CRANFIELD UNIVERSITY

Lim Chi Keong Reuben

# Improving Planned and Condition-based Maintenance Decision Support

School of Engineering
SOE Research

PhD
Academic Year: 2011 - 2014

Supervisor: PROF David Mba
Mar 2014

CRANFIELD UNIVERSITY

School of Engineering
SOE Research

PhD

Academic Year 2011 - 2014

Lim Chi Keong Reuben

# Improving Planned and Condition-based Maintenance Decision Support

Supervisor: PROF David Mba
Feb 2014

## EXECUTIVE SUMMARY

In both civil and military aviation, maintenance plays a large role in ensuring continued safe operation and accounts for a significant portion of operating costs. Typically, a conservative planned maintenance (PM) program is initially developed to ensure the aircraft reliability but this often leads to over-maintenance. With more in-service experience, operators seek to customize the maintenance interval accordingly in order to reduce workload and cost without compromising safety. With prevailing use of health usage monitoring systems (HUMS), the maintenance can even transit from PM to condition-based maintenance (CBM) where further safety and costs benefit may be reaped. Whilst some guidance for such changes exists, it remains challenging for maintainers in practice as suggested methods often require significant component failure or test data; which are unavailable or too expensive to obtain. As such, this research reviews the challenges faced by maintainers when extending PM intervals or implementing CBM and seeks ways to support decision making for the changes.

For PM, the challenge to extend the maintenance interval with little or no past failure is addressed. Existing reliability methods were reviewed and two improved methods to estimate the reliability lower confidence bounds were developed. The first approach adopts the use of Monte Carlo simulation applied to the Weibull plot equation while the second uses a probabilistic damage accumulation model together with bootstrap techniques. Both methods are used to assess the reliability of extending the replacement interval of a gearbox bearing and are shown to perform better than existing methods as they provide tighter reliability confidence bounds.

For CBM, a survey on sensor technologies and diagnostic algorithms showed that vibration-based sensor is most widely used to detect fault. The study then demonstrates a CBM implementation using vibration-based HUMS data from in-service helicopters. Analysis of the FFT spectra shows that the fault patterns corresponding to progressing stages of bearing wear can be clearly observed. The fault patterns are extracted as features for unsupervised classification using Gaussian Mixture Models and used to infer the different bearing health states. Signal detection theory was then applied onto the classified feature to determine the detection thresholds for fault diagnosis. A

simplistic prognostic model using trend extrapolation to determine the replacement lead-time is then performed and use for maintenance planning.

In an effort to ease the implementation of CBM, ways to improve prognostics application is explored. The Switching Kalman Filter (SKF) was adapted for both diagnostic and prognostic under an autonomous framework that requires little user input. The SKF uses multiple dynamical models with each one describing a different stage of bearing wear. The most probable wear process is then inferred from the extracted feature data using Bayesian estimation. As different stages of bearing wear can be tracked using the dynamical behavior of the measurements, pre-established threshold for fault detection is no longer required for diagnostics. The SKF approach provides maintainers with more information for decision-making as a probabilistic measure of the wear processes are available. It also offers the opportunity to predict RUL more accurately by distinguishing between the wear stages and performing prediction only when rapid and unstable wear is detected. The SKF approach is demonstrated using in-service feature data from the AH64D TRGB and the results have shown the proposed methods to be a promising tool for maintenance decision-making.

As an extension of research on methodologies to improve PM and CBM decision support, a thioether mist lubrication is explored for its feasibility as a backup lubrication system for helicopters. The aim is to reduce the mishap severity category which in turn eases the extension of PM interval or its replacement with a CBM task. An experimental setup was developed to test the thermal properties of a spur gearbox with thioether mist lubrication under various load and speed conditions and it was shown that only a very small volumetric flow of lubricant is required to preserve the gears from damage in oil starved environment. As such, a thioether based mist backup system can potentially reduce the risk of oil starvation failures significantly.

## CONTRIBUTIONS AND RELATED PUBLICATIONS

The key contributions and related publications from this research are outlined as follows:

1. Developed methods to estimate the lower confidence bound of rolling element bearing reliability without past failures for maintenance interval extension.
    - *Estimating bearing lower bound reliability without past failures*, (Published Jan 2013), IMechE, Part O, Journal of Risk and Reliability.
2. Demonstrated a CBM task development for a gearbox bearing using in-service HUMS data.
    - *HUMS experience on the RSAF AH64D Tail Rotor Gearbox*, (Presented 19 Feb 2013), AHS Airworthiness, CBM and HUMS Specialists' Meeting.
3. Validated the use of FFT spectral analysis for diagnosis of localized and generalized bearing damage for in-service helicopters.
    - *Diagnosis and prognosis of AH64D tail rotor gearbox bearing degradation*, (Published Aug 2013), Proceedings of the ASME, IDETC/CIE 2013.
    - *Bearing Time-to-Failure Estimation using Spectral Analysis Features*, (Published Feb 2014), Structural Health Monitoring.
4. Developed the use of Switching Kalman filters for diagnostics and prognostics application.
    - *Switching Kalman Filter for diagnostic and prognostic of bearing degradation*, (Published Mar 2014), Structural Health Monitoring.
    - *Condition Monitoring and Remaining Useful Life estimation using Switching Extended Kalman Filter*, (In Press), special issue for IJSEAM on "Intelligent assets through Condition monitoring and Condition-based maintenance.
    - *Fault detection and remaining useful life estimation using Switching Kalman Filters*, (Published Oct 2013), 8th World Congress on Engineering Asset Management.
    - *Switching Extended Kalman Filter for diagnostic and prognostic of bearing degradation*, (Under review), Mechanical System Signal Processing.

- *Fault detection and remaining useful life estimation using Switching Extended Kalman Filter,* (Under review), Reliability Engineering and System Safety.

5. Evaluated the feasibility of thioether based mist for backup lubrication system
    - *HELMGOP- Helicopter main gearbox loss of oil performance optimization,* EASA.2011/5, European Aviation Safety Agency.
    - *Reliability Assessment of Helicopter Main Gearbox Lubrication Systems and Optimisation in Run-dry Conditions,* 6th Annual Rotorcraft Symposium
    - *Feasibility of Thioether based mist lubrication for helicopters,* (Under review), IMechE Part J, Journal of Aerospace Engineering

**ACKNOWLEDGMENT**

I like to express my gratitude to my wife, Chai Meng, for her unwavering support and patience during the course of my studies and for taking care of the family in my absence.

I am grateful to PROF David Mba for his guidance and mentorship without which this research would not to come to its fruition. Besides expert knowledge, his broad view of issues in perspective has allowed much valuable insight to be gleaned.

Lastly, I like to thank the Republic of Singapore Air Force for the opportunity and sponsorship to pursue this post-graduate program.

# Table of Content

## List of Figures

## List of Tables

# 1 INTRODUCTION

When a helicopter is first introduced, its design airworthiness is assured through type certification which ensures that the helicopter and its sub-components are designed to operate safely within allowable condition. For continued airworthiness and reliability, an initial preventive maintenance program is developed before introduction into field service to ensure its continued airworthiness and reliability. The main aim of preventive maintenance is to avoid the failure of equipment before it actually occurs by restoring the equipment reliability through servicing tasks. For equipment with high consequence of failure, such as critical aerospace components, preventative maintenance plays a key role in ensuring continued safe operation. The types of preventive maintenance task further comprises of planned maintenance (PM) and condition-based maintenance (CBM).

PM is a set of fixed tasks carried out at periodic intervals to prevent unscheduled failure [1]. The maintenance tasks are developed based on Reliability Centered Maintenance (RCM) method to determine the maintenance requirements of the components based on their failure modes [2]. Based on the operating context of the mechanical components, the task can range from simple lubrication or inspection for defects to restoration works such as repairs or replacement of worn components. Components such as bearings and gears within transmission systems are critical but they cannot be easily accessed for visual inspection. Thus they are typically restored to stringent conditions during overhaul to assure its reliability till the next overhaul which can be over a long period of time. Therefore, the interval and type of PM tasks to be prescribed has to be balanced with the consequence of the component failure and the ease of inspection.

CBM is a comparatively new concept that has grown with the increasing use of Health and Usage Monitoring Systems (HUMS) on aircrafts. In CBM, the on-board HUMS would diagnose the component's health status through embedded sensors and trigger for maintenance actions only when potential failure or defect of the component is detected. Once a potential failure is detected, the remaining useful life of the component could also be predicted using the HUMS data; though this has yet to be widely practiced in industry. The use of embedded sensors in CBM has significant

advantage over PM for gearboxes as it allows the components within to be monitored much more frequently. Besides enhancing safety, CBM allows maintenance to be prescribed only when required thus alleviating maintenance effort and costs. These benefits however, come with high initial hardware costs and higher complexity in the maintenance program.

For a helicopter transmission system, planned maintenance is pre-dominantly adopted as it is a requirement for certification [3] although there is much interest to move towards CBM. The US Department of Defense (DoD) in particular is a strong proponent of CBM and has launched a CBM+ initiative to integrate maintenance processes and relevant technology in an effort to improve platform safety, reliability and availability [4]. With more in-service experience, operators and maintainers can revise their PM or transit to a CBM program accordingly to reduce maintenance effort and lower operating costs. Such initiatives have become increasingly important as operating cost can account for up to 65-80% [4] of the helicopter's entire life cycle and rises as the platform ages. In this research, methods to improve decision-making for both PM and CBM development and implementation with focus on helicopter transmissions are explored.

# 2 PROBLEM DESCRIPTION

This chapter elaborates on some problems faced by maintainers in extending their PM intervals and transiting from a PM to CBM program for HUMS enabled helicopters. Before decisions to change the existing maintenance program can be approved by the relevant continued airworthiness authorities, the changes have to be substantiated to ensure safety and reliability standards are not compromised. Guidance materials providing a logical framework and acceptable means of compliance for substantiation exists but it can still be challenging for maintainers to use them to support decision-making in practice. The specific challenges encountered are described herein.

## 2.1 Challenges in extending maintenance intervals

The maintenance intervals for transmission components are typically determined by the Original Equipment Manufacturer (OEM) through engineering analysis during initial certification of the aircraft. The initial interval tends to be conservative to ensure a high safety level but this also often leads to over-maintenance [5]. With in-service experience, maintainers seek to revise the intervals for component overhauls and retirement lives to reduce maintenance costs and efforts. In EASA's requirement for management of continuing airworthiness [6], reliability programs to revise maintenance tasks and interval based on component failure experience in operation are required.

The process to develop the PM task and its interval is well documented in MSG-3 ATA [2] for civil aircraft and the military adopts their own standards such as MIL-STD-3034 [1] and DEF-STAN-0045 [7] for the US and UK respectively. The process within these documents requires that the reduction in reliability levels from the extension is within acceptable safety level and this has to be substantiated by engineering analyses. For transmission components, the engineering assessments are commonly carried out using reliability models such as Weibull distributions fitted with in-service time-to-failure data. However, helicopter components especially Flight Safety Critical Aircraft Parts (FSCAP) such as gears and bearings within the transmission system have a very high reliability and defect occurrence tends to be rare.

As such, traditional reliability models relying on past failure data can seldom be applied in practice.

In recognition of such circumstances, the International Maintenance Review Board Policy Board (IMRBPB) released Issue Paper (IP)-44 [5] in 2007 to provide guidelines on maintenance interval extension. IP-44 states that extension of a maintenance task is allowed if it is substantiated from servicing records that no significant defects were found with 95% confidence level. This allows extensions if little or no failures were found but it still does not provide guidance on the period of extension, which remains largely based on qualitative assessments. For the military, the naval aviation RCM guideline recommends for 'age exploration' to be carried out to extend maintenance intervals or retirement lives [8]. Age exploration involves an incremental approach to extend the intervals coupled with appropriate inspections but it also does not have guidelines on how those intervals should be determined. As such, an approach to assess the risk in the extension quantitatively is still lacking and an alternative that is not reliant on past failure data is desired for maintenance decision support.

### 2.2 Challenges in transiting from PM to CBM

For CBM tasks development, the key guidance document is the AC29-MG15 titled "Airworthiness Approval of Rotorcraft HUMS" [9], which is the FAA's advisory circular on implementation of HUMS on aircraft. The equivalent document for the military can be found in ADS-79B [10] published by the US Army. Through the use of HUMS, PM can be replaced by CBM thus alleviating operating costs and maintenance effort. Besides cost benefits, CBM also provide higher safety benefits as defects in critical components can be monitored and rectified before catastrophic failure with high collateral damage occurs. A thorough description of CBM benefits can found in the HUMS toolkit developed by the International Helicopter Safety Team (IHST) [11].

The benefits of CBM are well recognized but before they can be reaped, the HUMS system, including the hardware, diagnostics and prognostics algorithms, has to undergo a rigorous approving process laid out in AC29-MG15. With respect to transiting from a PM to CBM program, the key requirement is for validation of "maintenance credits" - a

term commonly used in CBM for credits earned through alleviation of maintenance effort such as removal or extension of maintenance tasks. For this validation, the physics of failure of the monitored component has to be understood and the diagnostic capability of the algorithms used in the maintenance decision-making process has to be demonstrated through seeded tests or field defects. The latter requirement has proven to be challenging in field application as seeded test are expensive and as previously described, the failure of FSCAP components are rare for adequate field data to be collected. In lieu of this, there are currently no civil helicopters that are certified with maintenance credits based on AC29-MG15 and the FAA is still validating the approach based on the S-92 and BK-117C2 helicopter [12] at the time of writing this thesis.

For the military, the US Army is the operator of the largest fleet of HUMS enabled helicopters and they had published some of their experiences on their AH64D and H-60 helicopters [13; 14]. However, the findings in those case studies are isolated and lack repeatability to demonstrate the consistency of the HUMS performance. Therefore, evidence to demonstrate the effectiveness of the diagnostics algorithm used is very limited despite the huge amount of literature on CBM related research. In addition, the limited in-service case studies show only diagnostic capabilities with no prognostic application. Like diagnostic algorithms, most prognostics methods require a large number of seeded tests or field defects for training the algorithms before they can be applied. Validation through in-service field defects is even more difficult as the HUMS data throughout the degradation process has to be available. Furthermore, it is more difficult to integrate prognostics methods into the maintenance processes as significant expert knowledge is often required to select the appropriate method and for the HUMS data to be pre-processed accordingly. Therefore, it is desirable for both diagnostics and prognostics algorithms to be demonstrated to be effective for in-service applications and for prognostics algorithms' ease of use to be improved so that it can be adopted by maintainers more widely.

### 2.3 Limitations due to severity categorization

An underlying factor which has significant impact on a PM task extension or transition to CBM task is the mishap severity category of the system itself. Under the system

safety framework, the mishap severity of a component is assessed based on its probability of failure (or conversely its reliability) and the consequence of its failure and for risks to be managed accordingly. The guidelines for assessing the system (and subsystem) mishap severity are documented in FAA System Safety Handbook [15] and MIL-STD-882D [16] for both civil and military aviation respectively. The types of risk are typically identified using tools such as Fault Tree Analysis (FTA) or Failure Mode Effects and Criticality Analysis (FMECA) and these are performed before the aircraft can be certified. Based on these analyses, the risk of system mishaps are managed through planned maintenance to ensure continued airworthiness as mentioned above. Any risk that could not be mitigated through design, maintenance or procedures would then have to be accepted as residual risk.

Currently, both AC-29 MG15 and ADS-79B only allows maintenance tasks on systems with a 'Major' severity category to be replaced with a CBM task. 'Major' severity category denotes a 'significant reduction in safety margins and functional capability' in event of failure whilst 'Hazardous' severity category denotes a large reduction in both safety margins and capability. Most of the helicopter transmission system such as the gearboxes and lubrication systems however belongs to 'hazardous' category due to their direct impact on flight safety. Therefore, CBM can be developed for enhanced safety and some logistics planning benefits from prognostics but the PM tasks itself cannot be replaced. It is perhaps for this reason that several literature on helicopter diagnostics and prognostics were focused on the oil cooler fan bearing [17-19] where CBM implementation in compliance to AC-29 MG15 and ADS-79B is more likely. In this context, more systems on the helicopter could be candidates for CBM task replacement if the mishap severity category of its failure is relaxed to 'Major' category. For critical systems such as the main and tail rotor gearboxes, this will be unlikely as they have direct impact to flight safety and the consequences of failure will remain severe despite reducing its probability of failure. For auxiliary systems such as the accessory gearboxes or lubrication systems however, it is possible to reduce the impact of their failures by improving system reliability or introducing redundancies. The mishap severity of these systems could then be reduced which in turn allows CBM task to be implemented.

Besides easing CBM task replacement, the reduction of the mishap severity category eases PM task interval extension as well. In [8], the rigor of the inspections required during the age exploratory phase for interval extension is highly dependent on the mishap severity category of the system. With higher mishap severity, more engineering data is required to ensure safe operation in during the period of extension and laboratory testing is called for if in-service data is unavailable. However, reducing the mishap severity category for lubrication systems can be challenging as there had been several helicopter incidents related to oil starvation. As such, ways to reduce the risk of lubrication system failure is needed. Admittedly, reducing system risk of helicopter subsystems does not contribute towards maintenance decision support. It is included in this research as it serves to enable transition to CBM and/or PM task extension through design improvements.

# 3 AIM AND OBJECTIVES

With ever-increasing competition for resources, aircraft operators have to improve the way their fleets are maintained so that both cost and manpower efficiency can be achieved without compromising safety of flight. With the challenges discussed in the previous chapters, ways to overcome them is explored. The overall aim of this research is to improve maintenance decision support for extending preventive maintenance intervals and transiting towards a condition-based maintenance program for helicopter transmission. The specific objectives of this research are outlined as follows:

- Evaluate methods for assessing component reliability with little or no failure data
- Improve or develop methods for substantiating escalation of planned maintenance interval
- Survey the existing use of HUMS on in-service helicopters
- End-to-end development of a CBM task for an in-service helicopter in accordance with FAR29 MG-15 and ADS-79B
- Improve the ease of use of prognostic tools for maintenance
- Reduce the mishap severity of helicopter lubrication system to ease transition to CBM task or PM task extension

The scope to cover the full range of transmission components for the above objectives would be too wide. Therefore, the research is going to focus on rolling element bearings as they tend to have shorter service live compared to gears and shafts and is often the life limiting factor for transmission overhauls. Despite focusing on the bearing only, the challenges faced in practice are multi-faceted in nature. Therefore, this research will span across different fields of engineering disciplines including reliability, data processing and statistical computing to address the various issues. The maintenance data used in this research are gathered from the Republic of Singapore Air Force (RSAF) AH64D helicopter. It includes engineering information from technical manuals, HUMS data and limited failure data from maintenance records. These types of information are typically accessible by maintainers in practice and its use to improve maintenance decision-making is presented in this research.

# 4 ESTIMATING BEARING RELIABILITY WITHOUT FAILURE DATA FOR PM INTERVAL EXTENSION

In this chapter, existing methods to estimate a rolling element bearing's reliability without failure data are reviewed. Traditional reliability models uses historical time-to-failure data fitted onto a parametric distribution such as Normal, Exponential or Weibull and are most commonly used. As mentioned previously, the use of these methods can be challenging when there is little or no failure data available. For improving reliability assessment in such circumstances, alternative methods are explored. Specifically, approaches to quantitatively assess the risk from the extension of PM task are sought to support maintainers in decision-making.

## 4.1 Review of reliability methods requiring little or no failure data

Although limited, there are approaches that does not rely on past failure data for reliability assessment. Nelson's method is one such approach as described in [20]. In his method, truncated test data without failure cases is used to obtain a lower confidence bound for reliability. It assumes that the components lives follows a Weibull distribution and requires the Weibull slope parameter, $\beta$ to be known. In a maintenance scenario, this method can be applied where bearings replaced without defects in past scheduled maintenance are treated as truncated non-failure test data. For bearings, a value of 1.11 for $\beta$ is generally accepted based on observations by Palmgren and Lundberg [21]. A limitation of this method is that the lower confidence bound tends to be very wide. A small extension in interval will result in a large drop in reliability, which renders it unfavorable in maintenance decision-making.

The use of the 2-parameter Weibull plot equation shown in Eqn (1) is another approach which is widely adopted for bearing reliability assessment in the industry. In the absence of failure data, the bearing rated life, $L_{10}$ at which 10% of the bearings are expected to fail, is often used as the fatigue life of the bearing. In more safety critical applications such as aerospace or automobile components, the rated life adopted can be stringently reduced to $L_1$ or even $L_{0.1}$ [22]. The $L_{10}$ life of a bearing is often known else it can be calculated with bearing geometry and loading conditions using methods

described in DIN or BS 281. With $L_{10}$ known, the reliability of the bearing at any time can be determined using Eqn (1). Compared to Nelson's method, additional information, $L_{10}$ is used in the Weibull plot equation method. A drawback in the direct use of this method is that it assumes a fixed $\beta$. As mentioned by Vlcek et al in [23], different groups of the same bearings tested under similar condition can result in different $\beta$ values with widely varying $L_{10}$ lives. As such, the Weibull plot equation by itself does not account for variation in $\beta$ and does not provide a confidence bound in the reliability assessment. As such, the uncertainty in the reliability assessment is not available to maintainers from the direct use of the Weibull plot equation.

Besides time-to-failure distribution models, model or physics-based methods may be applied as well for reliability assessment. Model-based method adopts a known mathematical model of the system's degradation process such as the Paris's crack growth or bearing life equation. Depending on the model, either health (i.e. vibration levels) or usage type data (i.e. applied stress) can be used for reliability prediction. The bearing fatigue life equation using Miner's Rule is well established by Lundberg and Palmgren in [21] where the bearing life can be determined from the bearing geometry and normal contact loads. The use of Miner's Rule in evaluating accumulated fatigue damage is typically applied as a deterministic process where the damage per load cycle and the allowable damage threshold are fixed. However, this degradation process can also be described as probabilistic models. Liao et al [24] describes the probabilistic modeling of fatigue accumulation where both the accumulated damage and the fatigue life limit of the components are modeled as random variables with a known distribution. Reliability at a specified time is then assessed as the probability that the accumulated damage distribution is greater than the fatigue life distribution. A shortcoming of this approach is that both the distributions of the damage accumulation process and the allowable damage threshold are required and may not be readily available. A direct use of this method also does not provide confidence bound in the reliability assessment to handle uncertainties.

Other methods of bearing reliability assessment using little or no failure events in existing literature are limited. R. Sehgal et al [25] developed a procedure based on graph theory and matrix approach to model the reliability of the bearing by considering

the reliability of its elements and their connections. However, their method was focused on obtaining a reliability measure for bearing selection during design and does not model the bearing's reliability against usage time when it is in use.

### 4.2 Lower Bound Reliability Estimate of a Gearbox Bearing

In this study, two methods are proposed to improve the reliability assessment of a gearbox bearing. In the first proposed method, Monte Carlo simulation is applied with the Weibull plot equation to obtain the variation in the Weibull Slope parameter and then estimate the lower confidence bound. In the second proposed method, probabilistic damage accumulation model is applied to assess the bearing reliability and Bootstrap techniques are applied to obtain the lower confidence bound. The proposed methods are built on the work of Vlcek et al [23] and Rathod et al [26] for the two approaches respectively.

In addressing the variation of $\beta$ and $L_{10}$ in bearings, Vlcek et al [23] applied Monte Carlo simulation testing together with strict series reliability model to determine the variation as a function of bearing tested. They compared the results from their approach against bearing endurance sets (analyzed by Harris [27; 28]) and found that 98% of the bearings have life exceeding the minimum life predicted from the simulation. As such, their approach predicts the lower bound variation of the bearing life from the calculated $L_{10}$. A unique aspect of their approach is that it allows the distribution of $\beta$ and $L_{10}$ to be estimated without using failure data. A Weibull distribution is still assumed for the bearing life but the parameters are no longer fixed. It should be noted that the Monte Carlo based method predicts a statistical variation in the $L_{10}$ but not the minimum life itself where no failure occurs. A no-failure fatigue life for bearing has been approximated by Tallian [29] but it is not employed here as a three parameter Weibull distribution is assumed and still requires a significant amount of failure data. In the first proposed method, the Monte Carlo Simulation method by Vlcek et al [23] is adopted to estimate the confidence interval in $\beta$ and $L_{10}$. A lower bound reliability estimate is then obtained by using the lower bound set of $\beta$ and $L_{10}$ in the Weibull plot equation.

In the use of probabilistic damage accumulation models, Rathod et al [26] proposed a simplified approach where the damage accumulation distribution may be obtained from the component's fatigue life distribution. For a bearing, the fatigue life distribution could in turn be obtained from the Monte Carlo Simulation Testing described above. As such, both distributions could be estimated with limited available information. In addition, the confidence bound for the reliability for the probabilistic damage accumulation model are obtained through bootstrap techniques [30].

Both proposed methods are described here through a case study performed on data adapted from a helicopter gearbox bearing. The increase in risk by extending the bearing replacement interval from 2000hrs to 3000hrs is evaluated using both methods and compared against Nelson's method. The specification of this bearing is detailed in Table 1. It is assumed that 30 such bearings were previously replaced in the helicopter fleet during maintenance at 2000hrs with no defects found.

Table 1 Bearing Specifications.

| | |
|---|---|
| Bearing Type | Angular Contact Bearing |
| Material | AISI M50 |
| Number of Balls, $Z$ | 19 |
| Ball Diameter (in), $d$ | 0.375 |
| Pitch Diameter (in), $d_m$ | 2.559 |
| Contact Angle (deg), $\alpha$ | 30 |
| $L_{10}$ Life (hrs) | 14700 |
| $L_\beta$ Characteristic Life (hrs) | 111632 |
| Weibull Slope, $\beta$ | 1.11 |

### 4.3 Weibull Plot Equation with Monte Carlo Simulation

The cumulative distribution of bearing life can be represented by the Weibull plot equation given by:

$$lnln\left(\frac{1}{R(t)}\right) = \beta ln\left(\frac{t}{L_\beta}\right), \text{ where } 0 < L < \infty, 0 < R < 1 \quad (1)$$

where, $R(t)$ is the bearing reliability, $lnln(1/R(t))$ is the percentage of bearings surviving at time $t$. $L_\beta$ is the characteristic life of the bearing. $\beta$ is the Weibull slope. For bearing replacement interval at $L$ = 2000hrs and $\beta$ = 1.11, reliability of the bearing can

be obtained using Eqn. (1) to be $R$ = 0.9886. If the interval is extended to $L$ = 3000hrs, the reliability of the bearing reduces to $R$ = 0.9821. As mentioned, this approach is simple to apply but it assumes a fixed $L_{10}$ life and $\beta$ = 1.11. If there are variations in these parameters, the reliability assessment will change accordingly.

Vlcek et al [23] applied Monte Carlo simulation for "virtual bearing testing" to determine variations in bearing life for any number of bearings tested. Monte Carlo simulation is a class of computational algorithms broadly used to obtain numerical results through repeated random sampling. It is widely used in problems where closed form solutions or deterministic methods are unavailable. In their approach, Vleck et al have virtual bins containing the inner race, outer race and ball bearings. Within each virtual bin, a large number of the parts (i.e.1000) are assigned an order number (i.e. 1, 2, 3…1000). Each part is then assumed to have a life according to their order number correlated to the Weibull plot for the part (i.e. $L_{0.1}$, $L_{0.2}$. $L_{0.3}$...$L_{100}$). A bearing is then assembled by randomly selecting parts from each bin and the life of the assembled bearing is limited by the part with the lowest life as shown in Figure 1.

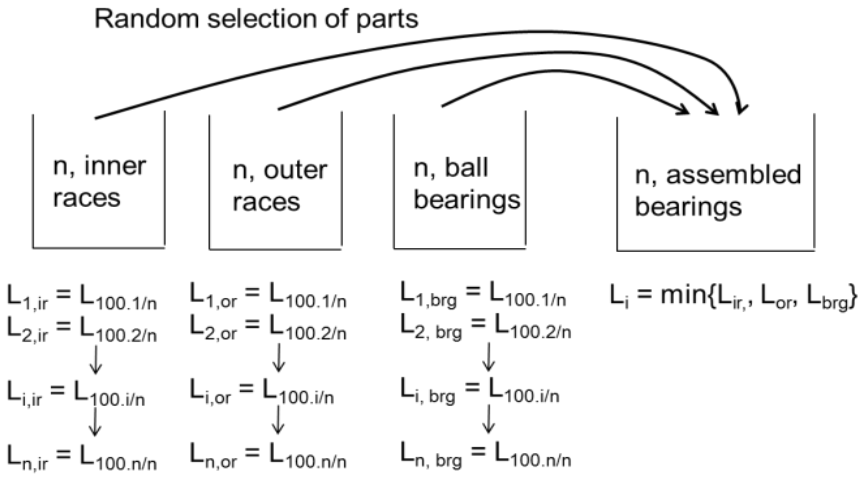


Figure 1 Virtual bearing assembly

This is repeated for N number of bearings tested and a set of $\beta$ and $L_{10}$ for the bearing set can be determined by fitting the set of N bearings' life back into the Weibull plot equation. This procedure is then repeated for multiple sets of bearings to obtain the distribution $\beta$ and the corresponding $L_{10}$. The (100C)th percentile of $\beta$ and $L_{10}$ can then

be determined from the distribution and used in Eqn. (1) to derive the lower bound reliability.

### 4.3.1 Determination of Bearing Inner & Outer Race Life

Before performing the Monte Carlo simulation, the rated life of the inner race and outer race are first calculated. Based on method described by Harris and Kotzalas [22], the inner race life, $L_{ir}$ to outer race life, $L_{or}$ can be determined to be:

$$\frac{L_{ir}}{L_{or}} = \left(\frac{C_{ir}}{C_{or}}\right)^3 \left(\frac{F_{or}}{F_{ir}}\right)^3 \approx \left[\left(\frac{1-\gamma}{1+\gamma}\right)^{1.39+\frac{1}{3}} \left(\frac{J_2}{J_1}\right)^2\right]^3 \tag{2}$$

where, $C_{ir}$ and $C_{or}$ are the dynamic load capacity of the inner and outer race. $F_{,ir}$ and $F_{or}$ are the applied load on the inner and outer race. $\gamma$ is given by $dcos(\alpha/d_m)$. $J_1$ and $J_2$ are the factors relating mean load on a rotating and non-rotating raceway to the maximum raceway normal load under nominal load distribution factor, $\varepsilon$ = 0.5. These may be obtained from reference tables in [22] based on the bearing geometry. Zaretsky's rule [23] states "For radially loaded ball and roller bearings, the life of the rolling element set is equal to or greater than the life of the outer race. Let the life of the rolling element set (as a system) be equal to that of the outer race." The relationship between the bearing rated life $L_{sys}$ and its sub-components based on strict series reliability model is then given by:

$$\begin{aligned}\left(\frac{1}{L_{sys}}\right)^{\beta} &= \left(\frac{1}{L_{ir}}\right)^{\beta} + \left(\frac{1}{L_{or}}\right)^{\beta} + \left(\frac{1}{L_{re}}\right)^{\beta} \\ &= \left(\frac{1}{L_{ir}}\right)^{\beta} + 2\left(\frac{1}{L_{or}}\right)^{\beta}, where\ L_{re} = L_{or}\end{aligned} \tag{3}$$

It should be noted that Zaretsky's rule differs from the method of Lundberg and Palmgren as the latter does not include the lives of the rolling elements as they have observed that the rolling elements have much longer lives compared to the races. From Eqn. (2) the life ratio of the inner to outer race is 0.349. Applying this ratio to Eqn. (3)

together with $L_{sys} = L_{10}$ = 14700hrs, $L_{ir}$ and $L_{or}$ are then determined to be 22719hrs and 65131hrs.

### 4.3.2 Monte Carlo Simulation Testing of Bearing Life

For the Monte Carlo testing, the number of bearings, N to be tested will affect the amount of variation in the rated life. With a higher number of bearings tested, the variation in the Weibull slope parameter, $\beta$ and the $L_{10}$ life will converge towards the assumed values as shown by Vlcek et al [23]. A low number of bearings tested will conversely result in huge variations. As such, the number is recommended to be representative of the number of bearings in the fleet or have undergone replacement. If there were any defective bearing found during the replacement, it can used to evaluate the lower bound variation of the bearing lives. For this particular investigation, $N$ is set at 30 which represents the number of bearings replaced in the fleet during past maintenance. The Monte Carlo Simulation is performed using script written in MATLAB and the virtual test of the 30 bearings was repeated 1000 times to estimate the variation in $\beta$ with convergence of the parameter achieved. Figure 2(a) shows the Weibull plots from Monte Carlo Testing of one set of 30 bearings from which a set of $\beta$ and $L_{10}$ is obtained. The results from the 1000 repetitions are shown in Figure 2 (b) from which distribution of $\beta$ and $L_{10}$ are obtained as shown in Figure 2 (c) and (d).

Assuming a normal distribution, the Weibull plot parameters from the testing are $\beta \sim N(1.187, 0.169)$ and $L_{10} \sim N(17340,5970)$; with a coefficient of variance, $c_v$ of ~0.344 obtained for $L_{10}$. This agrees well with experimental results by Rosado et al [31], where fatigue life testing on M50 bearings showed a $c_v$ of ~0.21 and ~0.36 for heavily and lightly loaded bearings respectively. By applying the Monte Carlo simulation, the variability in $\beta$ is considered and a lower confidence bound for the bearing reliability can be constructed. The lower 95% confidence bound for $\beta$ is 0.917 and the corresponding $L_{10}$ is 7520hrs. Using these parameters, the reliability of the bearing against time based on the Weibull plot equation and its corresponding 95% lower confidence bound can be obtained as shown in Figure 3. Based on this lower bound, the change in reliability by extending the replacement interval from 2000hrs to 3000hrs is from 0.9692 to 0.9557.

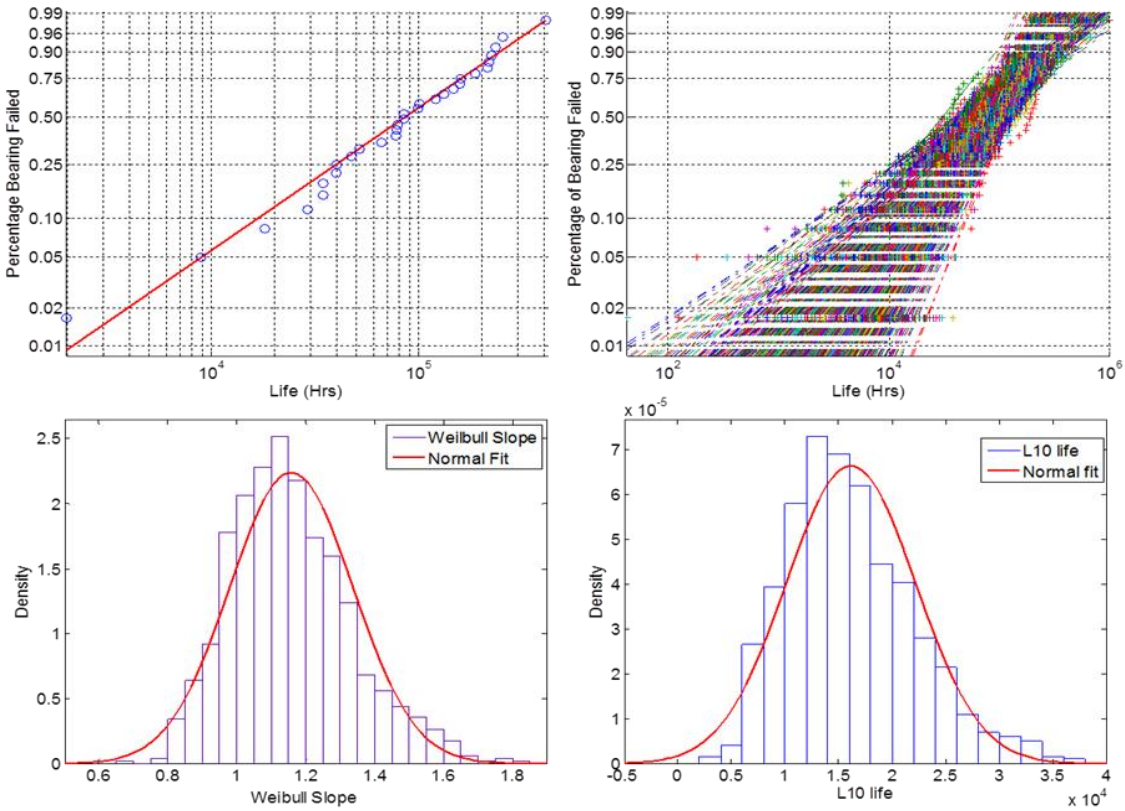


Figure 2 (a)Weibull Plot for Monte Carlo Simulation for one set of 30 Bearing Test, (b)Weibull Plot from Monte Carlo Test of 1000 set of 30 Bearing Tests (c) Distribution of $\beta$ from Monte Carlo Test and (d) Distribution of $L_{10}$ from Monte Carlo Test

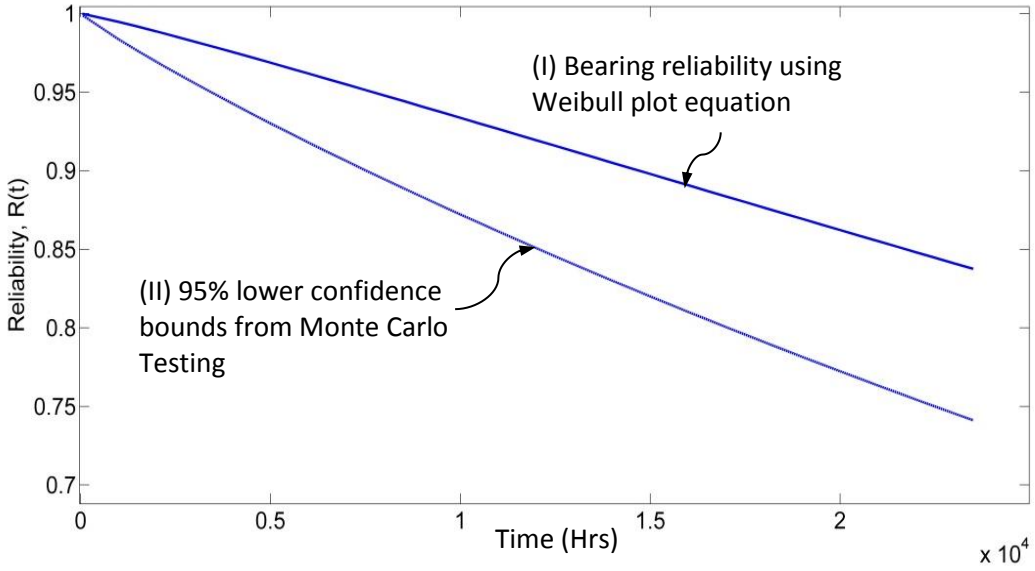


Figure 3 Reliability plot using (I) Weibull Plot Equation and (II) with 95% lower confidence bounds in Weibull Slope, β

### 4.4 Probabilistic Damage Accumulation (PDA) Model

The bearing fatigue life can be modeled using Miner's Rule where the bearing fails when the fatigue damage *D*, accumulates to the threshold, $D_{cr}$ =1 at its fatigue or $L_{10}$ life, given by:

$$D = \sum_{i=1}^{i=n} \frac{t_i}{L_{10,i}} \text{ , for } n \text{ load levels} \tag{4}$$

The damage accumulation from direct use of Miner's Rule however, is a deterministic approach which does not account for the stochasticity in the process. In Figure 4, the damage accumulation is modeled with a probabilistic approach where both the $D_c$ and *D* are treated as random variables with a probability density function (pdf). Reliability at a given time is assessed based on the overlap (or interference) between the two distributions as shown in Figure 4. In this way, the reliability of the bearing can be obtained from the damage accumulated from time in usage.

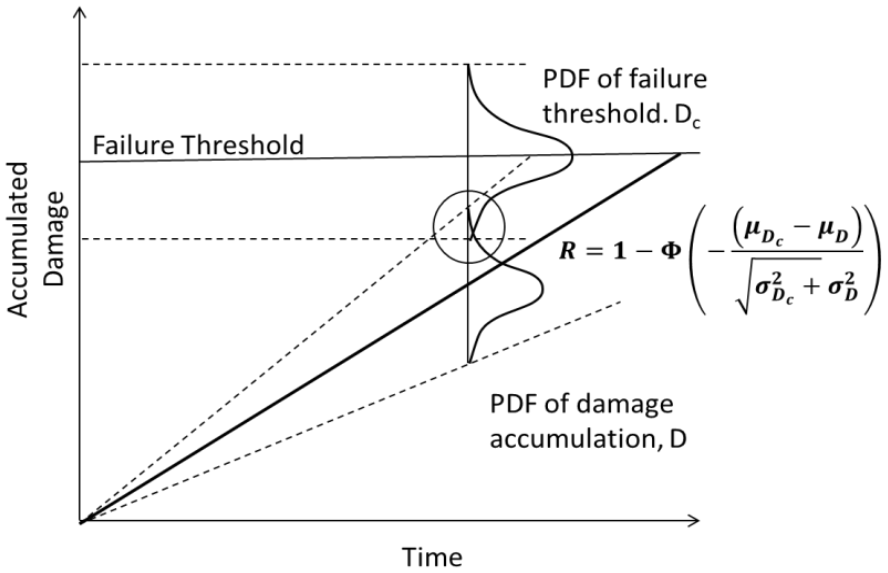


Figure 4 Probabilistic modeling of damage accumulation (Adapted from [26])

Rathod et al [26] proposed a methodology for such probabilistic modeling of fatigue damage accumulation. The method employs the linear damage accumulation model of Miner's Rule, a probabilistic S-N curve of the subject component and assumes a one to

one transformation between the damage accumulation pdf and the fatigue life pdf. The following assumptions have been made in applying the model:

(1) Fatigue failure occurs when damage accumulation, $D$ reaches the threshold damage, $D_c$, where $E[D_c] = 1$.
(2) The threshold damage or critical damage has the same distribution as the damage accumulation measure.
(3) When usage life is equal to the fatigue failure life, the variability of threshold damage accumulation, $\sigma_{Dc}^2$ is equal to the variability of damage accumulation measure, $\sigma_D^2$. The variability of damage accumulation measure continuously increases with usage life but when usage cycle reaches to fatigue failure level, the corresponding variability is assumed.

The S-N curve equation used in fatigue damage accumulation is given by,

$$N_f S^m = A \tag{5}$$

, where, $N_f$ is the cycle to failure at stress level, *S. m* is the slope parameter and *A* is the fatigue strength constant. The bearing life equation can be expressed in similar form as:

$$L_{10} F^p = C^p = constant \tag{6}$$

, where *F* is the applied load on the bearing, *C* is the bearing load capacity and *P* is the load life exponent and is typically 3 for angular contact bearings. In this case, the $L_{10}$ is treated as the fatigue life which is often practiced as mentioned above. It can be seen that the relationship between bearing life and applied load is similar to the S-N curve used in fatigue analysis. As such, the method proposed by Rathod et al [26] may be adapted for evaluating bearing reliability. For the general case under variable loading, the variation or standard deviation of the damage accumulation measure with usage life is given by:

$$\sigma_D = \sqrt{\sum_{i=1}^{n} \left( \frac{t}{L_{10,i}} \left( \frac{\sigma_{L_{10,i}}}{L_{10,i}} \right) \right)^2} \, , for\ n\ load\ levels \tag{7}$$

and the critical damage threshold is given by:

$$\sigma_{Dc} = \sqrt{\sum_{i=1}^{n}\left(\frac{t_i}{L_{10,i}}\left(\frac{\sigma_{L_{10,i}}}{L_{10,i}}\right)\right)^2}\ , for\ n\ load\ levels \tag{8}$$

, where $t_i$ is the total time spent at each different load levels and $L_{10,i}$ is the $L_{10}$ life at each different load level. The reliability of the bearing applying stress strength interference method is then given by:

$$R(t) = \Phi\left(\frac{\left(\mu_{D_c} - \mu_D\right)}{\sqrt{\sigma_{D_c}^2 + \sigma_D^2}}\right) = \Phi\left(\frac{\left(1 - \sum_{i=1}^{n}\frac{t}{L_{10,i}}\right)}{\sqrt{\sigma_{D_c}^2 + \sum_{i=1}^{n}\left(\frac{t}{L_{10,i}}\left(\frac{\sigma_{L_{10,i}}}{L_{10,i}}\right)\right)^2}}\right) \tag{9}$$

### 4.5 Reliability Assessment of Bearing using PDA model

Before the probabilistic damage model can be applied, the fatigue life pdf of the bearing is required. Without physical test data, the pdf of the bearing fatigue life (or the $L_{10}$ life) obtained from Monte Carlo testing previously is used. For $L_{10}$ ~N(17340,5970), $\sigma_{L10}$ = 5970hrs from Section 4.3.2, the critical damage threshold can be obtained from Eqn. (8) to be $\sigma_{Dc}$ = 0.344. In this case, $\sigma_{Dc}$ is equal to $c_v$ as there is only one load level. With these parameters, the reliability of the bearing against usage time can be obtained using Eqn. (9). For single load case, Eqn. (9) simplifies to:

$$R(t) = \Phi\left(\frac{\left(1 - \frac{t}{L_{10}}\right)}{\sqrt{\sigma_{D_c}^2 + \left(\frac{t}{L_{10}}\left(\frac{\sigma_{L_{10}}}{L_{10}}\right)\right)^2}}\right) \tag{9a}$$

The reliability of the bearing against time using this approach is shown in Figure 5. The change in reliability by extending the replacement interval from 2000hrs to 3000hrs is from 0.9947 to 0.9910.

### 4.6 Constructing Confidence Bounds Using Bootstrap Estimates

The reliability assessment using probabilistic damage accumulation model can be further improved upon by constructing confidence bounds using a parametric Bootstrap procedure. The Bootstrap technique was first introduced by Elfron and Tibshirani [30] and its application is as follows:

(a) For a given usage time, the pdf for $D$ and $D_c$ are obtained from Eqn. (4), (7) & (8)

(b) From each distribution of $D$ and $D_c$, random samples $d^*$ and $d_c^*$ of size 30 are drawn.

(c) From each sample set, the mean and standard deviation $\mu_{d^*}$, $\mu_{dc^*}$, $\sigma_{d^*}$ and $\sigma_{dc^*}$ are obtained.

(d) $\mu_{d^*}$, $\mu_{dc^*}$, $\sigma_{d^*}$ and $\sigma_{dc^*}$ are then applied in Eqn. (9) to obtain an estimated reliability, $R^*$.

(e) Steps (b) to (d) are repeated for a large number of times. In this example, it is repeated 2000 times. A distribution of $R^*$ is then obtained.

(f) The confidence interval can be estimated by taking the percentiles from the $R^*$ distribution.

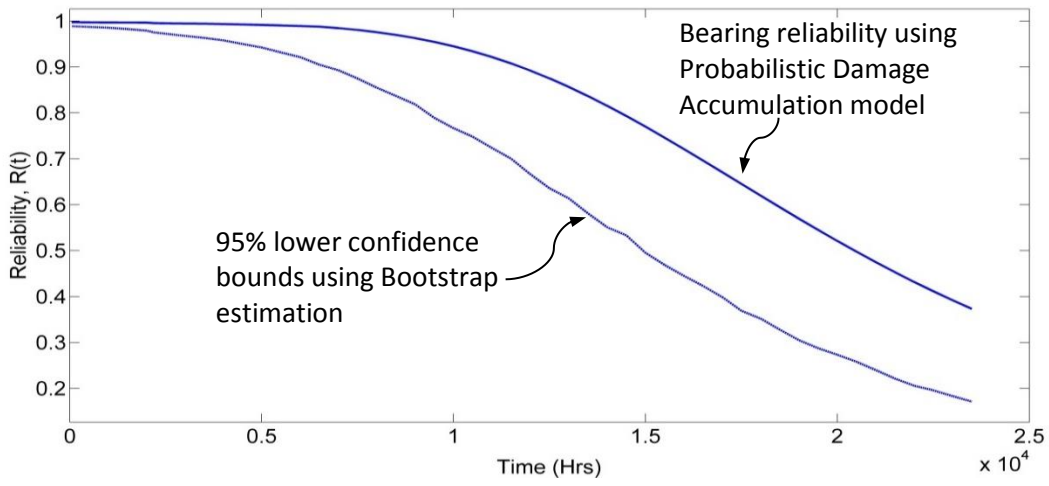


Figure 5 Reliability plot from PDA model with 95% lower confidence bound obtained using bootstrap estimation

The above steps provides the point estimate confidence interval of *R(t)* at a given usage time. They can be repeated across usage time to obtain the confidence bounds. The result is shown in Figure 5 for 95% confidence bounds. The change in reliability by extending the replacement interval from 2000hrs to 3000hrs is from 0.9792 to 0.9699.

### 4.7 Nelson's Method

Nelson [20] shows that the reliability can be assessed using a Weibull analysis method based on few or no failure. It is a method used to estimate the reliability and confidence limits that apply to little or no failures in testing with an assumed Weibull slope value for the test component. For a sample of non-failure components size, *n*, the lower bound reliability can be obtained as follows:

$$R(t) = \exp\left(\frac{-t^{\beta}\chi^2_{(C;2r+2)}}{2t'}\right) \tag{10}$$

where $\beta$ is the Weibull slope parameter, $\chi^2_{(C;2r+2)}$is the (100C)th percentile of the chi-square distribution with *(2r+2)* degrees of freedom, r is the number of failure cases and $t' = \sum_{i=1}^{n} t_i^{\beta}$. For $n$ = 30, $\beta$ =1.11 and $t$ = 2000hrs for each bearing, $t'$ = 138441hrs. For $C$ = 95% confidence limit and $r$ = 0, $\chi^2_{(0.95;2)}$= 5.992. For an extension of bearing replacement interval from 2000hrs to 3000hrs, the change in reliability using Eqn. (10) is from 0.9050 to 0.8550.

### 4.8 Comparison of Reliability Assessment Methods

The results from the various methods used to assess the gearbox bearing reliability are shown in Figure 6. It can be seen that Nelson's method gives a very wide lower confidence bound. In comparison to Nelson's method, both approaches using the Weibull plot equation and the PDA model yields narrower confidence bounds. It can be observed that the reduction in reliability for the PDA model is lower compared to the Weibull plot equation for low usage time and increases with usage time. This is because the variability in the damage accumulation in the PDA model increases proportionally with usage time. As the usage time increases, the variation in damage

accumulation increases and the reduction in reliability increase accordingly. For very high usage time, the PDA model can even be more severe compared to Nelson's method as shown in Figure 6.

It should be noted that all three approaches discussed here are very different in nature. Nelson's method used only information from bearing replacement hours and does not utilize any other information. The Weibull plot equation uses bearing geometry information with bearing life following a distribution. The PDA model assumes the $L_{10}$ life as the bearing fatigue life and damage accumulation follows Miner's Rule. In practice, these are the various reliability approaches that can be applied by maintenance engineers to assess the risk in extending bearing replacement interval. Depending on the extension period and type of information available, a conservative estimate can be obtained.

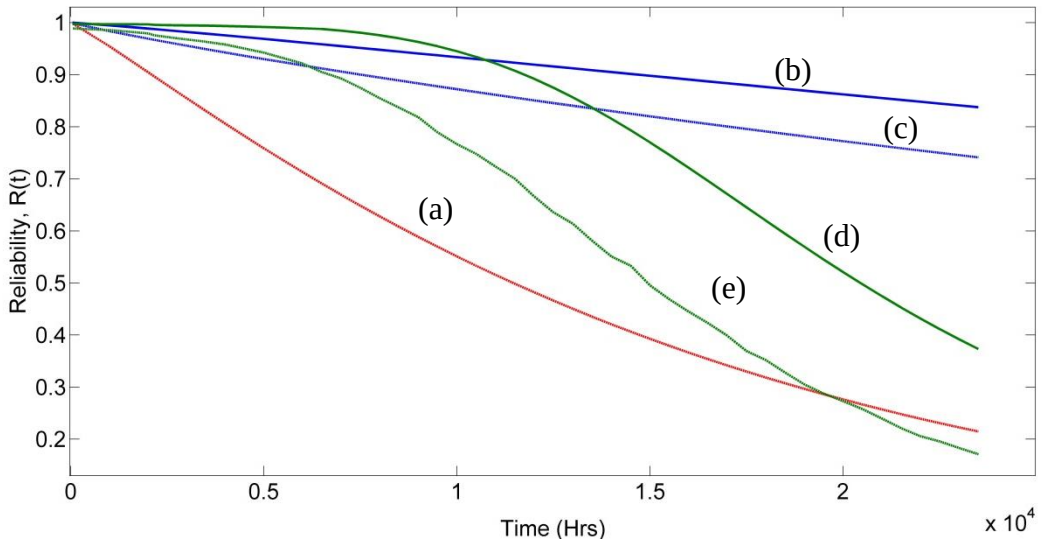


Figure 6 Reliability plot from (a) Nelson's Method (95% confidence) (b) Weibull plot eqn. (c) Weibull plot eqn. lower bound (95% confidence) (d) PDA model (e) PDA model lower bound (95% confidence)

Table 2 Comparison of Reliability Assessment Methods

| | Nelson's Method (95% confidence) | Weibull Plot Eqn | Weibull Plot Eqn lower bound (95% confidence) | PDA Model | PDA Model lower bound (95% confidence) |
|---|---|---|---|---|---|
| R at 2000hrs | 0.9050 | 0.9886 | 0.9692 | 0.9947 | 0.9792 |
| R at 3000hrs | 0.8550 | 0.9821 | 0.9557 | 0.9910 | 0.9699 |
| R at 8000hrs | 0.6280 | 0.9478 | 0.8945 | 0.9757 | 0.8548 |
| % Δ at 3000hrs | -5.52% | -0.66% | -1.39% | -0.37% | -0.95% |
| % Δ at 8000hrs | -30.61% | -4.13% | -7.71% | -1.91% | -12.70% |

For this bearing replacement interval extension from 2000hrs to 3000hrs, the results from the various methods are shown in Table 2. In practice, maintainers can choose between these methods for evaluating the reliability. Although it is the most conservative, Nelson's method is not desired as the lower bound is too wide. For the low usage time where PM interval of 2000hrs to 3000hrs being evaluated, the Weibull plot equation lower bound approach is recommended as it provides a more conservative assessment compared to the PDA model. However, if the usage time is higher with PM interval 8000hrs being evaluated, the PDA model lower bound would provide a more conservative result.

### 4.9 Chapter Summary

From the case study, a conservative estimate of the change in reliability from interval extension can be obtained with limited available information. Realistic figures for bearing geometry and replacement times have been used in the case study so that the reliability assessment is reflective of in-service application. It is recognized that this work is not substantiated by experimental or field data. However, the goal is to develop a conservative estimate for the purpose of maintenance interval extension. This work could be applied in maintenance interval escalation work where the amount of extension needs to be quantitatively substantiated instead of being based on expert opinions which can be subjective. In this study, the reliability of a single bearing is considered and shown. However, the use of Monte Carlo testing, together with strict

series reliability model can be applied to transmissions systems as well. In a separate work, Zaretsky et al has extended their Monte Carlo Testing approach to assess reliability of turboprop reduction gearboxes [32]. Similarly, the methods proposed in this paper may be extended for assessment of such higher assembly system as well. Despite the growing prevalence in the use of HUMS capabilities for CBM, such methods for extending PM remains useful as not all components on the helicopter can be monitored through HUMS and PM will continue to play a significant role in helicopter maintenance.

### 4.10 Related Publication

- Reuben Lim, David Mba, (2013), *Estimating Lower Bound Reliability without failure data*, IMechE Part O, Journal of Risk and Reliability

# 5 TRANSITION TO CONDITION-BASED-MAINTENANCE

With the increasing use of HUMS on helicopters, there is much interest to replace PM tasks with CBM ones. The transition from PM to CBM however does not occur with the embodiment of HUMS on the helicopter alone. Most helicopter platforms in-service today were not designed with HUMS in mind and the maintenance packages would still be largely based on PM. The CBM task has to be developed and demonstrated to be safe and reliable before it can replace the PM task. In this chapter, an overview of CBM is carried out and the technical requirements to develop CBM tasks are focused upon. It should be noted that a host of other practical issues such as developing the maintenance manuals, organization setup and personnel training will have to be carried out as well but these would not be covered in this study.

## 5.1 Background on CBM and HUMS Development

In accordance with [33], CBM is defined as "*maintenance performed as governed by condition monitoring programs*" while condition monitoring itself is defined as "*acquisition and processing of information and data that indicate the state of a machine over time*". The concept of CBM for helicopters itself is not entirely new as the use of magnetic chip detector in transmissions and engines and Spectrometric Oil Analysis (SOA) of the lubricating oil are common features in existing helicopter designs. Both methods monitors the conditions based on the accumulated debris in the detector or particles in the oil over time. However, the fidelity of these methods are rather low as they are often effective only when the damage is significant. When vibration sensors was used for health monitoring of helicopter transmissions operating in the North Sea in the 1990s [34], it led to a growth in the use of HUMS to not only detect but to diagnose the faulty component and extent of damage as well. Since then, the capability of such HUMS has progressed considerably with advances in the development of sensors, data acquisition and signal processing technologies. The progress is not limited to the HUMS hardware itself but extends to the concepts in which they are applied and standards to assure safe implementation as well. For implementation policies on helicopter, the UK Civil Aviation Authorities (CAA) was the key proponent in the 1990s in advocating the use of HUMS to monitor transmission

condition. It was also the UK CAA that developed the first guidance material, CAP 753 [35] on implementation of HUMS on helicopters which preceded the more comprehensive ones by the FAA and US Army. In more recent developments, IMRBPB has started to include Structural Health Monitoring (SHM) within the MSG-3 document to guide development of CBM tasks for airframes maintenance [36].

For a CBM task to be developed, various processes have to be carried out before the embedded sensor data can be meaningful to support maintenance decision-making. The data processing flow for a CBM task based on ADS-79B and ISO 13374-1:2003 is shown in Figure 7. It should be noted that the processes within state detection and health assessment are often referred to as diagnostics. A literature review of the existing technologies and methodologies associated with these processes is carried out in this chapter. It should be noted that the relevant literature in each process can be very large; especially for diagnostics and prognostics, and thus the review here is not exhaustive in itself.

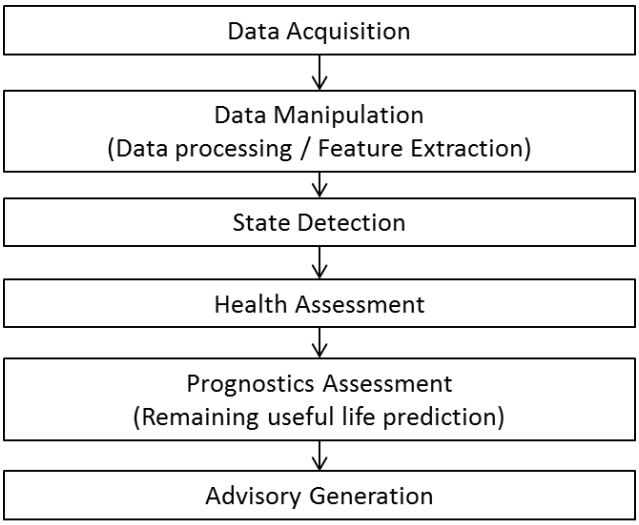


Figure 7 Data processing and information flow for CBM task [10; 37]

## 5.2 Data acquisition from HUMS Sensors

Through the years, a wide range of sensors have been developed for monitoring and fault detection and several surveys have been carried out by different authors and agencies on their effectiveness. The Federal Aviation Administration (FAA) carried out

one of the first survey for helicopter HUMS [38]. The National Aeronautics and Space Administration (NASA) performed several surveys [39-42] to apply HUMS ranging from gearbox to engine health monitoring. The Civil Aviation Administration (CAA) from the UK has also conducted a review of rotor system HUMS [43]. From those surveys, the types of sensors used for HUMS can be broadly grouped into (1) vibration, (2) oil-condition, (3) temperature and (4) strain-based sensors. Each group of sensors measures a different phenomenon exhibited when the component is behaving abnormally.

Vibration based sensors, typically accelerometers detects fault patterns in the vibration signal when defects are present. Amongst the sensors, the use of accelerometers is most commonly used for HUMS, especially for helicopter transmission systems. Oil-condition based sensors such as chip detectors detect abnormally large particles in the lubrication oil when excessive damage occurs. The inductance type debris detection which is based on oil-condition has been developed by Goodrich and GASTOP for HUMS application on the AH64 helicopter [44] and F22 aircraft [45]. It was shown to be effective in detecting gear and bearing damage but is unable to distinguish between the two when both share a common lubrication system [46]. Spectrometric oil analysis is another common oil-based technique used to determine the type and amount of metallic wear. Oil samples from in engines and gearboxes are regularly taken and sent to a laboratory where atomic emission spectrometer or atomic absorption spectrophotometer determines the type and concentration of particles in the oil [47]. Temperature based sensors measures increased operating temperature arising from friction due to abnormal wear. It is commonly used as safety alarms to detect rise in lubricating oil temperature in event of gearbox failure but not in HUMS as it is not suited more measuring the component temperatures itself. Strain based sensors does not directly detects defects but monitors applied loads on the component which causes the strain and is used for usage rather than condition monitoring. In [48], a wireless strain sensor from Microstrain® was installed and flight tested on the pitch link of the MH-60S helicopter. This allows the structural loads of the pitch link to be monitored and its fatigue live to be accurately assessed based on usage.

There are new sensors being developed and optic fibers and magnetic stress sensors are amongst them. Optical fiber strain sensor works based on the principal that the wavelength of the reflected light changes according to the induced strain on the optical fiber. The use of optical fiber strain sensors for structural health monitoring is gain popularity due to its light weight. In [49], NASA applied the use of optical fiber strain sensors to diagnose defects in the OH-58 planetary gearbox test rig. Magnetic stress sensor operates based on the principle of that the magnetic permeability of a ferromagnetic material changes as it is subjected to mechanical loading also known as the inverse magnetostrictive effect. The key advantage of this technique is that it does not require contact with the component it is monitoring. In [50], JENTEK Quadri-Directional Magnetic Stress Gage (QD-MSG™) is used to monitor applied torque, axial and bending loads on a rotating shaft.

### 5.2.1 HUMS airworthiness qualification

Besides ADS-79 which guides the CBM program, the HUMS sensor and data acquisition system has a host of airworthiness qualifications such as environmental, aircraft integration and safety standards to be met. A comprehensive review of these qualification requirements was carried out in [51] and shown in Figure 8. In lieu of these stringent requirements, helicopter HUMS qualification is a costly affair and only there is only a handful of commercially available HUMS system. It is interesting to note that the US Army has adopted ground rules that the HUMs system should not provide further cockpit indication that is not already present in the aircraft's baseline configuration [52]. The software certification and data processing standards are very high for cockpit indications so that information provided to pilots are accurate but they also require very stringent and costly testing. The ground rules meant that HUMS systems should be mainly used by maintainers on ground only and should not change the baseline risks that were accepted during the helicopter type certification. The HUMS which have been qualified to requirements in Figure 8 includes IMD-HUMS from Goodrich, ZING-HUMS (formerly known as IAC-HUMS) from Honeywell, IVHM from GE aviation and EuroHUMS from Eurocopter, all that of which uses accelerometers for condition monitoring.

| ASPECT OF CERTIFICATION | STANDARD/ GUIDANCE | COMMENTS |
|---|---|---|
| Software Development (on-board) | DO-178 | Software assurance level varies from A to D depending on criticality |
| Software Development (ground based) | DO-178 | Software assurance level varies from AL1 to AL6 depending on criticality |
| Hardware Development | DO-254 | Assurance level varies from A to E depending on criticality |
| Safety Consideration | ARP 4754/4761<br>MIL-STD-1629A | MIL-STD-1629A for FMECA |
| Environmental | DO-160 | Standard AMC |
| Integrity of Data | DO-200A | Applicable for data management system |
| BIT/BITE Function | MIL-STD-1591<br>MIL-STD-2165 | MIL-STD-2165 for testability |
| Integrated Diagnostics | ARINC 604<br>ARINC 624-1 | ARINC 604 for design and use of BITE<br>ARINC 624-1 for OMS design guidance |
| Verification & Validation guidance for ISHM | AC 29 M-15<br>SAE ARP 5783 | FAA Advisory circular for HUMS<br>Metric for evaluating diagnostics algorithms |
| Overall guidance for ISHM | ADS-79B | US Army guidance for HUMS |
| Interface across layers | OSA-CBM<br>OSA EAI | Implementation of the ISO-13374 functional specification |

Figure 8 Airworthiness certification basis for HUMS [51]

### 5.3 Data processing and Feature Extraction methods

Data processing is the analysis of the condition monitoring signals and interpreting it for decisions to be based upon. The meaningful descriptors of the processed signals in turns are referred to as features or condition indicators. The condition monitoring signals depends on the type of sensors being employed and vibration-based sensors is the most widely used in commercial HUMS system as shown above. There are several vibration-based data processing methods to date for faults diagnosis in rotating machines and they can be broadly classified into (1) time domain, (2) frequency domain and (3) time-frequency domain. A comprehensive review of these methods is discussed in [53] and the common techniques in each domain are discussed briefly here.

### 5.3.1 Time Domain Methods

Time domain methods use the descriptive statistics of the time-series vibration waveform itself to identify faults. As faults develops in such as spalls in bearings or shaft imbalance, the statistical behavior of the vibration waveform changes and is used to detect the fault. The common statistics used are the root-mean-square (RMS) and kurtosis as shown in Eqn (11-12) where $x_i$ is the vibration signal, $\bar{x}$ is the signal mean, $\sigma$ is the signal standard deviation and $N$ is the number of samples in the signal.

$$RMS = \sqrt{\frac{\sum_{i=1}^{N} {x_i}^2}{N}} \tag{11}$$

$$Kurtosis = \frac{\frac{1}{N}\sum_{i=1}^{N}(x_i - \bar{x})^4}{\sigma^4} \tag{12}$$

Common statistics however does not reveal the component that is defective within the gearbox or the nature of the defect. It can be unreliable when the measurement contains noise in the environment. As such, the time domain signals are processed to extract more information. Gear defects commonly use time synchronized signals to obtain features such as FM4 to detect damage in gear teeth. FM4 is calculated by the ratio of the kurtosis and the square of the difference signal and was developed to detect faults in gears. The difference signal is obtained by removing the primary gear mesh frequency and its harmonics from the acquired signal. Further descriptions of such features used for bearing and gear feature extraction are described in [54; 55].

Autoregressive (AR) modeling of the vibration time series is another popular time domain method and has been applied together with neural networks by Baillie D.C and Mathew J. to detect different types of bearing faults in [56]. The key advantage of AR modeling is that it does not require a long duration of the vibration waveform to be sampled. Besides trial-and-error applied in [56]**,** methods such as least-square estimation and Yule-Walker equations can be used to estimate the order of the AR model and its parameters but it can still be difficult to obtain an accurate model in practice.

### 5.3.2 Frequency Domain methods

For frequency domain methods, the use of the Fast Fourier Transform (FFT) of the vibration signal is widely used to identify fault frequencies in rotating machineries. The discrete form of the FFT is shown in Eqn (13), where $x_i$ is the vibration signal, *X(k)* is the frequency spectrum, *N* is the number of samples in the signal.

$$X(k) = \frac{1}{N}\sum_{i=0}^{N-1} x_i e^{-j\frac{2\pi ki}{N}} \tag{13}$$

When the vibration is digitally sampled, the sampled waveform will be discontinuous as the sample length will not be a multiple integer of the waveform's frequency content. This causes spectral leakage in the FFT and to address it, a windowing function is typically applied to weigh down the discontinuity [57]. For feature extraction, statistics such as the Root-Mean-Square (RMS) energy of the fault frequencies has been shown to be a simple but effective method to diagnose faulty bearings in [58; 59]. Bearing fault frequencies, however, can sometime be masked by more dominant gear mesh frequencies and envelope analysis is a popular technique used to identify them by demodulating the high frequencies impulses caused by the bearing faults [57]. Cepstrum analysis; the FFT of the log of the vibration spectrum, is another method used to obtain the bearing fault signal by separating the harmonics associated with the shaft and gears from the measured signal [57]. The key disadvantage of direct FFT spectra is that it can only handle stationary signals (i.e. the statistical properties of the signal does not change) as time information is lost during the transform. Another frequency domain method is Bispectrum analysis, which is a higher order statistics used to determine nonlinear interactions between frequencies within the signal [60]. It has been applied for fault detection in gearboxes [61; 62] and induction motors [63]. The disadvantage of this approach is the computational complexity and difficulty in the interpretation of the results [63].

### 5.3.3 Time-Frequency methods

Unlike frequency domain methods, Time-Frequency methods analyses waveforms in both time and frequency domain and is able to handle non-stationarity in the waveform signal. Common Time-Frequency methods which have been applied to fault detection for gears and bearings include the Short-time Fourier transform (STFT) [64] and Wigner–Ville distribution [65]. In STFT, FFT of small segments of the waveform signal is taken and combined so that changes in the spectrum can be viewed. This approach however suffers from low resolution due to the segmentation. Wigner-Ville distribution has higher resolution as it does not suffer from segmentation like the STFT but the result can be difficult to interpret due to interference from the transformation process [65]. Another time–frequency method is the wavelet transform which is actually a time-scale representation of the waveform signal. It has the advantage of producing high frequency resolution for low frequency signal and vice versa. The selection of wavelet is essential though and commonly used ones are Morlet and Haar [53]. In [66], wavelet analysis is adopted for fault detection of planetary gearbox using vibration data. In more recent developments, the Hilbert-Huang Transform (HHT) is an algorithm-based approach that empirically decomposes a waveform signal into its intrinsic oscillatory modes or Intrinsic Mode Functions (IMF)[67]. Hilbert transform is applied to these IMF to obtain the instantaneous frequency of the waveform as a function of time and has been applied to detect faults in bearings [68; 69]. The advantage of the HHT is that it's self-adaptive and does not require a wavelet selection. However, it lacks theoretical foundation and the processing time can be long due to the decomposition process [70].

## 5.4 State Detection

State detection serves to establish normal baseline condition of the monitored component and uses decision boundaries or threshold that distinguishes the component conditions when new measurements or features are acquired. In a way, state detection can be viewed as a classification problem to distinguish between the normal and abnormal condition of the component using the extracted features. The simplest approach to state detection is the commonly used statistical control process (SPC)

approach where an alarm is triggered if the signal is above a pre-defined threshold, i.e. 3σ from the mean. This is often used in initial stages of HUMS implementation where field and test data are not available for detection threshold setting.

For a two-class problem, the prediction outcomes can either be positive (P) or negative (N). If the outcome from a prediction is P when the true state is also P, then the prediction is a true positive (TP). If the true state is negative, then the prediction is a false prediction (FP) which causes false alarms. Conversely, a true negative (TN) is when both the prediction and the true state is negative and a false negative (FN) is when the prediction outcome is negative while the true state is positive. When defining the fault detection threshold, a balance needs to be achieved between the sensitivity to detect fault and the false alarm or false positive rates. If a threshold is set low, fault can be detected easily but this can lead to more false alarms arising from noise. Conversely, a high threshold will reduce false alarm but may lead to missed detection. From the regulatory guidance [9; 10], the false alarm rate has to be shown to be <5% from seeded testing or field data for the CBM task to be valid. In [14], Dempsey et al illustrated the use of such principals in signal detection theory for setting detection threshold. The statistical distribution or pdf of condition measurements (or extracted features) from normal and defective components are obtain and plotted as shown in Figure 9.

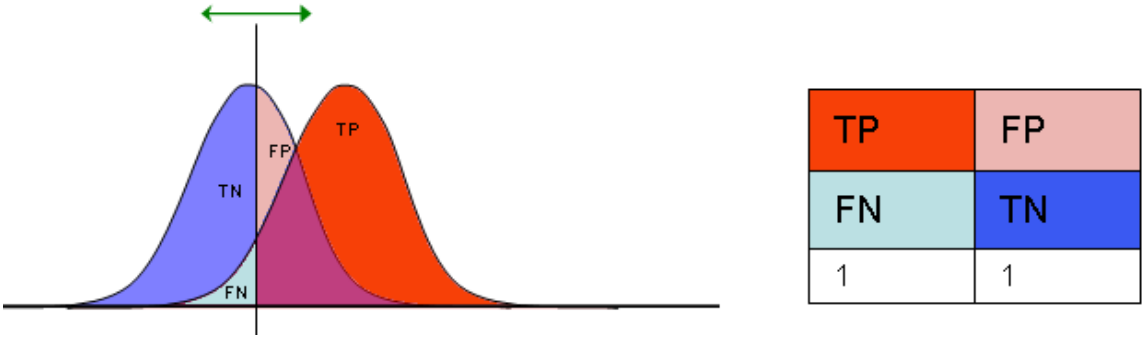


Figure 9 Statistical distribution of feature for normal and defective components in threshold setting [71]

The false alarm and missed detection rates with respect to the threshold can then be determined accordingly. Note that the probabilities of the pairs TP/FN and TN/FP each sum to one as shown in Figure 9. It can be seen from Figure 10 that increasing the

threshold to the right will reduce false alarm but increase missed detection accordingly. The use of receiver operating characteristic (ROC) curves is another common method used to display the same information [14] as shown in Figure 10. The advantage of using ROC curve is that it provides a more intuitive display of the performance of the thresholds by comparing the true and false positive detection rates.

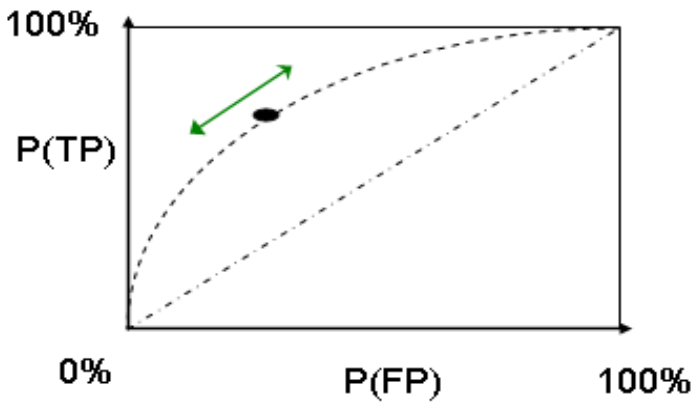


Figure 10 Receiver Operating Characteristic Curve [71]

## 5.5 Health Assessment

In state detection above, a single feature is used to classify a component into two (binary) state conditions; normal or abnormal. However, a component may be monitored by more than one sensor as a lone sensor may not adequately classify the health state. Signals from a lone sensor could also be processed in different ways to extract different features. Furthermore, a component can have multiple health states to be diagnosed due to different failure modes or stages of damage. Heath assessment is thus an extension of state detection and serves to diagnoses the current health of the component by considering all available state information from different sensors or features [37]. A good example of this is seen in [72] where both oil debris and vibration signals are used to detect gear damage. The classification problem however becomes multi-dimensional where different feature dataset has to be grouped corresponding to the component health states. Such a classifier; trained from available datasets, can then be used to map new data to determine the health state.

In the field of machine learning, there are many techniques available for multi-dimensional classification and there are two main groups depending on the nature of the classification problem. If the feature dataset used for training the classifiers are labeled, i.e. each instance within the set of training data corresponds to a known output (or a health state for diagnostic application), then supervised classification applies. If the dataset is not labeled, then the structure within the dataset has to be inferred and this is referred to as unsupervised classification. For diagnostic works on seeded tests, supervised classification is most commonly applied such as in [73-75] as the health states of the component can be observed directly and labeled against sensor measurements or extracted features as damage progresses. In datasets obtained from field defects however, the component's condition could not be observed directly i.e. the bearing within the helicopter transmission, and it can only be monitored indirectly via the sensor measurements. Therefore, the health states during damage progression are not available and unsupervised classification has to be applied to infer them instead.

### 5.5.1 Supervised Classification

There several types of supervised classifiers being used for bearing diagnostic and the most commonly applied ones are artificial intelligence (AI) based methods such as Artificial Neural Network (ANN) and Support Vector Machine (SVM). AI-based methods works by learning the patterns between the input features and output health states and can handle high non-linearity in the relationship. Other popular methods include Gaussian mixture models (GMM) and Hidden Markov Model (HMM) which are statistical based and k-nearest-neighbor (kNN) which is a non-parametric method. In [73], Gang N. et al applied a range of different classifiers; SVM, kNN, Linear Discriminate Analysis (LDA), Random Forest (a decision tree based classifier) and Kohonen Neural Network, on time domain statistical features from an induction motor. In that work, SVM, kNN and LDA outperforms the ANN based Kohonen Neural Network. In another comparative study conducted by Mendel et al [74], ANN, SVM and kNN were applied on features obtained from envelope analysis to diagnose bearing faults and it was concluded that ANN has slightly better performance. In [76], Gang Y. et al showed that GMM outperforms ANN in diagnosing bearing fault using wavelet transform features. As the nature of the experimental setup, defects and feature

extraction methods varies widely between studies, it can be difficult to compare the relative performance of one classifier against another as each classifier has its own strength and weaknesses.

### 5.5.2 Unsupervised Classification

In analyzing data from field defects, the transition in health states cannot be observed directly thus unsupervised classification is more suitably applied to investigate the health states prior to failure of the component using indirect sensor measurements and extracted features. The range of unsupervised classification algorithms are comparably smaller and are often an extension of supervised classification. The popular ones include hierarchical clustering, *k*-means, Self-Organizing Maps (SOM) and GMM. In all these algorithms, the aim is to find clusters within the datasets through similarity measures such as distances or density metrics between data instances within the data space. There are several types of distance metric used and the common ones are Euclidean and Mahalanobis distance. The advantage of Mahalanobis distance over standard Euclidean distance is that it takes into account the correlations between features in a multi-dimensional data set and is scale-invariant. A detailed description of such distance metrics can be found in [77].

Both hierarchical clustering and *k*-means are two common non-parametric, distance-based methods. Hierarchical clustering work by merging data instances into clusters such that the distance between classes are maximized while minimizing the within class distances. The hierarchy is built as clusters are progressively merged together at each agglomeration. This process is repeated till a specified number of clusters or distance between clusters is met. [78]. *k*-means clustering is similar but it operates by classifying observations into *k* clusters in which each observation belongs to the cluster with the nearest mean. Like hierarchical clustering, it too seeks to minimize the within cluster distances between observations [79].

GMM is a statistical-based approach that describes the dataset using a weighted sum of probability density function of multiple Gaussian distributions with each distribution representing a cluster or health state. In supervised classification above, GMM is

typically used to represent labeled classes with complex distributions as a weighted sum of Gaussian distributions. In unsupervised learning however, each weighted Gaussian distribution is assumed to represent a cluster. In [80], Roulias D. et al applied GMM to estimate the normal and abnormal health states in a run-to-failure bearing test. SOM is a type of ANN that performs unsupervised classification. It is commonly used to visualize high-dimensional features in low dimensions (typically two). The network or 'map' created can preserve the topological relationships within the feature set such that similar instances would be close together. In [81], Shahakupar S. et al compared the use of SOM with a hybrid approach of both *k*-means and hierarchical clustering for analyzing gene microarray data and found the latter to have better performance. Despite being termed "unsupervised", all these algorithms requires the number of clusters to be inferred to be pre-defined by user; with exception of hierarchical clustering which can also use a distance criterion.

### 5.5.3 Dimensionality Reduction

When the dimension of features is large (>10), the classification with good separation between groups can be difficult to achieve due to the infamous "curse of dimensionality" where the amount of data required for training increases exponentially with the data dimension. In such circumstances, data reduction methods such as Principle Component Analysis (PCA) and LDA are applied. PCA is more commonly adopted and it involves reducing the dimension of the features into a smaller set of principal components (PC) where the greatest variance in the data is projected onto the $1^{st}$ PC with decreasing variance in subsequent PCs. Most of the variance in the original features to be captured by a reduced number of PC and this allows better classification performance in the PC space [82]. In [83], Silvia M. Z. applied PCA to reduce a large dimension dataset of temperature, pressure and vibration measurements in a heat exchanger before classification is carried out for fault diagnosis.

## 5.6 Prognostics Assessment

Prognostics involve the analysis of symptoms (typically sensor measurements and extracted features) to predict future condition and remaining useful life (RUL) [33]. A

comprehensive survey of prognostics methods was conducted by different authors in [53; 84-86] and their works showed that there are three main approaches which are namely, (1) model-based, (2) AI-based and (3) statistical-based methods. Both AI and statistical based methods are also commonly referred to as data-driven methods [87] as HUMS data is often used. A brief overview of these methods is discussed herein.

### 5.6.1 Model-based methods

Model-based method uses mathematical representation of the system physical to determine the RUL based on its current state. The use of damage tolerance analysis in aircraft structures is one such example where the time to fracture can be calculated deterministically based on the current crack size, material properties and expected usage loading [88]. Such methods require does not require training data from past failure cases but knowledge of the systems degradation process is required. Model-based method is not widely applied in practice however as extensive testing is required to develop and validate the analytical models, especially when there are many different failure modes. In [89], a comprehensive bearing spall propagation model that is adaptive to usage loads was developed by Marble S. for RUL prediction. Besides physical models, known degradation behavior can be modeled as well. In [90], Gebraeel N. applied the use of exponential trend in the bearing vibration signal for RUL prediction. His method can be considered a hybrid of model and statistical-based method as the vibration signal is used to update the trend parameters to obtain a more accurate prediction. Similarly in [91], Li Y. used bearing vibration signal to adaptively update the parameters of the Paris Law equation for crack propagation in bearings.

### 5.6.2 Artificial Intelligence-based methods

While AI is used for classification of discrete states in health assessment, it is used for regression in prognostics to approximate continuous functions between feature data and a component's RUL. The technique employed remains largely unchanged from diagnostics applications with the key difference that extracted features are now mapped to the operating life of the component. As such, the output is no longer the health states but a predicted RUL instead. In [92], accelerated bearing run-to-failure test data was

used to train the ANN before it is applied for RUL prediction. In [93], ANN is used with logic rules to allow the RUL predict to be adaptive to environmental conditions. Besides ANN, SVM regression is another popular method used for prognostics and has the advantage of being less prone to over-fitting compared to ANN. In [94], SVM was applied for bearing RUL prediction. In general, AI-based prognostic methods has the same advantage of being able to model highly non-linear problems and suffer from the same drawback of requiring a large number of training data. For prognostic application, this drawback is more acute as datasets capturing the run-to-failure history of a component are even harder to come by. Hidden Markov Models (HMM) is another popular statistical-based AI technique that has been employed for prognostics. HMM models the condition of the component as a hidden state and they are inferred by observing the measurement data [95]. The advantage of HMM is that it can model different types of failures but again requires a lot of training data. In [96], Zhang X. H. et al applied HMM for bearing prognostics. In [97], Camci F. et al extended a variant of the HMM, the hierarchical HMM, for drill bit prognostics. The hierarchical HMM has improvements over the conventional HMM as it allows for multiple health states to be jointly represented to improve RUL estimation. HMM methods however have the same drawback of requiring a lot of training data.

### 5.6.3 Statistical-based methods

Statistical-based method uses the statistical behavior of the HUMS data to predict its future value. The trend extrapolation is the widest used of statistical methods for prediction of RUL [85]. A trend of the degradation path is obtained through regression methods and then extrapolated to determine the time taken to hit a pre-defined failure threshold [98; 99]. A weakness of this approach is that it assumes that usage condition and the degradation behavior paths between specimens remains the same. To account for variation in the degradation, Lu and Meeker [13] applied the use of random coefficient regression on several sets of run-to-failure test data for crack growth in aluminum to evaluate uncertainties in their degradation path trends. In that method, the regression parameters are treated as random variables and their distributions estimated using degradation paths from past failure cases. This approach however is seldom applied in practice due to limited data availability.

In [75], Siegel D. et al, applied the use of robust regression curve fitting approach to predict the RUL of helicopter bearings on a test bench. An exponential damage growth is adopted and regression is performed with each new measurement for its entire history since fault detection. A drawback of this approach is that a fixed degradation path model is assumed when it may vary under different failures modes or operating environment. The use of the entire measurement history can become unwieldy as the number of measurements becomes very large. This can also cause the regression model to be insensitive to more recent trends in the measurements. A way to overcome this is to use a sliding window of recent measurements for the regression so that it is more adaptive. However, a suitable window length to provide a robust yet adaptive prediction model would remain challenging.

A statistical-based approach that is comparatively recent is the use of recursive Bayesian techniques such as Kalman and particle filtering. In these techniques, the component condition or degradation path is modeled as a state-space model with known noise in the measurement data [100]. A more precise state is then recursively estimated with the noise filtered. Prognostic is performed by propagating the state into future time steps. KF has been applied by Lall P. et al to electronics prognostic for estimating remaining useful life of ball grid array connections [101; 102] and similarly by Celaya J. to electrolytic capacitors [103]. Particle filtering is based on the same concept but it allows for non-linear dynamical process and can accommodate non-Gaussian noise [104]. The use of particle filtering however can be very computationally intensive as Monte-Carlo simulation is employed heavily in the procedure to estimate the non-gaussian distributions.

A problem faced in practice that is common when applying all of the above statistical-based methods is that the RUL prediction accuracies tend to be poor in the initial period after fault detection. This is a reason why confidence bounds of the RUL prediction are an important piece of information to aid in decision-making. Expert input is then required to decide when the RUL prediction is reasonably accurate based on past experience or when an acceptable confidence bounds is achieved.

## 5.7 Advisory Generation

Advisory generation involves the use of the diagnostic and prognostic information gathered for maintenance decision-making and planning. It should be noted that even if highly accurate prediction of RUL is achieved, maintainers are unlikely to perform the task just before the predicted RUL as that can be disruptive to flight schedules. Instead, the RUL is often used to allow time for logistic preparation such as spares demand and the task itself would typically be align with other PM tasks to reduce aircraft downtime. In addition, the lower confidence bound of the RUL prediction would be of more importance to maintainers as it has higher safety implications on the aircraft.

## 5.8 CBM applications in helicopters

In this section, a survey of existing literatures on the use of HUMS for CBM is carried out. Only cases where the HUMS application is validated through component seeded tests or field defects are considered such that regulatory guidelines in [9; 10] are met. A survey of cases where faults were successfully detected on a helicopter platform is shown in Table 3 along with the associated data processing technique. It can be seen that all of the cases examined in this study were from military aircraft where HUMS system is most widely implemented. From these cases, the most commonly adopted feature for bearing defects is the vibration energy at the bearing defect frequencies obtained from the FFT spectra. For gear defects, the FM4 feature was the most adopted feature that is obtained from the kurtosis of the residual vibration signal. Although the number of cases in Table 3 shows that HUMS can be effective in detecting faults, the nature of the findings in those case studies are isolated and lack repeatability to demonstrate the consistency of both the sensor and the diagnostic algorithms. In lieu of this, more advance AI and statistical methods are not employed. No prognostics work was published as well in both civil and military helicopters. As such, the existing literature lacks case studies in which the end-to-end process involved in developing a CBM task is demonstrated. Such case studies are essential as it helps to build user confidence in the application of HUMS for CBM.

Table 3 Published works on helicopter fault detection from seeded testing or field defects

| Aircraft | System | Fault | Method | Feature | Validation | Ref. |
|---|---|---|---|---|---|---|
| AH64D | Honeywell | Aft and Fwd Hanger Bearing wear & corrosion | Frequency Analysis | Bearing Energy | Seeded testing | [14] |
| AH64D | Honeywell | Main Swashplate bearing broken cage & spalling | Frequency Analysis | Bearing Energy | Field defect | [13] |
| AH64D | Honeywell | Nose gearbox bevel gear | Time Analysis | FM4 | Field defect | [14] |
| AH64D | Honeywell | Tail Rotor gearbox bevel gear tooth crack | Time-Frequency analysis | - | Seeded testing | [105] |
| AH64D | Honeywell | APU Clutch failure | Time Analysis | Peak vibration | Field defect | [13] |
| H-60 | Goodrich | Tail Rotor Bearing | Frequency Analysis | Bearing Energy | Seeded testing | [17] |
| H-60 | Goodrich | Tail Rotor Gear scoring | Time Analysis | FM4 | Field defect | [14] |
| H-60 | Goodrich | Main gearbox Bevel Gear coating anomaly | Time Analysis | Residual Kurtosis | Field defect | [14] |
| H-60 | Goodrich | Hanger Bearing | Frequency Analysis | Bearing Energy | Seeded testing | [17] |
| H-60 | Honeywell | Oil Cooler Fan Bearing spalling and pitting | Frequency Analysis | Bearing Energy | Field defect | [18] |
| H-60 | - | Planet gear carrier crack | Time Analysis | FRMS[1], NSDS[2] | Seeded testing | [106] |
| AS332 | Eurocopter | Tail Driveshaft double Bearing spalling | Frequency Analysis | Bearing Energy | Field defect | [19] |
| OH-58 | - | Main gearbox input pinion tooth crack | Time Analysis | FM4 | Seeded testing | [107] |
| CH-47 | Honeywell | Main rotor swashplate bearing | Frequency Analysis | Bearing Energy | Seeded testing | [108] |

1: Filter Root Mean Square; 2: Normalised Sum of Difference Signal

### 5.9 Chapter Summary

From the above review, it can be seen that diagnostics and prognostics methods are very widely researched. Sensor technologies have comparatively less coverage in CBM

literature but it should be noted that an effective sensor reduces the need for complex diagnostic and prognostic algorithms to reduce measurement noise or infer the component state. As mentioned by Sikorska et al [85], most of the researches in diagnostics and prognostics are performed in a laboratory environment with seeded fault testing and there are little published works on complex components exposed to normal operating environments, especially for helicopters. Notably, there are very limited demonstrated cases whereby the diagnostics and prognostics methods are applied on in-service helicopters to develop a CBM task in compliance to regulatory guidance.

# 6 CBM TASK DEVELOPMENT FOR AH64D TAIL ROTOR GEARBOX BEARING

In this chapter, the end-to-end process for developing a CBM task is demonstrated using in-service maintenance and HUMS data collected on the AH64D helicopter belonging to the Republic of Singapore Air Force (RSAF). Vibration-based condition monitoring data from different helicopters with in-service defects found on their Tail Rotor Gearbox (TRGB) are correlated with tear down inspection findings. Frequency domain method using descriptive statistics of the FFT spectra is then extracted as features and despite its simplicity, it is shown to be effective for diagnostic and prognostic of bearing degradation within the gearbox in the field environment.

## 6.1 AH64D HUMS Description

The acquisition of data on the RSAF AH64D helicopter is carried out through the IAC-HUMS developed by Honeywell. It consists of 18 accelerometer sensor measuring vibration levels on different transmission and engine components as shown in Figure 11. The on-board systems measures vibration levels whenever the aircraft is on ground with rotors at flat pitch and rotating at 101% RPM (FPG101). This provides a controlled flight condition in which the vibration measurements are taken. The accelerometer measurement on the TRGB is acquired at a sampling rate of 48 kHz with a window size of 16,384 data points (~0.34 sec) and filtered using a Hanning window to reduce spectral leakage. Sixteen sets of measurements are further asynchronously averaged with no overlapping applied to reduce sporadic noise in the signal. Due to limited onboard data storage capacity, the time domain data are not stored for post flight downloads and only the FFT spectrums are available for further processing and analysis. Condition Indicators (CI) are processed on-board using both raw time domain data from the accelerometers and the Fast Fourier Transform (FFT) spectrum. For each flight flown, a text file containing the FFT data for all the accelerometers' measurements is generated and it may contain more than one set of measurement as the FPG101 condition may be met several times in a flight. As the number of files to extract is large, a MATLAB script is written to read and extract the FFT data for a selected sensor.

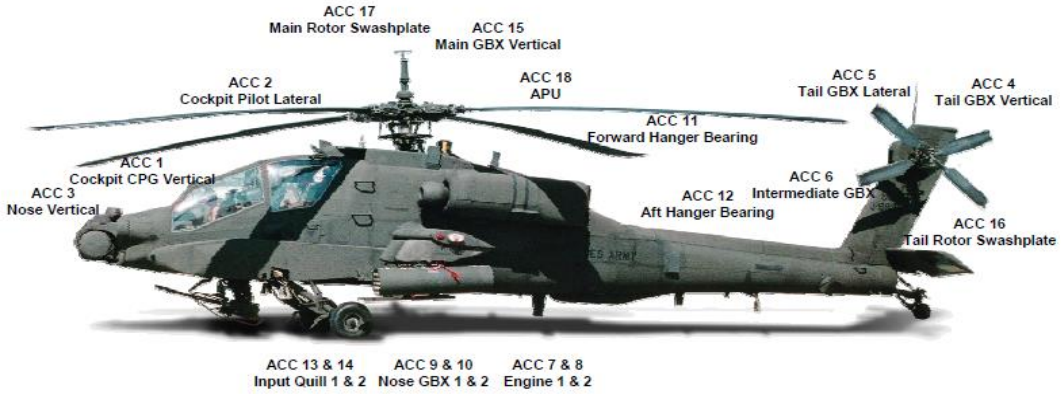


Figure 11 Accelerometer locations on AH64D HUMS

## 6.2 AH64D Tail Rotor Gearbox

The analysis discussed in this paper focuses on the TRGB output shaft thrust bearing in the AH64D helicopter. The TRGB is grease lubricated single stage gearbox and serves to transmit drive torque from the intermediate gearbox to the tail rotor system. An assembly drawing of the gearbox and the location of the accelerometers used to monitor this component are shown in Figure 12. For the TRGB, there are two accelerometers measuring both vertical and lateral directions.

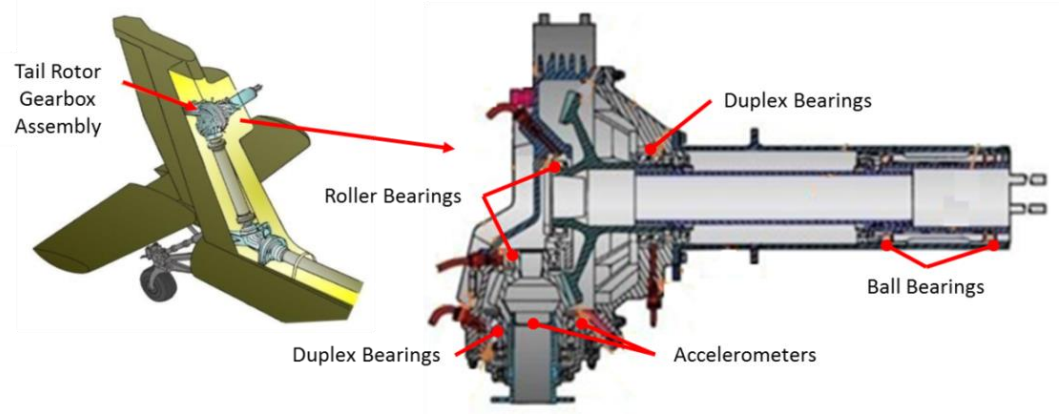


Figure 12 AH-64D tail rotor gearbox location & assembly layout [4]

### 6.2.1 TRGB Bearing Defect

From maintenance records, three TRGB was found with grease leaking from the output seals and upon further inspection, the output shaft was found with radial play. These

TRGBs for the three helicopters had accumulated 1204, 1171 and 962 flying hours respectively. They were then removed for disassembly and further teardown inspection. After completed disassembly, the ball bearings inside the outboard shaft were found to have extensive pitting on ball elements and spalling in the bearing races as shown in Figure 13. Pitting is a phenomenon in which small pits approximately 0.1 mm in depth are formed on the raceway surface by rolling fatigue and occurs earlier in the bearing damage progression stage. In comparison, spalling (a.k.a flaking in some literatures) is a phenomenon in which the bearing surface turns scaly and peels off due to repeated contact stress received on the raceway and rolling surface during rotation [109].

Figure 13 (Top left) Pitting on ball bearing elements, (Bottom left) Spalling on Inner race, (Top right) Spalling on Outer race, (Bottom right) Wear debris from removed grease sample

The running path of the wear pattern in the outer race is axially displaced and the pattern in the inner race is the widest in the radial load direction. This is characteristic of bearing wear under both axial and radial loads as guided in [110]. There was significant amount of wear debris found in the grease and it was evident that the radial play of the shaft was caused by the deteriorated bearing. Evidence of heat oxidation was found on the quill shaft as well. It is likely that the damage in the bearing is caused by corrosion which initiated from moisture intrusion through leaking output seals; a

common defect reported for the AH64D TRGB which was also reported in [105]. When grease leaks are found, the leak could not be repaired in the field and an unscheduled replacement of the TRGB has to be carried out. As the TRGB is a critical flight component, the replacement will require follow-on checks and functional check flights as well. Such unscheduled replacement causes aircraft unavailability and significant man-effort for recovery. As such, it is desired for such defects to be detected earlier and for the replacement to be performed during scheduled maintenance.

### 6.2.2 TRGB Gear Mesh Frequencies

As the amplitude of these gear mesh frequencies are much higher, they can 'drown' out features at other frequencies which may indicate the bearing's condition. In order to mitigate this, known Gear Mesh Frequencies (GMF) are identified from the FFT spectrum and removed if necessary. From [28], the GMF for the TRGB is calculated using Eqn (14) and the gear transmission data shown in Figure 14. The calculated GMF are listed in Table 4.

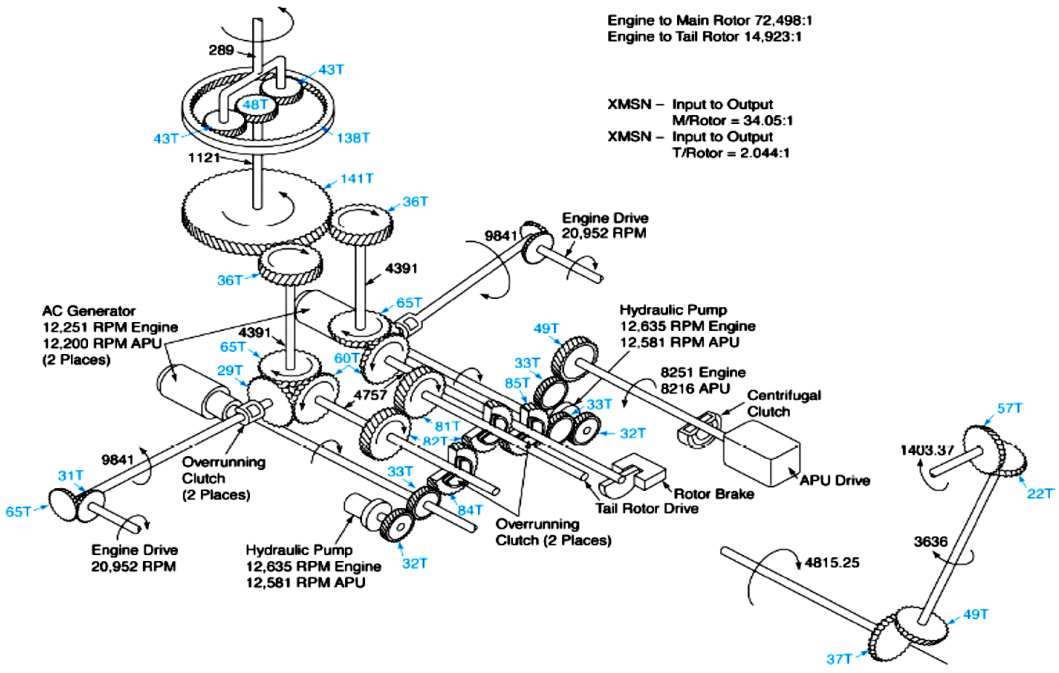


Figure 14 AH64D Transmission overview

$$Gear\ Mesh\ Frequency\ (GMF) = Number\ of\ gear\ tooth \times Shaft\ RPM \quad (14)$$

Table 4 Dominant Tail Rotor Shaft and Gear Mesh Frequencies

| Description | Frequency (Hz) |
|---|---|
| Tail Rotor Shaft Frequency | 23.6 |
| Tail Drive Shaft Frequency | 80.3 |
| Tail Gearbox Gear Mesh Frequency | 1347 |
| Intermediate Gearbox Gear Mesh Frequency | 3000 |

### 6.2.3 TRGB Bearing Defect Frequencies

As mentioned in brief earlier, the defect frequencies from the bearing are of interest in vibration analysis as defects developing within the bearing sub-components may exhibit itself with increasing amplitudes at these frequencies. From [29], the defect frequencies of a bearing with a given geometry are as shown in Eqn (15-18). The calculated defect frequencies for the tail rotor bearing are as shown in Table 5.

$$Ball\ Pass\ Frequency\ Inner\ Race\ (BPI) = \frac{Z}{2}\left(1 + \frac{d}{d_e}cos\alpha\right)f \quad (15)$$

$$Ball\ Pass\ Frequency\ Outer\ Race\ (BPO) = \frac{Z}{2}\left(1 - \frac{d}{d_e}cos\alpha\right)f \quad (16)$$

$$Ball\ Pass\ Frequency\ (BF) = \frac{1}{2}\frac{d}{d_e}\left(1 - \left(\frac{d}{d_e}cos\alpha\right)^2\right)f \quad (17)$$

$$Cage\ Frequency\ (FT) = \frac{1}{2}\left(1 - \frac{d}{d_e}cos\alpha\right)f \quad (18)$$

Table 5 Tail Rotor Gearbox Bearing Defect Frequencies

| Description | Defect Frequency (Hz) |
|---|---|
| Ball Pass Frequency Inner race, BPFI | 294 |
| Ball Pass Frequency Outer race, BPFO | 244 |
| Ball Pass Frequency, BF | 107 |
| Cage Frequency, BF | 12.8 |

## 6.3 Spectral Analysis of vibration data

The available measurements from the three helicopters prior to and after replacement of the defective TRGB were obtained for analysis. For two of the helicopters, gaps in the data history exist but it does not affect the study significantly as the trends from the FFT spectrum plots can still be clearly observed. HUMS data from another helicopter with a serviceable TRGB and with similar operating hours were also obtained for comparison with the defective gearbox. In this study, the HUMS data from the lateral accelerometer are used as its vibration signature showed a clearer response compared to the other two accelerometers. The reason for this is not investigated here though it is likely that there is less noise in the lateral direction in the environment. In their work using the AH64D Tail Rotor Test Rig, Goodman et al [4] has also observed that the lateral accelerometer is more sensitive to conditions within the TRGB.

Figure 15(a) shows the Time-Frequency plot of the acceleration FFT spectrums against flying hours for a serviceable TRGB. Figure 15(b) shows a magnified view at a lower frequency range that shows the evident vibration signatures which includes the Tail Rotor Shaft Frequency and its harmonics, the Gear Mesh Frequency (GMF) of the TRGB and the Intermediate Gearbox (IGB) frequency, together with their harmonics. Sidebands modulated at the Tail Rotor Pylon Shaft Frequency can also be observed surrounding the TRGB and IGB GMF. Figure 15(c) shows a snapshot of spectral plot at $t = 800\ FH$ and the mentioned frequency contents. From Figure 15, the magnitude of the spectral peaks at the GMF and their sidebands is stationary and does not show any trends over time for a serviceable TRGB.

In comparison, Figure 17 shows the similar Time-Frequency plots of the FFT spectrum against time for the defective TRGB #1. The dominant gear mesh and sideband signatures are also present but several fault patterns that differ from the serviceable TRGB plot are apparent. A key observation is that there are distinct changes in fault patterns at different frequency bands. From Figure 17(a), it can be seen that there is a steady increase in spectral peaks within the *0 - 5 kHz* band. This vibration energy at this lower frequency band then falls and this is followed by increase in spectral peaks at higher frequency band above *10 kHz*. Notably, spectral peaks spaced at the output shaft

BPFI and accompanied by sidebands modulated at 1 x tail rotor shaft frequency can be readily observed in Figure 17(b).Figure 17(c) shows the snapshot of the FFT spectra at 700FH. These fault patterns in the FFT spectrum are also consistently observed in TRGB #2 and #3, which shared the similar bearing defect findings as shown in Figure 18 and Figure 19.

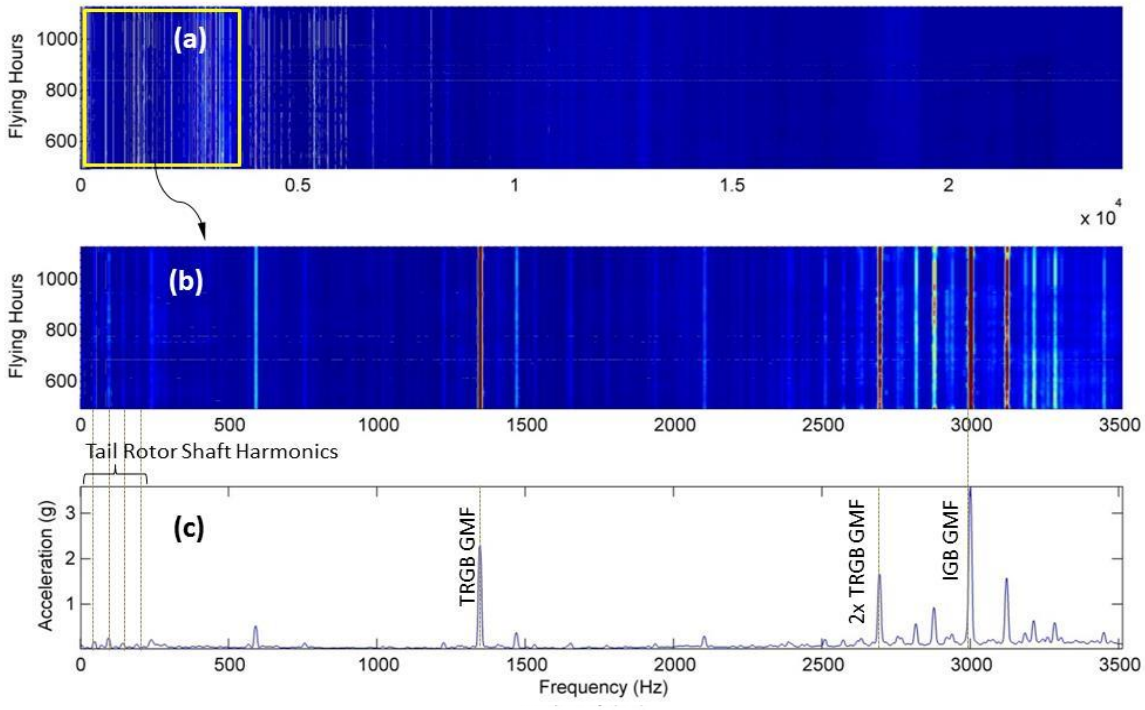


Figure 15 Time-Frequency plot of acceleration spectrum across flying hours for a serviceable TRGB

Based on the fault patterns seen in Figure 17 to Figure 19, the bearing has different stages of degradation. Such change in the pattern of the vibration spectrum as bearing deteriorates through different stages has been described in [111; 112] and shown in Figure 16 as follows:

- In Stage I, micro-defects and crack initiation causes ultra-high frequency activities. These activities are typically monitored using Acoustic Emission as such as in [113] rather than accelerometers.
- In Stage II, the micro faults develops into pits which begins to excite bearing element parts resulting in signals associated with their natural frequencies to be

appear. Enveloping analysis is commonly used to demodulate a selected high frequency bandwidth of the FFT spectra and extract the bearing defect frequencies in this stage.

- In Stage III, the pits become larger and the fundamental bearing defect frequencies and their harmonics can be observed from the FFT spectra. Depending on the extent of the damage, these frequencies can be modulated by the shaft frequency and be observed as sidebands as well.
- Stage IV is the final condition before bearing catastrophic failure. As the defect size becomes widespread or as multiple defects merges, the bearing elements vibrate more randomly with the higher clearances. The localized defects may also have 'smoothen' out due to wear which reduces the signature of the periodic vibration as described in [114]. As such, the distinct bearing defect frequencies diminishes as an increase in noise floor or 'haystack' rises in the higher frequencies ranges.

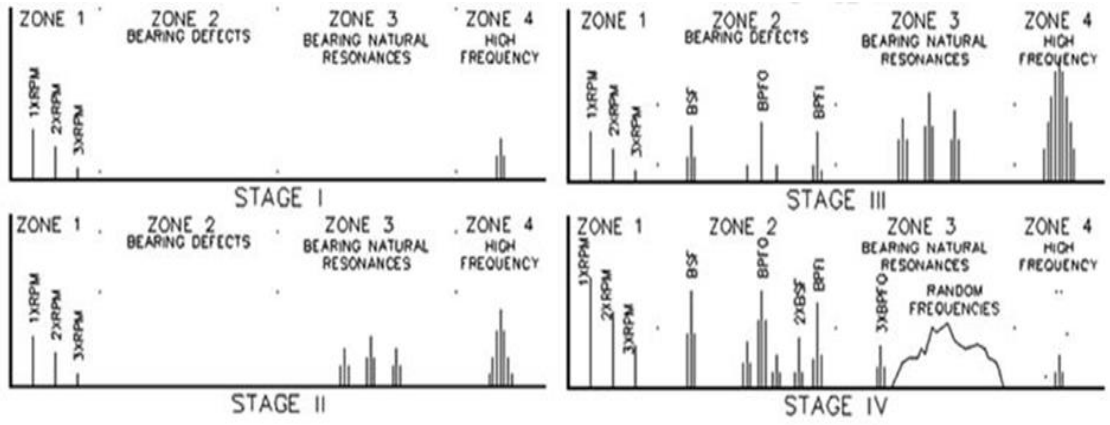


Figure 16 Fault patterns of bearing damage stages [111]

All three TRGB have the same reported defects of grease leak and free play in bearings. From their FFT spectrum plots shown in Figure 17 to Figure 19, the fault patterns for Stage III and IV bearing damage is evident. The fault patterns that suggest the transition from localized to generalized damage in the bearing.as follows:

- Stage III Fault Pattern. As localized damage initiated by corrosion grows within the bearing inner race, the ball bearing elements passing over the defect will cause periodic vibration that are observed as peaks spaced at the Ball Pass

Frequency Inner-race (BPFI). These peaks spaced at the output shaft BPFI and accompanied by sidebands modulated at 1 x tail rotor shaft frequency as shown in Figure 17(c). From [112], this is due to the bearing defect frequency acting as carrier frequencies for the shaft speed frequency. The presence of multiple harmonics of these peaks strongly suggests defects in the bearing inner race. This is further ratified from the severe spalling pits found in the inner race as seen in Figure 13 above.

- Transition from Stage III to Stage IV. As the defect size becomes widespread or as multiple defects merges, the damage becomes generalized and the bearing elements vibrate more randomly with the higher clearances. The localized defects may also have 'smoothen' out due to wear which reduces the signature of the periodic vibration as described in [114]. This damage mechanism can again be supported by Figure 13, where it can be seen that the damage is widespread in the bearing races and the rolling elements. As such, the distinct BPFI peaks in the lower frequency band falls as the vibration energy in the higher frequency band or 'haystack' rises as shown in Figure 17. This behavior has been described by Qiu H. in [112] but to the author's knowledge has not been demonstrated in published literature.

- Stage IV Fault Pattern. After the fall in BPFI peaks, several peaks are observed in the higher frequency range at 10 – 15 kHz. The frequencies at which these peaks occur are similar in Figure 17 to Figure 19. This suggests that the peaks are not due to random noise and should correspond to the natural frequency of the bearing sub-components, which resonates due to increased impacts within the bearing. Concurrently, a broad-based peak centered at 18.5 kHz with a bandwidth of 1.5 kHz was noted be have increased relative to the broad frequency band. This is again evidently consistent in the three gearboxes as seen in Figure 17 to Figure 19. These peaks should correspond to the natural frequency of the gearbox assembly which is also rising due to increased impacts within the bearing.

- Normal Condition after replacement. When the TRGBs were replaced, the fault patterns are no longer present and the spectrum reverts to that of a serviceable TRGB with the associated gear mesh frequencies as seen in Figure 17.

The repeatability in the fault pattern strongly suggests that the bearing damage progression for this failure mode is consistent and can be reliably monitored. Furthermore, the stages of the bearing degradation can be monitored through the fault patterns at different frequency bands. Due to in-availability of time domain data, envelope analysis to detect incipient fault in Stage II bearing damage cannot be carried out. However, it is evident that spectral analysis is effective in diagnosing a defective bearing in the TRGB.

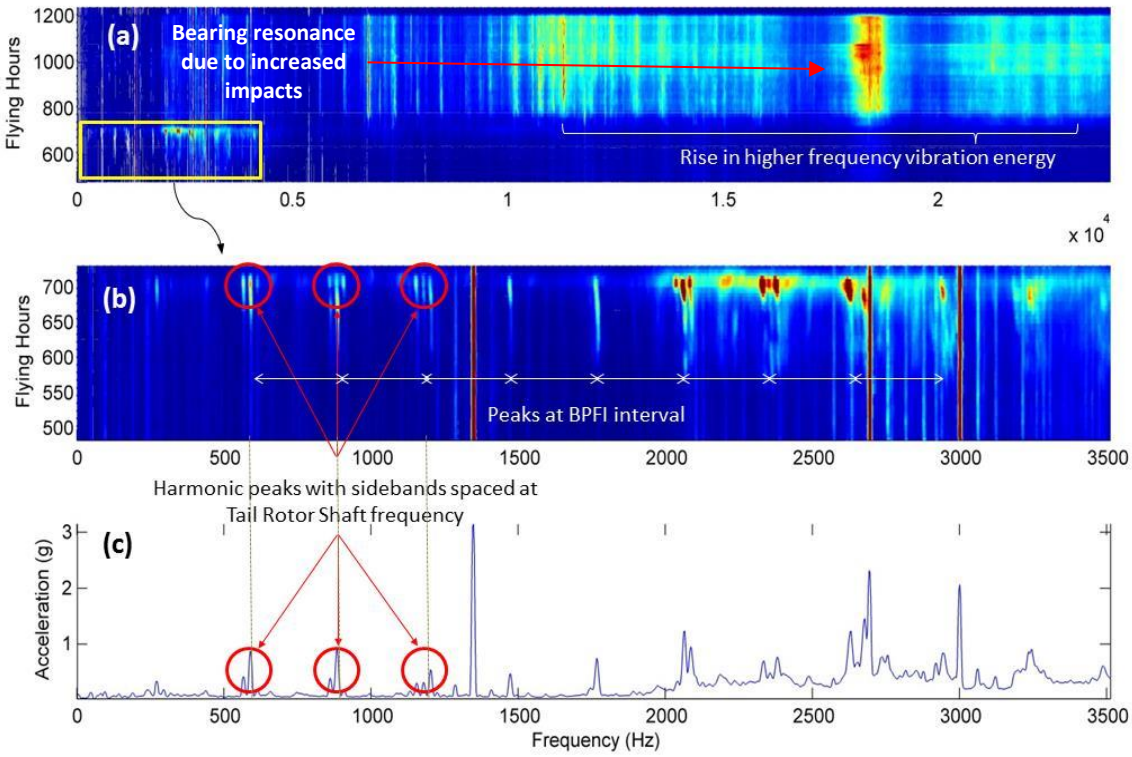


Figure 17 Time-Frequency plot of defective TRGB #1 acceleration FFT spectrum over time

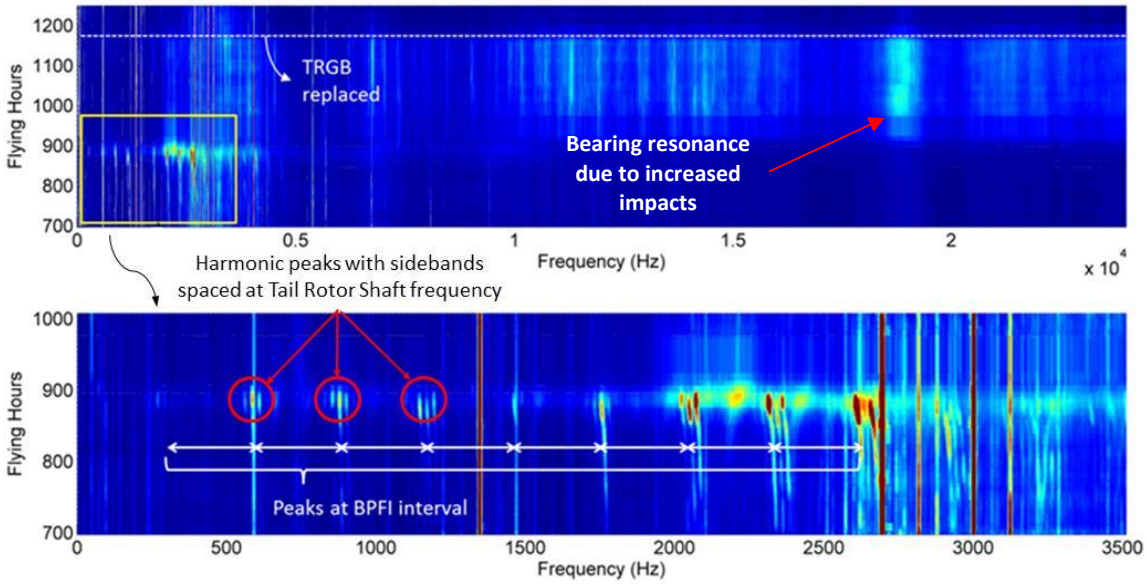


Figure 18 Time-Frequency plot of defective TRGB #2 acceleration FFT spectrum over time

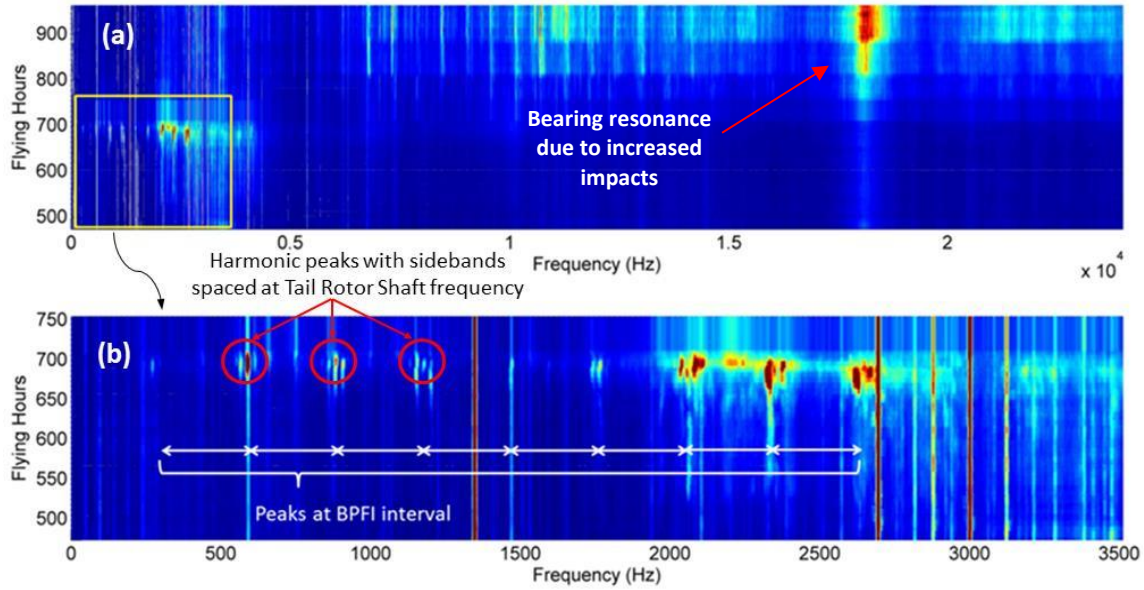


Figure 19 Time-Frequency plot of defective TRGB #3 acceleration FFT spectrum over time

## 6.4 Feature Extraction and Selection

From the spectrum plots, descriptive statistics are extracted as features for diagnosis of the bearing health state and prognosis of remaining useful life. As the fault patterns can be easily observed in the spectrum plots, the use of bearing RMS energy is an effective measure of the bearing condition. For the two degraded states; localized and generalized damage, a feature is developed for each state. Other statistical features such as the commonly used Peak amplitude, RMS energy and kurtosis of the entire FFT spectra were also extracted for comparison and selection. The various feature trend plots for the three TRGB are shown in Figure 20 to Figure 22.

### 6.4.1 Low Band Bearing Energy

For localized damage, the feature adopted is the RMS energy of the frequency magnitudes in the low frequency band of 250 – 2500 Hz to capture the peaks of the BPFI harmonics shown in Figure 17. A rejection band of 0 to 250 Hz and 1250 – 1600 Hz are applied to eliminate effects from the tail shaft frequency, the dominant gear mesh frequencies and sidebands calculated earlier. This feature captures the vibration caused by localised damage within the bearing. For the low band bearing energy shown in Figure 20(b) to Figure 22(b), it can be seen that it rises exponentially before falling back to normal levels. The exponential rise in vibration energy is commonly seen in bearing tests as shown by Harris and Kotzalas in [115]. The subsequent drop in the vibration energy is less frequently seen but tests on bearings performed by Dempsey et al [116] and Williams et al [114] had also shown similar drops. In all three instances, the low band bearing energy feature is encouraging as they rises monotonically to a similar peak value of ~1.5$g_{rms}$ which is desirable for setting the failure threshold in subsequent prognostic application. The key drawback of this feature is that it falls back to normal levels and this can give the false indication that bearing condition is normal. There are periods in the data history that are missing due to download errors or ground station unavailability but the overall trend of the data from the gearboxes can still be clearly observed.

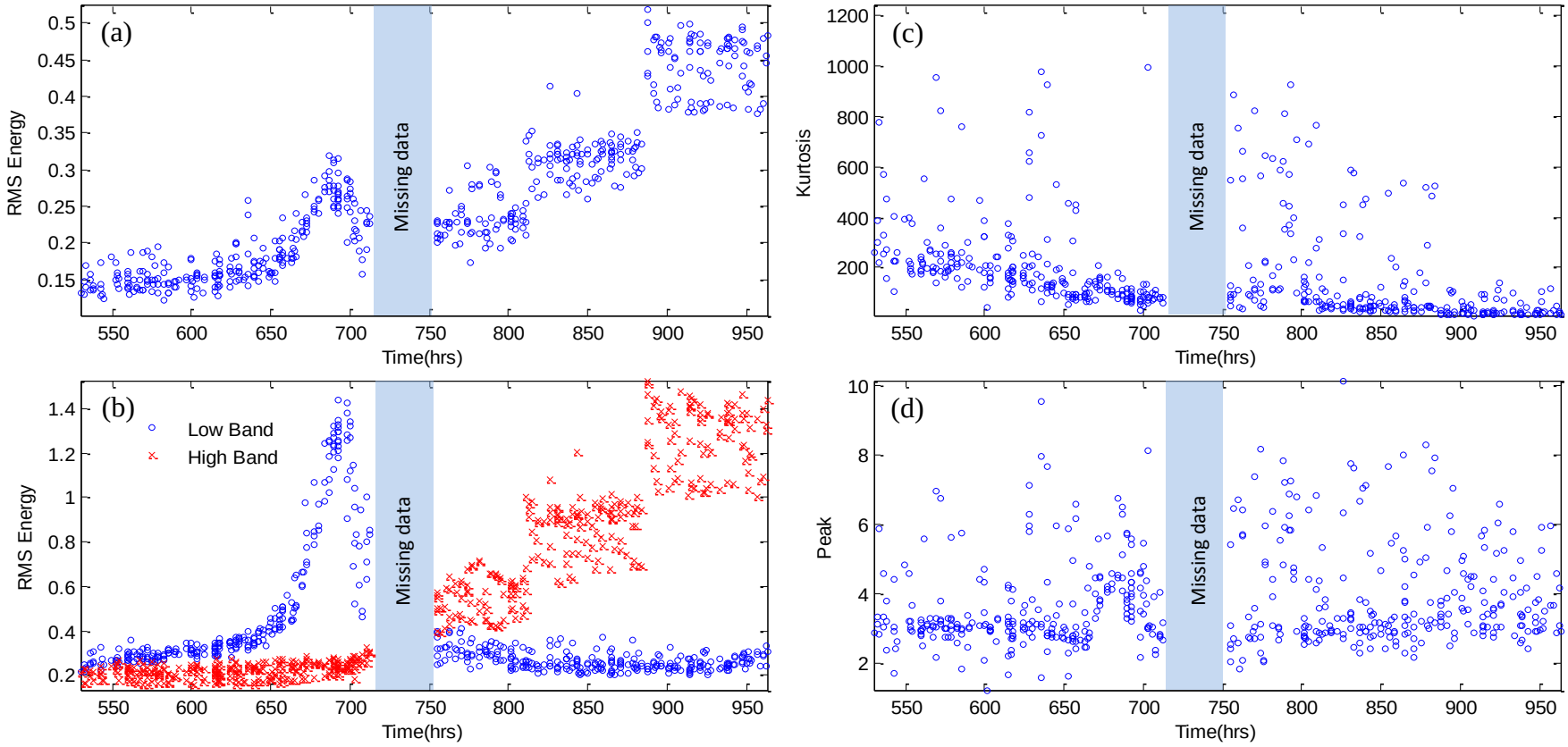


Figure 20 Extracted Features from TRGB 1.

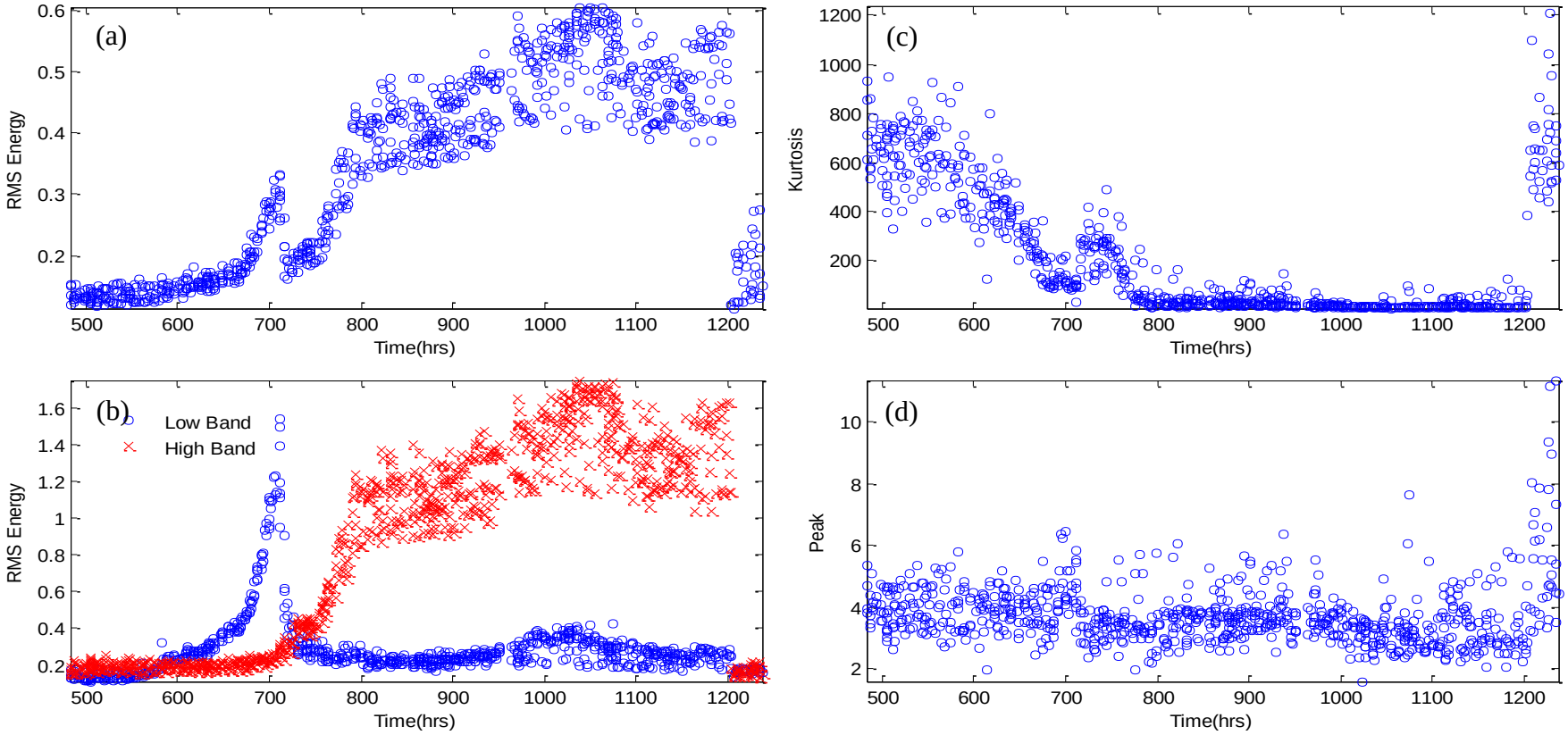


Figure 21 Extracted Features from TRGB 2.

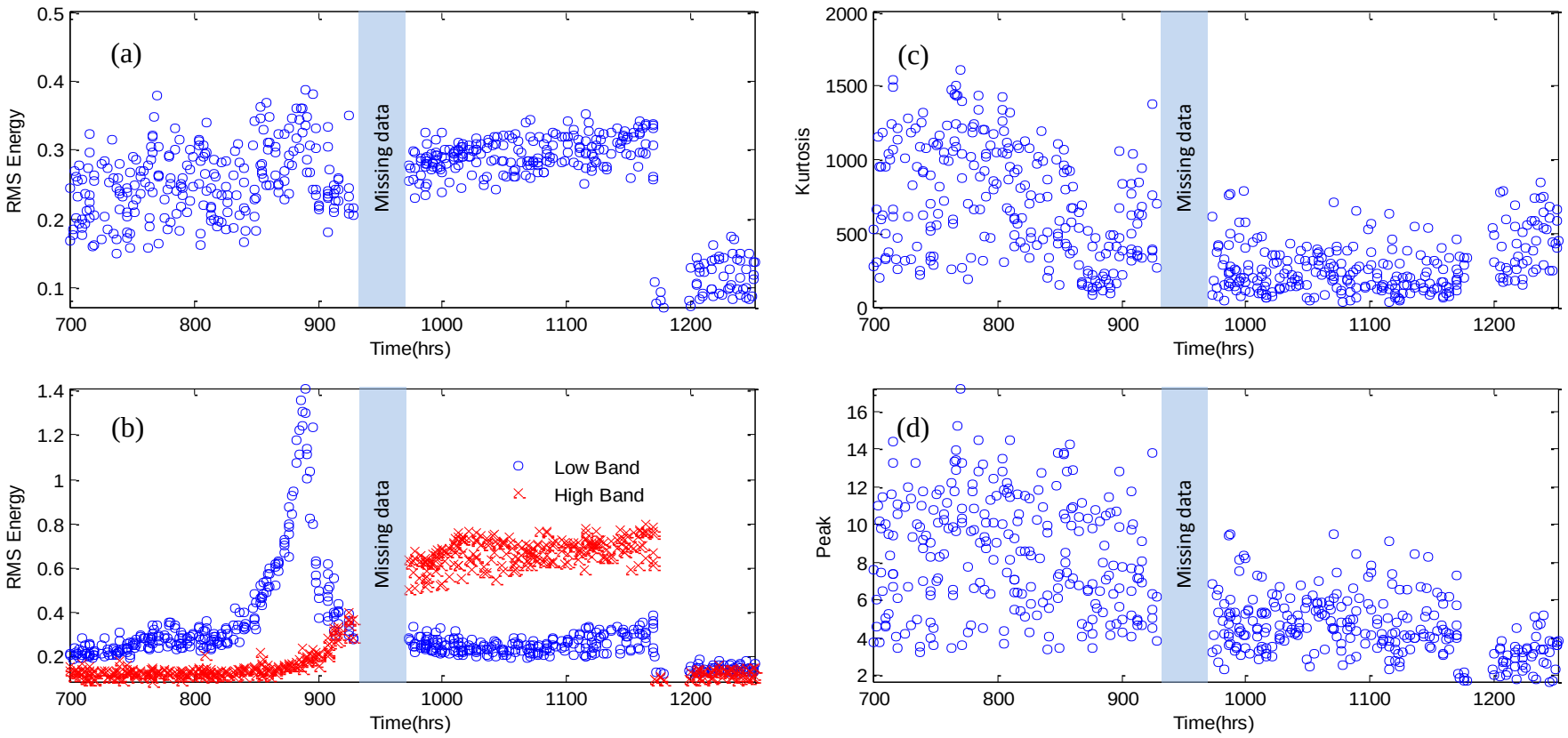


Figure 22 Extracted Features from TRGB 3.

### 6.4.2 High Band Bearing Energy

For generalized damage, the RMS energy of the frequency magnitudes in the high frequency band of 10 kHz to 24 kHz is adopted. It should be noted that this high frequency band is often the demodulation band used in envelope analysis to detect incipient defects. It is used here however as a measure of generalised damage in the bearing. For the high frequency band bearing energy shown in Figure 20(b) to Figure 22(b), it begins to rise when the low band energy reaches its peak. As the low band energy falls back to normal levels, the high band energy rises with increased scatter. The increased scatter is likely attributed to the increasing random impact from generalized bearing damage. The high band energy does not rise exponentially but rather saturates to form an 'S' shaped profile. From the high band energy trend, it can also be seen that the generalised damage can progress continuously as seen in TRGB 1 and 2 or in discrete stages can be observed for the progression in TRGB 3. It should be noted that the TRGBs are in service for considerable time after the low band energy have peaked and for the 'S' shape profile from the high band feature to be formed. The high band energy feature rises monotonically and is consistent between the three TRGBs. However, it only rises after the bearing damage has become widespread and is not able to detect the localised damage earlier. This in turn results in shorter detection lead-time for maintenance actions to be planned.

### 6.4.3 Feature Selection

In comparison, the RMS energy, kurtosis and peak amplitude of the entire FFT spectra does not perform as well compared to the high and low energy features. The RMS energy shown in Figure 20(a) to Figure 22(a) does not rise monotonically but rises to a peak before dipping and rising again. This pattern is attributed to the sum of the behavior of both low and high band bearing energy as described above. As such, the separate use of the low and high band energy features can better describe the bearing degraded state. Kurtosis shown in Figure 20(c) to Figure 22(c) exhibits a linearly decreasing when the low band energy is rising exponentially and flattens out when the high band energy rises. This linear trend is simpler compared to the low band bearing energy and could be more ideal to monitor localised damage but the magnitude and

scatter of the kurtosis value vary widely between the TRGBs which makes it a less consistent feature. The peak amplitude shown in Figure 20(d) to Figure 22(d) has the worst performance as it does not show any discernible trends. From the different features, the low and high band bearing energy features performs well for monitoring localised to generalised bearing damage. However, each feature by itself is unable to distinguish both localised and generalised damage and they have to be used together to monitor both damage states. Furthermore, their sequential behavior presents opportunities for effective diagnostic and prognostic model to be developed.

## 6.5 State Detection and Threshold Setting

For diagnosis of the bearing health state, thresholds to distinguish between normal, localised and generalised damage have to be established from the selected features. For signal detection theory described above to be applied, the features have to be labeled according to the relevant health states for the respective pdf to be obtained. However, this is not available from the field defect data as the actual transition time between normal, localised and generalised damage is not known. Therefore, unsupervised classification is applied to determine the cluster of features corresponding to the three bearing states. Figure 23 shows the plot of the Low and High Band Feature from the three defective TRGBs. It can be seen that both features are highly uncorrelated with each other and thus further processing using techniques like PCA to improve separability is not required. Inspection of the plot also shows that the amount of the feature instances and the densities from the different states varies; with normal condition features being the most abundant, followed by generalised damage and then localised damage being the most sparse as it's state is transitory. In this section, both *k*-means and GMM classifications are employed and compared for their performance.

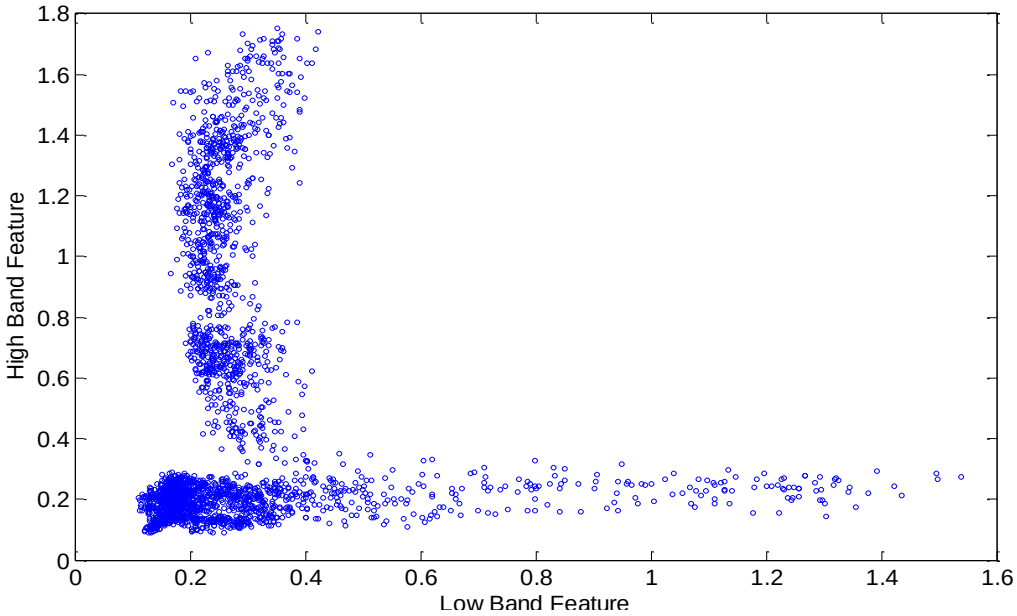


Figure 23 Plot of Low and against High Band Feature

### 6.5.1 *k*-means Description

As mentioned earlier, *k*-means is a popular distance-based unsupervised classification technique and is described in detail in [79]. *k*-means groups a set of *N* dimensional feature set into *k* clusters with each cluster parameterized by a vector mean, $\vec{\mu}^{(k)}$. Using Euclidean distance, the distance between any two instances, $\vec{x}$ and $\vec{y}$ in the feature set is defined by:

$$d(\vec{x}, \vec{y}) = \|\vec{x} - \vec{y}\| \tag{19}$$

The *k*-means algorithm begins by arbitrary assigning an initial mean vector, $\vec{\mu}^{(k)}$. An *assignment* step is then performed where each feature instance, $\vec{x}_i$, is assigned to a class *k* with their nearest mean such that,

$$k^i = argmin\{d(\vec{x}_i, \vec{\mu}^{(k)})\} \tag{20}$$

An *update* step is then carried out where the vector mean is adjusted to match the sample means of the instances that they are classified in. The *assignment* and *update*

steps are then iterated until there is no change to the assignment of the instances to the cluster means.

### 6.5.2 *k*-means Classification of high and low band feature

The *k*-means algorithm is applied to the low and high band features *(N = 2)* shown in Figure 23 with the number of clusters, $k = 3$ for the desired states; normal, localised and generalised damage. The result of the classification with the clusters and their means are shown in Figure 26. It can be seen that *k*-means can cluster the features into regions where the three bearing health states are expected. However, the state detection thresholds established from the classification boundaries are considerably high. For generalised damage especially, there is a significant portion of instances that are poorly classified as normal. The high detection thresholds will provide a good true positive rates but it causes a high false negative rates as well. In addition, the high detection thresholds will result in reduced remaining useful life prediction due to lower threshold to the defined failure limit.

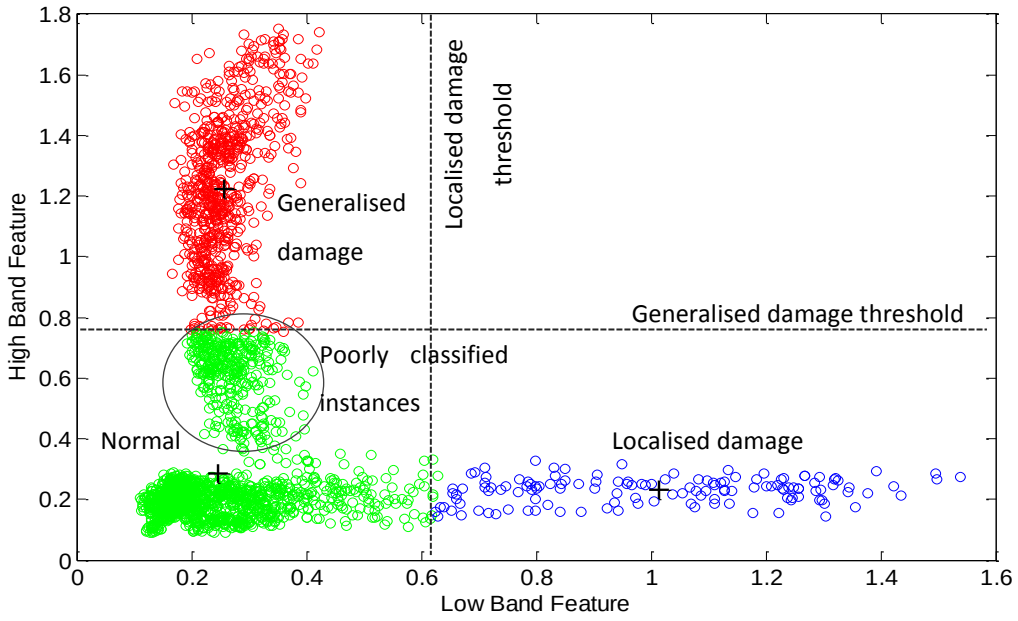


Figure 24 *k*-means classification of high and low band TRGB features

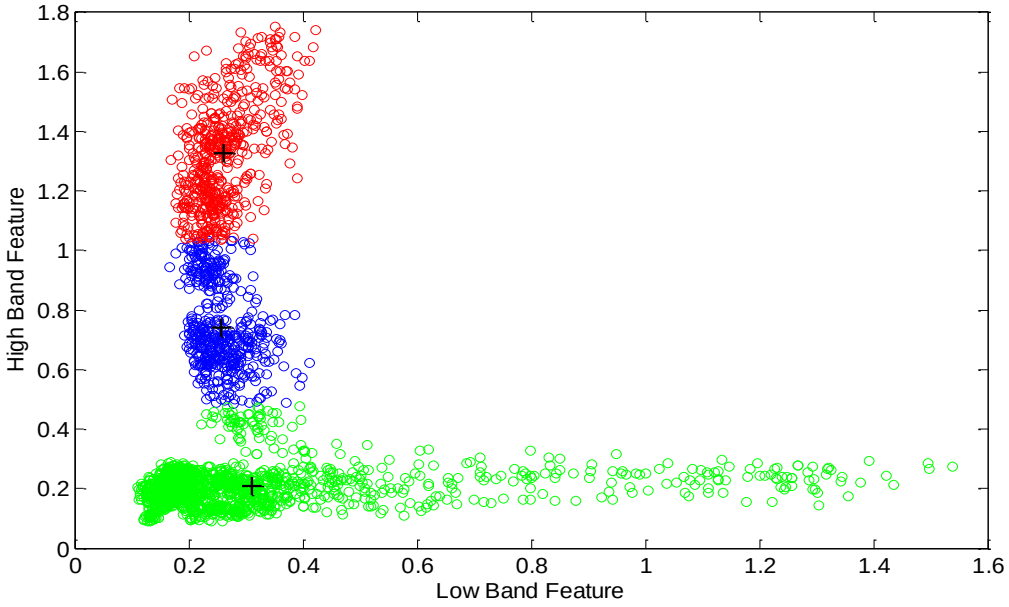


Figure 25 *k*-means misclassification due to poor initial condition

*k*-means performed poorly to this feature set as the clusters are elongated and varies in their densities. The *k*-means algorithm only considers the distance between the means and the feature instances and not the density or shape of the clusters. As such, instances that actually belong to the Generalised damage cluster are incorrectly assigned to the Normal condition cluster. *k*-means is also sensitive to the initial means and poor assignment can behave cause it to perform badly as shown in Figure 25. In this case, the entire localised damage instances are misclassified. Another disadvantage of *k*-means is that it is a hard classifier, i.e. the feature instances are classified into discrete clusters. Intuitively, instances near the classification boundaries should have some uncertainty of belonging in other clusters. As such, a direct application of *k*-means does not provide a probabilistic interpretation of which class the instance belongs too. Towards this end, a variant such as soft *k*-means may be applied. Soft *k*-means is an extension of the standard *k*-means where an additional stiffness parameter is added so that each instance now has a weightage of belonging to different clusters based on their distances to the cluster means [79]. However, the inability of *k*-means to represent the elongated clusters in this feature set still remain thus other methods are adopted.

### 6.5.3 Gaussian Mixture Model Description

Amongst the different unsupervised classification algorithms, GMM is adopted here as it is less susceptible to misclassification compared to *k*-means when the clusters sizes and densities vary widely [117]. As mentioned, GMM is a statistical model that uses a weighted sum of pdfs of multiple Gaussian distributions to describe the states in the feature space. A description of the methodology from [76] is adopted here. The GMM works by creating a model of each state which is written as:

$$\lambda = (w_i, \mu_i, \Sigma_i) \; for \; i = 1 \dots k \tag{21}$$

where $\lambda$ is the model, $w$ represent the weights assigned to the Gaussian means, $\vec{\mu}$ is the vectot mean, $\Sigma$ is the covariance matrix and $k$ is the number of states to be modeled. For a set of $N$ dimensional feature vector, $\vec{x}$ to be modeled with GMM, $P_i(\vec{x}), i = 1 \dots k$, are the pdfs of $\vec{x}$ generated from the *ith* component of GMM which is denoted by $\lambda_i$ and is given by:

$$P(\vec{x}) = P(\vec{x}|\lambda_i) = \frac{1}{\sqrt{(2\pi)^k |\Sigma_i|}} \exp\left\{-\frac{1}{2}\left[(\vec{x} - \overrightarrow{\mu_i})' {\Sigma_i}^{-1} (\vec{x} - \overrightarrow{\mu_i})\right]\right\} \tag{22}$$

A weighted sum of pdfs of all the $k$ components is then used to compute the probability that $\vec{x}$ belonged to model $\lambda$, where, $\sum_{i=1}^{k} w_i = 1$.

$$P(\vec{x}|\lambda) = \sum_{i=1}^{k} w_i P_i(\vec{x}) \tag{23}$$

For a given dataset, $X = \left(\overrightarrow{x_1}, \overrightarrow{x_2} \dots \overrightarrow{x_J}\right)$, the log-likelihood function of the model given the dataset is given by:

$$log\left(L(\vec{x}|\lambda)\right) = log \prod_{j=1}^{J} p\left(\overrightarrow{x_j}|\lambda\right) = \sum_{j=1}^{J} log\left(\sum_{i=1}^{k} w_i P_i\left(\overrightarrow{x_j}\right)\right) \tag{24}$$

The estimation of the model parameters ($\vec{\mu}$ and $\Sigma$) of the mixture of Gaussian distributions is carried out using the Expectation-Maximization (EM) algorithm. The EM algorithm is often used to obtain maximum likelihood estimation from incomplete data for various models such as HMM and GMM applied here. The incomplete data in this case is the unavailability of the label to determine which Gaussian mixture each data instance belongs to. The EM algorithm consists of an Expectation step and a Maximization step which are used iteratively to find the maximum likelihood of the parameters given the dataset using an initial estimate of the model parameters. *k*-means has been a popular method used to estimate the initial parameters before EM algorithm is applied. A detailed description of the EM algorithm can be found in [73; 118]. A drawback of GMM classification is that it can be sensitive to its initial parameters like *k*-means and there is no guarantee that it will converge to a global optimum and thus mis-classification due to local optimum can occur.

### 6.5.4 GMM Classification of High and Low Band Features

GMM classification is applied to the low and high band features *(N = 2)* with $k = 3$ for the desired states; normal, localised and generalised damage. The result of the GMM classification is shown in Figure 26 and the estimated parameters are shown in Table 6. Figure 27 demonstrates a case in which the GMM converged towards a local optimum for the TRGB feature set and it can be seen that the features are badly classified. As such, it is essential for the classification to be evaluated manually instead of an automated routine. From Figure 26, it can be seen that the separation of generalised damage state is good but less so between normal and localised damage. The small off - diagonal values of the covariance matrices again shows that both features are highly uncorrelated with each other. Detection thresholds for both localised and generalised damage can then be determined from the clustered datasets using Signal Detection Theory. As the two features are highly uncorrelated, it is simpler for the thresholds to be set independent of each other. The pdfs for both features for normal and damaged conditions are shown in Figure 28 with the associated ROC curves plotted in Figure 29.

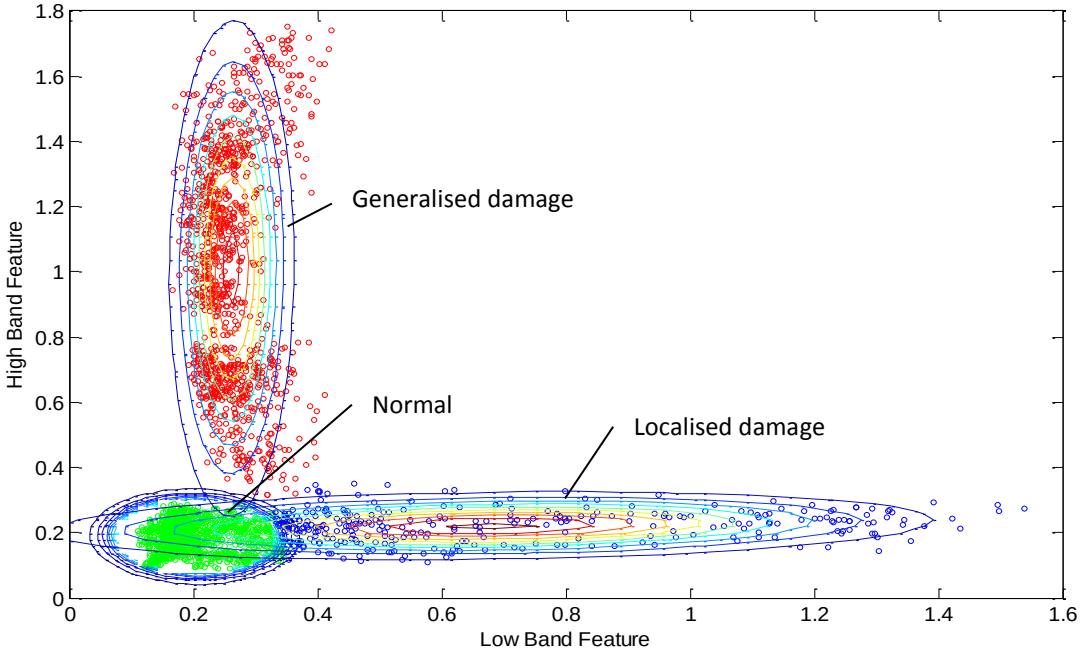


Figure 26 GMM classification of TRGB high and low band features

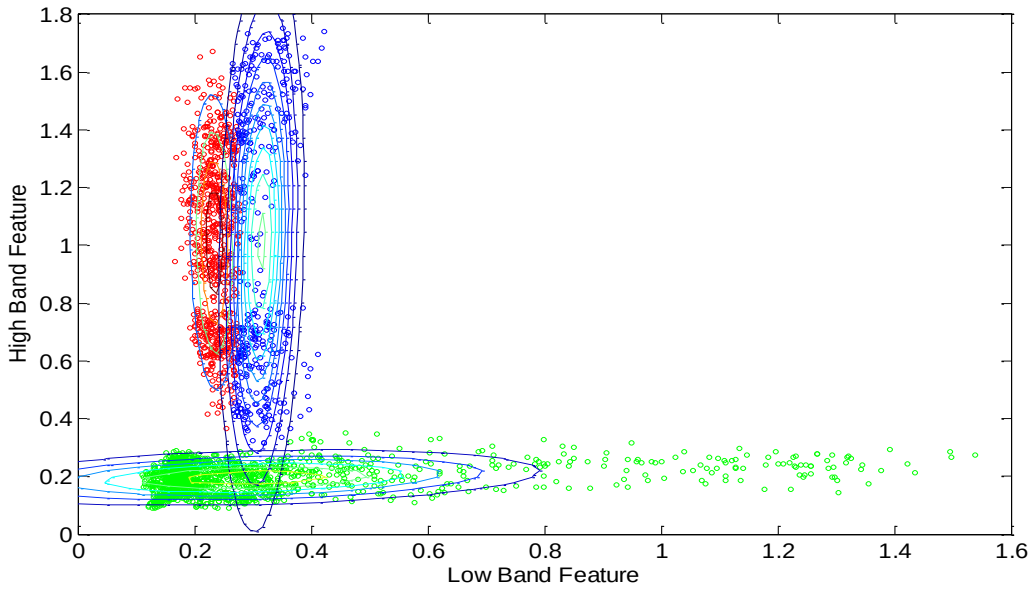


Figure 27 GMM Misclassification due to local optimum

Table 6 GMM Parameter Estimation

| $k$ | 1 (Normal) | 2 (Localised damage) | 3 (Generalised damage) |
|---|---|---|---|
| $w_k$ | 0.456 | 0.132 | 0.412 |
| $\overrightarrow{\mu_k}$ | [0.203 0.188]' | [0.648 0.221]' | [0.262 0.992]' |
| $\Sigma_k$ | $\begin{bmatrix} 0.0029 & \sim 0 \\ \sim 0 & 0.0021 \end{bmatrix}$ | $\begin{bmatrix} 0.118 & 0.004 \\ 0.004 & 0.0024 \end{bmatrix}$ | $\begin{bmatrix} 0.0023 & 0.0001 \\ 0.0001 & 0.1416 \end{bmatrix}$ |

### 6.5.5 Fault Detection Threshold

For the low band feature, a lognormal distribution provides a better fit to the feature histogram for both normal and damaged conditions. For the high band feature, normal distributions were used. For both the low and high band features, the threshold is set such that the probability of false alarm from a serviceable TRGB is below the 5% allowable limit as required in AC29-MG15 [9] and ADS-79B [10]. As seen in Figure 28, the damage detection thresholds are set at 0.32 g(rms) and 0.27 g(rms) for the low and high band features respectively. Based on these thresholds, the false alarm and true detection rate for the low band feature is 2.1% and 92.3%. For the high band feature, the false alarm and true detection rates are ~0% and 97.6%. From the ROC curve, it can again be seen that the high band feature has higher detection performance compared to the low band feature due to better class separation.

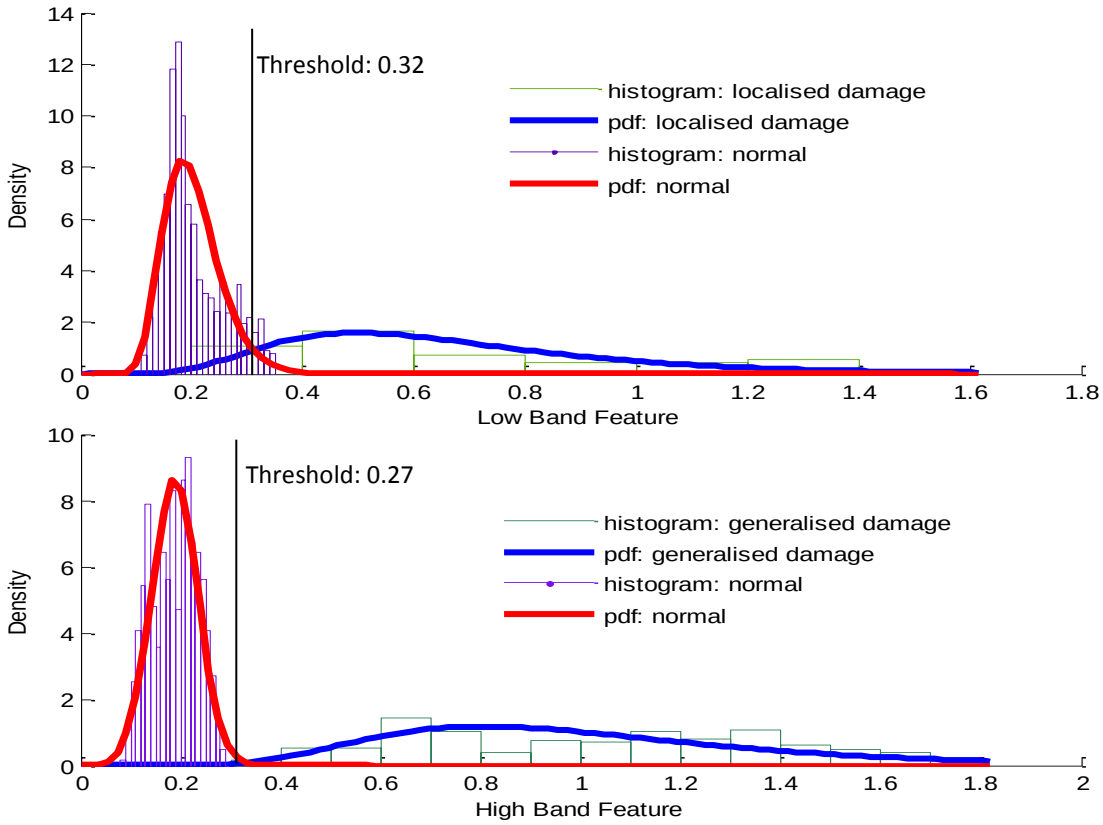


Figure 28 Probability distribution functions: (Top) Low Band Feature, (Bottom) High Band Feature

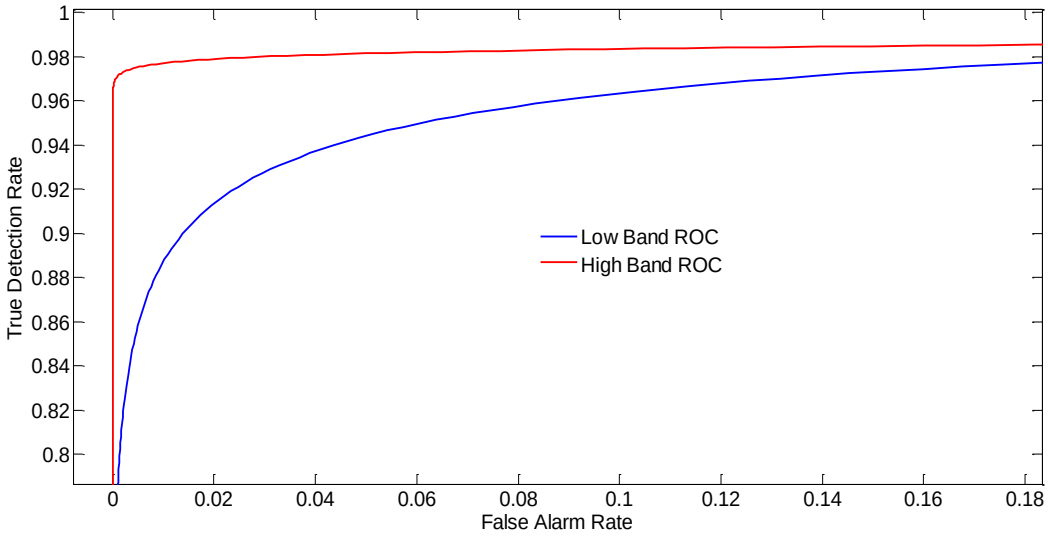


Figure 29 ROC Curve for both Low & High Band Features

## 6.6 Trend Extrapolation Prognostic Model

From the feature trend plots and detection thresholds, prognostic models can be developed to predict the time to failure of the output shaft bearing. Failure is defined to be the degraded state at which bearing replacement is desired. In this application, it is before the onset of widespread damage in the bearing that leads to grease leak. As both localised and generalized damage in the bearing can be monitored, different prognostic models can be developed depending on the damage states of interest. Figure 30 depicts the features at the different stages of bearing damage.

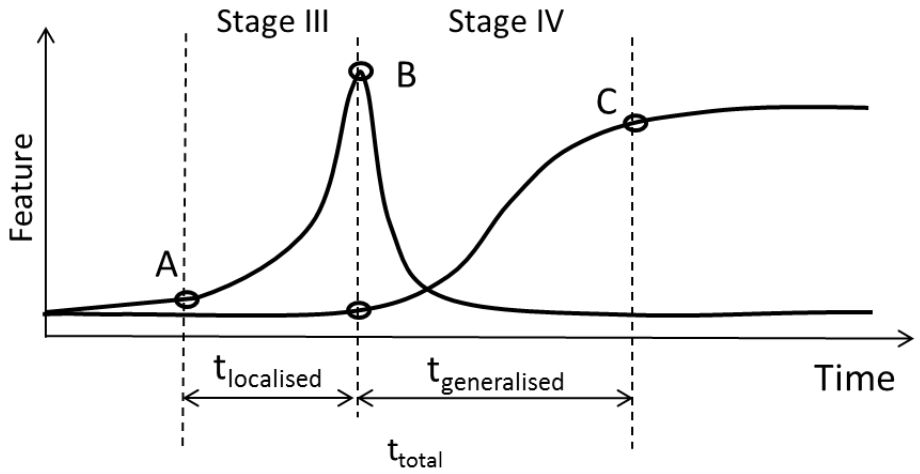


Figure 30 Features at different bearing damage states

If the failure condition used is localised damage, the prognostic model can be developed using low band energy feature with exponential damage growth models. Bearing replacement is then desired before the low band energy reaches an established threshold. Similarly, a prognostic model may be developed using the high band feature if the failure condition used is generalised damage. A model using both set of features can also be developed to maximize the period of notification prior to bearing replacement as seen in Figure 30. As the feature plots shows a similar trends across the TRGBs, a prognostic model using simple trend extrapolation is adopted here and this can be easily implemented without the use of complex algorithms. In most prognostic models, especially those developed from seeded faults or accelerated tests, damage is initiated during the start of the test and damage propagation begins immediately. In

practice, components are not seeded with defects and the time when damage initiates can vary widely and dependent on many operating and environmental factors. From Figure 20(b) to Figure 22(b), the variation in damage initiation can be observed as the time in which the low band feature begins to rise differs between TRGBs. Therefore, the time since new is not evaluated and the time from detectable damage to the defined damage state is adopted instead.

### 6.6.1 Localised Damage Model

For the localised damage prognostic model, only the monotonic rising portion of the low band energy data (Stage I and II of Figure 30) is considered. The exponential function in Eqn. (19) is used to fit the features and estimate the parameters. This is used to model the feature degradation path which is stationary for a serviceable gearbox followed by exponential increase as damage initiates and progresses. As the trend after damage detection is of interest, the curves are aligned at the time when the detectable damage threshold is expected to be crossed. Figure 20 shows the low band energy exponential fit for the three TRGBs.

$$f(t) = ae^{bt} + c \tag{19}$$

It can be seen from Figure 31 that the three plots correlate very closely which shows that the rate of progression for localised damage is consistent. A regression curve is fitted to the combined dataset to estimate the overall degradation path in Figure 32. The peak values for the low band energy are similar between the three gearboxes at 1.54 g(rms), 1.36 g(rms) and 1.44 g(rms) respectively. As such, the failure threshold for localised damage is conservatively set lower at 1.2 g(rms). From the regression fit in Figure 21, the time from detectable damage of 0.32 g(rms) to the defined localised damage threshold of 1.2 g(rms) is determined to be 61.1hrs with corresponding 90% confidence bounds of 58.1.hrs and 63.8hrs. The 90% confidence bounds was included to provide a probabilistic measure of the time-to-failure and is used to aid maintenance decision making.

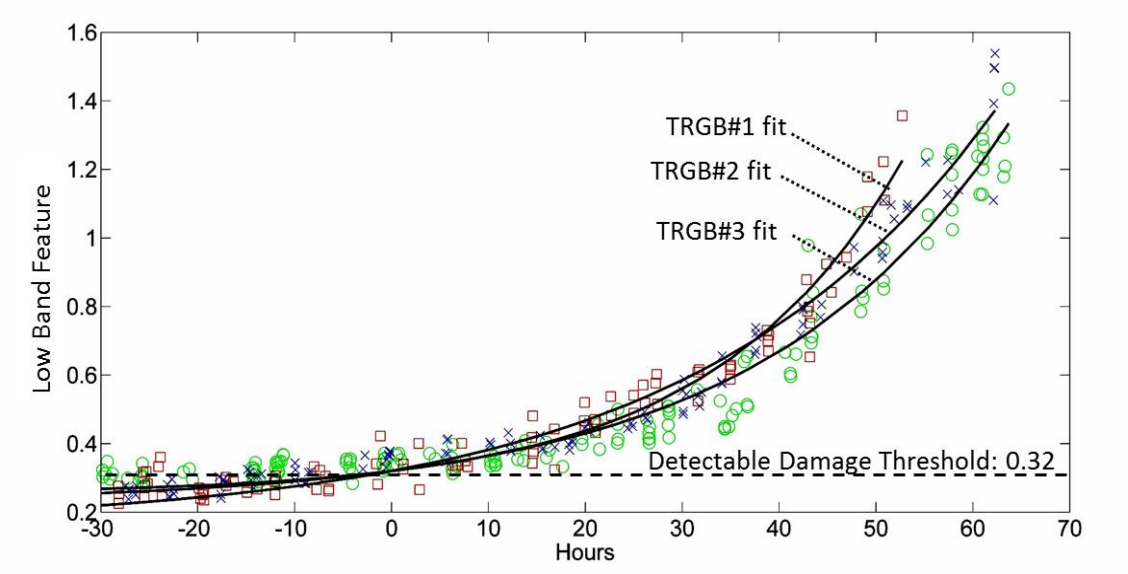


Figure 31 Exponential fit of the Low Band Energy feature: (□) TRGB#1, (x) TRGB#2, (o) TRGB#3

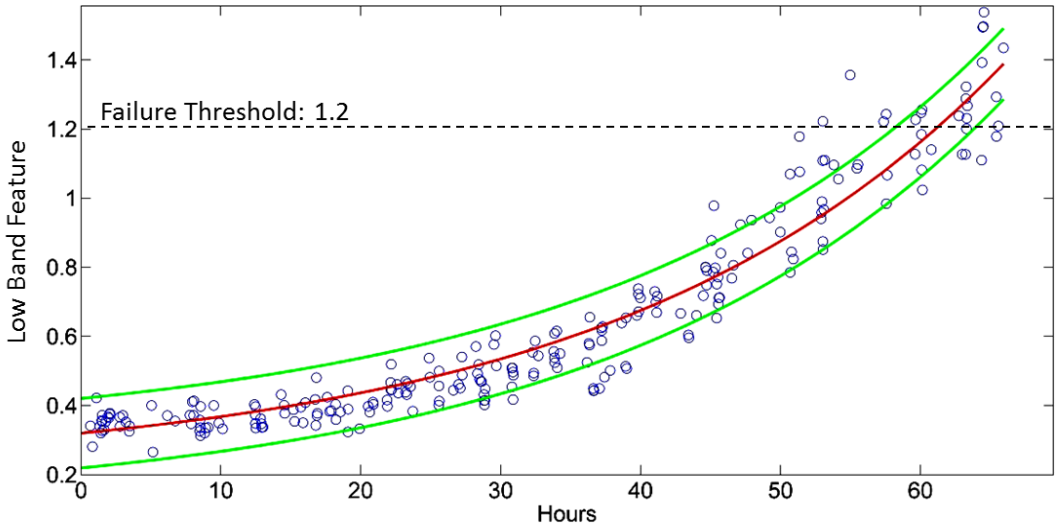


Figure 32 Exponential fit of the Low Band Energy feature datasets with 90% confidence bounds

### 6.6.2 Generalised Damage Model

In most condition based maintenance applications, the use of the low band vibration energy feature for prognostic of localised damage as described above would suffice. However, the behavior that the low band vibration energy fault pattern follows the low

band vibration energy is exploited here to further increase the lead time before bearing replacement. In Figure 20(b) to Figure 22(b), the high band feature data displays a 'S' shaped profile also known as a logistic function with increased scatter. As such, a 5-parameter logistic (5PL) regression model as shown in Eqn (20) is used for the fitting the high band feature data. The 5PL model is used as it can flexibly fit asymmetric trends in the data compared to standard logistic regression as described in [119]. However, it is noted that the 5PL curve can be difficult to fit as the initial estimate of each parameter has to be selected carefully. In this study, the initial estimates are selected through trial and error and adjusting the parameters based on their properties shown in Table 7.

$$f(t) = a + \frac{b}{\left(1 + \left(\frac{t}{c}\right)^{d}\right)^{e}} \tag{20}$$

Table 7 Properties of the 5PL parameters

| Parameter | Properties of curve |
|---|---|
| A | Lower asymptote |
| B | Upper asymptote |
| C | Affects the position of inflection point |
| D | Rate of change between asymptotes |
| E | Asymmetry factor |

The fitted curves for the three gearboxes are shown in Figure 33 and the fitted parameters. The curves are aligned at the time when the low band energy reaches their peak values. Unlike the low band feature curves, the high band curves do not have similar upper limits. The upper asymptotes for the three TRGBs from the fitted curves are 1.13 g(rms), 0.65 g(rms) and 1.17 g(rms) respectively. The logistic regression is performed on the combined high band feature dataset as shown in Figure 34. Due to the increasing variance (or heteroscedasticity) in the data, a robust regression is employed. The advantage of using robust regression here is that it is less susceptible to outliers skewing the regression fit. Outliers with high residuals will have lower weights thus reducing their effect. For a conservative estimate of the generalised damage failure time, a low failure threshold of 0.5 g(rms); which is below the upper asymptote of

TRGB#2 is set. Based on this threshold, the time to generalised damage failure after localised damage has formed is 72.4 hrs corresponding 90% confidence band of 14.3hrs and 133.6hrs. Compared to the localised damage model, the generalised damage model is less accurate as a prognostic tool as its confidence bounds is much wider. For high confidence levels above 90%, the generalised damage model would be ineffective as the confidence bounds would be too wide for any predictions on failure time to be made.

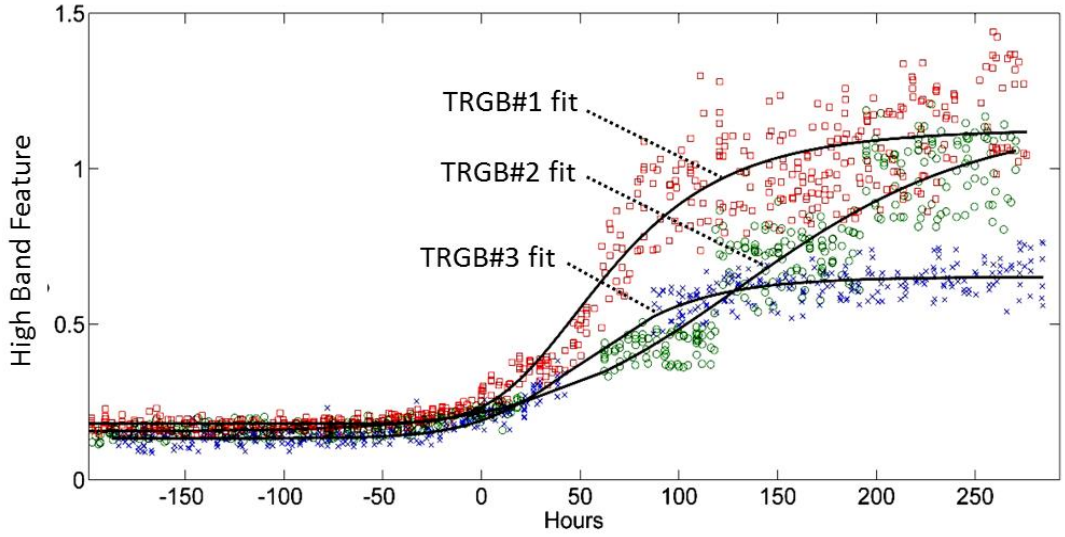


Figure 33 5PL fit of the High Band feature: (□) TRGB#1, (x) TRGB#2, (o) TRGB#3

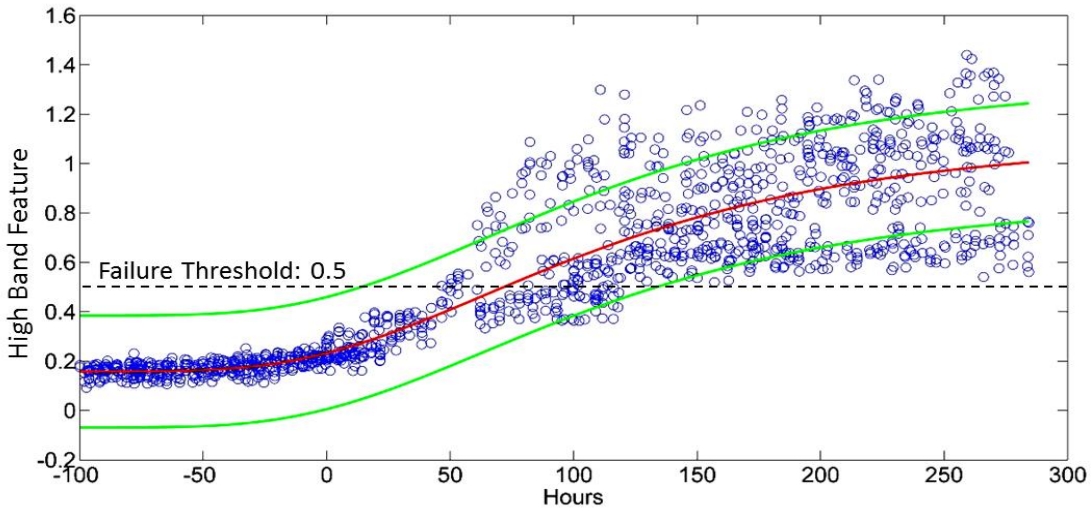


Figure 34 5PL fit of the High Band feature datasets with 90% confidence bounds

## 6.7 Implementation in maintenance

From analysis of the feature data, damage models for predicting localised and generalised damage were developed. Assuming a normal distribution, the localised damage failure time, $t_{localised}$ is $\sim N(61.6, 2.1^2)$ and the generalised damage failure time, $t_{generalised}$ is $\sim N(72.4, 47.7^2)$. The sum of both failure times can therefore be evaluated to be $t_{total} \sim N(134, 47.8^2)$. Using both the localised and generalised damage model, the total time from detection of localised damage to replacement of bearing from generalised damage is 134hrs with 90% confidence bounds of 72.7 hrs and 134.3 hrs. Using the lower confidence bound for conservative estimate, the use of both damage models allows a 25% improvement in the detection lead time compared to the use of the localised damage alone. From these findings, a CBM program for the TRGB output shaft bearing can be recommended. The bearing can be monitored for localised damage using the detectable damage threshold established. In event that this threshold is exceeded, the replacement of the TRGB can be planned for in the next 72.7 hrs. The high band feature can also serve as a diagnostic tool as the rise in magnitude and increased scatter of the feature are clear indication of generalised damage within the bearing. This redundancy can be useful for this TRGB application as the low band feature does not rise continuously and will drop back to normal levels after it peaks. If the HUMS data during the period in which the low band feature is unavailable, the high band feature can still provide indication of bearing damage at a later time.

## 6.8 Chapter Summary

In this chapter, the end-to-end process for development of a CBM task, including both diagnostics and prognostics application is shown. The operational HUMS data from three TRGBs found with damaged bearings were analyzed and correlated with their tear-down inspection findings. From analysis of their vibration spectrum, it was shown that there were fault patterns that distinguish the TRGBs with damaged bearings from serviceable ones clearly. Notably, this study on field data ratified experimental observations on fault patterns in vibration signals for bearing damage progression. It was shown that low-band vibration energy levels does not rise monotonously with bearing damage progression and can fall back to normal levels. Furthermore, high band

vibration energy rises following the fall in low band vibration energy. From these fault patterns in the vibration spectrum, the progression from initial damage to localised damage and subsequently generalised damage in the bearing can be inferred. Two features were then developed using the vibration energy from selected frequency bands in the vibration spectrum to monitor localised and generalised bearing damage. From the feature data, prognostic models based on trend extrapolation was developed. Using both damage models, the lead time between damage detection and bearing replacement is improved compared to use of the localised damage model alone. This study shows that bearing diagnostics and prognostic can be effectively carried out in the field environment using simple methods of FFT spectral analysis and trend extrapolation.

Admittedly, there are areas in the study that can ideally be further improved. To begin with, the lack of time domain data restricts the use of signal processing methods. As such, techniques such as envelope analysis which may better detect incipient fault and allow a longer detection lead-time cannot be used. Secondly, the TRGB is a relatively simple gearbox thus the results here cannot be extended to more complex designs such as epicyclic arrangements used in main gearboxes. Lastly, the prognostics using trend extrapolation assumes that the operating environment remains the same and the RUL prediction is not adaptive to changes. It also requires significant review and manipulation of the feature data for trends to be identified and then RUL prediction. In practice, an automated process to perform this is desired to reduce the need for trained HUMS analyst.

## 6.9 Related Publications

- *HUMS experience on the RSAF AH64D Tail Rotor Gearbox,* (Presented 19 Feb 2013), AHS Airworthiness, CBM and HUMS Specialists' Meeting.
- Reuben Lim, David Mba, (Published Oct 2013) *Diagnosis and prognosis of AH64D Tail Rotor Gearbox Bearing Degradation*, Proceedings of the ASME 2013,IDETC/CIE 2013.
- Reuben Lim, David Mba, *Bearing Time-to-Failure Estimation using Spectral Analysis Features*, (Accepted for publication), Structural Health Monitoring.

# 7 DIAGNOSTICS AND PROGNOSTICS USING SWITCHING KALMAN FILTERS

In this chapter, ways to improve the ease of use of prognostic applications by maintainers is explored. As mentioned several times in this research, the lack of failure cases for training data is a key constraint thus the use of AI and statistical methods such as ANN and HMM are not further investigated. Conversely, hybrid approaches using both model and statistical based methods is adopted as they require less training data. In particular, the use of recursive Bayesian estimation techniques such as Kalman Filtering is of interest as it is adaptive to change and can handle uncertainty in the degradation trend. In this chapter, the use of switching Kalman Filters (SKF) is investigated for both fault detection and RUL prediction of rolling element bearing. This approach and its benefits they are shown using both simulated and feature data extracted from the TRGB bearing.

## 7.1 Overview of Kalman Filtering

The Kalman Filter (KF) is a stochastic filtering process that recursively estimates the state of a linear dynamic system in the presence of Gaussian measurement and process noise by minimizing the mean squared error [85]. For non-linear systems, the Extended Kalman Filter (EKF), which is a linear approximation of the non-linear function is most commonly adopted [120; 121]. When the system is highly non-linear where linear approximation is inadequate, the Unscented Kalman Filter (UKF) can be adopted [120]. Kalman Filtering have been used in a wide range of applications from navigation and tracking to economic forecasting. In maintenance application, it has been applied to engine health diagnostic [122], rolling element bearings [123] and in recent years to electronics prognostic for estimating remaining useful life of ball grid array connections [101; 102; 124] and electrolytic capacitors [103]. The key advantage of using Kalman Filtering is that it accounts for measurement and system noise in the CM data and the system state and parameters of the degradation model can be adaptively estimated as they evolve through time [125]. It can also be used for prognostics by forecasting the system state into the future using the degradation model and the latest available measurement. The algorithms for both KF and EKF are described here.

### 7.1.1 Kalman Filter

The KF consists of a linear discrete state-space model describing the evolution of a process given by:

$$\begin{aligned} x_t &= A_{t-1}x_{t-1} + q_{t-1} \\ y_t &= H_t x_t + r_t \end{aligned} \tag{21}$$

where $x_t$ is the true but hidden state of the system and $y_k$ is the observable measurement of the state. The KF assumes a linear system dynamics and all noise follows a Gaussian distribution. $A$ is the fundamental matrix describing the system dynamics and $H$ is the measurement matrix. $q_{t-1} \sim N(0, Q_t)$ is the process noise and $r_{t-1} \sim N(0, R_t)$ is the measurement noise. The KF estimates the value of $x_t$, given the measurement, $y_t$ by filtering out the noise. This is carried out using the 'Prediction' and 'Update' steps also known as the Ricatti Equations [121] as follows.

*Prediction Step:*

$$\begin{aligned} &\text{Predicted state estimate:} && \hat{x}_t = A_t x_{t-1} \\ &\text{Predicted estimate covariance} && \hat{P}_t = A_t P_{t-1} {A_t}^T + Q_t \end{aligned} \tag{22}$$

*Update Step:*

$$\begin{aligned} &\text{Measurement residual:} && v_t = y_t - H_t \hat{x}_{t|t-1} \\ &\text{Residual covariance} && C_t = H_t \hat{P}_t H_t^T + R_t \\ &\text{Kalman Gain} && K_t = \hat{P}_t H_t^T {C_t}^{-1} \\ &\text{Updated state estimate} && x_t = \hat{x}_t + K_t v_t \\ &\text{Updated estimate covariance} && P_t = (I - K_t H_t)\hat{P}_t \end{aligned} \tag{23}$$

The Kalman filter can also perform prediction repeating the prediction step in Eqn. (22) using the last known state and covariance estimate without updating the state and covariance estimate.

### 7.1.2 Extended Kalman Filter

As mentioned, the EKF is a non-linear extension of the KF which uses linear approximation of the non-linear function to estimate the state mean and covariance

[120; 121]. The linear approximation performed through first and second-order taylor series expansion of the non-linear function is most commonly used and the first-order is adopted here. The discrete state-space model describing a non-linear process is given by:

$$\begin{aligned} x_t &= f(x_{t-1}) + q_{t-1} \\ y_t &= h(x_t) + r_t \end{aligned} \tag{24}$$

where $x_t$ is the true but hidden state of the system and $y_k$ is the observable measurement of the state. *f(.)* is the fundamental matrix describing the system dynamics and *h(.)* is the measurement matrix and both are functions assumed to be continuously differentiable. The 'Prediction' and 'Update' steps for the EKF are similar to the KF except that Jacobian matrixes of the non-linear functions are used instead as shown.

*Prediction Step:*

$$\begin{aligned} &\text{Predicted state estimate:} && \hat{x}_t = f(x_{t-1}, t-1) \\ &\text{Predicted estimate covariance} && \hat{P}_t = F(x_{t-1}, t-1)P_{t-1}F'(x_{t-1}, t-1) + Q_{t-1} \end{aligned} \tag{25}$$

*Update Step:*

$$\begin{aligned} &\text{Measurement residual:} && v_t = y_t - h(\hat{x}_{t-1}, t) \\ &\text{Residual covariance} && C_t = H(\hat{x}_t, t)\hat{P}_t H'(\hat{x}_t, t) + R_t \\ &\text{Kalman Gain} && K_t = \hat{P}_t H'(\hat{x}_t, t){C_t}^{-1} \\ &\text{Updated state estimate} && x_t = \hat{x}_t + K_t v_t \\ &\text{Updated estimate covariance} && P_t = (I - K_t H(\hat{x}_t, t))\hat{P}_t \end{aligned} \tag{26}$$

where *F(.)* and *H(.)* are the Jacobians of *f*(.) and *h(.)* are given by

$$\begin{aligned} F(x_{t-1}, t-1) &= \left.\frac{\partial f(x_{t-1}, t-1)}{\partial x}\right|_{\hat{x}_{t-1|t-1}} \\ H(\hat{x}_t, t) &= \left.\frac{\partial h(x_t, t)}{\partial x}\right|_{\hat{x}_{t|t-1}} \end{aligned} \tag{27}$$

### 7.1.3 Limitations of KF and EKF in prognostics

A limitation of both KF and EKF is that the system's degradation process has to be time-invariant else the model can be unstable and it's estimations divergent. However, the degradation process in components can be uncertain and evolve over time as seen in bearing wear tests [115]. For example, in Figure 35, the vibration measurement of a serviceable bearing can be stationary with measurement noise. When slow stable wear from damage such as surface pitting occurs, the vibration level gradually rises linearly. When the accumulated damage is severe, the vibration level then rises exponentially. This behavior can also be seen from our TRGB example in Figure 20, the Low Band Energy Feature exhibits stationary then linear trend before the exponential trend is apparent subsequently. The use of a single degradation model would result in inaccurate state estimates and cause RUL predictions to diverge or fluctuate depending on whether the underlying degradation process is under or over-fitted by the model. In lieu of this, model-based prognostic works often only use portion of the feature data that fits the assumed degradation model [80; 90; 102] and this data censoring has to be performed manually by a trained analyst. Furthermore, it is difficult to determine when the assumed model can be suitably applied in an automated application where the complete feature history prior to failure is unknown. Thresholds can be set adequately high to only identify unstable wear trends for the model to be fitted correctly such as in [80; 90] but this would reduce sensitivity for damage detection. As such, the use of KF or EKF for automated analysis can be challenging.

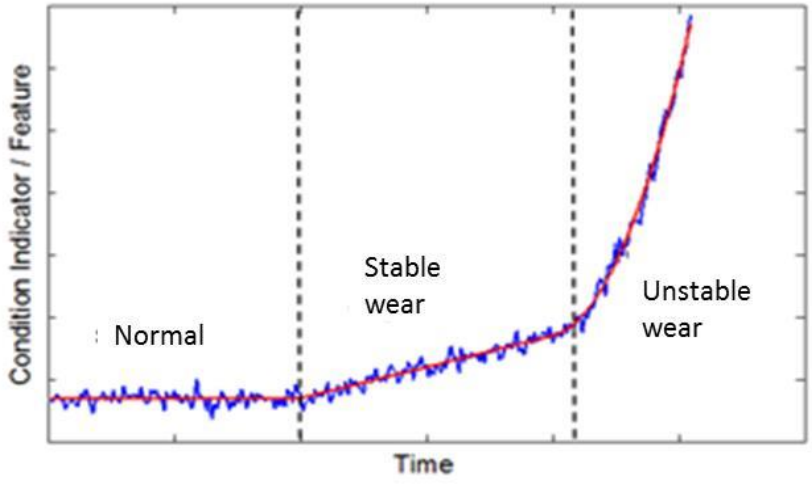


Figure 35 Evolution of degradation process across time

## 7.2 Switching Kalman Filter

For the different degradation process to be modeled, the Switching Kalman Filter (SKF) is adopted here. The SKF, also known as Linear Dynamic models [126] can track system with changes in their dynamical process. It is popularly used to track multiple moving targets but has also been applied in meteorology [127] and econometric [128]. In maintenance applications, it has been used to diagnostic sensors and actuator failures [129] and changes in non-linear stochastic control systems [130]. The switching Kalman filter can be represented as a dynamic Bayesian network as shown in Figure 36. In each time step, both the model switch variable, $S_t$ and state variable, $x_t$ are hidden and have to be inferred from the observations, $y_t$. For a system with multiple dynamics which are described with $n$ Kalman filters, the size of the belief state will increase exponentially at each time step to $n^t$. As such, inferring the probability of every state at each time step becomes intractable.

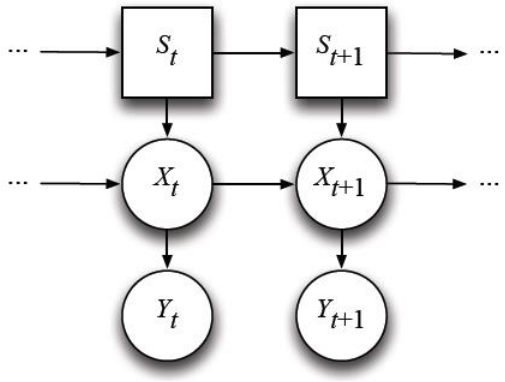


Figure 36 Dynamic Bayesian Network representation of a Switching Kalman Filter[126].

To overcome this problem, approximation method like the Generalised Pseudo Bayseian (GPB) algorithm as described in [120] was adopted. In each time step, the state and covariance estimates from all the filters in the previous time step are combined with weights assigned according to the mix probabilities of the model switch variable, $S_t^{i|j}$ and the model transition probability, $Z_{ij}$ as shown in Equation (28) and (29).

Model switching probabilities:

$$S_t^{i|j} = \frac{Z_{ij} S_{t-1}^i}{\sum_{i=1}^n Z_{ij} S_{t-1}^i} \tag{28}$$

Weighted state and covariance estimates:

$$\begin{aligned} \tilde{x}_{t-1}^j &= \sum_{i=1}^n S_t^{i|j} x_{t-1}^i \\ \tilde{P}_{t-1}^j &= \sum_{i=1}^n S_t^{i|j} \left\{ P_{t-1}^i + \left[ x_{t-1}^i - x_{t-1}^j \right] \left[ x_{t-1}^i - x_{t-1}^j \right]' \right\} \end{aligned} \tag{29}$$

With the weighted state and covariance estimates, the usual Kalman filter as shown in Equation (2) and (3) is carried out for each filter model with each yielding a predicted state, $\hat{x}_{t-1}^j$ and covariance, $\hat{P}_{t-1}^j$estimate. The likelihood of each filter is then determined with Equation (30) using their measurement residual,$v_t^i$. The probability of each model at the current time step can then be obtained as shown in Equation (31). The weighted state and covariance estimate update for the current time can also be determined using Equation (32). A detailed description of SKF is available in [126] and a good demonstration of SKF with use of GPB is shown in [131].

Likelihood of filter from measurement residual:

$$L_t^i = N(v_t^i; 0, C_t^i) \tag{30}$$

Probability of each model:

$$S_t^i = \frac{L_t^i \left( \sum_{i=1}^n Z_{ij} S_{t-1}^i \right)}{\sum_{i=1}^n \left( L_t^i \sum_{i=1}^n Z_{ij} S_{t-1}^i \right)} \tag{31}$$

The weighted state and covariance estimate update are computed as follows:

$$x_t = \sum_{i=1}^{n} S_t^i \, x_t^i$$
$$P_t = \sum_{i=1}^{n} S_t^i \left\{ P_t^i \left[ x_t^i - x_k \right] \left[ x_{t-1}^i - x_t \right]' \right\} \tag{32}$$

A concern from this method is that approximation of the posterior state by combining the previous time steps would contain error and that this error will accumulate over many time steps as highlighted in [126]. However, it was discussed in [126] and shown in [132] that the error will remain bounded regardless of the number of time steps. A detailed description of SKF is available in [126] and a good demonstration of SKF with use of GPB2 algorithm is shown in [131].

In the context of degradation modeling, SKF is applied to track the different type of bearing degradation processes shown in Figure 35 as it evolves. The different types of degradation behaviors have to be known or estimated based on heuristics. The SKF consists of multiple state-space models; each of them describing a different degradation behavior using the basic KF or EKF depending on its underlying function. The SKF then switches between these models based on their likelihood calculated from the feature data. In this way, the most probable degradation model is used for RUL prediction and reduces errors introduced when an unsuitable degradation model is used to model the feature data. It should be noted that by tracking the dynamical behavior of different degradation processes, fault detection can be performed without using pre-established detection thresholds. In addition, this approach helps maintainers to predict RUL more accurately by distinguishing between stable and unstable wear and performing prediction only when unstable wear is detected.

### 7.3 SKF formulation using EKF for tracking varying degradation processes

As mentioned earlier, the degradation processes has to be known for the SKF method to be applied. For bearing degradation, it is well established that its vibration-based signals is stationary when functioning normally and exhibits linear or exponential

trends when it degrades [92; 93; 133]. It is then assumed that the bearing degradation is monotonically increasing and it can evolve from normally operating to stable wear and then unstable wear or from normal operating to unstable wear directly. The dynamics of these processes are represented using a zero and first order polynomial and an exponential model respectively. A Kalman filter is built for each of them and they are used together in the SKF. The state equations for each KF are obtained by 'stacking' the base state, $x$ with the unknown and time-varying parameters, θ for the model. The $x$ and θ are then estimated by each KF using the updated measurements at each time step. The SKF then calculates the most likely model for the time step given the measurement. For the exponential filter, extended Kalman filter is applied due to its non-linear form. The state transition $F_i(.)$ is obtained from the Jacobian of the state equations using Eqn (4). It is assumed that the process noise entering the system only consists of zero mean white noise $q_a$ and $q_b$ which models the wear rate parameters $a_t$ and $b_t$ stochastically for both stable and unstable wear respectively. The state, transition and process noise covariance for each filter are shown below with subscripts 1, 2 and 3 denoting the zero, first order and exponential Kalman filters respectively.

*Zero Order polynomial model (Normal Operation)*

$$\begin{aligned} &\textit{State:} && x_t = x_{t-1} \\ &\textit{State Transition:} && F_{1,t} = 1 \\ &\textit{Process Noise:} && Q_{1,t} = 0 \\ &\textit{Measurement:} && y_t = x_t + r_t \\ & && H_{1,t} = 1 \end{aligned} \tag{33}$$

*1st Order polynomial model (Stable Wear)*

$$\begin{aligned} &\textit{State:} && x_t = x_{t-1} + a_{t-1}\Delta t \\ & && a_t = a_{t-1} \\ &\textit{State Transition:} && F_{2,t} = \begin{bmatrix} 1 & \Delta t \\ 0 & 1 \end{bmatrix} \\ &\textit{Process Noise:} && Q_{2,t} = \begin{bmatrix} 0 & 0 \\ 0 & q_a \end{bmatrix} \end{aligned} \tag{34}$$

*Measurement:*

$$y_t = x_t + r_t$$
$$H_{2,t} = [1 \quad 0]'$$

*Exponential model (Unstable Wear)*

*State:*

$$x_t = x_{t-1} e^{b_{t-1}\Delta t}$$
$$b_t = b_{t-1}$$

*State Transition:*

$$F_{3,t} = \begin{bmatrix} e^{b_{t-1}\Delta t} & x_{t-1}\Delta t e^{b_{t-1}\Delta t} \\ 0 & 1 \end{bmatrix} \quad (35)$$

*Process Noise:*

$$Q_{3,t} = \begin{bmatrix} 0 & 0 \\ 0 & q_b \end{bmatrix}$$

*Measurement:*

$$y_t = x_t + r_t$$
$$H_{3,t} = [1 \quad 0]'$$

*Model transition matrix*

$$Z = \begin{bmatrix} 0.999 & 0.0005 & 0.0005 \\ \sim 0 & 0.999 & 0.001 \\ \sim 0 & \sim 0 & \sim 1 \end{bmatrix} \quad (36)$$

*Initial model probabilities, state and covariance estimate:*

$$S_0 = [0.98 \quad 0.01 \quad 0.01]$$
$$x_0 = y_{0,}$$
$$a_0 = 0 \quad (37)$$
$$b_0 = 0$$
$$P_0 = I$$

For the SKF, the state transition matrix $Z$ is set such that the system tends to remain in its own state with $Z_{ii} \sim 1$. It is also assumed that the degradation rate can only progress i.e. from normal to stable and unstable degradation but not the reverse. However, $Z_{ij}$ is assigned a value approximately zero for $i>j$ as a value of zero can cause underflow problems in Eqn (31) when implemented as a software program. The initial model probability, $S_0$ is set with high probability that it's in normal condition. The initial state estimate, $x_0$ is initialized to the first measurement and initial parameters $a_0$ and $b_0$ are

zero. The initial covariance matrix, $P_0$ is set arbitrarily with an identity matrix, $I$. As such, the input parameters required using this approach are the measurement noise, $r_t$ and the process noise, $q$ from the respective models.

### 7.4 Diagnostic using SKF on simulated data

The SKF approach to diagnose the degradation processes is tested here using simulated data. Figure 37 shows three different evolving degradation processes; (1) normally operating to stable wear at $t = 150hrs$, (2) normally operating to unstable wear at $t = 150hrs$ and (3) normally operating to stable wear at $t = 100hrs$ and then unstable wear at $t = 200hrs$. The simulated degradation measurements are generated using the measurement equations from Eqn. (33-35) where they are modeled with a stationary, linear and exponential function respectively. An additive measurement noise, $r \sim N(0, 0.08^2)$ is added all three processes. For stable wear, a wear rate parameter, $a = 0.01$ is arbitrarily adopted with process noise, $q_a \sim N(0, 0.001^2)$. Similarly for unstable wear, an arbitrary wear rate parameter, $b = 0.04$ is adopted with process noise, $q_b \sim N(0, 0.004^2)$.

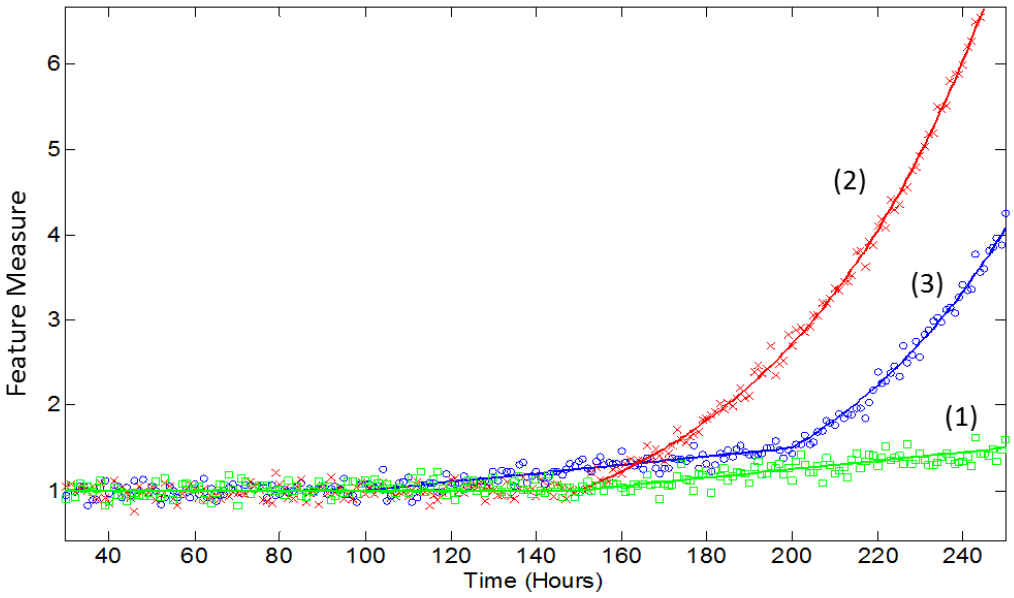


Figure 37 Simulated degradation processes with measurement & process noise: (1) normal to stable wear at t = 150hrs, (2) normally operating to unstable wear at t = 150hrs and (3) normally operating to stable wear at t = 100hrs & unstable wear at t

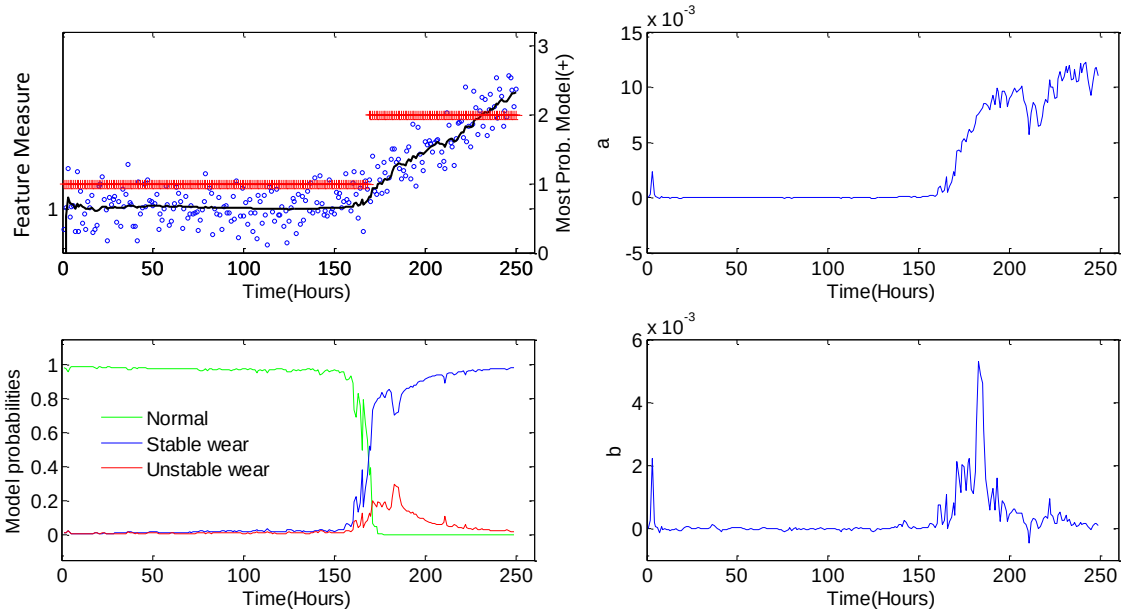


Figure 38 Normal to stable wear (Top left) Filtered state and most probable model, (Bottom left) Model probabilities, (Top & bottom right) Estimated parameters $a_t$ & $b_t$

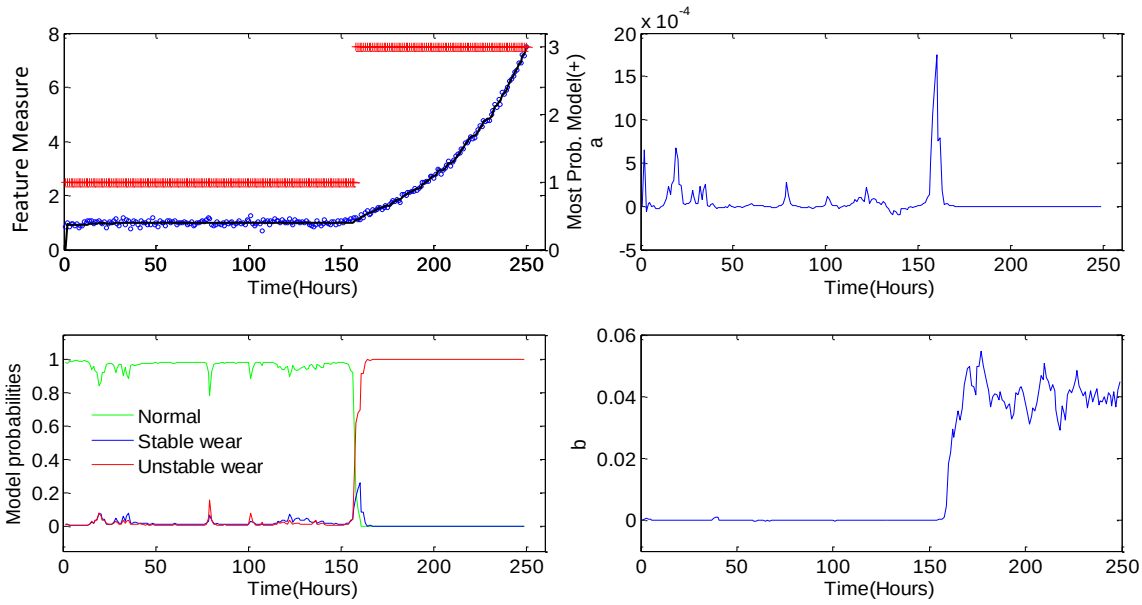


Figure 39 Normal to unstable wear (Top left) Filtered state and most probable model, (Bottom left) Model probabilities, (Top & bottom right) Estimated parameters $a_t$ & $b_t$

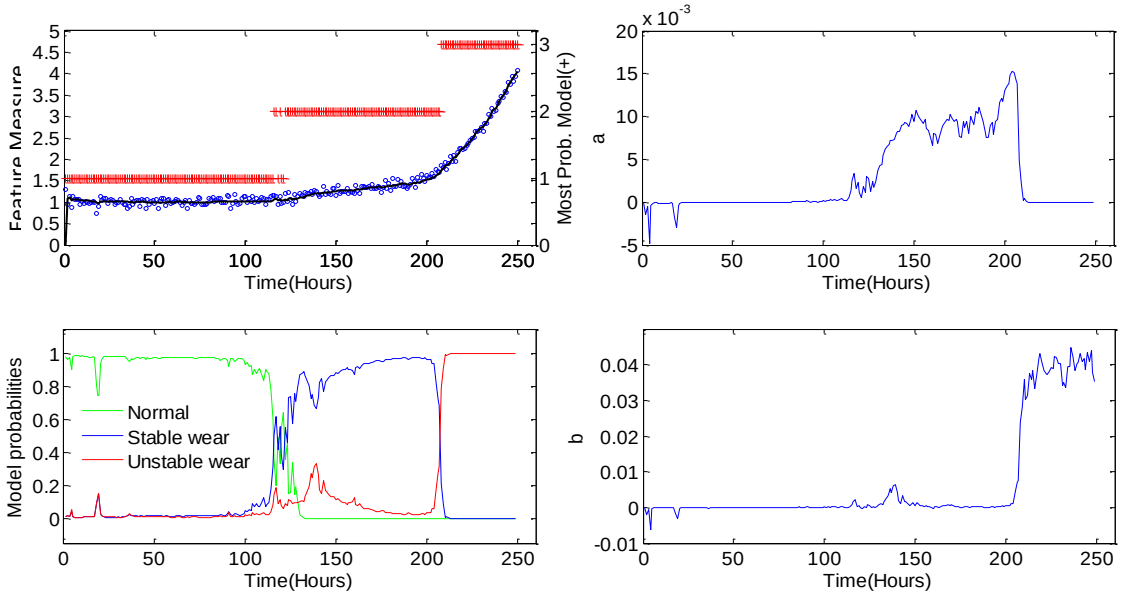


Figure 40 Normal to stable and unstable wear (Top left) Filtered state and most probable model, (Bottom left) Model probabilities, (Top and bottom right) Estimated parameters $a_t$ and $b_t$

Figure 38 to Figure 40 shows the SKF results in tracking the three evolving degradation processes. It can be seen that the SKF is able to track and estimate the most probable degradation process well using the dynamical behavior of the measurement. For normal to steady wear, the SKF detects the change at 169hrs compared to 150hrs. For normal to unsteady wear, the SKF detects the change at 158hrs compared to 150hrs. For normal to steady and then unsteady wear, the SKF detects the change at 116hrs and 208hrs compared to 100hrs and 200hrs respectively. The SKF lags behind the actual transition times as it is performing the estimation in real-time and requires adequate measurements from the dynamical process. In addition, it can estimate the wear rate parameters, *a* and *b* well at ~0.001 and ~0.04. It should be noted that the estimation will not converge towards the exact parameter value due to inherent noise added to the measurements. The ideal case where the dynamical models of the degradation processes and their measurement and process noise are known is shown here. In practice, the dynamical model has to been selected based on expert knowledge and the additive noise has to be estimated. In addition, the noises may be non-Gaussian which would not be handled well by using the KF and EKF formulation.

## 7.5 Case study on AH64D Helicopter Tail Rotor Gearbox Bearing

The SKF approach is applied to extracted feature data from the AH64D Tail Rotor Gearboxes (TRGB) to evaluate its performance for an in-service scenario. The Low Band Energy feature from the three TRGB obtained earlier is used here and shown in Figure 41 with the HUMS monitored hours shown instead of the component hours to simplify the time scale. A general trend of stationary, linear and then exponential rise can be seen across the TRGBs but the rate and duration in each stage differs between individual gearboxes.

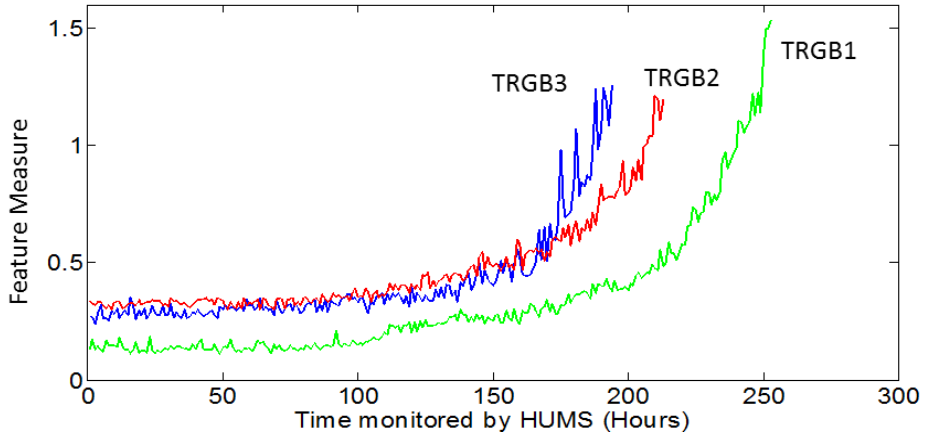


Figure 41 Feature plot for three defective TRGBs (HUMS monitored hours)

As the actual time at which the bearing degradation processes transits cannot be physically observed, they are inferred using piece-wise or segmented regression [134] on the complete feature data history of TRGB 1 as shown in Figure 42. The feature history is segmented into the three degradation stages with two transition times or "breakpoints". A zero, first order polynomial and an exponential regression is then applied to the segmented feature data for each degradation stage and the sum of their residuals is obtained. This procedure is repeated iteratively with different sets of transition times and the optimal piece-wise regression fit is obtained from the set of transition time that minimizes the sum of the residual. Figure 42 shows the transition times for the degradation stages are at 87hrs and 202hrs for TRGB 1. Although these may not be the real transition time, the piece-wise regression provides the optimal fit for each stage of the degradation process. In practice, the complete feature history will

not be available thus the SKF is used as a decision support tool to diagnose the different stages of bearing degradation. When the degradation is diagnosed to be fast and unstable, the RUL of the gearbox bearing is then estimated. RUL estimation is not carried out during the stable wear as the prediction will be over-estimated although maintainers may choose to monitor the condition more frequently. It is possible to establish a threshold limit for the onset of unstable wear if adequate past failure datasets are available to predict the time before unstable wear occurs. It is not considered here however as the goal is to estimate the RUL.

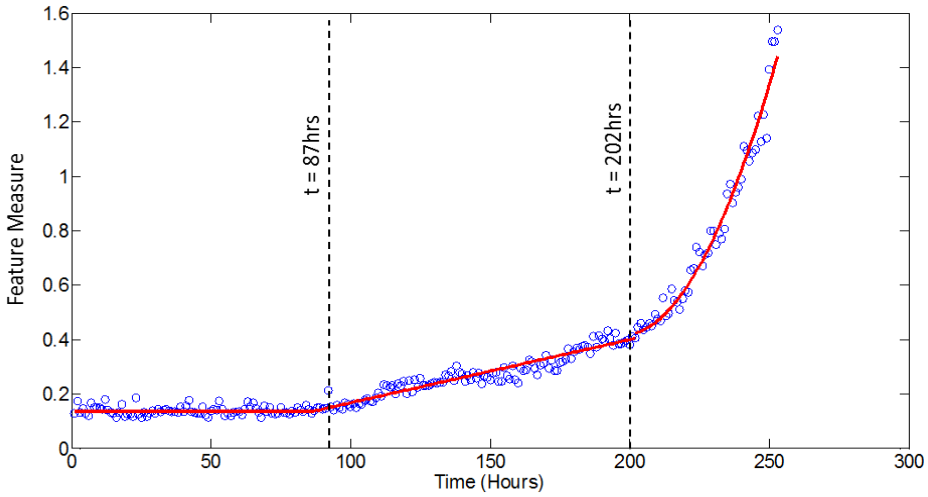


Figure 42 Degradation transitions for TRGB 2 inferred from piece-wise regression

### 7.5.1 Measurement Error

The measurement error for TRGB 2, $r = 3.2e\text{-}4$ is obtained by taking the variance of the stationary measurements when the TRGB is in a good condition. It is noted that the variance for this gearbox is smaller than the variance obtained across a range of gearboxes as obtained from the GMM classification above as shown in Table 6. In practice, the measurement noise is likely to vary between individual gearboxes due to slight differences in mechanical properties or transmission path.

### 7.5.2 Process Noise

The process noises, $q_a$ and $q_b$ contain the uncertainty of the filter in modeling the real world. The process noise acts as a 'memory' for the respective Kalman filter model

whereby setting it to zero will cause it to remember all past measurements and unresponsive to new ones and conversely, a high process noise will cause the filter to be more responsive to new measurements [121]. In the SKF approach, the process noise for the respective KF, is obtained by tuning the model with past similar defect cases and is assumed to be the same across gearboxes. The SKF formulations are applied with $q_a$ and $q_b$ set as a small percent of the estimated or guessed value of the parameters [135]. One approach is to estimate the parameters by regression analysis of the segment of the feature data which is inspected to fit the parameter's model. The SKF model is then simulated and the parameters tuned till the model is acceptably consistent. This is where training of the SKF is required from similar failure cases. In this study, the process noises were tuned using feature data from TRGB 2 and 3 before being tested on TRGB 1. $q_a$ and $q_b$ were estimated after tuning to be 1e-7 and 1.5e-6 respectively. Figure 43 and Figure 44 shows the results of the SKF on TRGB 2 and 3 after the process noise were tuned and it can be seen that the filter tracks the various stages of degradation consistently in both gearboxes. As the wear rate transits, the probability of the assumed model changes accordingly. Both the stable and unstable wear rate parameter, $a_t$ and $b_t$ are tracked.

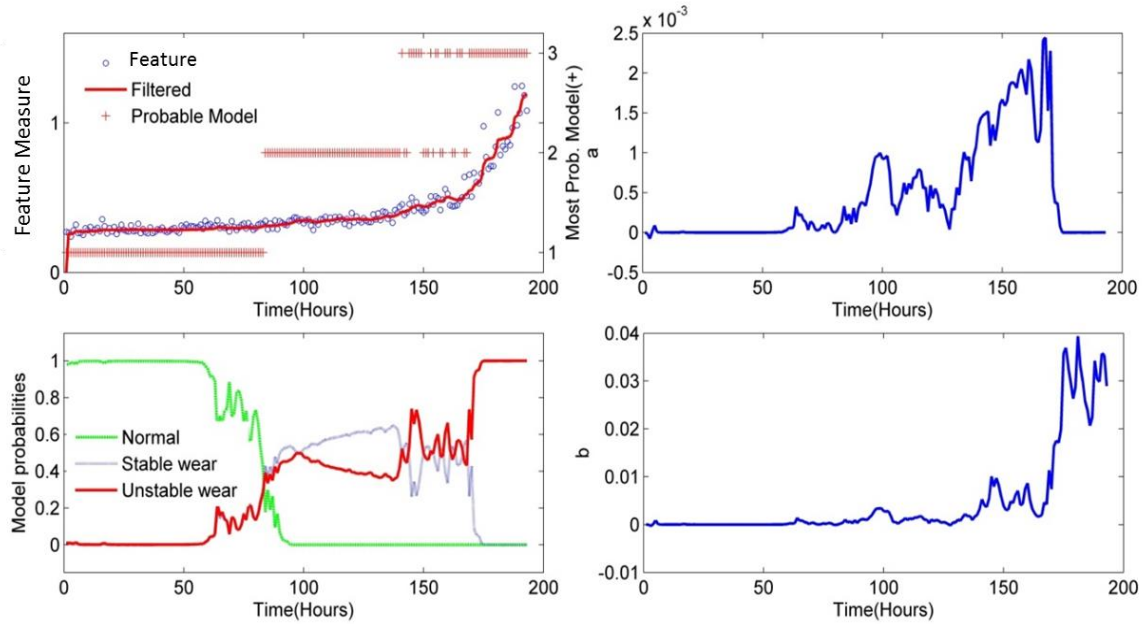


Figure 43(Top left) TRGB 2 Feature data, filtered state and most probable model, (Bottom left) Model probabilities, (Top & bottom right) Estimated parameters $a_t$ & $b_t$.

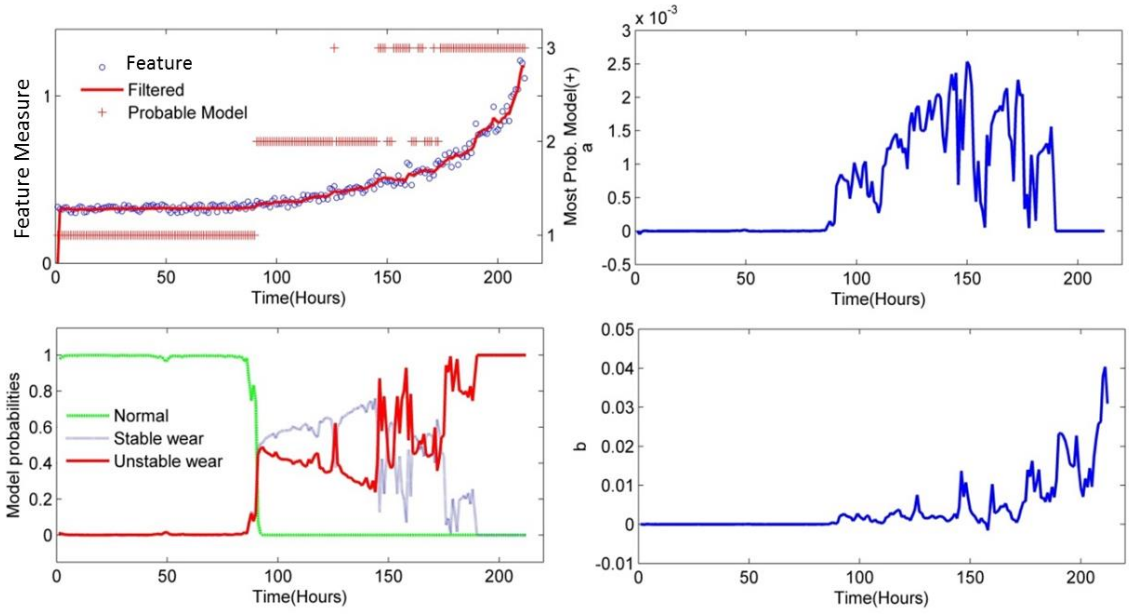


Figure 44 (Top left) TRGB 3 Feature data, filtered state and most probable model, (Bottom left) Model probabilities, (Top & bottom right) Estimated parameters $a_t$ & $b_t$.

In Figure 43, it can be seen that $b_t$ stabilizes as it approaches the failure time which is desirable as it allows a more accurate RUL prediction to be made. From both Figure 43 and Figure 44, it can be seen that the filter does not perform well when the feature data are not monotonously rising and it then has to take a longer time before it converges. Similar observations have also been made by [103] when using polynomial KF. This may cause false alarms of unstable wear in practice and a simple way to reduce it is for maintainers to evaluate the probabilities between stable and unstable wear.

### 7.6 SKF diagnostics on AH64D TRGB bearing

The formulated SKF model is applied to the TRGB 3 bearing feature data and the results are shown in Figure 45. The SKF can adaptively track the different bearing degradation processes with the process noise tuned from the other two gearboxes. When compared to the optimal solution from the piece-wise regression, the SKF transition time lags behind at 105hrs and 212hrs compared to 87hrs and 202hrs. However, it should be noted that the SKF is performing the estimation in real-time and requires adequate measurements from the dynamical process. When the features are not increasing monotonically at ~200hrs, the SKF has to take a longer time before it

converges as mentioned above. Instead of relying on the absolute value of the CM measurements, the SKF uses the dynamic behavior between the current and past measurement to diagnose the degradation state. Therefore, it is not dependent on a fixed threshold which are typically derived from statistical evaluation of large numbers of past failure cases.

Another key advantage of this technique for diagnosis is that it provides the probability of the degradation process the bearing is undergoing. In comparison, the widely used, statistical process control (SPC) approach only triggers when the feature is above a statistical limit and no further information is available. In practice, maintenance engineers using SPC would then wait for future measurements to confirm if it's a fault or false alarm before prescribing any tasks. The quantitative probability measure from the SKF allows more support for maintenance engineers as the probabilities of the bearing conditions can be compared in the event of an outlier measurement.

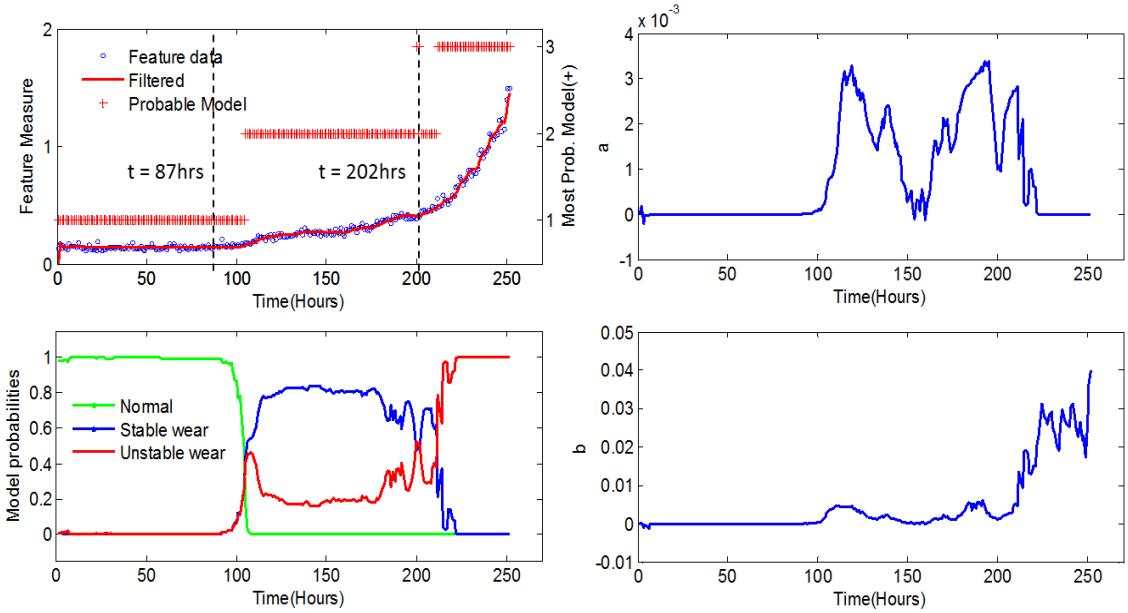


Figure 45 (Top left) TRGB 1 feature data, filtered state and most probable model, (Bottom left) Model probabilities, (Top & bottom right) Estimated parameters $a_t$ & $b_t$

## 7.7 Estimation of remaining useful life

For prognostics, the RUL of the bearing is predicted whenever an unsteady degradation is detected. The SKF infers the most probable dynamic model to be applied at each

time step for prediction. The RUL is predicted by propagating the weighted state and covariance estimates obtained from the Ricatti Equations at each time step and determining the time when the degradation state crosses the failure threshold. Figure 46 shows the RUL forecast when the SKF detects unstable degradation after 212hrs and it can be seen that the accuracy of the RUL estimate improves as time progresses.

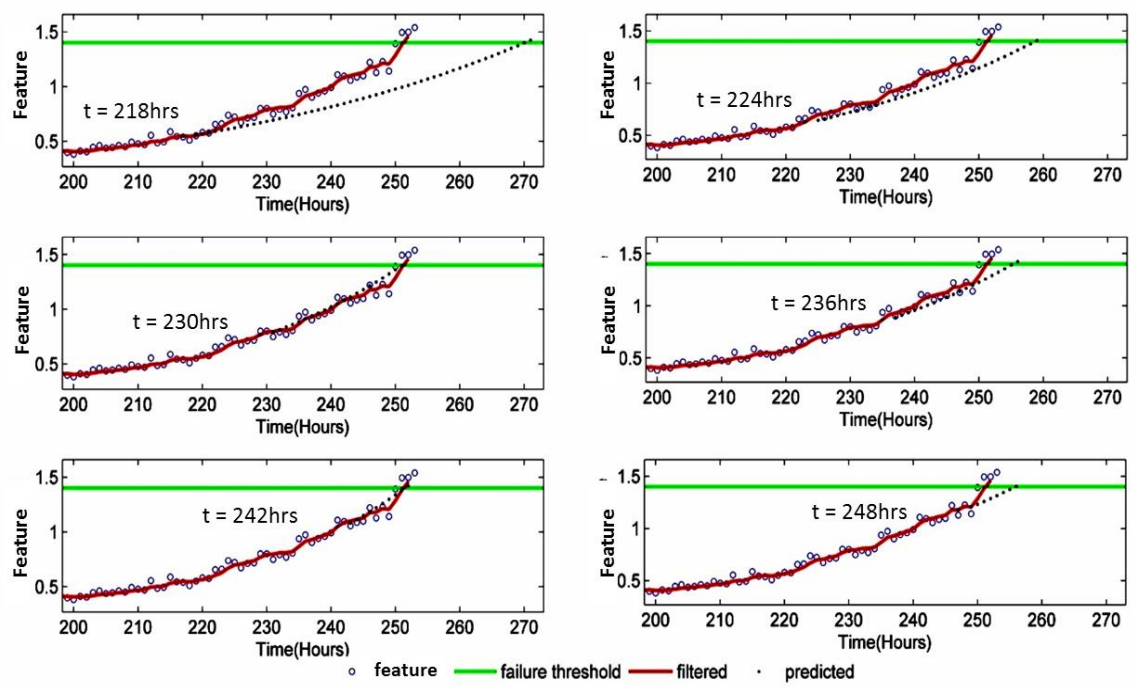


Figure 46 RUL forecast using SKF prediction at different times

### 7.7.1 Estimation of RUL confidence bounds using Monte Carlo Simulation

Besides the RUL prediction, its confidence bounds, particularly the lower bound, is important information for a prognostic tool to be effective in aiding decision support. In this section, the confidence bounds for the RUL predictions are estimated and a performance metric is used to assess the overall prognostic capability. For a linear process, such as in steady degradation, the RUL probability density function can be obtained with closed form solution using a special case of Bernstein distribution [136] after which the confidence bounds can be obtained. For non-linear unsteady degradation processes however, a closed-form solution is generally not available and a Monte-Carlo based approach [136; 137] has to be adopted to estimate the RUL

probability density function instead. At each of the prediction steps, a large $N$ sample of the states are randomly drawn from the multivariate distribution, $N(\vec{x}_t, P_t)$ and a RUL estimate is computed for each sample. A histogram can then be built from the $N$ number of RULs and be used to approximate the density function. Figure 47 shows the RUL histogram built from the Monte-Carlo method at different time steps and it can be seen that the accuracy and precision of the RUL estimates generally improves over time. From the RUL pdf, the 90% confidence bound for the RUL is obtained.

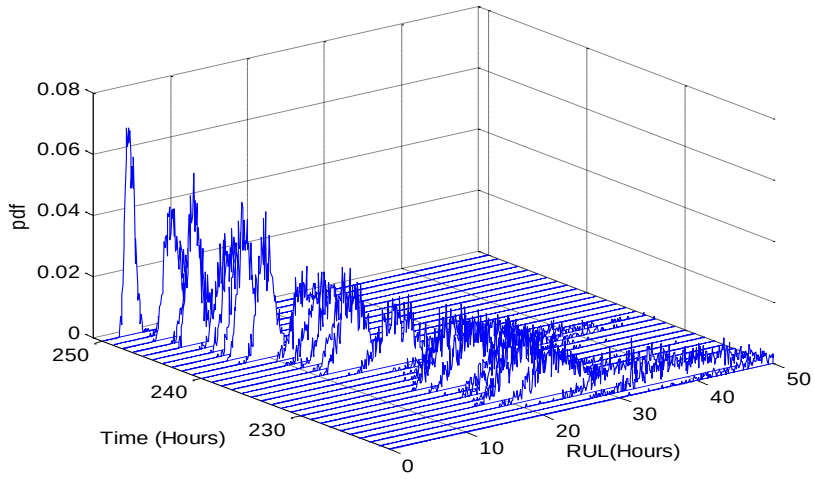


Figure 47 RUL probability density function from Monte Carlo Simulation

### 7.7.2 Prognostic performance metric

The α-λ metric [137] is applied to evaluate the performance of this prognostic evaluation as shown in Figure 48. The α-λ metric compares the actual RUL to the predicted RUL with converging α bounds that provides an accuracy region. The α bounds are application specific and a prediction is correct if it falls within the alpha bounds. Besides, the predicted mean RUL, the upper and lower 90% confidence bounds are shown as well. However, there are points on the RUL trajectory that lies outside the accuracy zone towards the end of useful life which is a behavior reportedly observed in [138] as well. This behavior could be attributed to unsteady vibration levels as the accumulated damage in the bearing becomes sizeable and could perhaps be addressed by lowering the failure threshold limit. Besides the RUL estimate, most of the lower

confidence bound, which is important for conservative estimate of the RUL prediction are close to the lower 30% accuracy bound as well

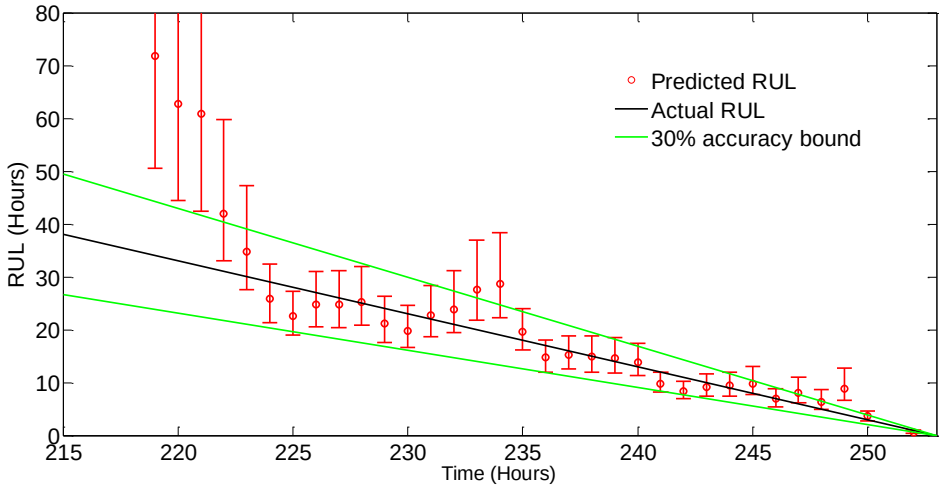


Figure 48 $\alpha$ - $\lambda$ performance metric with 90% confidence interval of predicted RUL

### 7.8 Comparison with use of EKF

Figure 49 shows a comparison between the SKF method with the use of EKF with exponential formulation alone. A log scale is applied on the feature measure axis so that difference in predictions between the two models can be seen more easily. It can be seen that the SKF can better adapt to the different degradation processes to yield lower prediction errors compared to the use of EKF alone. This is most prominent in the normal condition as the exponential model adopted in EKF is more susceptible to fluctuations compared to the stationary model used in the SKF. In the stable wear stage, the difference from both models is less as the non-linear EKF can approximate the linear behavior in this stage as well as the linear model in the SKF. In fact, the EKF prediction outperforms the SKF near transition points at 87hrs and 202hrs as the SKF takes time to converge. In steady degradation, the prediction RMS errors are at 4.9e-4 and 4.7e-4 for SKF and EKF respectively. In the time region away from the transition points between 110hrs to 190hrs, the prediction RMS error for SKF was lower at 3.9e-4 against 4.4e-4 for EKF respectively. More importantly, the EKF does not provide information on the degradation process compared to the SKF. As mentioned, this is useful as it informs maintainers of the probability that the RUL predicted will be

accurate. When the EKF is applied alone, there will significantly higher inaccurate RUL predictions in the normal and stable wear stages. Notably, both the SKF and EKF yield the same results as the state progresses into unstable degradation. This shows that the errors from the approximation using the GPB algorithm does not accumulate and the SKF will converge towards the underlying state as mentioned earlier,

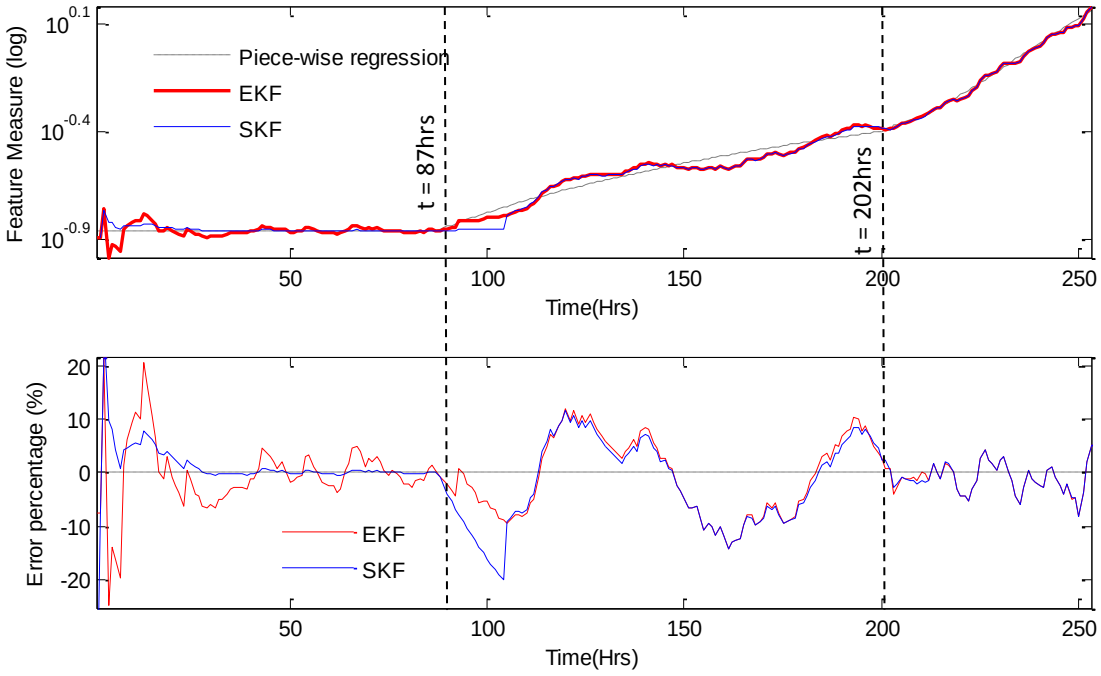


Figure 49 (Top) Comparison of predicted states using EKF with SKF (log scale) , (Bottom) Prediction errors of EKF and SKF from piece-wise regression

## 7.9 Polynomial Kalman Filter with SKF

In the above study, the unsteady wear was represented with an exponential function which was then modeled using EKF due to its non-linear state-space form. The EKF is also referred to as a suboptimal filter as the optimal solution is not guaranteed, unlike the linear Kalman Filter. In this section, the use of a 2$^{nd}$ order polynomial function in place of the exponential function is investigated. The advantage of using a higher order polynomial function is that its state-space form is linear and the simpler Kalman Filter can be applied. As such, there is no need to calculate the Jacobian matrix in each step which reduces the computational resources. The formulation and use of polynomial Kalman filter in the SKF approach is shown herein.

### 7.9.1 Polynomial Kalman Formulation

Like in the EKF formulation above, it is assumed that the bearing degradation is monotonically increasing and it evolves from serviceable to stable wear and then unstable wear. The dynamics of these processes are represented here using zero, first and second order polynomial Kalman filters respectively. The development and application of polynomial KF is extensively covered by Zarchan et al in [121]. As described in that reference, the state transition matrix for different order polynomials can be derived from the Taylor series expansion of the fundamental matrix. Similarly, the process noise covariance matrix for different order polynomials can be derived from the fundamental matrix as well. The fundamental state, state transition and process noise covariance matrices describing the polynomial filters are shown below with subscripts 1, 2, 3 denoting the zero, first and second order KF respectively. The model transition matrix, Z and the initial model probabilities, state and covariance estimates are the same from Eqn. (36) and Eqn. (37) above.

*Zero Order polynomial model (Normal Operation)*

$$
\begin{aligned}
&\textit{State:} && x_t = x_{t-1} \\
&\textit{State Transition:} && F_{1,t} = 1 \\
&\textit{Process Noise:} && Q_{1,t} = 0 \\
&\textit{Measurement:} && y_t = x_t + r_t \\
& && H_{1,t} = 1
\end{aligned}
\tag{38}
$$

*1st Order polynomial model (Stable Wear)*

$$
\begin{aligned}
&\textit{State:} && x_t = x_{t-1} + \dot{x}_{t-1}\Delta t \\
& && \dot{x}_t = \dot{x}_{t-1} \\
&\textit{State Transition:} && F_{2,t} = \begin{bmatrix} 1 & \Delta t \\ 0 & 1 \end{bmatrix} \\
&\textit{Process Noise:} && Q_{2,t} = q_s \begin{bmatrix} \frac{\Delta t^3}{3} & \frac{\Delta t^2}{2} \\ \frac{\Delta t^2}{2} & \Delta t \end{bmatrix}
\end{aligned}
\tag{39}
$$

*Measurement:* $y_t = x_t + r_t$

$$H_{2,t} = [1 \quad 0]'$$

*2nd Order polynomial model (Unstable Wear)*

*State:*

$$x_t = x_{t-1} + \dot{x}_{t-1}\Delta t + \ddot{x}_{t-1}\frac{\Delta t^2}{2}$$

$$\dot{x}_t = \dot{x}_{t-1} + \ddot{x}_{t-1}\Delta t$$

$$\ddot{x}_t = \ddot{x}_{t-1}$$

*State Transition:*

$$F_{3,t} = \begin{bmatrix} 1 & \Delta t & \frac{\Delta t^2}{2} \\ 0 & 1 & \Delta t \\ 0 & 0 & 1 \end{bmatrix} \tag{40}$$

*Process Noise:*

$$Q_{3,t} = q_s \begin{bmatrix} \frac{\Delta t^5}{20} & \frac{\Delta t^4}{8} & \frac{\Delta t^3}{6} \\ \frac{\Delta t^4}{8} & \frac{\Delta t^3}{3} & \frac{\Delta t^2}{2} \\ \frac{\Delta t^3}{6} & \frac{\Delta t^2}{2} & \Delta t \end{bmatrix}$$

*Measurement:* $y_t = x_t + r_t$

$$H_{3,t} = [1 \quad 0 \quad 0]'$$

### 7.9.2 Application of Polynomial Kalman Filter on AH64D dataset

The SKF approach using polynomial KF formulation is applied to the same AH64D dataset to compare its performance against the use of EKF formulation. Just as before, the process noise, $q_s$ is set initially as a small percentage of the measurement error, $r$ and tuned with the feature data from TRGB 2 and 3 and till the model is acceptably consistent yet responsive to changes in the degradation processes. It should be noted that both first and second order KF shares the same $q_s$ that need to be estimated and tuned here. This is in contrast with the EKF approach described above where the process noises for the linear and exponential filter are estimated separately. The value of $q_s$ tuned from the two gearboxes is 5e-8. Figure 50 and Figure 51 shows the results of the SKF on TRGB 2 and 3 after the process noise were tuned and it can be seen that the SKF model can again track the various stages of bearing degradation consistently in both gearboxes.

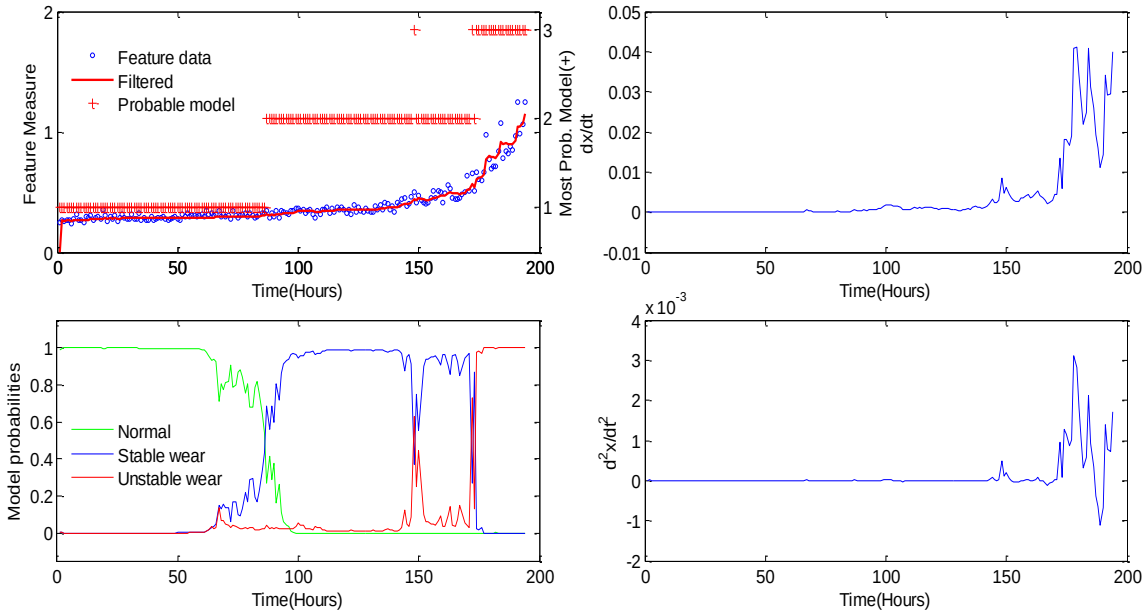


Figure 50 (Top left) TRGB 2 feature data, filtered state and most probable model using polynomial SKF, (Bottom left) Model probabilities, (Top & bottom right) Estimated states $\dot{x}$ & $\ddot{x}$

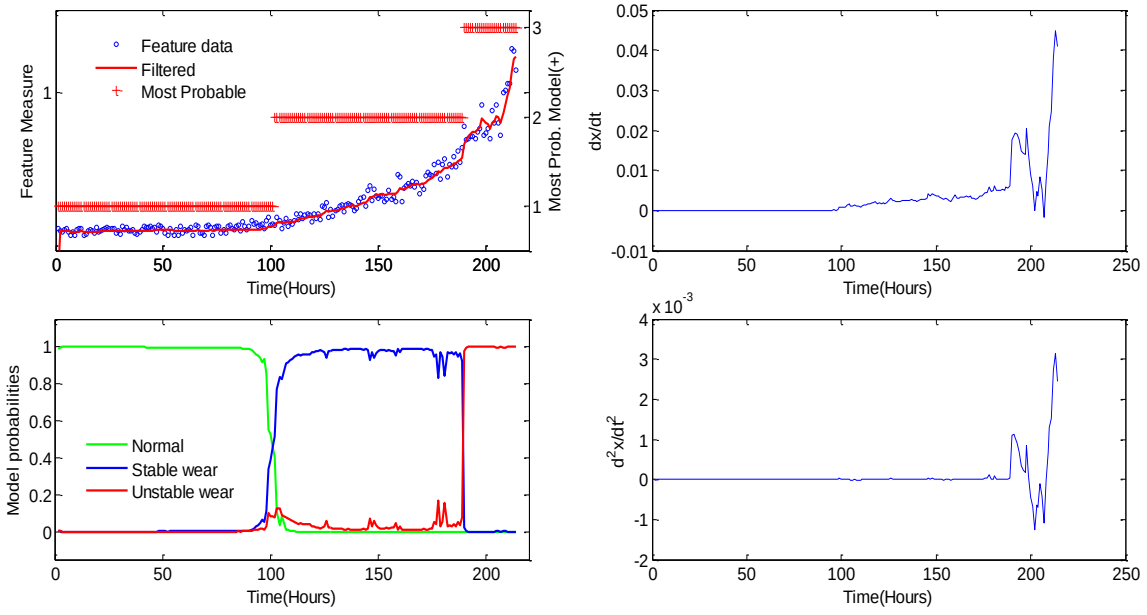


Figure 51 TRGB 3 feature data, filtered state and most probable model using polynomial SKF, (Bottom left) Model probabilities, (Top & bottom right) Estimated states $\dot{x}$ & $\ddot{x}$

With lower $q_s$, the second-order filter will fit the measured data better compared to the first-order filter and the model becomes more sensitive to changes but also prone to over-fitting. Conversely, the first-order filter will fit better with higher $q_s$ and is less sensitive to changes. This means that a lower *qs* can detect unstable wear earlier but this can be undesirable as second-order growth tends to over-predict RUL during initial stages. Note that this describes the effect of the process noise on the probabilities between the first and second order KF within the SKF method. Within each filter, a high process noise will cause the filter to be more responsive to new measurements as mentioned above. The upper and lower right plots shown in Figure 50 and Figure 51 shows the rate and acceleration of the feature and not parameters like in the EKF plots above. The process noise, $q_s$ is then applied to TRGB 1 with the same measurement noise, $r$ = 3.2e-4 applied and the results is shown in Figure 52.

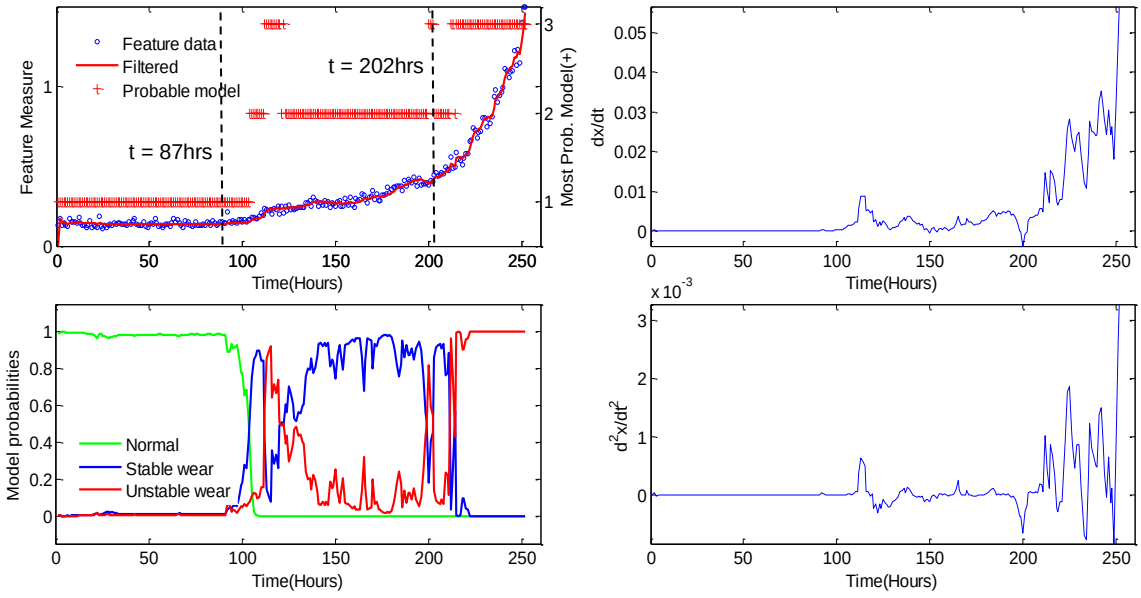


Figure 52 (Top left) TRGB 1 feature data, filtered state and most probable model using polynomial SKF, (Bottom left) Model probabilities, (Top & bottom right) Estimated states $\dot{x}$ & $\ddot{x}$

From Figure 52, it can be seen that the polynomial SKF can generally track the different stages of bearing wear well in comparison with the non-linear EKF formulation shown in Figure 45. The polynomial SKF tracks a change from normal to stable wear at 104hrs but it incorrectly tracks the bearing wear to be unstable at 112hrs before converging back at 122hrs. It then tracks stable to unstable bearing wear at

212hrs. The tracking error is again attributed to non-monotonous trends within the data during which the higher order filter is more dominant. However, this also shows that the SKF can correct itself with more measurement updates. The erroneous tracking of the bearing wear stage can cause false alarms in practice. A way to mitigate this is for maintainers to decide by evaluating the available model probabilities between stable and unstable wear. For example, a decision threshold may be designed such that unsteady degradation is considered only when the model probability is $>0.95$. As such, the unsteady degradation will only be triggered when the likelihood of the degradation model has converged. The drawback however is that lead-time for RUL prediction will be reduced as longer time is required to reach convergence.

### 7.9.3 Estimation of Remaining Useful Life using Polynomial Kalman Filter

The RUL is predicted in the same way by propagating the predicted state and covariance estimates obtained at each time step using and determining the time when the wear state crosses the damage limit threshold.

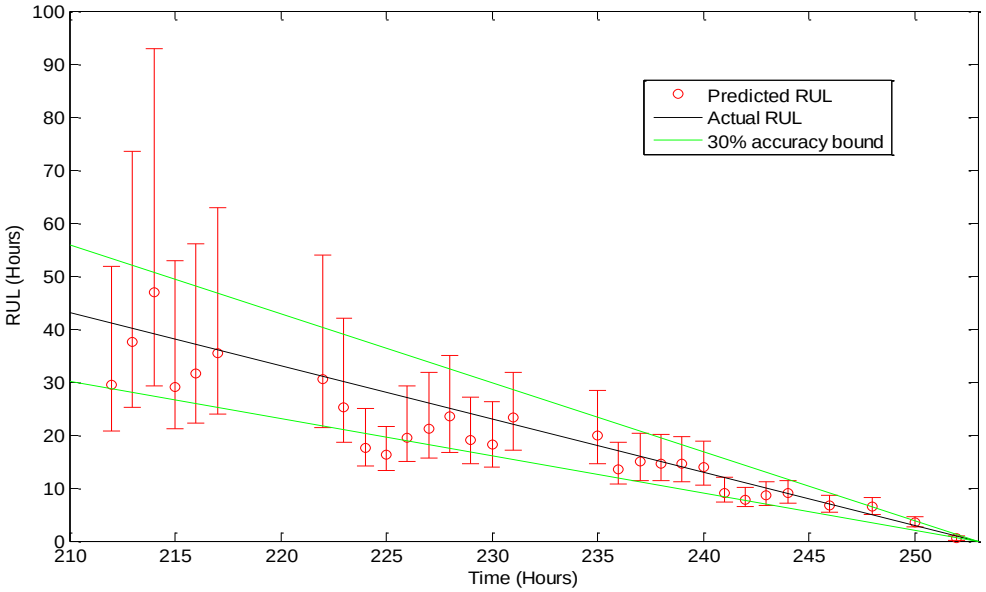


Figure 53 $\alpha$ - $\lambda$ performance metric with 90% confidence interval of predicted RUL using polynomial Kalman filter

Figure 53 shows the RUL forecast when the SKF detects unstable wear after 212 and it can be seen that the accuracy of the RUL does improve as time progresses. However,

the RUL prediction cannot be carried out at points where the feature wear rate, $\dot{x}$ are negative as seen in the time period ~200hrs and ~232hrs from the upper and lower right plots of Figure 52. The bearing wear state propagated from these points does not increase and will not cross the damage threshold; thus a RUL cannot be obtained. As such, there are corresponding missing predictions at those times in Figure 53.

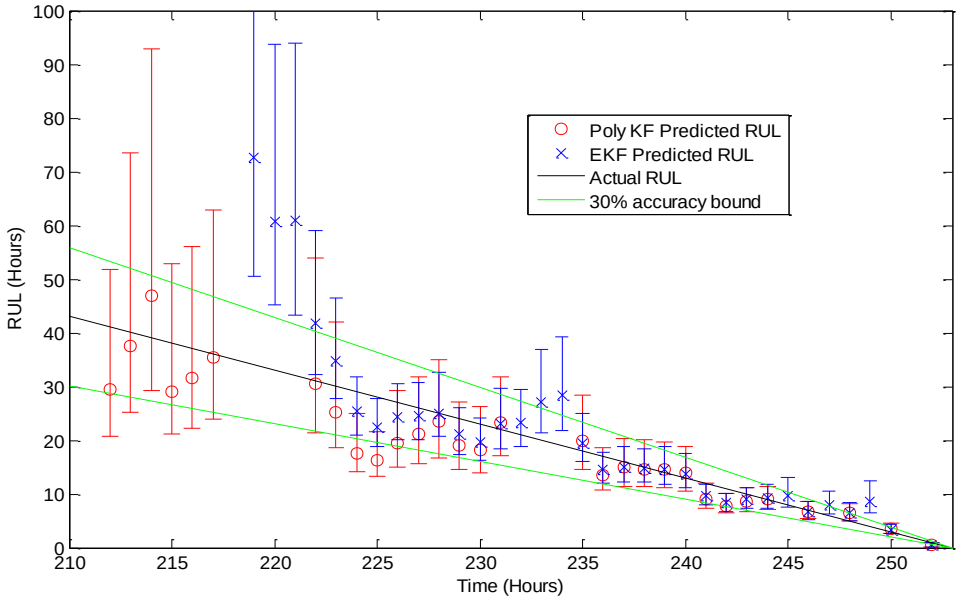


Figure 54 Comparison of α - λ performance between polynomial Kalman filter and Extended Kalman filter

Figure 54 shows the comparison of the RUL prediction between the polynomial SKF with the non-linear EKF formulation. While the polynomial SKF tracks the unsteady wear earlier, it is unable to provide predictions when the data trend is non-monotonous. It is noted however that the predictions from the EKF formulation tends to be poor at the non-monotonous trends also, thus the benefit may not be very significant. As time progresses towards the actual failure time, both formulations give close and accurate prediction of the RUL. Therefore, the use of polynomial SKF can be a good option as it is simpler to implement and require less computational resource due to its linear form.

## 7.10 Chapter Summary

In this study, the use of SKF is applied for fault detection and remaining useful life estimates of a helicopter gearbox bearing is presented with promising results. The SKF model allows for degradation processes to evolve through time from which the underlying dynamical process would be inferred accordingly. This approach can provide maintainers with more information for decision-making as a probabilistic measure of the bearing degradation process and its health status is available. From the prognostic performance metric, it was shown that the RUL estimates have high accuracy when it is inferred that the degradation process is likely to be unstable. This in turn can provide maintainers with higher confidence on the predicted RUL for maintenance planning. The SKF was developed with EKF formulation used for unsteady wear. An alternative formulation using polynomial KF was also investigated for use in the SKF model. It was found that the performance of the polynomial KF is comparable with EKF formulation and it offers the advantage of simpler implementation due to its linear form.

## 7.11 Related Publications

- *Switching Kalman Filter for diagnostic and prognostic of bearing degradation*, (Published Mar 2014), Structural Health Monitoring.
- *Condition Monitoring and Remaining Useful Life estimation using Switching Extended Kalman Filter,* (In Press), special issue for IJSEAM on "Intelligent assets through Condition monitoring and Condition-based maintenance.
- *Fault detection and remaining useful life estimation using Switching Kalman Filters*, (Published Oct 2013), 8th World Congress on Engineering Asset Management.
- *Switching Extended Kalman Filter for diagnostic and prognostic of bearing degradation,* (Under review), Mechanical System Signal Processing.
- *Fault detection and remaining useful life estimation using Switching Extended Kalman Filter,* (Under review), Reliability Engineering and System Safety.

# 8 REDUCING RISK OF LUBRICATION SYSTEM FAILURES IN HELICOPTERS

The lubrication system in a helicopter performs the function of lubricating and cooling the transmission bearings and gears. It is a critical system and up to 31% of transmission related accidents has been estimated to be related to the lubrication system. A review of recent transmission related helicopter accidents showed that backup lubrication system were not effective as they have common causal failure modes with the main system. As such, lubrication systems and associated subsystem such as the accessory gearbox would have 'Hazardous' mishap severity category despite having redundancies built in. An independent backup lubrication system would reduce failure risk significantly but this can be impractical based on existing oil-based designs.

In the 1990s, the National Aeronautics and Space Administration (NASA) proposed a thioether-based mist lubrication which showed promising results of thermal stability and low gear wear after the gearbox was run dry. In their investigation, the thioether liquid is misted and delivered in an airstream to gears operating at such high temperatures that the molecules react on the wearing surfaces to generate a lubricious deposit which provides effective lubrication. The key benefit of the thioether is that a very low volume is required which is ideal in developing a compact backup lubrication system. In section, the thermal properties of a thioether-based lubrication in a gearbox run-dry situation are investigated. An experimental test rig was setup to measure the gear temperature profiles between thioether and a conventional oil lubricant under different load and speed conditions. The work here is largely exploratory to study the feasibility of thioether based mist lubrication as a viable candidate for a backup lubrication system in helicopters.

## 8.1 Review of helicopter gearbox failure incidents

In an early survey conducted in the 1990s, it was shown that an estimated 31% of transmission related helicopter accidents are attributed to the failure of the lubrication system [139]. In the past five years, there were also a number of helicopter accidents

that were due to lubrication failures as shown in Table 8. The key failure modes of lubrication systems from Table 8 can be summarized to be either loss of oil or loss of oil pressure. Class A helicopters such as the S-92 and the AS332 have backup lubrication systems but they did not serve their function in these instances as they have common causal failure modes with the main system. Common causal failures exist when both main and backup systems share the same critical component or subjected to the same failure modes. The designs of backup systems can vary widely between helicopters but most have common causal failures. An independent backup lubrication system would reduce failure risk significantly but this can be impractical based on existing oil-based designs due to costs and impact on payload. In this section, a brief review of existing backup lubrication system design is carried out and the feasibility of a thioether-based backup system is investigated.

Table 8 Lubrication System related incidents

| Date | Helicopter | Country | Description | Reference |
|---|---|---|---|---|
| Apr-05 | Sikorsky S-92A | Norway | Failure of drive of MGB oil pump | [140] |
| Jan-08 | Sikorsky S-92A | Malaysia | MGB input module overheating that led to slow oil leak | [140] |
| Feb-08 | Schweizer 269D-1 | UK | Seizure of MGB pinion outer bearing due to oil starvation | [141] |
| Mar-09 | Sikorsky S-92A | Canada | Total loss of MGB oil due to fracture of oil filter bowl fixing titanium studs | [140] |
| Apr-09 | Aerospatiale AS332 L2 | UK | Loss of MGB oil due to MGB case rupture | [142] |
| May - 12 | Eurocopter EC225 | UK | Loss of oil pressure due to fracture of shaft driving the oil pumps | [143] |

## 8.2 Survey of existing Lubrication System Design

The function of a lubrication system is to lubricate and cool the transmission bearings and gears. Most helicopter lubrication system designs consist of oil sump, oil pump, filters and coolers. The oil pump is driven by the helicopter's accessory gearbox and provides the pressure head to distribute the oil from the sump to the oil gallery. The oil is then filtered and cooled in heat exchangers before being recycled. Common safety measures include oil filters, metallic chip detection, oil temperature and pressure sensors. The difference in design between helicopters depends largely on the layout and redundancies of these common components. For the AS332, the lubrication system

consists of a single oil sump, an oil pump and single oil cooler. Its backup features a single emergency pump that activates upon low oil pressure in the main oil pump [144]. This design has many common causal failures as both systems share the same oil sump and filters. The Sea King main lubrication system consists of a single oil sump, two oil pumps and single oil cooler. It has an independent backup lubrication system consisting of an independent oil sump and emergency pump [145]. The use of independent system is also adopted in the military attack helicopter WAH64D, which consists of a dual redundancy system with each system consisting of an oil sump, an oil pump and oil cooler [146]. Although it can provide higher reliability, an independent system which duplicates the main system is costly to install, takes up significant payload space and is still exposed to the same type of risks. As such, an emergency lubrication system that is light weight and simple in design would be desirable.

## 8.3 Thioether based Backup Lubrication System

NASA had proposed a thioether-based mist lubrication which showed promising results of thermal stability and low gear wear under run-dry conditions [147]. In their method, the thioether liquid is misted and delivered in an airstream to gears operating at such high temperatures that the molecules of the thioether react on the wearing surfaces to generate a lubricious deposit which provides effective lubrication. The key advantage of this approach is that only a very small volume of thioether is required thus the reservoir volume can be kept small. In their work, a flow rate of only 15ml/hr is required to cool a set of gear mesh [147]. This delivery mechanism overcomes the need to recycle the lubricant which further reduces the dependency on the reliabilities of oil pumps and oil coolers. NASA concluded that the use of oil mist-based lubrication was suitable for emergency lubrication [148]. In this investigation, a similar experimental setup is proposed to investigate the effectiveness of a thioether based mist lubrication by comparing the temperature profile of the gears with oil dip and oil mist lubrication.

## 8.4 Lubrication System Test Rig Description

An overview of the experimental setup is depicted in Figure 55 and the physical laboratory setup shown in Figure 56. The rig consists of a gearbox and the mist

lubrication mechanism. The gearbox arrangement is shown in Figure 57. The gear set employed in this test is made of case carburized steel with specifications shown in

Table 9. An AC three-phase electrical motor (1.1 KW) with speed of 690 rpm was employed to drive the gearbox. A simple mechanism that permitted a pair of coupling flanges to be rotated relative to each other, thereby applying a pre-torsional load, was employed to apply torque load onto the gears. The lubricating oil used in the gearbox was Aeroshell 555 which is a common aerospace lubricant for helicopter gearboxes in accordance with DOD-L-85734 and DEF STAN 91-100.

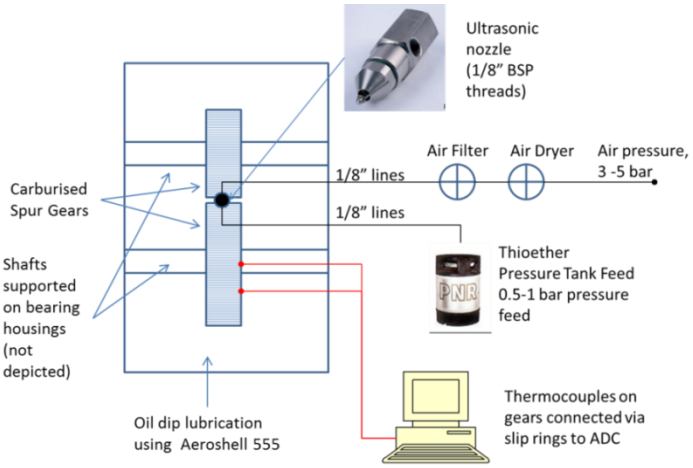


Figure 55 Experimental rig overview

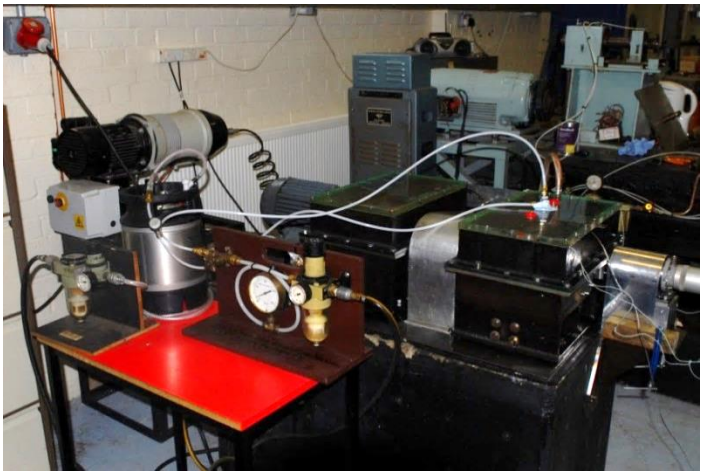

Figure 56 Laboratory setup

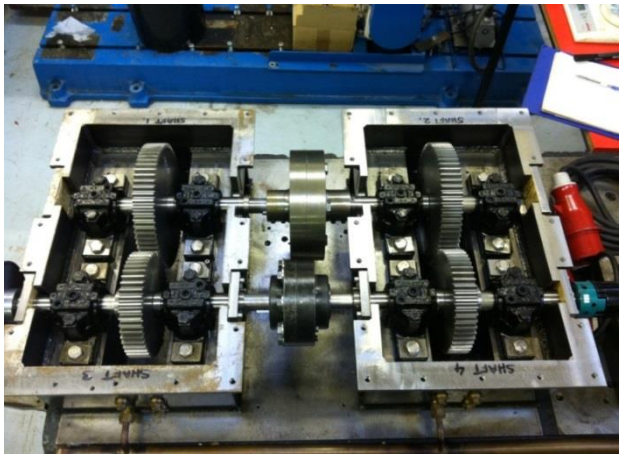

Figure 57 Gearbox back-to-back arrangement

Table 9 Pinion and Gear Specification

| | Spur |
|---|---|
| Number of teeth, pinion: gear | 49: 65 |
| Base diameter, pinion: gear (mm) | 138.13: 183.24 |
| Pitch diameter, pinion: gear (mm) | 147: 195 |
| Tip diameter, pinion: gear (mm) | 153: 201 |
| Root diameter, pinion: gear (mm) | 139.5: 187.5 |
| Contact Ratio | 1.33 |
| Module (mm) | 3 |
| Addendum modification coefficient | 0 |
| Surface roughness, Ra (μm) | 0.8, 2.00 |
| Face width (mm) | 15,30 |
| Pressure Angle (degree) | 20 |
| Helix Angle (degree) | 0 |
| Modulus of Elasticity (Gpa) | 228 |

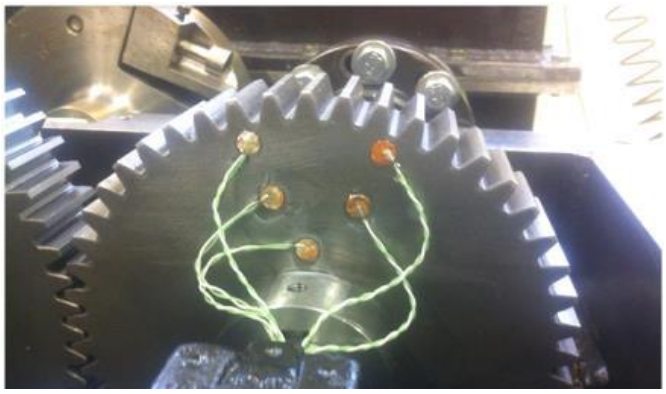

Figure 58 Thermocouple arrangement on gear face

Type K thermocouples were used to measure the temperature profiles with connection via slips rings in the shaft to an analog to digital converter. Five thermocouples are

arranged radially as shown in Figure 58 to obtain the temperature profile across the gear face. In addition, two additional thermocouples are used to monitor the ambient temperature within the gearbox and the gearbox temperature itself. To create the thioether mist, ultrasonic nozzles (PNR Part number: MAD 0331 B1) were used to dispense the lubricant so that a low flow rate and fine misting can be achieved. The system consists of a liquid pressure tank for the lubricant and requires filtered and dry air pressure supplied through a compressor as shown in Figure 2. In [147], the thioether used is a blend of 4 compounds, (a) 1,1-thiobis [3-phenoxybenzene]; molecular weight, (b) 1-phenoxy-3-[[3-(phenylthio) phenyl[thio]benzene, (c) 1,1-thiobis [3-(phenylthio) benzene] and (d) 1,3,-bis (phenylthio) benzene. In this experiment, the thioether used is Poly(oxy-1,2-ethanediyl), α-butyl-ω-hydroxy-,mixed ethers with 2-ethyl-1-hexanol and 2,2'-thiobis[ethanol] or its product name Vulkanol OT, a commercially available compound.

## 8.5 Experimental Test Plan

The gearbox is run under different conditions of lubrication, torque, speed and rotation direction as shown in Table 3. The rotational direction of the gears used in the experiment is defined in Figure 59. For each test run, the temperatures and time taken for the gear temperatures to stabilize were measured. In addition, the gears were inspected for damage after the run. In this experiment, the gear temperature is considered to have stabilized when the rate of temperature rise is <0.2°C/min (<12°C/hr).

Table 10 Test Runs and Conditions

| Test | RPM | Torque | Lubrication | Mist Rate | Gear Rotation |
|---|---|---|---|---|---|
| 1 | 690 | 100Nm | Oil Dip | NA | CW |
| 2 | 690 | 100Nm | Thioether Mist | 12-15ml/hr | CW |
| 3 | 690 | 100Nm | Thioether Mist | 12-15ml/hr | CCW |
| 4 | 1420 | 100Nm | Thioether Mist | 12-15ml/hr | CCW |
| 5 | 1420 | 100Nm | Oil Dip | NA | CCW |
| 6 | 1420 | 280Nm | Oil Dip | NA | CCW |
| 7 | 1420 | 280Nm | Thioether Mist | 12-15ml/hr | CCW |
| 8 | 1420 | 280Nm | Oil Mist | 12-15ml/hr | CCW |
| 9 | 1420 | 280Nm | Pressurised Air | NA | CCW |

For all the tests with mist lubrication, the liquid pressure and air pressure system was adjusted to deliver thioether or oil at an approximate rate of 12-15ml/hr which is similar to the rate employed in [147]. By comparing Test 1, 4 and 6 against Test 3, 5 and 7, the performance of oil dip lubrication against thioether mist lubrication under increasing speed and torque load conditions are evaluated. When comparing Test 1 with Test 3, it is assumed that gear rotation direction does not affect the temperature profile in oil dip lubrication. Test 2 and 3 compares the effects of gear rotation direction on thioether lubrication. Test 7, 8 and 9 offer comparisons of the performance of thioether mist against oil mist lubrication and pressurized air cooling.

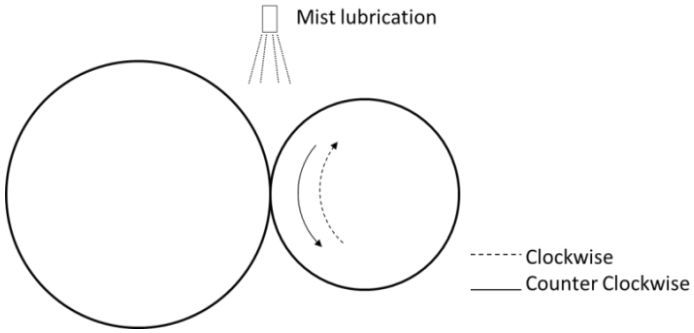


Figure 59 Gear Rotation Direction

## 8.6 Test Results and Discussion

A summary of the test duration and the stabilization temperatures reached in the tests runs are shown in Table 11 and the discussions are as follows.

Table 11 Test Runs Stabilisation Temperature (°C)

| | | Gear Inner | Gear Middle | Gear Outer | Gearbox Air | Gearbox Casing |
|---|---|---|---|---|---|---|
| 1 | 4.90 Hrs | 43.6 | 43.3 | 45.5 | 53.1 | 41.1 |
| 2 | 0.90 Hrs[1] | 55.3 | 56.6 | 64.4 | 43.3 | 27.3 |
| 3 | 2.56 Hrs | 56.0 | 56.8 | 62.6 | 56.5 | 38.4 |
| 4 | 1.54 Hrs | 78.9 | 80.3 | 81.5 | 81.6 | 51.7 |
| 5 | 2.99 Hrs | 59.4 | 59.7 | 59.0 | 68.6 | 48.3 |
| 6 | 1.82 Hrs | 64.3 | 64.8 | 64.5 | 69.7 | 47.2 |
| 7 | 1.50 Hrs | 90.9[2] | 95.0 | 95.9 | 87.5 | 51.6 |
| 8 | 2.37 Hrs | 98.5 | 104.4 | 103.7 | 99.4 | 59.8 |
| 9 | 0.17 Hrs[1] | 71.7 | 119.7 | 121.4 | 63.7 | 22.8 |

[1] Test terminated to prevent gear damage
[2] Last temperature shown after thermocouple dislodged at 1.36Hrs

### 8.6.1 Comparison of thioether mist lubrication against oil dip lubrication

The comparison of the measurement from the thermocouple nearest the gear teeth between thioether against oil lubrication is shown in Figure 61. It should be noted that gear temperatures will rise rapidly in the absence of lubrication as shown in Test 9. It can be seen that the thioether mist lubrication is effective in lubricating and cooling the gears to allow the gear temperatures to stabilize. The stabilization temperatures however are higher compared to oil dip lubrication. From Test Run 1 and 3, the temperature stabilizes at 45.5°C and 62.6°C for the oil-dip and thioether mist respectively. With higher speeds and torque loads in Test run 4, 5, 6 and 7, the stabilization temperatures are higher for thioether mist lubrication reaching 78.9°C and 90.0°C compared to 59.4°C and 64.3°C using oil dip lubrication. The temperature profile normalized to the thioether mist lubrication stabilization temperature is shown in Figure 62. It is shown that the stabilization temperature for oil lubrication is approximately 30% lower compared to thioether mist lubrication. Besides higher stabilization temperature, the stabilization time is longer for thioether mist lubrication as well. As shown in Figure 62, at condition of 690RPM and 100Nm, the thioether mist took 50mins to achieve temperature rate of <0.2°C/min compared to 19mins for oil-dip lubrication. At higher speed of 1420RPM, the thioether mist took 70mins to 53 mins for oil-dip lubrication. At both higher speed and load of 1420RPM and 280Nm, the thioether also took a longer duration of 76mins compared to 45mins using oil lubrication. Inspection of the gears after the tests showed only minor scuffing on the gear teeth surfaces as shown in Figure 63. A brownish lubricious layer of residue is found on the gear teeth surface after the test with thioether mist. Despite the higher stabilization temperature and longer stabilization duration, thioether mist lubrication is shown to be effective as there were no wear on the gears after operating for duration over 30mins.

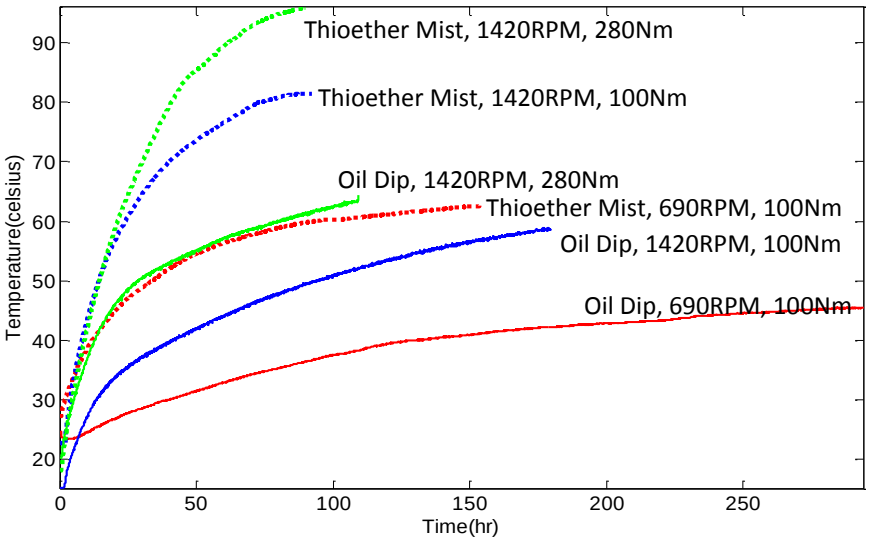


Figure 60 Normalised Temperature profile of oil dip against thioether mist lubrication under different speed and torque load conditions (Gear outer temperature shown)

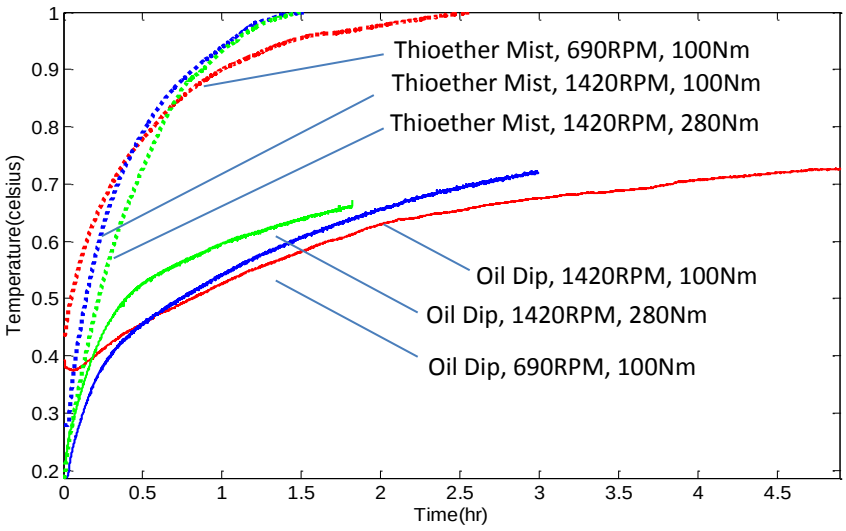


Figure 61 Temperature rate comparison (Gear outer temperature rate shown)

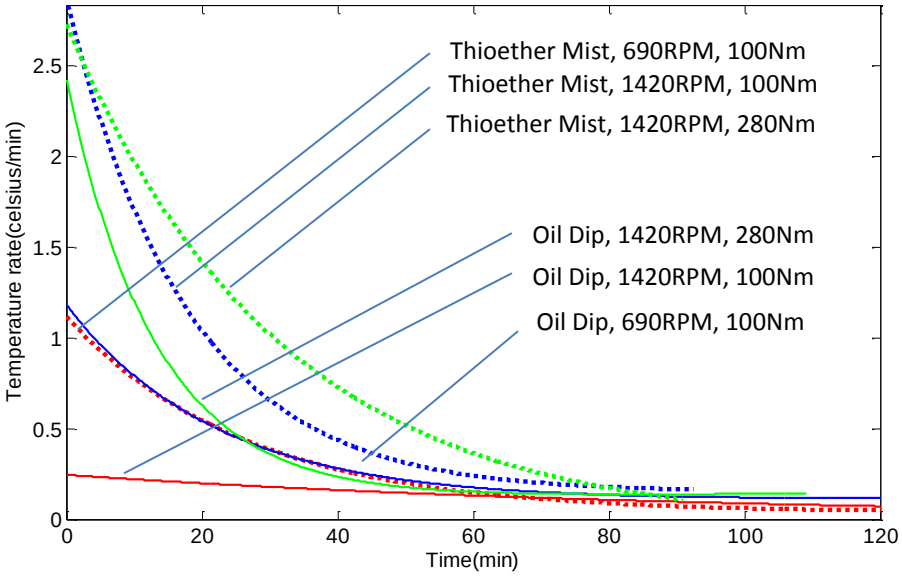


Figure 62 Normalised Temperature profile of oil dip against thioether mist lubrication under different speed and torque load conditions (Gear outer temperature shown)

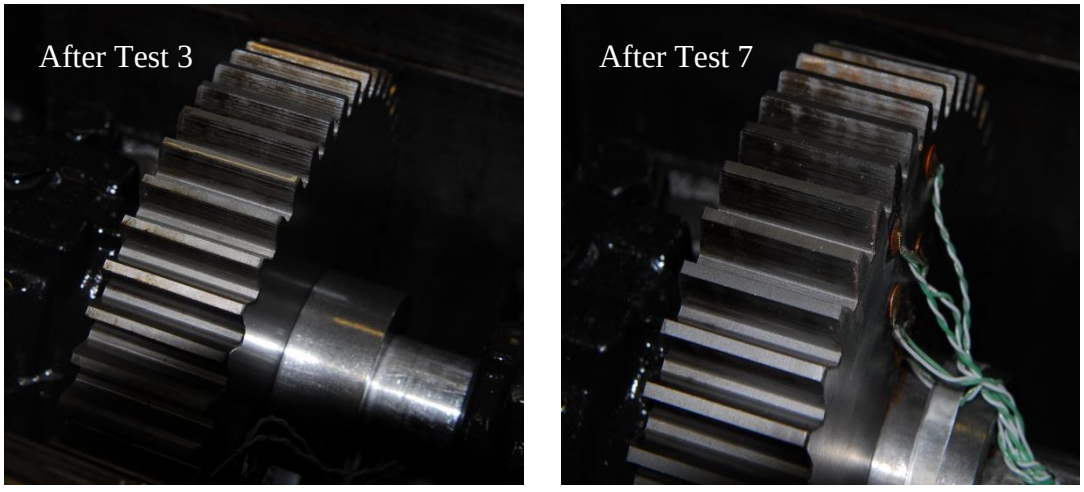


Figure 63 Undamaged gear with lubricious residue on teeth surface after Test 3 and Test 7

### 8.6.2 Comparison of gear rotation direction on thioether lubrication

The effect of gear rotational direction when using thioether lubrication is shown in Figure 64. It is clearly seen that the mist lubrication is not effective when it is sprayed after the gear teeth meshing as the temperature rises sharply and steadily. When the rotational direction is changed, the mist lubrication improves significantly and the

temperature rise rate decreases and stabilizes. As mentioned previously, the poor performance of the mist lubrication when it is sprayed after the gear teeth meshing could be attributed to the fling off of the thioether lubricant as the gear teeth emerges from the mesh. This can be a disadvantage for a mist lubrication based system as the spray nozzle has to be placed before the gear teeth meshes. For main rotor gearboxes which are pre-dominantly planetary gearbox in design, this drawback may not be significant as the mist lubricant can be applied on the static ring gear. However, further work to validate the performance on a planetary gearbox would be required.

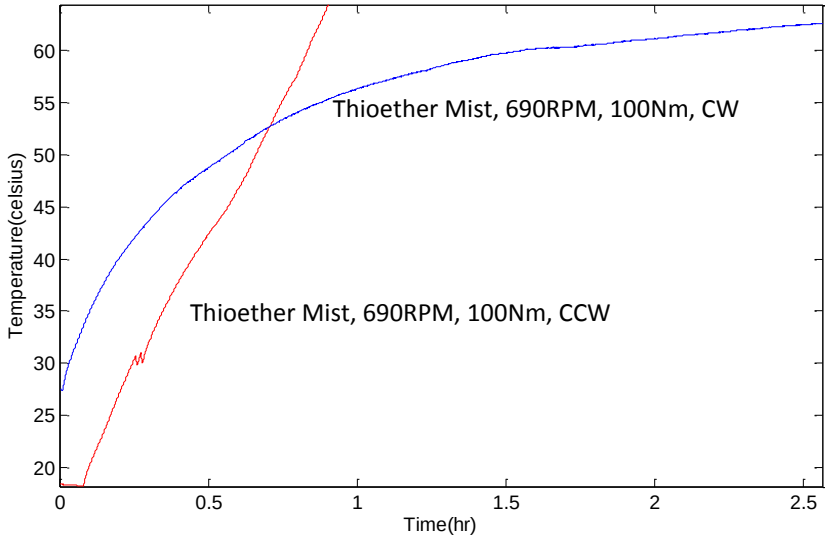


Figure 64 Effect of gear rotation direction on Thioether mist lubrication (Gear outer temperature shown)

### 8.6.3 Comparison of mist lubrication at using oil and thioether and air cooling

The effect of using the Aeroshell 555 lubricating oil as a mist lubricant is also investigated in this study. This is to evaluate if the mist lubrication would be effective if an alternative lubricant is used. Like the thioether, a low flow rate of 12-15ml/hr of oil is used in Test 8 under similar speed and torque load conditions from Test 7. The temperature profile comparison between oil and thioether mist lubrication is shown in Figure 65.

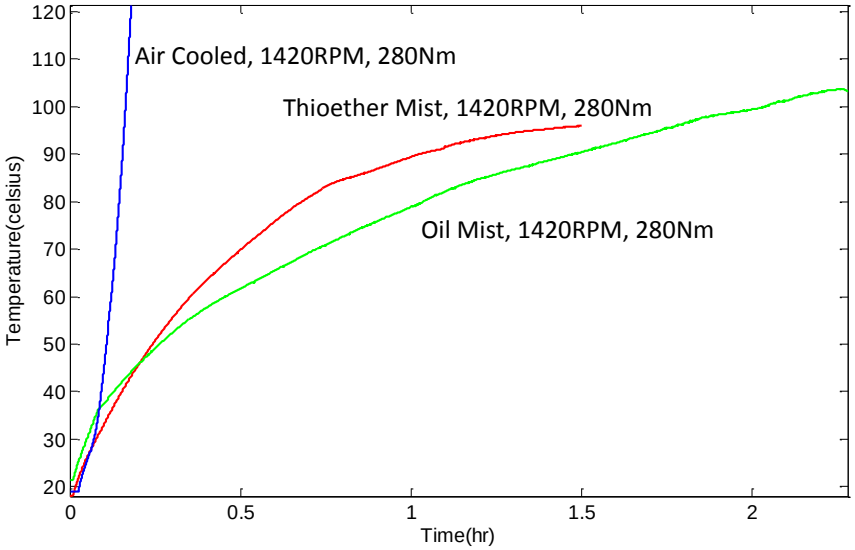


Figure 65 Temperature profile for mist lubrication using oil, thioether and air cooling (Gear outer temperature shown)

It can be seen that the temperature rise rate is slower for oil mist lubrication. However, thioether mist lubrication stabilizes at a lower temperature of 95.9°C compared to 103.7°C for the oil mist. This shows that oil is also a viable candidate for use in mist lubrication. As a comparison, the temperature profile of the gears subjected to only pressurized air cooling is also shown in Figure 65 where it rises very rapidly. For a thirty minutes requirement for the gearbox to operate safely after loss of primary lubrication as per Certification Specification 29.927, the gear mesh in the test requires only approximately 6 – 8ml of oil or thioether lubricant to prevent the gears from overheating and excessive wear.

### 8.6.4 Tests Summary

A summary of the key observations made from the above tests are as follows:

- Thioether mist lubrication does provide adequate lubrication for the gears to achieve stabilization temperatures although the stabilization temperatures reached will be higher compared to oil dip lubrication.
- The rotational direction of the gears has a significant impact on the performance of the mist lubrication. The mist lubrication is not effective when it is not

sprayed into the meshing gear teeth (CCW in the experiment) and this could be due to the thioether being “fling off” the gear surface.

- When oil is applied as a mist-lubrication, the rate of temperature rise of the gears is slower but it stabilizes at temperatures higher compared to thioether mist lubrication.
- The use of mist lubrication can significantly reduce gear wear by providing either thioether or oil lubricant at very low flow rate. This allows the lubricant reservoir of an emergency or backup lubrication system to be kept small.

## 8.7 Chapter Summary

From this study, it can be concluded that thioether-based mist lubrication is comparable to oil based lubrication in a gearbox run-dry situation. From the tests, oil mist lubrication shows a slower rate of temperature rise but it stabilizes at temperatures higher compared to thioether mist lubrication. With the very low lubricant flow rate of a mist lubrication system, the lubricant reservoir can be kept small which is ideal for a backup lubrication system. Notably, this allows a thioether-based backup lubrication to be developed that is independent of the main oil-based lubrication system. They do not share common failures modes and this in turn provides a lower risk of oil starvation failure. The mishap severity category for the lubrication system and associated subsystem such as the accessory gearbox may be relaxed and eases the implementation of PM interval extension or CBM task development. It is noted however that the lubricant to be used is not restricted to thioether however as oil can also be applied with mist lubrication. A drawback of the mist lubrication system is that the rotational direction of the gears has a significant impact on performance. The mist lubrication is not effective when it is sprayed after the gear teeth meshing and may complicate the spray system design.

## 8.8 Acknowledgement

The work performed in this chapter was supported by the European Aviation Safety Agency under Research Project EASA.2011/5 [149].

## 8.9 Related Publications

- *HELMGOP- Helicopter main gearbox loss of oil performance optimisation,* EASA.2011/5, European Aviation Safety Agency.
- *Reliability Assessment of Helicopter Main Gearbox Lubrication Systems and Optimisation in Run-dry Conditions,* 6th Annual Rotorcraft Symposium
- *Feasibility of Thioether based mist lubrication for helicopters*, (Under review), IMechE Part G, Journal of Aerospace Engineering

# 9 CONCLUSION

In the research conducted, methods to improve the maintenance decision support for extending PM task interval and implementing CBM tasks has been developed. Notably, this research has placed much emphasis on issues relevant to practitioners in aircraft maintenance and the methods proposed are meant to be implementable in practice. Towards this end, this research considers the resources available to and constraints faced by maintainers; such as information availability and regulatory requirements. The challenges faced by maintainers are multi-faceted in nature and different methods are proposed to address them accordingly. The outline of the research to improve maintenance decision support is shown in Figure 66.

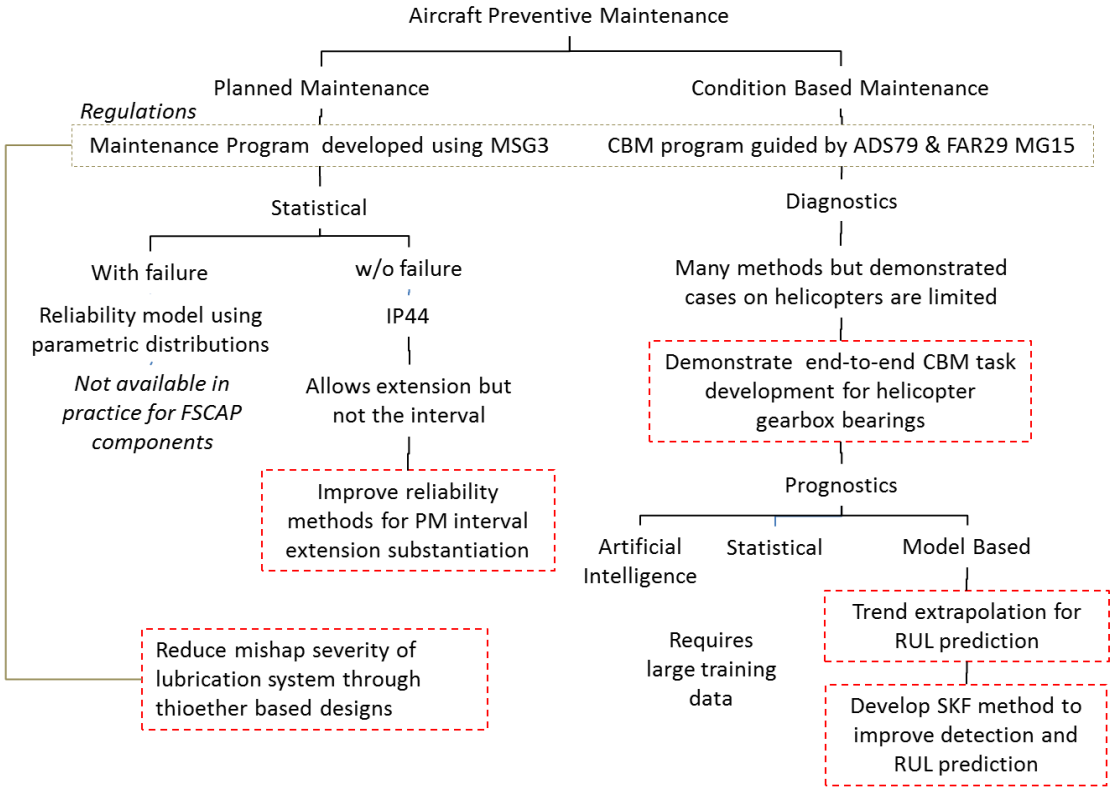


Figure 66 Outline of research to improve maintenance decision support

The key contributions from this research; as described the objectives includes, (1) Improvement of reliability methods for substantiating escalation of planned maintenance interval, (2) Demonstrated an end-to-end development of a CBM task for an in-service helicopter, (3) Improved the ease of use of prognostic tools by adapting

the use of SKF and (4) Evaluated the feasible use of thioether based lubrication system to reduce mishap severity.

For PM, the challenge to extend the maintenance interval with little or no past failure has been addressed. Existing reliability methods were reviewed and two improved methods to estimate the reliability lower confidence bounds were developed. The first approach adopts the use of Monte Carlo simulation applied to the Weibull equation while the second uses a probabilistic damage accumulation model together with bootstrap techniques. Both methods are used to assess the reliability of extending the replacement interval of a gearbox bearing and are shown to perform better than existing methods. Based on the developed methods, maintainers is able to make better use of data from their maintenance reports or usage monitoring system to make better decision when extending PM task interval.

For CBM, a survey on sensor technologies and diagnostic algorithms showed that vibration-based sensors and simple FFT spectra statistics is more widely used in helicopters and has shown successful implementation in the field. However, a comprehensive CBM task development using in-service HUMS data in accordance of regulatory guidance is lacking. The study then demonstrates the development of a CBM task using HUMS data from the RSAF AH64D helicopters TRGB. Analysis of the FFT spectra shows that the fault patterns corresponding to progressing stages of bearing wear can be clearly observed. Notably, it is shown that vibration energy at lower frequency band does not rise monotonously with damage progression and that high-band vibration energy rises following the fall in low-band vibration energy. These fault patterns can be observed consistently across defective bearings from different helicopters and validates observations from past experimental investigations by other authors [111; 112]. The study has used both high and low band energy features independently to track localised and generalised bearing damage. The fault patterns are extracted as features for diagnosis and a prognostic model using trend extrapolation is developed to determine the replacement lead-time for maintenance planning. From this work, an end-to-end CBM task development; for both diagnostic and prognostics, in compliance to the regulatory guidance has been demonstrated and shows the feasibility in field applications.

As prognostic is critical in reducing maintenance effort in CBM, methods to improve prognostics with little or no failure data are further researched. A model-based, Bayesian estimation method, the Switching Kalman Filter (SKF) was adapted for both diagnostic and prognostic under a single framework. The SKF uses multiple dynamic models each describing a different degradation process. The most probable degradation process is then inferred from the extracted feature data using Bayesian estimation. By using the dynamic behavior, pre-established fault detection threshold is no longer required thus less failure data is needed. This approach also provides maintainers with more information for decision-making as a probabilistic measure of the degradation processes are available. This helps maintainers to predict RUL more accurately by distinguishing between the degradation states and performing prediction only when unstable degradation is detected. The use of SKF also allows uncertainty in the damaged progression to be handled and a probability distribution function of the remaining useful life to be obtained. The SKF approach is demonstrated using the in-service feature data and is shown to be a promising tool for maintenance decision-making. Using this approach, prognostics task can be made easier as a significant part of an analyst's decision making process has been automated. A drawback of this approach however, is that the dynamic behavior of the system's degradation process as to be known. It is also not fully autonomous as the noise process has to be tuned manually based on past failures.

As an extension of this research, an effort to reduce the mishap severity category for helicopter lubrication system is explored so that PM interval extension and replacement with CBM task can be implemented more easily. Thioether based mist lubrication is investigated and it was shown from experiments that only a very small volumetric flow of lubricant is required to preserve the gears from damage in oil starved environment. A thioether mist lubrication system can thus be designed to be independent of the main lubrication system as the thioether lubricant does not need to be recycled and does not share the same causal failure modes as the main system.

In this research, the challenges faced by maintainers on implementing PM task extension and CBM task replacement have been investigated and ways to address them has been proposed. It is hopeful that the case studies presented and methodologies

proposed in this research would be useful to practitioners in some ways towards improving helicopter maintenance.

## 10 FUTURE WORKS

In this study, the use of Monte Carlo simulation with Weibull equation and the PDA model using Miner's Rule is applied to evaluate the reliability of extending bearing PM interval. As both Weibull equation and Miner's Rule are commonly adopted in many engineering applications, the methods developed to improve the reliability assessment can be applied to those applications as well. The use of Monte Carlo simulation may be applied together with reliability methods such as Reliability Block Diagrams (RBD) to evaluate the risk of complex systems. Besides its use with strict series reliability model, ways to expand the method to include parallel and independent systems can be investigated. This would allow the method to be applied to more complex systems. For the PDA model, the current work is based on a uniform load. The method can be further improved upon by applying it to systems with variable load conditions which is more reflective of systems in practice. The PDA model allows for the sequence of variable loading to be considered and its performance in modelling the effects would be examined.

The SKF is adapted in this study for diagnostics of the bearing degradation state and estimation of the RUL. The concept however can be applied to other systems to determine their degradation state based on their dynamic behavior. Several rotating machinery components such as gears and shafts may also adopt this method to help maintenance decision-making as most of them exhibits regression trends in their degradation path as well. The SKF approach is not limited to mechanical systems however as it can be used to model the dynamical behavior of control systems as well. In addition, the concept also allows for other recursive Bayesian estimation methods can be explored to be used with the SKF concepts as well. As mentioned before, for applications with highly non-linear behavior or the noise present is non-Gaussian, the EKF may not provide a good estimate. In further development, the use of other non-linear filters such as Unscented Kalman Filter and the Particle Filters would be explored. In addition, the bearing degradation problem described here consists of a

single sensor measurement. In further works, the SKF would be expanded to model systems with multiple sensor measurement inputs and this would allow sensor fusion for better health assessment as well.

For the use of thioether mist design for helicopter backup lubrication system, the initial results are promising although the work is very much in its infancy. Beyond this stage, the thioether mist lubrication should be tested on a helicopter gearbox test rig to determine their performance in a representative environment. Notably, the testing in this work is based on a commercially available thioether compound which is not the same as that used in the experiments by NASA. A comparison of the performance between the compounds and their performance against oil-mist lubrication should be evaluated should the opportunity arises. Besides the performance of the mist lubrication itself, practical aspects of the installation design within a complex gearbox assembly needs to be considered as well. For the backup system to be independent, the spray nozzle and the thioether delivery system has to be separate from the main lubrication system so that there are no common failure modes. The pressure source for the spray mist has to be independent as well. Therefore, these sub-systems have to be integrated within the gearbox and the feasibility of the designs has to be assessed.

**References**


[1] Department of Defense, (2011), *Reliability-Centered Maintenance (RCM) Process*, Standard Practice ed., USA.

[2] Air Transportation Association, (2007), *ATA MSG-3 Operator/ Manufacturer Scheduled Maintenance, Development*, USA.

[3] European Aviation Safety Agency, (2010), *Airworthiness and Environmental Certification - Part 21*.

[4] Department of Defense (2008), *Condition Based Maintenance Plus DoD Guidebook,* , USA.

[5] IMRBPB, (2008), *Evolution / Optimization Guidelines, Issue Paper 44 (Issue 2)*, USA.

[6] European Aviation Safety Agency, (2009), *Continuing Airworthiness Requirements - Part M*, Luxembourg.

[7] UK MOD, (2008), *DEF STAN 00-45 : Reliability Centered Maintenance (Issue 1)*, UK.

[8] NAVAIR (2005), *Guidelines for the Naval Aviation Reliability-Centered Maintenance Process*, NAVAIR 00–25–403, USA.

[9] Federal Aviation Adminstration, (2003), *AC29 MG15 - Airworthiness Approval of Rotorcraft Health Usage Monitoring Systems (HUMS)*, USA.

[10] Army Aviation and Missile Life Cycle Command, ( 2011), *ADS-79B-HDBK (2011), Handbook for Condition Based Maintenance Systems for US Army aircrafts*, USA.

[11] Lau, S., Brisbois, F., Gregoire, J., Hasty, T., Alamond, J., Antolick, L. and Green, D. (2013), *Health and Usage Monitoring Systems Toolkit,* , International Helicopter Safety Team, USA.

[12] Smith, K. (2011), "Health Usage Monitoring System", *AHS/AIAA Specialist's Meeting: Airworthiness, CBM & HUMS*, 7 Feburary, Huntsville, Alabama, AHS, USA, .

[13] Abdel, B., William, R., Lester, E. and Lemuelle, E.,Grant (2006), "On-board vibrations and health monitoring systems: An approach to achieve condition-based maintenance (CBM)", *IMAC-SEM, Journal of the American Helicopter Society,* .

[14] Paula, J., Dempsey, Jonathan, A., Keller and Daniel, R., Wade (2008), *Signal Detection Theory Applied to Helicopter Transmission Diagnostic Thresholds*, NASA/TM—2008-215262, NASA, USA.

[15] Federal Aviation Adminstration (2000), *System Safety Handbook*, , FAA, USA.

[16] Department of Defense (2000), *Standard Practice for System Safety*, MIL-STD-882E, USA.

[17] Baker, C., Marble, S., Morton, B. P. and Smith, B. J. (2007), "Failure Modes And Prognostic Techniques For H-60 Tail Rotor Drive System Bearings", *Aerospace Conference, 2007 IEEE*, pp. 1.

[18] Paula, J.,Dempsey, Jeremy, B., Daniel, R.,Wade and Nathan, B. (2010), "Comparison of Test Stand and Helicopter Oil Cooler Bearing Condition Indicators<br />", *American Helicopter Society 66th Annual Forum*, May 11-13, Phoenix, AZ, USA, .

[19] Viniacourt, F., Yew-Kwong Woon, Ghelam, S., Morel, H., Chan, K. and Shy Chuan (2011), "Operational data Processing and Techniques for Innovative MAINTenance (OPTIMAINT)", *Prognostics and System Health Management Conference (PHM-Shenzhen), 2011*, pp. 1.

[20] Lu, M.,Wei and Wang, C.,Julius (2008), "Weibull Data Analysis with Few or no Failures", in *Recent Advances in Reliability and Quality in Design, Springer Series in Reliability Engineering* , Springer eBook ed, Springer, .

[21] Lundberg, G. and Palmgren, A. (1947), "Dynamic Capacity of Rolling Bearings", *Acta Polytech., Mechanical Engineering Series*, vol. 1, no. 3.

[22] Harris, T.,A. and Kotzalas, M.,N. (2006), "Fatigue Life: Basic Theory and Rating Standards", in *Essential Concepts of Bearing Technology, Fifth Edition*, fifth ed, CRC Press, Boca Raton, FL, pp. 195.

[23] Vlcek, B.,L., Hendricks, R. C. and Zaretsky, E.,V. (2004), *Determination of Rolling Element Fatigue Life from Computer Generated Bearing Tests*, NASA TM 2003-212186, NASA, USA.

[24] Liao Min, Xu Xiaofei and Yang Qing-Xiong (1995), "Cumulative fatigue damage dynamic interference statistical model", *International Journal of Fatigue*, vol. 17, no. 8, pp. 559-566.

[25] Sehgal, R., Gandhi, O. P. and Angra, S. (2000), "Reliability evaluation and selection of rolling element bearings", *Reliability Engineering & System Safety*, vol. 68, no. 1, pp. 39-52.

[26] Rathod, V., Yadav, O.,Prakash, Rathore, A. and Jain, R. (2011), "Probabilistic Modeling of Fatigue Damage Accumulation for Reliability Prediction", *International Journal of Quality, Statistics, and Reliability*, vol. 2011.

[27] Harris, T.,A. and McCool, J.,I. (1996), "On the Accuracy of Rolling Bearing Fatigue Life Prediction", *Journal of Tribology*, vol. 118, no. 2, pp. 297-309.

[28] Harris, T.,A. (1995), *Establishment of a New Rolling Bearing Contact Life Calculation Method*, N68335-93-C-0111, U.S. Naval Air Warfare Center, Aircraft Division Trenton, Trenton.

[29] Tallian, T. (1962), "Weibull Distribution of Rolling Contact Fatigue Life and Deviations Therefrom", *Weibull Distribution of Rolling Contact Fatigue Life and Deviations Therefrom*, vol. 5, no. 1, pp. 183-196.

[30] Elfron, B. and Tibshirani, R. (1994), *<br />An Introduction to the Bootstrap*, Chapman & Hall, London.

[31] Rosado, L., Forster, N.,H., Thompson, K.,L. and Cooke, J.,W. (2009), Rolling Contact Fatigue Life and Spall Propagation of AISI M50, M50NiL, and AISI 52100, Part I: Experimental Results", *Tribology Transactions*, vol. 53, no. 1, pp. 29.

[32] Zaretsky, E.,V, Lewicki, D.,G., Savage, M. and Vlcek, B.,L (2007), "Determination of Turboprop Reduction Gearbox System Fatigue Life and Reliability", *STLE Tribology Transaction*, vol. 50, no. 4, pp. 507.

[33] BS ISO 13372:2012 (2004), *Condition monitoring and diagnostics of machines — Vocabulary*, British Standards.

[34] Brian, D., Larder (1999), "Helicopter HUM/FDR: Benefits and Developments", *Annual forum; 55th, American Helicopter Society*, American Helicopter Society, pp. 1839.

[35] Safety Regulation Group, Civil Aviation Authority, (2012), *CAP 753 - Helicopter Vibration Health Monitoring (VHM)*, UK.

[36] Joint Industry Proposal (Airbus, Boeing, Bombardier, Embraer, Gulfstream) , *Further advanced definition of Structural Health Monitoring (SHM)/Addition to MSG-3*, available at: http://easa.europa.eu/certification/products/docs/imrbpb-issue-papers/IP%20105.pdf (accessed Nov/11).

[37] ISO 13374-1:2003, (2003), *Condition monitoring and diagnostics of machines — Data processing, communication and<br />presentation — Part 1: General guidelines*.

[38] Larry, M., Barbara, M., Jeff, F. and Diane, W. (1991), *Rotorcraft Health and Usage Monitoring Systems - A Literature Survey*, FAA/RD-91/6, FAA, USA.

[39] James, J.,Zakrajsek, Paula, J.,Dempsey, Edward, M.,Huff, Michael, A., Robab, S., Alan, D., Vermont, P.,Ephraim, Paul, G. and Harry, J.,Decker (2006), *Rotorcraft Health Management Issues and Challenges*, NASA/TM—2006-214022, NASA, USA.

[40] Irebert, R.,Delgado, Paula, J.,Dempsey and Donald, L.,Simon (2012), *A Survey of Current Rotorcraft Propulsion Health Monitoring Technologies*, NASA/TM—2012-217420, NASA, USA.

[41] Raylund, R., Harold, S. and James, C. (1996), *Feasibility Study of a Rotorcraft Health and Usage Monitoring System (HUMS): Results of Operator's Evaluation*, DOT/FAA/AR-95/50, NASA, USA.

[42] Paula, J.,Dempsey, David, G.,Lewicki and Dy, D.,Le (2007), *Investigation of Current Methods to Identify Helicopter Gear Health*, 20070018745, Glenn Research Center, USA.

[43] S, S., B, D.,Larder and R, E.,Brown (2008), *HUMS Extension to Rotor Health Monitoring*, CAA Paper 2008/05, CAA, UK.

[44] B, H. and D, M. (1998), *In-line oil debris monitor (ODM) for helicopter gearbox condition assessment*, 19980624-0079, USA.

[45] Miller, J. L. and Kitaljevich, D. (2000), "In-line oil debris monitor for aircraft engine condition assessment", *Aerospace Conference Proceedings, 2000 IEEE*, Vol. 6, pp. 49.

[46] Paula, J.,Dempsey, David, G.,Lewicki and Harry, J., Decker (2004), *Investigation of Gear and Bearing Fatigue Damage Using Debris Particle Distributions*, NASA/TM—2004-212883, NASA, USA.

[47] Joint Oil Analysis Program (2012), *Joint Oil Analysis Program Manual*, NAVAIR 17-15-50.1, U.S Navy, USA.

[48] Steven, W., Arms, Christopher, P., Townsend1, Jacob, H., Galbreath, Stephen, J., DiStasi, Daniel, L. and Nam, P. (2011), "Flight Testing of Wireless Sensing Networks for Rotorcraft Structural Health and Usage Management Systems", *7th DSTO International Conference on Health & Usage Monitoring*, Feb 28 - Mar 3, Melbourne, Australia, .

[49] Jason, S. K., Paul, D., Samuel, David, G., Lewicki, Kelsen, E., LaBerge, Ryan, T., Ehinger and Jason, F. (2011), *Fiber Optic Strain Sensor for Planetary Gear Diagnostics*, NASA/TM—2011-217123, NASA, USA.

[50] Neil, G., Yanko, S., Todd, D., Scott, D., David, G., Darrell, S. and Vladimir, Z. (2008), "Magnetic Stress Gages for Torque and Load Monitoring in Rotorcraft", *Presented at the American Helicopter Society (AHS)*, April 29-May 1, Montreal, Canada, USA, .

[51] Matthias, B. and Partha, P.,Adhikari (2012), "Simulation Framework and Certification Guidance for Condition Monitoring and Prognostic Health Management", *European Conference of Prognostics and Health Management Society 2012,* Sep 23 - 27, Minneapolis, Minnesota, USA, .

[52] Lewis, W. (2011), "ADS 79-B Handbook for CBM Systems on US Army Aircraft", *AHS Airworthiness, CBM and HUMS Specialist Meeting,* 7 Feburary, Redstone, AHS, .

[53] Jardine, A. K. S., Lin, D. and Banjevic, D. (2006), "A review on machinery diagnostics and prognostics implementing condition-based maintenance", *Mechanical Systems and Signal Processing,* vol. 20, no. 7, pp. 1483-1510.

[54] P., V., M., K. and R., Š (2005), "Condition Indicators for Gearbox Condition Monitoring Systems", *Acta Polytechnica Vol. 45,* vol. 45, no. 6, pp. 35-43.

[55] Kurfess, T., Billington, S. and Liang, S. (2006), "Advanced Diagnostic and Prognostic Techniques for Rolling Element Bearings", in Wang, L. and Gao, R. (eds.) Springer London, , pp. 137-165.

[56] Baillie, D. C. and Mathew, J. (1996), "A comparison of autoregressive modeling techqniues for fault diagnosis of rolling element bearings", *Mechanical Systems and Signal Processing,* vol. 10, no. 1, pp. 1-17.

[57] Randall, R. B. (2011), *Vibration-based condition monitoring : industrial, aerospace, and automotive applications,* Wiley, Chichester.

[58] Uluyol, O., Kim, K. and Hickenbottom, C. (2011), "Update to a Systematic Approach to Bearing Health Monitoring", *7th DSTO International Conference on Health & Usage Monitoring,* 28th Feb - 3rd Mar, Melbourne, Australia, .

[59] Dykas, B., Krantz, T.,L., Decker, H. and Lewicki, D.,G (2009), *Experimental Determination of AH64 Apache Tailshaft Hanger Bearing Vibration Characteristics with Seeded Faults,* ARL-TR-4865, Army Research Laboratory, USA.

[60] Mendel, J. M. (1991), "Tutorial on higher-order statistics (spectra) in signal processing and system theory: theoretical results and some applications", *Proceedings of the IEEE,* vol. 79, no. 3, pp. 278-305.

[61] Guoji, S., McLaughlin, S., Yongcheng, X. and White, P. (2014), "Theoretical and experimental analysis of bispectrum of vibration signals for fault diagnosis of gears", *Mechanical Systems and Signal Processing,* vol. 43, no. 1–2, pp. 76-89.

[62] Liu Yibing, Zhou Yanbing, Xin Weidong, He Ying and Fan Pengqi (2010), "Bispectrum analysis for feature extraction of pitting fault in wind turbine gearbox", *Mechatronics and Automation (ICMA), 2010 International Conference on,* pp. 1298.

[63] Byungchul Jang, Changyong Shin, Powers, E. J. and Grady, W. M. (2004), "Machine fault detection using bicoherence spectra", *Instrumentation and Measurement Technology Conference, 2004. IMTC 04. Proceedings of the 21st IEEE*, Vol. 3, pp. 1661.

[64] Cocconcelli, M., Zimroz, R., Rubini, R. and Bartelmus, W. (2012), "STFT Based Approach for Ball Bearing Fault Detection in a Varying Speed Motor", in Fakhfakh, T., Bartelmus, W., Chaari, F., et al (eds.) Springer Berlin Heidelberg, , pp. 41-50.

[65] Zhou, Y., Chen, J., Dong, G. M., Xiao, W. B. and Wang, Z. Y. (2011), "Wigner–Ville distribution based on cyclic spectral density and the application in rolling element bearings diagnosis", *Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science*, vol. 225, no. 12, pp. 2831-2847.

[66] Jing Yu, Yip, L. and Makis, V. (2010), "Wavelet analysis with time-synchronous averaging of planetary gearbox vibration data for fault detection, diagnostics, and condition based maintenance", *Mechanical and Electronics Engineering (ICMEE), 2010 2nd International Conference on*, Vol. 1, pp. V1-132.

[67] Huang, N. E., Shen, Z., Long, S. R., Wu, M. C., Shih, H. H., Zheng, Q., Yen, N. C., Tung, C. C. and Liu, H. H. (1998), "The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis", *Proceedings of the Royal Society of London.Series A: Mathematical, Physical and Engineering Sciences*, vol. 454, no. 1971, pp. 903-995.

[68] Pan, M.,Chun and Tsao, W.,Chang (2013), "Using Appropriate IMFs for Envelope Analysis in Multiple Fault Diagnosis of Ball Bearings", *International Journal of Mechanical Sciences*, .

[69] Rai, V. K. and Mohanty, A. R. (2007), "Bearing fault diagnosis using FFT of intrinsic mode functions in Hilbert–Huang transform", *Mechanical Systems and Signal Processing*, vol. 21, no. 6, pp. 2607-2615.

[70] Junsheng, C., Dejie, Y. and Yu, Y. (2006), "Research on the intrinsic mode function (IMF) criterion in EMD method", *Mechanical Systems and Signal Processing*, vol. 20, no. 4, pp. 817-824.

[71] Wikipedia (2013), *Receiver Operating Characteristic*, available at: http://en.wikipedia.org/wiki/File:Receiver_Operating_Characteristic.png (accessed 5/19).

[72] Dempsey, P.,J. and Afjeh, A.,A. (2002), *Integrating Oil Debris and Vibration Gear Damage Detection Technologies Using Fuzzy Logic*, NASA/TM—2002-211126, NASA, USA.

[73] Gang, N., Jong-Duk, S., Achmad, W., Bo-Suk, Y., Don-Ha, H. and Dong-Sik, K. (2007), "A Comparison of Classifier Performance for Fault Diagnosis of Induction Motor using Multi-type Signals", *Structural Health Monitoring*, vol. 6, pp. 215-229.

[74] Mendel, E., Rauber, T.,W., Varej, F.,M. and Batista, R.,J. (2009), "Rolling Element Bearing Fault Diagnosis in rotating machines of oil extraction rigs", *17th European Signal Processing Conference*, August 24-28, Glasgow, .

[75] Siegel, D., Ly, C. and Lee, J. (2012), "Methodology and Framework for Predicting Helicopter Rolling Element Bearing Failure", *Reliability, IEEE Transactions on*, vol. 61, no. 4, pp. 846-857.

[76] Yu, G., Li, C. and Sun, J. (2010), "Machine fault diagnosis based on Gaussian mixture model and its application", *The International Journal of Advanced Manufacturing Technology*, vol. 48, no. 1-4, pp. 205-212.

[77] McCune, B., Grace, B.,James and Dean, L. (2002), "Distance Measures", in *Analysis of Ecological Communities*, MjM Software Design ed, MjM Software Design, USA, pp. 45.

[78] Manning, D.,Christopher, Raghavan, P. and Schütze, H. (04/07/2009), *Hierarchical clustering*, available at: http://nlp.stanford.edu/IR-book/html/htmledition/hierarchical-clustering-1.html (accessed June 2013).

[79] MacKay,J.,C.,David (2003), "Probabilities and Inference", in *Information Theory, Inference, and Learning Algorithms*, 4th ed, Cambridge University Press, UK, pp. 281.

[80] D Roulias and T H Loutas and,V.Kostopoulos (2012), "A hybrid prognostic model for multistep ahead prediction of machine condition", *Journal of Physics: Conference Series*, vol. 364, no. 1, pp. 012081.

[81] Shahapurkar, S. S. and Sundareshan, M. K. (2004), "Comparison of self-organizing map with K-means hierarchical clustering for bioinformatics applications", *Neural Networks, 2004. Proceedings. 2004 IEEE International Joint Conference on*, Vol. 2, pp. 1221.

[82] M Pirra and E Gandino and A Torri and L Garibaldi and,J.M.Machorro (2011), "PCA algorithm for detection, localisation and evolution of damages in gearbox bearings", *Journal of Physics: Conference Series*, vol. 305, no. 1, pp. 012019.

[83] Zanoli, S. M. and Astolfi, G. (2012), "Application of a mahalanobis-based pattern recognition technique for fault diagnosis on a chemical process", *Control & Automation (MED), 2012 20th Mediterranean Conference on*, pp. 1347.

[84] Heng, A., Zhang, S., Tan, A. C. C. and Mathew, J. (2009), "Rotating machinery prognostics: State of the art, challenges and opportunities", *Mechanical Systems and Signal Processing*, vol. 23, no. 3, pp. 724-739.

[85] Sikorska, J. Z., Hodkiewicz, M. and Ma, L. (2011), "Prognostic modelling options for remaining useful life estimation by industry", *Mechanical Systems and Signal Processing*, vol. 25, no. 5, pp. 1803-1836.

[86] Gorjian, N., Ma, L., Mittinty, M., Yarlagadda, P. and Sun, Y. (2009), "A review on degradation models in reliability analysis", *Proceedings of the 4th World Congress on Engineering Asset Management*, Athens, Springer.

[87] Si, X., Wang, W., Hu, C. and Zhou, D. (2011), "Remaining useful life estimation – A review on the statistical data driven approaches", *European Journal of Operational Research*, vol. 213, no. 1, pp. 1-14.

[88] Department of Defense, (2005), *Aircraft Structural Integrity Program (ASIP)*, Standard Practice ed., USA.

[89] Marble, S. and Morton, B. P. (2006), "Predicting the remaining life of propulsion system bearings", *Aerospace Conference, 2006 IEEE*, pp. 8 pp.

[90] Gebraeel, N. (2006), "Sensory-Updated Residual Life Distributions for Components With Exponential Degradation Patterns", *Automation Science and Engineering, IEEE Transactions on*, vol. 3, no. 4, pp. 382-393.

[91] Li, Y., Billington, S., Zhang, C., Kurfess, T., Danyluk, S. and Liang, S. (1999), "Adaptive prognostics for rolling element bearing condition", *Mechanical Systems and Signal Processing*, vol. 13, no. 1, pp. 103-113.

[92] Gebraeel, N., Lawley, M., Liu, R. and Parmeshwaran, V. (2004), "Residual life predictions from vibration-based degradation signals: a neural network approach", *Industrial Electronics, IEEE Transactions on*, vol. 51, no. 3, pp. 694-700.

[93] Shao, Y. and Nezu, K. (2000), "Prognosis of remaining bearing life using neural networks", *Proceedings of the Institution of Mechanical Engineers, Part I: Journal of Systems and Control Engineering*, vol. 214, no. 3, pp. 217.

[94] Chuang Sun and Zhousuo Zhang and,Zhengjia He (2011), "Research on bearing life prediction based on support vector machine and its application", *Journal of Physics: Conference Series*, vol. 305, no. 1, pp. 012028.

[95] Rabiner, L. (1989), "A tutorial on hidden Markov models and selected applications in speech recognition", *Proceedings of the IEEE*, vol. 77, no. 2, pp. 257-286.

[96] Xing-Hui Zhang and Jian-She Kang (2010), "Hidden Markov models in bearing fault diagnosis and prognosis", *Computational Intelligence and Natural Computing Proceedings (CINC), 2010 Second International Conference on*, Vol. 2, pp. 364.

[97] Camci, F. and Chinnam, R. B. (2010), "Health-State Estimation and Prognostics in Machining Processes", *Automation Science and Engineering, IEEE Transactions on*, vol. 7, no. 3, pp. 581-597.

[98] Usynin, V.,Alexander (2007), *A Generic Prognostic Framework for Remaining Useful Life Prediction of Complex Engineering Systems* (PhD thesis), University of Tennessee - Knoxville, USA.

[99] ISO 13381-1:2004, ( 2006), *Condition monitoring and diagnostics of machines — Prognostics*.

[100] Jouni, H., Arno, S. and Simo, S. (2011), *Optimal Filtering with Kalman Filters and Smoothers*, Department of Biomedical Engineering and Computational Science, Aalto University School of Science, Finland.

[101] Lall, P., Lowe, R. and Goebel, K. (2010), "Prognostics using Kalman-Filter models and metrics for risk assessment in BGAs under shock and vibration loads", *Electronic Components and Technology Conference (ECTC), 2010 Proceedings 60th*, pp. 889.

[102] Lall, P., Lowe, R. and Goebel, K. (2012), "Extended Kalman Filter Models and Resistance Spectroscopy for Prognostication and Health Monitoring of Leadfree Electronics Under Vibration", *Reliability, IEEE Transactions on*, vol. 61, no. 4, pp. 858-871.

[103] Jose, R.,Celaya, Chetan, K., Gautam, B., Sankalita, S. and Kai, G. (2011), "A Model-based Prognostics Methodology for Electrolytic Capacitors Based on Electrical Overstress Accelerated Aging", *Annual Conference of the Prognostics and Health Management Society 2011*, Vol. 2, .

[104] Zio, E. and Peloni, G. (2011), "Particle filtering prognostic estimation of the remaining useful life of nonlinear components", *Reliability Engineering & System Safety*, vol. 96, no. 3, pp. 403-409.

[105] Nicholas, G., Abdel, B., Vytautas, B., Ronak, S. and Yong-June, S. (2009), "CBM Component Testing at the University of South Carolina: AH64 Tail Rotor Gearbox Studies", *American Helicopter Society Technical Specialists' Meeting on Condition Based Maintenance*, Feb 10-11, Huntsville, AL, USA, .

[106] Blunt, D. M. and Keller, J. A. (2006), "Detection of a fatigue crack in a UH-60A planet gear carrier using vibration analysis", *Mechanical Systems and Signal Processing*, vol. 20, no. 8, pp. 2095-2111.

[107] Harry, J.,Decker and David, G.,Lewicki (2003), *Spiral Bevel Pinion Crack Detection in a Helicopter Gearbox*, NASA/TM—2003-212327, Glenn research Center, USA.

[108] Jonathan, A.,Keller and Paul, G. (2005), "Inserted Fault Vibration Monitoring Tests for a CH-47D Aft Swashplate Bearing", *American Helicopter Society 61st Annual Forum*, Jun 1 – 3, Grapevine, TX, American Helicopter Society International, USA, .

[109] JTEKT Corporation (2009), *Ball & Roller Bearings: Failures, Causes and Countermeasures*, B3001 E P, Koyo, Japan.

[110] BS ISO 15243:2004 (2004), *Rolling bearings — Damage and failures — Terms, characteristics and causes*", *British Standards Institution*, .

[111] STI Field Application Note (2012), *Rolling Element Bearings*, REB, Sales Technology, Inc., League City TX.

[112] Qiu, H., Luo, H. and Eklund, N. (2009), *On-board bearing prognostics in aircraft engine: Enveloping analysis or FFT?* AFRL-RX- WP - TP -2009-4141, Air Force Research Laboratory, USA.

[113] M., E. and D., M. "Assessment of natural crack initiation and its propagation in slow speed bearings", *Nondestructive Testing and Evaluation*, vol. 24, no. 3, pp. 261-275.

[114] Williams, T., Ribadeneira, X., Billington, S. and Kurfess, T. (2001), "Rolling Element Bearing Diagnostics in Run-to-Failure Lifetime Testing", *Mechanical Systems and Signal Processing*, vol. 15, no. 5, pp. 979-993.

[115] Kotzalas, M. N. and Harris, T. A. (2001), "Fatigue Failure Progression in Ball Bearings", *Journal of Tribology*, vol. 123, no. 2, pp. 238-242.

[116] Paula, J., Dempsey, Nathan, B., Chris, H., Jeremy, B. and Daniel, R., Wade (2010), "Correlate Life Predictions and Condition Indicators in Helicopter Tail Gearbox Bearings", *American Helicopter Society 66th Annual Forum*, May 11-13, Phoenix, AZ, USA, .

[117] LaPaugh, A. (2008), *Clustering Algorithms:K-means*, available at: https://www.cs.princeton.edu/courses/archive/spring08/cos435/Class_notes/clustering2_toPost.pdf (accessed 05/23).

[118] Moore, A. (2004), *Clustering with Gaussian Mixtures*, available at: http://www.autonlab.org/tutorials/gmm14.pdf (accessed 05/23).

[119] Gottschalk, P. G. and Dunn, J. R. (2005), "The five-parameter logistic: A characterization and comparison with the four-parameter logistic", *Analytical Biochemistry*, vol. 343, no. 1, pp. 54-65.

[120] Särkkä, S. (2011), "*Optimal Filtering*", in *Bayesian Estimation of Time-Varying Systems: Discrete-Time Systems*, , pp. 36-38.

[121] Zarchan, P. and Musoff, H. (2005), "Polynomial Kalman Filters", in *Fundamentals of Kalman Filtering: A Practical Approach*, 2nd ed, American Institute of Aeronautics and Astronautics, USA, pp. 156.

[122] Simon, D. (2008), "A comparison of filtering approaches for aircraft engine health estimation", *Aerospace Science and Technology*, vol. 12, no. 4, pp. 276-284.

[123] Christoph, A., Robert, S. and Uwe, K. (2012), "Unscented Kalman Filter with Gaussian Process Degradation Model for Bearing Fault Prognosis", *European Conference of the Prognostics and Health Management Society*, Vol. 3, July 3-5 2012, Dresden, Germany, .

[124] Lall, P., Junchao Wei and Goebel, K. (2012), "Comparison of Lalman-filter and extended Kalman-filter for prognostics health management of electronics", *Thermal and Thermomechanical Phenomena in Electronic Systems (ITherm), 2012 13th IEEE Intersociety Conference on*, pp. 1281.

[125] Jose, R.,Celaya, Abhinav, S. and Kai, G. (2012), "Uncertainty Representation and Interpretation in Model-based Prognostics Algorithms based on Kalman Filter Estimation", *Annual Conference of the Prognostics and Health Management Society 2012*, Vol. 3, Minneapolis, USA, .

[126] Kevin, M. (1998), *Learning Switching Kalman Filter Models*, 98-10, Compaq Cambridge Research Lab Tech Report.

[127] Manfredi, V., Mahadevan, S. and Kurose, J. (2005), "Switching kalman filters for prediction and tracking in an adaptive meteorological sensing network", *Sensor and Ad Hoc Communications and Networks, 2005. IEEE SECON 2005. 2005 Second Annual IEEE Communications Society Conference on*, pp. 197.

[128] Lim, Y. and Cheng, S. (2012), "Knowledge-driven autonomous commodity trading advisor", *2012 IEEE/WIC/ACM International Conference on Intelligent Agent Technology*, Macau, .

[129] Youmin Zhang and Jin Jiang (2001), "Integrated active fault-tolerant control using IMM approach", *Aerospace and Electronic Systems, IEEE Transactions on*, vol. 37, no. 4, pp. 1221-1235.

[130] Xudong Wang and Syrmos, V. L. (2008), "Interacting multiple particle filters for fault diagnosis of non-linear stochastic systems", *American Control Conference, 2008*, pp. 4274.

[131] John, Q. , *SKF Demo*, available at: http://www.cit.mak.ac.ug/staff/jquinn/software.html (accessed October/21).

[132] Boyen, X. and Koller, D. (1998), "Tractable inference for complex stochastic processes", *Proceedings of the Fourteenth conference on Uncertainty in artificial intelligence*, Jul-98, Madison, Wisconsin, USA, pp. 33.

[133] Sutrisno, E., Oh, H., Vasan, A. S. S. and Pecht, M. (2012), "Estimation of remaining useful life of ball bearings using data driven methodologies", *Prognostics and Health Management (PHM), 2012 IEEE Conference on*, pp. 1.

[134] H, P.,Ritzema (1994), "Frequency and Regression Analysis. Chapter 6", in *Drainage Principles and Applications*, International Institute for Land Reclamation and Improvement (ILRI), Wageningen, The Netherlands, pp. 175-224.

[135] Bar-Shalom, Y., Li, X. and Kirubarajan, T. (2001), "Estimation with applications to tracking and navigation", in Wiley, New York, pp. 441-459.

[136] C, J.,Lu and W, Q.,Meeker (1993), "Using Degradation Measures to Estimate a Time-to Failure Distribution", *Technometrics*, vol. 35, no. 2, pp. 161-173.

[137] Saxena, A., Celaya, J., Balaban, E., Goebel, K., Saha, B., Saha, S. and Schwabacher, M. (2008), "Metrics for evaluating performance of prognostic techniques", *Prognostics and Health Management, 2008. PHM 2008. International Conference on*, pp. 1.

[138] Saxena, A., Celaya, J., Saha, B., Saha, S. and Goebel, K. (2009), "On Applying the Prognostic Performance Metrics", *International Conference on Prognostics and Health Management (PHM)*, .

[139] Astridge, D.,G (1992), "Safety Assessment of Transmission Health and Usage Monitoring Systems", *Safety Assessment Methods for Transmission Systems*, 1 Apr, Solihill, West Midlands, IMechE, UK, .

[140] Transportation Safety Board of Canada (2013), *Aviation Investigation Report A09A0016*, available at: http://www.bst-tsb.gc.ca/eng/rapports-reports/aviation/2009/a09a0016/a09a0016.asp (accessed Feb/22).

[141] UK Air Accidents Investigation Branch (2004), *G-TAMA*, EW/C2008/02/04, AAIB, UK.

[142] Jarvis, M.,P and Sleight, P. (2011), *2/2011 Aerospatiale (Eurocopter) AS332 L2 Super Puma, G-REDL*, AAR/2011, UK AAIB, UK.

[143] UK Air Accidents Investigation Branch (2012), *EC225 LP Super Puma, G-REDW*, S2/2012, AAIB, UK.

[144] Eurocopter *AS332 Maintenenace Manual*, .

[145] UK MOD "Seaking Maintenenace Manual", .

[146] Agusta Westland *WAH-64D Maintenance Manual*, .

[147] Morales, W., Handschuh, R.,F. and Krantz, T.,L. (2007), *Feasibility Study of Vapor-Mist Phase Reaction Lubrication Using A Thioether Liquid*, NASA/TM—2007-215035, NASA, USA.

[148] Mcgrogan, F. (1976), *Oil-air mist lubrication for helicopter gearing*, NASA-18538, NASA, USA.

[149] Mba, D., Place, S., Rashid, H. and Lim, R. (2012), *HELMGOP- Helicopter main gearbox loss of oil performance optimisation*, EASA.2011/5, European Aviation Safety Agency.