaluation, 2005. **24**(1): p. 11-28.
35. Badi, M.N.M., S.N. Engin, and D. Schonfeld, *Fault classification of a model drive-line using time domain data*, in *COMADEM*. 1996: Sheffield, UK.
36. Al-Balushi, K.R. and B. Samanta, *Gear fault diagnosis using energy-based features of AE signals.* Journal of Systems and Control Engineering, 2002. **216**: p. 249-264.
37. Al-Balushi, K.R. and B. Samanta, *Gear Fault Diagnostics using Wavelets and Artificial Neural Networks*, in *COMADEM*. 2000.

38. Sentoku, H., *Acoustic Emission in Tooth Surface Failure Process of Spur Gears.* Journal of Acoustic Emission, 1998. **16**(1-4): p. 19-24.
39. Singh, A., D.R. Houser, and S. Vijayakar, *Early detection of gear pitting* in *ASME Power transmission and gearing conference*. 1996. p. 673-678.
40. He, D., R. Li, and E. Bechhoefer, *Split Torque Type Gearbox Fault Detection using Acoustic Emission and Vibration Sensors*, in *International Conference on Networking, Sensing and Control (ICNSC)*. 2010: Chicago.
41. Raad, A., et al., *On the comparison of the use of AE and vibration analysis for early gear fault detection*, in *The 8th Western Pacific Acoustics Conference*. 2003: Melbourne, Australia.
42. Yoshioka, T. and T. Fujiwara, *New acoustic emission source locating system for the study of rolling contact fatigue.* Wear, 1982. **81**(1): p. 183-186.
43. Physical Acoustic Corporation, *PCI-2 Based AE System*. 2007, Mistrass.
44. Suman, A., et al. *Study on amplification of acoustic emission signals during tensile deformation of Aluminium*. 2007.
45. Miyachika, K., S. Oda, and T. Koide, *Acoustic Emission of Bending Fatigue Process of Spur Gear Teeth.* Journal of Acoustic Emission, 1995. **13**(1-2): p. 47-53.
46. Singh, A., D.R. Houser, and S. Vijayakar, *Detecting gear tooth breakage using acoustic emission: a feasibility and sensor placement study.* Journal of Mechanical Design, 1999. **121**: p. 587-593.
47. Hamzah, R.I.R. and D. Mba, *The influence of operating condition on acoustic emission (AE) generation during meshing of helical and spur gear.* Tribology International, 2009. **42**: p. 3-14.
48. Vinogradov, A., *On shear band velocity and the detectability of acoustic emission in metallic glasses.* Scripta Materialia, 2010. **63**(1): p. 89-92.
49. Tandon, N. and B.C. Nakra, *Defect Detection of Rolling Element Bearings by Acoustic Emission Method.* Journal of Acoustic Emission, 1990. **9**(1): p. 25-28.
50. Morhain, A. and D. Mba, *Bearing defect diagnosis and acoustic emission.* Journal of Engineering Tribology, 2003. **217**(4): p. 257-272.

51. Li, X., *A brief review- acoustic emission method for tool wear monitoring during turning.* International Journal of Machine Tools & Manufacture, 2002. **42**: p. 157-165.
52. Sharma, V.S., S.K. Sharma, and A.K. Sharma, *An approach for condition monitoring of a turning tool.* Journal of Engineering Manufacture, 2007. **221**: p. 635-646.
53. Quadroa, A.L. and J.R.T. Brancob, *Analysis of the acoustic emission during drilling test*, in *24th International Conference on Metallurgical Coatings and Thin Films*. 1997. p. 691-695.
54. Carpinteri, A., G. Lacidogna, and N. Pugno, *Structural damage diagnosis and life-time assessment by acoustic emission monitoring.* Engineering Fracture Mechanics, 2007. **74**: p. 273-289.
55. ISO, *13381: Condition monitoring and diagnostics of machines - Prognostics, General guidelines* 2004.
56. Sheppard, J.W., M.A. kaufman, and T.J. Wilmer, *IEEE standards for prognostics and health management.* IEEE Aerospace and Electronic Systems Magazine, 2009. **24**(9): p. 34 - 41.
57. Jardine, A., D. Lin, and D. Banjevic, *A review on machinery diagnostics and prognostics implementing condition-based maintenance.* Mechanical Systems and Signal Processing, 2006. **20**(7): p. 1483-1510.
58. Lee, J., et al., *Intelligent prognostics tools and e-maintenance.* Computers in Industry, 2006. **57**(6): p. 476-489.
59. Heng, A., et al., *Intelligent condition-based prediction of machinery reliability.* Mechanical Systems and Signal Processing, 2009. **23**(5): p. 1600-1614.
60. Heng, A., et al., *Rotating machinery prognostics: State of the art, challenges and opportunities.* Mechanical Systems and Signal Processing, 2009. **23**(3): p. 724-739.
61. Wu, W., J. Hu, and J. Zhang, *Prognostics of Machine Health Condition using an Improved ARIMA-based Prediction method*, in *IEEE Conference on Industrial Electronics and Applications*. 2007. p. 1062-1067.
62. Bond, L.J., et al., *Damage assessment technologies for prognostics and proactive management of materials degradation.* Nuclear Technology, 2011. **173**: p. 46-55.

63. Farrar, C.R. and N.A.J. Lieven, *Damage prognosis: the future of structural health monitoring.* Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 2007. **365**(1851): p. 623-632.
64. Farrar, C.R., et al., *Damage Prognosis - Current Status and Future Needs*. 2003, Los Alamos National Labratory.
65. Sikorska, J.Z., M. Hodkiewicz, and L. Ma, *Prognostic modelling options for remaining useful life estimation by industry.* Mechanical Systems and Signal Processing, 2011. **25**(5): p. 1803-1836.
66. Si, X.-S., et al., *Remaining useful life estimation – A review on the statistical data driven approaches.* European Journal of Operational Research, 2011. **213**(1): p. 1-14.
67. Hines, J.W. and A. Usynin, *Current computational trends in equipment prognostics.* International Journal of Computational Intelligence Systems, 2007. **1**(1): p. 94-102.
68. Dragomir, O.E., et al., *Review of prognostic problem in condition-based maintenance*, in *European Control Conference*. 2009.
69. Bond, L.J. and R.M. Meyer. *Online monitoring to enable improved diagnostics prognostics and maintenance*. in *ICI*. 2011. Daejeon, Korea.
70. Greitzer, F.L., et al., *Development of a framework for predicting life of mechanical systems: Life Extension Analysis and Prognostics (LEAP)*, in *International Society of Logistics*. 1999 Las Vegas, Nevada.
71. Greitzer, F.L. and T.A. Ferryman, *Predicting Remaining Life of Mechanical Systems*, in *Intelligent Ship Symposium IV* 2001.
72. C. Tan, P. Irving, and D. Mba, *A comparative experimental study on the diagnostic and prognostic capabilities of acoustics emission, vibration and spectrometric oil analysis for spur gears.* Mechanical Systems and Signal Processing, 2007. **21**(1): p. 208-233.
73. Li, Y., et al., *Adaptive prognostics for rolling element bearing condition.* Mechanical Systems and Signal Processing, 1999. **13**(1): p. 103-113.
74. Qiu, J., *Damage Mechanics Approach for Bearing Lifetime Prognostics.* Mechanical Systems and Signal Processing, 2002. **16**(5): p. 817-829.
75. Amiable, S., et al., *A comparison of lifetime prediction methods for a thermal fatigue experiment.* International Journal of Fatigue, 2006. **28**: p. 692-706.

76. Ray, A. and S. Tangirala, *Stochastic modeling of fatigue crack dynamics for on-line failure prognostics.* IEEE Transactions on Control Systems Technology, 1996. **4**(4): p. 443-451.
77. Coble, J. and J.W. Hines, *Fusing Data Sources for Optimal Prognostic Parameter Selection*, in *Sixth American Nuclear Society International Topical Meeting on Nuclear Plant Instrumentation, Control, and Human_Machine Interface Technologies NPIC&HMIT*. 2009: Knoxville, Tennessee.
78. Coble, J., *Merging Data Sources to Predict Remaining Useful Life – An Automated Method to Identify Prognostic Parameters*, in *Nuclear Engineering*. 2010, University of Tennessee.
79. Lu, C.J. and W.O. Meeker, *Using degradation measures to estimate a time-to-failure distribution.* Technometrics, 1993. **35**(2): p. 161-174.
80. Liao, H. and E.A. Elsayed, *Reliability inference for field conditions from accelerated degradation testing.* Naval Research Logistics 2006. **53**: p. 576–587.
81. Gebraeel, N.Z., et al., *Residual Life Predictions From Vibration-Based Degradation Signals - A Neural Network Approach.* IEEE transactions on Industrial Electronics, 2004. **51**(3): p. 694-700.
82. Huang, R., et al., *Residual life predictions for ball bearings based on self-organizing map and back propagation neural network methods.* Mechanical Systems and Signal Processing, 2007. **21**(1): p. 193-207.
83. Herzog, M.A., T. Marwala, and P.S. Heyns, *Machine and component residual life estimation through the application of neural networks.* Reliability Engineering & System Safety, 2009. **94**(2): p. 479-489.
84. Bird, L., et al., *Policies and market factors driving wind power development in the United States.* Energy Policy, 2005. **33**(11): p. 1397-1407.
85. Arabian-Hoseynabadi, H., H. Oraee, and P.J. Tavner, *Failure Modes and Effects Analysis (FMEA) for wind turbines.* International Journal of Electrical Power & Energy Systems, 2010. **32**(7): p. 817-824.
86. Spinato, F., et al., *Reliability of wind turbine subassemblies.* IET Renewable Power Generation, 2009. **3**(4): p. 387–401.

87. Reinertsen, R., *Residual life of technical systems; diagnosis, prediction and life extension.* Reliability Engineering and System Saftey, 1996. **54**: p. 23-34.
88. Ray, A. and M.-K. Wu, *Fatigue damage control of mechanical systems.* Smart Material Structure 1994. **3**: p. 47-58.
89. Ray, A. and C.F. Lorenzo. *Damage-mitigating control - an interdisciplinary thrust between controls and material science*. in *American Control Conference*. 1994. Baltimore, Maryland.
90. Ray, A. and J.C. Newman. *Stochastic modeling of fatigue damage dynamics for failure prognostics and risk analysis*. in *American Control Conference*. 1995. Seattle, Washington.
91. Ray, A. and S. Tangirala. *A nonlinear stochastic model of fatigue crack length for on-line damage sensing*. in *35th Decision and Control Conference*. 1996. Kobe, Japan.
92. Holmes, M., S. Tangirala, and A. Ray. *Life-extending control of a reusable rocket engine*. in *American Control Conference*. 1997. Albuquerque, New Mexico.
93. Patankar, R., A. Ray, and A. Lakhtakia, *A state-space model of fatigue crack growth.* International Journal of Fracture, 1998. **90**: p. 235-249.
94. Tangirala, S., et al., *Life-extending next term control of mechanical structures - experimental verification of the concept.pdf.* Automatica, 1998. **34**(1): p. 3-14.
95. Ray, A. and R. Patankar, *A stochastic model of fatigue crack propagation under variable-amplitude loading.* Engineering Fracture Mechanics, 1999. **62**: p. 477-493.
96. Ray, A. and S. Phoha, *Stochastic modeling of fatigue crack damage for information-based maintenance.* Annals of Operations Research, 1999. **91**: p. 191-204.
97. Ray, A. *A state-space model of fatigue crack growth for real-time structural health monitoring*. in *The 19th Digital Avionics Systems Conferences*. 2000. Philadelphia.
98. Zhang, H., A. Ray, and S. Phoha, *Hybrid life-extending control of mechanical systems - experimental validation of the concept.* Automatica, 2000. **36**: p. 23-36.
99. Patankar, R. and A. Ray, *State-space modeling of fatigue crack growth in ductile alloys.* Engineering Fracture Mechanics, 2000. **66**: p. 121-151.

100. Ray, A. and R. Patankar, *Fatigue crack growth under variable-amplitude loading- Part I:Model formulation in state-space setting.* Applied Mathematical Modeling, 2001. **25**: p. 979-994.
101. Ray, A. and R. Patankar, *Fatigue crack growth under variable-amplitude loading- Part II – Code development and model validation.* Applied Mathematical Modeling, 2001. **25**: p. 995-1013.
102. Sastry, V.V.S. and A. Ray. *Online monitoring of fatigue crack damage for life extending control of aircraft structure*. in *American Control Conference*. 2001. Arlington, Virginia.
103. Gupta, S., et al., *Identification of statistical patterns in complex systems via symbolic time series analysis.* ISA Transactions, 2006. **45**(4): p. 477-490.
104. Khatkhate, A., et al. *Life-extending control of mechanical systems using symbolic time series analysis*. in *American Control Conference*. 2006. Minneapolis, Minnesota.
105. Gupta, S., *Behavioral paatern identification for structural health monitoring in complex systems*, in *MechanicalEngineering*. 2006, The Pennsylvania State University.
106. Gupta, S. and A. Ray, *Real-time fatigue life estimation in mechanical structures.* Measurement Science and Technology, 2007. **18**(7): p. 1947-1957.
107. Bae, S.J. and P.H. Kvam, *A Nonlinear Random-Coefficients Model for Degradation Testing.* Technometrics, 2004. **46**(4): p. 460-469.
108. Ardsomang, T., J.W. Hines, and B.R. Upadhyaya. *Heat Exchanger Fouling and Estimation of Remaining Useful Life* in *Annual Conference of Prognostics and Health Management Society*. 2013.
109. Gebraeel, N. and J. Pan, *Prognostic Degradation Models for Computing and Updating Residual Life Distributions in a Time-Varying Environment.* IEEE Transactions on Reliability, 2008. **57**(4): p. 539-550.
110. V. Wowk, *Machinery Vibration: Balancing*. 1995, New York: McGraw-Hill.
111. W.S. Liu and J. Cela, *Count Data Models in SAS*, in *SAS Global Forum*. 2008.
112. D. Lambert, *Zero-Inflated Poissoin Regression, with an application to defects in manufacturing.* Technometrics, 1992. **34**(1): p. 1-14.

113. Guikema, S.D. and J.P. Coffelt, *Modeling count data in risk analysis and reliability engineering*, in *Handbook of Performability Engineering* K.B. Misra, Editor. 2008, Springer: London.
114. Montgomery, D.C., *Introduction to statistical quality control*. 1985, New York: Wiley.
115. Leger, R.P., W.J. Garland, and W.S. Poehlman, *Fault detection and diagnosis using statistical control charts and artificial neural networks.* Artificial intelligence in engineering, 1998. **12**(1): p. 35-47.
116. Benneyan, J.C., *Use and interpretation of statistical quality control charts.* International Journal for Quality in Health Care, 1998. **10**(1): p. 69-73.
117. Woodall, W.H. and D.C. Montgomery, *Research issues and ideas in statistical process control.* Journal of Quality Technology, 1999. **31**: p. 376-386.
118. S. He, W. Huang, and W.H. Woodall, *CUSUM Charts for Monitoring a Zero-inflated Poisson Process.* Quality and Reliability Engineering International 2011. **28**(2): p. 181-192.
119. Leger, R.P., W.J. Garland, and W.F.S. Poehlman, *Fault detection and diagnosis using statistical control charts and artificial neural networks.* Artificial Intelligence in Engineering, 1998. **12**: p. 35-47.
120. Das, N., *A note on the efficiency of nonparametric control chart for monitoring process variability.* Economic Quality Control, 2008. **23**(1): p. 85 - 93.
121. Liu, R.Y., *Control Charts for Multivariate Processes.* Journal of the American Statistical Association, 1995. **90**(432): p. 1380-1387.
122. Y. Dai, C. Zhou, and Z. Wang, *Multivariate CUSUM control charts based on data depth for preliminary analysis* 2004, The Natural Sciences Foundation of Tianjin
123. Nosonovsky, M. and P.K. Rohatgi, *Biomimetics in Materials Science; Self-Healing, Self-Lubricating, and Self-Cleaning Materials*. 2012, New York: Springer.
124. Hager, M.D., et al., *Self-Healing Materials.* Advanced Materials, 2010. **22**(47): p. 5424-5430.
125. Elber, W., *Fatigue crack closure under cyclic tension* Engineering Fracture Mechanics, 1970. **2**: p. 37-45.

126. Zapatero, J., B. Moreno, and A. González-Herrera, *Fatigue crack closure determination by means of finite element analysis.* Engineering Fracture Mechanics, 2008. **75**: p. 41-57.
127. Meggiolaro, M.A. and J.T.P.d. Castro, *On the dominant role of crack closure on fatigue crack growth modeling.* International Journal of Fatigue, 2003. **25**(843-854).
128. Silva, F.S., *Crack closure inadequacy at negative stress ratios.* International Journal of Fatigue, 2004. **26**: p. 241-52.
129. Lee, S.Y., *Effects of Overload and Under-load on Internal Strains/Stresses and Crack Closure during Fatigue-Crack Propagation*, in *Material Science and Engineering*. 2009, University of Tennessee.
130. Sehitoglu, H., *Crack Opening and Closure in Fatigue,.* Engineering Fracture Mechanics, 1985. **21**(2): p. 329 339.
131. Liu, J.Z. and X.R. Wu, *Study on fatigue crack closure behavior for various cracked geometries.* Engineering Fracture Mechanics, 1997. **57**: p. 475-491.
132. Newman, J.C. and W. Elber, *Mechanics of Fatigue Crack Closure*, in *ASTM*. 1988: Philadelphia.
133. Sun, Y., *The Study of Crack Closure Phenomenon Following One Tensile Overload*. 2007, University of Tennessee.
134. Daniewicz, S.R. and S. Ismonov, *Simulation and comparison of several crack closure assessment methodologies.* International Journal of Fatigue, 2010. **32**(2): p. 428-433.
135. Lee, C.S., C.G. Park, and Y.W. Chang, *Precise determination of fatigue crack closure in A1 alloys.* Materials Science and Engineering, 1996. **A216**: p. 131-138.
136. Lee, J.-M. and B.-H. Choi, *Experimental observation and modeling of the retardation of fatigue crack propagation under the combination of mixed-mode single overload and constant amplitude loads.* International Journal of Fatigue, 2009. **31**(11-12): p. 1848-1857.
137. Lee, S.Y., et al., *A study on fatigue crack growth behavior subjected to a single tensile overload: Part II. Transfer of stress concentration and its role in overload-induced transient crack growth.* Acta Materialia, 2011. **59**(2): p. 495-502.

138. Lee, S.Y., et al., *A study on fatigue crack growth behavior subjected to a single tensile overloadPart I. An overload-induced transient crack growth micromechanism.* Acta Materialia, 2011. **59**(2): p. 485-494.
139. Chang, H., et al., *Acoustic emission study of fatigue crack closure of physical short and long cracks for aluminum alloy LY12CZ.* International Journal of Fatigue, 2009. **31**(3): p. 403-407.
140. Melek, W.W., et al., *Comparison of Trend Detection Algorithms in the Analysis of Physiological Time-Series Data.* IEEE Tramsactions on biomedical engineering, 2005. **52**(4): p. 639-651.
141. Lim, J.-H. and D.H. Park, *Trend change in mean residual life.* IEEE Tramsactions on Reliability, 1995. **44**(2): p. 291-295.
142. Park, D.H., *Testing whether failure rate changes its trend.* IEEE Tramsactions on Reliability, 1988. **37**: p. 375-378.
143. Taghipour, S. and D. Banjevic, *Trend analysis of the power law process using Expectation–Maximization algorithm for data censored by inspection intervals.* Reliability Engineering & System Safety, 2011. **96**(10): p. 1340-1348.
144. Vaurio, J.K., *Identification of process and distribution characteristics by testing monotonic and non-monotonic trends in failure intensities and hazard rates.* Reliability Engineering and System Safety, 1999. **64**: p. 345-357.
145. Louit, D.M., R. Pascual, and A.K.S. Jardine, *A practical procedure for the selection of time-to-failure models based on the assessment of trends in maintenance data.* Reliability Engineering & System Safety, 2009. **94**(10): p. 1618-1628.
146. J. Vierta and J.K. Vaurio, *Testing statistical significance of trends in learning, ageing and safety indicators.* Reliability Engineering and System Safety, 2009. **94**: p. 1128-1132.
147. Kvaløy, J.T., B.H. Lindqvist, and H. Malmedal. *A statistical test for monotonic and non-monotonic trend in repairable system*. in *European Conference on Safety and Reliability*. 2002. Torino.
148. Onoz, B. and M.c. Bayazit, *The Power of Statistical Tests for Trend Detection.* Turkish journal of engineering & environmental sciences, 2003. **27**: p. 247 - 251.

149. Mcfaddent, P.D. and S. J.D., *Model for the vibration produced by a single point defect in a rolling element bearing* Journal of Sound and Vibration 1984. **96**: p. 69-82
150. Wang, Y.-F. and P.J. Kootsookos, *Modeling of Low Shaft Speed Bearing Faults for Condition Monitoring.* Mechanical Systems and Signal Processing, 1998. **12**(3): p. 415-426.
151. Konstantinov, B. and T. Yoshida, *Real-time qualitative analysis of temporal shapes of (bio) process variables.* AIChE Journal 1992. **38**(11): p. 1703-1715.
152. Flehmig, F., R.V. Watzdorf, and W. Marquardt, *Identification of trends in process measurements using a wavelet transform.* Computers & Chemical Engineering, 1998. **22**(1): p. 491–496.
153. Wu, J.-L. and P.-C. Chang, *A Trend-Based Segmentation Method and the Support Vector Regression for Financial Time Series Forecasting.* Mathematical Problems in Engineering, 2012. **2012**: p. 1-20.
154. Peng, Z.K., P.W. Tse, and F.L. Chu, *An improved Hilbert–Huang transform and its application in vibration signal analysis.* Journal of Sound and Vibration, 2005. **286**(1-2): p. 187-205.
155. Yang, W.-X., *Interpretation of mechanical signals using an improved Hilbert–Huang transform.* Mechanical Systems and Signal Processing, 2008. **22**(5): p. 1061-1071.
156. Žvokelj, M., S. Zupan, and I. Prebil, *Multivariate and multiscale monitoring of large-size low-speed bearings using ensemble empirical mode decomposition method combined with principal component analysis.* Mechanical Systems and Signal Processing, 2010. **24**(4): p. 1049-1067.
157. Taghizadeh, J. and M.A. Najafabadi, *Classification of acoustic emission signals collected during tensile tests on unidirectional ultra high molecular weight polypropylene fiber reinforced epoxy composites using principal component analysis.* Russian Journal of Nondestructive Testing, 2011. **47**(7): p. 491-500.
158. Johnson, M., *Waveform based clustering and classification of AE transients in composite laminates using principal component analysis.* NDT & E International, 2002. **35**(6): p. 367-376.

159. Bozdogan, H., *Akaike's information criterion and recent developments in information complexity.* Journal of mathematical psychology, 2000. **44**(1): p. 62-91.
160. Hyvärinen, A. and Oja E., *Independent component analysis: algorithms and applications.* Neural Networks, 2000. **13**(4-5): p. 411-430.
161. Vamoş, C., *Automatic algorithm for monotone trend removal.* Physical Review E 2007. **75**.
162. Vamoş, C. and M. Craciun, *Automatic Trend Estimation*. SpringerBriefs in Physics. 2012, New York: Springer.
163. Wu, J.-L., P.-C. Chang, and Y.-F. Pan, *Building a Trend Based Segmentation Method with SVR Model for Stock Turning Detection.* World Academy of Science, Engineering and Technology, 2012. **65**.
164. Chung, F.-L., et al., *An evolutionary approach to pattern-based time series segmentation.* Evolutionary Computation, IEEE Transactions on, 2004. **8**(5): p. 471-489.
165. Wu, S.-j., et al., *A Neural Network Integrated Decision Support System for Condition-Based Optimal Predictive Maintenance Policy.* IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans, 2007. **37**(2): p. 226-236.
166. Haddad, G., P.A. Sandborn, and M.G. Pecht, *An Options Approach for Decision Support of Systems With Prognostic Capabilities.* IEEE Tramsactions on Reliability, 2012. **61**(4): p. 872-883.
167. Wang, J.J., et al., *Review on multi-criteria decision analysis aid in sustainable energy decision-making.* Renewable and Sustainable Energy Reviews, 2009. **13**: p. 2263-2278.
168. Pohekar, S.D. and M. Ramachandran, *Application of multi-criteria decision making to sustainable energy planning - A review.* Renewable and Sustainable Energy Reviews, 2004. **8**: p. 365-381.
169. Konak, A., D.W. Coit, and A.E. Smith, *Multi-objective optimization using genetic algorithms: A tutorial Original Research Article.* Reliability Engineering & System Safety, 2006. **91**(9): p. 992-1007.

170. A. Gomez, et al., *Multiobjective genetic algorithm strategies for electricity production from generation IV nuclear technology* Energy Conversion and Management, 2010. **51**(4): p. 859-871.
171. Kusiak, A. and H. Zheng, *Optimization of wind turbine energy and power factor with an evolutionary computation algorithm.* Energy, 2010. **35**(3): p. 1324-1332.
172. Hawkins, T., et al. *Wind Turbine Power Capture Control With Robust Estimation*. in *ASME Dynamic Systems and Control Conference*. 2010. Cambridge, Massachusetts: ASME.
173. Furuta, H., et al., *Optimal bridge maintenance planning using improved multi-objective genetic algorithm.* Structure and Infrastructure Engineering, 2006. **2**(1): p. 33-41.
174. Orcesi, A.D. and D.M. Frangopol, *Probability-based multiple-criteria optimization of bridge maintenance using monitoring and expected error in the decision process.* Structural and Multidisciplinary Optimization, 2011. **44**(1): p. 137-148.
175. Sardiñas, R.Q., M.R. Santana, and E.A. Brindis, *Genetic algorithm-based multi-objective optimization of cutting parameters in turning processes.* Engineering Applications of Artificial Intelligence, 2006. **19**: p. 127-133.
176. Deb, K., et al., *A Fast and Elitist Multiobjective Genetic Algorithm: NSGA-II.* IEEE Tramsactions on Evolutionary Computation, 2002. **6**(2): p. 182-197.
177. Balaban, E. and J.J. Alonso. *An Approach to Prognostic Decision Making in the Aerospace Domain*. in *Annual Conference of the Prognostics and Health Management Society* 2012.
178. Iyer, N., K. Goebel, and P. Bonissone, *Framework for Post-Prognostic Decision Support*, in *IEEE Aerospace Conference*. 2006.
179. Feldman, K., T. Jazouli, and P.A. Sandborn, *A Methodology for Determining the Return on Investment Associated With Prognostics and Health Management.* IEEE Transactions on Reliability, 2009. **58**(2): p. 305-316.
180. Haddad, G., P. Sandborn, and M. Pecht, *Using Real Options to Manage Condition-Based Maintenance Enabled by PHM*, in *IEEE International Conference on Prognostics and Health Management*. 2011: Denver, Colorado.
181. Rosas, P., *Dynamic influences of wind power on the power system*, in *Electric Power Engineering*. 2003, Technical University of Denmark.

182. *Wind Energy Conversion Systems: Technology and Trends*, ed. S.M. Muyeen. 2012, NewYork: Springer.
183. Abdin, E.S. and W. Xu, *Control Design and Dynamic Performance Analysis of a Wind Turbine-Induction Generator Unit.* IEEE Transactions on Energy Conversion, 2000. **15**(1): p. 91-96.
184. Li, Y., et al., *Adaptive Proghnostics for Rolling Element Bearing Condition.* Mechanical Systems and Signal Processing, 1999. **13**(1): p. 103-113.
185. R.M. Sundar, J.P.S., *Performance of wind turbines in a turbulent atmosphere.* Solar Energy, 1983. **31**(6): p. 567-575.
186. Systems, P.S., *Wind Turbine Control Toolbox*, in *Wind Turbine Module*. 2009, Princeton Satellite Systems: Princeton, New Jersey.
187. Niknam, T., A. Kavousifard, and J. Aghaei, *Scenario-based multiobjective distribution feeder reconfiguration considering wind power using adaptive modified particle swarm optimisation.* IET Renewable Power Generation, 2012. **6**(4): p. 236.
188. Grid, N. *Electricity Transmission Operational Data*. 2013; Available from: http://www2.nationalgrid.com/UK/Industry-information/Electricity-transmission-operational-data/Data-Explorer/.
189. Hines, J., et al., *Technical Review of On-Line Monitoring Techniques for Performance Assessment, Volume 2: Theoretical Issues.* US NRC, 2008.
190. Baraldi, P., F. Mangili, and E. Zio, *Investigation of uncertainty treatment capability of model-based and data-driven prognostic methods using simulated data.* Reliability Engineering & System Safety, 2013. **112**(0): p. 94-108.
191. Byington, C.S., et al. *A model-based approach to prognostics and health management for flight control actuators*. in *IEEE Aerospace Conference*. 2004.
192. Daigle, M. and S. Sankararaman. *Advanced Methods for Determining Prediction Uncertainty in Model-Based Prognostics with Application to Planetary Rovers*. in *Annual Conference of the Prognostics and Health Management Society*. 2013.
193. Orchard, M., et al. *Advances in uncertainty representation and management for particle filtering applied to prognostics*. in *Prognostics and Health Management, 2008. PHM 2008. International Conference on*. 2008. IEEE.

194. Sankararaman, S. and S. Mahadevan, *Uncertainty quantification in structural damage diagnosis.* Structural Control and Health Monitoring, 2011. **18**(8): p. 807-824.
195. Saxena, A., et al. *Metrics for evaluating performance of prognostic techniques*. in *Prognostics and Health Management, 2008. PHM 2008. International Conference on*. 2008. IEEE.
196. Sun, J., et al., *Prognostics uncertainty reduction by fusing on-line monitoring data based on a state-space-based degradation model.* Mechanical Systems and Signal Processing, 2014. **45**(2): p. 396-407.
197. Tang, L., et al. *Methodologies for uncertainty management in prognostics*. in *Aerospace conference, 2009 IEEE*. 2009. IEEE.
198. Wang, P., B.D. Youn, and C. Hu, *A generic probabilistic framework for structural health prognostics and uncertainty management.* Mechanical Systems and Signal Processing, 2012. **28**: p. 622-637.
199. Frangopol, D.M., *Life-cycle performance, management, and optimisation of structural systems under uncertainty: accomplishments and challenges 1.* Structure and Infrastructure Engineering, 2011. **7**(6): p. 389-413.

VITA

Seyed A. Niknam was born in Tehran, Iran. He graduated with a Bachelor of Science degree in Mechanical Engineering in June 2003. While an undergraduate, He worked as design engineer for a private company in Tehran for two years. He then attended Brunel University in London where he received his Master of Science degree with Distinction in Advanced Manufacturing Systems in June 2006. Then he started working as the instructor of Computer Aided Design in the collaborative program between the University of Birmingham and Amirkabir University of Technology. At the same time, he was working as quality engineer and sales engineer for promoting metrology equipment including Taylor-Hobson products.

In 2008, he joined Wichita State University and started working as graduate teaching assistant for the popular course of Engineering Graphics. He moved to The University of Tennessee in August 2009 and started working at the PROaCT Lab. For three years he worked on DOE and NSF related projects with focus on the reliability of wind turbine drive train. He graduated with a Master of Science degree in Reliability and Maintainability Engineering in May 2011. From January 2012 to August 2014, he was responsible for teaching the core courses in the Reliability and Engineering Program at The University of Tennessee, Knoxville. He completed his PhD in Industrial and System Engineering in August 2014.

As a graduate student, Seyed received the Chad and Ann Holiday Fellowship in Fall 2013. He is a member of American Society of Quality (ASQ), American Society of Mechanical Engineers (ASME), and The Institute for Operations Research and the Management Sciences (INFORMS).