



UNIVERSITY OF PATRAS
MECHANICAL ENGINEERING AND AERONAUTICS
DEPARTMENT

**Methodologies for Remaining Useful Life
Estimation with Multiple Sensors in Rotating
Machinery**

PhD dissertation

Author: Dimitrios Roulias

Patras, June 2014



European Union
European Social Fund

This research has been co-financed by the European Union (European Social Fund – ESF) and Greek national funds through the Operational Program "Education and Lifelong Learning" of the National Strategic Reference Framework (NSRF) - Research Funding Program: Heracleitus II. Investing in knowledge society through the European Social Fund.



Ευρωπαϊκή Ένωση
Ευρωπαϊκό Κοινωνικό Ταμείο



Με τη συγχρηματοδότηση της Ελλάδας και της Ευρωπαϊκής Ένωσης



UNIVERSITY OF
PATRAS

University of Patras

Mechanical and Aeronautics Engineering Department

[This page intentionally left blank]

University of Patras
Department of Mechanical Engineering and Aeronautics
Division of Applied Mechanics, technology of materials and Biomechanics

Patras, 14/07/2014

To the Department of Mechanical Engineering and Aeronautics

Subject: Approval of PhD Thesis of Mr. Dimitrios ROULIAS.

The 7-member examination board for the judgment of PhD Thesis of **Mr. ROULIAS Dimitrios** under the title: **"Methodologies for remaining useful life estimation with multiple sensors in rotating machinery"**, which consists of: Professor V. Kostopoulos, Department of Mechanical Engineering and Aeronautics University of Patras, Professor S. Fassois, Department of Mechanical Engineering and Aeronautics University of Patras, Professor David Mba, Chair in Rotating Machines Technology, Energy and Power Division, School of Engineering, Cranfield University, Lecturer TH. Loutas, Department of Mechanical Engineering and Aeronautics University of Patras, Professor A. Dentsoras, Department of Mechanical Engineering and Aeronautics University of Patras, Professor P. Groumpos, Department of Electrical and Computer Engineering University of Patras, Professor N. Fakotakis, Department of Electrical and Computer Engineering University of Patras, during its today assembly has concluded *unanimously* that the subject of his PhD Thesis is prototype, original and of substantial contribution to science and has given the grade of *excellent*.

THE EXAMINATION BOARD

V. KOSTOPOULOS

S. FASSOIS

D. MBA

T. LOUTAS

A. DENTSORAS

P. GROUMPOS

N. FAKOTAKIS



Handwritten signatures of the examination board members in blue ink, positioned over a dotted line. The signatures are: V. Kostopoulos, S. Fassois, D. Mba, T. Loutas, A. Dentsoras, P. Groumpos, and N. Fakotakis.

[This page intentionally left blank]

*To my parents,
Panagiotis and Evaggelia*

Abstract

The focus of this thesis was the development of failure prognosis methods (prognostics) in rotating machinery with use of multiple sensors digital signal processing and machine learning techniques. The motivation stems from the void in literature concerning prognostics in meshing gearboxes. Moreover, there are several but inconclusive works regarding bearing prognosis.

Few research groups have studied multi-hour fatigue gear experiments and this was one of the contributions of this thesis. Moreover, the study expanded beyond the sheer application of vibration monitoring with the addition of an Oil Debris Monitoring probe (ODM) as well as Acoustic emission (AE).

The method of AE monitoring is, once again, proposed as a robust technique for failure prognosis being better correlated with gear pitting level compared to the classic vibration monitoring technique. Moreover, judging from ODM recordings the gear pitting comprises of two phases i) a linear phase, with an almost constant pitting rate and ii) a very short non linear phase where the pitting rate increases exponentially, an explicit indication of a critical failure.

Multi-hour gear experiments that are close to real scale applications are very demanding in time as well as in invested capital. To bypass this shortfall a gear failure like simulation is proposed. The simulation framework is based on real life experiments and is applied to assess a number of data-driven Remaining Useful Life (RUL) estimation techniques namely i) Proportional Hazards Model (PHM), ii) ϵ - Support Vector Regression ϵ -SVR and iii) Exponential extrapolation based on bootstrap sampling.

In the current thesis a feature extraction scheme for prognosis is proposed and assessed based on time domain, frequency domain statistical features and Wavelet Packet (WP) energy derived from AE and vibration recordings. ICA is proposed as a preferable fusion technique for gear failure prognostics. Application of ICA for feature fusion provided a clear improvement regarding the earlier presented bootstrap extrapolation technique.

Bearings are also taken into account since they are closely connected to gearboxes. In the current thesis a wavelet denoising method is proposed for bearing vibration recordings aiming to the improvement of the diagnostic and prognostic potential of vibration. Finally the importance of data fusion is highlighted in the case of bearings. It is observed that a feature extraction scheme can generalize the application of prognostics, even in cases where RMS may yield no important degradation trend.

Περίληψη στα Ελληνικά

Η παρούσα εργασία εστιάζεται στην ανάπτυξη μεθοδολογιών πρόβλεψης τελικής αστοχίας σε περιστρεφόμενα συστήματα με χρήση πολλαπλών αισθητήρων και μεθόδων μηχανικής μάθησης και επεξεργασίας σήματος. Το κίνητρο προήλθε από το κενό που υπάρχει στη βιβλιογραφία όσον αφορά την προγνωστική σε κιβώτια ταχυτήτων. Η προγνωστική σε έδρανα έχει μεν μελετηθεί αλλά σε μικρό βαθμό και η παρούσα εργασία έρχεται να συμβάλλει και σε αυτό τον τομέα.

Στα πλαίσια αυτής της εργασίας εκπονήθηκε ένας αριθμός πειραμάτων κόπωσης κιβωτίων ταχυτήτων. Η μελέτη επεκτάθηκε πέραν της παρακολούθησης κατάστασης με τη μέθοδο των κραδασμών και συγκεκριμένα μελετήθηκαν καταγραφές σωματιδίων σιδήρου στο λιπαντικό (ODM) καθώς και Ακουστική Εκπομπής (AE). Η μέθοδος AE ευρέθη πιο στενά συσχετισμένη με τη σταδιακή υποβάθμιση της ακεραιότητας του κιβωτίου ταχυτήτων σε σχέση με τις καταγραφές κραδασμών. Επίσης με βάση τις καταγραφές του αισθητήρα σωματιδίων σιδήρου διακρίθηκαν δύο στάδια υποβάθμισης i) μια γραμμική περιοχή με σχεδόν σταθερό ρυθμό απελευθέρωσης υλικού από την επιφάνεια των δοντιών και ii) μια σύντομη αλλά έντονα μη γραμμική αύξηση στο ρυθμό αυτό πολύ κοντά στο τέλος της λειτουργίας του κιβωτίου.

Τα πολύωρα πειράματα κόπωσης σε γρανάζια είναι πολύ απαιτητικά. Για να παρακαμφθεί αυτή η δυσκολία αναπτύχθηκε ένα φαινομενολογικό μοντέλο για αναπαραγωγή χρονοσειρών που ομοιάζουν σε καταγραφές γραναζιών σε κόπωση. Το μοντέλο αυτό στηρίχθηκε σε πραγματικά πειράματα κόπωσης. Έτσι έγινε δυνατό να εξεταστούν και να συγκριθούν ένας αριθμός μεθοδολογιών εκτίμησης εναπομένουσας ζωής και συγκεκριμένα i) Proportional Hazards Model (PHM), ii) ε- Support Vector Regression ε-SVR και iii) Exponential extrapolation βασισμένο σε μια διαδικασία bootstrap sampling.

Στην παρούσα μελέτη προτείνεται ένα σύνολο παραμέτρων προερχόμενο από το πεδίο της συχνότητας, του χρόνου και κυματοπακέτων. Αυτό, συνδυαζόμενο με μια διαδικασία σύμπτυξης δεδομένων (ανάλυση σε πρωταρχικές και ανεξάρτητες συνιστώσες) αξιοποιείται για πρόγνωση σε γρανάζια σε κόπωση. Η τεχνική ανεξάρτητων συνιστωσών προτείνεται σαν προτιμότερη από τη σκοπιά της προγνωστικής καθώς βελτιώνει την εκτίμηση της εναπομένουσας ζωής.

Η εργασία επεκτάθηκε και σε έδρανα κύλισης. Προτάθηκε μια διαδικασία wavelet denoising η οποία ενισχύει τόσο τη διαγνωστική όσο και την προγνωστική δυνατότητα του αισθητήρα κραδασμών. Τέλος, η σημασία της εξαγωγής παραμέτρων υπογραμμίζεται και στην περίπτωση της προγνωστικής σε έδρανα. Συνδυάζοντας πολλαπλές παραμέτρους και

αισθητήρες κραδασμών μαζί με ένα μοντέλο ϵ -SVR παρέχεται ένα ολοκληρωμένο μοντέλο πιθανοτικής εκτίμησης εναπομένουσας ζωής σε έδρανα κύλισης ακόμα και σε περιπτώσεις που η τιμή RMS των κραδασμών δεν παρέχει πληροφορία.

Acknowledgements

The present thesis was carried out in the Applied Mechanics Laboratory, Mechanical Engineering & Aeronautics Department, University of Patras during 2008-2014. It has been co-financed by the European Union (European Social Fund – ESF) and Greek national funds through the Operational Program "Education and Lifelong Learning" of the National Strategic Reference Framework (NSRF) - Research Funding Program: Heracleitus II. Investing in knowledge society through the European Social Fund.

Now having completed this thesis I would like to sincerely thank:

My supervisor *Professor Vassilis Kostopoulos* for the collaboration, his guidance and the many opportunities he has offered me during my stay in Applied Mechanics Laboratory,

Professor Spilios Fassois for his guidance and cooperation during this thesis. Also for his excellent undergraduate lecture series of “Stochastic signals and Systems” from which I was greatly benefited. I was honored to have him in my supervisory committee,

Professor David Mba for his guidance and cooperation during this thesis. I was honored to have him in my supervisory committee,

Professor Nikos Fakotakis, Petros Groumpos and Argyris Dentsoras for their fruitful comments and for partaking in the examination board. I was truly honored,

Lecturer Professor Theodoros Loutas for his guidance, cooperation and constant motivation throughout this thesis. Also for partaking in the examination board.

I would also like to thank *Nikos* for his friendship, scientific and moral support all these years. I would also like to thank all the staff in AML for their friendship and support in all necessities of this thesis through these years and particularly for their help with the experimental part of this thesis.

List of Tables

Name	Page
Table 3.1 : Research based on simulated failure data.	43
Table 4.1 . Experimental configurations.	49
Table 4.2 . Gear failure experiments.	66
Table 4.3 . Selected model for AE-ODM functional relation.	75
Table 5.1 . Simulated Gear failure experiments.	88
Table 6.1 . WP bandwidth.	109
Table 6.2 . Extracted features.	109
Table 6.3 . Selected prognostic features.	110
Table 6.4 . Prognostic performance for fused health metrics.	116
Table 7.1 . Bearing failure data bases.	122
Table 7.2 . Operating conditions.	124
Table 7.3 . DWT level and corresponding performance.	132
Table 7.4 . WP bandwidth.	136
Table 7.5 . Selected model parameters.	138

List of Figures

Name	Page
FIG 1.1. Crude oil production level in the last decade.	2
FIG.1.2. All fossil fuel production profiles.	2
FIG 1.3. Graphical representation of the evolution in maintenance policies.	4
FIG. 1.4. Graphical representation of CBM health monitoring.	5
FIG. 1.5. Representation of a condition monitoring scheme.	6
FIG. 2.1. Functional layers in the CBM system architecture.	11
FIG. 2.2. WT sub-system failure rate and downtime per failure for two surveys including over 20000 turbine years of data.	13
FIG. 2.3. Comparison of failure contribution of WT generator components taken from Alewine (2010).	14
FIG. 2.4. Frequency of various failure root causes as a percentage.	15
FIG. 2.5. BPFO = ball pass frequency, outer race, BRFI =ball pass frequency, inner race, BSF = ball spin frequency; FTF = fundamental train frequency (cage frequency).	16
FIG. 2.6. On the left a velocity probe cross section. A piezoelectric accelerometer is presented on the right.	17
FIG. 2.7. A graphic representation of the vibration response formation in a meshing machine.	18
FIG. 2.8. A slight deviation in shaft rotating frequency (upper graph) may completely mask the underlying gear fault signature (Jafarizadeh et al. (2008)).	20
FIG. 2.9. Acoustic emission of continuous type (a) obtained from acoustic emission sensor mounted on a gear box case. Acoustic emission of burst type (b) coming from composite material. The material was tested for fatigue. The vertical axis is in Volts, the horizontal axis is the time steps.	24
FIG. 2.10. A cross section of an ODM (left) and principle of operation (right).	26
FIG. 2.11. Principle of operation for a closed loop Hall effect electric current sensor.	28
FIG. 3.1. The problem of prognosis.	32
FIG. 3.2. A training testing scheme for an on-line prognosis model.	33
FIG. 3.3. A typical vibration RMS signature for a gear failure history taken from Gasperin et al. (2011a). His predictions are accurate after 3/4 of the gear life. In fig. 3.3 a linear and a short, non-linear, degradation phase is denoted by us. In Chapter 4 gear failure experiments will exhibit a similar pattern.	36
FIG. 4.1. Photographs of the gearbox fatigue test rig. On the right, a close up of the	47

master/slave gear pair.

FIG. 4.2. Vibration level (left) for different set up configurations plotted along with lubricating oil temperature (right). The recordings are obtained from vibration channel 1 (x-axis vibration).	49
FIG. 4.3. AE levels (left) for different configurations plotted along with lubricating oil temperature (right).	50
FIG. 4.4. The aggregated mass of ferrous particles and its rate of particles production for lifelong experiment III.	53
FIG. 4.5. The aggregated mass of ferrous particles and its rate of particles production for lifelong experiment VI.	54
FIG. 4.6. Sensor measurements from a gear lifelong test with artificially induced gear cracks.	55
FIG. 4.7. AE P-P amplitude for experiment III (left) and experiment VI (right).	56
FIG. 4.8. AE RMS for experiment III (left) and experiment VI (right).	57
FIG. 4.9. AE Waveforms for lifelong experiment III (upper plots) and their respective PSDs (lower plots). The waveforms are taken from 10%, 60% and 90% of operational time.	58
FIG. 4.10. AE Waveforms for lifelong experiment VI and their respective PSDs. The waveforms are taken from the 10%, 60% and 90% of operational time.	59
FIG. 4.11. AE P-P amplitude for experiment III (left) and experiment V (right).	60
FIG. 4.12. AE RMS for experiment III (left) and experiment V (right).	60
FIG. 4.13. Typical AE channel 2 waveform (top) and its corresponding PSD (bottom) from experiment V.	61
FIG. 4.14. AE P-P amplitude (on the left) and RMS (on the right) for experiment III.	62
FIG. 4.15. Typical AE waveform (upper plots) and its corresponding PSD (lower plots) corresponding to experiment III.	63
FIG. 4.16. P-P amplitude for experiment III (left) and experiment VI (right).	64
FIG. 4.17. RMS for experiment III (left) and VI (right).	64
FIG. 4.18. P-P amplitude for experiment III (left) and VI (right).	65
FIG. 4.19. RMS for experiment III (left) and VI (right).	65
FIG. 4.20. P-P amplitude (left) and RMS (right) for vibration z-axis sensor, experiment III.	66
FIG. 4.21. The data series on the left comprises missing values. On the right a remedy is applied.	70
FIG. 4.22. Simulation results for AE-ODM relation.	75
FIG. 4.23. The basic idea behind the synthesis of lifelong RMS histories.	76

FIG. 4.24. Three synthesized Fe-rate histories.	79
FIG. 4.25. Four synthesized failure histories.	80
FIG. 5.1. The flowchart depicts the general gear failure prognosis framework of this chapter. This chart is motivated by the findings of Chapter 4.	84
FIG. 5.2. Test results regarding hazard rate and reliability function on three out of sample failure histories.	91
FIG. 5.3. Two RUL prediction experiments. Upper graph corresponds to realisation 16 and lower graph to realisation 17, table 5.1. Actual time of failure is 169 and 192 time samples for the upper and lower graph respectively.	92
FIG 5.4. On the left the significance of support vectors is denoted regarding a linear regression task. On the right the significance of the ϵ -Loss function.	95
FIG. 5.5. Mean square error (MSE) (left) and cross correlation coefficient (CCC) (right) for various values of ϵ and C model parameters. The exact values of the respective indexes are shown on the previous page.	98
FIG. 5.6. Two prediction results with the probabilistic ϵ -SVR technique. The triangular-marker line corresponds to the predictions. The 95% confidence bounds (upper and lower) are also depicted. The straight line corresponds to the ARUL for comparison. The corresponding MSE and CCC scores are also denoted on the graph.	100
FIG 5.7. Diagram for the bootstrap sampling process.	101
FIG. 5.8. Visualization of the interpolation process.	102
FIG. 5.9. Predicted time of failure for two test gear failure histories.	104
FIG. 6.1. Figure a) is derived from vibration ch1 (experiment V), b) vibration ch3 (experiment V), c) vibration ch1 (experiment VI), d) vibration ch2 (experiment VI), e) AE ch1 (experiment III), f) vibration ch2 (experiment III).	111
FIG. 6.2. A 2-D data set (large circles) is projected on the derived maximum variance directions denoted with two vector arrows. Their projections are denoted with the smaller dots.	113
FIG. 6.3. Reduced health monitoring metrics via PCA.	115
FIG. 6.4. Reduced health monitoring metrics via ICA.	116
FIG. 6.5. Quantitative estimation of “Time of failure” along with appropriate error bounds via a boxplot.	117
FIG. 6.6. The proposed data fusion with probabilistic extrapolation prognosis scheme.	118
FIG. 7.1. RMS of four bearings being tested simultaneously under accelerated fatigue	

conditions. In this test bearing 1 (top-most curve) was found severely damaged in the final visual inspection. The RMS recordings of bearings 2, 3 and 4 are much correlated indicating that the corresponding recordings were severely affected from bearing 1 failure vibrations. This experiment provided only 1 valid failure history. 123

[Fig. 7.2](#): Photos of degraded bearings from [Nectoux et al. \(2012\)](#). 124

[FIG. 7.3](#). Typical RMS timeseries for two lifelong bearing experiments. a) and c) RMS histories from case 1 and 2 bearings respectively (see [table 7.2](#)), b) and d) zoom in the change-point due to critical fault occurrence. 125

[FIG. 7.4](#). Two “irregular” RMS histories derived from Case 1, ([fig. 7.4a](#)) and Case 2 ([fig. 7.4b](#)) with reference to [table 7.2](#). On the left, a sudden increase in RMS at ~270mins was not proven detrimental to the bearing. In fact the particular fault seems to wear out with the bearing revolution. On the right, few increasing and decreasing trends make any effort of diagnosis or prognosis practically impossible. 127

[FIG. 7.5](#). A graphic representation of DWT. 128

[FIG. 7.6](#). On the left, the original RMS from three bearing failure histories and on the right the corresponding RMS after the denoising of vibration waveforms. 130

[FIG. 7.7](#). Upper figures (a, b) correspond to original RMS from a bearing failure history, b) being a zoom in axis of a). Lower figures (c, d) correspond to WD processed versions of the latter. 131

[FIG. 7.8](#). Flow chart of the proposed preprocessing technique. 132

[FIG. 7.9](#). MSE (right) and CCC (left) for various values of ϵ and C model parameters. The exact values of the respective indexes are shown above. 133

[FIG. 7.10](#). Dash dot straight line is the actual RUL of the testing component and it is plotted for comparison. MSE and CCC are computed on the original (scaled) output and the normalised actual RUL. Figure a) corresponds to the original waveforms and b) to WD-preprocessed waveforms. 134

[FIG. 7.11](#). Spectral flatness as a function of operational time for three different lifelong rolling element bearing tests. 137

[FIG. 7.12](#). Prediction results for “case 1” data base (a) and “case 2” database (b). 95% confidence intervals ($\pm 1.69\sigma$) are added on each prediction curve. The straight line corresponds to the actual remaining life for each component. 139

[FIG. 7.13](#). Flow chart of the proposed multi-parametric bearing RUL estimation scheme which is described in [Section 7.4](#). 140

Table of Contents

Title	Page
Abstract.....	i
Abstract in Greek.....	iii
Acknowledgements.....	v
List of Tables.....	vii
List of figures.....	viii
Table of contents.....	xii
1. Introduction.....	1
1.1 Background and motivation.....	1
1.2 Maintenance strategies.....	3
1.3 Scope of this research and objectives.....	6
1.4 Important contributions of this thesis.....	9
2. Overview of Bearings and Gearboxes Monitoring Techniques.....	10
2.1 Introduction.....	10
2.2 Importance of gearboxes and bearings.....	12
2.3 Gearbox modes of failure.....	15
2.4 Bearings modes of failure.....	16
2.5 Monitoring techniques.....	17
2.5.1 Vibration.....	17
2.5.2 Acoustic emission.....	23
2.5.3 Oil Debris Monitoring.....	26
2.5.4 Motor Current Signature.....	28
3. Remaining Useful Life Estimation in Engineering Systems.....	29
3.1 Terminology.....	29
3.2 Literature review.....	33
3.3 Discussion.....	42

4. Gear failure experiments	45
4.1 Introduction.....	45
4.2 Assumptions.....	46
4.3 Experimental Set up.....	47
4.4 Effect of contextual parameters in the sensor recordings.....	48
4.5 Lifelong gear fatigue tests.....	51
4.5.1 ODM recordings.....	52
4.5.2 Acoustic emission recordings.....	56
4.5.3 Vibration recordings.....	63
4.5.4 Discussion on the experiments.....	66
4.6. Preprocessing of the health metrics.....	67
4.7 Correlations between sensors.....	70
4.8 Synthesized health indexes.....	76
4.9 Conclusion.....	78
5. Estimating the Remaining Useful Life in Gearboxes	83
5.1 Introduction.....	83
5.2 Proportional hazards model.....	84
5.2.1 Theory.....	84
5.2.2 Application in gear failure data.....	88
5.3 Probabilistic ϵ -SVR.....	93
5.3.1 Theory.....	93
5.3.2 Application in gear failure data	97
5.4 Probabilistic exponential extrapolation scheme.....	101
5.4.1 Application on gear failure data.....	103
5.5 Discussion.....	105
6. Estimating the Remaining Useful Life in Gearboxes with feature fusion	107
6.1 Introduction.....	107
6.2 Feature extraction.....	108
6.2.1 Results.....	115
6.3 Conclusion.....	118

7. Remaining useful life estimation of bearings.....	121
7.1 Introduction.....	121
7.2 Bearing failure data bases.....	122
7.3 Mono-parametric bearing failure prediction.....	126
7.3.1 Application of ϵ -SV regression.....	132
7.4 Poly-parametric bearing failure prediction.....	135
7.4.1 Feature extraction.....	135
7.4.2 Prediction results.....	138
7.5 Conclusion.....	140
8. Conclusion and Closure.....	141
Appendix A.....	144
Appendix B.....	149
References.....	152
List of publications.....	161
Short C.V.....	163

"Essentially, all models are wrong, but some are useful"

George E. P. Box

CHAPTER 1

INTRODUCTION

1.1 Background and motivation

September 5th 2013 had nothing uncommon regarding the events that took place that day. Taking a look at the world news one could find sad events such as a tragic collapse in a Bangladesh clothing-factory, few protests across Europe regarding educational reforms, a big military parade in Moscow and the news regarding the price of oil. At that day crude-oil had never been more expensive¹, ever. The price of oil is controlled of course by a number of economic and political parameters. Among others, the market mechanisms of supply and demand and various competitions between oil producing countries. However, those factors have always existed throughout the 20th and the beginning of 21th century.

Liquid fuel production the last decade has never been higher in history. This can be attributed to a number of reasons such as the introduction of new oil extracting technologies

¹ <http://www.zerohedge.com/news/2013-09-05/day-history-crude-oil-has-never-been-more-expensive>

(fracking) which have been around for a while but have become an economically profitable technology for the first time, again due to the high price of liquid fuels. The second alarming characteristic of the fuel production curve is its trend in this past decade. It doesn't exist. Although technology of 21st century pushes fuel production to levels unseen in history, the production seems stagnant for a decade and fluctuates around a mean value (fig. 1.1)².

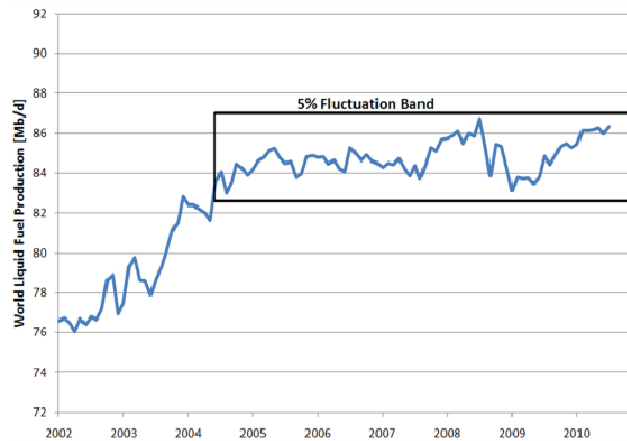


FIG 1.1. Crude oil production level in the last decade.

It may be that we are past the peak historic oil production. This fact is also highlighted in the controversial Hubbert curve (Fig. 1.2) (Hubbert (1956)) first presented by Shell geologist M. King Hubbert in 1956. This symmetric logistic distribution curve predicts depletion in natural resources and has been confirmed in the case of USA drilled liquid oil production which actually peaked in 1970.

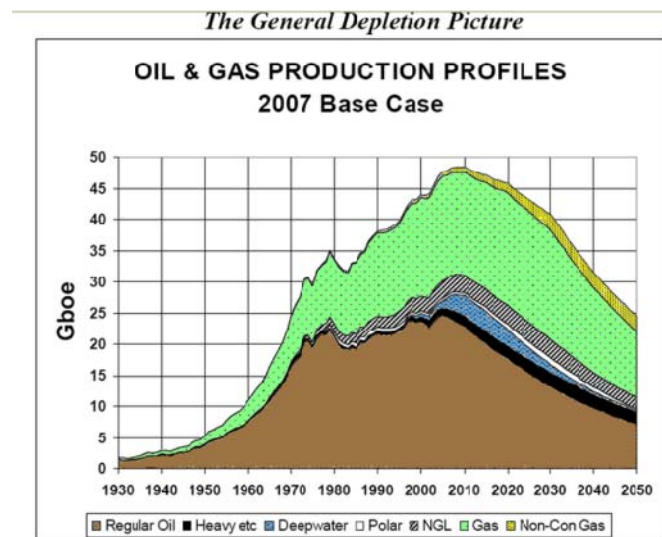


FIG. 1.2 All fossil fuel production profiles.

The latter evidence is indicative of a situation where new fossil fuels are being difficult to discover. “Easy” fuel sources have been depleted or are very close to depletion. The Deep

² <http://www.zerohedge.com/news/guest-post-global-oil-risks-early-21st-century>

Water Horizon incident in 2010 tragically demonstrated that most of the oil left is deep offshore or in any case very difficult to reach even with 21st century state-of-the art drilling technology with a high risk. High end extracting technology is itself very costly in capital and energy contributing to the final product cost level (Kothamasu et al. (2006)).

We began our overview with liquid fuel extraction because in modern world fossil fuel is life. It is transportation and medicine and consumer products and everything. The cost of reaching to underground oil reserves is easily transformed to the cost of everyday living in our modernized world. At the heart of modernized societies lies the constant need for natural resources which are not limited to liquid fuel. Important minerals such as copper and iron and even gold and silver are also becoming harder and costlier to extract and this doesn't appear to be a secluded incident but rather a constant condition. It is the view of the author of this thesis that in a world of hashier natural resources societies and must become more economic and more efficient.

Maintenance in industry is a key factor for efficiency and economy. In fact it is reported that 15-40% of manufacturing costs are due to maintenance actions. This is a vast economic amount. It is more than certain that improvements in maintenance policies can have a great positive impact in economic industrial operation and this is also stated in the relative literature (Byington and Garga (2001), Huynh et al. (2012)). More efficient and economic maintenance actions can result in lower operational costs, higher availability of the industrial equipment and optimized operational time intervals between replacement actions. Timely maintenance actions can contribute to avoiding catastrophic events that may compromise a whole line of production and induce high or unbearable costs. Companies themselves are becoming more concerned of the higher operational cost combined with higher consumer expectations of the 21st century. Needless to say that an efficient reliability framework and early equipment flaw diagnosis can increase personnel safety and save lives in critical modern industrial applications such as nuclear industry or deep off shore oil drilling.

1.2 Maintenance strategies

In this sub-section a brief overview of the existing maintenance strategies will be given. We are particularly interested in the currently applied maintenance strategies and in the way those strategies can be improved towards higher efficiency. A good point of reference will be a diagram³ that pictures the various strategies in terms of sophistication and abstract term "efficiency".

As depicted in [fig 1.3](#), the most naive strategy is "don't fix". This trivial non-maintenance strategy is applied in cheap everyday equipment. For instance, if a lawnmower

³ <http://www.plant-maintenance.com/articles/ConMon21stCentury.shtml>

breaks down you don't fix it, you buy a new. The first level of maintenance strategy is the reactive. A maintenance action takes place only after a failure event. This is a frequent action in mechanical equipment such as car vehicles. It is frequently met in literature as corrective maintenance. The third strategy shown in the figure is planned or preventive maintenance. This kind of maintenance is applied in critical mechanical systems where a fatal flaw may have unpleasant consequences.

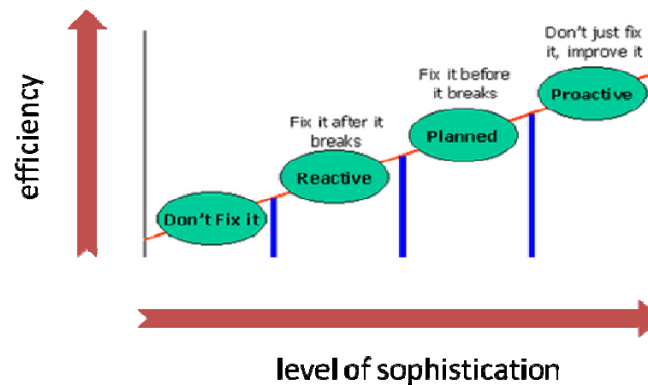


FIG 1.3. Graphical representation of the evolution in maintenance policies.

In this case the decision for maintenance can be taken based on various criteria.

- Physics based criteria:

A physical model can be applied to find the failure threshold of a mechanical structure or system. This can be done numerically with the help of commercial FEM programs such as ANSYS. Other physics based laws for RUL estimation are the L10 law and Paris-Erdogan law which can be combined with FEM analysis to provide a first result. It is the view of the author that these models have limited application particular in rotating machinery operating under (more or less) constant conditions.

- Reliability or aging base criteria

Reliability criteria consider data from past failures from other similar components or of the same component being under repair actions. These can give a good estimation in cases where the stochasticity is limited or the structure is working under static conditions. Several models of reliability have been proposed in literature regarding mechanical systems with most important the Weibull process. A review of such models can be found in [Hokstad \(1997\)](#). The disadvantage of these models is that they don't account for any change in operating parameters of the particular asset being either external (load, rotating speed) or internal (appearance of an unprecedented fault). Therefore it may result in excessive maintenance actions. Excessive maintenance actions of certain machinery components can, themselves, have an impact on the

overall machinery condition. Moreover, increased stopping times in industrial equipment can result in significant economic losses. Therefore reliability models are nowadays deemed insufficient for a variety of critical mechanical applications.

- Condition based monitoring

This is a relatively new area in maintenance policies. In general, this strategy assumes an operating parameter which is very well correlated to the actual degradation level of the particular asset. The asset is considered to have failed when this parameter crosses a predefined threshold as shown in [fig 1.4](#).

The advantage of CBM is that it is application specific, in contrast to reliability decision making. The issue here is to find a) a reliable condition parameter and b) a reliable “failure” threshold.

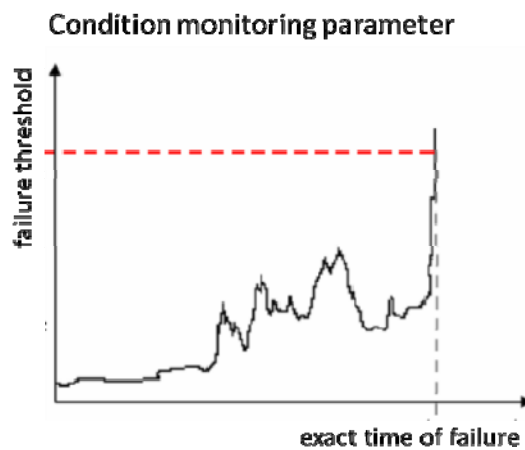


FIG. 1.4. Graphical representation of CBM health monitoring.

Finally a statistical method that combines reliability and condition based indexes has been proposed by Cox ([Cox \(1972\)](#)) and has been applied in few cases for RUL estimation.

The last two decades we see industry to shift towards CBM solutions. For instance, OLF (the Norwegian Oil Industry Association) has stated that “one key function in cutting costs is optimization of maintenance efforts. The future of maintenance management will be to turn away from offshore-stationed personnel carrying out periodically predetermined work schedules, or unplanned maintenance tasks when equipment fails, towards heavy use of condition-based maintenance (CBM)” ([Oljeindustriens Landsforening \(2005\)](#)). Diagnosis and prognosis are the two main aspects of CBM.

Diagnosis is the set of actions related to indentifying a particular flaw and its cause in an operating mechanical component. Diagnosis is closely related to non destructive testing, their main difference being the on-line application of the former within a CBM framework. Diagnosis of a structural flaw may or may not lead to an immediate shut down and depends

on the application and experience. In other words, in a mechanical system the identification of a flaw may not mean its failure. It should also be noted that not all single components have similar flaw signatures. Moreover, particular components may have more than one modes of failure thus contributing to the complexity of its diagnosis.

Prognosis on the other hand is the knowledge of future condition of a particular asset. The European standard on maintenance terminology does not define “prognosis”. This term doesn’t appear in International Federation of Automatic Control (IFAC) keyword list. This is evidence that the area is quite new compared to machine diagnostics.

Machine prognosis is performed through sensor signals monitoring and therefore it is closely related to diagnosis. The final outcome of prognosis is the knowledge of the component’s remaining operational life. Remaining useful life or time to failure is the time interval between the present inspection of the machines operating parameters and the point of failure. RUL is probably the most important outcome of the CBM strategy.

The question “when will it break” is central in maintenance and it is frequently accompanied with “how confident am I” regarding this prediction.

Both diagnosis and prognosis contribute to a robust CBM scheme. The focus of this research will be on RUL estimation along with confidence intervals for mechanical applications and to a lesser extend to diagnosis of particular failure modes.

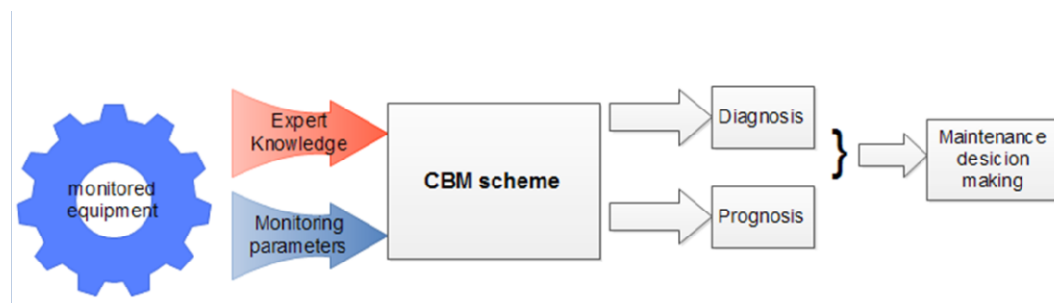


FIG. 1.5. Representation of a condition monitoring scheme.

1.4 Scope of this research and objectives

The main objective of this research is the development of remaining useful life estimation schemes for rotating machinery. The study will focus on gearbox and bearing RUL estimation. It is natural to study both equipment since you can rarely have the one (gearbox) without the other (rolling bearings) in any existing rotating mechanical system. The RUL estimation of these equipment is a challenging task since they exhibit various modes of failure. RUL estimation of these components is important because they are at the heart of a variety of industrial applications.

The objective of this thesis is fulfilled taking certain steps.

a. A review in the modes of failure of those two components is essential in developing a RUL estimations scheme. Unless the degradation mechanisms are understood no step towards reliable CBM can be taken. A brief review of the relative literature is presented. Various monitoring techniques are also highlighted. RUL estimation also depends on the accuracy of the sensors. Their strengths and their flaws are important when setting up a CBM strategy and therefore it is important to reach the current level of knowledge. Finally certain mathematical tools that are central in data driven RUL and are frequently met in literature are presented. These comprise of signal processing techniques as well as probabilistic RUL estimation models.

b. A gearbox data base is developed within the framework of the current research. The experimental setup is comprises laboratory level gearbox. However it gives an insight of the degradation mechanisms that take part during gearbox operation and gives further insight regarding newer and less applied monitoring techniques, particularly ODM and AE.

c. Lifelong gearbox experiments are rare and expensive to obtain. Therefore it was natural to develop a framework for lifelong gearbox degradation data simulation. The simulation focuses on the stochastic characteristics of the best performing monitoring technique. The assessment for the best monitoring technique and the estimation of its stochasticity takes place on the actual lifelong gearbox experiments performed in the laboratory.

d. A number of RUL techniques are adapted to gearbox failure data. They are assessed for their accuracy and their effectiveness.

e. Numerous lifelong rolling data bases exist on the web. They are assessed and the best is chosen for further experimentation.

f. In the current framework acoustic emission, vibration and oil debris are all tested for optimum CBM on an actual gearbox.

g. Signal processing takes place for feature extraction and feature reduction in the direction of multi sensor CBM and sensor fusion. PCA and ICA techniques are assessed in this direction not for diagnosis but for gear RUL estimation.

The latter actions which comprise the current thesis are organized in chapters as follows.

In **Chapter 1**, a brief introduction is given regarding the whole thesis. The motivation is presented along with few starting comments regarding CBM. The objectives of this thesis and certain new contributions are also highlighted.

A general review in RUL estimation techniques is provided in **Chapter 2**. Various monitoring techniques that exist in literature regarding gearboxes and rolling element bearings are highlighted. Moreover a brief overview in reference to RUL mathematical tools is given in **Chapter 3**. Stochastic modeling and regression are at the heart of RUL estimation

in actual engineering systems and is frequently coupled with signal processing of the underlying waveform data. The weight will be given in most recent RUL techniques.

A presentation of the gearbox experimental setup is given in **Chapter 4**. The recordings and some first observations regarding the experiments are given in this text. Moreover a simulation framework is presented in order to leverage our lifelong data.

The data presented in chapter 3 will be applied for gearbox RUL estimation in **Chapter 5**. This chapter covers the issue of mono-parametric RUL estimation. The simulated measurements of **Chapter 3** will make it possible to adapt and compare several RUL estimation methods. Among them, Cox regression and probabilistic ϵ -Support Vectors Regression is applied for the first time in gearbox prognostics.

Multi-sensor and multi-parameter monitoring is assessed in **Chapter 6** for degradation monitoring and prediction. AE, vibration and ODM are combined for the first time. PCA and ICA are assessed in their role as parameter fusion techniques. ICA is found superior and the resulting feature is applied for on-line RUL estimation. This time in actual degradation data. Few of the results presented in this chapter are published in [Loutas et al. \(2011\)](#).

Rolling element bearings are the subject of **Chapter 7**. Several research groups have uploaded bearing degradation data on the web. The various open access data bases are looked into and the most appropriate is chosen for further experimentation on RUL prediction. A processing technique is for the first time applied in rolling bearings to provide a more reliable threshold for novelty detection. More precisely, wavelet denoising of vibration waveforms is used to dispose off signal components irrelevant to critical bearing failure events. Feature extraction and dimensionality reduction is also discussed. In the context of feature extraction a new monitoring feature is assessed (spectral flatness). Finally a non-linear support vector regression scheme is proposed as an optimal solution for RUL estimation in bearings. A number of the results presented in this chapter have been published in [Loutas et al. \(2013\)](#) and [Roulias et al. \(2013\)](#).

Chapter 8 concludes the current thesis. The most important aspects are highlighted and few ideas for future research are provided.

It is worth mentioning that the final decision for maintenance is based on a number of factors and not only on the output of a condition monitoring system. The choice of the maintenance personnel, maintenance tools, logistics support and spares are important aspects in the final component replacement decision and are not covered in the current thesis.

1.5 Important contributions of this thesis

1. A lifelong gearbox database is produced. This data base comprises of AE, ODM and vibration full operational life recordings ([Chapter 4](#)).

2. An AE health index simulation framework is developed for degrading gearboxes under constant operating conditions ([Chapter 4](#)).

3. The functional connection between AE and ODM recordings is established for the first time ([Chapter 4](#)).

4. A Cox proportional hazards model with error bounds with AE as covariate is developed and assessed for gearboxes for the first time ([Chapter 5](#)).

5. Probabilistic ϵ -SVR are assessed for gearboxes for the first time ([Chapter 5](#)).

6. A bootstrap method is proposed for error bounds estimation of the latter method in the total absence of failure time prior ([Chapter 5](#)).

7. For the first time to the author's knowledge AE is applied for accurate quantitative RUL estimation ([Chapter 6](#)).

8. An efficient multisensory ICA-exponential extrapolation scheme is proposed for gear life prediction with error bounds ([Chapter 6](#)).

9. Few higher order moments of gearbox vibration spectrum are highlighted for the first time as adequate prognostic features ([Chapter 6](#)). For bearings, Weiner entropy is highlighted for the first time as a good monitoring index of natural degradation ([Chapter 7](#))

10. A RUL estimation scheme using multiple features and support vector regression is developed and assessed ([Chapter 7](#)).

CHAPTER 2

Overview of Bearings and Gearboxes

Monitoring Techniques

2.1 Introduction

CBM in rotating machinery comprises of several distinct steps. The accuracy of each step is important for the effectiveness of the whole scheme. [Lee J. et al. \(2004\)](#) list the particular steps of CBM as follows.

- a) Data acquisition to obtain data relevant to system health. Data are acquired via the appropriate sensors. The sensors are adjusted to the nature of the particular failure modes.
- b) Data processing. This way meaningful conclusion can be extracted from the sensor recordings so that the technician may make a maintenance decision.
- c) Data fusion and models. The fusion could incorporate a number of sensors or a number of different models.

d) Finally, the maintenance decision which will be made by the technician. The decision could be taken by an expert or by reasoning software. In the second case, this would yield an automated CBM system.

Vachtsevanos et al. (2009) propose a similar approach to CBM steps as shown in fig 2.1. In this figure each CBM step is highlighted. Below each graphic square Vachtsevanos has underlined possible flaws that may ultimately result in an erroneous maintenance decision. It is important that the accuracy and efficiency of each step are assessed.

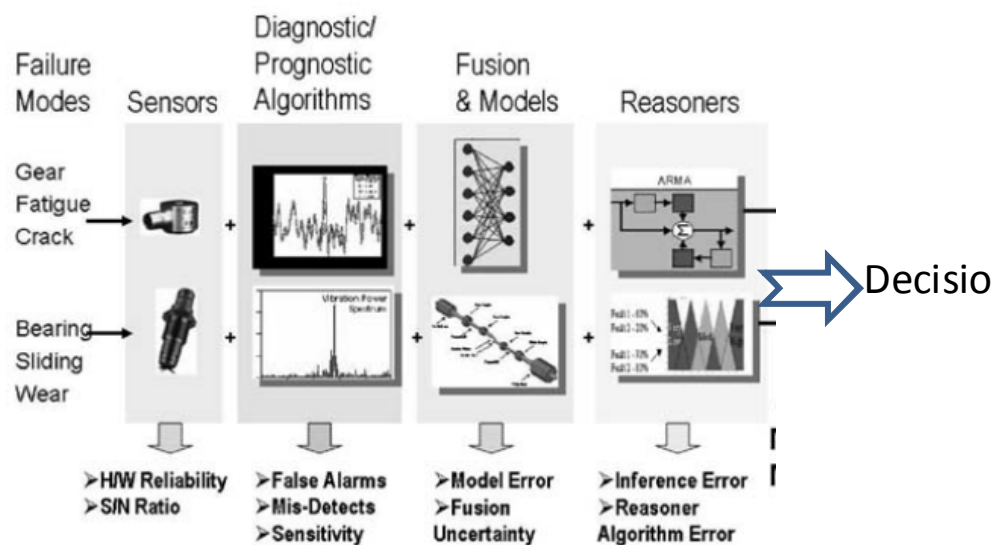


FIG 2.1. Functional layers in the CBM system architecture [Vachtsevanos et al. (2006)].

The level of the current knowledge regarding CBM is presented in the current section. Two particular aspects of CBM, monitoring techniques and data processing for the purpose of machinery RUL estimation are presented.

Various monitoring techniques exist in industrial practice as well as in scientific literature. Their desired characteristics as listed in Fantoni (2004) are:

- ✓ Non-intrusive
- ✓ Non-destructive
- ✓ Reproducible
- ✓ Applicable to a wide range of temperatures, dose-rates, pressures etc
- ✓ Sensitive to degradation, especially when close to failure
- ✓ Cost effective

Among the most frequently applied for gearbox and rolling bearings monitoring which comply with the latter characteristics are:

- Vibration Measurement and Analysis

- Oil Analysis and Tribology
- Acoustic emission
- Motor Current Analysis
- Infrared Thermography
- Acoustic Noise
- Temperature of the lubricating oil.
- Sound camera

Some of the above monitoring techniques are of waveform type, such as vibration and acoustic noise monitoring. Others are of single value type such as lubricating temperature and there is one 2-D representation (sound camera). Not all of the above pose interest from a prognostics point of view and therefore for the current thesis. Only the first four techniques, as listed above, are going to be looked into.

2.2 Importance of gearboxes and bearings

Gearboxes are essential mechanical components. Their success in transmitting and/or alternating mechanical torque has made them a key component in critical applications in helicopters, wind turbines (WT), the mining and cement industry. Gearbox design has not changed substantially for the past many decades and is not bound to change the following years. Few efforts have taken place in the direction of replacing gears, particularly concerning the design of large and costly off shore wind turbines. For instance, ScanWind has been testing gearless 3.5-megawatt wind turbines on the Norwegian coast since 2003 (MIT Technology Review, September, 2009). However, decades will pass until gearboxes are effectively replaced in energy applications, even if an efficient design has been devised. Gearboxes in their various forms are still deemed irreplaceable for torque transmission and alternation.

A single gearbox is a very costly part particularly when large industrial applications are considered. The transmission system in large off-shore wind turbines may account alone for about 13% of the overall cost of a 5MW wind turbine⁴.

Gearboxes frequently work under high thermal and mechanical stress conditions as well as moisture or chemically oxidating environments. Moreover, they comprise of a number of moving parts and subsystems. Therefore they are by nature prone to failures. Let's take the example of wind turbine application. Wind turbine gearboxes failures have been covered statistically in great extend. A survey that took place within the framework of Crabtree et al. (2010) points out the frequency and severity of various wind turbine failure modes. [Figure 2.2](#) depicts these results which come from two different projects. Although drive train fault is not

⁴ EWEA. The Economics of Wind Energy. March 2009

the most frequent failure mode it results in the worst downtime per failure and is therefore the most damaging economically failure mode. The severity of gearbox failure has also been pointed out much earlier by McNiff et al. (1990). As stated in Renewable Energy World (September 2008) “Gearbox failures account for the largest amount of downtime, maintenance, and loss of power production. These costly failures can total 15-20% of the price of the turbine...”. In another report in North American Windpower (June 2006) “One recent report says retrofits of gearboxes were necessary at all 30 turbines in an offshore farm - after less than two years. In another case many turbines were already on their second or third gearbox retrofit after only 5 years. Gearbox replacement costs are as much as \$300,000 per failure”. Moreover, there are issues connected to accessibility of the maintenance staff, especially concerning Wind Turbines (WT) and off shore WTs in particular, which involves carrying the engineering maintenance personnel and possibly a number of tools and spare parts. And WT is just an example. The final cost for gearbox maintenance or even replacement can be very high for a wide variety across industrial applications.

Failure Rate and Downtime from 2 Large Surveys of European Wind Turbines over 13 years

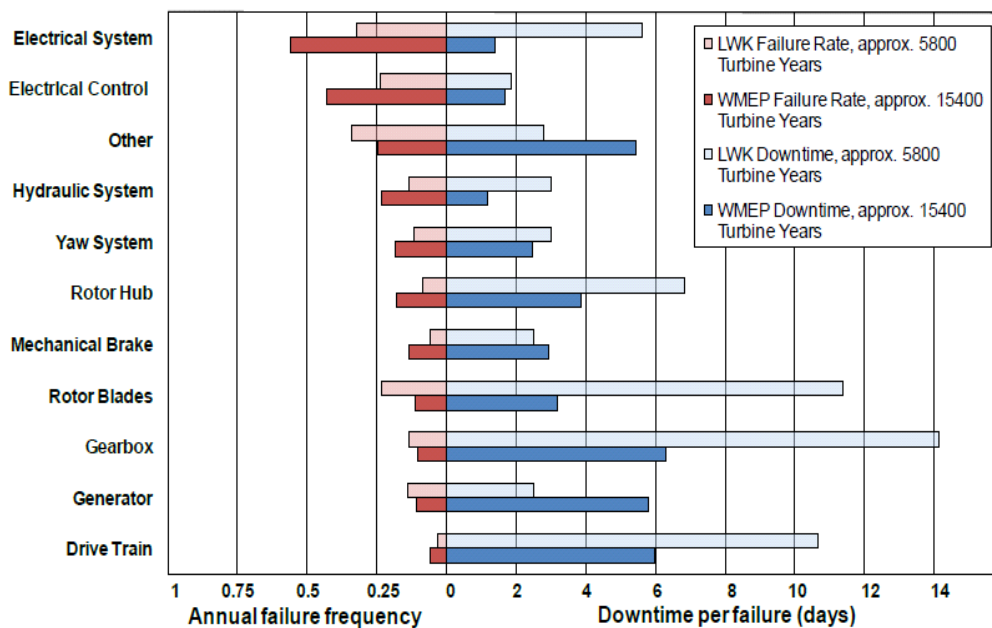


FIG 2.2. WT sub-system failure rate and downtime per failure for two surveys including over 20000 turbine years of data.

There have also been some reports relative to aviation. It has been stated that an estimated 20% of total helicopter accidents are attributed to gearbox failure or malfunction (Tan (2005)).

Bearings are another key component of rotating machinery. Their function is to constraint relative motion of parts to only a desired motion and reduces friction between moving parts. The root cause of a number of failures can be traced back to the drive train bearing system. Again we will use the WT example for justification. In a survey that took part by [Alewine \(2010\)](#) identifies bearings as the number one root cause for WT failures.

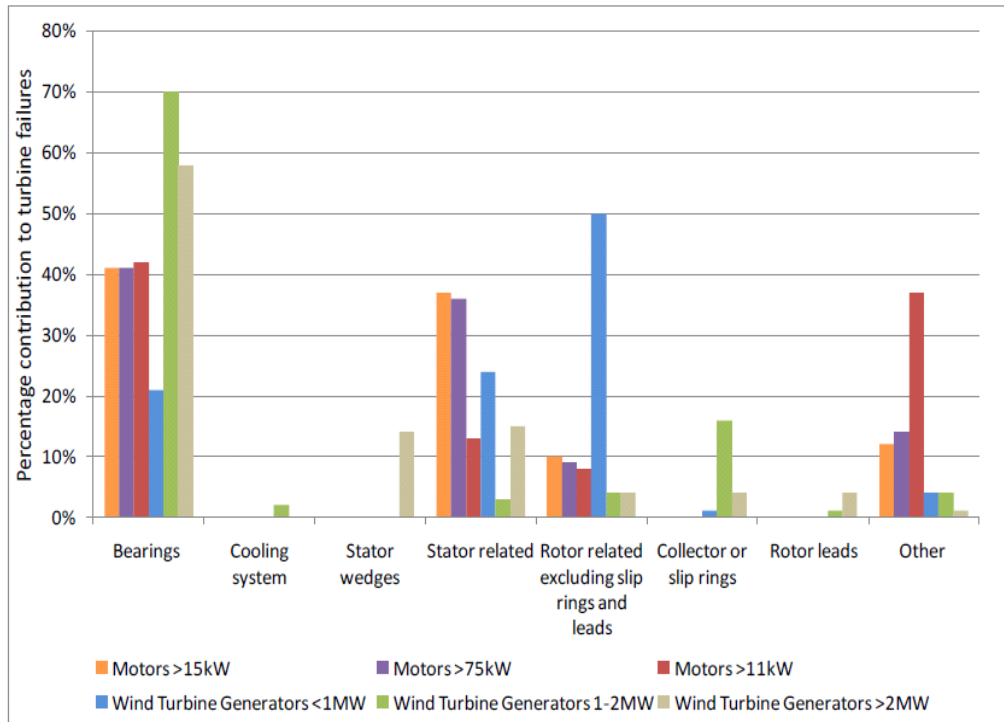


FIG. 2.3. Comparison of failure contribution of WT generator components taken from [Alewine \(2010\)](#).

The survey results are summarized in [fig. 2.3](#). Furthermore, as stated in [Tavner \(2008\)](#), the largest percentage of root causes of failures, not only on WTs but also in other industrial applications where electrical machines are used can be traced back to bearings. His research findings are summarized in [fig. 2.4](#).

In induction motors, the most common failure is also related to bearings followed by stator winding failures and rotor bar failures as reported in [Ocak et al. \(2005\)](#).

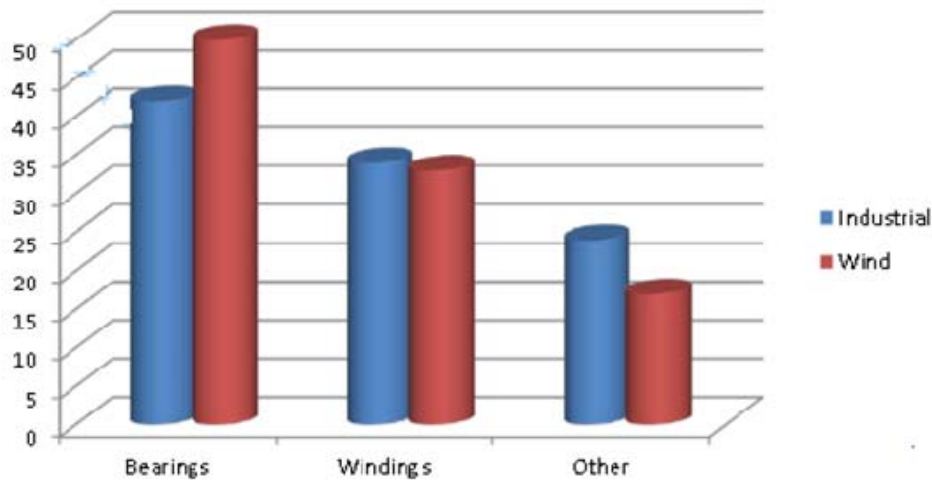


FIG. 2.4. Frequency of various failure root causes as a percentage

To sum up, the importance of gearbox and REB monitoring is evident. The economic stakes attached to gearbox and bearing mechanical systems are high. These two components are going to be the focus of this research and they are going to be assessed separately in the following chapters.

2.3 Gearbox modes of failure

When input and output axes are parallel to each other two types are the most common in practice, spur and helical gears. In spur gears the length of contact is constant whereas for the spur gears it gradually increases and decreases in length over a particular tooth. Sliding and rolling are the main types of stress on the meshing gear teeth. There are two types of fatigue resulting in different types of failure. Bending fatigue is the main cause for crack initiation near the tooth root which eventually leads to complete tooth fracture. On the other hand, surface fatigue, which is the most frequent mode of gear failure, is the cause of surface pitting, spalling and tooth wear (Fernandes and McDulling (1997)).

Pitting is considered as the initial phase of spalling and occurs constantly beneath the surface of the tooth during gear meshing process. Spalling is the next level of fatigue failure and it manifests as small craters on the surface of the tooth. Pitting and spalling eventually lead to material detachment from the surface of the tooth. Finally, the process of fatigue wear characterizes evident changes in the shape of the tooth cross section. The lubricating oil plays a significant role in the tooth contact process. When the two sliding bodies meet, an area of very high pressure is created locally. A very thin oil film forms between these two surfaces of the scale of micrometers. This film distorts the meshing teeth pressure profile and eventually

the contact surfaces themselves. The elastic distortion due to the lubricating oil is frequently referred as elasto-hydrodynamic (EHL) regime.

2.4 Bearings modes of failure

Two are the most common bearing designs, ball and roller bearings, both sharing a common operating mechanism. The ball bearing consists of an inner race, the balls and the outer race. The main method of lubrication is grease lubrication which is particularly recommended for moderate applications with speed parameter less than 600000 DN^5 .

Rolling element bearings can malfunction in a number of ways. Inner race, outer race, ball spalling and cage faults can arise due to poor lubrication, improper assembly or just because of operating fatigue. After all every bearing has an expected operation life typically in thousands of hours or in millions of revolutions. [Randall \(2004\)](#) summarizes the various modes of bearing failure. Each mode corresponds to a natural frequency. This frequency appears in the vibration spectrum and most often it is obscured by noise and modulated by the gear meshing frequency and its harmonics. The most common means of extracting these signatures from vibration recordings is through the method of enveloping which will be briefly covered in a following paragraph. It should be noted that, regarding bearings, the focus of this research is in prognosis and to a lesser extend diagnosis or failure analysis.

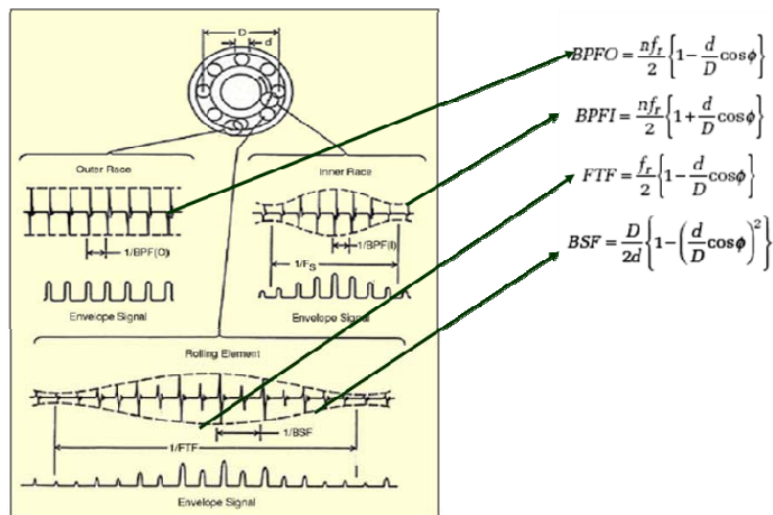


FIG. 2.5. BPFO = ball pass frequency, outer race, BPFI = ball pass frequency, inner race, BSF = ball spin frequency; FTF = fundamental train frequency (cage frequency).

⁵ DN is a product where D is the bearing mean diameter and N is the rotational speed in rev/min.

2.5 Monitoring techniques

2.5.1 Vibration

Vibrations monitoring is one of the oldest and more widely applied techniques for rotating machinery CBM and NDT. Typically vibration recordings range from tens of Hz to 20 KHz, depending on the particular fault signature and dominating resonant frequencies of the structure. Vibration recordings have rich information content related to the nature of the rotating frequency (oil whirl signatures, rotating frequency and its harmonics, tooth meshing frequency and its harmonics) as well as unprecedented fault signatures (gear faults, bearing faults etc). In the past vibration recordings were obtained through velocity probes. A velocity probe cross section is depicted in [fig. 2.7](#). These sensors capture the first derivative of the strain induced by the machine operation. They were able to capture very low frequency signatures of up to 1/3 of the rotating frequency. The dominant frequency in a rotating system is the meshing frequency. The meshing frequency always refers to a particular gear and it is given by $f_m = N_t * f_r$ where N_t is the number of gear teeth and f_r is the gear rotating frequency.

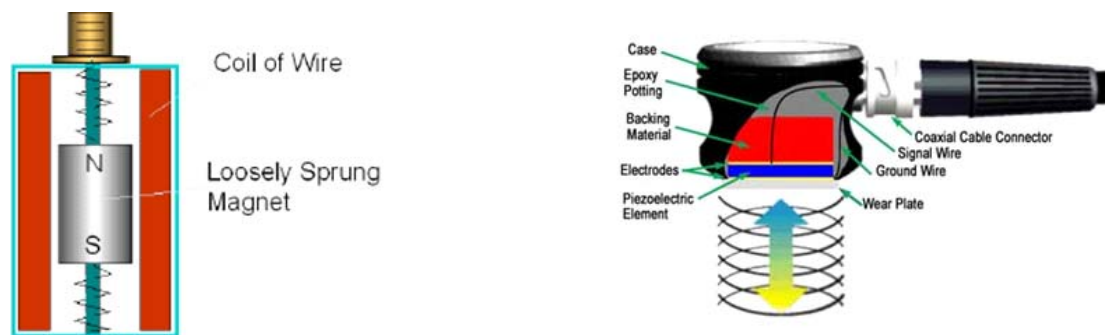


FIG 2.6. On the left a velocity probe cross section. A piezoelectric accelerometer is presented on the right.

Incipient gearbox and bearing faults typically manifest in the gear meshing frequency and its multiples. This is most commonly in the band of few kHz. A piezoelectric accelerometer is the appropriate sensing device for monitoring this kind of faults.

[Randall \(2004a, b\)](#) has highlighted the most important aspects regarding vibration monitoring of rotating machinery. His focus is mostly on diagnostics. He states that different modes of gearbox failure manifest in different frequency bands. Regarding gears, Randall underlines the importance of the load on the axis relative to the monitoring ability of the vibration technique.

[Jardine et al. \(2006\)](#) presents an extensive review regarding machinery diagnosis strategies involving vibration monitoring until that date. [Randall and Antoni \(2011\)](#) give an excellent tutorial regarding the various modes of failure in bearings and the means to identify

them from a very early stage. They review a variety of signal processing for failure signature extraction all based on vibration monitoring.

A general representation of a rotating machine vibration response is the following

$$x = (d + w + e_0) * h \quad (2.1)$$

Where

x : is the accelerometer recorded signal

d : is the deterministic component of the machine operation. This is attributed to the driving torque and gives rise to the oil whirl signature, rotating frequency and its harmonics as well as tooth-meshing frequency and its harmonics

w : is an unprecedented faulty signature. It manifests as a transient signal. It may be a frequency modulated signal, amplitude modulated signal or a combination of the latter.

e_0 : is a contextual noise component. Most usually it can be assumed to be a normally (Gaussian) distributed variable with mean zero and some constant variance σ^2 .

h : returns the effect of the transmission path. The transmission path is the rotating machine structure that intervenes between the source of vibration and the sensor interface. The $*$ sign stands for convolution in time domain. The latter combination of signal sources is graphically depicted in [fig. 2.8](#).

The signal is recorded by digital means. The sampling frequency is normally set to capture a particular mode of failure. The maximum sampling should theoretically be twice the highest faulty signal frequency according to Nyquist. In practice the sampling frequency is 3 to 4 times this criterion. Eventually the signal x is saved as a discrete time series and it is processed as such.

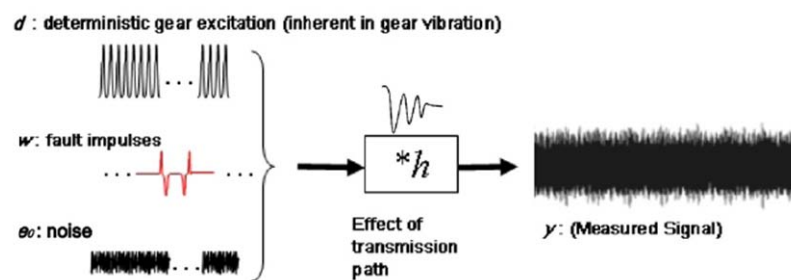


FIG. 2.7. A graphic representation of the vibration response formation in a meshing machine.

The general idea behind vibration CBM is to pinpoint and isolate the particular effect of w , which, most often, is masked by the addition and convolution schematically shown in [Figure 2.8](#). This isolation can be performed with techniques such as Time Synchronous Averaging (TSA), envelope analysis, Spectral Kurtosis (SK) and Wavelet Denoising.

TSA: The most common way to cancel the noise is performing time synchronous averaging (TSA). This method transforms the recorded signal from the time domain to the angular domain, the angle being measured on a particular gear. The process resembles a comb filter and was introduced by [McFadden \(1987\)](#). Since then it has been proven a very successful tool in gearbox diagnosis.

$$x_{averaging}(t) = \frac{1}{N} \sum_{n=0}^{N-1} x(t+nT) \quad (2)$$

Where N is the number of averages and T is the shaft rotational period

The weakness of this method is that it requires a totally synchronous signal. Otherwise smearing is introduced that screens the desired frequency peaks. The result of a small axis speed deviation on the vibration signature is depicted in [Figure 2.8 \(Jafarizadeh et al. \(2008\)\)](#). [Jafarizadeh et al. \(2008\)](#) have proposed a remedy for asynchronously averaged signals. They developed a correlation factor that corrected the shaft revolution signal duration. The weakness of their method is the demand for an outside tachometer signal adjacent to the shaft of interest. The tachometer made it possible to map the recorded signal to the angular domain. The application of tachograph can be difficult in industrial applications where the rotating gear of interest may be very difficult to reach. [Combet and Gelman \(2007\)](#) proposed a blind TS averaging without the need of a tachometer signal. However their method was limited to a very small deviation from the average shaft rotating frequency typically smaller than 5%.

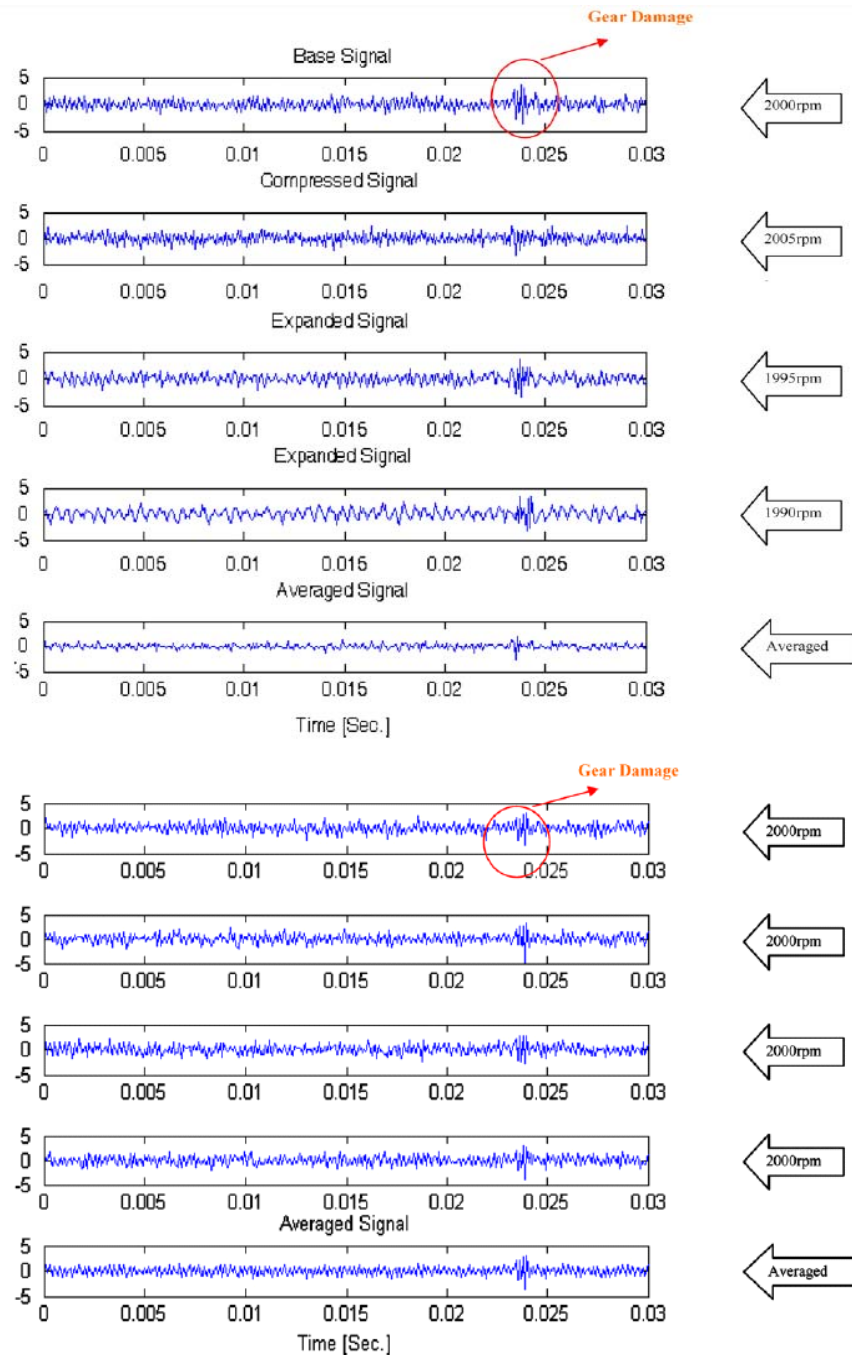


FIG. 2.8. A slight deviation in shaft rotating frequency (upper graph) may completely mask the underlying gear fault signature (Jafarizadeh et al. (2008)).

It should be mentioned that the current thesis is not concerned with asset diagnosis which in general is a section of high maturity in academia. The focus of this thesis is prognosis. However, as shown earlier in the introducing chapter, diagnosis and prognosis are connected. An expert should know at which point of degradation his asset is in order to make a life prediction. This means that a prognosis measure is a good diagnosis index. The ways

diagnosis and prognosis are combined within a CBM framework can vary. One should be careful though that a good diagnosis measure may not always be a good prognosis. Particular measures of failure such as kurtosis (Eq. 2.3) with proven damage detection ability are obsolete in predicting the failure point of a machine or its RUL.

$$\text{Kurtosis} = \frac{\int_{\text{time_window}} (x(t) - \hat{\mu})^4}{\sigma^4} \quad (2.3)$$

$$\text{Root_Mean_Square} = \frac{\int_{\text{time_window}} (x(t) - \hat{\mu})^2}{\text{time_window}} \quad (2.4)$$

Where *time_window* is the length of the vibration time series under consideration, σ is the time series sample deviation and $\hat{\mu}$ is the estimated mean of the timeseries.

On the other hand vibration RMS (Eq. 2.4) which is possibly the oldest vibration monitoring parameter can be applied efficiently for damage diagnosis and prognosis as will be shown in Chapter 4.

Frequency Analysis: Fourier transform and its practical implementation Fast Fourier Transform (FFT) are powerful tools for machinery diagnosis. Vibration spectrum most often possesses a strong noisy content. Due to its high noise content an estimator of the vibration spectrum is used such as Power Spectral Density (PSD). PSD is the spectrum of the auto-covariance function as shown in Eq. 2.6. In discrete time- discrete frequency form

$$S[k] = \sum_{n=0}^{N-1} x[n] e^{-j2\pi \frac{kn}{(N/2)}}$$

$$PSD[k] = \sum_{n=0}^{N-1} \gamma[n] e^{-j2\pi \frac{kn}{(N/2)}} \quad (2.6)$$

$$\gamma[n] = \frac{1}{N} \sum_{i=1}^{N-l} (x[i] - \hat{\mu})(x[i+n] - \hat{\mu})$$

Where n is an integer index, x is the digitized signal and N is the total number of samples in x timeseries, γ is an estimator of the auto-covariance function $\hat{\mu}$ is an estimate of the timeseries mean. The relations are taken from Fassois (2005).

In the section of bearing faulty modes it was shown that particular faults give rise to different frequencies. These faulty signatures are obscured by the normal operational frequencies (gear meshing) of the rotating system. Techniques such as envelop analysis are applied to extract these signatures.

Envelope Analysis: Faulty signatures in rotating machines are amplitude modulated with the normal operation frequency, in our case the meshing frequency and its harmonics. The result is the translation of the faulty signature in the frequency domain and the appearance of modulated sidebands around the modulating signal. The isolation and amplification of the modulated signal is performed with envelop analysis. Envelop analysis takes the vibration recording and checks the presence of a specific signature by i) band-pass filtering for modulation band isolation ii) Hilbert transform to isolate only a single side band iii) calculation of the analytical signal which is the inverse Fourier transform of the modulated side band and iv) rectification and smoothing of the analytical signal to extract the periodicity of the faulty signature. The difficult task of pinpointing the optimal band for demodulation is effectively performed with the help of Spectral Kurtosis (SK) and its graphical approach, the kurtogram ([Antoni \(2006\)](#)).

SK: has been proven an excellent measure for detecting incipient faults in bearings. Antoni has given a theoretical basis for SK ([Antoni \(2006\)](#)) as well as a Matlab code (*Fast_Kurtogram.m*) to efficiently perform amplitude demodulation with sequential band pass filters and SK. SK has found application in bearing fault diagnosis.

Wavelet Denoising: Wavelet denoising is a technique to dispose of noise from the vibration recordings and enhance the signature of an incipient fault. To this end the signal is split into frequency bands with discrete wavelet decomposition (DWT). Then each high frequency band is thresholded according to the method of [Donoho \(1995\)](#). Finally the residual signal is reconstructed according to DWT theory. The faulty signature, if existent, is supposed to be more distinct in the denoised form. [Peng and Chu \(2004\)](#) provide a bibliographic reference in wavelet denoising applications in mechanical systems. The particular work also covers general applications in machinery diagnostics in reference to wavelets. [Loutas and Kostopoulos \(2012\)](#) provide a more up to date list of existing applications of wavelets in CBM of gearboxes, bearings and electric motors.

Cyclostationarity: Let's focus again on [eq. 2.1](#) in an effort to characterize the signal x . In machinery that rotates under constant conditions, there will be always deviations from exact periodicity. In gearboxes this is attributed to small speed fluctuations that normally exist in real life industrial applications. The hidden periodicities of a gearbox can be extracted through TSA, as mentioned earlier.

In bearings, the slip of the rollers inside the bearing cage introduces additional deviation from periodicity. As stated in [Randall et al. \(2001\)](#) signals from bearing faults end up being asynchronous to the shaft rotation. This means that even after compensation of speed fluctuation with TSA bearing faulty signatures are difficult to be extracted. [Randall et al. \(2001\)](#) propose higher order statistical tools in the framework of cyclostationarity analysis to

handle these kind of faulty signatures. The signals whose auto-covariance (which is a second order statistic) is periodic function are referred to as cyclostationary or higher order periodic).

To conclude this section, the advantage of vibration monitoring is its proven efficiency in fault diagnosis over the years. Moreover, numerous methodologies exist to extract faulty signatures not only in academic literature but also in practical applications. The disadvantage of vibration monitoring is the high noise component especially when the machine operates in an industrial environment and the prevalence of tooth meshing frequency harmonics in the recordings. The faulty signature is often obscured by these factors.

2.5.2 Acoustic emission

Acoustic emission (AE) manifests as stress waves travelling inside a material. The source of these waves is the material itself and particularly its inner structure. Outside mechanical stresses that change the material's inner structure are the source of these stress waves. Sources of AE in mechanical equipment can be impact, friction, bubble formation and collapse or collapse of vortices.

Since AE is connected to the inner structure of the material its energy is normally concentrated in the range of 100 to 200 kHz and this is the main advantage of AE monitoring. Due to this fact, high background noise and structural resonances are unlikely to affect the recordings. Another aspect that proves AE a better choice in many applications is the fact that it is non-directional in contrast to vibration monitoring which requires the application of three axis sensory system to fully cover all degrees of freedom.

The AE sensors are not much different from their vibration accelerometer counterparts. They consist of a metallic case, a piezoceramic element and damping material. Structurally they are much smaller in size than vibration piezoelectric sensors. The main difference to their counterpart is the different resonant frequency band. AE signals that are captured within this band (hundreds of kHz) are normally quite faint. Therefore it is common that a preamplifier is also used to enhance the signal prior to the Data Acquisition equipment.

The (AE) recordings may be of burst, continuous or mixed type as shown [fig. 2.10](#). The main drawback of this sensing technique is that it is subjected to attenuation while travelling through the conducting material. Therefore, AE sensors should be placed very close to the particular component that needs to be monitored. Another shortfall is that AE events may be produced from sources irrelevant to damage progression. Therefore care should be taken in order to correlate the respective AE recordings to the underlying degradation mechanism.

There is an increasing interest in AE monitoring of rotating. Several works have investigated the correlation of AE events to various types of gearbox mechanical stresses and

degradation levels. Bending experiments in non meshing gearboxes (Singh et al. (1999), Miyachika et al. (1995)) have highlighted AE effectiveness in early crack identification under various conditions. Several works have taken place regarding AE monitoring of meshing spur gears with seeded and natural faults (Singh et al.(1999), Raad et al.(2003), Tandon and Mata (1999)). All latter experiments simulated actual industrial working conditions and their findings highlighted the superiority of AE monitoring both for seeded and natural fatigue faults. Tan et al. (2007) have conducted a thorough research regarding the monitoring ability of AE in natural spur gear fatigue. They measured the level of pitting by Spectrometric Oil Analysis at regular –though quite sparse-- intervals during the gear operation. Their results showed that the recorded AE activity was linearly correlated to the actual pitting rate of an operating gearbox. Their findings were confirmed for a variety of loading conditions.

Loutas et al. (2006) and Loutas et al. (2009) have also experimented with acoustic emission in rotating machinery. The most significant outcome of their study was the verification that RMS values of AE signals can discriminate successfully between different damage and loading states in an operating gear box.

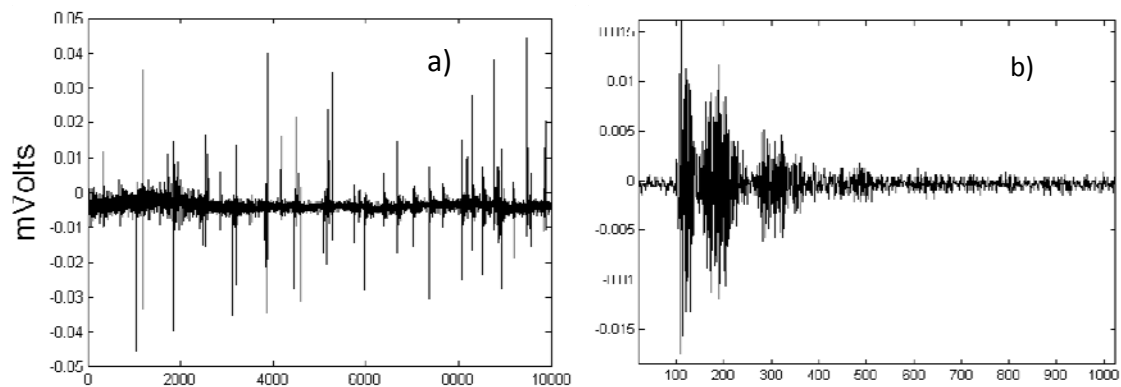


FIG. 2.9. Acoustic emission of continuous type (a) obtained from acoustic emission sensor mounted on a gear box case. Acoustic emission of burst type (b) coming from composite material. The material was tested for fatigue. The vertical axis is in Volts, the horizontal axis is the time steps.

Moreover they applied signal processing techniques and Discrete Wavelet Transform (DWT) in particular to enhance the AE fatigue monitoring ability. Furthermore they developed a new, low cost type of slip ring which could contribute to the further application of acoustic emission method in rotating machinery (Loutas et al. (2008)). Eftekharijad (2010) has contributed a thorough investigation concerning AE monitoring in helical gears confirming the effectiveness of the monitoring method.

AE has also found application in bearings monitoring. Among the most influential studies that compare AE and vibration monitoring in bearings was Shiroishi and Li (1997). They conducted experiments with inner and outer race seeded defects. His results underline

the potential of AE in comparison to vibration. [Al-Dossary et al. \(2008\)](#) also conducted experiments relative to bearing diagnosis. They experimented with a variety of defect sizes and bearing defect modes. With the application of AE they were able to characterize the size of the seeded fault with spectacular accuracy.

As has been noted, AE has several limitations. [Sikorska et al. \(2008\)](#) have provided an overview on acoustic emission condition monitoring advantages and flaws. They state that noise is the most significant infliction with AE recordings. Noise stems from

- (a) residual, steady background noise from electronic components;
- (b) distorted signal components due to exceeding the limits of components in the signal acquisition chain;
- (c) aperiodic transient events, such as valve activity or electro-magnetic interference (EMI) and radio frequency interference (RFI);
- (d) quasi-periodic transients events such as rubbing of seals or bearing faults (when not looking for incipient seal or bearing faults);
- (e) low frequency modulation of the background continuous AE level, also common in reciprocating machinery signals or due to misalignment;
- (f) genuine AE from mechanisms of no interest to the analysis being performed. These also include AE activity associated with the operation of healthy machines or fluid flow through pipes.

AE monitoring ability in rotating machinery is promising and still not fully investigated. It is inherently relieved from gear meshing periodic components and its signal to noise ratio is proven higher than conventional vibration recordings. AE is an optimized monitoring method and application of signal processing doesn't improve dramatically its diagnostic ability. [Eftekharnjad \(2010\)](#) applied spectral kurtogram in raw AE waveforms of a defected helical gear. The results didn't provide additional information. On the other hand, signal band pass filtering actually resulted in better resolution of AE events and this was also observed with an increase in crest factor. [Kilundu et al. \(2011\)](#) applied cyclostationarity theory to increase the resolution of the 1-D burst signatures. It is the author's view that AE is already an optimized monitoring technique and sophisticated signal processing does not improve dramatically its diagnostic potential.

In the current thesis AE is going to be applied for machinery life prognosis. To the authors knowledge this is the first time effort of AE application for point of failure prediction and RUL estimation.

2.5.3 Oil Debris Monitoring (ODM)

As it was mentioned earlier in this chapter various degradation modes appear due to the constant operation of a gearbox. These degradation modes include pitting of the gear tooth surface, pitting and spalling that can result in cracking initiation. These mechanisms are due to the friction between the gear teeth and immediately result in loss of material from the teeth surface. In our case this material is ferromagnetic and can be captured with an OD monitoring sensor. OD sensor consists of three coils that surround a magnetically and electrically inert section of tubing. The two outside field coils are driven by a high frequency alternating current source such that their respective fields are nominally opposed or cancel each other at a point inside the tube and just under the center sensor coil. The lubricant carrying the released ferrous particles passes through the magnetic coils. The particles are filtered after the ODM coils. This way they do not re-enter the lubricant circuit and only the newly generated particles are counted. The field coil produces a magnetic field that alters when the amount of ferrous particles passing through the circuit changes. This is measured as a small electric pulse which is captured inductively by the smaller “sense” coil. The amplitude of the pulse is proportional to the amount and mass of particles passing through the coils. Sensor is sensitive to particles from a certain size threshold. It has been shown the aggregate mass of ferrous particles can be correlated to the structural degradation of the gearbox (Dupuis (2011)).

Devices that measure other metallic particles, such as Cr, Cu and Mg have also been applied for engines and rotating machinery diagnostics. These devices are described as Spectrometric Oil Analyses devices. These elements exist inside the material alloy and are released with the degradation progress of the asset. Jardine et al. (1987) assessed the effect of non-Ferrous particles.

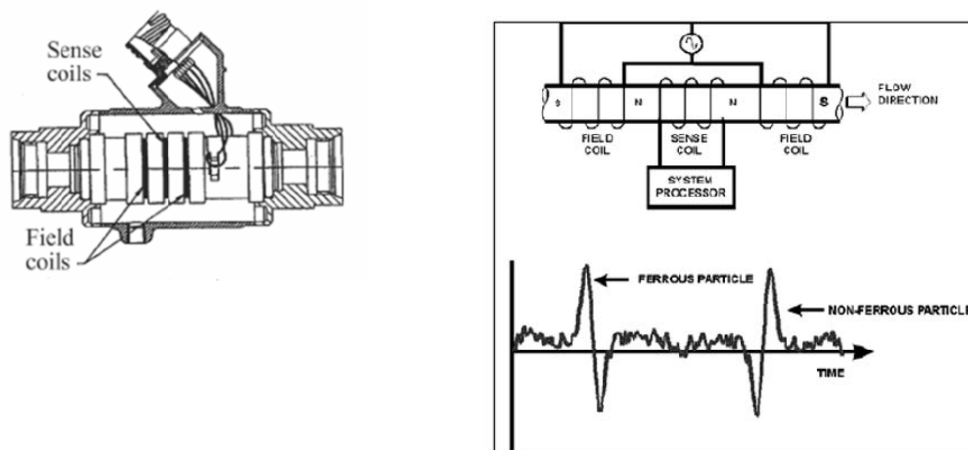


FIG. 2.10. A cross section of an ODM (left) and principle of operation (right).

They concluded that non ferrous materials make no independent contribution to the overhaul maintenance decision making. That does not mean that the latter elements are not diagnostic of engine wear but rather that they are closely related to the Fe particles level. Other metrics deriving from ODM, such as the rate of ferrous particles generated and their size can also provide insight regarding the actual health condition of the machine as shown in [Loutas et al. \(2011\)](#).

A shortfall of ODM is that it is very sensitive to start up and shut down of a machine. Such events generate big amounts of ferrous particles which are not correlated to the health state of the machine. Moreover, the lubrication filter may not trap all ferrous particles ([Miller and Kitaljevich \(2000\)](#)) sending them back to the lubrication circuit. Finally, depending on the geometry of the gearbox case, ferrous material can be trapped inside the case again resulting in poor measurements. All the latter issues have to be addressed in order to obtain a safe debris monitoring.

[Dempsey et al. \(2003\)](#) at GRC/NASA have conducted extensive experimental work and published interesting results from gear testing at a special test rig utilizing vibration and OD measurements. With the goal to improve the performance of the current helicopter gearbox health monitoring systems, they have tested gears at high shaft speed for multi-hour periods and correlated special parameters-features extracted from the vibration recordings with the Fe debris mass accumulated during the tests. They have integrated their results in a fuzzy logic based health monitoring system with satisfactory performance for a series of tests. [Dupuis \(2010\)](#) has experimented with real size turbine gearboxes and bearings. He correlated the OD recordings with damage size. Moreover he produced some empirical decision thresholds regarding the cumulative ferrous particles mass. These thresholds were used for gearbox and bearing diagnosis and maintenance decision making. Finally he delved into RUL prediction. He used the ODM count recordings to estimate two moving average predictions, a short term and a long term. He then calculated a weighted average of the latter to produce the Fe-count extrapolated point to which he would assign a RUL estimation. His results were relatively satisfactory. [Bolander \(2009\)](#) also utilized ODM recordings rather to diagnose the level of damage and not for actual prognosis. He used the ODM recordings to feed a physics based prediction model, therefore creating an ensemble condition based-physics model for prediction with relative success.

One of the current thesis contributions is to further assess the applicability of ODM to gearbox fatigue with a focus on failure prognosis.

2.5.4 Motor Current signature

As has been mentioned earlier in this chapter, gearbox and bearing failures are closely linked to the structural integrity of electrical machines and particularly of the asynchronous type. Any slight change in the structural integrity of the rotating machine is supposed to create changes in the electrical dissipated power (or produced in case of a generator) of the machine. The electric power for a three phase machine is given as

$P = \sqrt{3} V I \cos\theta$. V is the instant voltage difference between 2 phases, I is the instant stator current value and $\cos\theta$ is the machine power coefficient. For machines which are linked to the electric grid, the voltage is generally steady. Any change in the condition of the machine is largely expected to appear as a small change in the instant power dissipated and therefore in the current signature. Closed loop Hall Effect sensors are used to measure currents¹ since

- They have very good accuracy
- They have an almost flat frequency response for a wide frequency band, making possible to capture potential high frequency transients
- Negligible dependence on temperature
- Linearity

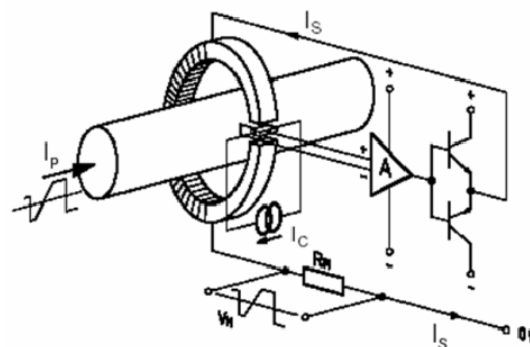


FIG. 2.12. Principle of operation for a closed loop Hall effect electric current sensor.

The most notable advantage of current signature monitoring is that it is applicable to machinery that is hard to reach or the attachment of a conventional vibration accelerometer is impossible. Moreover their recordings are characterized by no attenuation of the underlying signal and they have a very low cost in comparison to AE or vibration sensors.

Several works have explored the effectiveness of current signature analysis in bearing as well as in gearbox fatigue monitoring (Eftekharijad (2010), Janjarasjitta et al.(2008), Lu et al. (2012), Lu et al. (2012), Kar and Mohanty (2006)), all latter works focusing principally on mechanical fault diagnosis. Their work dealt with the sidebands of the first current harmonic that arise with the appearance of a fault.

The assessment of current signature analysis will not extend further because this monitoring method is not taken into account in the current thesis.

CHAPTER 3

Remaining Useful Life Estimation in Engineering Systems

3.1 Terminology

The terms “prognostics” derives from Greek word προ-γνώση meaning the knowledge of the future. It is one of the most intriguing issues in CBM even impelling some researchers towards lyricism (“the holy Grail of RAMS⁶ analysis” by [Zio and Pelsoni \(2013\)](#), “the Achilles’ heel of CBM” ([Vachtsevanos and Wang \(2001\)](#)).

There seems to be no consensus regarding the term “prognostic” and “prognosis”. In the current thesis “prognosis” comprise a set of actions that provide a good RUL estimate along with adequate confidence intervals for a given operating asset.

In the current chapter we will present works that assume a constantly operating mechanical asset. The choice of a proper degradation metric is the first step for prognosis.

⁶ Reliability Availability Maintainability Safety

This degradation metric is extracted from sensors that monitor the asset's operating condition. It may be single valued or a vector of values. It is worth mentioning that in the current thesis framework both single value and multiple dimension measurements are going to be assessed in the analysis of [Chapter 5 and 6](#).

Throughout the asset operating life this metric is sampled at intervals coinciding with the sampling intervals of the sensors. The corresponding sampling period will define the diagnostic and prognostic resolution of the overall CBM scheme. Let us now be more specific.

Supposing this period is t_s . A basic assumption which is made forces t_s to be much shorter than the total operating life of the asset, that is $t_s \ll t_{op}$. This is essential in order to capture the dynamic characteristics of the degradation process. Let us also define the metric timeseries as $y(t_i)$, $t_i = i * t_s$, $i = 0, \dots, (N-1)$ where N is the total number of historic samples and $N * t_s = t_{op}$. The sampling is assumed homogenous that is $t_s = t_{i+1} - t_i$. Several metrics can be defined regarding the asset operating history as well as the condition monitoring feature.

Time to failure (T_{fail}): It is the time at which the asset can no longer perform the task it was originally designed for. There are various failure processes. For instance, an asset may gradually pass from every degradation stage until it fails therefore defining a “soft failure” mode. In the case of gearboxes, a “soft failure” is induced due to the gradual surface pitting of the meshing gear teeth. In contrast, a “hard failure” occurs when the asset suddenly passes from a very low level of degradation to final failure with one jump. Regarding gearboxes, this could be a gear tooth unexpectedly being chopped off during operation.

Remaining useful life (RUL): Let's assume that at some observation point $t_i = t_{obs} < T_{fail}$ a total number of n_{obs} feature observations is available $[y_1, y_2, \dots, y_{n_{obs}}]$. In this context, RUL can be defined as the conditional expected time of to failure, given current working age

$$\widehat{RUL}(t_{obs}) = E[T_{fail}(t_{obs}) - t_{obs} | T_{fail} > t_{obs}, y_i, 0 \leq t_i \leq t_{obs}] \quad (3.1)$$

Time of failure is assumed in the current context as the total operating time until the feature $y(t_i)$ crosses a user defined threshold. Since these feature measurements are discrete they define a set of numbers and the failure time can be defined as $T = \sup \{t_i : y(t_i) \equiv y_i < y_{thres}\}$ where y_{thres} is the defined failure threshold.

Failure threshold: The selection of a failure threshold on the measured condition feature is in general a difficult though a very significant practical issue. ISO 2372 and ISO 10816 are frequently adopted for defining adequate vibration threshold levels. In general failure thresholds require significant experience on the particular operating asset. Defining exact, robust and reliable failure threshold of particular components is outside the scope of

this thesis. All prediction models in literature provide a prediction pdf that in most cases incorporate operational uncertainties and this will also be shown later in this chapter. Several researchers have added stochasticity in the value of y_{thres} (Nystad et al. (2012), Si et al. (2013)).

Defining a prediction model: The current study will focus on data-driven prediction models. This fault model is needed to propagate the current failure state until the failure threshold is reached. A set of assumptions is essential in order to perform an assessment. The set of assumptions will define the output of the model and for the current thesis they are going to be listed in [Chapter 4](#).

After the definition of the prediction problem a particular model should be chosen. A number of predictive models have been proposed in literature and their choice depends on the particular application and end-user demands. Sikorska et al. (2011) have presented a literature research on data driven models spanning until the year 2010. The implementations that are presented comprise different levels of complexity and each one has different strengths and limitations. According to Sikorska the family of prediction models less encountered in literature are extrapolation models, although these are the most frequently used in industrial applications. This thesis will try to fill this gap by contributing an improved extrapolation model based on multiple sensors fusion ([Chapter 6](#)).

Past sensor data are gathered that best describe the degradation of the particular asset. These data may come from other component failure histories, the particular component which is monitored or a combination of the latter. These recordings are bagged together and divided in training and testing sets. The training set is used to define the model parameters. The testing set is used to measure the final accuracy of the prediction model on unseen data.

The sensing method most commonly applied for prediction in rotating machinery is vibration acceleration. Feeding the whole waveform in the prediction model is computationally expensive. It is more convenient to extract some statistic feature from the raw waveform. The most frequent statistic used in literature is the waveform RMS value and less often the maximum peak-to-peak amplitude. Other statistics and combinations are less common in literature. Acoustic Emission (AE) and Oil Debris Monitoring (ODM) features are also rarely applied for gearbox life prediction. The current thesis is going to contribute to this direction assessing the prediction accuracy of AE and ODM. Moreover a fusion scheme is going to be proposed for gearbox feature extraction in [Chapter 6](#). In [Chapter 7](#) a combination of multiple statistic indexes is proven to be highly effective in predicting bearing RUL.

Confidence intervals: A point estimate regarding the future health of a mechanical asset is rarely of significance especially when the health measurements are noisy or when its operation comprises uncertainties. Confidence intervals are very important in maintenance decision making especially in critical areas (i.e. the nuclear industry). Acceptable confidence

bounds on our estimation are given by industrial practice. 90% and 95% confidence are the typically applied to assess the goodness of the estimation. These intervals incorporate the uncertainty in our predictions.

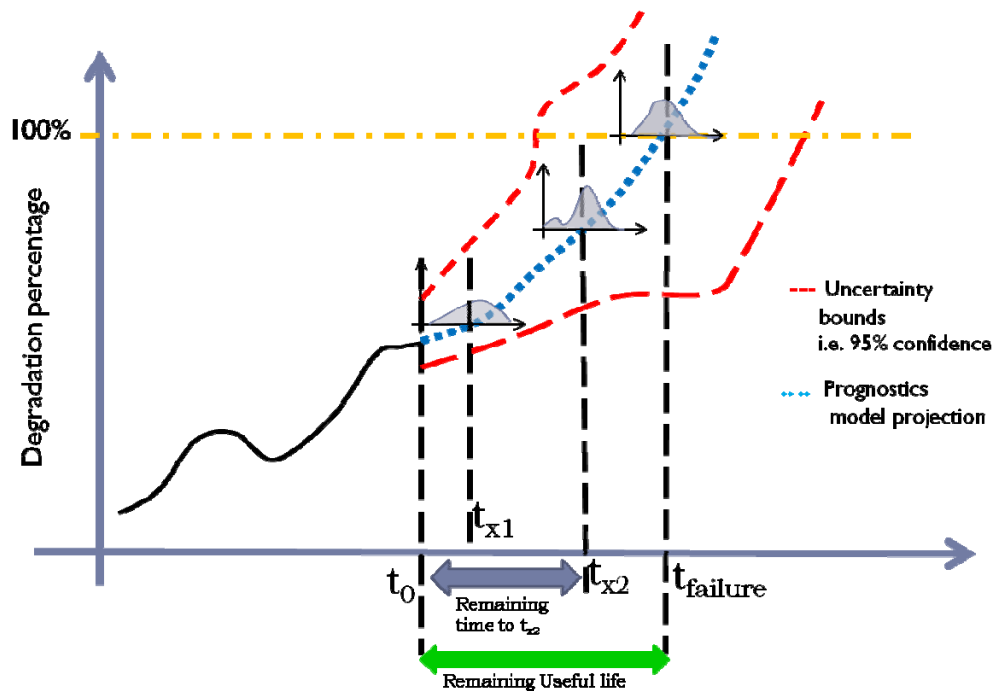


FIG. 3.1. The issue of prognosis in engineering systems.

Whether a predictive model is acceptable or not depends on the accuracy of its prediction. Most often, if the prediction is within acceptable bounds the model can be considered acceptable even if the point prediction does not coincide with the actual remaining life of the machine (Fig. 3.1). Prediction accuracy is the closeness of the estimated value to the actual value of the predictor. Precision, on the other hand, is a measure that adds the estimation distribution information to the performance measurement. Precision reveals how closely clustered together are the estimations (Dragomir et al. (2009)). High precision means narrower confidence bounds.

For a prediction model to be acceptable, high accuracy and precision is required but not too high to allow generalization for our model. Saxena et al. (2010) try to put some order in the field of prognostics performance metrics. They list a number of performance measures associated to the latter notions.

Sources of uncertainty: Baraldi et al. (2013) summarize sources of uncertainty that appear during RUL prediction in mechanical devices.

i) Randomness in the future degradation dynamics of the equipment. This randomness stems from changes in contextual parameters, such as load, temperature or the appearance of a sudden failure.

ii) Modeling error, i.e., inaccuracy of the prognostic model used to perform the prediction. Model parameters are trained on a set of past data. Every mechanical component is in principle different from any other and this variation is modeled as uncertainty in model parameters.

iii) Uncertainty in current and past equipment degradation data, which are used by the prognostic model to elaborate the RUL prediction. These data are usually acquired by sensors with some measurement noise or derived from diagnostic systems assessing the equipment health state with some degree of uncertainty.

iv) Other sources of uncertainty relative to the failure threshold (which has been covered earlier) or the point of degradation initiation.

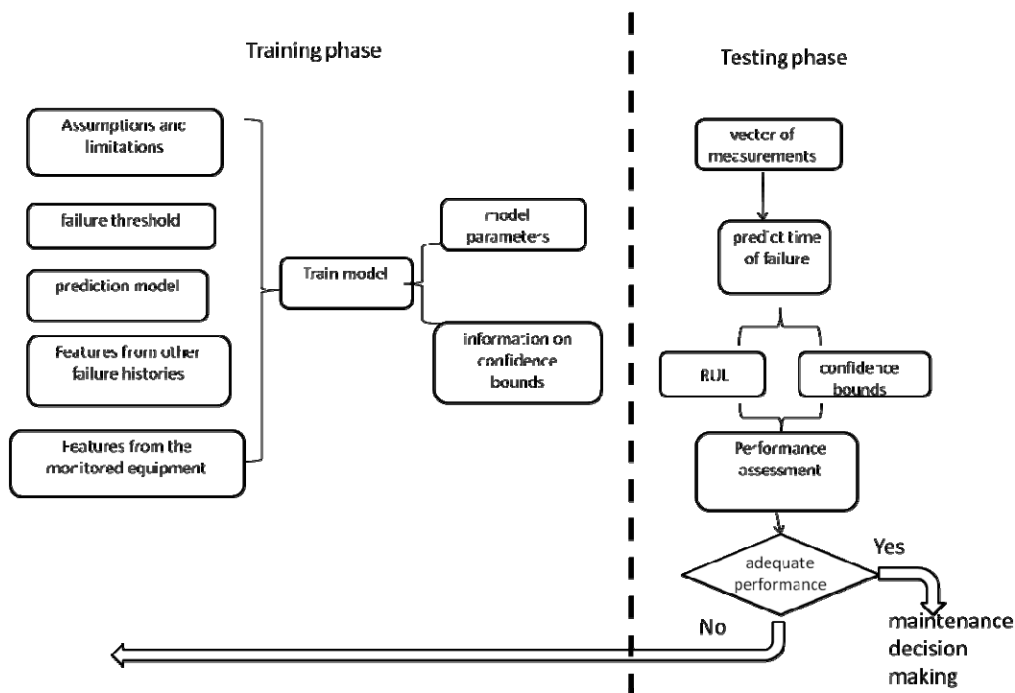


FIG. 3.2. A training/testing scheme for an on-line prognosis model.

3.2 Literature review

The last few years numerous researchers have published works on RUL estimation of industrial assets. Also numerous classifications have been proposed regarding these models such as the one proposed by Sikorska et al. (2011). The following review will roughly follow the categorization of Sikorska.

One approach to the problem of life prediction is by building a model on the past measurements of solely the component under study. This is a reasonable approach. Every mechanical component is unique from a structural point of view. This has the implication that small structural differences may induce different scales of the health metric which can limit the generalisation of a prognostic model. Since we are concerned with stochastic CBM

variables the simplest model to explain the timeseries is a linear autoregressive model. Normally the stochastic feature has a non-stationary trend which can be included in the linear model therefore yielding an ARIMA (Auto-Regressive Integrated models with Moving Average) mathematical structure. The number of model coefficients is decided with various information criteria such as Akaike Information Criterion (AIC) or Bayesian Information Criterion (BIC) (see [Fassois \(2005\)](#)). The final estimation of model parameters can be made with the well known Yule-Walker equations on the training data set. [Tse and Atherton \(1999\)](#) experimented with ARIMA models. They moved even further to prove that for non-linear timeseries other mathematical structures, namely Feed Forward and Recursive Artificial Neural Networks (FFANN and RANN), performed better in terms of prediction accuracy than ARIMA models. The monitored component of their case study was defective gearbox of a compressor in a chemical plant running for 140 days passing from “normal operation” to “defective operation” and finally “critical failure”. In both cases the one step ahead and two step ahead prediction of the monitoring variable value was adequate. [Vachtsevanos and Wang \(2001\)](#) have compared an ARIMA model with a Wavelet Neural Network (WNN) regarding bearing failure prognosis. In this work a failure threshold was placed and an actual RUL estimation was derived. They showed that the WNN performed better than ARIMA in accuracy but only for the very final stages of the bearing life where the degradation dynamics are changing non-linearly.

[Roulias et al. \(2011\)](#) have experimented with ARMA/GARCH multi step-ahead prediction model comparing it with an FFANN scheme. In their work they showed that the former required half the number of parameters to produce an equally accurate model compared to the latter. [Roulias et al. \(2012\)](#) combined FFANN and embedding dimension theory to develop an optimal multi step-ahead prediction scheme on a gearbox condition metric. They also showed that different operating regions correspond to different optimal step-ahead prediction models. [Wang et al. \(2004\)](#) moved even further to propose an Adaptive Neuro Fuzzy Inference System (ANFIS) for predicting the value of the monitoring parameter. Their model proved more efficient in adapting to sudden changes in health dynamics than RANN provided sufficient training data were provided. Their case study comprised various seeded and natural pitting faults on a meshing gearbox.

The methane pump lifelong monitoring stemming from the work of Bo-Suk Yang was the basis for a number of works relating to feature value prediction. [Pham and Yang \(2010\)](#) combined ARIMA with heteroscedasticity to perform prediction. Heteroscedasticity assumes a timely correlation not only on the estimated value average but also the underlying noise variance. [Gang and Yang \(2009\)](#) applied Demster Shafer regression for multistep ahead prediction. [Liu et al. \(2013\)](#) applied probabilistic ϵ -Support Vector Regression for condition metric value prediction along with meaningful confidence bounds. The component under

study was derived from a nuclear factory facility. They tried to predict the future index value along with confidence intervals 1 hour ahead in the future.

The latter researches assess the ability of data-driven machine learning methods to capture slowly or fast varying dynamics of an operating component. Since the output of all these models is neither RUL nor T_f estimation the question is what is the utility of these models, apart from being very interesting academic exercises? If the underlying parameter is not guaranteed not to scale among failure histories, what is the use of knowing the value of this metric. Moreover, what is the point of knowing the value of a parameter in the future in the absence of some failure threshold? Finally, is there any point in training a model on “healthy” measurements and expect to predict any future progression.

One answer is to train a model on healthy measurements and based on this model to measure the deviation at each time in the future. When this deviation is large enough (i.e. $DEV > 2 * \text{model variance}$) for several continuous samples then this is a confident indication of damage initiation. This is a task that almost all the latter models can perform and lies in the field of novelty detection or hypothesis testing.

It is worth mentioning that one weakness of the latter procedures is that long term predictions are made based on other intermediate predictions by propagating the solution until it reaches a “failure” threshold. On the other hand, as pointed out by Enrico Zio⁷ the industry is becoming reluctant to prediction methods that involve intermediate predictions for RUL estimation. In general it is the view of the author of this thesis that the utility multistep ahead prediction models are limited and in most cases serve as academic exercises.

Many works have taken into account a state-space representation of the degradation process accompanied with a failure threshold. These assume an underlying state variable x that is non-observable and it may (Orchard and Vachtsevanos (2009), Baraldi et al. (2013)) or may not (Gasperin (2011a, b), Miao et al. (2013)) be explicitly connected to an actual physical process. The feature measurement y is related to the underlying process. A general representation of a state space model is depicted in Eq. 3.1.

$$\begin{aligned} \mathbf{x}[n+1] &= f(\mathbf{x}[n]) + \mathbf{w}[n] \\ \mathbf{y}[n] &= g(\mathbf{x}[n]) + \mathbf{q}[n] \end{aligned} \quad (3.1)$$

In the latter relationship the first equation is considered the –hidden- state transition relation and the second is the observation equation. $x[n]$ is a markov process conditional on $x[n]$ (high correlation of length “1”) and $y[n]$ ’s are independent with each other and only dependent on $x[n]$.

⁷ Personal communication, PHM conference, Milan, September 2013.

Gasperin et al. (2011a) have assumed a linear state space model (Kalman filter). They assumed an idle, 2-D, state variable and connected it to the vibration RMS measurement. Therefore f becomes a 2x2 matrix A and g a 2x1 vector C . They also assumed w a 2-D and q 1-D zero-mean gaussian noise, the two mutually independent. The component under study was a gear operating under constant conditions with a configuration for accelerated fatigue pitting of the gear. The model matrixes and variances were optimized each time a new measurement was acquired with the EM algorithm. The particular realization for Kalman filtering was the Rauch-Tung-Striebel (RTS) algorithm. The algorithm provided A , C and covariance matrices estimations and was applied on a sliding window of specific length every time a new measurement was obtained yielding a time updating linear model. The estimation of RUL was based on a predefined threshold on the vibration RMS measurements. Their predictions naturally depended on the linearity of the measurement parameter and they were better near the end of the asset's operating life. The respective confidence intervals were calculated with Monte-Carlo sampling from the posterior of RMS implying that they drew lines with the slope given by vector C and the initial point given by the posterior given by the variance of q (probably with a version of Matlab© *randn* command). Moreover they don't justify the necessity for a Kalman filter compared to the application of a simpler ARIMA model on the measured parameter. Gasperin et al. (2011b) experimented on variable load conditions the latter procedure on the same test rig testing providing adequate results.

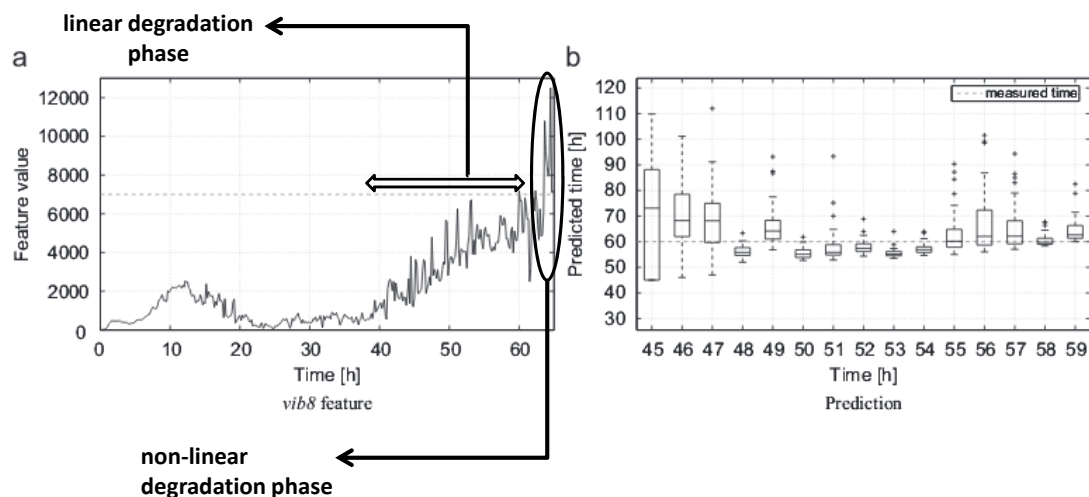


FIG. 3.3. A typical vibration RMS signature for a gear failure history taken from Gasperin et al. (2011a). His predictions are accurate after 3/4 of the gear life. In fig. 3.3 a linear and a short, non-linear, degradation phase is denoted by us. In Chapter 4 gear failure experiments will exhibit a similar pattern.

Reuben and Mba (2014) have applied Kalman filtering on bearing RUL forecasting. They divided the failure histories into three regions of degradation and incorporated this

knowledge into RTS algorithm. This made the prediction model to adapt better to the particular failure process. Their prediction results showed adequate accuracy.

Gebraeel and his group have also experimented on bearings (Gebraeel et al. (2005), Gebraeel et al. (2009)). They also chose to experiment with vibration RMS as the health index. This time they incorporated knowledge of historic failures having the luxury of 50 bearing failure histories. The experiments took place on a specially designed bearing fatigue degradation rig. They accepted that different bearings exhibit different failure modes and therefore conducted classification of degradation histories prior to testing each corresponding prediction model. In this direction they experimented with a) a linear model b) an exponential model with two parameters and a gaussian noise component whose logarithm is zero mean gaussian noise with some variance σ^2 and c) an exponential model with two parameters and a noise component whose logarithm is correlated gaussian noise (Brownian) with zero mean and linearly increasing variance ($\sigma^2=t.\sigma_0^2$). Models b) and c) are specifically designed to meet certain degradation modes with high efficiency or at least this is what Gebraeel alleges. The parameters value and their variance - for each model a), b), and c) - were calculated on a *training* data set with a simple interpolation scheme. At each *testing* failure history the respective RUL was calculated along with confidence intervals, in closed form, applying ARIMA model (a), an exponential model with adapting parameters (b) and Brownian motion model (c). The question is, why choose a model with Brownian noise. The answer is that certain bearing life histories did not exhibit a strictly monotonic degradation until failure. This element of bearing degradation is also pointed out in PRONOSTIA bearing data base (see in later paragraph of this section and Chapter 7 of this thesis). The respective variances were updated via Bayes' equation. In the process they contributed closed form variances for RUL estimation provided a specific failure threshold. Moreover, they proved that in model a) assumption, prior (training set) variance and measurement (testing bearing) variance did not yield any correlation regarding the final a posteriori stochasticity of the model parameters. In other words the final uncertainty in terms of RUL variance was just the sum of the latter variances (training histories and testing bearing). On the other hand correlations emerged as cross terms in the calculation of RUL variance in the case of model b) and c). Finally with the set of the latter assumptions they proved that bearing RUL distribution followed a Bernstein pdf. Of course they were not the first to apply the Bernstein model for confidence interval prediction. In fact they built on the work of Ahmad and Sheikh (1984) who were probably the first to derive such a relation. They also studied Bernstein pdf in quite an extend.

In general, Gebraeel's work was very well established mathematically yielding numerous measurements of uncertainty in elegant closed forms for each of the (a), (b) and (c) cases. On the other hand:

- i) the model success depended on the particular bearing degradation mode.

ii) the resulting extrapolation model posterior parameters were only slightly (less than 5%) affected by the testing bearing measurements. In a real application we would expect the on-line measurements to have more impact on the model parameters.

iii) They (along with numerous other researchers) assumed gaussianity of the latent parameters throughout their study which can be kind of restricting regarding confidence intervals. In fact, an underlying heavy tail measurement distribution of parameters may lead to underestimation of confidence bounds.

vi) The resulting RUL is, similar to Gasperin, better when closing to the asset's end of life (closer than half way to failure time). Moreover the respective confidence intervals in most cases are so wide they are practically obsolete.

The same group has applied ANNs (Gebraeel et al. 2004) on the same data base. They used characteristic frequency amplitudes (see Section 2.3) to create a metric that combines information from various bearing failure modes. The predicted T_f was again satisfactory.

Di Maio et al (2012) have also worked on the bearing database of Gebraeel. They have applied Bayesian relevance vector machines (RVM) to add sparsity to an exponential extrapolation model. This way they decreased the model complexity without an impact in its efficiency. Moreover, assuming Gaussianity, confidence intervals are added to RUL prediction. However, the confidence intervals were too narrow. Contrary to the Gebraeel's work they only used past measurements from the particular operating component and did not incorporate any prior distribution on their exponential extrapolation model parameters.

Si et al. (2013) have also experimented on bearing prediction degradation. Their model predicted the drift of a gyroscope in degrees per hour, due to bearing fatigue. The gyro drift stems from an integration and therefore it is naturally modelled as a Wiener process. In the particular work both the drift coefficient and the process noise component were assumed as random walk variables. The most interesting attribute of this research is the comparisons with other Wiener degradation models, an element rarely found in literature. On the other hand the application of the particular gyro is mostly related to weapon system application which is sad.

Aging theory has been applied for RUL estimation in many instances. Its most notable utility is in cases that a component undergoes several state transitions before indicated completely failed. Tobon-Mejia et al. (2012) and Medjaher et al. (2012) experimented on PRONOSTIA FEMTO-ST data base. In their work they adapted a Mixture of Gaussians Hidden Markov⁸ Models (MoG-HMMs) for RUL prediction along with the respective confidence intervals. MoG-HMMs are used to model non-cyclostationary stochastic signals

⁸ A system that has the Markov property is a discrete time system whose dynamics are controlled by the relation $P(X[n+1] | X[n], X[n-1], X[n-2], \dots, X[1]) = P(X[n+1] | X[n])$. In other words the transition from one system state to the other does not depend on other previous states or the system has memory "1".

and they resemble Kalman filtering, the difference being that the underlying state variable x (see Eq. 3.1) is not stochastic ($q=0$) and takes values from a set of discrete values, the “state vocabulary”. This way they added more sophistication in the extraction of the degradation signal. What they did was the following:

A few bearing failure histories were kept for training. On each failure history an HMM was adjusted yielding an optimal state sequence, state duration distribution Gaussian kernel parameters and state transition probabilities. All the latter parameters were calculated based on Expectation Maximisation of the corresponding likelihood function. The number of states was defined from expert knowledge. In other words each Markovian state was connected with an a priori known bearing degradation level. Then the mean and variance of each state is evaluated over all failure histories. On the testing history they performed MoG clustering at each time sample. Once the current state is defined the mean and variance of the remaining state sequence is evaluated to produce the final RUL. This is calculated according to aging theory equations with the state transition probabilities provided from the HMM training process. Their final results are not very accurate. The quality of their results is limited by their database which is in general very inhomogeneous. Moreover, the use of Markov property is not sufficiently justified for their case study and in what way their approach differs from a simple MoG clustering approach plus a aging model for RUL prediction. We can guess this is for the same reason Gebrael has assumed Brownian noise. Bearings (and PRONOSTIA in particular) tend to have abnormal degradation trends as will be highlighted in Chapter 7 of the current thesis. This element is not sufficiently covered in the latter research but it was clarified after personal communication with the group⁹.

Dong and He (2007) also utilised aging models for RUL estimation. Contrary to the implementation PRONOSTIA group, this time the state transition probabilities were not implicitly defined from the forward backward equations but through hypothesis testing. Their study assessed a back hoe pump under various levels of degradation. Discrete and distinguishing degradation levels were defined during the training phase. This was done through supervised training of HSMM models on the raw vibration waveform responses. In the next step of the training phase, the derived HSMM (Hidden Semi Markov Models) were used for hypothesis testing (*supervised* classification) on a number of full failure histories. The hypothesis test was performed at every raw waveform via the likelihood test. A Markov left to right (constantly degrading) transition model was assumed on the identified state sequences. The mean state duration and corresponding variances were calculated on the corresponding a training set.

⁹ Personal communication with Medjaher in PHM conference, September 2013, Milan

Hypothesis testing took place at a testing failure history. After the degradation level was identified aging model was applied (as above with [Tobon-Mejia et al. \(2012\)](#)) to measure the expected RUL mean and variance. It repeated that HMM and their generalisation, HSMM, was adjusted on the vibration waveforms which, for the case study, were *non-cyclostationary*. Therefore the latter diagnosis-prognosis method cannot not be applied on bearing or gearbox vibration responses. [Peng and Dong \(2011\)](#) used the same data set to adapt a variety of time dependent Markov transition matrices. The final RUL with time dependent Markov transitions was (naturally) found superior to the one calculated in [Dong and He \(2007\)](#). [Moghaddass \(2013\)](#) does the same as the latter researchers only in an unsupervised manner. It should be mentioned that he performs all his tests on simulated degradation data of an aircraft turbomachine. The simulation code was based on physics first principles and was borrowed from NASA Prognostic Data Repository.

In the last decade a lot of work has been focused on Particle Filtering (PF) techniques for prognosis in engineering systems. The motivation was provided by the excellent tutorial on PF by [Arulampalam et al. \(2002\)](#). In this work stochastic models were assumed of the form of [Eq. 3.1](#). It was proven that PF gave more accurate predictions regarding the underlying degradation state in comparison to Kalman filtering predictions. It should be mentioned though that in his experiments he studied explicitly non linear state-space models and the findings are not generally applicable. In other words whether PFs are worth applied or not should be validated at each case study.

The main novelty of PF techniques is that instead of using closed form gaussian distributions to evaluate uncertainty intervals, discrete realisations of the underlying distribution is applied. To this end a number of “particles” are used to discretise the distribution interval. A weight is appointed to each particle to simulate the probability intensity at each discrete point. A Bayesian update equation and an effective sampling-resampling process makes possible the effective projection of point predictions as well as valid confidence intervals.

The most prominent application of PFs was in [Orchard and Vachtsevanos \(2009\)](#). They predicted the RUL distribution for a growing crack in a planetary gear plate in an operating UH-60 helicopter. For this purpose they seeded a crack on the planetary gearbox under consideration. To identify and predict the time of failure they applied

- i) an ANSYS model to accurately update on-line (or whenever deemed necessary) the state model variables based on the applied load.
- ii) a strain gage on the induced crack to accurately measure its length
- iii) a (Paris-Erdogan) state Markov model which is *linear* with respect to the length of the crack.

iv) the non linearity was introduced on the observation equation. The observation sequence corresponded to an extracted vibration feature. At each particular time intervals a strain gage was sampled to account for any misinformation of the vibration condition monitoring feature.

Their results were very interesting. Their predictions were accurate regarding the length of the crack as well as the RUL of the component. Moreover, they proved that a PF application was inferior to Kalman filtering application. On the other hand, with the application of so many subsystems (i to iv) it would be a miracle not to catch the exact crack evolution. It is worth mentioning that a part of the study took place on synthesized crack growth data.

[Sun et al. \(2012\)](#) have also experimented with PFs this time assuming a simple linear state-space model. The failure timeseries were obtained from C-MAPSS software simulating turbofan degradation during operation, a software that has been also been incorporated in the work of [Moghaddass \(2013\)](#).

Enrico Zio and his colleagues have attributed a great deal in PF Sequential Monte Carlo analyses. [Baraldi et al. \(2013a\)](#) have explored bootstrap methods and compared them to PF approach for RUL estimation of a turbine blade based on synthesized creep data. [Baraldi et al. \(2013b\)](#) and [Zio and Pelsoni \(2011\)](#) borrowed the synthesis model of Orchard and tested a logit model and an ensemble of ANNs respectively to better capture the non-linearity of the observation relationship. [Wei et al. \(2013\)](#) tested the application of ϵ -SVR to account for the non-linearity of degradation processes. They tested their model on simulated data, car engines reliability data and turbochargers in diesel engines.

In general ϵ -SVR has been explored recently for mechanical systems RUL estimation as already shown by [Wei et al. \(2013\)](#) and [Liu et al. \(2013\)](#). [Benkedjouh et al. \(2013\)](#) have explored the latter method on their data base (PRONOSTIA bearing data base). They extracted wavelet packet energy percentages at each band to use as a feature vector and applied a non-linear feature fusion technique (ISOMAP). [Loutas et al. \(2013\)](#) have also experimented on PRONOSTIA data base. They trained a probabilistic ϵ -SVR model and adapted confidence intervals with Bayesian equations and gaussian priors. They tested its efficiency on an unseen experiment on RUL prediction along with its distribution. The quality of their results depended on the similarity between the training and testing experiments.

Cox proportional hazards model ([Cox \(1972\)](#)) has been rarely applied for actual predictions -- a brief presentation of these models is going to be provided in [Chapter 5](#)--. The outcome of this method is the calculation of the hazard and reliability function at every step of the asset operation. The latter are probability measures met in Reliability Analysis (RA). RUL is indirectly found from these measures contrary to the methods that have been disrobed until this point. [Tran et al. \(2012\)](#) have experimented with Cox model on the methane pump

failure data base of Bo-Suk Yang. In their work they state that their model predicted the RUL of the particular asset. What it does though is to predict the value of the reliability function and not the actual time to failure. Moreover their model is i) very elaborate, ii) it uses multistep ahead prediction which, as has been mentioned earlier is not considered a reliable strategy to make predictions and iii) it is only useful in the very last stages of the asset degradation. To be just to the group though, it must be mentioned that the data base of Bo-Suk Yang is very difficult to work with and probably its value lies in that it is obtained from a real life, critical, industrial operating.

[Banjevich and Jardine \(2005\)](#) built on the findings of [Jardine et al. \(1987\)](#). The latter had already devised a method to compute the hazard from spectrometric oil analysis of a gas turbine based on a Weibull proportional hazards model. The former assumed a markov chain on the sequence of failure states and indirectly estimated the RUL of the turbine. [Ghodrati et al. \(2012\)](#) are some of the few to apply Cox's model for gradually degrading mechanical equipment. Their focus was on the hydraulic system of a load-haul-dump. They used three external covariates for their model namely dust, temperature and skill of the operator –it is unknown how this latter could be measured--. As a next step they applied a Weibull distribution as their base distribution function with parameters derived from expert knowledge -- rather than implicit computation from a maximum likelihood scheme for instance. Their results were adequate. [Nystad \(2012\)](#) has experimented with natural gas export compressors. In his aging models he incorporated an appropriate technical condition index. The corresponding index is a smoothed compound of several covariates from bearing monitoring techniques. Moreover he incorporates replacement actions to bring his model closer to actual industrial conditions. His model gives adequate RUL predictions though without providing error bounds on his aging model. He only places error bounds on a condition based only prediction model that he fits on past failure histories with a power law trend analysis plus a gamma prior. The additional model is trained on the technical condition index.

3.3 Discussion

This thesis' focus is on RUL point estimation along with error bounds. The focus is on important rotating machinery parts, namely gearboxes and bearings. In this direction three predictive models are going to be assessed for mono and multi-parametric RUL prediction. All proposed models are deemed state of the art and beyond. To make valuable assessments on these models a pool of failure histories is needed. Failure histories are difficult to attain though when referring to industrial components. The exact replication of such events in laboratory is insufficient by nature and expensive in energy as well as in capital. Therefore, many works have experimented with replicated or simulated data, some of them depicted in

table 3.1. These groups have bypassed the lack of actual data by simulated measurements. The measurements are taken either from implied stochastic update models or FEM software packages. All models were chosen so that they were as close to the physics of the degradation process as possible. Motivated by these efforts, the author of this thesis will contribute a synthesis framework for gearbox failure data (Chapter 4). This framework will be applied in order to assess the three proposed predictive models (Chapter 5).

Research group	Replicated	To test
Orchard and Vachtsevanos (2009)	A linear gamma process with non linear observation relation	PF with explicit dynamic model relation
Zio and Pelsoni (2011)	Borrowed from Orchard and Vachtsevanos (2009)	A logit dynamic model with PF
Wang et al. (2011)	Temperature of a nuclear reactor tank (completely non linear dynamic relations)	Linear stochastic model with correlated (Brownian) noise component with an updated drift coefficient
Sun et al. (2012)	Damage level of an airplane turbo engine (software C-MAPSS borrowed from NASA repository)	Application of PF and Linear state-space combination for RUL prediction
Moghaddass (2013)	Damage level of an airplane turbo engine (software C-MAPSS borrowed from NASA repository)	Time dependent Hidden Semi-Markov state sequence modeling for prediction
Baraldi et al. (2013)	Creep of a turbine blade	i)PF with explicit model relation ii) exponential degradation model
Zhang et al. (2014)	Vibration level of a degrading rolling bearing	No application

Table 3.1. Research efforts regarding prognosis on simulated data.

Cox regression models are rarely applied in literature. Reasons for this could be their indirect manner of computing RUL of a device, the lack of a means to compute confidence bounds on the prediction and the scarcity of reliability data, an element crucial for these

models to provide a valid prediction. This thesis will contribute in this direction with a model trained on a single degradation parameter for RUL prediction. Moreover a method for error bound estimation is going to be proposed for the first time to the author's best knowledge. The whole methodology is going to be applied in gearboxes for the first time, to the author's knowledge.

Support Vectors was the hot spot of prognosis in 2013. The author's research group contributed to this direction and particularly in bearings prognostics with a methodology for RUL estimation along with error bounds ([Loutas et al. \(2013\)](#)).

Extrapolation models are rarely studied in literature. However industry seems to trust these more than propagation models i.e. multistep ahead prediction models or machine learning regression models. The author proposes a simple bootstrap extrapolation method which, linked to a feature fusion block, can provide adequate predictions along with error bounds for gearboxes in total absence of a failure prior. The feature fusion block is based on Independent Components Analysis (ICA) studied for the first time for its prediction potential in gearboxes. A number of results were published in [Loutas et al. \(2011\)](#). Finally the proposed feature extraction method along with machine learning regression model is going to be studied for the first time in bearings in [Chapter 7](#).

CHAPTER 4

Gear Failure Experiments

4.1 Introduction

As shown in [Chapter 3](#), few scientific works are concerned with multi-hour gear fatigue experiments. This fact was the motivation for developing a gearbox fatigue set up for multi-hour fatigue experiments. In the current chapter the multi-hour test set up is going to be presented along with a summary of the obtained sensor recordings.

Multi-hour gear experiments are a daunting task. They require effort and intensive care and they are expensive in capital and energy. On the other hand a study of prognostic techniques is insufficient without an adequate pool of failure experiments. To bypass this shortfall a framework for simulating failure signal timeseries is proposed which is built upon actual experimental recordings. This is motivated by a number of similar works in the field of machine health prognosis ([Section 4.8](#)).

In general, the organization of this chapter will be as follows.

Quoting [MacKay \(2003\)](#) “one must make assumptions in order to make inferences”. The set of assumption for this thesis and this chapter in particular is going to be presented in

[Section 4.2](#). [Section 4.3](#) provides a brief presentation of the experimental set-up. Prior to gear failure tests it is deemed necessary to examine the effect of contextual parameters on the sensor recordings. An overview of these effects is provided in [Section 4.4](#). The effect of contextual parameters will help in the definition of the gear fatigue test configurations. In fact a number of failure tests took place in the framework of this research. Representative recordings are going to be presented in [Section 4.5](#) Since these data are going to be used in a prognosis module a raw data preprocessing step is essential. This preprocessing is presented in [Section 4.6](#). Correlations among sensor recordings are discussed in [Section 4.7](#). Finally a gear failure simulation framework is going to be presented and justified in [Section 4.8](#). This simulation scheme will be utilized in the prognosis study of [Chapter 5](#). The most important points of this chapter are going to be repeated and highlighted in [Section 4.9](#).

4.2 Assumptions

This section provides the set of assumptions that will define the scope of this thesis. Assumptions related to notations are included in the Acronyms & Abbreviations section. It should be noted that “device” and “component” are going to be used interchangeably in the text and they are going to signify a single component (a spur gear or a bearing element).

- As has been recursively mentioned during the first three chapters of this dissertation the subject of the thesis is rotating systems (gearboxes and rolling bearings) with a focus on health prognosis.
- Within the thesis framework these devices are supposed to operate under constant external conditions. More specifically constant load is going to be assumed on the axis plus constant rotating speed and constant lubricating oil temperature.
- All devices under consideration are going to operate from healthy condition to complete failure.
- The devices are non repairable during operation.
- Only fatigue degradation is going to be assessed. This is found in literature as “soft failure”. Sudden jump from perfectly normal to critically damaged condition (for instance broken gear tooth incident) are not going to be assessed in this thesis. The main reason is that their assessment requires reliability data regarding their appearance and such data were not available in framework of this thesis.
- The device is under condition monitoring. A condition indicator is recorded at regular intervals. As described in [Chapter 3](#), these intervals should be much shorter than the total operation life of the asset. The condition or health indicator can be a single value or a combination of features. The health indicators are most commonly statistical

metrics such as rms, kurtosis etc. They are in general correlated to the health level of the asset and are of stochastic nature.

As has been mentioned, all these assumptions will hold throughout this thesis unless stated otherwise.

4.3 Experimental set up

To the author's knowledge no free-access archives of actual multi-hour lifelong gearbox experiments exist. There is an important reason for this. Lifelong gearbox experiments are very expensive in capital as well as in energy. This lack of multi-hour data histories motivated the development of the current experimental set up and an overview of the gear failure data base is going to be provided in [Section 4.5](#).

[Figure 4.1](#) shows the experimental setup used for the gears testing. The test rig consists of two gears made from 045M15 steel with a module of 3 mm, pressure angle 20° , which have 53 and 25 teeth. The driving/master gear had a face width of 5mm and the driving/slave gear a width of 2mm. The purpose of this configuration was to simulate accelerated gear fatigue. The axes of the gears are supported by two ball bearings each.

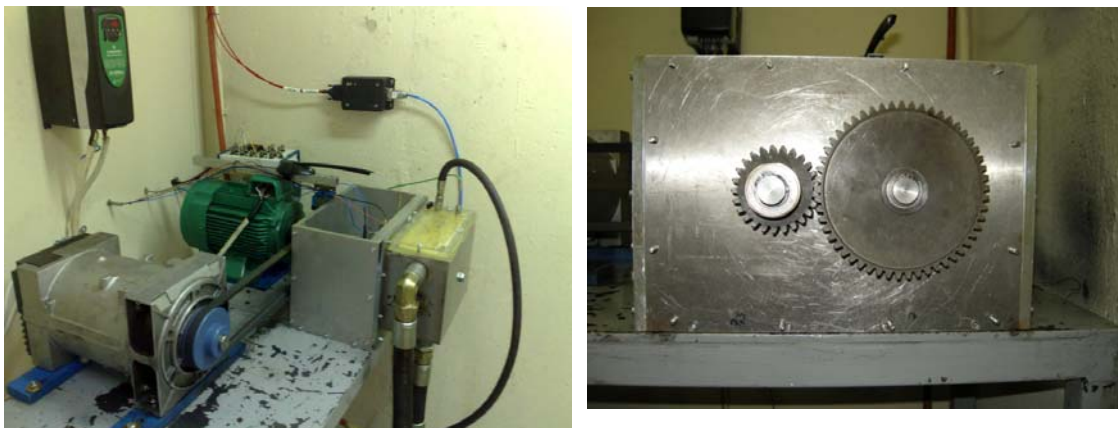


FIG. 4.1. Photographs of the gearbox fatigue test rig. On the right, a close up of the master/slave gear pair.

All the above are settled in an oil basin in order to ensure proper lubrication. The gear box is powered by a motor and consumes its power on a generator. Their characteristics are as follows:

- 1 stage gearbox with two gears (25 and 53 teeth)
- 3- phase 5 hp motor (220V, 9A, 50Hz, 1400 rpm) controlled by inverter
- Single phase generator with continuous power consumption control (4.2 KVA, 3000 rpm, 50Hz).
- The oil pump is of wet type without oil recirculation
- The shafts are ball bearing supported.

This setup is an evolution of the setup used in a previous work (Loutas et al. (2008)). The oil output at the bottom of the gearbox case provides more realistic ODM measurements as the debris circulation is ensured. Moreover simpler and faster open-close procedure is accomplished.

Three non-destructive techniques monitored the gearbox during operation namely vibration, acoustic emission and oil debris monitoring. A 3-axis Bruel & Kjaer accelerometer was used for the vibration monitoring both mounted upon the gearbox case. The sampling frequency used was 50 kHz and recordings of 1 sec duration were obtained. Three acoustic emission sensors by Physical Acoustics Corporation (USA) with a frequency range of 100-800 kHz recorded continuous AE activity at a sampling rate of 2 MHz. The AE recordings duration is 100 msec. AE as well as vibration is recorded every five minutes. A special innovative device was designed in order to mount the AE sensor (channel 2) in a rotating component avoiding thus the costly solution of the slip-ring that is typically used in literature. Its monitoring efficiency is going to be assessed in a following section.

Oil debris data were collected using a commercially available oil debris monitoring (ODM) sensor (MetalSCAN by GASTOPS USA, fig. 2.11). The sensor is in-line with the oil circuit. The oil circulation is assisted by a pump. Its operation is based on the changes in the electromagnetic field caused by the metal particles passing through the sensor. The sensor measures the number of particles as well as their size from 225 to 1000 μm approximate diameter. A filter after the ODM sensor holds all particles larger than 70 μm not allowing them to re-enter the oil circuit.

The recording of all the above data is realized by a National Instruments NI-6070 IMS/SEC FIREWIRE data acquisition card and is assisted by special software in-house developed in Labview.

4.4 Effect of contextual parameters in the sensor recordings

The contextual parameters of the set up impact the level of vibration and therefore the sensor recordings. In order to study this impact a number of experiments were conducted. More precisely four experiments took place ranging different axis load and different axis rotating speed. The particular experiment configurations are reported in Table 4.1. The results are summarized in fig. 4.2 and fig. 4.3.

Axis speed-axis load configuration	Speed (Hz)	Load (kW)	Duration of experiment
“Low”-”Low”	25	0.2	~3 hrs
“High”-”Low”	50	1	-“-
“Low”-”High”	25	0.7	-“-
“High”-”High”	50	2.5	-“-

Table 4.1. Experimental configurations.

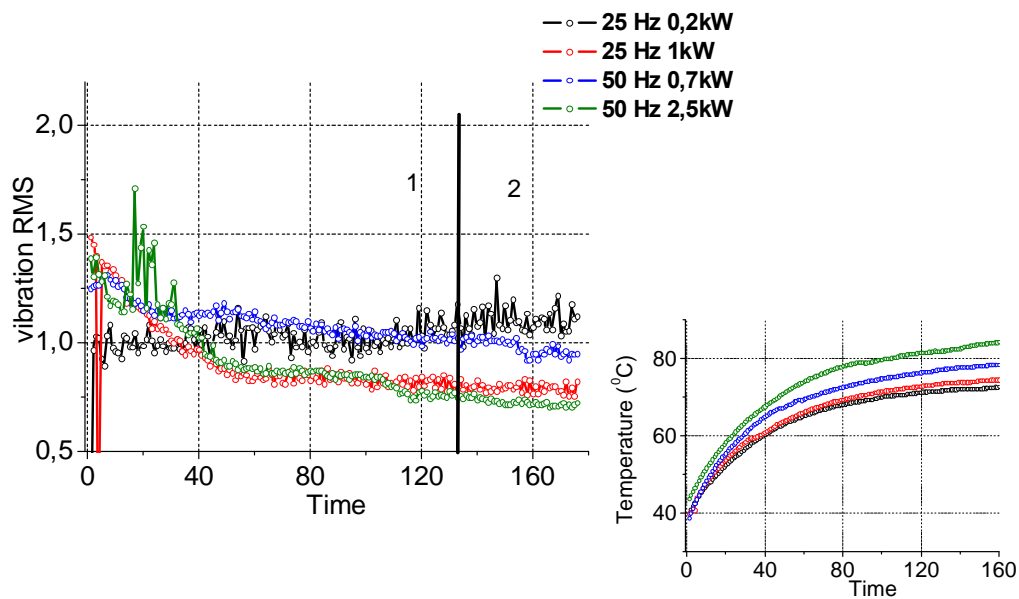


FIG. 4.2. Vibration level (left) for different set up configurations plotted along with lubricating oil temperature (right). The recordings are obtained from vibration channel 1 (x-axis vibration).

At first let's focus on the temperature recordings in [fig. 4.2](#). It is observed that oil temperature depends on rotating speed as well as the axis load. It comprises a transient phase that seems to last for two to three hours depending on the experimental configuration. The longest transient seems to correspond to the “high”-“high” configuration spanning to more than 2 hours. The steady state temperature is also dependant on the particular configuration.

Vibration level, as shown in [fig. 4.2](#), varies with axis load, rotating frequency and lubricant temperature. The latter factors are not easily de-correlated except for the beginning of the experiment where temperature is similar for all configurations. Two regions are roughly distinguished in [fig. 4.2](#) (left graph) with respect to vibration RMS. Region “1” involves the transient stage of gearbox operation. In this interval recordings are greatly

affected by the rise of the lubricant temperature, except perhaps for the “low” load - “low” speed configuration. This trend is referred in literature as run-in effect. It is evident that higher load and speed leads to more intensive run-in-effect. In the steady state condition the final level of vibration depends on the particular set up with the “high”-“high” configuration corresponding to the lowest level.

Figure 4.3 depicts AE level derived from the same experiments. The particular AE sensor is mounted on the shaft of the driving gear hopefully catching AE events from the respective gear. It seems that AE is also prone to run-in effects particularly for the high speed configurations. However it seems that the run-in effect is shorter in the case of AE spanning to approximately 28 minutes of operation in all cases. The shorter duration of run-in phase is noted as another advantage of AE monitoring compared to vibration. In the steady state (region “2”) condition the configuration with the lowest AE production is again the “high speed”-“high load” condition with little difference from the other three configurations.

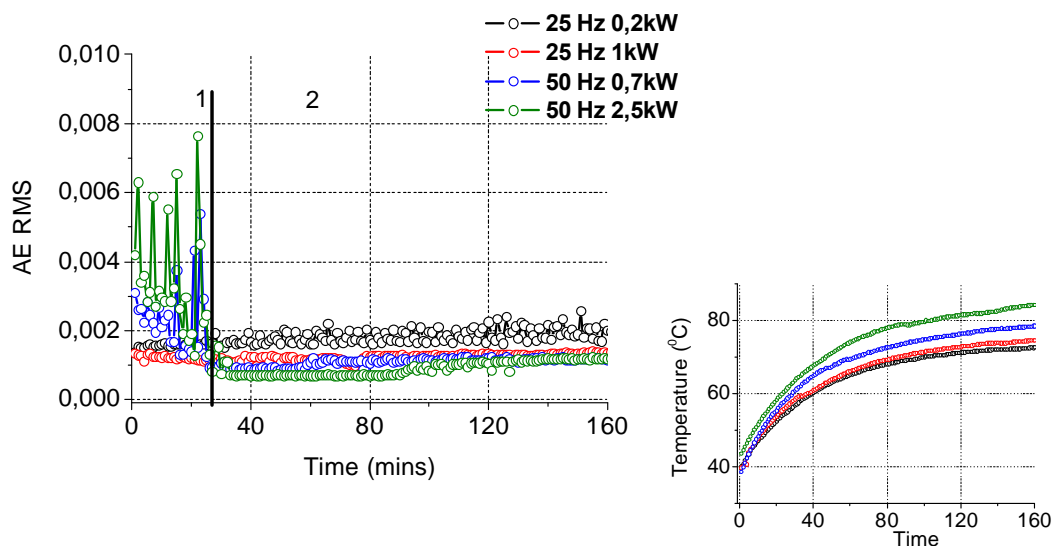


FIG. 4.3. AE levels (left) for different configurations plotted along with lubricating oil temperature (right).

The run in phase is largely due to the decrease in lubricant oil viscosity and it is described as elastohydrodynamic (EHL) regime. In EHL a very thin film of oil is formed, in the scale of few micrometers that takes part in the AE as well as vibration waveform formation. The theoretical basis for EHL and the underlying cause and effect relations are beyond the scope of this thesis.

To sum up, it seems that higher speed and load conditions result in lower levels of vibration in the steady operation state. Lower vibration level results in lower distress of the

test rig. In our case this was important since several gear fatigue test were going to take place in the particular test rig. After all, a high load profile is always desirable so that the gear teeth are in excellent contact and therefore the useful signal is correctly recorded. Also, higher load would utterly result in faster degradation of the tested gear.

Another observation is that AE recordings have a much shorter transient phase than vibration as depicted in the above figures. This is a desirable characteristic for a reliable gearbox health metric and to the author's knowledge this is the first time that this difference is being highlighted.

Following the latter study the gear failure tests are chosen to be carried out in the most extreme load and speed conditions notably the “high”-“high” configuration according to [table 4.1](#).

4.5 Lifelong gear fatigue tests

Performing lifelong gearbox experiments proved to be a daunting task. In order to study the effect of pitting fatigue the gearbox is meant to operate continuously for a few days under extreme load and rotating speed conditions. Six lifelong gearbox experiments took place in the framework of this thesis. Of those six, five were kept for further study. One experiment endured multiple DAQ malfunctions and other technical issues during its operation and was therefore excluded. [Table 4.2](#) provides a brief overview of the conducted experiments.

An overview of the experimental recordings is going to be provided in the following sections. Considering the waveform type monitoring techniques RMS, peak to peak (P-P) amplitude and PSD plots are going to be depicted. It should be mentioned that P-P amplitude will denote the maximum minus the minimum value of a waveform.

A final point should be mentioned. Regarding the conducted experiments, a minimum of start-stop events took place so that the influence of fatigue would effectively be isolated and correlated to sensor recordings. As will be shown even these few start-stop events were very influential. This is the reason why no visual control took place throughout these experiments. An effort to correlate waveform sensor recordings to the actual gearbox situation would be inefficient without any visual inspection or some other means of actual health measurement. For an indication of the actual level of degradation we resort to ODM recordings. ODM recordings from lifelong gearbox operation are rarely exhibited in literature and this is one of the important contributions of this thesis.

The particular configuration induced high vibration to the test rig. The result was that after the sixth failure experiment the rig was in a bad condition for further gear fatigue experiments.

	Total operating time	Mode of failure	Status
Experiment I	150 hours	Tooth pitting fatigue	Acoustic Emission channel 3 detachment, vibration y-axis malfunction
Experiment II	132 hours	-“-	Multiple DAQ malfunctions
Experiment III	162 hours	-“-	All sensor channels valid
Experiment IV	121 hours	-“-	Vibration z-axis malfunction
Experiment V	149 hours	-“-	All sensor channels valid
Experiment VI	92 hours	-“-	Vibration z-axis malfunction

Table 4.2. Gear failure experiments.

4.5.1 ODM recordings

Oil debris data were collected using a commercially available oil debris monitoring (ODM) sensor. The sensor is in-line with the oil circuit and its circulation is assisted by an electric pump. The principle of operation has been described in [Section 2.4](#). This sensor measures the number of particles as well as their size from 225 to 1000 μm approximate diameter. A filter after the ODM sensor holds all particles larger than 70 μm not allowing them to re-enter the oil circuit and affect the measurements. Typical recordings are depicted in [fig. 4.4](#) and [4.5](#). For the sake of conciseness only two failure histories are depicted. The observations that are made hold for the rest of the valid failure histories.

[Figure 4.4](#) corresponds to recordings derived from experiment III in particular. This experiment ended after 160 hours of almost continuous operation under constant load and shaft rotation conditions. At the final visual inspection multiple gear teeth were completely worn out.

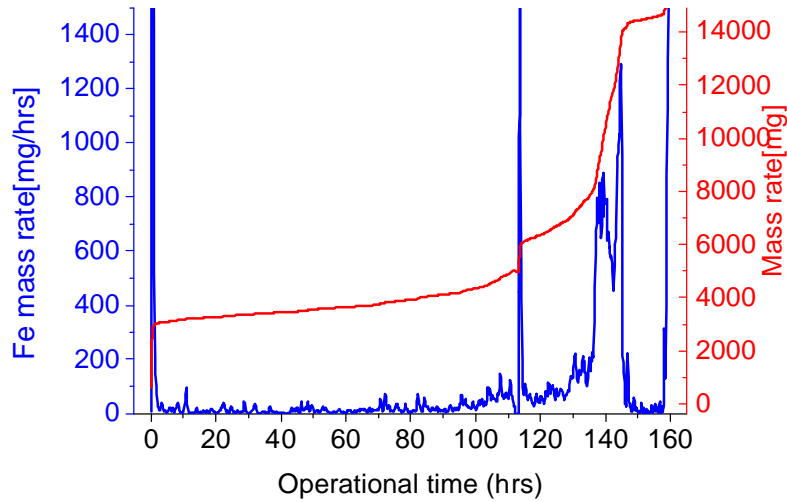


FIG. 4.4. The aggregated mass of ferrous particles and its rate of particles production for lifelong experiment III.

Stopping the machine in regular intervals for visual inspection was deemed impractical. One reason was that start-stop events had an impact at the recordings. ODM sensor in particular proved very sensitive in start-stop events. The lack of regular visual inspection is the reason that the exact point of failure was not possible to be defined with confidence. However, continuous ODM recordings can provide some insight. The detection of a critical failure point can be based on the rate of ferrous particles creation. To be more precise let's focus again on [fig. 4.4](#). Ferrous particles production rate seems to increase violently between 130 and 140 hours of operation. Until that instant though no significant change in the dynamics of particle production is noted. This is evidence that the dynamics of the degradation changes significantly and implying a point of critical failure. From this observation we can deduce that a simple heuristic threshold in Fe-mass rate could provide a confident diagnosis threshold for a critical failure diagnosis. The exact threshold value on Fe-mass rate could be defined after inspection of the five conducted gear failure experiments.

Fe aggregate mass recordings evidently possess diagnostic potential. This metric's trend seems to agree with a gradual pattern for the gearbox. Another point that should be made concerns the influence of start-stop events. In [fig. 4.4](#) two spikes can be observed on Fe-rate curve. These spikes coincide with the stop-start of the machine in the beginning of the experiment and at 113 hrs and do not correspond to any degradation mechanism or failure mode. This is a drawback of ODM sensor recordings. ODM recordings are very sensitive to start-stop events and this should be taken into account when performing a diagnosis task.

[Figure 4.5](#) corresponds to the lifelong experiment IV. This experiment ended after 92 hours of almost continuous operation under constant load and shaft rotation conditions. The

mode of gear failure was, again, fatigue wear and the history of degradation is well depicted in ODM recordings.

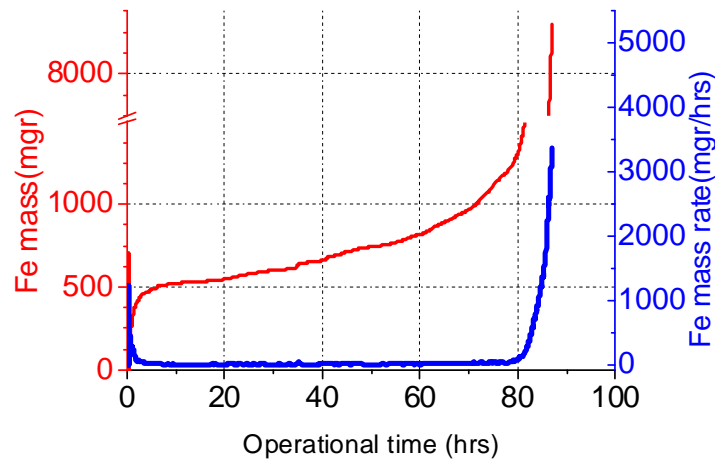


FIG. 4.5. The aggregated mass of ferrous particles and its rate of particles production for lifelong experiment VI.

This time no stop events occurred in the middle of the experiment for inspection. Experiment IV ran without stop until total wear of the driving gear. The critical failure must have occurred much earlier than the final point of inspection. Judging by the Fe mass rate curve, this could be placed as early as 85 hours of operation or even earlier.

A few conclusions can be drawn from the latter graphs. First of all, it is observed that the ferrous particle creation dynamics is almost steady until some point. At that instant the degradation gains momentum and particle creation grows hyper-exponentially. Another observation is that Fe-mass rate is an inferior change point detection method in comparison to Fe aggregate mass recording. It was very hard to place a definite failure threshold on aggregate ODM mass because it exhibited very different scales among the conducted experiments. Namely, the aggregate mass ranged from 3000 to 9000 mg in the vicinity of the critical failure point. On the other hand ODM Fe mass rate recordings scaled less from experiment to experiment and the change point was much easier to identify. It is concluded that Fe mass rate is a more confident change-point detection method than Fe mass.

Bearing in mind the latter presentation we can summarize some disadvantages relative to the method of ODM.

- It is very sensitive to start-stop events. These events should be taken into account and their effect should be decorrelated from the final diagnostic inspection.
- Fe-Mass recordings have a big scale of variability. Therefore change point detection cannot be made with confidence.

- The recordings are sensitive to the set-up configuration. Ferrous particles can be trapped inside the gearbox oil bath. In a laboratory test bench this risk can be minimized by introducing simple oil bath geometry. But in real scale applications this may not be the case and a careful mechanical design should be taken into account.
- The lubrication filter may not trap all ferrous particles sending them back to the lubrication circuit.
- It is very expensive. The particular product by GASTOPS used in the experiments I-VI costs more than 30000 Euros and this is only for a medium scale application. Moreover in a real life application the labor cost for mounting of the equipment should be included.

Finally ODM cannot capture abrupt, “hard”, structural changes. This is worth mentioning here although the specific failure mode is beyond the scope of the current study. We will elaborate our statement by presenting a single graph:

A lifelong experiment with a different configuration had taken place in the laboratory (Loutas et al. (2009)). A single gear tooth crack was simulated at the beginning of the experiment. This was done by seeding a transverse tooth crack. The goal of this experiment was to assess the ability of various monitoring methods to capture a chopped tooth event. The results are depicted in fig. 4.6.

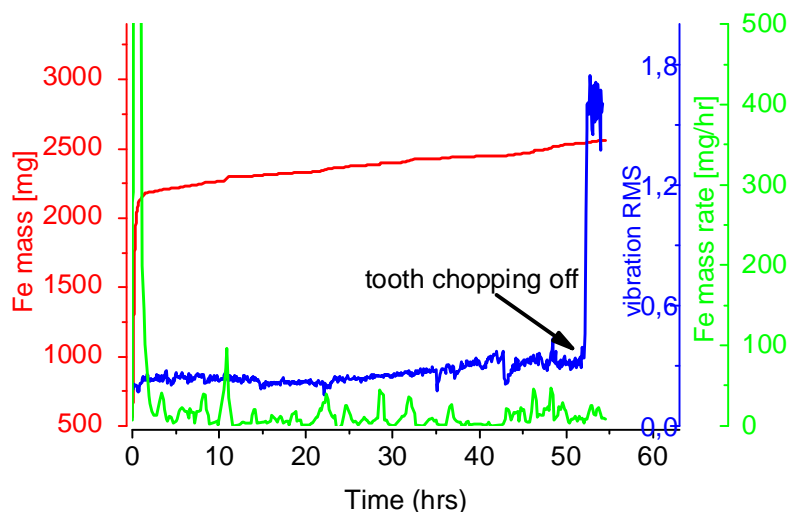


FIG. 4.6. Sensor measurements from a gear lifelong test with artificially induced gear cracks.

Figure 4.6 depicts the vibration RMS level, the ODM Fe mass and Fe mass rate recordings in a common graph. Vibration monitoring proved very efficient in capturing the chopped tooth at about 54 hours. The same stands for AE sensors. Looking at the ODM recordings though, no change is noticed at that time instant. No evident change in Fe mass

rate level or its volatility and no change to the trend of aggregate Fe mass is observed. This is evidence that ODM is not suited for sudden gearbox structural changes (“hard failure” incidents).

To sum up, ODM recordings, with all its drawbacks, is deemed an efficient means of gearbox fatigue monitoring. The CBM designer should combine ODM with some other monitoring technique to derive optimal results. This belief will be also highlighted in the following sections.

4.5.2 Acoustic emission recordings

In this section AE recordings are going to be presented. For the sake of succinctness, recordings coming from lifelong experiments III and IV (Table 4.1) are going to be depicted. To assess the efficiency of the recordings RMS and peak to peak (P-P) amplitude are going to be plotted against operational time. Moreover the spectrum of each sensor waveform will be shown whenever it is deemed necessary.

Acoustic emission channel 1

The first sensor which is going to be presented is AE ch1 sensor. It is mounted on the shaft of the driving gear hopefully catching AE events due to the gradual degradation of the driving gear. The exact point of placement was decided based on a number of Hsu-Nielsen source trials¹⁰.

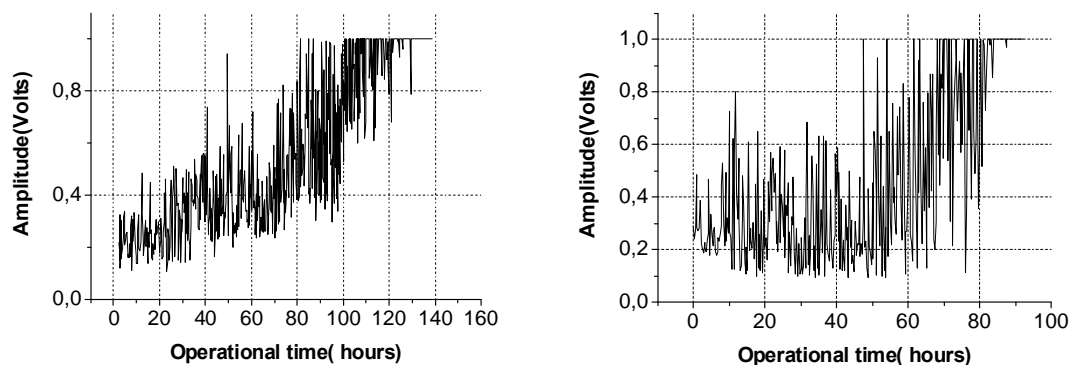


FIG. 4.7. AE P-P amplitude for experiment III (left) and experiment VI (right).

The AE channel 1 P-P amplitude can in general follow the gradual gear degradation. However its volatility is annoyingly intensive. Another point that can be made is that close to the end of the experiments, there is evidence of waveform clipping effect. This is probably

¹⁰ Hsu-Nielsen test involves breaking a pencil lead at several positions of the test rig in order to determine the optimum sensor location. The location is based on the minimum signal attenuation criterion.

due to insufficient adjustment of the AE sensor preamplifier causing the recordings to be outside the dynamic range A/D decoder.

RMS of AE waveforms (fig. 4.8) seem to be a better candidate for gearbox degradation monitoring. This fact seems to hold for all lifelong experiments. From a diagnostics and prognostics point of view there seems to be an issue with RMS histories and this is the small but apparent scale difference. RMS from experiment III ranges from 0.01 to 0.1 Volts and for experiment VI from 0.002 to 0.07. Slightly different scales were noted in all lifelong experiments.

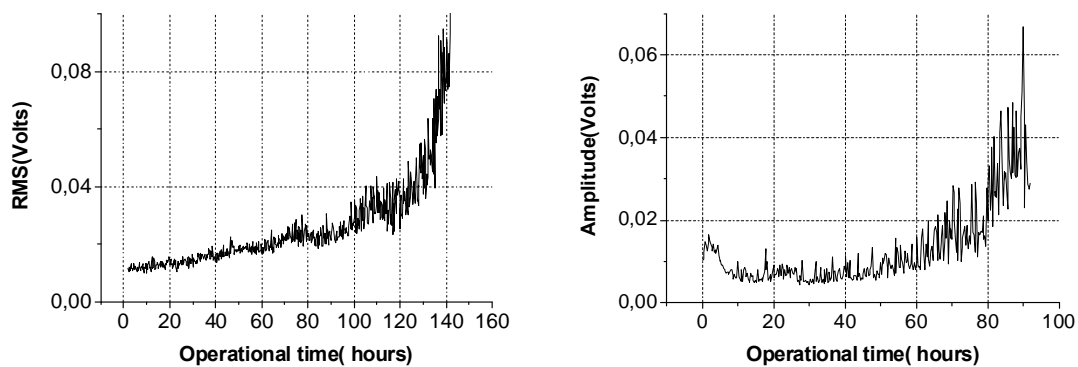


FIG. 4.8. AE RMS for experiment III (left) and experiment VI (right).

The clipping effect noted in fig. 4.7 induces undesired DC and low frequency components in AE spectrum. This is dealt with in a later time with the application of a highpass butterworth filter. The filter is designed to have a very low cut off frequency. For the case of AE this could be 10 kHz. Butterworth belongs to the class of infinite impulse response (IIR) filters and can achieve a steep cut-off frequency slope with only a small number of filter coefficients. During the design care is taken so that the choice of filter coefficients would not produce an unstable filter. Instability is a serious side effect of the IIR property of butterworth filters. The filtering process was implemented in Matlab© software using *butter* command. Figures 4.9 and 4.10 provide a few AE waveforms and their corresponding power spectral density. The waveforms are taken from subsequent time instants from the respective experiments. It should be mentioned both fig 4.9 and 4.10 depict high pass filtered waveforms as earlier reported. They comprise the expected AE form and nothing specific can be mentioned regarding the timeseries waveforms.

Regarding the spectrum of AE, it is observed that the energy is concentrated in the band 50-200 kHz. With the progression of the experiments a low energy “tail” appears in AE spectrum in the frequency range 200 kHz to 500 kHz. These tails are highlighted with an ellipse in the corresponding plots. This is evidence that the frequency band 200-500 kHz

poses diagnostic and possibly prognostic potential. This is going to be further investigated in Chapter 5.

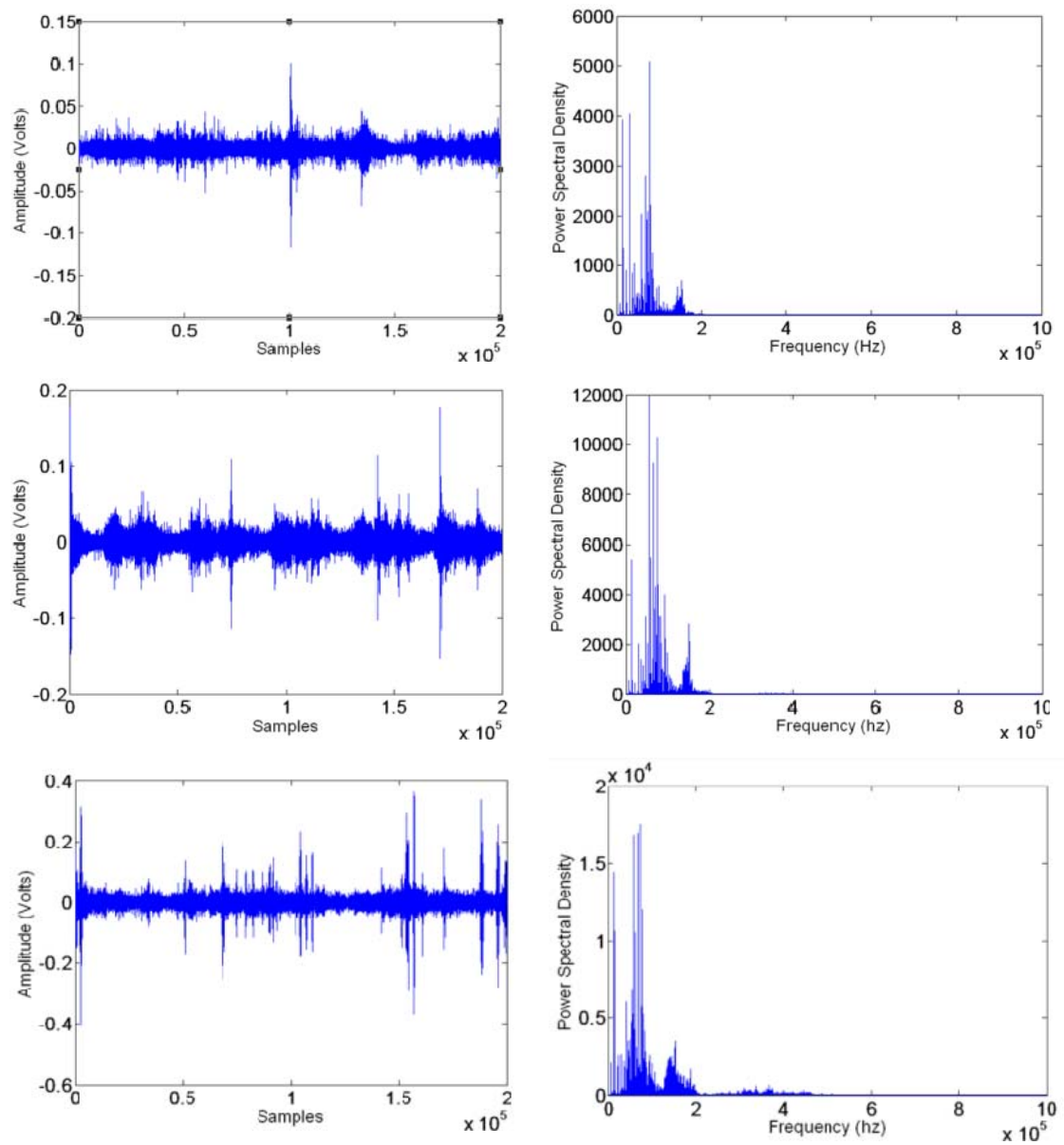


FIG. 4.9 AE Waveforms for lifelong experiment III (left graphs) and their respective PSDs (right graphs). The depicted waveforms, from upper to lower, correspond to 10%, 60% and 90% of operational time.

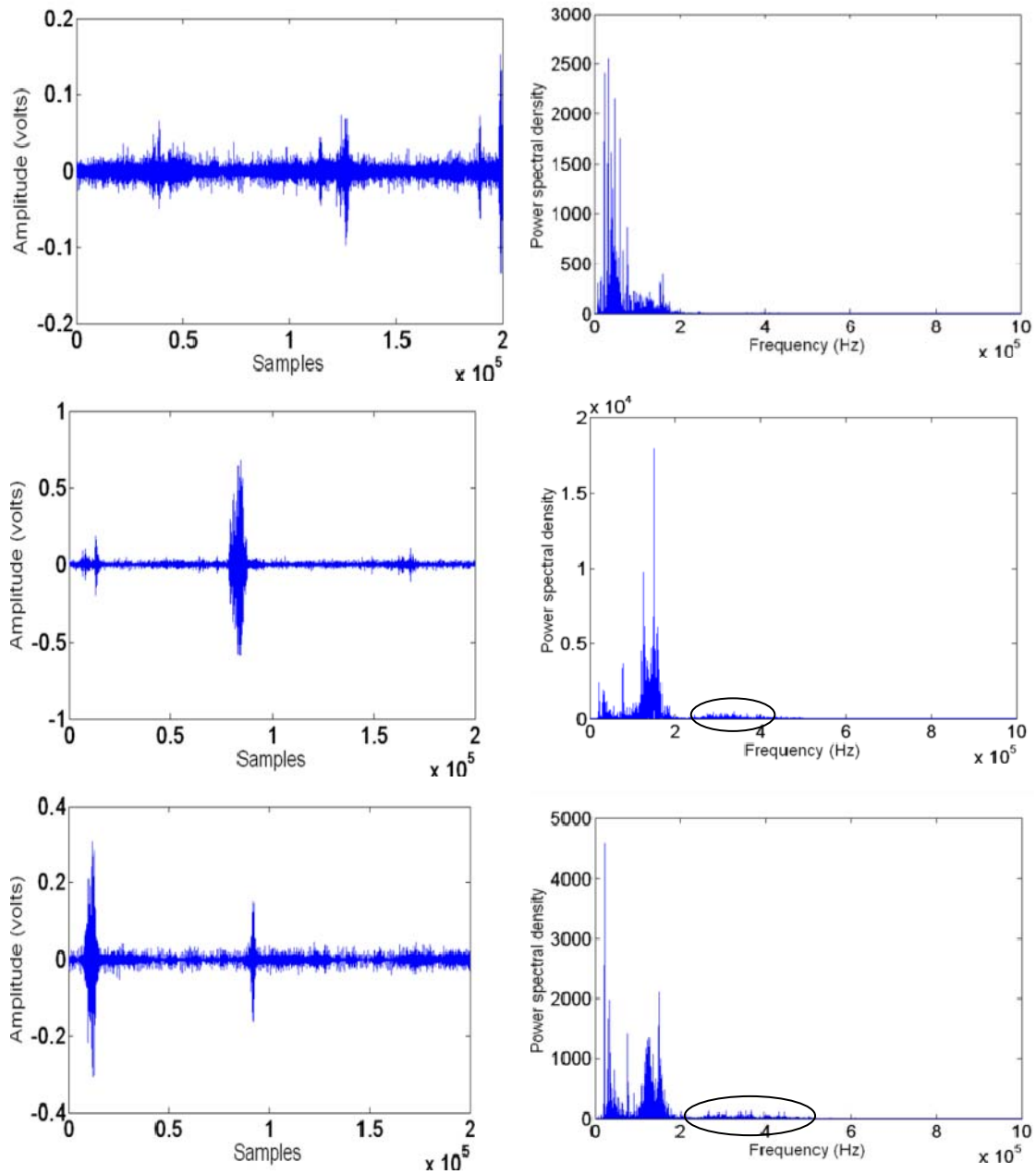


FIG 4.10. AE Waveforms for lifelong experiment VI and their respective PSDs. The waveforms are taken (from upper to lower plot) from the 10%, 60% and 90% of operational time.

To sum up, AE channel 1 seems to be a valid sensor for gear health monitoring. AE P-P amplitude seems very difficult to work with. On the other hand RMS is a much better diagnostics and prognostics candidate metric. Finally, a poor calibration of the preamplifier induced unwanted frequency components that were easily dealt with a high pass filter. This waveform preprocessing did not substantially influence the RMS trend in both lifelong experiments history. The recordings of this sensor will be kept for further study.

Acoustic emission channel 2

This sensor was in friction contact with the rotating gear. The specific application of this sensor is discussed elsewhere (Loutas et al. (2008)). This section will assess the efficiency of the particular sensor set up in connection to meshing gearbox fatigue.

Let's focus on experiment III recordings. A cease in these recordings is directly observed at 110 hours. The reason for this is not clear but can be attributed to electronics malfunction. Up to that point the waveform P-P amplitude does not yield any interesting trends again in contrast to RMS recordings. RMS seems to capture some gradual degradation initiating at around 50 hours. In general, however, no interesting trends arise in experiment V including AE P-P amplitude and RMS. Some typical waveforms that arise from that sensor from both experiments are presented in fig. 4.11 with their respective spectrum.

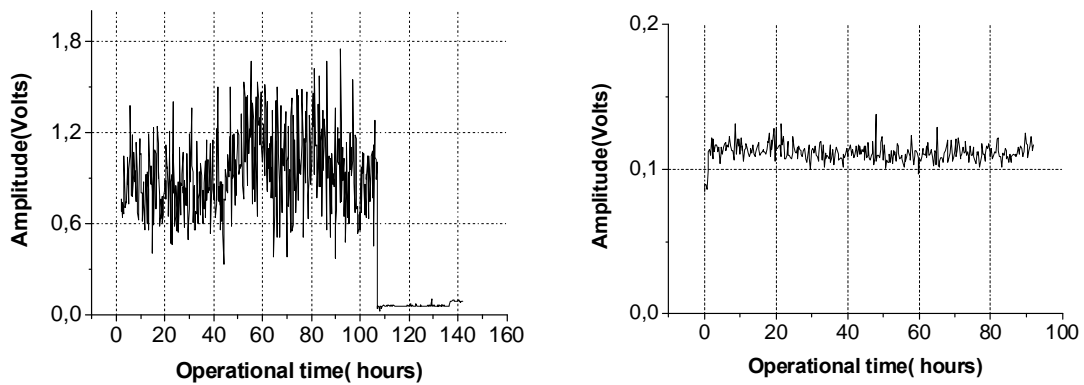


FIG 4.11. AE P-P amplitude for experiment III (left) and experiment V (right).

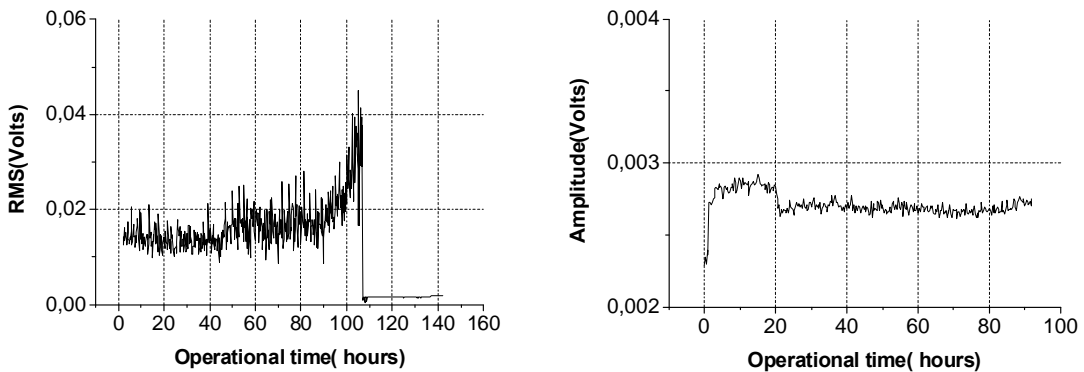


FIG 4.12. AE RMS for experiment III (left) and experiment V (right).

The spectrum of the recordings is concentrated in the band 20-120 kHz. The spectral distribution does not yield any change throughout the operation history. This observed in all lifelong gear experiments. The particular recordings are regarded obsolete and will not be accounted further for this study.

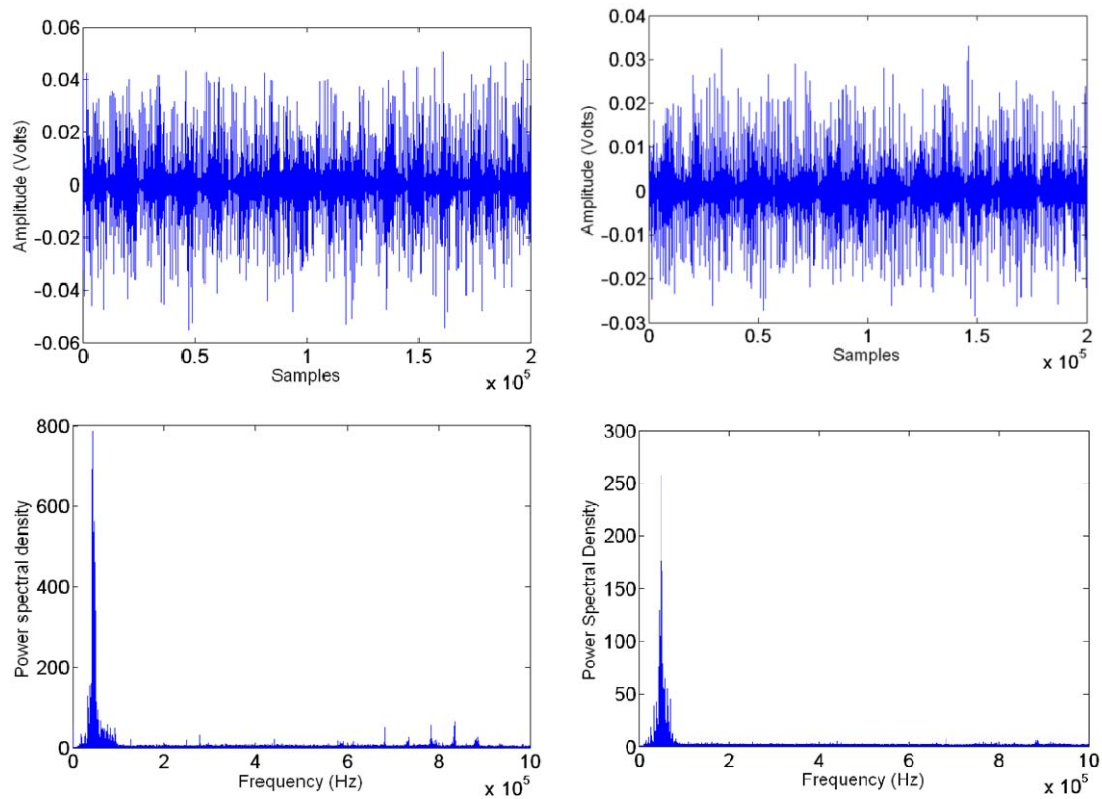


FIG 4.13. Typical AE channel 2 waveform (top) and its corresponding PSD (bottom) from experiment V.

Acoustic emission channel 3

This sensor is mounted on the bearing of the driving gear. Its purpose was to capture a possible bearing flaw. The existence of a bearing flaw could influence the measurements of the other sensors. Its application was important in order to de-correlate such an influence.

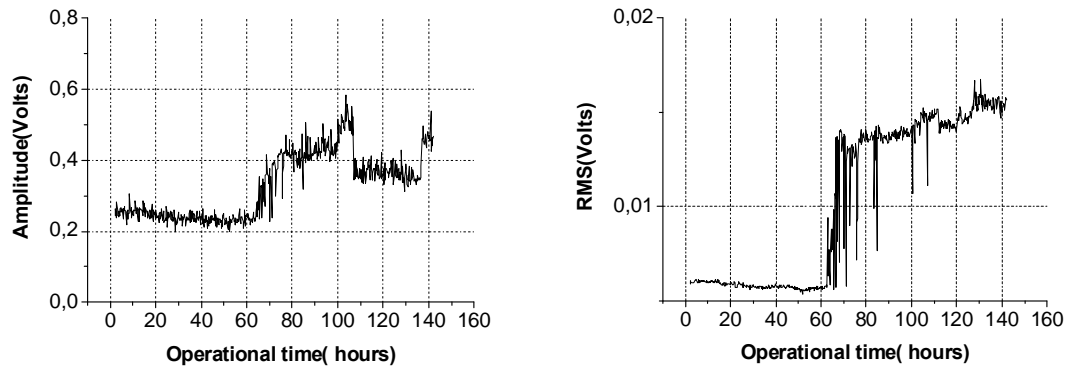


FIG. 4.14. AE P-P amplitude (on the left) and RMS (on the right) for experiment III.

Only experiment III depicted some interesting trend. As depicted in [fig 4.14](#) there is a jump in the recordings at 60 hours. This jump is observed in P-P amplitude as well as in RMS. Again, RMS is more emphatic. There is no doubt that AE channel has caught a bearing fault at that particular instant. The validity of AE recordings is also reflected in [fig. 4.15](#). The time as well as frequency domain have the expected AE form. It should be mentioned that the bearings were of closed form therefore their thorough visual inspection was impossible.

Two inspections one at 110 hours and one at the end of the experiment (at 162 hours) found that the bearing was operational, that is it was revolving without some evident constraint. This indicates that a) AE could capture a very small flaw in the bearing probably at the instant it appeared and b) this small fault didn't have a significant impact on the dynamics of the gear operation. Moreover, if one observes the RMS history of AE channel 1 sensor, no correlation is evident in the vicinity of 60 hours. The same holds for AE channel 2 recordings where a small change at 45 hours ([fig. 4.12](#)) cannot be connected to AE channel 3 recordings. All the latter constitute evidence that the appearance of the bearing fault can be assumed to have a negligible impact in the gear degradation dynamics. [Figure 4.15](#) depicts the recordings from two time instants in gearbox operating life. The recordings are taken from experiment III. The time instants are at 20% and 80% of the gearbox operating life. The change in the AE energy is evident both in the time and frequency domain. No change in the frequency domain distribution is noted though. The energy of the signal is concentrated between 20 and 100 kHz, the same way as AE channel 2 recordings. No other experiment gave interesting information regarding AE channel 3 sensor.

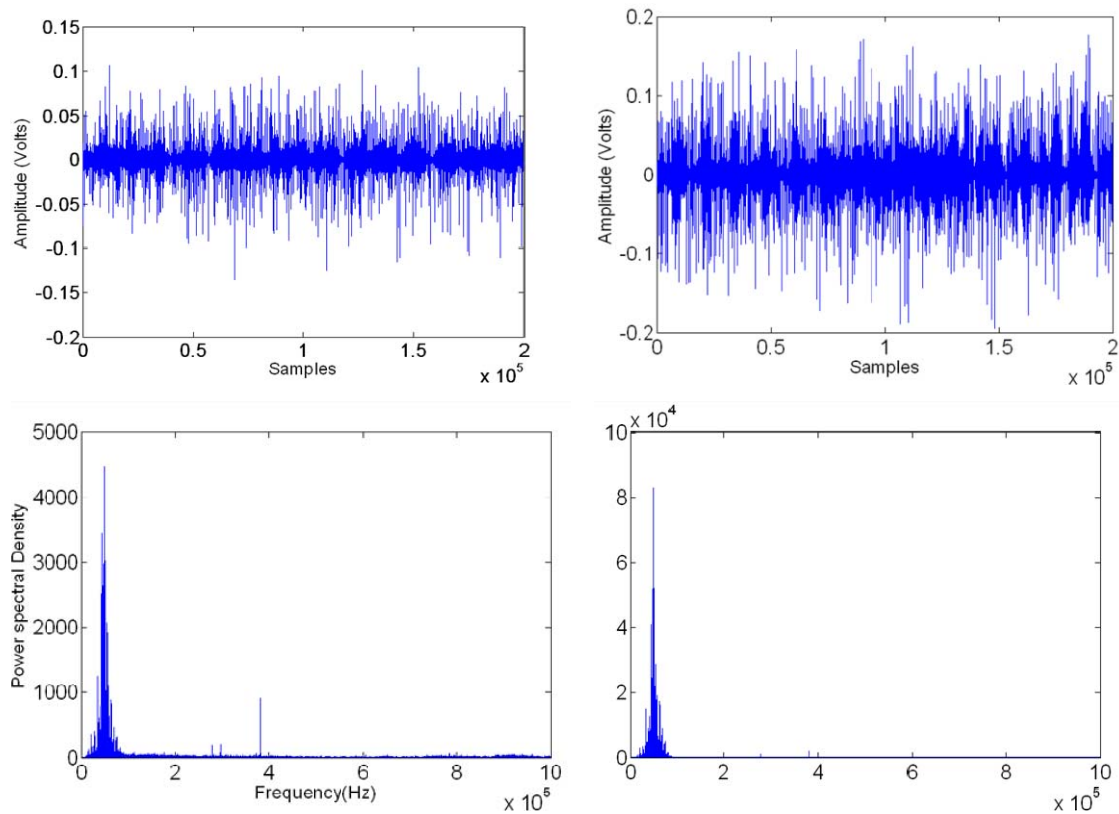


FIG. 4.15. Typical AE waveform (upper plots) and its corresponding PSD (lower plots) corresponding to experiment III.

4.5.3 Vibration recordings

As was earlier mentioned, a single three-axis sensor is mounted on the gearbox case. For the sake of coherence to the previous sections focus will be given in experiments III and VI. However the observations generalize to whole failure data base.

Vibration recordings assessment is nothing new in the area of gearbox CBM and will therefore be as brief as possible. The main objectives are to validate the recordings, assess and compare them to AE channel 1 which monitors the degrading gear.

Vibration X-axis

Figure depicts the maximum P-P amplitude for the particular sensor regarding experiment III and IV. Fig. 4.16 depicts the respective RMS histories.

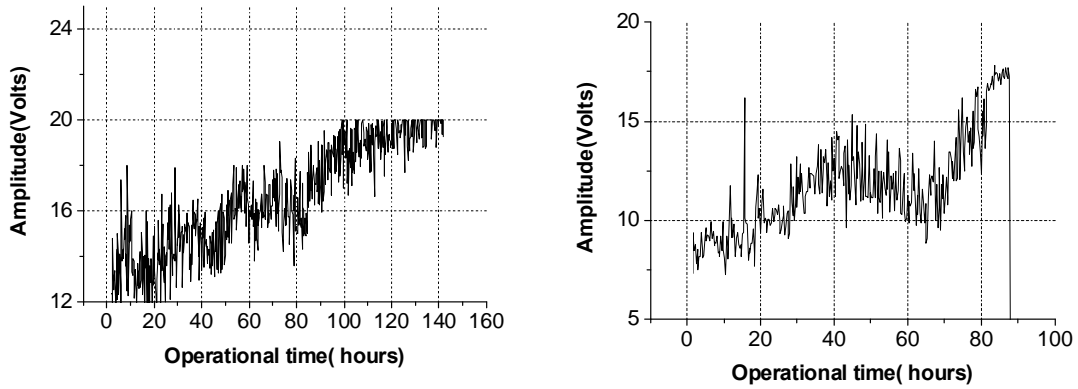


FIG. 4.16. P-P amplitude for experiment III (left) and experiment VI (right).

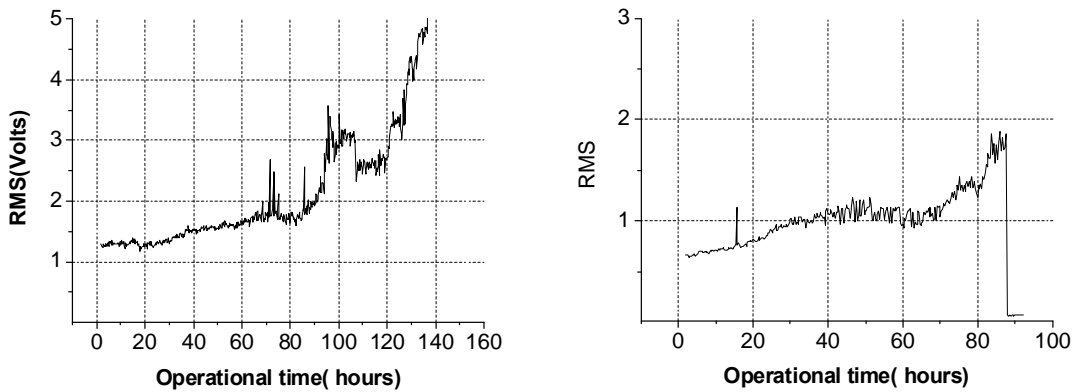


FIG 4.17. RMS for experiment III (left) and VI (right).

We will first focus on experiment III. At a very initial stage (close to 100 hours) there was the first evidence of sensor clipping effects due to insufficient calibration. These effects were detrimental in P-P amplitude recordings rendering them obsolete from a diagnostic point of view. On the other hand, RMS recordings were less affected. The monotonic trend of RMS after 100 hours is a consequence of the gear degradation and it can be correlated with it. At the vicinity of 110 hours the clipping effect was dealt with a highpass filter preprocessing. The cut-off frequency was placed low at 100 hz. Close to 140 hours, though, the increase in vibration level had a detrimental effect in vibration recordings. Most probable the interval after 136 hours will be disposed off and not taken into account at all.

A sudden jump in RMS recordings in 110 hours coincides with a stop event. This is evidence that the level of vibration is sensitive to stop-start events in contrast to AE recordings for instance, as depicted in earlier sections.

The sensor is better calibrated for experiment VI and that's why there is such a difference in the scale compared to experiment III. Also, no clipping effect is observed this

time. Near the end of this experiment (after 88 hours) a DAQ malfunction deprived us of the very last waveform recordings.

Vibration Y-axis

Figure 4.18 depicts the maximum P-P amplitude for the particular sensor regarding experiment III and IV. Figure depicts the respective RMS histories.

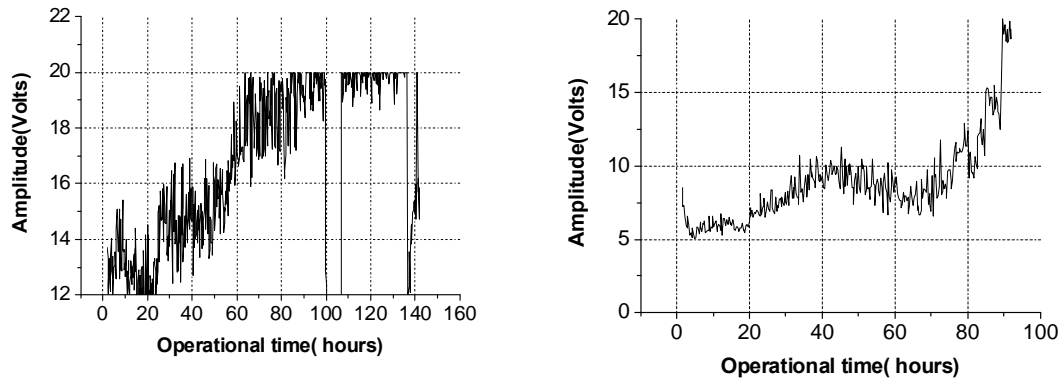


FIG 4.18. P-P amplitude for experiment III (left) and VI (right).

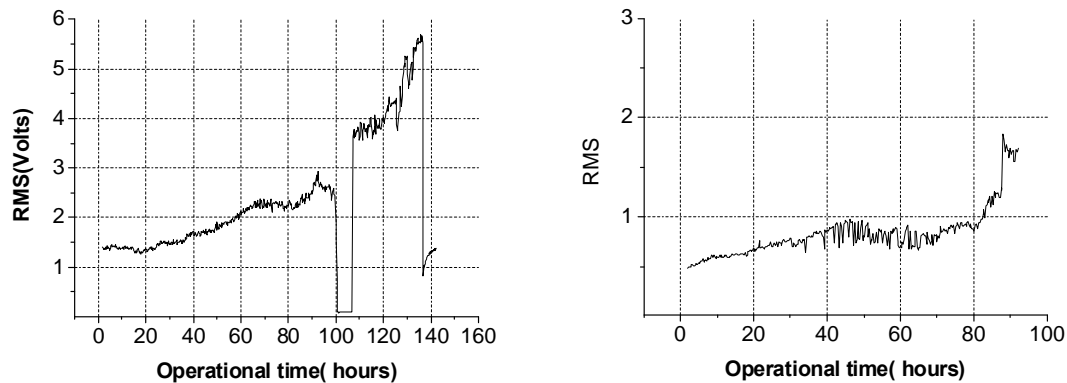


FIG 4.19. RMS for experiment III (left) and VI (right).

The poor sensor calibration had an even more profound impact on y-axis vibration recordings. RMS recording constitutes a good degradation metric and it is not so much affected by the bad calibration. RMS increases constantly with the operational life of the gearbox. A sensor malfunction caused several recordings (from 100 hours to 110 hours) to be lost. This fact can easily be dealt with the missing data theory as will be shown in Section 4.4.

Experiment VI seems not to have such issues. On the other hand the trend in P-P amplitude as well as in RMS is not very satisfactory. It only catches the final stage of the gearbox degradation but not hint for the earlier stages of fatigue failure.

Vibration Z-axis

This is the last vibration sensor and also the last sensor to be reported in this thesis. Again its RMS and P-P amplitude are shown in [fig. 4.20](#). It should be noted that experiment VI did not yield any valid waveform measurements probably due to DAQ system malfunction. Therefore, only the recordings from experiment III are presented.

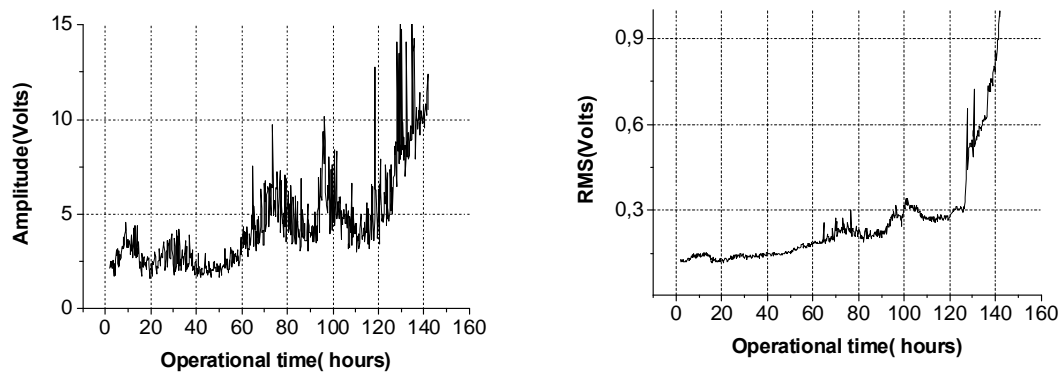


FIG. 4.20. P-P amplitude (left) and RMS (right) for vibration z-axis sensor, experiment III.

In this case no evident clipping effect is noted. P-P amplitude is very inhomogenous and no conclusion can be derived. On the other hand, RMS is again a safe condition monitoring index. The mono-tonicity of the metric is interrupted due to the stop event at 110 hrs. Again, the influence of a stop event in the vibration recordings is evident. Finally, the sudden jump in RMS recordings at around 136 hours coincides with the sudden change in Fe-particles rate of production in the corresponding [fig. 4.4](#).

4.5.4 Discussion on the experiments

Several gear failure experiments were conducted in the framework of this thesis. The difficulty of multi-hour full lifelong fatigue experiments was highlighted. Few recording shortfalls were related to DAQ malfunction and poor sensor calibration. Regardless of these minor shortfalls several conclusions can be drawn:

- AE has a shorter transient run-in phase compared to vibration monitoring as depicted in [Section 4.4](#). This is highlighted for the first time to the author's knowledge.

- ODM recordings seem to be a very good depiction of the degradation history of the gearbox. Based on these recordings the degradation dynamics of the gearbox can be divided into two regions. The first region is characterised by a relatively steady rate of Fe-particles production. When the fatigue degradation has progressed sufficiently the dynamics of Fe-particle release change dramatically. It is assumed that the gear can operate with relative success even when it enters the second phase of degradation but only for a short while. The

exact “critical failure” threshold cannot be accurately defined and it is not highly important. For the experiments at hand a threshold value of about 450 mg/hr can be defined heuristically.

- ODM recordings seem to be poor indicator of a “hard” failure event such as a gear tooth being chopped off. To the author’s knowledge this is the first time this shortfall of ODM is highlighted.

- RMS is a much better descriptive parameter than P-P amplitude for the waveform sensors. It comprises a lower noise content as well as better monotonicity. Finally its trend seems to be less affected by possible sensor poor calibration that result in waveform clipping.

- AE channel 1 sensor seems to follow very well the gradual degradation of the asset. Its recordings are mixed burst-like and continuous type mode. Some clipping effect was dealt with a high pass IIR filter that did not affect the RMS form. Furthermore it scales among the lifelong fatigue experiments but not dramatically. Finally, there were hints that there is diagnostic potential in the band 200-500 kHz. This can be exploited with a signal processing technique such as Wavelet packets and it will be covered in [Chapter 6](#).

- AE channel 2 did not exhibit significant trends relative to gear degradation. This could be attributed to the oil film that intervenes between the sensor interface and the surface of the rotating gear. Therefore, for the current study it is completely rejected.

- AE channel 3 is proven efficient in capturing small bearing flaws during operation. Its recordings helped in measuring the effect of a small bearing flaw in the level of AE and vibration recordings. An inspection of RMS histories verified the slight impact the bearing fault had on the gear dynamics in experiment III. The particular sensor did not yield any important information in the rest of the fatigue experiments.

- Vibration recordings were found to follow, in general, the degradation level of the machine. In certain cases poor sensor calibration led to extensive clipping effect, particularly near the end of the experiment. Two remedies are proposed for this issue 1) high pass filtering with a very low cut-off frequency to minimize the clipping effect influence on the waveform frequency content and 2) disposing off a part of the final observations if the vibration recordings when deemed excessively corrupt. The clipping does not have a detrimental impact on RMS history and it is therefore deemed valid for prognostic purposes. After all poor sensor calibration can appear in real life applications and is unavoidable in long term monitoring of a device whose critical failure threshold is not explicitly defined.

4.6. Preprocessing of the health metrics

The obtained gear failure recordings are going to be studied for prognostic purposes. To this a series of preprocessing is going to be applied on all sensor timeseries. The parameter

that will comprise the prediction timeseries will be the RMS unless explicitly defined otherwise.

Clipping effect reduction. The exact point of failure of the machine and therefore the final level of vibration was not a priori known. This led to poor calibration of the waveform sensors which, in a few cases, resulted in clipping effect. These effects were handled as described in a previous section i) high pass filtering with a cut off frequency of 10 kHz for AE channel 1 recordings and 200 Hz for vibration ii) dispose of measurements which were deemed excessively corrupt.

Run-in period disposal. In rotating machinery, vibration recordings are highly correlated with contextual parameters and most notably the machine temperature. The effect of temperature rise was neglected with the disposition of a certain time interval at the beginning of the experiment. Typically the first few hours will be neglected.

Time synchronous averaging. TSA proved impossible to implement on the derived vibration waveforms. One reason was the lack of a tacho-signal to synchronise the gearbox revolutions. A blind phase deconvolution scheme ([Combet and Gelman \(2007\)](#)) for TSA was applied on vibration waveforms but the results were difficult to evaluate. The reason could be higher speed deviations than the ones permitted by the algorithm (the algorithm assumes <5% speed deviation ratio). In any case, sheer RMS seems to hold an adequate information content relative to damage progression as shown in previous sections.

Outlier detection. Outlier detection is an issue for any robust time series model. There is no explicit definition of when a time series sample is an outlier or not. Generally, such a recording stems neither from the normal physical condition of the system under consideration nor contextual noise. It can be attributed to sensor problems, DAQ malfunctioning or some other event that is not relevant to the operation of the asset and it generally does not follow the statistical distribution of the bulk of the data. The outlier detection and cancelling that took place is based on the work of [Menhold et al. \(1999\)](#) and is completely causal.

The outlier decision threshold. The outlier decision threshold should be of statistical nature. Most often, the standard deviation of a moving window is defined as a good means to distinguish between normal and abnormal samples. This is not very robust though. In [Menhold et al. \(1999\)](#) a threshold based on the median absolute deviation is applied (Eq. (4.1)) divided with 0.6745. This standardizes Median Absolute Deviation (MAD) in order to make the scale estimate consistent with the standard deviation of a normal distribution

$$\text{MAD} = \frac{\sum_{k=0}^{N-1} |x_k - d_N|}{N} \quad (4.1)$$

Referring to Eq. 4.1, x_k is the k th sample of the timeseries window of width N and d_N is the median of the respective timeseries. If the next sample is larger than three times the normalised MAD then it is replaced by the median of the windowed timeseries else it is left untouched.

The window length. The length N of the time series window in Eq. 4.1 should also be defined. Too long a window wouldn't take into account the local dynamics of the phenomenon. On the other hand, too short a window could allow outliers be accepted as valid recordings, especially in cases that patches of outliers exist in the timeseries. The window should have length longer than the larger outlier patch.

Missing values. Frequently in timeseries studies we experience loss of data. This can be efficiently handled with application of sampling theory. In our case vibration channel 2 (fig. 4.20) experienced such a shortfall. In order to replace the missing values two estimations are required, the trend and the additive noise.

Trend estimation. We choose the boundary points and we fit a function that more or less describes the local progression of the timeseries. An single exponent is chosen and its parameters are found with a non-linear optimization scheme. In Matlab© environment the *fit* function helps us to find the required parameters using the Levenberg-Marquadt (LM) optimization algorithm. LM is an improvement over gradient descent algorithm that gives good answers for small and medium size non-linear parameter optimisation problems with the criterion of the minimum square error.

Noise estimation. First we choose a small time window prior to the data loss. We dispose of any local trends with a simple differentiation and subtract the mean from this timeseries window. The result will hopefully comprise stationary data. A test for stationarity can be made over the differentiated data window by autocorrelation function plot.

We approximate our stationary (noisy) data with a mixture of two Gaussians. The motivation comes from Lloyd (1982) stating that a mixture of gaussian can approximate any pdf with arbitrary accuracy. Equation 4.2 depicts the pdf deriving from two Gaussians.

$$p(x|\theta) = c_1 N(x, m_1, \sigma_1) + c_2 N(x, m_2, \sigma_2) \quad (4.2)$$

where θ is the vector of pdf parameters, c_1, c_2 are the mixing constants, m and σ are the mean and variance, x is the sample vector value and N is the gaussian pdf

$$N(x|m,\sigma)=\frac{1}{\sigma\sqrt{2\pi}}\exp\left(-\frac{1}{2}\frac{(x-m)^2}{\sigma^2}\right) \quad (4.3)$$

The optimization of these parameters is performed on the noisy samples that precede the acquisition void. The optimization scheme is based on the standard EM algorithm (Appendix B). Once the parameters have been defined a sampling scheme will provide the synthesized noise data. The *random* Matlab© command will provide us with a number of noisy samples from the mixture of the two Gaussians. These samples will be added to the deterministic exponential trend from the earlier step.

The latter process takes place for the data series of fig. 4.19 and the result is depicted in fig. 4.21.

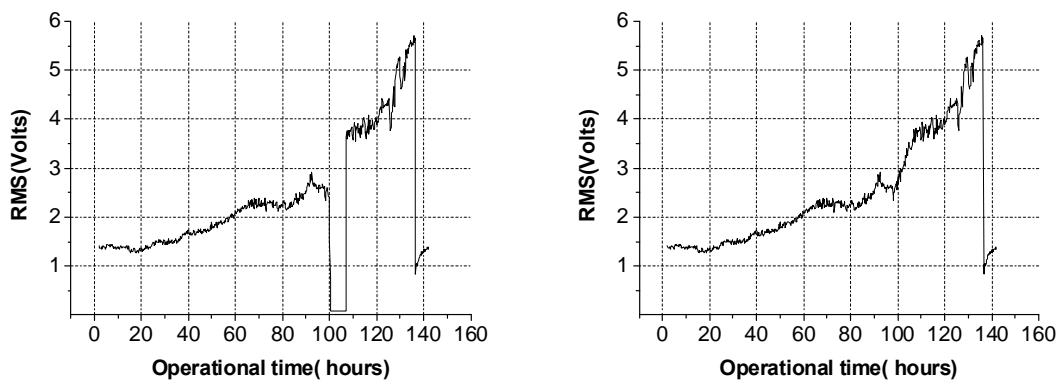


FIG 4.21. The data series on the left comprises missing values. On the right a remedy is applied.

Stop-start events. As has been noted, stop-start events have an impact on ODM recordings. We choose to dispose a certain number of recordings comprising about half an hour right after each machine re-initialization.

4.7 Correlations between sensors

The correlation among waveform sensors is evident in all failure histories. AE sensor-1 as well as the three-axis vibrations sensor seem to carry a similar amount of fatigue related information. The question is, if one were to choose a single RMS indicator to monitor the gearbox fatigue which would be the best choice. To answer this question we calculate the similarity of each sensor timeseries to a health signal of reference. This signal is supposed to be a very accurate depiction of the underlying gear degradation process and is none other than

ODM aggregate mass. This metric is assumed to be an accurate recording of pitting fatigue. We will limit the machine operation to the point that Fe-mass rate crosses the 450 mg/hr threshold. This particular threshold is derived after observation of all failure histories at hand. We calculate the similarity of each sensor timeseries with ODM recording in mean square terms. The Coefficient of Cross-Correlation (CCC) is our similarity criterion

$$CCC_{xy} = \frac{E\{(x - m_x)(y - m_y)\}}{\sigma_x \sigma_y} \quad (4.4)$$

where m is the sample mean and σ the sample standard deviation of the respective timeseries. CCC is frequently referred to as Pearson coefficient in literature. A CCC value “1” would imply perfect linear relation between the two compared timeseries whereas “0” would imply independence.

CCC	Exp. I	Exp. II	Exp. III	Exp. IV	Exp. V	Exp. VI
ODM Mass/ AE channel 1	0.755	N/A	0.9381	0.751	0.8505	0.7513
ODM Mass/ vib x-axis	0.6991	N/A	0.6928	0.797	0.8272	0.7407
ODM Mass/ vib y-axis	N/A	N/A	0.7101	0.881	0.4747	0.6611
ODM Mass/ vib z-axis	0.5104	N/A	0.6521	N/A	0.8488	N/A

$p < 0.05$ for null hypothesis in every case. N/A corresponds to loss of sensor data.

Table 4.2

Table 4.2 depicts the CCC scores of each waveform sensor RMS. It seems that AE has the best monitoring potential in four out of five failure histories since it is a better correlated with ODM aggregate mass. Since the gear tooth pitting is directly connected to ODM aggregate mass, AE is again proven the best monitoring technique for the particular failure mode. It is worth mentioning that ODM and RMS of waveform recordings had different time scales. In order to make valid comparisons a matching of the time timeseries scale was required. We were reluctant to apply an interpolation scheme since this would act as a finite impulse response filter and therefore induce within timeseries correlations possibly distorting the sensor information content. The best strategy was to simply find the ODM sample with the closest timestamp to each corresponding waveform RMS sample.

The findings of [Table 4.2](#) reconfirm the work of [Tan et al. \(2007\)](#). In this study spectrometric oil analysis was used to compare the level of Fe particles in the lubricant oil with the level of AE. Excellent correlation is not observed though in each failure history. Deviations from linearity can be attributed to non-linearity of the transmission path and possible AE attenuation during transmission through the bearings. Moreover, the transmission path could undergo small changes from experiment to experiment due to the replacement actions explaining the deviations shown in [Table 4.2](#).

As a next step we aim to find the functional relation between the AE recordings and the level of fatigue of the gearbox. Again, the level of fatigue will be demonstrated by the ODM aggregate mass. In the framework of this thesis we will search for the relation between AE activity and an indirect index of pitting, the ODM aggregate mass recording. To this end we will resort to the methods of pattern recognition.

Pattern recognition for regression. Suppose we have a set of examples lying in the N-dimensional space (I^N) and a set of outputs of M dimensions (O^M). Regression is the set of actions that appoint an appropriate function f that maps the examples to our outputs when the outputs belong to the real numbers ($O^M \in R^M$).

$$I^N \xrightarrow{f} O^M \quad (4.5)$$

In order to define a proper regression function we need to take a number of steps as shown below.

Definition of inputs. Our inputs will be the RMS values which are obtained from the gear failure experiments. The AE recordings are bagged together to form the “input” set and are normalized in the scale from [-1,1]. In the experiments that took place AE RMS scaled significantly less than vibration in all cases. Here lies another advantage of the AE compared to vibration health monitoring from the scope of machine learning. A limited variance in scales is desirable because it results in better generalisation of the trained models.

Definition of outputs. ODM aggregate mass recordings are kept as our output space. The recordings are scaled with their mean maximum value within the lifespan of each gearbox. Again they are bagged together to form the “output” set they are scaled in the [-1,1] region. In the end what we have is 1-dimensional normalized inputs (AE RMS) and 1-dimension normalized outputs (ODM aggregate mass) which will produce our training and testing sets.

Definition of regression scheme. We need to find a (linear or non-linear) mapping that best describe our outputs in terms of our inputs. [Tan et al. \(2007\)](#) have proposed a linear mapping. We suspect the same therefore we will experiment with linear models.

Linear regression. Ridge regression is an improvement over the least squares problem for linear regression. More specifically, suppose we need to express a relationship $\mathbf{y}=\mathbf{a}^T.\mathbf{x}$. In terms of least squares we search for the minimum of the quantity $\sum\|\alpha\mathbf{x}_i - \mathbf{y}_i\|^2$ where the latter sum is over all input-output $(\mathbf{x}_i, \mathbf{y}_i)$ couples in our data base and the term inside the sum is the Euclidian norm. Ridge regression penalizes the size of regression coefficients (in our case \mathbf{a} 's) providing a more robust optimum solution to the least squares problem. The quantity that is optimized according to this method is $\sum\|\alpha\mathbf{x}_i - \mathbf{y}_i\|^2 + \lambda\|\alpha\|^2$ where λ is a regularization term and takes a small value ($\lambda\in[0.2,0.3]$). The exact value of λ is appointed after a few trial and error runs on a separate validation set using Matlab©.

Further on, we wish to explore non-linear mappings that possibly explain better the AE-ODM mass relation.

Feed Forward Artificial Neural Networks. FFANN is a notion derived from biology, in particular the organisation of the human central neural system. Similarly to biological neural systems, an artificial NN comprises of numerous interconnected units called the neurons. A neuron is a mathematical function. It takes the weighted sum of its inputs and maps it to an interval, based on a (linear or nonlinear) transfer function. Several aspects of the ANN should be determined to define an efficient classification scheme

1. The architecture.

The FFANN has one input layer with as many units as the input dimension. In our case this will be a single.

An ANN with one hidden layer can approximate any function that contains a continuous mapping from one finite space to another¹¹. The number of hidden-layer neurons is going to be defined by a number of bootstrap experiments.

The FFANN has one output layer with dimension 1 corresponding to the dimension of our outputs.

2. The standard hyperbolic tangent is most frequently applied as the non-linear term when referring to FFANN models. For the purposes of regression the output layer will be a linear function. The input-output pairs are adjusted to the scale of the hyperbolic function and the linear output function that is [-1,1] in both cases.

3. The FFANN weights are going to be optimized by minimizing an appropriate objective function. Again this function is none other than the sum of square error between the predicted and the actual outputs. The comparison takes place on a “testing” set.

¹¹ <http://www.heatonresearch.com/node/707>

4. The optimization algorithm is the Levenberg-Marguadt (LM). LM is in fact an improved steepest descent iteration algorithm. This algorithm performs very well for non-linear regression and is especially suited for FFANN adjustment. Its fast convergence is ensured partly by its clever structure and partly by the normalization of input-output training sequence. The normalization of input output pairs is essential so that the largest to smallest eigenvalue ratio of the objective function Hessian is closer to one. The closer proximity of the Hessian eigenvalues results in faster convergence to the family gradient descent iteration algorithms where the each iteration step is found by calculation of the Hessian matrix --or pseudo Hessian in the case of LM (Epleman (2007)).

Radial Basis Function (RBF) networks. A second candidate to explore non-linearity in our data-base is the RBF. Like neural networks, radial basis networks comprise a weighted sum of non-linear mappings. The transfer function has the form $\varphi(\|x - x_0\|)$ where x_0 is the “centre” of the function and $\|\cdot\|$ is the Euclidian distance (radius) from the centre. The function φ can take various forms. In our case the standard gaussian kernel will be applied for experimenting, $\varphi(x|x_0) = \exp\left(-\frac{\|x-x_0\|^2}{\sigma^2}\right)$ where x_0 is first determined with a few runs of k-means algorithm that provide a first estimate for a EM like iterative algorithm to adjust all x_0 and σ 's of the network. Again, the appropriate number of RBF's is defined with cross-validation.

RBFs as well as FFANNs comprise some stochasticity in their optimization process. In both cases the initial parameter values are distributed randomly across the input layer space. This stochasticity is dealt with a multiple bootstrap training and testing process.

Boot-strap simulations. Bootstrap methods are used to create many training sets from a relatively limited distribution of measurements. Each training set that will be created will comprise of a small number of input-output pairs (for example 200) which are randomly sampled from the data base at hand. Then we repeat the sampling process create a small number of training sets for example 10. We choose the *randi* Matlab© function to sample from the data pool because it generates integer samples from a uniform pdf without inserting bias. This way each set is representative of the underlying database. Each bootstrap training set is used to adjust a FFANN or a RBF.

Another set is used for feeding the adjusted model. This latter set is the testing set and is created the same way. The result is later compared with the actual training paradigms label. The variance or mean square error of the predicted and actual values is used as the similarity criterion and it will define the best AE-ODM mapping model.

One set of 200 pairs is created with bootstrap sampling for training and one for testing. The process took place 10 times for each realization of FFANNs and RBFs. The number of hidden units varied in the case of FFANN and RBF.

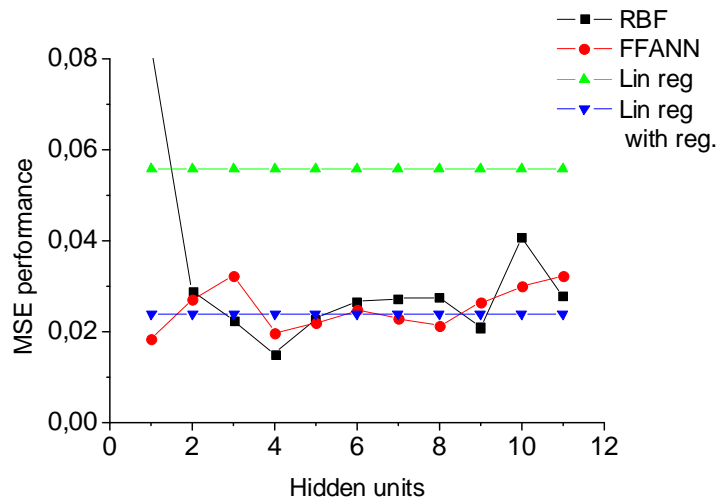


FIG. 4.22. Simulation results for AE-ODM relation.

Figure 4.22 summarizes the simulation results. In the figure, linear regression and linear regression with regularization (Ridge method) are shown as lines for comparison. The best performance of all tested models is achieved with RBF with 4 hidden units. The difference in performance with higher RBF schemes is not significant. We choose the smallest model that explains our database. The performance of RBF with two units is evidently higher than the Ridge implementation for several RBF realizations as well as most FFANN realisations. This implies that a non-linear relation can better link AE and ODM aggregate mass recordings. The relationship between AE and ODM aggregate mass recordings is established for the first time to the author’s knowledge, builds on the work of Tan et al. (2007) and is highlighted for the first time to the author’s best knowledge.

Best descriptor of AE-ODM relation	Number of hidden units	Mean Square Error
RBF	4	0.015

Table 4.3. Selected model for AE-ODM functional relation.

4.8 Synthesized health indexes

In the previous section a relation between AE and ODM recordings was established along with error estimation. We can build further upon these findings towards synthesizing lifelong gearbox data. The idea is to create an adequate pool of gearbox lifelong histories that can comprise the basis for RUL prediction model development and assessment. Several other authors have used synthesized measurements to make inference and assess life prediction models (Zio and Pelsoni (2011), Moghaddass (2013), Zhang et al. (2014), Baraldi et al. (2013), Wang et al. (2011), Orchard and Vachtsevanos (2009)). Those researchers focused on some first physical principles to synthesize the lifelong timeseries. The motivation for such an action is basically the scarcity of valid real life degradation data of mechanical devices.

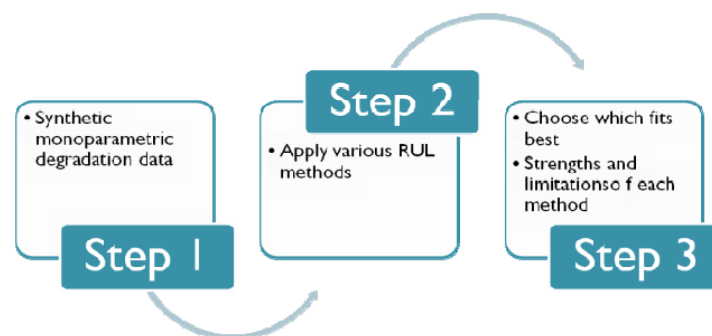


FIG. 4.23. The basic idea behind the synthesis of lifelong RMS histories.

We need to build degradation gearbox sequences from synthesized ODM recordings. The idea is to:

1. Synthesize ODM Fe particles production rate recordings that are very close to the histories at hand.
2. Integrate the latter quantity to produce ODM aggregate mass recordings.
3. Produce AE RMS histories based on the optimal model of [Section 4.5](#).

Synthesize ODM Fe-mass rate. Based on the experiments at hand we can define two gearbox operating regions. The first region is characterized by steady Fe particles production rate. This coincides for instance with the first 130 hrs for experiment III and to the first 78 hrs for experiment VI ([fig 4.4, 4.5](#)). This region will be referred as the linear operating region. The second region is characterized by a dramatic change in iron particle formation. The asset is supposed to operate for a short within this strongly non-linear regime. We will further assume a failure when Fe mass production rate crosses a deterministic threshold.

ODM Fe mass can be simulated as a gamma process, its derivative to being a gamma random variable.

$$\Gamma(x|\alpha,\beta) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x} \quad (4.5)$$

where α is the gamma pdf shape parameter and β is the gamma scale parameter. An α close to 1 implies a Poisson like process and while it increases the pdf becomes more bell like.

The shape and scale parameters for this gamma process can be calculated based on the experiments at hand. A number of Fe mass rate measures from the conducted experiments and a gamma pdf is fitted with Matlab *gamfit* tool. This tool estimates the gamma parameters with the EM algorithm ([Appendix B](#)) along with 95% confidence intervals.

Different experiments are expected to have different average rate of particle production. To parameterize this variability it is assumed that the (positive) average rate of particle production is again a random gamma variable. Its parameters are defined heuristically by a few trial and error efforts. Alternatively, the confidence intervals calculated from the fitting process of the earlier paragraph can be used. This can be done by sampling these gamma parameters from a normal distribution with mean the *gamfit* maximum likelihood estimate and a variance estimated from the *gamfit* confidence interval. The sampling would take part once in every experiment.

As mentioned earlier, a critical point (CP) is observed in all experiments which distinguishes two operating regions. This CP defines the “beginning of the end” for the meshing gear and normally it is of stochastic nature and is assumed as a gamma random variable. The distribution of CPs can only be defined on a population of failed gears. Since such a database was not possible to create, the shape and scale parameters of the corresponding pdf are chosen arbitrarily. Weibull process is frequently used to simulate failure times in engineering components. Weibull pdf has a similar form to gamma distribution and it could be used instead for sampling CPs.

The non-linear degradation region is characterized by a strong trend. An exponential function is insufficient to simulate this accelerated process. A hyper exponential function is chosen for this purpose.

4.8.1 Synthesis results

The model which will generate the lifelong failure histories is shown below. Equation 4.6 depicts the corresponding Fe particles production rate.

$$\begin{aligned} \text{ODM}_{\text{rate}}(k) &= St_1 \cdot \exp(St_2 \cdot (t[k])^3) + St_3[k] \\ St_1 &\sim \Gamma(1.5, 0.05) \\ St_2 &\sim \Gamma(24.4, 0.05) \\ St_3 &\sim \Gamma(0.8, 12) \end{aligned} \tag{4.6}$$

$k=0,1,\dots,N$ samples, $t[k]$ is the time stamp of the corresponding sample, St_i 's are the variables that introduce the model uncertainty as described in the previous section. More precisely:

St_1 is the intensity of particles production. For each life history this is assumed a constant variable but changes from history to history to simulate differences in machine configurations. It is sampled from a gamma distribution each time a new history is being simulated.

St_2 defines the appearance of a CP and it is defined arbitrarily, only to match the chosen time scale parameter t . Again, this is sampled only once from a gamma distribution before the degradation history is simulated. The sampling takes place in each simulated history.

St_3 simulates the peaks that appear in the particles production rate. The respective parameters are fitted on actual measurements as mentioned in the previous section. This value is positive and is sampled from a gamma distribution every at each time point. In all latter cases the sampling is made simple with the use of *gamrnd* Matlab© function.

The chosen non-linearity was chosen to be an exponential function with a time scale power of 3. This is chosen bigger than 1 to match the actual measurements as mentioned in the previous section.

The failure threshold is defined as 350 assuming milligrams of ferrous particles per hour. Once this threshold is crossed the AE histories simulation comes to end. Figure 4.24 depicts 3 synthesized curves.

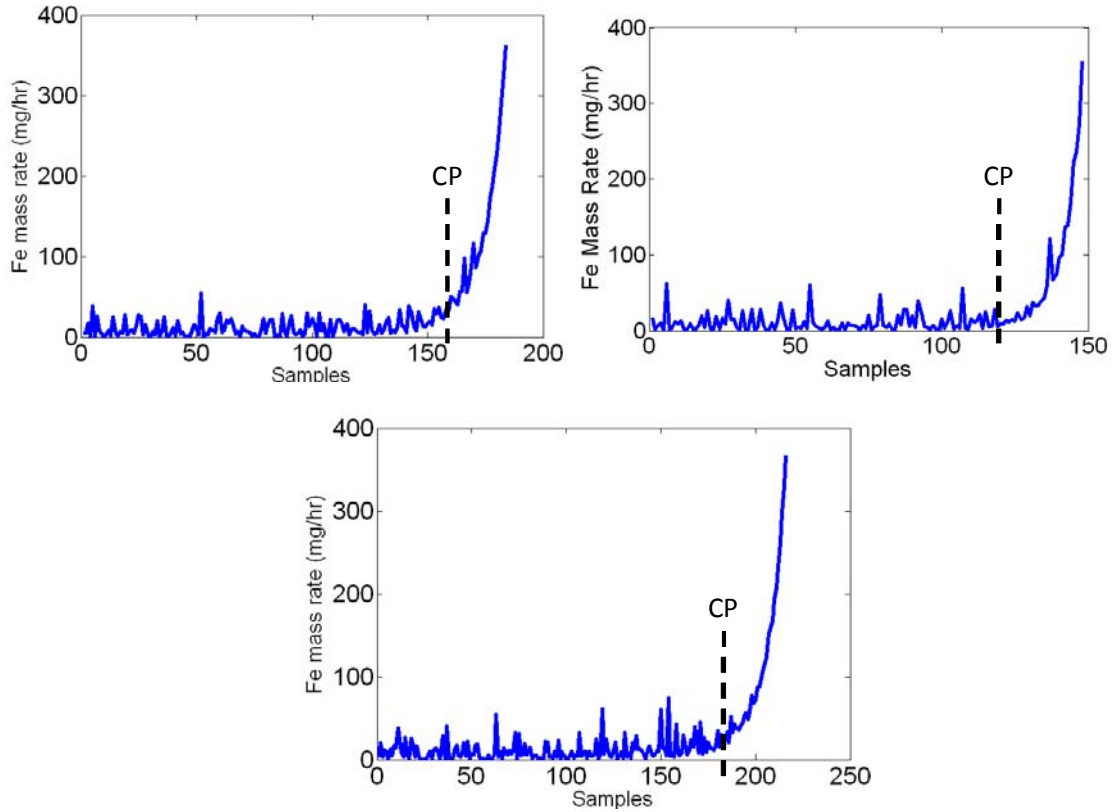


FIG 4.24. Three synthesized Fe-rate histories.

Naturally the ODM aggregate mass is given by the following relation

$$\text{ODM}(k + 1) = \text{ODM}(k) + \text{ODM}_{\text{rate}}(k) \quad (4.7)$$

assuming $\text{ODM}(0)=0$

Equation 4.7 defines a Markov process related to the aggregate Fe mass.

AE RMS lifelong curves are only a few steps way. First the ODM curves are normalized in the region $[-1,1]$. Then AE RMS histories are produced as follows.

$$\text{RMS}_{\text{AE}}(k) = \sum_{i=1}^4 \exp\left(\frac{\text{ODM}[k] - \text{kernel_center}_i}{\sigma_i}\right) + w[k] \quad (4.8)$$

$$w \sim N(0, \sqrt{0.015})$$

Where the first term of eq. 4.7 is the optimal RBF network from Section 4.5, x is the synthesized ODM measurements scaled to $[-1,1]$ and N is the normal pdf with zero mean. The standard deviation of w is the square root of variance (mean square error) estimated in the same Section 4.5. Figure 4.25 depicts some synthesized RMS histories which are truncated to -1 . An observation made on fig. 4.25 is that the maximum of each curve is not the same in

each realization. This small difference in scale accounts for the different scales encountered in the actual experiments.

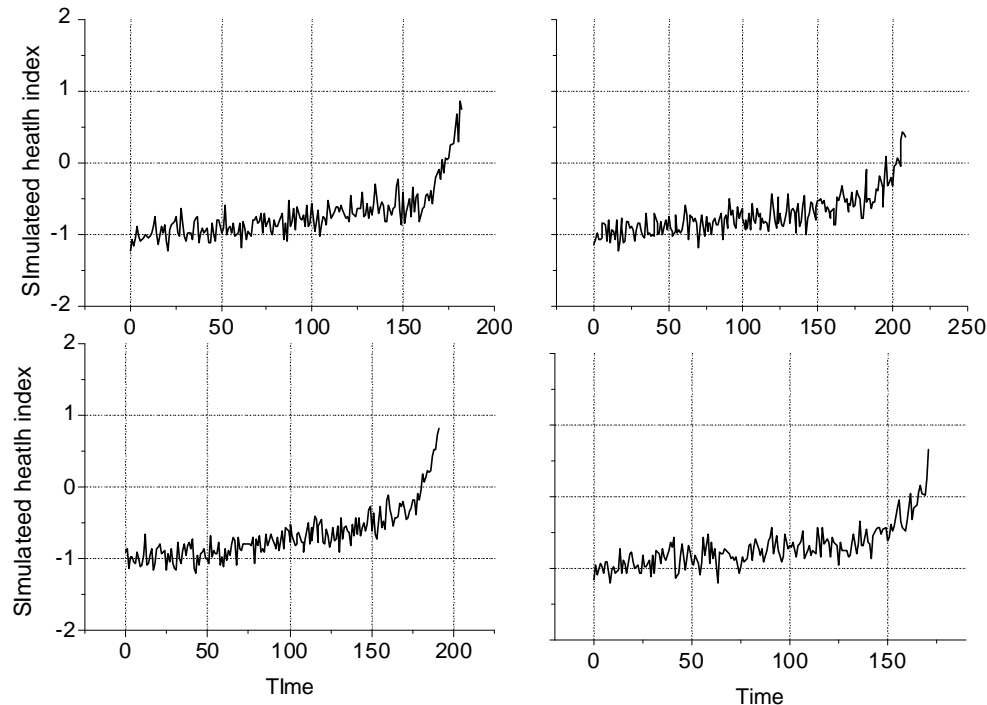


FIG. 4.25. Four synthesized failure histories.

4.9 Conclusion

❖ Several experiments took place for the study of gear pitting fatigue.

A number of experiments took place to measure the run-in effect under different load and speed configurations. The basic result was that AE exhibits a shorter transient phase in comparison to vibration. To the author's knowledge this is the first time this characteristic of AE is highlighted.

Few lifelong experiments took place in the framework of the current thesis. After observation of all derived failure histories ODM was confirmed as an excellent measurement of gearbox degradation level although not for every possible mode of failure. ODM fails to capture sudden events such as chopped gear tooth incident. To the author's knowledge this is the first time this characteristic of ODM is highlighted. Therefore a combination of ODM with some other, waveform type, sensing is proposed for robust gear health monitoring.

❖ Fe aggregate mass was well correlated to the operating duration and therefore the assumed- health level of the equipment. ODM mass was also highly correlated with AE channel 1 and the 3-axis vibration sensor gear failure recordings. Two dynamic gear

degradation regions could be defined on ODM Fe-mass rate curve: one corresponding to “linear” degradation and another with highly non linear dynamics. The component is assumed to operate in the former and for a very small duration in the latter until it crosses a failure threshold. This observation is coherent with other research such as [Gasperin et al. \(2011a\)](#) as denoted in [Chapter 3](#) and [fig. 3.3](#).

❖ AE monitoring is highlighted as the most accurate degradation monitoring technique for a meshing gearbox scoring the highest similarity criterion, a result coherent to the work of [Tan et al. \(2007\)](#). Furthermore a non linear relation was established between AE and ODM relation providing a -marginally- better performance compared to linear -ridge- regression.

❖ Finally, this chapter contributed a framework for simulating AE RMS measurements that correspond to gearbox failure fatigue. To our best knowledge such an effort is made for the first time although it is motivated by numerous other researches as denoted in [Section 4.6](#). The reasons to simulate AE and not ODM measurements are summarized in [Section 4.3.1](#).

Various prediction methods are going to be applied, some for the first time and their efficiency towards RUL is going to be assessed in the following chapter.

CHAPTER 5

Estimating the Remaining Useful Life in Gearboxes

5.1 Introduction

Chapter 5, 6 and 7 are the core of this thesis. They comprise the predictions on synthesized as well as real life data from gearboxes and bearings. The proposed integral CBM scheme (prognosis and diagnosis) is depicted in fig. 5.1.

Chapter 5 covers the issue of gearbox prognostics with a single monitoring parameter. In particular, RMS is solely going to be studied. Due to the scarcity of full life fatigue experiments, synthesised timeseries are going to be applied for this study. The particular synthesis process has been discussed in Section 4.8. Three prediction methods are going to be adapted on this database.

Age-dependent model is going to be applied and assessed for RUL prediction in Section 5.2. The particular Cox proportional hazards model is going to be assessed for the first time for gear failure fatigue data.

A regression tool is going to be assessed in Section 5.3 and more particularly a ϵ -Support Vector regression scheme. Finally an extrapolation model is going to be adapted in

Section 5.4. All methods are applied for the first time to predict gearbox RUL on mono-parametric data. The strengths of each method as well as their limitations are going to be highlighted.

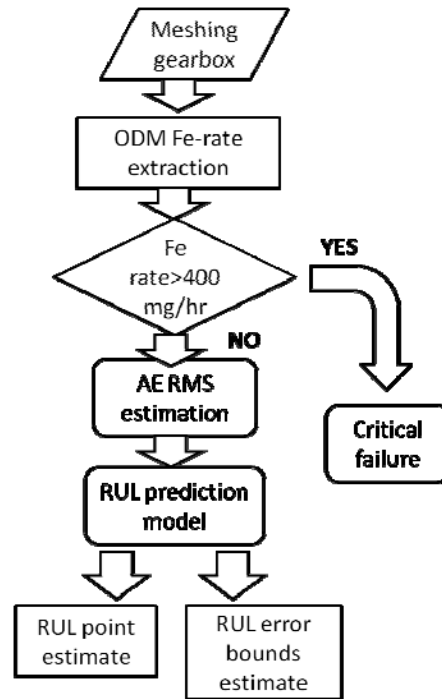


FIG. 5.1. The flowchart depicts the general gear failure prognosis framework of this chapter.

Section 5.5 concludes this chapter by repeating the main contributions and findings.

5.2 Proportional Hazards Model (PHM)

5.2.1 Theory

In this section a brief overview of the Proportional Hazards Model (PHM) is going to be presented which is based on the work of Cox (1972). Proportional hazards model is a semi-parametric model which has mainly found application in survival analysis of medical records. Its application in quantitative RUL prediction in mechanical systems literature is relatively limited though and non-existent for gear failures in particular.

The hazard function plays a central role in PHM. Typically it is the probability that a component will fail immediately after a time instant t provided it has survived until that instant.

$$h(t) = \lim_{\Delta t \rightarrow 0^+} \frac{\Pr(t \leq T_f \leq t + \Delta t \mid t \leq T_f)}{\Delta t} \quad (5.1)$$

Cox proves that when interval Δt becomes infinitesimal this probability can be written as

$$h(t) = h_0(t) \exp\left(\sum_i \gamma_i Z_i(t)\right) \quad (5.2)$$

Where h_0 is a baseline hazard rate and the functional inside the exponential summarizes the effect of various covariates on the hazard function, each covariate weighted with a γ_i parameter. These covariates can be external variables such as load, temperature, humidity, number of repairs of the equipment etc.. Internal covariates are based on the response of the structure to an external cause, for instance vibration and AE RMS or some other statistical index, ODM recordings etc. It is worth mentioning that the functional form of the exponent as well as the baseline function may have any form according to Cox. In the end, an optimization algorithm makes sure that the underlying function explains well the failure data base.

The reliability function stems from $h(t)$ and it is defined as the probability that a component survives more than t , $S(t) = \text{pr}(t \leq T_f)$. Reliability function takes the value 1 at the beginning of the operational life or some other value smaller than 1 if the asset has a non-zero probability to break down at the very beginning of its operation. In the current context it will be assumed that $S(t=0)=1$. Reliability function is a strictly monotonically decreasing function of time.

Assuming that failure in each subsequent operating interval $\Delta t = t_k - t_{k-1}$ is an independent variable then $S(t) = \lim_{\Delta t \rightarrow 0} \prod_{k=0}^{r-1} \{1 - h(t_k) \Delta t\}$ for r operating time intervals. Taking $h(t)$ to be integrable and adding the dependence of the covariates vector Z , Cox ends up with the following expression

$$S(t | Z(t)) = \exp\left(-\int_0^t h(s | Z(s)) ds\right) \quad (5.3)$$

The probability that the component will survive briefly after t provided it has survived until t is called the conditional reliability function and it is given by

$$S(t + \Delta t | t \leq T_f) = \frac{S(t + \Delta t)}{S(t)} \quad (5.4)$$

[Jardine et al. \(1987\)](#) assumed a Weibull basis hazard function for mechanical systems. This yields

$$h(t) = \frac{\beta}{\eta} \left(\frac{t}{\eta}\right)^{\beta-1} \exp(\mathbf{Y}^T \mathbf{Z}(t)) \quad (5.5)$$

where β and η are shape and scale parameters of the Weibul pdf. The respective parameters can be estimated through an optimization technique. We need an objective function for the optimization process. Since we are dealing with probability measures the

objective function of interest will be the likelihood function of the model optimised over the failure histories at hand. The likelihood function comprises of two factors:

- L_1 giving the probability that every respective component has survived up until its T_f and it is given by

$$L_1 = \prod_{i=1}^N S(t_i | \beta, \eta, \boldsymbol{\gamma}) = \prod_{i=1}^N \exp\left(-\int_0^{t_i} \frac{\beta}{\eta} \left(\frac{s}{\eta}\right)^{\beta-1} \exp(\boldsymbol{\gamma}^T \mathbf{Z}(s)) ds\right)$$

Where the i index corresponds to time intervals where no failure has been observed. The product L_1 takes place on all intervals being observed in all

failure histories, therefore $N = \sum_{j=1}^q (\text{num_intervals_per_history}_j)$ q being

the number of failure histories (components that were observed to have failed).

- L_2 giving the probability that a component fails at time T_f and is given by

$$L_2 = \prod_{i=1}^q h(t_i | \beta, \eta, \boldsymbol{\gamma}) = \prod_{i=1}^q \frac{\beta}{\eta} \left(\frac{t_i}{\eta}\right)^{\beta-1} \exp(\boldsymbol{\gamma}^T \mathbf{Z}(t_i))$$

One of the strengths of PHM is that it is able to incorporate both failure histories with an explicit failure event as well as “censored” failure histories where the failure is not directly observed. The effect of censored histories is not going to be studied in the current framework.

The total likelihood function is given by the product of L_1 and L_2 . Both factors being products of probabilities, they can become annoyingly small in a scale that cannot be managed by computer software. The standard process is to take the logarithm of this quantity therefore transforming the products to sums.

Eventually the quantity to be optimised is

$$\text{LogL} = \text{Log}(L_1) + \text{Log}(L_2) \Rightarrow$$

$$\text{LogL} = \sum_{N_{\text{interv}}} \int_0^{t_i} \frac{\beta}{\eta} \left(\frac{s}{\eta}\right)^{\beta-1} \exp(\boldsymbol{\gamma}^T \mathbf{Z}(s)) ds + \quad (5.6)$$

$$q \ln\left(\frac{\beta}{\eta}\right) + (\beta-1) \sum_q \ln\left(\frac{t_i}{\eta}\right) + \sum_q \boldsymbol{\gamma}^T \mathbf{Z}(t_i)$$

There are two ways to calculate the underlying parameters $\boldsymbol{\theta} = (\beta, \eta, \boldsymbol{\gamma})$. One is to take the derivative of the objective function with respect to each parameter in $\boldsymbol{\theta}$ and find the zeroth point with a Matlab® root finding function. Another way is to keep the objective function as

is and use direct search methods to find the optimum such as Genetic Algorithms (GA). The latter method will be chosen for the following study.

Calculation of expected RUL:

The latter equations (eq. 5.1- 5.6) merely provide probability of survival measures for the underlying component. The Estimated Remaining Useful Life (ERUL) or mean-time-to-failure of a component can be indirectly calculated at every time instant t_0 with the aid of the reliability functions. ERUL at t_0 can be thought as the mean of all survival times longer than t_0 . In order to compute the mean we need a survival probability density f under the condition that the asset has survived up until t_0 . This will be given with the aid of eq. 5.3, 5.4.

$$f(t|t_0) = -\frac{dS(t|t_0)}{dt} = -\frac{d\left(\frac{S(t)}{S(t_0)}\right)}{dt} = \frac{1}{S(t_0)} \frac{dS(t)}{dt} \quad (5.7)$$

Now that we have an expression for a conditional pdf we can derive an expression for the estimated time to failure provided the component has survived up until t_0 .

$$\begin{aligned} \text{ERUL}(t_0) &= E\{T_f - t_0 | T_f \geq t_0\} = \int_{t_0}^{\infty} (u - t_0) f(u | t_0) du = \\ &= \frac{1}{S(t_0)} \left\{ \int_{t_0}^{\infty} (u - t_0) \left(-\frac{dS(u)}{du} \right) du \right\} = \\ &= \frac{1}{S(t_0)} \left[(u - t_0) \frac{dS(u)}{du} \Big|_{t_0}^{\infty} + \int_{t_0}^{\infty} (u - t_0)' S(u) du \right] = \\ &= \frac{\int_{t_0}^{\infty} S(u) du}{S(t_0)} = \frac{\int_0^{\infty} S(u) du - \int_0^{t_0} S(u) du}{S(t_0)} \end{aligned}$$

The final expression can be written as

$$\text{ERUL}(t_0) = \frac{\text{MTTF}_0 - \int_0^{t_0} S(u) du}{S(t_0)} \quad (5.8)$$

Where MTTF is the mean time to failure for a component at $t=0$. Assuming a Weibull basis hazard function, this parameter is given by

$$MTTF_0 = \eta \cdot \Gamma\left(1 + \frac{1}{\beta}\right) \quad (5.9)$$

Where Γ is the gamma function. Eq. 5.8 is the important relation for RUL estimation.

5.2.2 Application in gear failure data

In this part we will assess the ability of Weibull PHM to produce meaningful RUL estimations in gearboxes. The actual failure histories at hand of table 4.1 are deemed too few to test PHM. In order to bypass this shortfall we will assume several life histories which will be produced by the stochastic model presented in Section 4.8 (Eq. 4.8).

Fifteen life histories are produced to test the application of PHM. RMS is going to be used as covariate. In each history the simulation stops when the Ferrous particles mass rate exceeds 400 assuming mg/hr. Since the failure threshold is set on this value, RMS histories exhibit slightly different scales (as shown in fig. 4.24) ranging from [-1, 0.56] to [-1, 0.98]. This is coherent to the actual gear failure experiments conducted in this thesis (as shown in fig. 4.8). Table 5.1 depicts the resulting operational time for each of the 15 simulated gear failure histories.

Simulated history	Operational time (in arbitrary units)
1-training	172
2-training	189
3-training	219
4-training	150
5-training	199
6-training	181
7-training	184
8-training	212
9-training	166
10-training	220
11-training	210
12-training	215
13-training	171
14-training	163
15-training	166
16-testing	172
17-testing	192

Table 5.1. Simulated Gear failure experiments.

Prior to fitting a PHM model, the form of the model should be defined and justified. First PHM is proven to perform better than a sheer Weibull process. Then the PHM is tested on few failure histories.

Weibull process vs Weibull PHM process fit:

A Weibull process assumes a Weibull hazard function that is

$$h(t) = \begin{cases} \frac{\beta}{\eta} \left(\frac{t}{\eta} \right)^{\beta-1}, & t > 0 \\ 0, & \text{elsewhere} \end{cases}$$

We can estimate optimal β and η parameters of the Weibull process by taking [eq. 5.6](#) and neglecting the effect of the covariate vector $Z(t)$. Genetic Algorithm (GA) ([Mitchell \(1996\)](#)) optimisation is applied to minimize the respective negative log-likelihood. GA performs stochastic search for the minimum of a function. An alternative would be to estimate the partial derivatives of the likelihood function, with respect to every parameter to be optimised, and search the roots of the resulting linear system. The GA process was opted due to its direct search property -it does not require the estimation of the gradient to reach the optimum but only uses the function values at particular random points-. Moreover almost in every run of the algorithm it computes an adequate –meaning good enough- solution. Furthermore it is especially efficient in optimizing functions that comprise many dimensions and non-linearities and the likelihood function is classified as such. An issue with the algorithm is the definition of the search interval for every underlying parameter, a demand that normally needs some experience or good insight relative to the underlying data base. For the current study the intervals for GA parameter search are defined as follows

$$\beta \in [-2, 2]$$

$$\eta \in [50, 300]$$

$$\gamma \in [2, 10]$$

In each run the algorithm returns an adequate but not optimal estimate for global minimum. Also the algorithm may be trapped in local minimum. Therefore several runs take place on the training data set (about 20). In the end the median of each parameter is returned as the optimum estimate. The model likelihood is calculated on the latter median values. The same is performed for the case where the RMS covariate is added. The optimization results are depicted in [table 5.2](#). The variances of [table 5.2](#) are the sample variances after the 20 consecutive runs of GA optimization algorithm.

	Weibull process	Weibull PHM with RMS as a covariate
Parameter estimation	$(\hat{\beta}, \hat{\eta}) = [10.3, 196.76]$ $\text{Var}(\hat{\beta}) = 7.2$ $\text{Var}(\hat{\eta}) = 9.7$	$(\hat{\beta}, \hat{\eta}, \hat{\gamma}) = [1.01, 163.4, 7.92]$ $\text{Var}(\hat{\beta}) = 0.0036$ $\text{Var}(\hat{\eta}) = 3122$ $\text{Var}(\hat{\gamma}) = 0.38$
Likelihood	LogL=-68.65	LogL=-33.89

Table 5.2. GA optimisation results.

Now we test the hypothesis that the covariate is statistically insignificant. To this end we apply the likelihood ratio test (Casella and Berger (2001)). Likelihood test compares two models, the one being an extended version of the other. The test statistic is calculated on the Loglikelihoods of the respective models on the particular database and it is given by

$$D = 2(\text{LogL}_{\text{alternative}_m} - \text{LogL}_{\text{null}_m}) \quad (5.10)$$

The probability distribution of the statistic is chi-square with degrees of freedom the $df = df_2 - df_1$ where df_1 and df_2 are the degrees of freedom of the null and augmented model respectively.

In our case we test the hypothesis that the augmented PHM explains significantly better against the null hypothesis that $\hat{\gamma} = 0$ (sheer Weibull process). The result is $D = 69.6$ with $df = 3 - 2 = 1$ degrees of freedom and the alternative is accepted with $p \ll 0.05$ over the $\hat{\gamma} = 0$ hypothesis. The parameters values of the accepted model are depicted in table 5.2. We observe that $\hat{\beta} = 1.01 > 1$ which implies a very weak effect of aging process. This is equivalent to a slowly increasing hazard and a slowly decreasing RUL with the aging of the component. A $\beta = 1$ would imply a constant hazard rate and therefore insignificance of the effect of aging in the estimation of hazard. The parameter η provides the characteristic scale of the distribution. $\hat{\eta} = 163.4$ means that merely 63.2% failures occur before 163.4 time samples. The MTTF is calculated on the median of the calculated parameters through eq. 5.9

and is found 163.15 time samples. Observing [table 5.1](#) it is evident that this value is well below the sample mean and leads to an underestimation of the true $MTTF_0$. This result is on the safe side since an underestimation of the remaining asset life cannot lead to a catastrophic decision. On the other hand it results in suboptimal RUL estimations.

RUL Prediction Results

[Fig. 5.2](#) summarises the results on three unseen failure histories. The reliability probabilities are depicted in particular. For convenience the latter curves are plotted with the median of the PHM parameters as estimated with the algorithm of paragraph 5.2.1. It is observed that both hazard rate and reliability diverge only slightly from their initial values, 0 and 1 respectively, with operating time. This outcome was expected due the fact that the Maximum Likelihood (ML) estimate of β parameter is very close to one and therefore the aging affects only slightly the latter probability metrics. Moreover, the condition monitoring parameter has great impact on the latter probability estimates but only close to each component failure point.

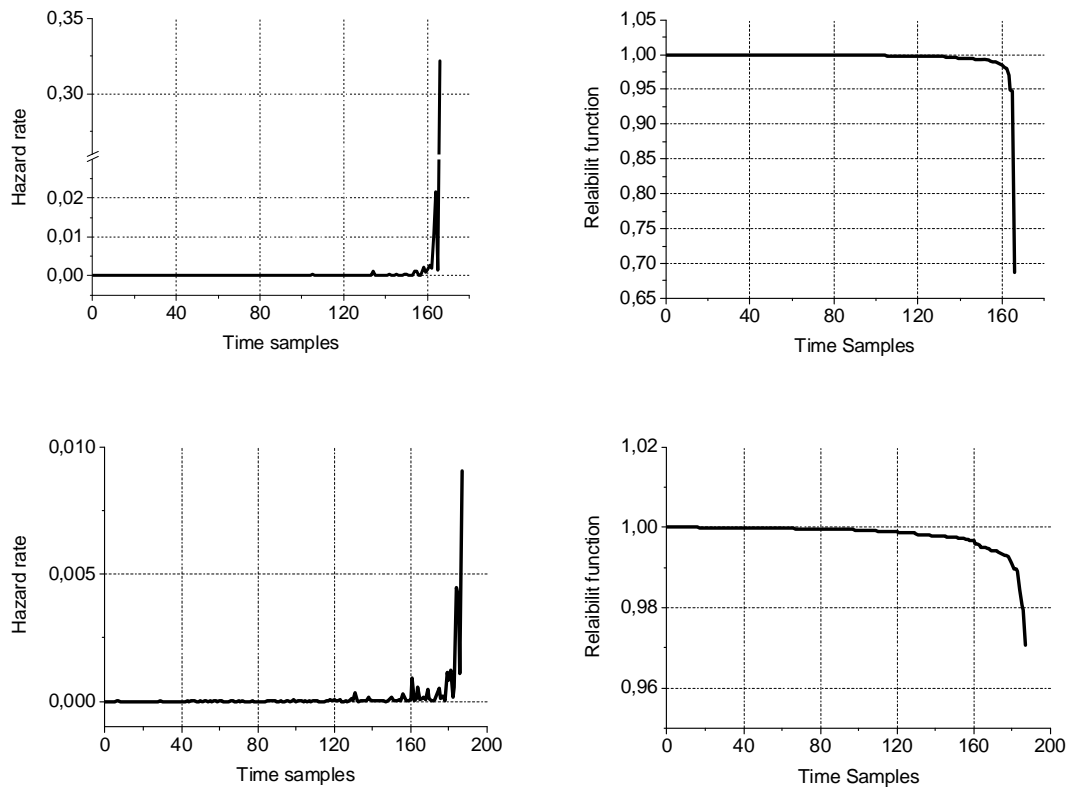


FIG. 5.2. Test results regarding hazard rate and reliability function on two out of sample failure histories. Upper graph corresponds to realisation “16” and lower graph to realisation “17”, [table 5.1](#).

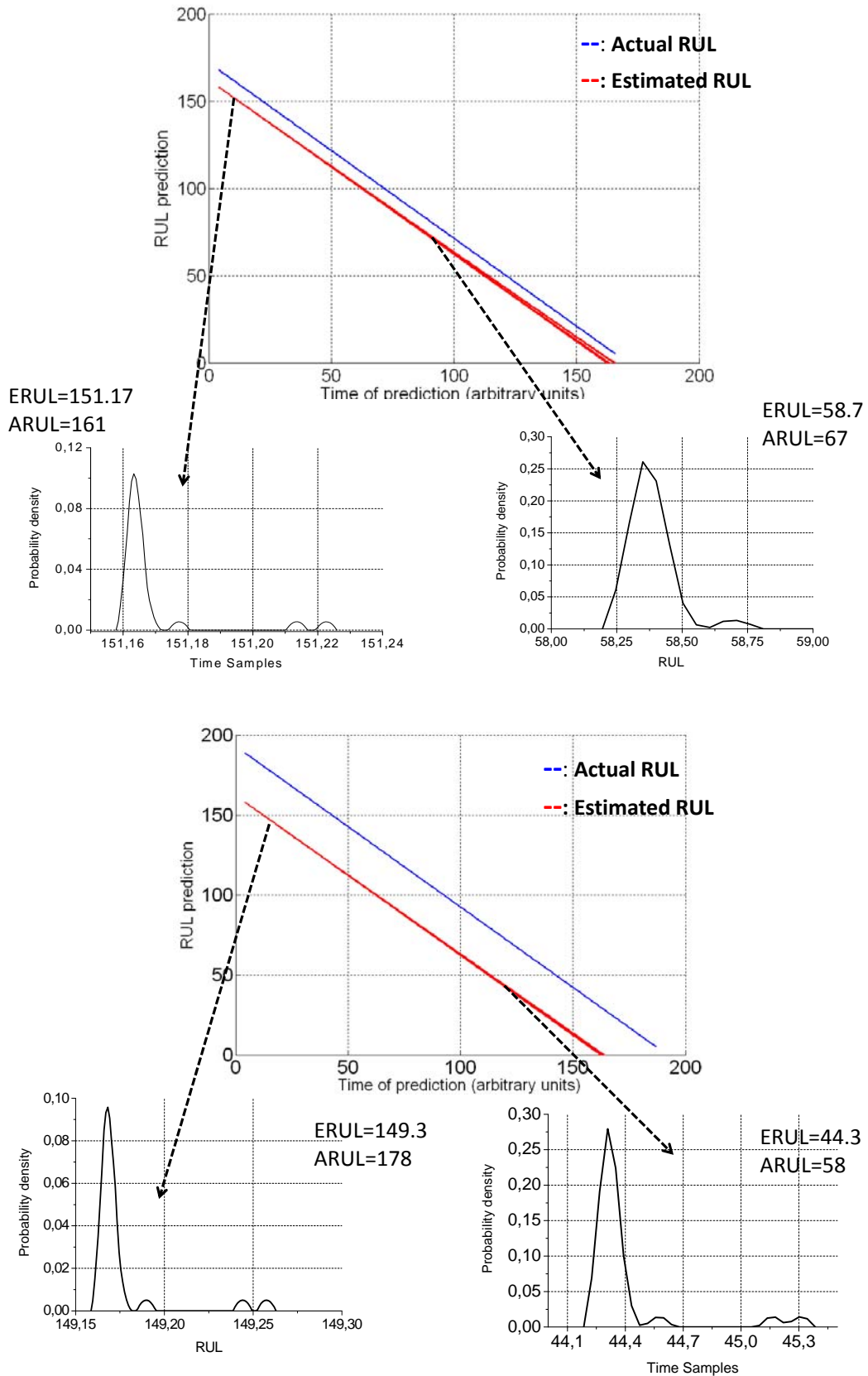


FIG. 5.3. Two RUL prediction experiments. Upper graph corresponds to realisation 16 and lower graph to realisation 17, [table 5.1](#). Actual time of failure is 169 and 192 time samples for the upper and lower graph respectively.

It is evident that both hazard rate and reliability function poses diagnostic value since they diverge sharply close to the end of life of the corresponding component. However, the important metric is RUL. To this end we calculate the parameter $MTTF_0$ and apply eq. 5.7 for each. $MTTF_0$ from the median PHM parameters is found to be 163.15.

The model from table 5.2 is applied for quantitative RUL predictions on two out of sample failure histories. These histories are derived from the model of eq. 4.28. The actual and predicted RUL is calculated on certain time instants and the results are depicted in fig. 5.2. 20 PHM model realisations are produced by the GA optimization process. Each realisation is forwarded in eq. 5.8. The corresponding predictions are plotted and interpolated in the RUL graph of fig 5.3. It is observed that despite the variance of PHM parameters (see table 5.2), the RUL prediction variability is very small, so small that it is almost obsolete. Figure 5.3 also depict the distribution of the interpolated predictions at particular time of observation. For better visualization an Epanechnykov kernel (Epanechnykov (1969)) has been fit on the distribution points yielding the corresponding pdf graphs.

All predictions seem to fall in an almost straight line, parallel to the Actual life line of each asset yielding an exceedingly high precision although the parameter variances (table 5.2) seemed not to be insignificant. A deviation from that line appears only close to the end of life of each assumed component. Moreover the condition monitoring index is more influential only close to the asset's end of life.

To sum up, the particular Cox regression model for gear failure prediction yields

- i) adequate but suboptimal RUL point predictions.
- ii) with very narrow error bounds that become wider with operational time
- ii) a negligible influence of the CBM feature except for the vicinity of the asset's end of life.

5.3 Probabilistic ϵ -SVR

5.3.1 Theory

Support vector machines are a concept introduced by Vapnik (1995) and became very popular the last years both for classification and regression. In principle, SVR is advantageous -as opposed to conventional regression techniques- as it maps in a non-linear fashion the original input data to a higher dimension in Hilbert space via kernel functions (Smola and Scholkopf (2004)). Supposed a set of n observations is provided. In the most general case, each observation lies in an M -dimensional space, $\mathbf{x}_i \in \mathbb{R}^M$, $i=1..n$. At the same time, a desirable mapping is provided for each observation vector. Let's assume a set of target values $y_i \in \mathbb{R}$ that comprises the vector of mappings. Regression is the problem of finding a

transformation f such that $X_{NxM} \xrightarrow{f} Y_{Nx1}$ in the best possible way. A good answer to the latter problem is given in the support vector context. The problem, in its simplest formulation, is stated as follows.

Given a parameterization

$$f(\mathbf{x}) = \mathbf{w}^T \mathbf{x} + b \quad (5.11)$$

find \mathbf{w} which minimizes the following functional

$$\min_{\mathbf{w}} \left\{ \frac{1}{2} \mathbf{w}^T \mathbf{w} + C \frac{1}{n} \sum_{i=1}^n \max(|y_i - f(\mathbf{x}_i)| - \varepsilon, 0) \right\} \quad (5.12)$$

The second term of the functional is an ε -sensitive cost function and that is where the ε -SVR name comes from. To add generality and flexibility to the method, slack variables ξ_i and ξ_i^* are incorporated. The optimization problem is modified as follows

$$\min \left\{ \frac{1}{2} \mathbf{w}^T \mathbf{w} + C \cdot \sum_{i=1}^N (\xi_i + \xi_i^*) \right\} \quad (5.13)$$

with the following constraints

$$\begin{aligned} y_i - \mathbf{w}^T \mathbf{x}_i - b &\leq \varepsilon + \xi_i \\ \mathbf{w}^T \mathbf{x}_i + b - y_i &\leq \varepsilon + \xi_i^* \\ \xi_i, \xi_i^* &\geq 0, \quad i = 1, \dots, n \end{aligned}$$

which is a quadratic optimization problem with linear constraints. The term $\sum_{i=1}^N (\xi_i + \xi_i^*)$ at [eq. 5.13](#) corresponds to an upper bound to the number of points not belonging to function f where N is the number of points that reside outside the ε margin. Constant C regulates the trade-off between a smoother or a more accurate interpolation function and it is a parameter provided by the user. A large C reduces the training error but increases the model complexity, i.e. the number of support vectors, and therefore the training time and possibly generalization ability. A smaller C results in a less complex model with higher mean square error but possibly resulting in better generalization. The final, dual optimization problem can be formulated as follows:

Find α_i and α'_i , $i=1, \dots, n$ such that the quadratic form J is maximized,

$$J(\alpha_i, \alpha'_i) = \sum_{i=1}^n y_i (\alpha_i - \alpha'_i) - \varepsilon \sum_{i=1}^n (\alpha_i + \alpha'_i) - \frac{1}{2} \sum_{i,j} (\alpha_i - \alpha'_i) (\alpha_i - \alpha'_i) (\mathbf{x}_i^T \mathbf{x}_j) \quad (5.14)$$

subject to

$$\sum_{i=1}^n (\alpha_i - \alpha'_i) = 0$$

$$0 \leq \alpha_i, \alpha'_i \leq C, i=1 \dots n$$

where α_i and α'_i , are the Lagrange multipliers. The objective function J at eq. 5.14 is convex. This means that a unique deterministic solution to the optimization problem can always be found. Figure 5.4 shows a two-dimensional space with a proposed linear interpolation function used for regression whereas in fig. 5.4 the linear ϵ -insensitive loss function can be seen. Various loss functions exist and the interested reader can refer to Gao et al. (2002).

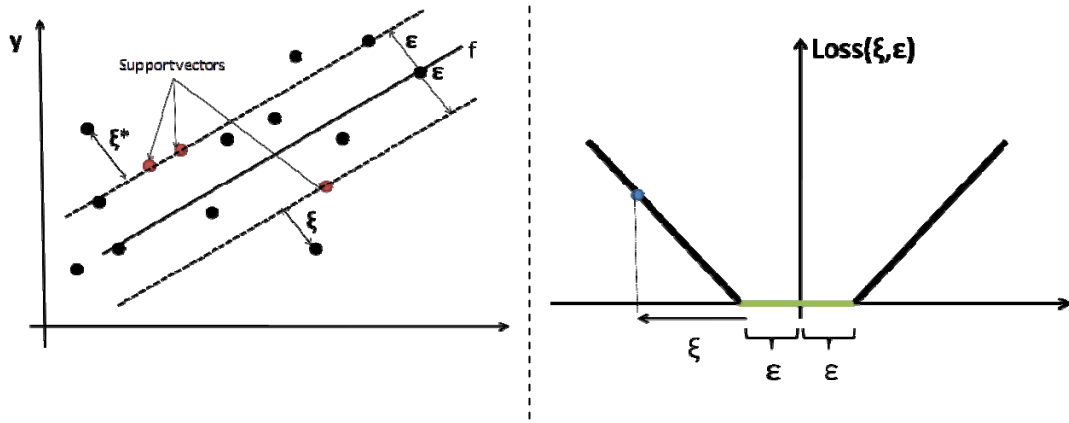


FIG 5.4. On the left the significance of support vectors is denoted regarding a linear regression task. On the right the significance of the ϵ -Loss function.

By solving the optimization problem in eq. 5.14 we get the SVM regression function

$$f(\mathbf{x}) = \sum_{i=1}^n (\alpha_i - \alpha'_i)(\mathbf{x}_i^T \mathbf{x}) + b \quad (5.15)$$

As in the case of SVMs for classification, an extension to non-linear SVR can be achieved using the kernel trick leading to the following solution:

$$f(\mathbf{x}) = \sum_{i=1}^n (\alpha_i - \alpha'_i)K(\mathbf{x}_i, \mathbf{x}) + b \quad (5.16)$$

where $K(\cdot, \cdot)$ is the kernel function. The kernel trick maps the original feature space to a higher feature space where the regression problem can potentially be solved with better efficiency.

There are various possible kernel choices such as the polynomial kernel of degree d

$$K(\mathbf{x}_i, \mathbf{x}) = \left((\mathbf{x}_i^T \mathbf{x}) + 1 \right)^d$$

or the radial basis function kernel

$$K(\mathbf{x}_i, \mathbf{x}) = \exp\left(-\gamma \|\mathbf{x}_i - \mathbf{x}\|^2\right), \quad \gamma \in \mathbb{R}^+ \quad (5.17)$$

Prediction error bounds

A probabilistic treatment of the data is not inherent within the classical SVR framework. The algorithm itself is deterministic in nature and does not provide error bounds at the predicted points. However, a requirement that always arises in machinery prognostics is that of quantifying the uncertainty for the model's RUL prediction providing not only the mean estimation but also a confidence associated with the prediction. In the field of CBM [Gao et al.\(2002\)](#) provide an estimation of the prediction error variance in a SVR framework for diagnostic purposes in critical components in nuclear industry. In RUL predictions, probabilistic SVR has not been yet utilized in the literature as far as we know.

The prediction error for each prediction $f(\mathbf{x}_i)$ in general, consists of two terms. The first term is an estimation of the data inherent "noise" and accounts for the model's temporal uncertainty. In order to find a measure of this temporal noise component, a Bayesian approach is applied. This Bayesian noise estimation takes place on the training data set and it assumes similar noise characteristics throughout the training data. Assuming $\Pr(f(\mathbf{x}_i) | X)$ is the posterior probability density function of the predictions $f(\mathbf{x}_i)$ at the point \mathbf{x}_i given the training dataset $X=\{\mathbf{x}_i, f(\mathbf{x}_i)\}$, then with the application of Bayes rule

$$\Pr(f(\mathbf{x}_i) | X) = \frac{\Pr(X | f(\mathbf{x}_i))\Pr(f(\mathbf{x}_i))}{\Pr(X)} \quad (5.18)$$

$\Pr(X | f(\mathbf{x}_i))$ is the likelihood of the model and is assumed to be an exponential function over the SVR loss function (as depicted in [Fig. 5.4](#)). $\Pr(f(\mathbf{x}_i))$ is the prior probability density function which is assumed a zero mean Gaussian process. $\Pr(X)$ is called the evidence of the model. It is actually the likelihood of the model integrated over all training set model prediction values. The exact calculation of every term of Eq. (5.18) is impossible unless a good choice of a prior took place. The most frequent choice for a closed form pdf is the normal distribution. This way an estimation of the posterior probability as a Gaussian probability distribution with variance $\sigma_{C(\varepsilon)}^2$

$$\sigma_{C(\varepsilon)}^2 = \frac{2}{C^2} + \frac{\varepsilon^2(3 + \varepsilon C)}{3(\varepsilon C + 1)} \quad (5.19)$$

C and ε are the hyperparameters of the SVR model.

The second term of the prediction uncertainty is associated to the unit to unit variability. Assuming a normal prior, Gao provides the following form for the unit to unit uncertainty level.

$$\sigma_K^2(f(x_j)) = \sigma_{M+1,M+1}^2 - \mathbf{\Sigma}_{(1:M),M+1}^T \mathbf{\Sigma}_{(1:M),(1:M)}^{-1} \mathbf{\Sigma}_{(1:M),M+1}$$

Where

$$\mathbf{\Sigma}_{M+1,M+1} = \begin{pmatrix} \mathbf{\Sigma}_{M,M} & \mathbf{\Sigma}_{(1:M),M+1} \\ \mathbf{\Sigma}_{(1:M),M+1}^T & \sigma_{M+1,M+1}^2 \end{pmatrix}$$

is the covariance matrix. This is calculated on the M model parameters which are the support vectors plus the j th prediction point \mathbf{x}_j^* . The covariance matrix has dimensions $(M+1) \times (M+1)$ where M is the number of the marginal vectors of the particular model. $\sigma_{M+1,M+1}^2$ is a scalar and is the value of the covariance matrix at position $(M+1, M+1)$. $\mathbf{\Sigma}_{(1:M),M+1}$ is a vector positioned at the right end of the covariance matrix and $\mathbf{\Sigma}_{1:M,1:M}$ is the sub-matrix spanning from $(1:M, 1:M)$. The variance term due to model uncertainty and the term due to feature space noise are superimposed at every prediction point. Assuming a mutual independence between the temporal and unit to unit variability we can calculate the total variance at each test point as the sum of the latter two terms

$$\sigma^2(f(x_j)) = \sigma_K^2(f(x_j)) + \sigma_{C(\varepsilon)}^2 = \tag{5.20}$$

$$\frac{2}{C^2} + \frac{\varepsilon^2(3 + \varepsilon C)}{3(\varepsilon C + 1)} + \sigma_{M+1,M+1}^2 - \mathbf{\Sigma}_{(1:M),M+1}^T \mathbf{\Sigma}_{(1:M),(1:M)}^{-1} \mathbf{\Sigma}_{(1:M),M+1}$$

5.3.2 Application in gear failure data

We assess the effectiveness of ε -SVR on the simulated data base of [Section 5.2](#) (see [table 5.2](#)). 15 failure histories are chosen to train and validate the model which is tested on 2 out of sample failure histories.

The input data will comprise of the simulated RMS histories (1-D data). As presented in [Section 4.8](#) the data are inherently scaled with a small unit to unit variability concerning the maximum of the AE RMS value. The output is the corresponding normalized Actual Remaining Useful Life (ARUL) of each component $ARUL \in [0,1]$.

The optimization problem of [eq. 5.15](#) is solved with the Sequential Minimisation Optimisation (SMO) algorithm realized in libsvm Matlab® toolbox. More particularly we choose to extend to a non-linear SV problem since we know a priori that there is a non linear

underlying mechanism associated with gear fatigue mechanism. To this end we choose the radial basis (RB) kernel to augment our model (eq. 5.17). Parameters C and ε are tuned on a validation set comprising of 5 out of the 15 training histories via a grid search. For the grid search the latter parameters take the following values.

$$\varepsilon \in [0.0001, 0.0005, 0.001, 0.005, 0.01, 0.05, 0.1, 0.5]$$

$$C \in [0.0001, 0.0005, 0.001, 0.005, 0.01, 0.05, 0.1, 0.5, 1, 3, 5, 10, 20, 30]$$

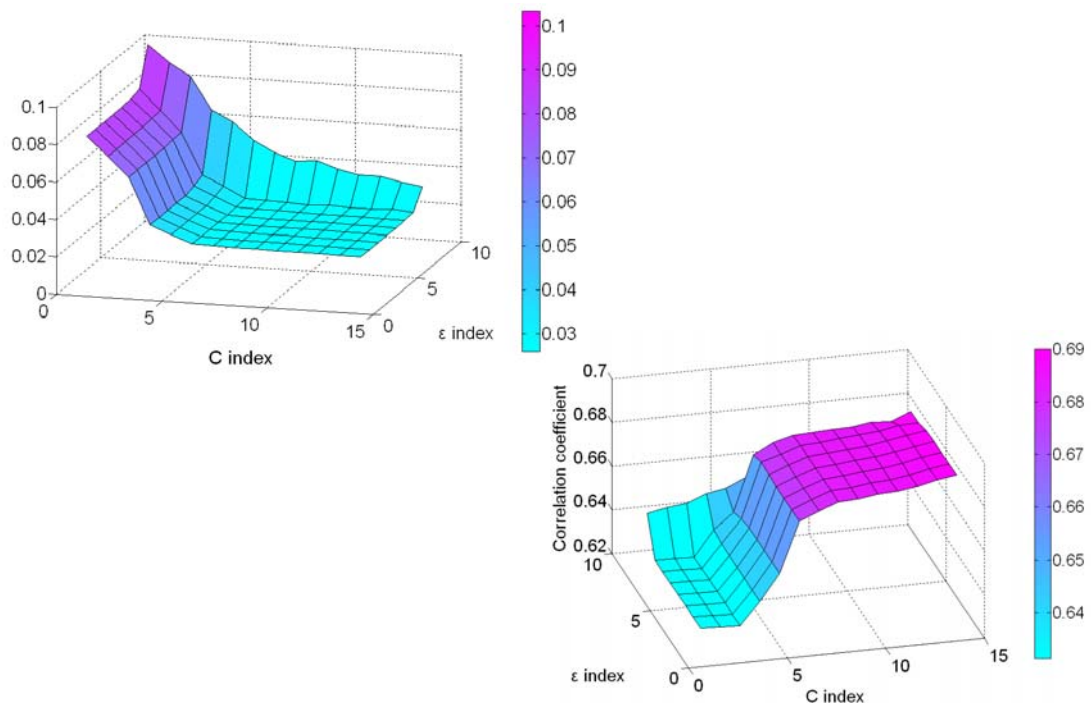


FIG. 5.5. Mean square error (MSE) (left) and cross correlation coefficient (CCC) (right) for various values of ε and C model parameters. The exact values of the respective indexes are shown on the previous page.

We need a model that performs well on the validation set in means of MSE and CCC. To this end, and based on fig 5.5, we choose $C=10$ and $\varepsilon=0.1$. The latter yield an $MSE \cong 0.2$ and $CCC \cong 0.68$ on the validation set. The value of the RB γ parameter is chosen 0.5. Experimenting with different values of γ parameter did not yield significantly different results. Table 5.3 summarises the training and validation results.

The final optimized model	
Number of training histories	10(training) +5(validation)
Number of training samples	2112(training)+851(testing)
C	10
ϵ	0.1
γ	0.1
MSE (validation set)	0.2
CCC (validation set)	0.68
Number of support vectors (training set)	275

Table 5.3

One can observe in [table 5.3](#) that the number of support vectors which coinciding with the number of model parameters is annoyingly high. In fact, almost one out of ten training samples is kept by the SMO optimization algorithm as a support vector. It should be mentioned though that SVRs are not designed with a demand of sparsity (for a sparse solution to the problem of SVRs see [Di Maio et al. \(2012\)](#)). We will not expand to the issue of SMO sparsity and will only study the accuracy and error bound prediction of the algorithm.

To test our SVR model we choose two unseen failure histories. The output of the model though is the predicted normalized RUL and therefore scaled in the interval [0,1]. To convert this value to the appropriate time scale we multiply the result with the mean life of the training data set. We may use the results from [table 5.1](#) assuming a sheer Weibull degradation process and obtain $MTTF_0=186.4$. Error bounds are calculated according to [eq. 5.20](#) on every measurement sample. The covariance matrices are computed taking into account the RB kernel function. The final results are plotted in [fig. 5.6](#) for two cases.

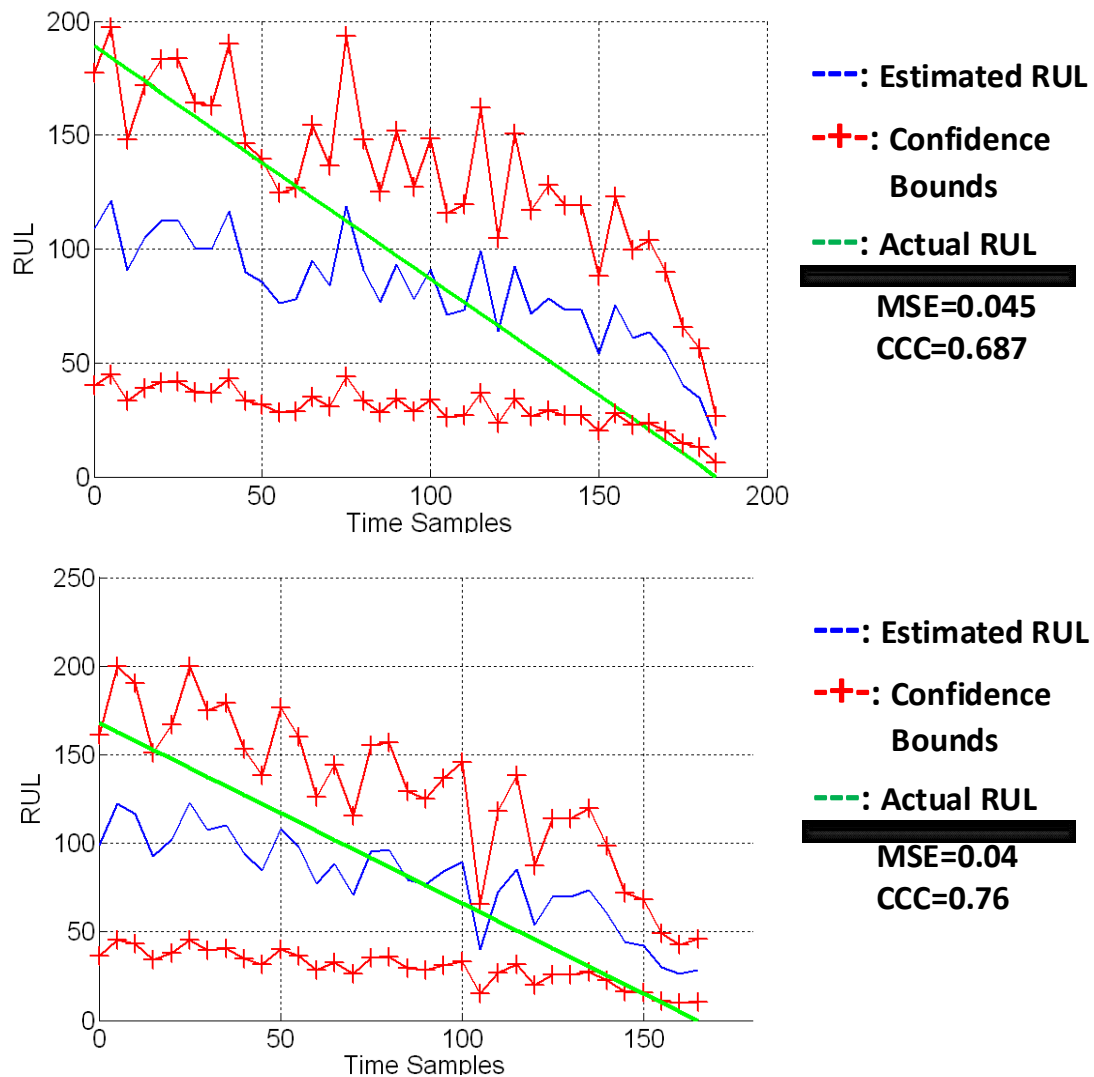


FIG. 5.6. Two prediction results with the probabilistic ε -SVR technique. The 95% confidence bounds (upper and lower) are also depicted. The straight line corresponds to the ARUL for comparison. The corresponding MSE and CCC scores are also denoted on the graph.

The error bounds (cross-marker line) correspond to a distance 1.96σ from the point prediction. This is roughly a 95% confidence interval since we have assumed a gaussian framework.

The prediction is well below the actual value for most of the component's operational life. This is a prediction on the safe side since under estimation of RUL does not produce catastrophic maintenance decisions. Moreover the algorithm produces meaningful predictions from the very beginning of the failure experiment. The $\pm 1.96\sigma$ bounds include the ARUL line although they are annoyingly wide. This means that the precision of the prediction model is low. Very close to the life end the predictions become more accurate and the error bounds become narrower.

5.4 Probabilistic exponential extrapolation scheme

In this section, a bootstrap method for historic data multiplication is proposed. Let's take the available timeseries $x(t_i)$, $i=0, \dots, (n_{\text{obs}}-1) \cdot t_s$. If the feature is densely sampled ($t_s \ll t_{\text{fail}}$), it can be assumed that there exists a small discrete time window of length n_{int} within which the dynamics of the asset's deterioration mechanism do not substantially change and the same should hold for the measured condition monitoring feature. In other words, within this small interval, the health measurements are assumed as stochastic variables that are distributed identically.

Figure 5.7 depicts a model timeseries. In this case the small time interval is defined as $n_{\text{int}}=3$ samples. These successive windows are truncated and put side by side, creating a vector timeseries \mathbf{x}_{new} of size $n_{\text{int}} \times \lfloor n_{\text{obs}} / n_{\text{int}} \rfloor$, the symbol $\lfloor \cdot \rfloor$ meaning rounding to the nearest integer lying in the interval $(-\infty, n_{\text{obs}} / n_{\text{int}}]$. Next, a sample is randomly drawn from each vector of the new data set, together with its time stamp, as shown in fig. 5.7. When the process of sampling $\lfloor n_{\text{obs}} / n_{\text{int}} \rfloor$ consequent values is finished the result will comprise a snapshot of the original dataset bearing the same dynamics. Running this algorithm k times, a pool of k timeseries can be created to form a statistically adequate distribution. Obviously, a maximum of $n_{\text{int}}^{\lfloor n_{\text{obs}} / n_{\text{int}} \rfloor}$ timeseries snapshots can be created, potentially providing a very dense distribution.

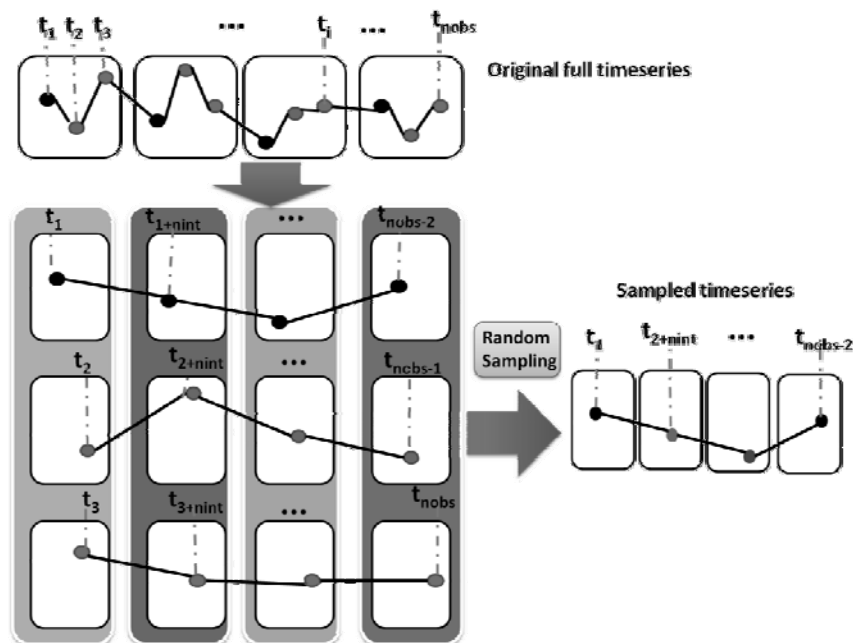


FIG 5.7. Diagram for the bootstrap sampling process

Remaining useful life distribution

Figure 5.8 depicts the conceptual time progression of a degradation signal taken from a model operating asset. The time scale is in arbitrary units. An observation window is defined on this condition monitoring feature. On this window, a number of exponential functions are trained according to the strategy of the earlier paragraph. Then these exponentials are extrapolated as shown in the figure. The \widehat{RUL} distribution can be defined in two steps.

Step 1: The intersection points of the future operational time line (vertical line) with the extrapolated exponentials give the distribution of the condition monitoring feature (CMF) value provided that this time point is reached. Thus our best knowledge can be summarized in a discrete empirical density function $p(\text{CMF}|t=t_i) = 1/k_{\max} \sum_{k=1}^{k_{\max}} \delta(y - \text{CMF}_k)$ where k_{\max} is the number of sample points and $\delta(y - \text{CMF}_k)$ equals to one when $y = \text{CMF}_k$. When a point estimator, such as the sample median, passes a critical degradation threshold it will mean that the critical degradation point has been reached.

Step 2: The intersection points of the exponential functions with a horizontal (see fig. 5.8 “degradation level”) line provide another distribution. This will be the \widehat{RUL} estimator distribution provided that the particular degradation level has been reached. Again our best knowledge can be summarized in a discrete empirical density function $p(\widehat{RUL}|y=y_{\text{thres}}) = 1/k_{\max} \sum_{k=1}^{k_{\max}} \delta(t - t_k)$ where y_{thres} is a degradation level threshold and $\delta(t - t_k)$ equals to one at the k intersection time stamp, $t=t_k$. For the current study, the important “degradation level” line is the one that defines the failure threshold level and it is usually decided from a priori expert knowledge.

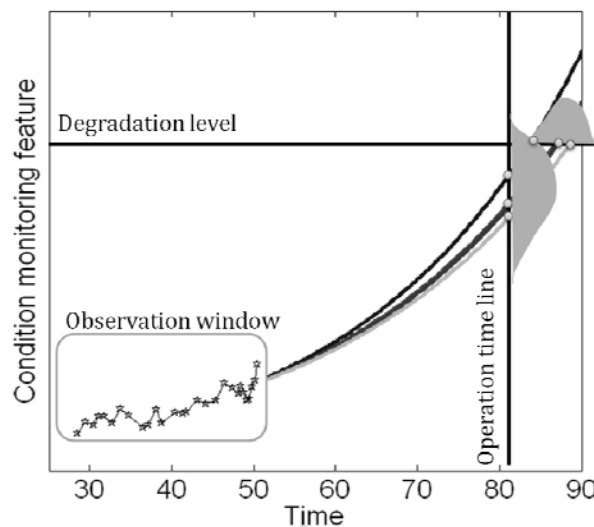


FIG. 5.8. Visualization of the interpolation process

5.4.1 Application in gear failure data

The strength of the latter method is that it does not require any prior knowledge regarding the failure mechanism of the particular monitored asset. This is very useful in cases where the historic failure data are very limited, that is in cases where [Section 5.2](#) and [5.3](#) strategies are not applicable. This is the case with very large and critical industrial components. On the other hand it is only able to incorporate the temporal component of the uncertainty which, with the strategy of the earlier paragraph, it transforms to a RUL distribution at every observation instant. The proposed method is going to be quite handy when we

The latter method is going to be applied for RUL estimation on the same two test failure tests as in the earlier sections. An exponential model is going to be fitted at each observation instant.

$$f(t) = \alpha + \beta \exp(\gamma t)$$

[Gasparin et al. \(2011a\)](#) have applied linear functions to derive their estimates on gearbox failure data. However, linear functions are deemed rather confining by the author of the current thesis. Exponential functions provide more flexibility particular for systems that exhibit some non linearity. As has been shown gearbox exhibit strong non linear degradation close to the end of life as shown in [Chapter 4](#).

According to the bootstrap method of the previous paragraph, 35 exponential functions are going to be fitted at each observation instant with short time interval (sample interval) of length $n_{int}=3$ subsequent samples and an interpolation interval of $n_{obs}=20$ samples. This way only the local dynamics of the degradation are kept for RUL estimation. The α , β , γ parameters of the exponent are derived with an improved gradient descent algorithm, the Levenberg-Marquadt (LM).

Prediction Results

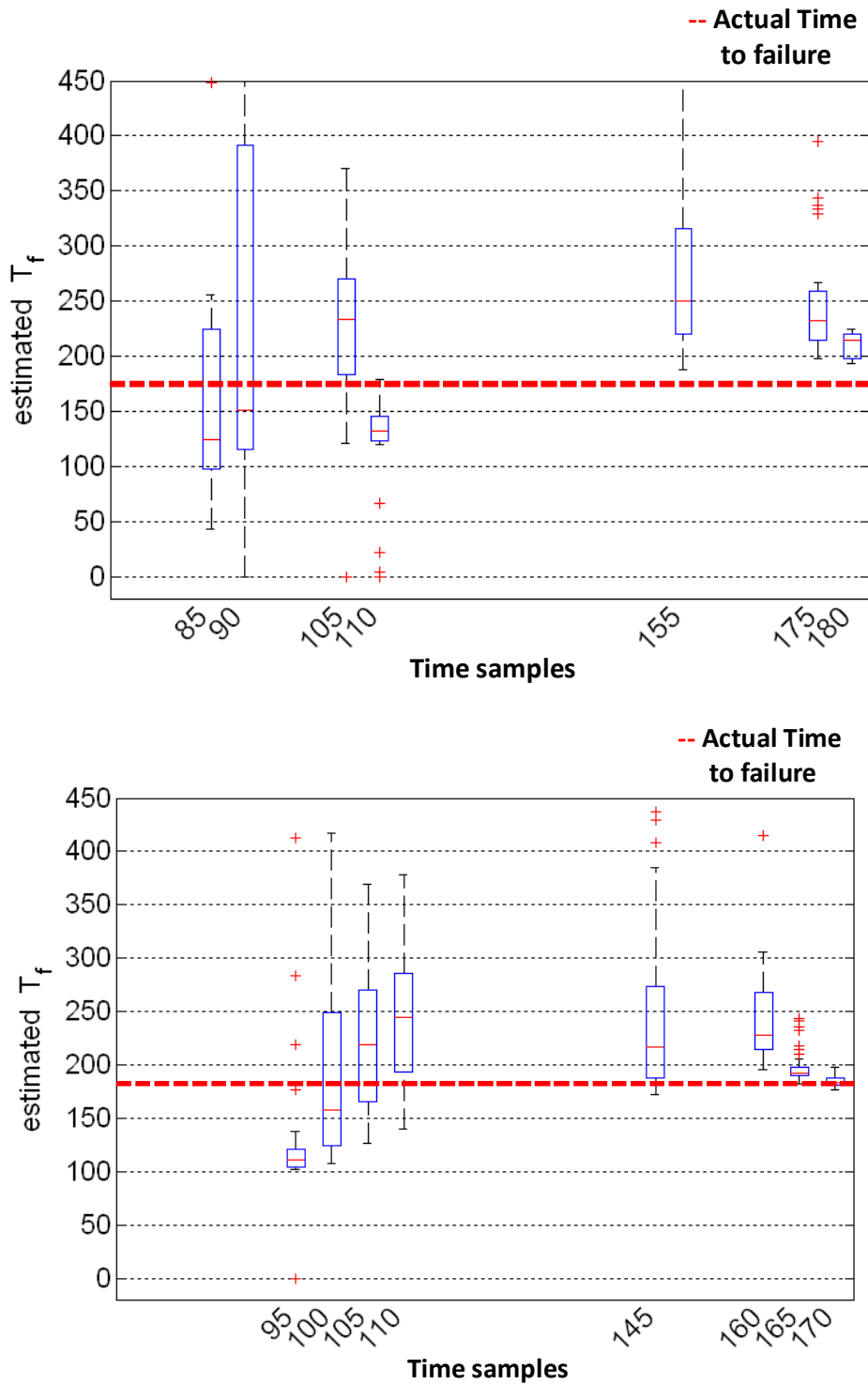


FIG. 5.9. Predicted time of failure for two test gear failure histories.

Fig. 5.9 depicts the predictions made with the extrapolation strategy on two test failure histories. The results are presented in boxplot. The median of the prediction provides the RUL point estimation. The box edges provide 25 and 75 percentile of the prediction

variance. The whiskers provide the maximum and minimum points not considered outliers by the plotting tool. It seems that the method provides a reasonable prediction along with meaningful error bounds as soon as ~60% of the total operational life for both failure histories. The predictions become more consistent near the end of life of the components where the exponential function effectively captures the end of life especially in “history 3” (for $t_{\text{obs}}=160, 165, 170$, with actual $T_f=172$). On the other hand it is observed that in most cases error bounds are exceedingly wide. Moreover, the exponents failed to provide a good prediction for a wide operating interval (110-155 time samples for “history 1”, 110-145 time samples for “history 3”) and thus the void in the prediction graphs of [fig. 5.9](#). This void is attributed to the inability of the exponential functional to reach the failure threshold provided the respective time window.

5.5 Discussion

Three RUL prediction methods have been assessed for gear failure-like data. Each method exhibited various strengths and limitations. Two are tested for the first time for gear failure data. The third method is an extrapolation scheme enriched with a bootstrap sampling method that bypasses the lack of a prior on asset failure histories. This method for producing confidence intervals in total absence of a prior is proposed for the first time in the context of gear failure prognostics.

- ❖ An age dependent model was studied in [Section 5.2](#) and more particularly a Weibull aging PHM. To the author’s knowledge this is the first time that such a model is assessed in a gearbox failure framework. The predictions of this model fell on a straight line, almost parallel to the actual life line. The prediction curve diverged only slightly to the actual value near the end of life of each tested component. This was due to the impact of the condition monitoring feature. The RUL prediction intervals were very narrow although the variance of the corresponding model parameters was not insignificant. Also, PHMs need an adequate pool of failure events to derive a meaningful MTTF for ERUL estimation as depicted in [eq. 5.8](#). It is the view of the author that PHMs could provide more information relative to gear degradation if more covariates were incorporated in the model.
- ❖ Probabilistic ε -SVR was tested on the simulated data base in [Section 5.3](#). This model is derived from machine learning literature. To the author’s knowledge this is the first time that such a Probabilistic ε -SVR is studied in a gearbox failure framework. Contrary to the aging model results, the estimated prediction intervals (95 percentile) was wide enough and included the actual RUL almost throughout the component’s life, probably too wide though. Moreover the RUL predictions followed in general well the level of degradation although the accuracy was small particularly in the

linear degradation phase giving meaningful estimation right from the beginning of the asset operation. In fact it underestimated the actual RUL in most of the components life, a result which is on the safe side of CBM decision systems. The algorithm became more effective very close to the time of failure of the component. Also ϵ -SVRs also require some adequate failure events pool.

- ❖ An updating probabilistic extrapolation model was investigated in [Section 5.4](#). The method for acquiring the prediction interval was based on bootstrap sampling of past measurements of the asset's health metric. This method is proposed for the first time to the author's knowledge. The application of this method on gear failure data showed some potential providing reasonable predictions as soon as in 60% of operational life. However, the linearity of the gearbox degradation metric (in ~80% of the asset's operating life) proved a limiting factor in most tested histories. The biggest advantage of this method is that it does not require any prior knowledge of the asset degradation process and only requires limited information regarding the failure threshold. Due to this advantage the extrapolation family of predictions is going to be further investigated in [Chapter 7](#).
- ❖ From the study of this chapter the preferable mono-parametric prognostics algorithm would be ϵ -SVR because it produces meaningful predictions from the very beginning of the lifelong experiments with adequate error bounds. In absence of failure prior though, it seems that an extrapolation scheme is more applicable although its performance is quite poor. In the next we will discuss an improvement of this method by adding feature extraction and feature fusion. All in all, it seems that the incorporation of a single health metric from a single sensor can only take us so far regarding the RUL estimation in gearboxes. In the next chapter the potential of multiple sensors and multiple health metrics combination is going to be assessed and compared to the results of [Chapter 5](#).

CHAPTER 6

Estimating the Remaining Useful Life in Gearboxes with feature fusion

6.1 Introduction

Feature extraction and feature fusion has been shown to enhance machine learning algorithms and methods ([Guyon and Eliseef \(2003\)](#)). The examples in machine learning and engineering literature incorporating feature fusion in classification tasks are numerous. On the other hand the application of feature fusion for machinery prognostics is very limited if non-existent. This chapter is going to contribute to this area in the context of gear failure prognostics.

This chapter is based on the gear failure data base developed in the current framework and was briefly presented in [Section 4.5](#). Several features are going to be extracted from waveform type sensors in [Section 6.2](#). The features that can better explain the gradual

gear degradation are going to be kept for further processing. Two basic fusion techniques, Principal and Independent Components Analysis (PCA and ICA) are also presented in this section. Both are going to be assessed for prognostic purposes, ICA proving to be a better candidate. Finally, an extrapolation prediction scheme is going to be applied, taken from [Section 5.4](#). This scheme is chosen because it can provide adequate results in cases of limited failure histories such as our own. [Section 6.3](#) concludes this chapter highlighting the more important findings and contributions. Some results from this chapter have been published in [Loutas et al. \(2011\)](#).

6.2 Feature extraction

Several features were extracted from the failure gear data base. Time domain, frequency domain and wavelet domain features were extracted from the most interesting waveform sensors, AE channel 1 and xyz-axis vibration. The list of time and frequency domain features that were extracted can be found in Appendix A.

The motivation for testing wavelet transform in our waveforms came from the observation that certain frequency bands showed diagnostic potential as depicted in [fig 4.10](#). Moreover, [Eftekhrejad \(2011\)](#) has already shown that wavelet transform provide diagnostic information relative to gear fatigue damage progression. To extract the power content of particular bands we could either apply:

- i. a Short Time Fourier Transform (STFT) on the timeseries and estimate the power within a particular frequency band or
- ii. Wavelet Packet (WP) or Discrete Wavelet Decomposition which is an elegant means to extract frequency band information based on the theory of filter banks ([Mallat \(1999\)](#)). Wavelet transform of a 1-D real signal is the atomic decomposition of the signal in scaled and shifted versions of a basis function called the mother wavelet.

$$W(\alpha, \beta) = \alpha^{-1/2} \int_{-\infty}^{\infty} s(t) \psi^* \left(\frac{t - \beta}{\alpha} \right) dt \quad (6.1)$$

In practical signal processing a discrete version of wavelet transform is often employed by discretizing the dilation (parameter α in [eq. \(6.1\)](#)) and translation (parameter β in [eq. \(6.1\)](#)). In general, the procedure becomes much more efficient if dyadic values of these parameters are used leading to the well known DWT. The wavelet packet transform (WPT) is an optimal sub-band tree structuring that yields a tree structure with each “branch” being a band-passed and down-sampled instance of the original signal.

We choose WP over windowed FFT because it does not require the definition of an appropriate window function. Moreover, it is more elegant plus Matlab Software provides a

fast and efficient algorithmic implementation of the decomposition process, the Fast Wavelet Transform. In the current framework we choose 3 levels of WPT decomposition based on the ‘db10’ mother wavelet. This yields 2^3 band passed waveforms according to [table 6.1](#). Finally the energy of each band is extracted and normalized with the total energy of the respective waveform. This way we get features inherently scaled in [0,1].

WP terminal node	AE frequency band (KHz)	Vibration frequency band (KHz)
1	[0, 156.2]	[0, 3.125]
2	[156.2, 312.5]	[3.125, 6.25]
3	[312.5, 468.7]	[6.25, 9.375]
4	[468.7, 625]	[9.375, 12.5]
5	[625, 781.2]	[12.5, 15.625]
6	[781.2, 937.5]	[15.625, 18.7]
7	[937.5, 1093.7]	[18.7, 21.87]
8	[1093.7, 1250]	[21.87, 25]]

Table 6.1. WP bandwidth.

Finally, [table 6.2](#) summarizes the extracted features.

Sensors	1 AE and 3 vibration
Time domain features	6
Frequency domain features	12
WP domain features	8
Total extracted features	26x4=104
Per failure experiment	

Table 6.2. Extracted features

All features are extracted from waveforms coming from actual meshing gearboxes.

It is worth mentioning that gear specific features, such as FM0, FM4 (see [Dempsey \(2003\)](#)) was impossible to obtain due to the inability to extract a meaningful TSA signal from the vibration waveform as has been noted in [Section 4.6](#).

Choosing the best performing features.

Not all of 104 features of [table 6.2](#) are fit for prognostic purposes. To this end we propose the following feature acceptance method.

First of all we want to choose a likeness criterion. This could easily be the coefficient of cross correlation (CCC). The question is with what curve we are to compare the extracted feature set. We propose to compare the particular features with the normalized remaining life line for every failure experiment. The more similar a feature is with the latter life line, the better it describes the decrease of the operational life and thus it constitutes a good prognostic metric. We further choose to place a lower acceptance threshold of $CCC=0.67$. Every feature that gives a $CCC>0.67$ is accepted for further study. The particular criterion is applied on three experiments, experiment III, V and VI, as denoted in [table 4.1](#). And the reason is that they were the more complete from a data acquisition point of view. The list of features that passed the life correlation criterion is shown in [table 6.3](#).

Sensor	Accepted features
AE ch1	TD1, TD2, FD3, FD4, FD5, FD9, FD10, WP1
Vib x-axis	TD1, FD1, FD3, FD4, FD5, FD9, FD10, WP2, WP4, WP5
Vib y-axis	TD1, FD1, FD3, FD4, FD5, FD8, FD9, FD10, WP2, WP4
Vib z-axis	TD1, FD1, FD3, FD4, FD5, FD9, FD10, WP2, WP4, WP5
Total selected features	38

Table 6.3. Selected prognostic features.

Some of the latter features are depicted in [fig. 6.1](#).

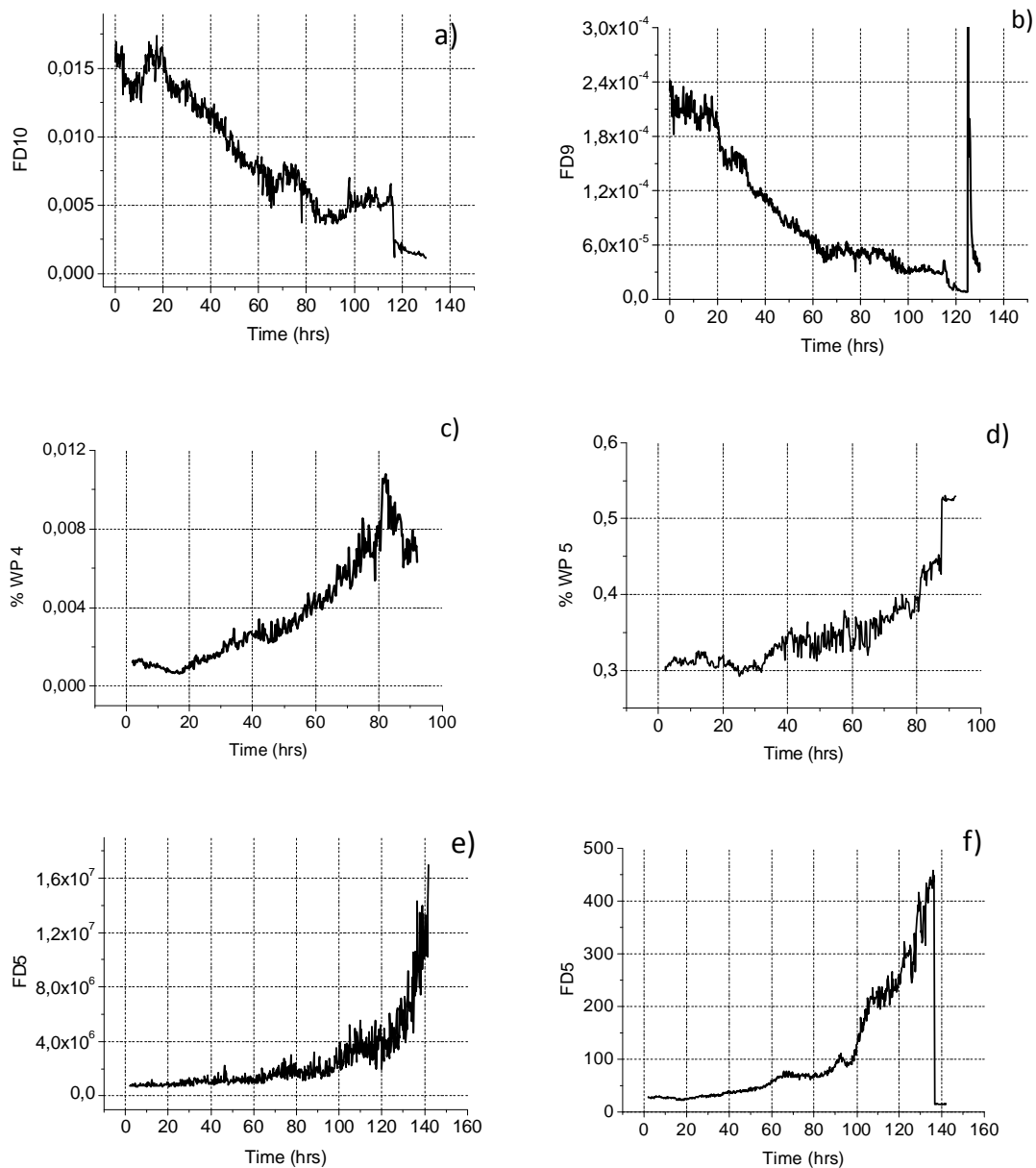


FIG. 6.1. Figure a) is derived from vibration ch1 (experiment V), b) vibration ch3 (experiment V), c) vibration ch1 (experiment VI), d) vibration ch2 (experiment VI), e) AE ch1 (experiment III), f) vibration ch2 (experiment III).

There were several features that could catch the critical failure point of the asset but were not good descriptors of the reduction of RUL such as kurtosis and skewness of time and frequency domain thus highlighting the fact that a good diagnostic feature is not necessarily a

good RUL prediction metric. It is worth mentioning that some features, such as FD9 and FD10 are highlighted for the first time as good gear failure predictors.

The final number of 38 features is deemed excessive. Moreover correlations were observed among the extracted features. In order to derive a more concise feature set we further process the extracted features with tools from the field of pattern recognition. After all the various problems arising from high dimensionality of feature space has been discussed in various contexts (Bishop (1995), Guyon and Eliseef (2003)). In the current work, PCA and ICA are going to be assessed for enhancing the gear failure feature set towards efficient gear failure prognosis.

Principal Components Analysis

Let's assume that a number of monitored signals are provided, $\mathbf{x}_{M \times N} = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N]$ being the matrix that holds N M -dimensional signals. In this case reducing the dimension M can be thought as discovering a low dimensional co-ordinate system in which we can approximately represent the data. Then we can express each space point \mathbf{x}_i as

$$\mathbf{x}^i \approx \mathbf{c} + \sum_{j=1}^R \mathbf{y}_j^i \mathbf{b}^j \equiv \tilde{\mathbf{x}}^i \quad (6.2)$$

Where R is the reduced number of dimensions, \mathbf{b} is a vector of R weights that act on a low dimensional space representation \mathbf{y} of \mathbf{x} and \mathbf{c} is constant point in the linear subspace and in some contexts it is omitted. Eq. 6.2 describes the reconstruction of the original data point \mathbf{x}_i . The main problem in PCA is what should those directions \mathbf{b} optimally be? One good choice is to find directions along which the energy, or variance, is maximised. Once these directions are defined then the new transformed data are simply the corresponding projections of the original data on these linear space directions as depicted in Fig. 6.1

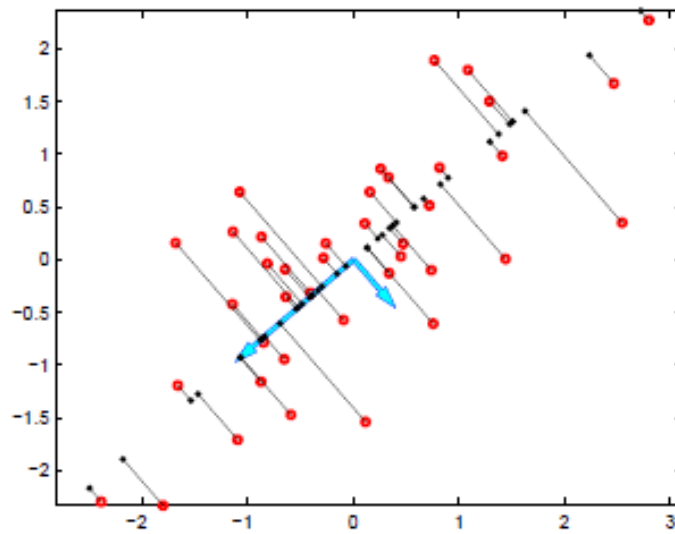


FIG. 6.2. A 2-D data set (large circles) is projected on the derived maximum variance directions denoted with two vector arrows. Their projections are denoted with the smaller dots.

PCA algorithm provides optimal directions in mean square terms and it is realized in the following steps.

- Assuming an original data set $\mathbf{x}_{M \times N}$

Step 1: After forming the data matrix of interest, centre the data points to zero subtracting the mean and scale each signal to some standard interval i.e. $[-1,1]$. This way “small” signals containing significant information will not be masked by “larger” signals, in terms of energy that is.

Step 2: Let’s assume $\Sigma = E\{\mathbf{x}\mathbf{x}^T\}$ is the data covariance matrix. Then find the eigenvalues of the covariance matrix and sort them in decreasing order. This can be done with the Singular Value Decomposition method which is able to handle singularities in the covariance matrix. The eigenvalues are the variances of the underlying linear subspace. Then keep only the variances (and therefore directions) that sum up to a percentage of the total variance or energy, typically 95% of the total variance.

Step 3: The approximate “reduction” of each original data point will be

$$\tilde{\mathbf{x}}_i = \mathbf{E}\mathbf{D}^{-1/2}\mathbf{E}^T \mathbf{x}_i \quad (6.3)$$

where \mathbf{D} , \mathbf{E} are the eigenvalue decomposition matrices from Step 2. In Matlab environment the decomposition takes place with the Singular Value Decomposition (SVD) algorithm. $\mathbf{E}\mathbf{D}^{-1/2}\mathbf{E}^T$ is also called the whitening matrix because it de-correlates the original signal sources.

Independent Components Analysis

Let's assume that a number of monitored signals are provided. It is assumed that those signals are a linear combination of some original, statistically independent sources. These original sources are mixed through a linear transformation which in matrix notation is $\mathbf{x}=\mathbf{A}\mathbf{s}$, where $\mathbf{x}=[x_1, x_2, x_3 \dots]$ is the matrix of the observed mixed signals, $\mathbf{s}=[s_1, s_2, s_3 \dots]$ is the unknown original source signals and \mathbf{A} is an unknown full rank matrix called the mixing matrix.

The original sources are considered independent. Independence of a number of signals means that the information carried by one signal cannot be inferred from the others.

In ICA we search for an invertible matrix \mathbf{W} such that $\mathbf{y}=\mathbf{W}\mathbf{x}$ where \mathbf{y} is an estimation of the initial source matrix \mathbf{s} and \mathbf{x} is the mixed signal matrix.

In order for the ICA algorithm to be utilized, the removal of any correlation among the data is needed. This procedure is called "whitening". After de-correlating the data, the original sources can be acquired through a simple rotation. The rotation matrix, \mathbf{W} , is calculated iteratively. The criterion of the iteration update is the maximization of "non-normality" of the signals' distribution. There are several measures of "non-normality". Statistical kurtosis is one such measure. Maximizing kurtosis is a way to achieve statistical independence between signals. Another criterion is the minimization of mutual information. The calculation of mutual information is based on the signals' cross entropy estimation. The cross entropy or neg-entropy of two probability distributions p_1 and p_2 of a random variable y is defined as:

$$J(y) = -\int p_1(y) \log p_2(y) dy \quad (6.4)$$

Since straight forward cross entropy calculation is difficult, especially when the probability densities are unknown, several indirect estimations of cross entropy exist. One such reduces the latter formula to:

$$J(y) \approx [E\{G(y)\} - E\{G(y)\}]^2 \quad (6.5)$$

Where y is assumed to be of zero mean and unit variance and v is assumed to be a Gaussian distribution with zero mean and unit variance. $E\{\}$ is the average operator. A proper selection of the function G can provide a very good estimation of neg-entropy.

ICA algorithm is applied in order to find an invertible transformation \mathbf{W} that minimizes this mutual information function J . In other words, this is equivalent of *finding*

directions in which the cross entropy is minimized. The steps of a fixed point ICA algorithm, which is proposed by Hyvärinen (1999) are described briefly.

Fastica algorithm

After whitening our data matrix then

Step 1: Choose an initial (random) weight vector \mathbf{w}

Step 2: Let $\mathbf{w}_{\text{new}} = \mathbf{E}\{\mathbf{x}g(\mathbf{w}_{\text{old}}^T \cdot \tilde{\mathbf{x}})\} - \mathbf{E}\{g'(\mathbf{w}_{\text{old}}^T \cdot \tilde{\mathbf{x}})\}\mathbf{w}_{\text{old}}$. This is an approximate

Newton iteration step

Step 3: Let $\mathbf{w}_{\text{new}}^{\text{norm}} = \frac{\mathbf{w}_{\text{new}}}{\|\mathbf{w}_{\text{new}}\|}$

Step 4: If $\|\mathbf{w}_{\text{new}}^{\text{norm}} - \mathbf{w}_{\text{old}}\|^2 < \epsilon$ then demixing matrix $\mathbf{W} = \mathbf{A}^{-1}$ is obtained, in an opposite case the procedure returns to step 2. ϵ is a sufficiently small number defined by the user as a stopping criterion. g is a non-linear estimate of entropy.

It is worth mentioning that although ICA is characterised as an unsupervised fusion method, due to its stochasticity in Step 1 it is deemed necessary for an expert to judge whether the decomposition provides meaningful results. Sometimes it takes more than one run of the latter algorithm to derive meaningful independent sources.

6.2.1 Results

The features presented in table 6.3 are arranged in an appropriate data matrix. In order to reduce the dimensionality of this feature space we applied both PCA and ICA. To test the potential of generalisation for each method the algorithms were run on a training experiment, in our case experiment V from table 4.1. The resulting transformation matrices were applied on a testing data set, in our case experiment III.

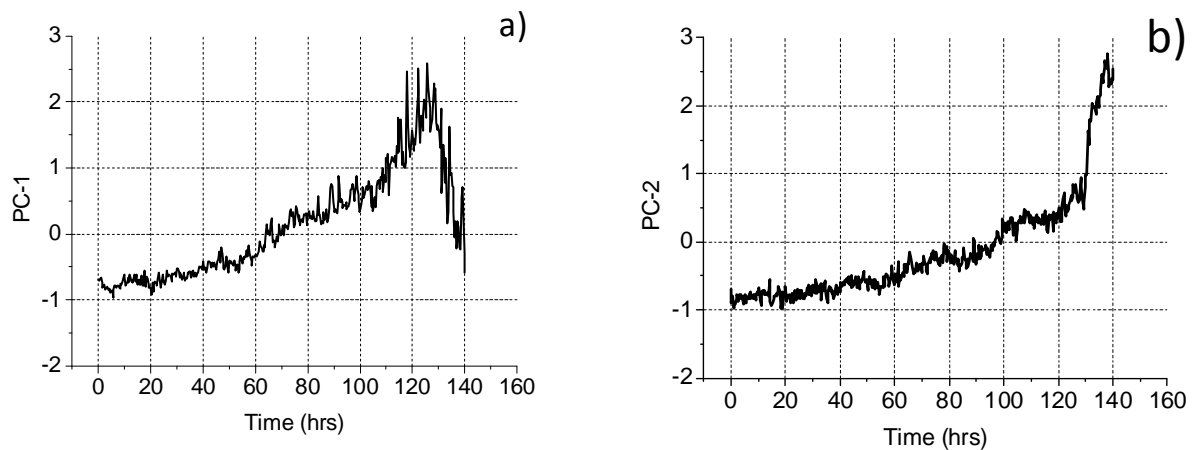


FIG. 6.3. Fused health monitoring metrics via PCA.

Applying the PCA algorithm on the training data set it was found that the first two principal components held 92% of the total variance. The whitening matrix (eq. 6.3) was also kept and was applied on the testing data set. The resulting principal components from experiment III are depicted in fig. 6.3.

ICA transformation matrix was also defined on the training data set. It should be mentioned that ICA optimization algorithm is applied on the whitened data which are provided from the PCA algorithm. The demixing IC matrix is applied on the testing data set and the results are depicted in fig. 6.4. IC-1 metric seems to pinpoint the critical failure point of the gear whereas IC-2 holds the information which is relevant to the gradual, hour-per-hour degradation of the asset. It seems that IC technique is able to distinguish among these two factors.

The question now is which of these transforms is preferable as a preprocessing prior to a life prediction model. To answer this we can resort to the similarity of each component to the normalized operational life of the particular component. The correlation coefficient metric is going to measure this similarity. The closest to “1” this metric is the better this metric can explain the demise of the operational time. Table 6.4 summarises the CCC performance of each derived metric (from figs. 6.3 and 6.4).

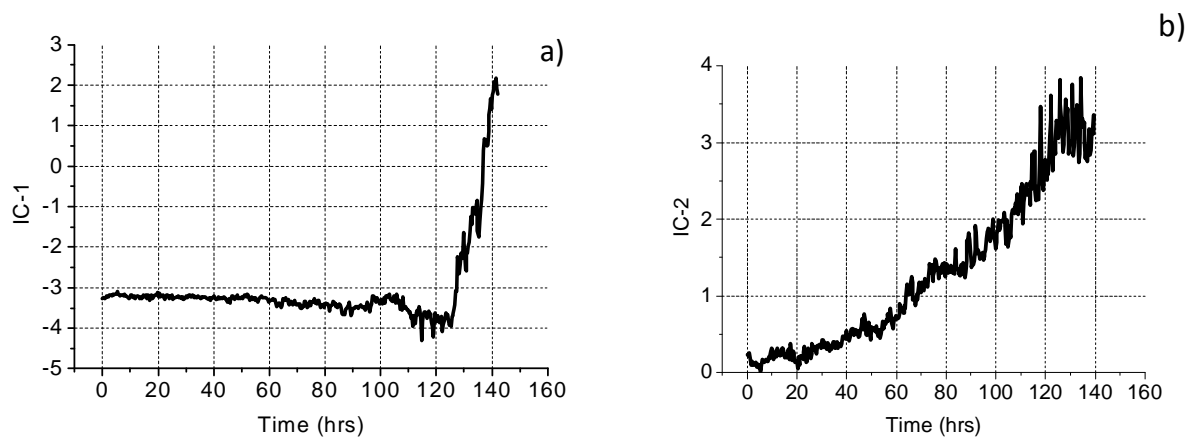


FIG. 6.4. Fused health monitoring metrics via ICA.

Extracted health metric	Performance
PC-1	0.73
PC-2	0.87
IC-1	0.34
IC-2	0.94

Table 6.4. Prognostic performance for fused health metrics.

Table 6.4 shows that IC-2 metric performs better in explaining the gradual reduction of the asset’s operational life. This is evidence that ICA is a better feature reduction technique compared to PCA with respect to prognosis.

To derive some quantitative life predictions based on IC-2 health metric we choose a prognostic method from Chapter 5. The preferable method is the bootstrap extrapolation as presented in Section 5.4. That’s because i) IC-2 seems to be non-linear and an exponential trend extrapolation seems a good choice and ii) due to the limited failure data base we have to assume a total absence of failure prior and thus inability to apply Cox aging model or ϵ -Support Vectors (as discussed in Sections 5.3 and 5.4).

In total 35 exponential functions are going to be fitted at each observation instant with short time interval (sample interval) of length $n_{int}=3$ subsequent samples and an interpolation interval of $n_{obs}=35$ samples. This way only the local dynamics of the degradation are kept for RUL estimation. The failure threshold is defined as $thres=3.5$. The α, β, γ parameters of the exponent are derived with the Levenberg-Marquadt algorithm. The results are depicted in fig. 6.5 in boxplot form.

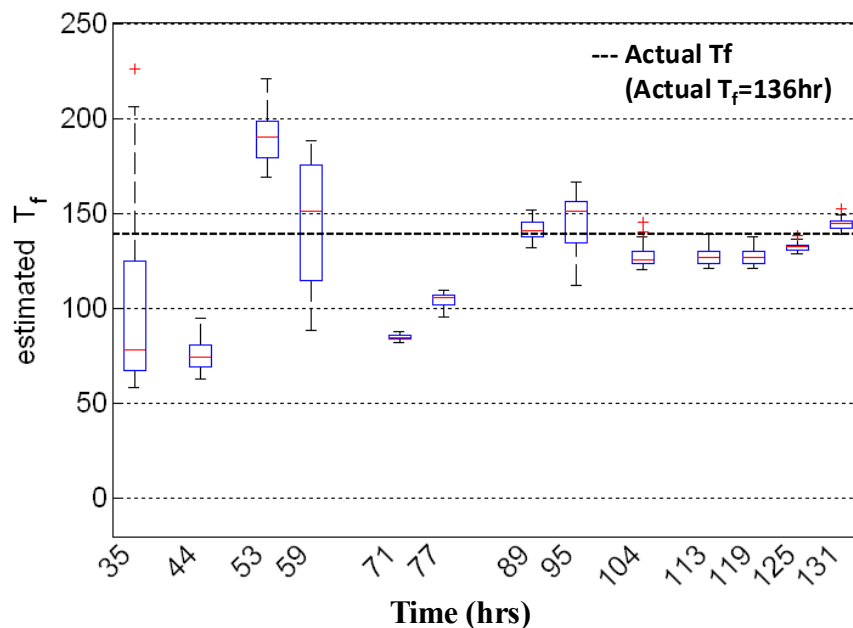


FIG. 6.5. Quantitative estimation of “Time of failure” along with appropriate error bounds via a boxplot.

ICA coupled with a simple extrapolation method can provide very good prediction results. As shown in fig. 6.5 this combination may provide adequate predictions as early as 43% of operational time life (at 53 hrs). Moreover meaningful upper and lower error bounds

are derived from the prediction algorithm although quite narrow close to the asset's end of life.

We can compare [fig 6.5](#) with the one derived from the gear failure mono-parametric analysis, take for instance the results from [fig. 5.8](#). Mono-parametric predictions produce wide error bounds and are relatively accurate close to the end of the asset's operational life. On the other hand, the proposed poly-parametric ICA fusion technique produces higher accuracy and precision even earlier than half way the total operational life. This is evidence that data fusion with ICA is a preferable technique for prediction purposes. This gives rise to an integral gear failure prognosis scheme which is summarized in [fig. 6.6](#). This scheme is proposed for the first time and its basic strength is that it can provide very good predictions and error bounds in cases where a very limited (possibly only one in order to derive the PCA, ICA fusion matrices) failure data base is available.

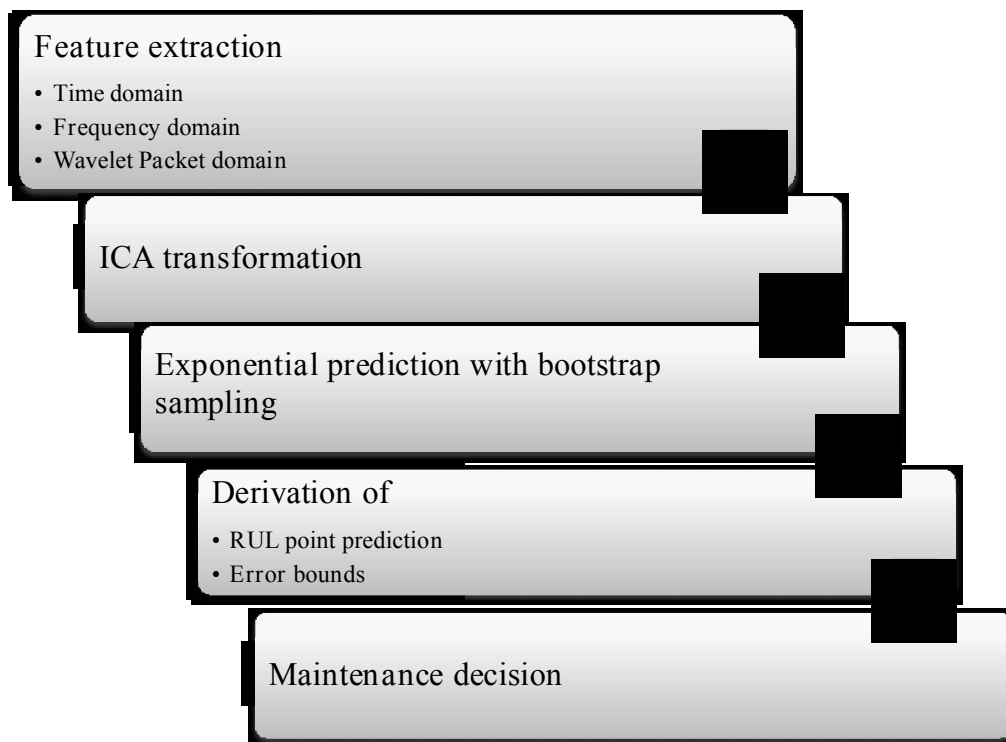


FIG 6.6. The proposed data fusion with probabilistic extrapolation prognosis scheme.

6.3 Conclusion

In this chapter the issue of feature extraction and feature fusion for gear failure prognosis was covered.

- ❖ A large number of features were extracted and the most predictive were kept through a process presented in [Section 6.2](#). Among them some features were highlighted for

the first time for their prognostic potential and particularly some higher order moments of the vibration spectrum as shown in [fig. 6.1](#).

- ❖ Both PCA and ICA were tested for feature fusion. To the author's knowledge this is the first time these techniques are assessed for gear failure prognosis. From the study there was clear evidence that ICA is superior in terms of gear RUL estimation than PCA as presented in [Section 6.2](#).
- ❖ The derived health metric, IC-2, is fed into the probabilistic extrapolation model of [Section 5.4](#) yielding an improved failure prediction algorithm. It is worth mentioning that the particular health metric takes into account multiple sensors, namely AE, vibration x, y and z axis, an element that is rarely if ever found in literature for gear failure prognosis. The whole prediction algorithm is summarized in [fig. 6.6](#) and it is proposed for the first time for gear failure to the author's best knowledge.

CHAPTER 7

Estimating Remaining Useful Life of Bearings

7.1 Introduction

The root cause of a number of failures and gearbox in particular can be traced back to the drive train bearing system (Alewine (2010)). Therefore it is natural to cover the issue of bearing failure prediction in the thesis framework. In the current chapter bearing failure prognosis is going to be assessed individually.

In contrast to gear failure data, bearing failure repositories are easier to come by. Section 7.2 is devoted in presenting and finally choosing a particular bearing failure repository for prognosis purposes. Having established an adequate bearing failure data pool we continue to mono- and multi-parametric bearing RUL estimation. Section 7.3 proposes a mono-parametric bearing failure prediction scheme based on the ϵ -SVR method already presented in Chapter 5. Moreover an enhancement is proposed based on wavelet denoising technique that results in better bearing prognosis and diagnosis. Section 7.4 moves even further by proposing a feature extraction scheme coupled with ϵ -SVR prediction model. The feature extraction scheme is adopted from Section 6.2. Section 7.5 concludes this chapter by repeating and highlighting the most important contributions. Part of the results from this chapter have been published in Roulias et al. (2013) and Loutas et al. (2013).

7.2 Bearing failure data bases

A number of open access gear failure repositories can be found in the web. Three of these repositories were looked into by the author of this thesis to choose the most appropriate for studying bearing RUL. These repositories are reported in [table 7.1](#).

Data Base	Research group	Identity	Number of tested components	Goal
Case Western	Case Western Reserve University	Vibration recordings from various seeded fault conditions with particular torque and axis speed configurations	35	Early diagnosis of particular failure modes
PRONOSTIA	FEMTO-ST	Vibration recordings from natural degradation of healthy bearings subject to various torque and speed conditions	17	Bearing RUL estimation under constant torque and rotating speed conditions
IMS Bearing Data	NSF I/UCR Center for Intelligent Maintenance Systems	Vibration recordings from natural degradation of healthy bearings subject to a particular torque and speed condition	12	Bearing RUL estimation under constant torque and rotating speed conditions
Purdue bearing data base¹²	University of Purdue, Department of industrial engineering	Vibration recordings from natural degradation of healthy bearings subject to a particular torque and speed condition	34	Bearing RUL estimation under constant torque and rotating speed conditions

Table 7.1. Bearing failure data bases.

¹² Not open source

The last repository was not considered in this thesis because it was not open source. Case Western was not considered in the thesis framework because it did not comprise natural lifelong degradation but merely seeded or simulated fault experiments. These damages were induced by scratching machining with an electrical discharge. IMS data base was interesting since it comprised what a number of naturally damaged bearing histories. However this data base had two basic flaws:

i) It was very in-homogenous, possibly due to the appearance of various failure modes. Moreover, in one test, comprising 4 tested bearings, the application of the feature extraction process of [Section 6.2](#) yielded no information whatsoever. More precisely no critical point close to failure and no correlation to the decrease of operational life was noted in any extracted timeseries.

ii) Each time 4 bearings were tested side by side in the particular IMS test rig. In each experiment, one or two bearings were proclaimed damaged out of the four simultaneously tested bearings. However, recordings from all four bearings showed a simultaneous increase in RMS (and various other statistical indices) as shown in [fig. 7.1](#). In other words, the particular test rig correlated the recordings of all four tested bearings leading to false conclusions.

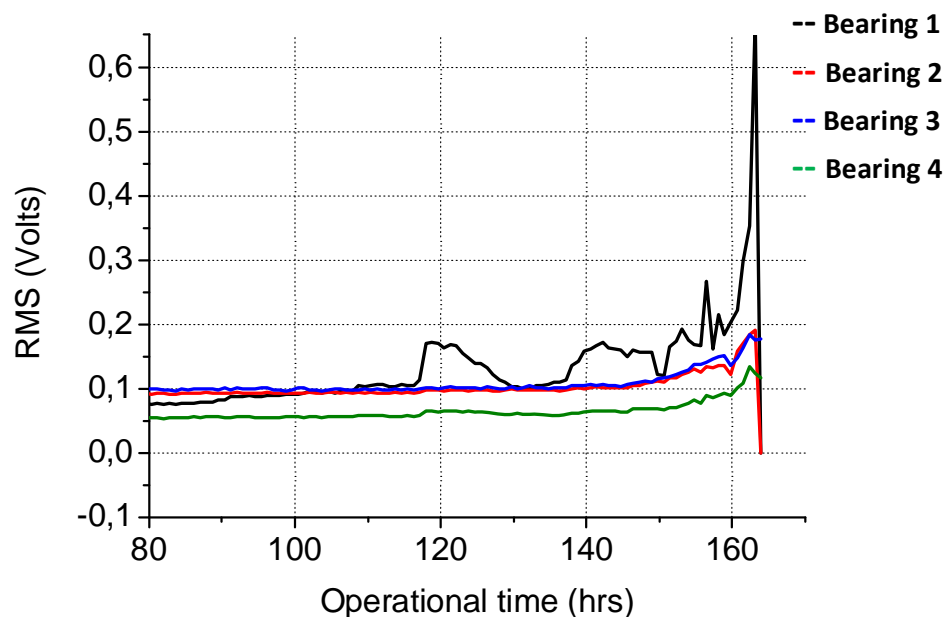


FIG. 7.1. RMS of four bearings being tested simultaneously under accelerated fatigue conditions. In this test bearing 1 (top-most curve) was found severely damaged in the final visual inspection. The RMS recordings of bearings 2, 3 and 4 are much correlated indicating that the corresponding recordings were severely affected from bearing 1 failure vibrations. This experiment provided only 1 valid failure history.

PRONOSTIA seemed to be the best available choice. The particular data come from the publicly available PRONOSTIA platform (Nectoux et al. (2012)) and comprise a number of run-until-failure experiments in rolling element bearings in various loading cases. Two vibration sensors were mounted on the bearing case for the degradation monitoring task, one recording the vertical and the other recording the horizontal axis acceleration signal. The sampling rate was 25.6 kHz, the length of every sampled acceleration waveform was 2560 samples whereas recordings took place periodically every 10 seconds. All tests were stopped when the vibration root mean square (RMS) acceleration crossed the 20g threshold. Constant loading and shaft speed conditions were kept for all experiments. The data were organized so that they cover three cases of load-speed conditions. These three set-ups are summarized in table 7.2. Photographic evidence was also provided as a visual assessment of the final rolling element bearing condition. It is evident in Fig. 7.2 that the respective 20g threshold was in general an adequate failure criterion.

Operating conditions	Shaft speed (rpm)	Load (N)	Number of components tested
Case 1	1800	4000	7
Case 2	1650	4200	7
Case 3	1500	5000	3

Table 7.2. Operating conditions



Fig. 7.2: Photos of degraded bearings from Nectoux et al. (2012)

Based on the extracted vibration RMS histories, several points can be made regarding the latter experiments. To this end, RMS of the horizontal axis accelerometer is plotted against operational time. In fig. 7.3 two discrete operational time regions are clearly evident. In the first operational region the vibration RMS level is low and seemingly constant. This

region spans from 0 to 368 minutes in [fig. 7.3a](#) and from 0 to 138 minutes in [fig. 7.3c](#). In this interval, the degradation of the bearing starts with minute cracks in its microstructure or the surface due to fatigue. These cracks progressively propagate to the surface of the raceway creating small pits and dislodging of small pieces of material. However, very few evidence for this gradual degradation can be found in the presented vibration RMS histories. At some point a knee in the time plot of RMS indicates the appearance of a critical fault. [Figure 4b](#) and [4d](#) focus at the time of this event. The dynamics of the degradation at this operating region change dramatically, based on the RMS histories. In the current framework we will choose to keep only the first operating region. We can distinguish among the two operational regions by placing a threshold in RMS. This threshold is going to define a conservative “end of life” point. More precisely, the end of useful life is defined as the point that the dynamics of the operating bearing change and it switches to another operational mode, as shown in [fig. 7.3b](#) and [7.3d](#).

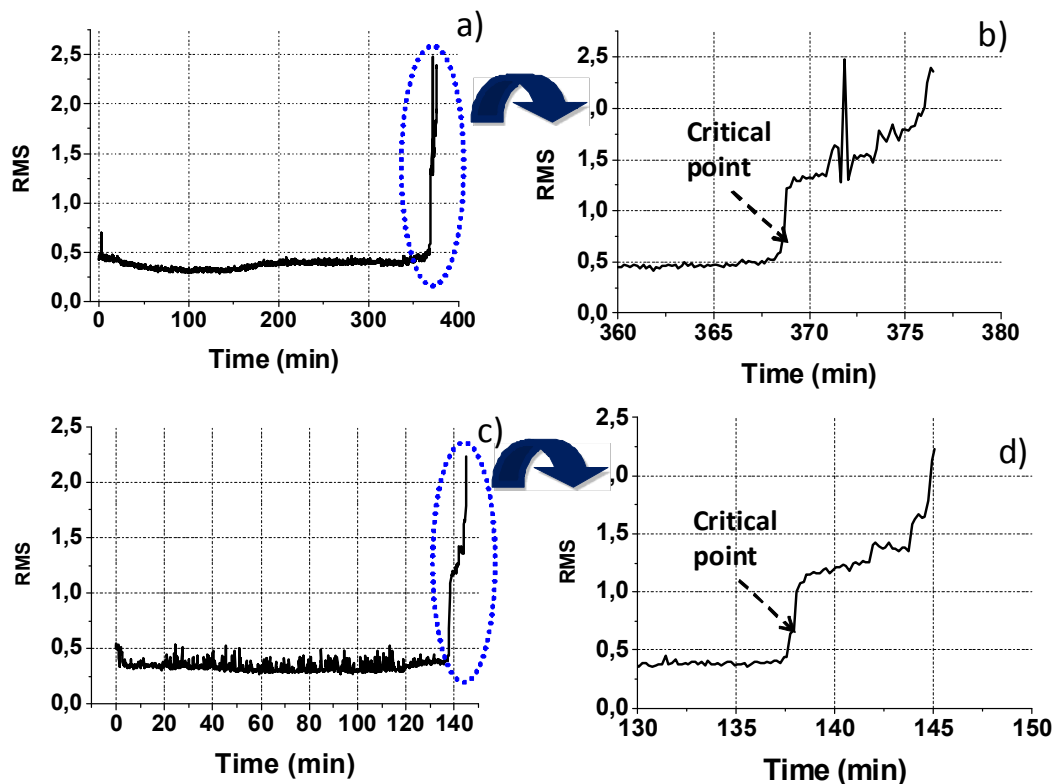


FIG. 7.3 Typical RMS timeseries for two lifelong bearing experiments. a) and c) RMS histories from case 1 and 2 bearings respectively (see [table 7.2](#)), b) and d) zoom in the change-point due to critical fault occurrence.

Moreover, as noted by the FEMTO group ([Nectoux et al.\(2011\)](#))

- The theoretical models based on frequency signatures to detect bearings' faults (such as the inner and outer races and the cage faults) do not work. Indeed, frequency signatures are difficult to obtain due to the fact that the degradation may concern all the components of the test bearing at a same time.

- Existing reliability laws for bearings' life duration, such as the L10, do not give same results than those obtained by the experiments (theoretical estimated life durations are different from those given by the experiments).

The latter observations make the issue of prognosis particularly challenging. Moreover, data-driven models seem to be the only choice towards robust and reliable RUL estimation. In the following sections feature extraction, feature fusion and prediction model adjustment are going to take place on the particular bearing data base. The particular methods are adopted from [Chapter 5](#) and [6](#).

7.3 Monoparametric bearing failure prediction

We will begin our study with single parameter monitoring. We will limit our study to waveform RMS for prognosis as we did in [Chapter 4](#) and [5](#). By plotting the RMS time lines for all 14 PRONOSTIA bearing histories several observations were made:

i) In several cases RMS gave no significant trend except for the very end of the bearing operational life, as shown in [fig. 7.3a](#), [7.3c](#).

ii) There were very few cases that RMS gave a monotonically increasing trend throughout the bearing life span. Also there were cases where RMS yielded a monotonically decreasing trend.

iii) There were cases that RMS yielded a completely irregular pattern as shown in [fig. 7.4a](#), [b](#). The reason for these irregular trends is possibly the appearance of small pitting on the bearing structure which becomes worn out with the bearing revolution. In fact the bearings seem to heal itself¹³ in particular cases.

iv) In a few cases RMS gave no information regarding the bearing health.

¹³ This was pointed out by the FEMTO-SH group in a personal communication, PHM Conference Milan, September 2013.

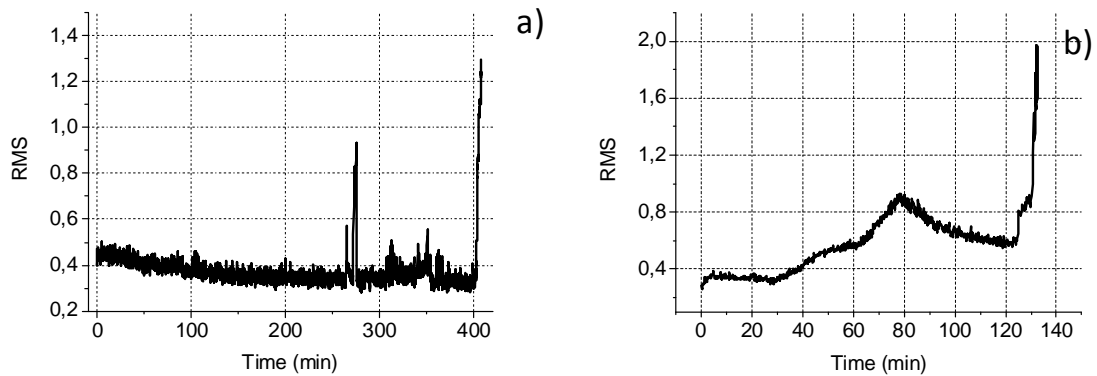


FIG. 7.4. Two “irregular” RMS histories derived from Case 1, (fig. 7.4a) and Case 2 (fig. 7.4b) with reference to table 7.2. On the left, a sudden increase in RMS at ~270mins was not proven detrimental to the bearing. In fact the particular fault seems to wear out with the bearing revolution. On the right, few increasing and decreasing trends make any effort of diagnosis or prognosis practically impossible.

This begs the question if there is some preprocessing that could smooth out or ease many of these irregularities. We tested the Wavelet Denoising (WD) technique and its effect on the RMS of the extracted recordings. The findings were quite interesting as will be depicted in the following paragraph.

Wavelet Denoising Theory

Suppose we need to extract an unknown signal f from a noisy timeseries $s(t_i)$

$$s(t_i) = f(t_i) + e(t_i) \quad (7.1)$$

where t_i is the discretised time and $e(t_i)$ is an additive noise component. Concerning physical sources the noise component is most often assumed a zero mean gaussian process with variance σ^2 independent and identically distributed (i.i.d.) samples. Kernel estimators or spline estimators can be applied in order to smooth out the noisy portion of the mixture. However they cannot handle well local signal structures. Fourier based signal processing can also dispose of a portion of the noisy component. This is done by applying linear filters at particular frequency bands. This method however is not efficient when the Fourier spectra of the signal and that of noise overlap.

Wavelet transform is the approximation of a function or signal, in mean square terms, with another function called a mother wavelet and DWT is its discrete form. In practical signal processing, the discrete version of wavelet transform is often employed by discretizing the wavelet dilation and translation parameters (see eq. 6.1). The procedure becomes even more efficient if dyadic values of these parameters are used. Practically relevant to dyadic

DWT is multi-resolution analysis (MRA) described by Mallat (1999). MRA is a time efficient digital signal processing method which is connected to the theory of filter banks. It yields a tree structure for the signal decomposition that splits the signal into segments with different frequency resolutions. Fig. 7.5 depicts a 2 level DWT tree. Each “branch” comprises of two mirror filters, a low pass (LP) and a high pass (HP) filter.

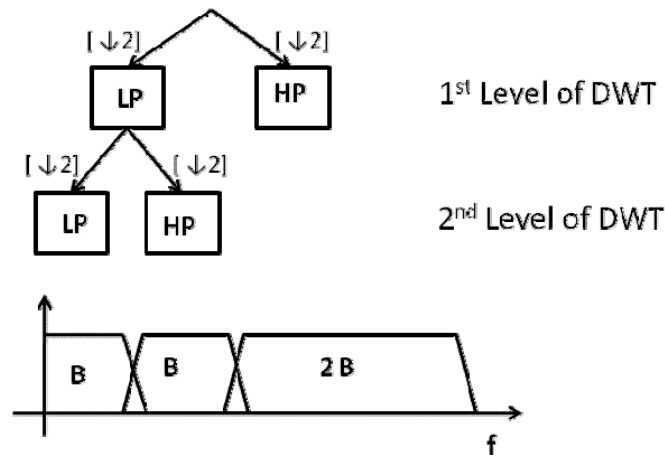


FIG 7.5: A graphic representation of DWT

The technique proposed by Donoho is based on the idea of thresholding the wavelet coefficients at each wavelet decomposition level. Those wavelet coefficients that have small absolute value are considered noisy and are thresholded. On the other hand, coefficients having large absolute value are considered as important information. The act of removing the small absolute value coefficients and reconstructing the signal should produce a signal with less noise which would yield a good approximation of the original signal f (eq. 7.1). The wavelet denoising (or wavelet shrinkage) technique is in general as follows:

Wavelet denoising Algorithm

- 1) Split the signal into segments with different frequency resolution via DWT

$$\tilde{\Theta} = \mathbf{W} \cdot \mathbf{Y} \tag{7.2}$$

Where \mathbf{W} is the DWT matrix, $\mathbf{Y} = \{y_1, y_2 \dots y_n\}$ is the original signal and $\tilde{\Theta} = \tilde{\theta}_{i,j}$ is the wavelet coefficient i of the j decomposition level.

- 2) Apply an appropriate threshold to the high frequency components of $\tilde{\Theta}$ known in literature as the details

$$\hat{\theta}_{i,j} = \beta_{i,j} \cdot \tilde{\theta}_{i,j} \quad (7.3)$$

Where $\beta_{i,j}$ is a shrinkage factor for the coefficient i at level j .

3) Reconstruct the denoised signal using the inverse DWT formula

$$\hat{\mathbf{Y}} = \mathbf{W}^{-1} \cdot \hat{\mathbf{\Theta}} \quad (7.4)$$

The expert should choose a wavelet basis as well as the depth of levels of decomposition in order to form the decomposition procedure of eq. (7.2) and a threshold method as described in eq. (7.4).

Thresholding with neighbouring blocks

Cai and Silverman (1998) have proposed an interesting threshold method. In that method the shrinkage factor at eq. (7.4) is not applied at each coefficient at each level. On the contrary it is applied in a group of successive coefficients and the threshold is calculated within adjacent blocks of coefficients. The method resembles a sliding window that calculates a threshold value within a neighbourhood and applies this threshold in the central block of this neighbourhood of coefficients. More specifically:

Neighcoef denoising algorithm

- 1) Perform DWT of eq. (7.3) with a certain wavelet basis and number of levels of decomposition in order to acquire the wavelet coefficients
- 2) At each level of decomposition, group the wavelet coefficients into disjoint blocks $\mathbf{b}_{i,j}$ of length $L_0 = \lceil (\log n) / 2 \rceil$. Each block is extended by an amount of $L_1 = \max(1, \lceil L_0 / 2 \rceil)$ points in each direction to form overlapping blocks $\mathbf{B}_{i,j}$ of length $L = L_0 + 2L_1$.
- 3) Within each block indexed i at level j , $\mathbf{b}_{i,j}$, estimate the coefficients via the shrinkage rule of eq. (7.4). The shrinkage factor $\beta_{i,j}$ is chosen with reference to the coefficients of the larger block $\mathbf{B}_{i,j}$

$$\beta_{i,j} = \max(0, (1 - \lambda L \sigma^2 / S^2)) \quad (7.5)$$

Where $k=1,2,\dots, L_1$ is the index of each data point within the bigger block $\mathbf{B}_{i,j}$,

$$S^2 = \sum_{(j,k) \in \mathbf{B}_{i,j}} \tilde{\theta}_{j,k}^2 \text{ and } \sigma^2 \text{ is the variance of the extended block and } \lambda=4.5053.$$

The latter method is shown to outperform other older threshold criteria such as Rigsure, NisuShrink and others (Cai and Silverman 2000). Therefore it will be the preferred WD method.

Results

In the current WD framework, 'db10' wavelet basis is chosen as a wavelet basis. Experimenting with a variety of wavelet bases no significant difference. Moreover, concerning eq. (7.6), the median absolute deviation (MAD) of the extended block is used as a more robust dispersion estimator instead of the sample standard deviation. The NeighBlock thresholding scheme was applied to all Detail wavelet coefficients. The results on the extracted feature (rms) are depicted in fig.7.6.

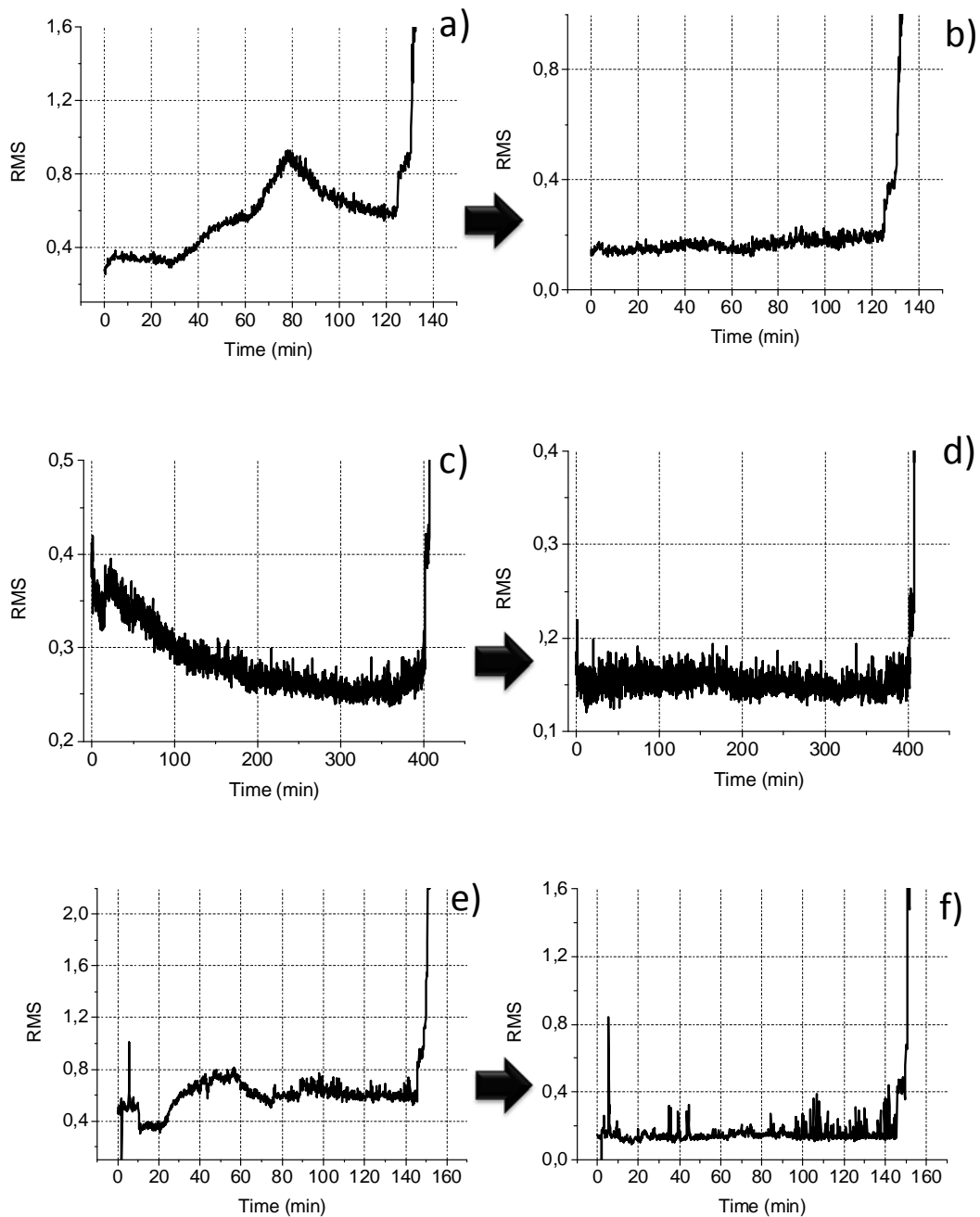


FIG. 7.6. On the left, the original RMS from three bearing failure histories and on the right the corresponding RMS after the denoising of vibration waveforms.

The effect of WD preprocessing is evident in [fig. 7.6](#). In all three cases ([fig. 7.6a, c, e](#)) various irregular trends appear throughout the bearing operational life. These trends obscure the bearing failure diagnosis. However, after the preprocessing these trends are effectively canceled out. Few outliers in [fig. 7.6f](#) can be effectively handled with to an outlier detection and cancelation scheme presented in [Section 4.6](#).

On the other hand the final critical failure is not masked by WD preprocessing. It seems that WD disposes of signal components that do not produce a critical failure and enhance the the final critical error jump in RMS measurements.

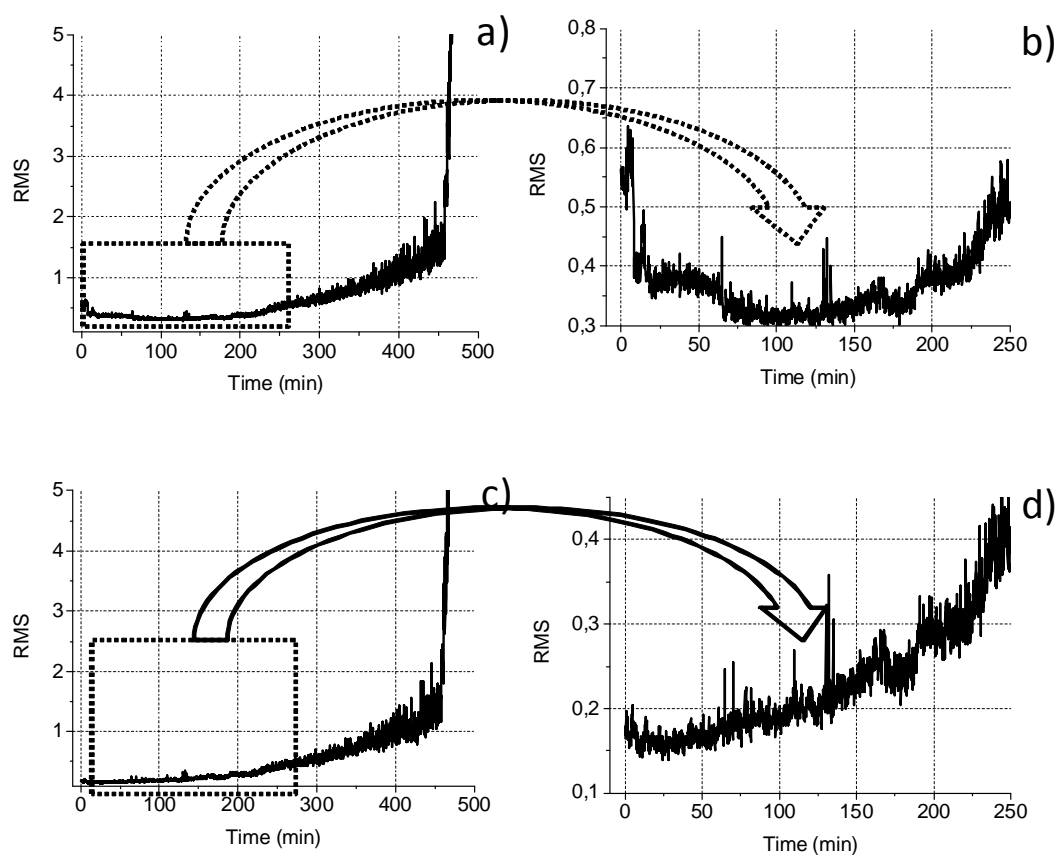


FIG. 7.7. Upper figures (a, b) correspond to original RMS from a bearing failure history, b) being a zoom in axis of a). Lower figures (c, d) correspond to WD processed versions of the latter.

Another group of bearing failure histories yielded an almost exponential functional form. This exponential functional form has been pinpointed in literature ([Gabraeel et al. \(2005\)](#), [Harris \(2001\)](#) and others). This time the final failure is completely distinguishable but there seems to be some bath-tab irregularity that spans almost half of the total operational life,

as depicted in [fig. 7.7b](#). WD results in smoothing out this extensive run in effect spanning the first 100 minutes (~20% of the bearing operational life). Moreover WD clears out the irrelevant component without masking the trend or the final critical failure point, as depicted in [figs. 7.7c](#) and [7.7d](#).

A final issue regarding WD remains to be covered and that is the depth (number of levels) of DWT chosen for decomposition. Regarding the first group of bearing histories (as depicted in [fig. 7.6](#)), experimenting with various levels of DWT did not yield significant differences. Regarding the second group (degradation mode similar to [fig. 7.7](#)) small differences were noted. In those cases WD-RMS index was compared to the normalized actual remaining life for various WD decomposition levels yielding the following results.

Level of decomposition	1	2	3	4	5
Correlation coefficient	0.6813	0.6784	0.7123	0.7506	0.7402

Table 7.3. DWT level and corresponding performance.

Based on [table 7.3](#), 4 levels of decomposition were opted for WD algorithm. The preprocessing flow chart is visualized in [fig. 7.8](#).

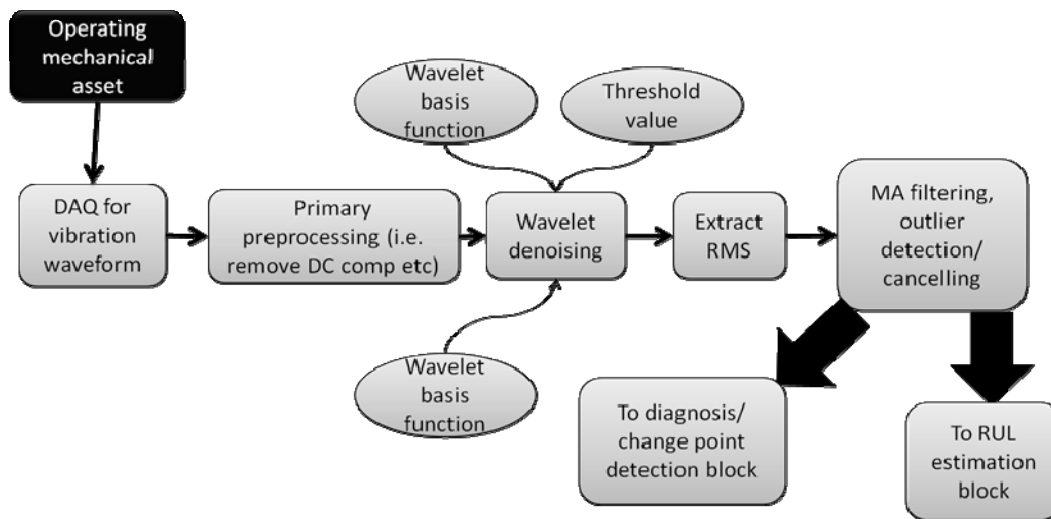


FIG. 7.8. Flow chart of the proposed preprocessing technique.

7.3.1 Application of ϵ -SV regression

The terminal effect of WD preprocessing can be highlighted by applying an actual prediction algorithm. Take for instance the ϵ -SVR prediction algorithm as presented in [Section 5.3](#) with the Radial Basis kernel. The choice of RBF is justified in [Loutas et al. \(2013\)](#).

After a visual inspection we ended up with 3 histories whose RMS follow an exponential like form, similar to [fig. 7.7a](#). Moreover we define a conservative failure threshold at ~ 2.5 Volts RMS.

Due to the limited number of failure histories we are inclined to keep two histories to train the prediction model one for testing. Prior to feeding the training histories in the SVR training algorithm they are scaled in the interval $[-1, 1]$. Normalised actual RUL is used as the corresponding target vector for the regression process. The testing sequence is therefore scaled in $[0, 1]$. We can change the scale of the output by multiplication with a Mean Time To Failure (MTTF). We can deduce MTTF from the remaining failure histories of “Case 1”.

The values of C and ϵ , and γ parameters remain to be defined. C and ϵ are search for via grid search, according to [Section 5.3](#). C , ϵ are assigned the following values:

$$\epsilon \in [0.0005, 0.001, 0.005, 0.01, 0.05, 0.1, 0.5]$$

$$C \in [0.01, 0.05, 0.1, 0.5, 1, 3, 5, 10, 20, 30, 50, 100, 200]$$

The corresponding performance is measured on the training set due to the limited number of experiments at hand (2 for training and 1 for testing).

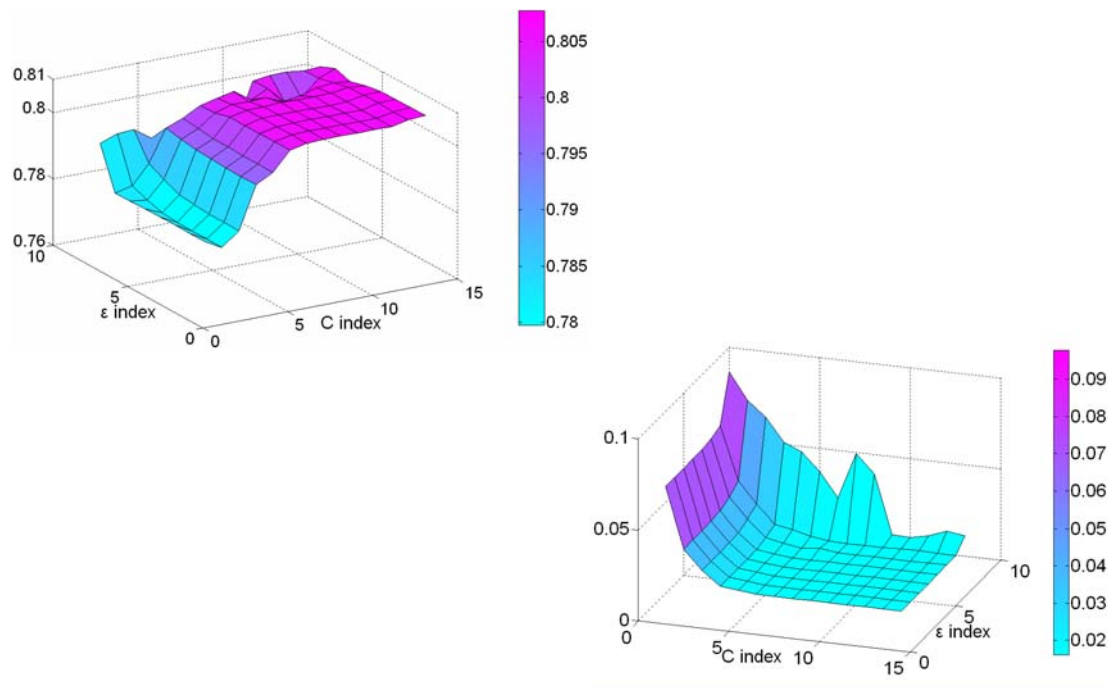


FIG. 7.9. MSE (right) and CCC (left) for various values of ϵ and C model parameters. The exact values of the respective indexes are shown above.

Experimentation with the RBF γ parameter did not yield much different results in the interval $\gamma \in [0.1, 1]$ thus a value of $\gamma = 0.5$ was chosen which gave arbitrarily good performance. The latter training-testing process took place for two cases,

- a) the model input being RMS and
- b) the model input being the RMS after the WD process. The prediction results are shown in [fig. 7.10](#). The output of the testing set was multiplied with the mean estimated life of the Case 1 and 2 failed components from [table 7.2](#).

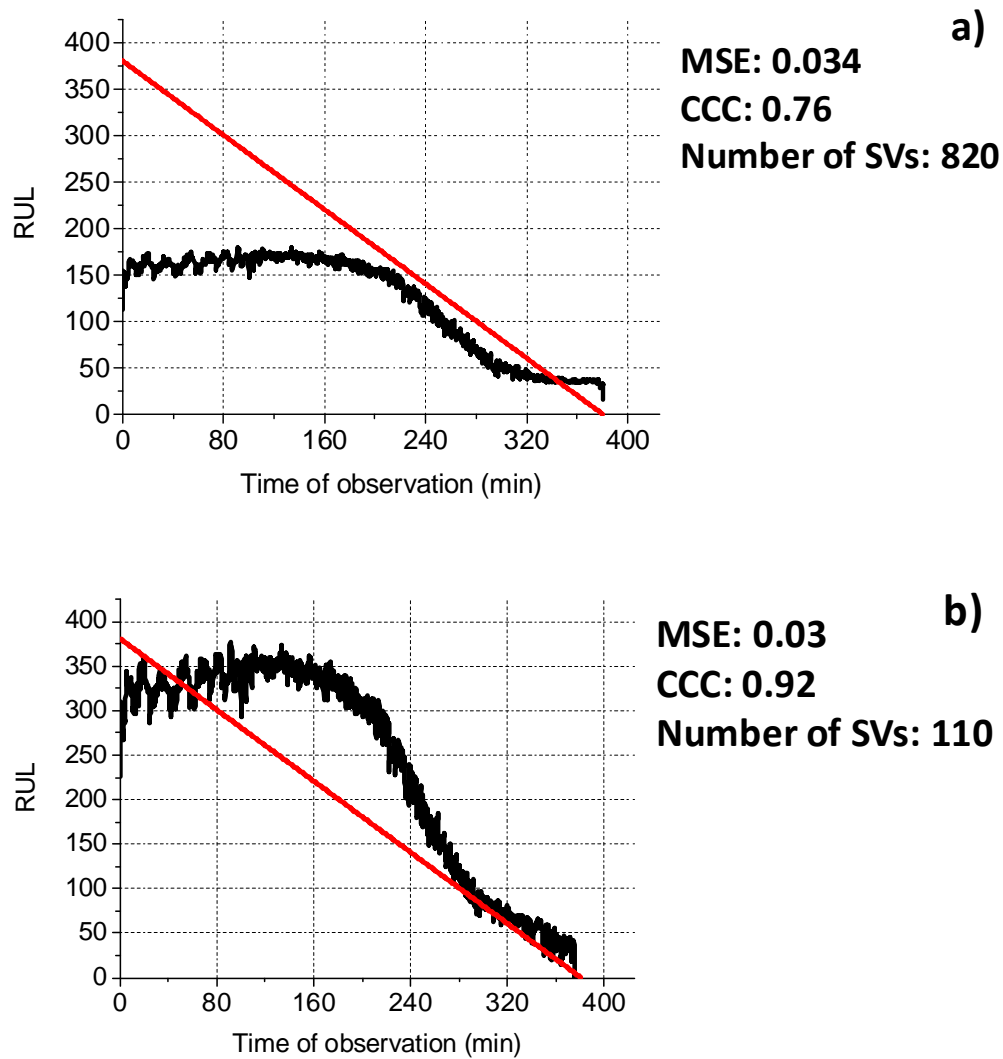


FIG. 7.10. Dash dot straight line is the actual RUL of the testing component and it is plotted for comparison. MSE and CCC are computed on the original (scaled) output and the normalised actual RUL. Figure a) corresponds to the original waveforms and b) to WD-preprocessed waveforms.

It is evident that in the case of RMS –WD the accuracy of predictions is higher. At the same time, the number of support vectors in the case depicted in [fig. 7.10a](#) is much higher than the one of [fig. 7.10b](#) (the total number of training points were ~3000). Although the original ϵ -SVR by construction does not provide sparse data representation, keeping all other factors constant, WD preprocessing results in better model parsimony.

From the latter study it is clear that WD results in a more predictive RMS index. It should be mentioned though that:

i) the study was based on a very limited failure data set. Therefore more bearing failure experiments are needed to further validate the effect of this method.

ii) the particular exponential degradation was observed only in a limited number of the total failure histories and is probably due to the particular bearing failure mode. Therefore an addition of a diagnosis/failure mode classification block seems to be necessary.

Could this intermediate diagnostic block be ruled out with a clever multiparametric analysis? Could a feature extraction and fusion process improve the latter mono-parametric WD-RMS+ ϵ -SVR prediction algorithm making it more generic? The following section is going to address these issues.

7.4 Poly-parametric bearing failure prediction

7.4.1 Feature extraction

A number of features are extracted from the vibration waveform recordings. The feature extraction block for prognosis has already been presented in [Section 6.2](#). The statistical features that were extracted can be found in Appendix A. Wavelet Packet (WP) domain features are also considered as described in [Section 6.2](#). The number of WP decomposition levels was chosen as 3, a choice that agrees with FEMTO-SH group ([Benkedjouh et al. \(2013\)](#)). The terminal WP nodes correspond to frequency bands that are depicted in [table 7.4](#).

WP terminal node	Vibration frequency band (KHz)
1	[0, 1.6]
2	[1.6, 3.2]
3	[3.2, 4.8]
4	[4.8, 6.4]
5	[6.4, 8]
6	[8, 9.6]
7	[9.6, 11.2]
8	[11.2, 12.8]

Table 7.4. WP bandwidth.

PRONOSTIA provides two vibration sensor recordings, a x - and y - axis. It was observed that both sensors held information relative to the bearing degradation and both were kept for RUL prognosis. The whole process yielded $26 \times 2 = 52$ features. Each timeseries feature was compared with the actual, normalised, life line. Those features who gave a CCC score better than 0.67 were kept for further experimentation. The whole process took place for “low” and “medium” load cases of [table 7.2](#). The “high” load case was not considered since the data base comprised of only 3 bearings. Finally, few failure histories that gave no important or very irregular information (see for instance [fig. 7.3a](#)) in any extracted feature were opted out.

Wiener entropy or spectral flatness

Wiener entropy or spectral flatness is a feature associated to the power spectrum of a time-series. It is a measure of the information content of the noisy time-series and its expression is given by the ratio of the geometrical mean of the power spectrum coefficients to its arithmetic mean:

$$WE = \frac{\exp\left(\frac{\sum_{k=0}^{N-1} \log(S(\omega_k))}{N}\right)}{\frac{\sum_{k=0}^{N-1} S(\omega_k)}{N}} \quad (7.6)$$

Where $S(\omega_k)$ is the power spectrum calculated at each frequency ω_k . Spectral flatness inherently takes values from in the range $[0,1]$. Flatness equal to one implies a totally uniform spectrum therefore white noise whereas zero flatness implies a monotone signal i.e. a pure sinusoid.

The vibration signal obtained from an accelerometer mounted on a bearing case is supposed to increase its flatness while degradation is progressing. In the beginning of the operation, acceleration signal should contain a modulated band that stems from the normal operation of the motor. While bearing degradation is progressing, the acceleration signal becomes “infested” with higher harmonics and resonances. Fig. 7.11 shows various spectral flatness evolutions versus test times for three bearing degradation experiments and verifies the increasing character due to degradation. Interestingly enough their evolution is monotonic with time as expected and additionally they possess relatively small volatility.

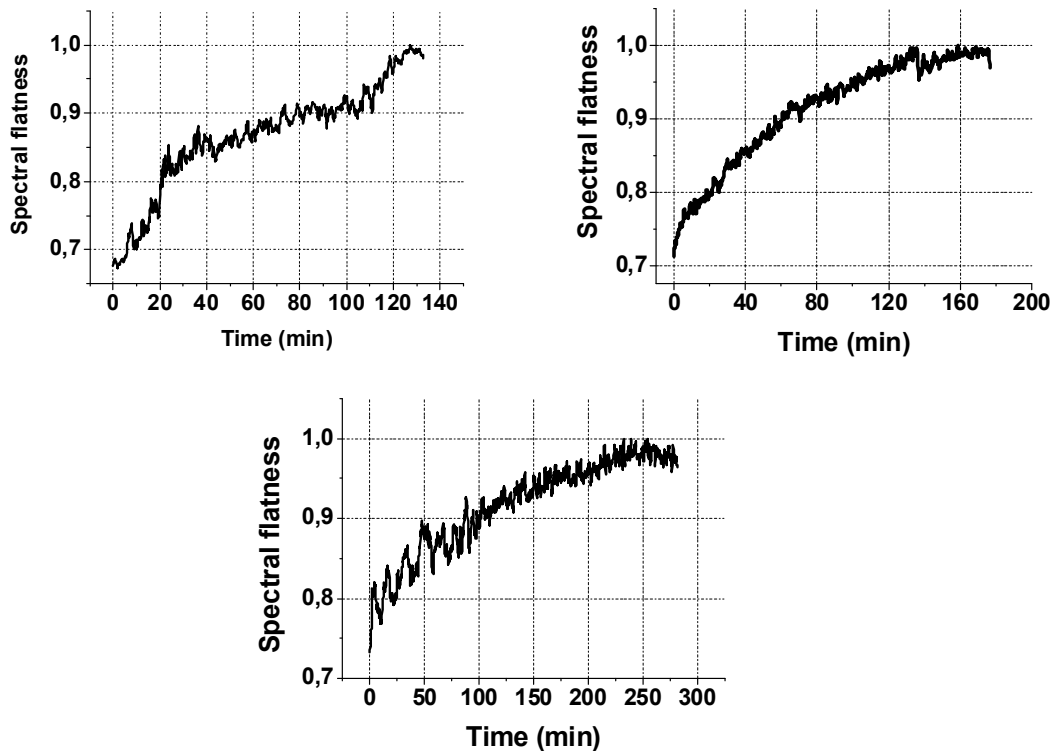


FIG. 7.11. Spectral flatness as a function of operational time for three different lifelong rolling element bearing tests.

After the feature selection process the selected prognostic features were the 8 WP features, as described in table 7.4 plus the Weiner entropy for both sensors yielding a total of 18 prediction features.

7.4.2 Prediction results

The hyperparameters tuning process took place separately for “case 1” and “case 2” training data sets via a grid search. The tuning results however were similar for both cases implying a generic connection between the data. Consequently the choice of model parameters was kept the same for both operating cases. [Table 7.5](#) summarizes the results.

Kernel type	C	ϵ	γ	Number of Support Vectors for “case 1”	Number of Support Vectors for “case 2”
$K(\mathbf{x}_i, \mathbf{x}) = \exp\left(-\gamma \ \mathbf{x}_i - \mathbf{x}\ ^2\right)$	10	0.01	0.05	135	24

Table 7.5. Selected model parameters.

The number of support vectors differed in the two cases under consideration. However the percentage of support vectors to the total number of training points was similar for both “case 1” and “case 2” experiments. Almost 2.5% of the training data points are support vectors for each case, satisfying the parsimony condition. The final testing of the ϵ -SVR model is performed on the testing data set. The prediction output of the ϵ -SVR model for every data input is the percentage of RUL taking values in the range [0,1]. To obtain absolute times this value was multiplied with the mean operational time obtained from “case 1” failure histories. The ability of the SVR model to predict the RUL is evaluated by means of MSE and SSC. [Fig. 7.12a](#) and [7.12b](#) depict the final RUL prediction results for “case 1” and “case 2” operational set up respectively. In each graph, the actual RUL as a function of time is also plotted. Prediction error bounds are added to each prediction point separately according to the study in [Section 5.4](#). These error bounds correspond to a 95% confidence interval assuming gaussian prior on our data.

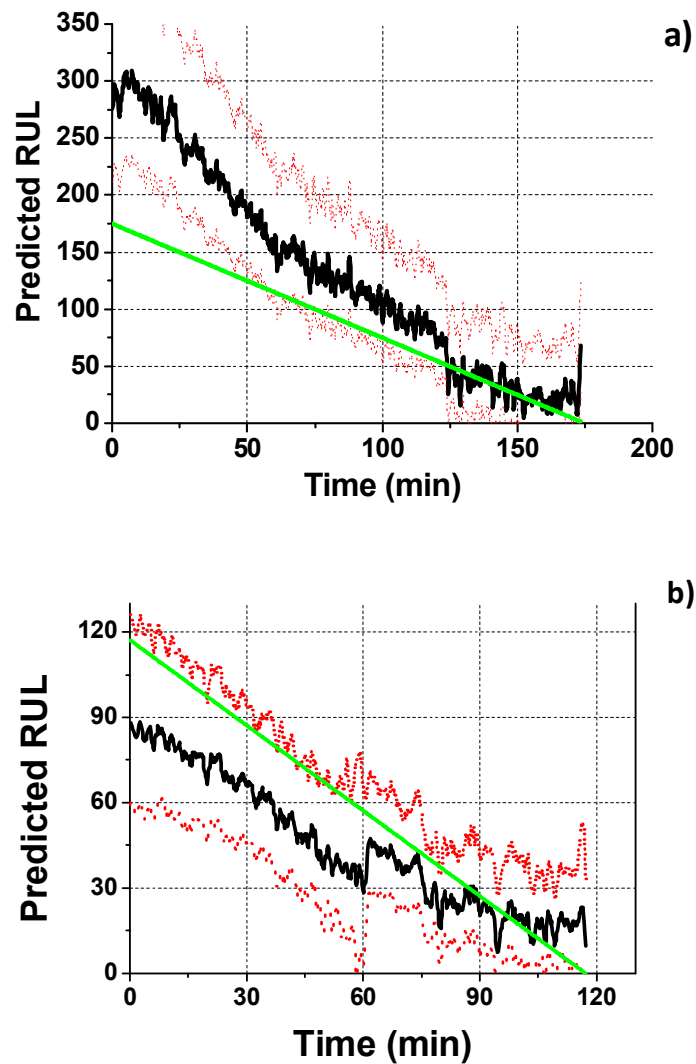


FIG. 7.12. Prediction results for “Case 1” data base (a) and “Case 2” database (b). 95% confidence intervals ($\pm 1.69\sigma$) are added on each prediction curve. The straight line corresponds to the actual remaining life for each component.

In general it was observed that the addition of a feature extraction step made the ϵ -SVR prediction scheme more general. More specifically, the prediction process gave meaningful results for several testing sets derived from “low” and “medium” configurations. Considering multiple statistical features provide more generalising prediction schemes compared to mono-parametric RMS prediction.

It is worth mentioning that PCA and ICA fusion techniques were attempted for the latter data base. They gave identical prediction results compared to the no-data-fusion case. For PRONOSTIA bearing data base, it seems that a feature reduction technique does not result in significantly better performance.

7.5 Conclusion

The current chapter covered RUL prognosis in bearings. The methodology that was applied was largely adopted from [Chapter 5](#) and [6](#) and the resulting predictions were in general adequate. PRONOSTIA database was utilized for experimenting.

❖ In the current chapter a new preprocessing technique was tested for enhancing the predictive ability of vibration recordings. This technique is based on the theory wavelet denoising. The improvement was proven using the ϵ -SVR prediction technique of [Chapter 5](#).

❖ Moreover, it was shown that the use of multiple statistical features could broaden the applicability of ϵ -SVR in cases where a sheer RMS gave no important degradation trend. The latter bearing RUL estimation scheme is proposed for the first time and is summarised in a flow chart in [fig. 7.13](#).

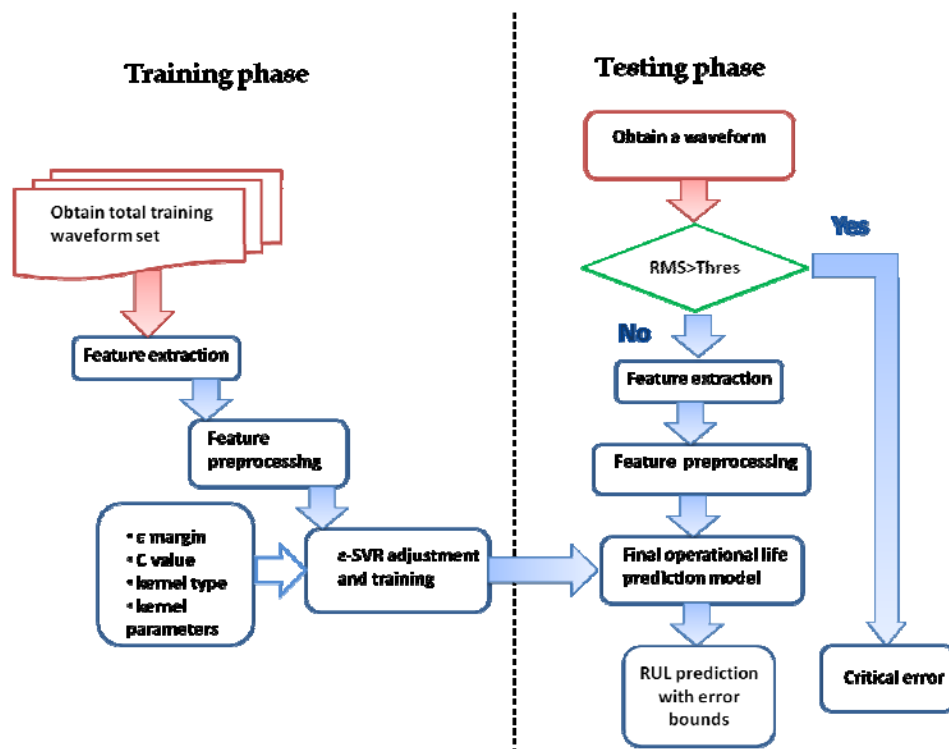


FIG. 7.13. Flow chart of the proposed multi-parametric bearing RUL estimation scheme which is described in [Section 7.4](#).

CHAPTER 8

Conclusions and Closure

The object of the current thesis was the development and study of prognostic techniques in rotating machinery and more particularly bearings and gears. To this end data-driven methodologies were studied in particular.

For gearbox in particular, there is a big void in literature, a void that the current thesis comes to fill in. The knowledge concerning bearing prognosis is wider but certainly not conclusive.

A gear failure data base was developed in the current framework to study gear pitting fatigue under constant conditions. Few research groups have been occupied and published multihour gear failure experiments ([Dempsey et al. \(2000\)](#), [Gasperin et al. \(2011\)](#)) therefore this was one of the contributions of the current thesis.

A number of prognostic methodologies were studied for gearboxes. Among others a new bootstrap trend extrapolation technique is proposed for RUL point estimation along with

meaningful confidence intervals. This technique is particularly directed to cases where very limited information exists concerning the break down of a component. This is the case for very large and costly mechanical equipment. Moreover, the positive influence of multiple features and sensors combination was highlighted to the direction of a more robust and efficient RUL estimation In meshing gears. This improvement was measured on the proposed bootstrap extrapolation method.

A gear failure-like health index simulation framework was proposed. The particular framework was designed to bypass the lack of actual gear failure histories related to gear pitting fatigue. This framework made it handy to develop and study a number of prognostic methodologies, based on the simulated health index.

AE is highlighted as a robust health monitoring technique for gearboxes. That's due to the shorter transient run in phase it exhibits related to vibration recordings. Moreover, in the steady working condition it seems to be less influenced by rotating speed and axis load configuration. Finally it seems to be better correlated than vibration with the ODM Fe-particle mass recordings and therefore the gear fatigue pitting rate.

The Wavelet Denoising method increases the prognostic as well as diagnostic potential of vibration RMS level in bearings. Moreover, concerning bearing prognostics, feature extraction makes it possible to apply prognostics even in cases where RMS gives no failure trend.

Furthermore ϵ -SVR was highlighted as a highly efficient method for prognosis in bearings as well as in gearboxes.

The models that were developed in the current thesis can be applied in the energy industry such as the gearboxes of wind turbines and critical rotating components in helicopters, oil refineries and cement industries.

Proposals for future research

The gear failure experiments were conducted on soft steel spur gear to accelerate degradation. However, the particular material is not exactly the one that is used in practical gear applications. Thus experiments with hardened steel are deemed helpful to provide a better insight to the actual gear failure mechanisms as well as to confirm the applicability of the proposed prognostics techniques.

An interesting and necessary expansion of the current study is to perform gear failure experiments with alternating torque or/and speed profile. This way the gear fatigue can be studied under more realistic conditions.

Weiner dynamic models, although widely applied for bearing failure are not yet studied for gear failure.

ODM was evidently efficient in capturing gear pitting however its cost is high. We propose the development of a low cost ODM sensor for future research.

Bearing diagnosis based on AE has been proven its efficiency. The turbomachinery group, Energy and Power division of the University of Cranfield has contributed a great deal in this section. On the other hand, bearing quantitative RUL estimation with AE is yet to implemented and studied.

Finally a rolling bearing test rig can be designed to test bearings of larger size than the ones of PRONOSTIA and thus come closer to realistically critical applications. Of course this requires the collaboration of more than one research partner so that the design and cost of construction can be covered.

Appendix A

A.1 The Expectation Maximisation Algorithm

Theorem: Suppose we have two random variables X and Y with distributions $P(X)$ and $P(Y)$ respectively. In the general case these variables are assumed continuous. Let's assume that these two variables are connected via a model θ . Let $P'(X) \equiv P_{\theta'}(X)$ and $P'(Y) \equiv P_{\theta'}(Y)$ the new distributions that arise from a different model θ' .

$$\text{If } \left(\sum_x P(X|Y) \ln \frac{P'(X,Y)}{P(X,Y)} > 0 \right)$$

$$\text{Then } P'(Y) > P(Y)$$

In other words, if we find a model θ' for which the first inequality holds then the observed data Y are better explained, they are more probable, with the new model.

Proof: We know that $\sum_x P(X|Y) = 1$ because $P(X|Y)$ is a probability function that exhausts the events-space X .

Then we can write

$$\ln P'(Y) - \ln P(Y) = \sum_x P(X|Y) \cdot \ln P'(Y) - \sum_x P(X|Y) \cdot \ln P(Y)$$

And due to the law of conditional probability

$$\begin{aligned} \ln P'(Y) - \ln P(Y) &= \sum_x P(X|Y) \cdot \ln \frac{P'(X,Y)}{P'(X|Y)} - \sum_x P(X|Y) \cdot \ln \frac{P(X,Y)}{P(X|Y)} = \\ &= \sum_x P(X|Y) \cdot \ln \frac{P'(X,Y)}{P'(X|Y)} + \sum_x P(X|Y) \cdot \ln \frac{P(X|Y)}{P(X,Y)} = \\ &= \sum_x P(X|Y) \cdot \ln \frac{P'(X,Y)}{P(X,Y)} + \sum_x P(X|Y) \cdot \ln \frac{P(X|Y)}{P'(X|Y)} \end{aligned}$$

The second term of the last relation is the cross entropy of distribution $P(X|Y)$ and $P'(X|Y)$, with the help of Jensen's inequality, is found at least equal to the entropy of $P(X|Y)$. The entropy of P is a non-negative quantity and equals zero when, for some event, $P=1$. Thus we end up

$$\ln P'(Y) - \ln P(Y) \geq \sum_x P(X|Y) \cdot \ln \frac{P'(X,Y)}{P'(X|Y)}$$

And therefore $P'(Y) > P(Y)$. It is worth mentioning that $P(X,Y)$ reads "the probability of X and Y " and $P(X,Y) = P(X)P(Y)$.

Application: Estimating the means and variances of k -Gaussians.

We want to estimate the k means and variances of m observations. Let's assume for simplicity a 1-D set of observations $X = \{x_i, i=1, \dots, m\}$. The parameters of interest will be $\theta = [\mu_1, \mu_2, \dots, \mu_k, \sigma_1, \sigma_2, \dots, \sigma_k]$. Let's denote $Z = [z_1, z_2, \dots, z_m]$ the set of labels that indicate from which gaussian were each observation produced. Now we propose a joint event set Y that combines the observation and labeling information $Y = [y_1, y_2, \dots, y_m]$ with $y_i = (x_i, z_i)$. From EM theorem we know that

$$\text{if } \sum_Y P(Y|X) \ln P'(X,Y) > \sum_Y P(Y|X) \ln P(X,Y)$$

$$\text{then } P'(Y) > P(Y)$$

Therefore the quantity to be maximized with respect to θ' is $J = \sum_Y P(Y|X) \ln P'(X,Y)$. In other words, we assume that the best estimate maximizes the posterior probability, we follow a Maximum A Posteriori (MAP) path.

Since Y completely determines X and $P(Y|X) = 0$ for all Y inconsistent with X , the problem is equivalent to maximising

$$J = \sum_Y P(Y|X) \ln P'(Y) = \frac{1}{P(X)} \sum_Y P(Y, X) \ln P'(Y)$$

Which again is equivalent to maximising $J = \frac{1}{P(X)} \sum_Y P(Y) \ln P'(Y)$

Since Y are random independent events $P'(Y) = \prod_{i=1}^m P'(z_i, x_i)$. Thus

$$J = \frac{1}{P(X)} \sum_Y P(Y) \ln \prod_{i=1}^m P'(z_i, x_i) = \frac{1}{P(X)} \sum_Y P(Y) \sum_{i=1}^m P'(z_i, x_i)$$

We can group the Y in the above sum by unique (x_i, z_i)

$$J = \frac{1}{P(X)} \sum_{j=1}^k \sum_{i=1}^m \ln P'(x_i, z_i = j) \sum_{\substack{z \\ z_i=j}} P(Y) =$$

$$\frac{1}{P(X)} \sum_{j=1}^k \sum_{i=1}^m \ln P'(x_i, z_i = j) P(X, z_i = j)$$

We want to maximize the quantity J with respect to the updated model parameters θ'

First we update the means

$$\frac{\partial J}{\partial \mu'_j} = \frac{1}{P(X)} \frac{\partial}{\partial \mu'_j} \sum_{j=1}^k \sum_{i=1}^m P(X, z_i = j) \ln P'(x_i, z_i = j)$$

$$= \frac{1}{P(X)} \frac{\partial}{\partial \mu'_j} \sum_{j=1}^k \sum_{i=1}^m P(X, z_i = j) \ln [P'(z_i = j) P'(x_i | z_i = j)]$$

$$= \frac{1}{P(X)} \frac{\partial}{\partial \mu'_j} \sum_{j=1}^k \sum_{i=1}^m P(X, z_i = j) \left(-\frac{k}{2} \ln(2\pi\sigma_j) - \frac{(x_i - \mu'_j)^2}{2\sigma_j^2} \right)$$

Where σ_k are the standard deviations of the model from the last iteration. The latter quantity is zero at its optimum. Thus we obtain the update relation for the gaussian means, after dropping the terms that are independent of μ'_j .

$$\frac{1}{P(X)} \sum_{i=1}^m P(X, z_i = j) \frac{\partial}{\partial \mu'_j} \left(-\frac{(x_i - \mu'_j)^2}{2\sigma_j^2} \right) = 0 \Rightarrow$$

$$\frac{1}{\sigma_j^2 P(X)} \sum_{i=1}^m P(X, z_i = j) (x_i - \mu'_j)^T = 0 \Rightarrow$$

$$\mu'_j = \frac{\sum_{i=1}^m P(X, z_i = j) x_i}{\sum_{i=1}^m P(X, z_i = j)} = \frac{\sum_{i=1}^m P(X | z_i = j) x_i}{\sum_{i=1}^m P(X | z_i = j)}$$

In other words, the updated means are found by weighting each observation with the probability estimation of the current model. In a similar manner the update relation for the respective variances

$$\sigma_j'^2 = \frac{\sum_{i=1}^m P(X | z_i = j) (x_i - \mu'_j)^2}{\sum_{i=1}^m P(X | z_i = j)}$$

EM is widely applied in numerous applications with underlying stochasticity due to its iterative nature. There is an estimation step, when we estimate the prior $P(X | z_i = j)$ and a maximisation step, when we update the underlying parameters. The success of EM is due to the iterative nature of the algorithm.

It is worth mentioning that EM almost always converges to an optimum after very few iterations. The final value though is not guaranteed to be a global optimum. Even this way the final parameter values are in practice a good answer to the underlying stochastic problem. EM has been effectively applied in problems such as Hidden Markov Model (HMM) and Kalman filtering parameter estimate.

Appendix B

List of extracted features.

Time domain	
$TD_1 = \sqrt{\frac{\sum_{n=1}^N (x(n))^2}{N}}$	Root Mean Square
$TD_2 = \max x(n) $	Maximum absolute amplitude
$TD_3 = \frac{\sum_{n=1}^N (x(n))^3}{(N-1)TD_1^3}$	skewness
$TD_4 = \frac{\sum_{n=1}^N (x(n))^4}{(N-1)TD_1^4}$	kurtosis
$TD_5 = \frac{TD_2}{TD_1}$	Crest factor
$TD_6 = \frac{TD_2}{\frac{1}{N} \sum_{n=1}^N x(n) }$	Impact factor
Frequency domain	
$FD_1 = \frac{\sum_{k=1}^K (s(k) - TD_2)^3}{K(\sqrt{TD_2})^3}^{14}$	Skewness of spectrum
$FD_2 = \frac{\sum_{k=1}^K (s(k) - TD_2)^4}{K \cdot TD_2^2}$	Kurtosis of spectrum
$FD_3 = \frac{\sum_{k=1}^K f_k s(k)}{\sum_{k=1}^K s(k)}$	Centre of spectrum

¹⁴ TD_2 is used as the energy factor for normalizing the sum of the denominator. This in accordance to Parseval's theorem stating that the total spectral energy is equal to the energy of the timeseries.

$$FD_4 = \sqrt{\frac{\sum_{k=1}^K (f_k - FD_4)^2 s(k)}{K}}$$

Second moment of spectrum with respect to the centre

$$FD_5 = \sqrt{\frac{\sum_{k=1}^K f_k^2 s(k)}{\sum_{k=1}^K s(k)}}$$

normalised second moment of spectrum

$$FD_6 = \sqrt{\frac{\sum_{k=1}^K f_k^4 s(k)}{\sum_{k=1}^K f_k^2 s(k)}}$$

normalised fourth moment of spectrum

$$FD_7 = \frac{\sum_{k=1}^K f_k^2 s(k)}{\sqrt{\sum_{k=1}^K s(k) \sum_{k=1}^K f_k^4 s(k)}}$$

-

$$FD_8 = \frac{FD_4}{FD_3}$$

Shape factor of spectrum with respect the spectrum centre of gravity

$$FD_9 = \frac{\sum_{k=1}^K (f_k - FD_3)^3 s(k)}{K \cdot FD_4^3}$$

Skewness with respect the spectrum centre of gravity

$$FD_{10} = \frac{\sum_{k=1}^K (f_k - FD_3)^4 s(k)}{K \cdot FD_4^4}$$

Kurtosis respect the spectrum centre of gravity

$$FD_{11} = \left| \frac{\sum_{k=1}^K (f_k - FD_3)^{1/2} s(k)}{K \sqrt{FD_4}} \right|$$

-

$$FD_{12} = e^{\frac{\sum_{k=0}^{K-1} \log(s(k))}{K} - \frac{\sum_{k=0}^{K-1} s(k)}{K}}$$

Spectral flatness

$s(k)$ is the energy of the k spectral line, $k=0..K-1$,
 f_k is the k th frequency label with $f_0=0$ and
 $f_k=F_s/2$, F_s being the sampling frequency

References

- Ahmad M., Shekh A.K., (1984), Bernstein Reliability Model: Derivation and Estimation of Parameters, *Reliability Engineering* 8; 131-148
- Alewine, K. (2010), Wind Turbine Generator Failure Modes Analysis and Occurrence, Windpower 2010, Dallas, Texas, May 2010.
- Raad A., Zhang F., Randall B., (2003), On the comparison of the use of AE and vibration analysis for early gearbox detection, the eighth pacific acoustics conference, Melbourne australia.
- Antony J., (2006) The spectral kurtosis: a useful tool for characterizing non-stationary signals, *Mechanical systems and signal processing* 20
- Arulampalam MS, Maskell S, Gordon N, Clapp T (2002). A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking. *IEEE Trans Signal Process* ;50(2):174–88.
- Banjevic D., Jardine A.K.S. (2005), Calculation of reliability function and remaining useful life for a marcof failure process, *IMA Journal of Management Mathematics* 17, 115–130
- Baraldi P, Compare M, Saucò S, Zio E (2013), Ensemble neural network-based particle filtering for prognosis, *Mechanical systems and signal processing* 41(1-2) 288-300
- Baraldi P., Mangili F., Zio E. (2013), Investigation of uncertainty treatment capability of model-based and data-driven prognostic methods using simulated data, *Reliability Engineering & System Safety* Volume 112, April, Pages 94–108
- Benkedjough T, Medjaher K, Zerhouni N, Rechak S, (2013) Remaining useful life estimation based on non-linear feature reduction and support vector machines, *Engineering applications of artificial intelligence* 26(7), 1751-1760
- Bishop C.H. (1995), *Neural networks for pattern recognition*, Oxford university press
- Bolander N., Qiu H., Eklund N., Hindle E., Rosenfeld T. (2009), “Physics based remaining useful life prediction for aircraft engine bearing prognosis,” in *Proc. Annu. Conf. Prognostics Health Manage. Soc.*
- Byington C.S., Garga A.K. (2001) *Handbook of multisensory data fusion*, CRC Press, Boca Raton, FL, D. L. Hall and Linas, eds.pp 23-1-23-31
- Cai T., Silverman B.W.(1998), Incorporating information on Neighboring coefficients into wavelet estimation. Technical report, Department of statistics, Purdue University, West Lafayette, USA
- Cai T., Silverman B.W. (2000) Finite performance of NeighBlock and NeighCoeff estimators. *Sankhya* 63, 127-148.

Casella G.; Berger R. L. (2001) *Statistical Inference*, Second edition. ISBN 978-0534243128 (page 375)

Crabtree, C.J., Feng, Y., Tavner, P.J. (2010), Detecting Incipient Wind Turbine Gearbox Failure: A Signal Analysis Method for Online Condition Monitoring, Scientific Track Proceedings, European Wind Energy Conference 2010, Warsaw,

Combet F., Gelman L. (2007), An automated methodology for performing time synchronous averaging of a gearbox signal without speed sensor, *Mechanical systems and signal processing* 21(6) 2590-2606

Cox D.R. (1972), Regression models and life-tables. *Journal of the Royal Statistical Society B-34* (2): 187–220,

Di Maio F., Tsui K.L., Zio E. (2012), Combining Relevance Vector Machines and exponential regression for bearing residual life estimation", *Mechanical Systems and Signal Processing*, vol.31 , pp. 405-427,

Dempsey P.J. (2003), Integrating oil debris and vibration measurements for intelligent machine health monitoring, Technical report, NASA/TM- 2003-211307.

Dong M, He D (2007) A segmental semi-Markov model (HSMM)-based diagnostics and prognostics framework and methodology, *Mechanical systems and signal processing* 21 2248-2266.

Donoho D., 1(995), De-noising by soft thresholding. *IEEE Transactions on Information Theory* 38(2), 613—627.

Dragomir O.E., Gouriveau R., Dragomir F., Minca E., Zerhouni N.(2009), Review of prognostic problems in Condition- based maintenance, European Control Conference (ECC).

Dupuis R. (2010), Application of Oil Debris Monitoring For Wind Turbine Gearbox Prognostics and Health Management, Annual Conference of the Prognostics and Health Management Society.

Eftekharnjad B (2010), Condition monitoring of gearboxes using Acoustic Emission, PhD thesis, Cranfield university.

Epanechnikov, V.A. (1969). "Non-parametric estimation of a multivariate probability density". *Theory of Probability and its Applications* 14: 153–158

Epleman M. (2007) OE 511/Math 652: Continuous Optimization Methods, Section 1

Fantoni P.F. (2004). NPP Wire System Aging Assessment and Condition Monitoring-A State of The Art Report. HWR-746 OECD Halden Reactor Project. Norway.

Fassois S (2005), Stochastic signals and systems (in greek), Editions of the University of Patras.

Fernandes P.J.L., McDulling (1997), Surface contact fatigue failures in gears, *Engineering failure analysis*, 4(2), pp. 99-107.

Gasperin M., Juricic D., Boskoski P. (2011), Model-based prognosis of gear health using stochastic dynamical models, *Mechanical systems and signal processing*, 25 (2), 537-548.

Gasperin M., Juricic D., Boskoski P. (2011), Model-based Prognostics under no-stationary operating conditions, *Annual conference of the prognostics and health management society*

Gao J.B., Gunn S.R., Harris C.J., Brown M. (2002), A probabilistic framework for SVM regression and error bar estimation, *Machine Learning*, vol.46, pp. 71-89.

Gebraeel N., Lawley M., Liu R., Parmeshwaran V. (2004), Residual life predictions from vibration-based degradation signals: a neural network approach, *Industrial Electronics, IEEE Transactions on* 51 (3), 694-700.

Ghodrati B., Kumar U., Ahmadzadeh F. (2012), Remaining useful life estimation of mining equipment: a case study, in *Proceedings of the International Symposium on Mine Planning and Equipment (MPES '12)*.

Guyon, I., Elisseeff, A. (2003). An Introduction to Variable and Feature Selection, *Journal of Machine Learning Research* 3, pp. 1157-1182.

Harris T.A., *Rolling bearing Analysis*, 4th. Edition, Wiley, New York

Hyvärinen A., Karhunen J., Oja E., *Independent Component Analysis* 2001 John Wiley & Sons

Hokstad P (1997), The failure intensity process and the formulation of reliability and maintenance models. *Reliability Engineering and System Safety*, 58, pp 69-82. 1997

Hubbert M.K. (1956), *Nuclear Energy and the Fossil Fuels. Drilling and Production Practice* (1956) American Petroleum Institute & Shell Development Co. Publication No. 95, See pp 9-11, 21-22.

Huynh K.T., Barros A., Brenguer C. (2012). Maintenance decision-making for systems operating under indirect condition monitoring: Value of online information and impact of measurement uncertainty. *IEEE Transactions on Reliability*, 61(2):410–425.

Jafarizadeh M.A., Hassannejad R., Etefagh M.M., Chitsaz S. , (2008) Asynchronous input gear damage diagnosis using time averaging and wavelet filtering, *Mechanical Systems and Signal Processing* 22 172–201

S. Janjarasjitta, H. Ocak, K.A. Loparo (2008) Bearing condition diagnosis and prognosis using applied nonlinear dynamical analysis of machine vibration signal *Journal of Sound and Vibration* 317 112–126

Jardine A. K. S., Anderson P. M. , Mann D. S. (1987), Application of the weibull proportional hazards model to aircraft and marine engine failure data, *Quality and Reliability Engineering International* Volume 3, Issue 2 pages 77–82, April/June 1987

Kilundu B., Chiementin X., Duez J., Mba D. , Cyclostationarity of Acoustic Emissions (AE) for monitoring bearing defects, *Mechanical Systems and Signal Processing*, Volume 25, Issue 6, August 2011, Pages 2061-2072

Kothamasu R., Huang S.H., Verduin W.H. (2006). System health monitoring and prognostics - a review of current paradigms and practices. *International Journal of Advanced Manufacturing Technology*, 28(9):1012–1024,.

Lee J., Abujamra R., Jardine A.K.S., Lin D., Banjevic D.,(2004) An integrated platform for diagnostics, prognostics and maintenance optimization, *The IMS '2004 International Conference on Advances in Maintenance and in Modeling, Simulation and Intelligent Monitoring of Degradations*, Arles, France

Liu J., Seraoui R., Vitelli V., Zio E., (2013), Nuclear power plant components condition monitoring by probabilistic support vector machine *Annals of Nuclear Energy* Volume 56, June 2013, Pages 23–33

Lloyd S.(1982), Least squares quantization in PCM. *IEEE Transactions on Information Theory* , 28(2):129–137

Loutas T.H., Kalaitzoglou J., Sotiriades G., Kayias E., Kostopoulos V., (2006), Diagnosis of artificial gear defects on single stage gearbox using acoustic emission, *Advanced Materials Research* 13, 415-420

Loutas T.H., Kalaitzoglou J., Sotiriades G., Kostopoulos V. (2008), A novel approach for continuous acoustic emission monitoring on rotating machinery without the use of slip ring, *Journal of Vibration and Acoustics* 130 (6), 064502

Loutas T.H., Sotiriades G., Kalaitzoglou I., Kostopoulos V. (2009), Condition monitoring of a single-stage gearbox with artificially induced gear cracks utilizing on-line vibration and acoustic emission measurements, *Applied Acoustics* 70 (9), 1148-1159

Loutas T.H., Roulias D., Pauly E., Kostopoulos V. (2011), The combined use of vibration, acoustic emission and oil debris on-line monitoring towards a more effective condition monitoring of rotating machinery, *Mechanical Systems and Signal Processing*, vol. 25, no. 4 , pp. 1339-1352,

Loutas T.H., Kostopoulos V. (2012), Utilising the wavelet transform in condition-based maintenance: a review with applications, In *Advances in Wavelet Theory and Their Applications in Engineering, Physics and Technology*

Lu, Gong X., Qiao W.(2012), Current-Based Diagnosis for Gear Tooth Breaks in Wind Turbine Gearboxes, *Dingguo Energy Conversion Congress and Exposition (ECCE)*, 2012 IEEE

Mallat, S.G. (1999) *A Wavelet Tour of Signal Processing*, Academic Press,

McFadden, P. (1987), A revised model for the extraction of periodic waveforms by time domain averaging, *Mechanical Systems and Signal Processing* 1 (1), pages 83-95

McNiff B.P., Musial W.D., Errichello R (1990), Variations in gear fatigue life for Different wind turbine breaking strategies, *AWEA wind Power 90*, Anonymous solar energy research institute

Menold P.H., R.K. Pearson F., Allgower, (1999) Online outlier detection and removal, *Proceedings of the 7th Mediterranean Conference on Control and Automation (MED99)*, Haifa, Isreal, pp. 1110-1133

Medjaher K., Tobon-Mejia D.A., Zerhouni N. (2012), Remaining useful life estimation of critical components with application to bearings, *IEEE Transactions on Reliability*, 61(2), pp. 292-302.

Miao Q., Xie L., Cui H., Liang W., Pecht M. (2013), Remaining useful life prediction of lithium-ion battery with unscented particle filter technique, *Microelectronics reliability* 53 (2013) 805-810.

Miller J.L., Kitaljevich D. (2000), In line oil debris monitoring for aircraft engine condition assessment, *IEEE aerospace conference*, Big sky, Montana USA

Miyachikaka K., Oda S., Koide T. (1995), acoustic emission in bending fatigue process of spur gear teeth, *journal of acoustic emission* 13, (1/2), pp S47-S53

Mohanty A.R., Kar C. (2006), Fault detection in a multistage gearbox by demodulation of motor current waveform. *Industrial electronics, IEEE transactions on*, 54(3) pp. 1285-1297

Moghaddass R. (2013), *Equipment degradation diagnosis and prognostics under a multistate deterioration process*, phd thesis University of Alberta, Canada

Nectoux P., Gouriveau R., Medjaher K., Ramasso E., Morello B., Zerhouni N., Varnier C. (2012), "PRONOSTIA: An Experimental Platform for Bearings Accelerated Life Test", PHM 2012 - 2012 IEEE Int. Conference on Prognostics and Health Management: Enhancing Safety, Efficiency, Availability, and Effectiveness of Systems Through PHM Technology and Application.

Niu G., Yang B.S. (2009) Dempster–Shafer regression for multi-step-ahead time-series prediction towards data-driven machinery prognosis, *Mechanical Systems and Signal Processing*, Vol. 23 (3) , pp. 740-751

Nystad B.H, Gola G., Hulsund J. E.(2012), Lifetime models for remaining useful life estimation with randomly distributed failure thresholds, *European Conference of Prognostics and Health Management Society*

Nystad B.H.(2008), Technical Condition Indexes and Remaining Useful Life of Aggregated Systems, phd Thesis submitted in Faculty of Engineering Science and Technology Department of Marine Technology Norwegian University of Science and Technology Trondheim, June 2008

Ocak H., Loparo K. (2005), HMM-based fault detection and diagnosis scheme for rolling element bearings, *Journal of Vibration and Acoustics*, vol. 127, pp. 299–306,

Oljeindustriens Landsforening (2005) Integrated Work Processes. Future work processes on the Norwegian Continental Shelf. The Norwegian Oil Industry Association.

Peng Z.K., Chu F.L. (2004), Application of the wavelet transform in machine condition monitoring and fault diagnostics: a review with bibliography, *Mechanical Systems and Signal Processing* 18 (2004) 199-221

Peng Y., Dong M. (2011), A hybrid approach of HMM and grey model for age-dependent health prediction of engineering assets, *Expert Systems with application* 38(10) 12946–12953

Pham H.T., Yang B.S. (2010), Estimation and forecasting of machine health condition using ARMA/GARCH model, *Mechanical Systems and Signal Processing*, Volume 24, Issue 2, February, Pages 546-558

Randall R.B., Antoni J., Chobsaard S. (2001), The relationship between spectral correlation and envelope analysis in the diagnostics of bearing faults and other cyclostationary machine signals, *Mechanical Systems and Signal Processing* 15 (5), 945- 962

Randall R.B. (2004), State of the Art in Monitoring Rotating Machinery , SOUND AND VIBRATION/MAY

Reuben L.C.K., Mba D. (2014), Diagnostics and prognostics using switching Kalman filters, Structural Health Monitoring 1–11

Roulias D., Loutas T.H., Kostopoulos V., The application of information processing techniques towards an effective on line prognostic scheme for rotating machinery, 5th international conference on emerging technologies on Non-Destructive Testing, September 2011, Ioannina

Roulias D., Loutas T.H., Kostopoulos V., A hybrid prognostic model for multistep ahead prediction of machine condition, accepted for presentation at COMADEM 2012, the 25th International Congress on Condition Monitoring and Diagnostic Engineering Management, Huddersfield, UK, 18-20 June 2012 (IOP Conference Series).

Saad Al-Dossary , Raja Hamzah R.I., Mba D. (2009), Observations of changes in acoustic emission waveform for varying seeded defect sizes in a rolling element bearing applied acoustics, Volume 70, Issue 1, Pages 58–81

Saxena A., Celaya J., Saha B., Saha S., Goebel K. (2010), “Metrics for offline evaluation of prognostic performance,” Int. J. Prognostics and Health Manage.

Shiroishi J., Li Y., Liang S., (1997), Bearing condition diagnostics via vibration and acoustic emission measurements, mechanical systems and signal processing 11(5) pp. 693-705

Si X.S., Wang W., Hu C.H., Chen M.Y., Zhou D.H. (2013), A Wiener-process-based degradation model with a recursive filter algorithm for remaining useful life estimation, Mechanical Systems and signal processing 35, 219-237

Sikorska J.Z., Mba D (2008), Challenges and obstacles in the application of acoustic emission to process machinery, J Zm, , Proc. IMechE Vol. 222 Part E: J. Process Mechanical Engineering

Sikorska J.Z., Hodkiewicz M., Ma L. (2011), Prognostic modeling options for remaining useful life estimation by industry, Mechanical Systems and Signal Processing 25 (2011) 1803–1836

Singh A, Houser DR, Vijayakar S. (1999), Detecting Gear tooth breakage using acoustic emission: A feasibility and Sensor placement study, Journal of Mechanical Design, transactions of the ASME, 121(4) pp.587-593

Smola A. J., Schölkopf B. (2004), A tutorial on support vector regression, *Statist. and Computing*, vol. 14, pp. 199–222.

- Sun J., Zuo H., Wang W., Pecht M. (2012), Application of a state space modeling technique to system prognostics based on a health index for condition-based maintenance, *Mechanical Systems and Signal processing* 28, pp. 585-596
- Tan, C.K. (2005), An investigation on the diagnostic and prognostic Capabilities of Acoustic emission in spur gear, PhD thesis, Cranfield University
- Tan, C.K., Irving, P., Mba, D (2007), A comparative experimental study on diagnostic and prognostic capabilities of acoustic emission, vibration and spectrometric oil analysis for spur gears, *mechanical systems and signal processing*, 21(1) pp. 208-233
- Tandon N., Mata S. (1999), Detection of defects in gears by acoustic emission measurements, *journal of acoustic emission*, 17 (1-2) pp 23-27
- Tavner, P. J. (2008), Review of Condition Monitoring of Rotating Electrical Machines, *IET Electric Power Applications*, Vol. 2, No. 4, pp. 215-247,
- Tobon-Mejia D.A., Medjaher K., Zerhouni N., Tripot G. (2012), A Data-Driven Failure Prognostics Method Based on Mixture of Gaussians Hidden Markov Models, *IEEE Transactions on Reliability*, 61(2), 491 - 503.
- Tran T.V., Pham H.T., Yang B.S., Nguyen T.T.(2012), Machine performance degradation assessment and remaining useful life prediction using proportional hazard model and support vector machine, *Mechanical system and signal processing* 32, pp. 320-330
- Tse P.W., Atherton D.P.(1999), Deterioration using vibration based fault trends and recurrent neural networks, *Journal of vibration and acoustics* Vol.121,355-362
- Vachtsevanos G, Wang P (2001), FAULT PROGNOSIS USING DYNAMIC WAVELET NEURAL NETWORKS, *AUTOTESTCON Proceedings*, IEEE Systems Readiness Technology Conference
- Vachtsevanos G., Lewis F. L., Roemer M., Hess A., Wu B. (2006), *Intelligent Fault Diagnosis and Prognosis for Engineering Systems*, John Wiley and sons
- Wang Q.W., Golnaraghi M.F., Ismail F., Prognosis of machine health condition using neuro-fuzzy systems, *Mechanical systems and signal processing* 18(2004) 813-831
- Wang W., Carr M., Xu W., Kabbacy K. (2011), A model residual life prediction based on Brownian motion with an adaptive drift, *Microelectronics reliability* 51; 285-293
- Wei Z, Tao T, ZhuoShu D, Zio E, A dynamic particle filter-support vector regression method for reliability prediction, *Reliability engineering and system safety* 119, 109-116
- Vapnik V.(1995), *The Nature of Statistical Learning Theory. Statistics for Engineering and Information Science*, Second ed. : Springer—Verlag,

Zhang X., Kang J., Zhao J., Teng H (2014), Bearing run to failure data simulation for condition based maintenance, TELKOMNIKA Indonesian Journal of Electrical Engineering Vol.12, No.1, pp. 514 ~ 519

Zio E., Piloni G. (2011) Particle filtering prognostic estimation of the remaining useful life of non-linear components, Reliability engineering and systems safety 96, 403-409

<http://icmu.nyc.gr/%20Six-Tech-Advancements-Changing-the-Fossil-Fuels-Game>

Journal publications

- 2014 **Roulias D**, Loutas TH, Kostopoulos V, Estimation of gearbox remaining useful life utilizing an ensemble of exponentials approach, *Quality and Probability Engineering International* (Wiley) ([submitted under review](#))
- 2013 TH Loutas, **D Roulias**, G Georgoulas, Remaining Useful Life Estimation in Rolling Bearings Utilizing Data-Driven Probabilistic E-Support Vectors Regression, *IEEE Transactions on Reliability*, 62 (4) , art. no. 6645455 , pp. 821-832
- 2013 **D Roulias**, T Loutas, V Kostopoulos, A statistical feature utilising wavelet denoising and neighblock method for improved condition monitoring of rolling bearings, *Chemical Engineering Transactions* 33 , pp. 1045-1050
- 2012 TH Loutas, A Panopoulou, **D Roulias**, V Kostopoulos, Intelligent health monitoring of aerospace composite structures based on dynamic strain measurements, *Expert Systems with Applications* 39 (9), 8412-8422
- 2011 A Panopoulou, **D Roulias**, TH Loutas, V Kostopoulos, Health monitoring of aerospace structures using fibre bragg gratings combined with advanced signal processing and pattern recognition techniques, *Strain* 48 (3), 267-277
- 2011 TH Loutas, **D Roulias**, E Pauly, V Kostopoulos, “The combined use of vibration, acoustic emission and oil debris on-line monitoring towards a more effective condition monitoring of rotating machinery”, *Mechanical Systems and Signal Processing*, Volume 25, Issue 4, May 2011, Pages 1339-1352
- 2011 A Panopoulou, T Loutas, **D Roulias**, S Fransen, V Kostopoulos, Dynamic fiber Bragg gratings based health monitoring system of composite aerospace structures, *Acta Astronautica* 69 (7), 445-457

Conference Proceedings publications

- 2011 A Panopoulou, TH Loutas, V Kostopoulos, Gomez-Molinero Vcente, S. Fransen, **D Roulias**, Dynamic Fibre bragg grating system for the damage detection of composite reflector antenna, 62nd International Astronautical Congress 2011
- 2011 TH Loutas, **D Roulias**, V Kostopoulos, Condition monitoring of rotating machinery combining multi-sensor data from vibration, acoustic emission and oil debris on-line monitoring. 7th National Conference on NDT of the Hellenic Society of NDT
- 2010 A Panopoulou, **D Roulias**, TH Loutas, N Chrysochoidis, V Kostopoulos, Advanced signal processing and pattern recognition techniques for health monitoring of aerospace structures ,9th HSTAM INTERNATIONAL CONGRESS ON MECHANICS

Presentations in Conferences

- 2013 **Roulias D**, Loutas T, Kostopoulos V, A statistical feature utilising wavelet denoising and neighblock method for improved condition monitoring of rolling bearings 2013 Prognostics and System Health Management Conference
PHM-2013 Milan 8-11 September, 2013
- 2012 **Roulias D**, Loutas TH, Kostopoulos V, A hybrid prognostic model for multistep ahead prediction of machine condition, accepted for presentation at COMADEM 2012, the 25th International Congress on Condition Monitoring and Diagnostic Engineering Management, Huddersfield, UK, 18-20 June 2012 (IOP Conference Series).
- 2011 **Roulias D.**, Loutas, T.H., Kostopoulos, V., The application of information processing techniques towards an effective on line prognostic scheme for rotating machinery, 5th international conference on emerging technologies on Non-Destructive Testing, September 2011, Ioannina

CURRICULUM VITAE

DIMITRIOS ROULIAS

ADDRESS AND PERSONAL DETAILS

Address: 67 Epidavrou Str.
Patras, 26442 – Greece
Tel: +30972574772
E-Mail: mroulias@yahoo.com
droulias@mech.upatras.gr
Nationality: Hellenic
Date of Birth: 19 September 1982, Preveza
Marital Status: Single
Military Service: Fulfilled
Current Status: Research Engineer
WebPage:

- https://www.researchgate.net/profile/Dimitris_Roulias2
- <http://scholar.google.gr/citations?user=9DZc0n4AAAAJ&hl=en&oi=ao>

EDUCATION

03/2008- today **PhD candidate**
Applied Mechanics Lab., Mechanical Eng. & Aeronautics, University of Patras
PhD Thesis: “**Methodologies for remaining useful life estimation with multiple sensors for rotating machinery**”. Prof. V. Kostopoulos
10/2000- 07/2007 **MEng & BEng in Electrical Engineering**, (6.97/10, Very Good), University of Patras.
Diploma Thesis: “**The effect of photoionisation in streamer propagation**”, 2007
03/2005- 06/2005 **Student Visitor Researcher**, Fellowship of the Greek State
Laboratoire des gaz et des plasmas Universite de Pau (UPPA)
“**Numerical estimation of ionisation phenomena in gases**”. Prof. J F Loiseau
2000 **High School**, Preveza, (19.2/20, Excellent)

MEMBER

- **Member of Technical Chamber of Greece** (since 2008)

RESEARCH INTERESTS

Strong theoretical background in the following research topics

- **Stochastic time series analysis for machinery failure diagnostics and prognostics: Kalman filters, linear ARMA processes, Particle filters, Random walk processes, Markov chains**
- **State of the art digital signal processing : Wavelets and wavelet lifting scheme, Teager Huang transform, Empirical Mode Decomposition, Higher order spectral analysis, cyclostationary signal analysis, kurtogram and envelop analysis, blind source separation (ICA, PCA, NNMF, ISOMAP)**
- **Machine learning for rotating machinery diagnosis and prognostics: clustering for large data, support vectors, soft computing techniques, Fuzzy inference systems**
- **Acoustic Emission, vibration and sound analysis for machinery diagnostics and prognostics**
- **Diagnostics and prognostics in rotating machinery and aerospace structures**

COMPUTER USE

- General use: OFFICE / ORIGIN / TecPlot
- Fluent in: MATLAB
- Good knowledge of FORTRAN, C, C++, ASSEMBLY
- Good knowledge of Labview

LANGUAGE KNOWLEDGE

- Greek (native)
- English (SOL Cambridge Proficiency)
- French

OTHER CERTIFICATIONS

- Energy Inspector

OTHER ACTIVITIES

Boxing, Economics, Dancing



European Union
European Social Fund

This research has been co-financed by the European Union (European Social Fund – ESF) and Greek national funds through the Operational Program "Education and Lifelong Learning" of the National Strategic Reference Framework (NSRF) - Research Funding Program: Heracleitus II. Investing in knowledge society through the European Social Fund.

