

Tribological assessment of oil condition sensors for marine launching equipment

Angel Torres Pérez

A thesis submitted in partial fulfilment of the requirements of
Bournemouth University for the degree of Doctor of Philosophy

Bournemouth University

April 2011

Copyright statement

This copy of the thesis has been supplied on condition that anyone who consults it is understood to recognise that its copyright rests with its author and due acknowledgement must always be made of the use of any material contained in, or derived from this thesis

Abstract

Problems in hydraulic systems associated with hydraulic fluids are an important area of research. Time has proven that contamination control in hydraulic and lubricating systems are fundamental to reliability and performance. Contamination control is needed to guarantee the quality of the oil and machine reliability. Among the several condition monitoring techniques, oil and wear analysis are the most effective for contamination control in hydraulic equipment. An emerging maintenance philosophy is oil condition based maintenance, in which the active monitoring of oil parameters determines the machine health and the variable service intervals. This new maintenance philosophy requires sensor technology able to provide real time indicators of the status of the system.

Nowadays, oil condition monitoring sensors are not a mature technology and their performance has not been widely assessed under controlled experimental conditions. This research explains the physical fundamentals of commercially available sensors and it discusses and evaluates their effectiveness under controlled experimental conditions.

Tribology tests were performed for sliding contacts as they are the most predominant type of contacts within hydraulic systems. Results reveal that several characteristics of the sensors must be improved for a more meaningful output and for an earlier detection of abnormal trends which are typical indicators of the onset of faults.

Finally, of all the future trends of oil condition monitoring sensors, the novel methodology to calculate the particle size distribution from ferrous debris density measurements is the most useful and important contribution to knowledge of this research. The proposed method when compared to current technology would bring a new type of particle counter that could break the technological size limitation of particle counters based on magnetometry (the most extended type in industrial and military machinery), leading to earlier fault detection. Improvements of this methodology would allow further development of low cost particle counters in the micrometer and submicrometer range which can be widely applied in many industrial processes and scientific disciplines.

Acknowledgment

I would like to express my gratitude to the following for their contributions and help during the course of this work.

Professor Mark Hadfield and Steve Austen for their role as supervisors and for giving me the opportunity of doing this research.

Dr. Gonzalo García-Atance Fatjó for his constant interest in my research and valuable comments.

Staff from Tolpuddle House for their help and support, especially thanks goes to Kevin Smith and Ken Millard.

I wish to acknowledge the financial support provided by the Engineering and Physical Sciences Research Council (EPSRC) and Royal National Lifeboat Institution (RNLI).

Especial thanks also go to my parents and my sister for their support and encouragement during these years. I could not have completed this document without their advice, particularly when I faced adverse situations which have strongly undermined my motivation to finish this thesis.

Publications Arising From This Research

Journal Articles

A. Torres Pérez, A. Hernández Battez, G. García-Atance, J.L. Viesca, R. González, M. Hadfield 2010. Use of optical profilometry in the ASTM D4172 standard. *Wear* (In Press)

A. Torres Pérez, G. García-Atance Fatjo, M. Hadfield, S. Austen, 2010. A model of friction for a pin-on disc configuration with imposed pin rotation. *Journal of Mechanism and Machine Theory* (In Press)

A. Torres Pérez, M. Hadfield, S. Austen, 2010. Determination of particle size distribution from ferrous debris density measurements. *Sensors and Actuators A* (Submitted)

Conference articles and presentations

A. Torres Pérez, M. Hadfield, S. Austen, 2011. Oil quality sensors based on changes of complex permittivity: State of the art and future trends, The Eighth International Conference on Condition Monitoring and Machinery Failure Prevention Technologies. Cardiff: MFPT and BINDT (2011).

A. Torres Pérez, M. Hadfield, S. Austen, 2010. On-line Ferrous Debris Density monitoring in sliding area contacts under boundary lubrication regime, The Seventh International Conference on Condition Monitoring and Machinery Failure Prevention Technologies [CD-ROM]. Stratford-Upon-Avon: MFPT and BINDT, (2010), ISBN 978-1-901892-33-8.

A. Hernández Battez, A. Torres Pérez, G. García-Atance, J.L. Viesca, R. González, M. Hadfield, 2009. Advantages of using optical profilometry in the ASTM D4172 standard. STLE 65th Annual Meeting & Exhibition.

Conference presentations

A. Torres Pérez, 2010. Future trends of condition monitoring of tribosystems using oil quality and wear debris sensors. Mission of tribology research 19 (IMEche)

Available on-line in: www.imeche.org/events/s1544

A. Torres Pérez, 2010. Characterisation of the rheological properties of lubricants for EHL film thickness prediction. TriboUK, Imperial College, London, 22-23 April 2010.

A. Torres Pérez, 2009. Output characterisation of oil condition monitoring sensors in terms of the lubrication regime. Mission of tribology research 18 (IMEche)

Available on-line in: www.imeche.org/events/s1473

Author's Declaration

This thesis contains the original work of the author except where otherwise indicated

Table of Contents

Abstract	I
Acknowledgment	II
Publications Arising From This Research	III
Author's Declaration	V
List of Figures	XI
List of Tables	XVIII
List of Equations	XX
Nomenclature	XXIV
Abbreviations and Acronyms	XXIX
1 Introduction	1
1.1 Research Objectives and Scope	7
1.2 Outline of the Report	8
2 Theoretical background	9
2.1 Review of lubrication theory	9
2.2 Review of flash temperature in sliding contacts	13
2.3 Reliability and reproducibility of bench tests	14
2.4 Bench test results analysis using lubrication theory	14
2.4.1 A real example under ASTM D4172	15
2.4.1.1 Experimental details	16
2.4.1.2 Results and discussion	17
2.4.1.3 Conclusions of the example	26
3 Wear prediction under the statistical approach	28
3.1 Real Area of Contact, an important concept for devising adhesive wear models	28
3.2 Probability of Asperity Contact	29

3.3	Volume of Adhesive Wear in Lubricated Contacts	32
3.4	Model of friction for a pin-on-disc configuration with imposed pin rotation.	34
3.4.1	Introduction	34
3.4.2	Kinematic analysis of the pin-on-disc configuration with imposed rotation of the pin.	35
3.4.3	A probabilistic friction force model for pin-on-disc configuration.	37
3.4.4	Normalised results of the preliminary probabilistic friction model.	40
3.4.5	Validation of the preliminary model with experimental data.	45
3.4.6	Improved probabilistic model for pin-on-disc configuration.	54
3.4.7	Possible applications of the improved model	58
3.4.8	Validation of the preliminary model with experimental data.	59
3.4.9	Conclusions	69
4	Oil condition monitoring techniques	71
4.1	Maintenance strategies and techniques	71
4.2	Condition-Based Predictive Maintenance	73
4.3	Oil condition monitoring	75
4.3.1	Process viscometers	76
4.3.1.1	Rheology	76
4.3.1.2	Devices to measure viscosity	80
4.3.1.3	Principles of measurement of AST100	81
4.3.1.4	Commercial alternatives to AST100 viscometer	84
4.3.1.5	The effectiveness of process control viscometers in hydraulic oils contaminated with water	85
4.3.2	Sensors to monitor the water content in oil	87
4.3.3	Tan Delta and Impedance Spectroscopy sensors	92
4.3.3.1	Introduction to Impedance Spectroscopy (IS) in lubrication oils	92
4.3.3.2	Current commercial technology	95
4.3.3.3	Single frequency circuits based on bridges and Tan δ sensors	97
4.3.3.4	New ideas for the development of a low-cost and highly-integrated oil condition IS sensor	105
4.3.4	Particle counters and ferrous debris density sensors	108
4.3.4.1	Ferrous debris density technology	110
4.3.4.2	Particle counters technology	115

4.3.4.3	The influence of particle deposition in ferrous debris density measurements	120
4.3.4.3.1	Test Methodology	122
4.3.4.3.2	Results	124
4.4	Design of the oil condition monitoring rig	129
4.4.1	Selection criteria of oil condition sensors	129
4.4.2	Test rig design	130
5	<i>Determination of particle size distribution from ferrous debris density measurements</i>	135
5.1	Classification of a single particle	136
5.2	Motion of particles under gravity through a fluid	138
5.3	Determination of the drag coefficient for a given particle diameter	145
5.4	Methodology for the determination of the particle size distribution from particle density measurements	150
5.4.1	Numerical model	150
5.4.1.1	Input Parameters	151
5.4.1.2	Determination of the drag coefficients for a given particle diameter	152
5.4.1.3	Elapsed time to settle down for each particle diameter	154
5.4.1.4	Construction of the M matrix for the range of simulating times and particle diameters	155
5.4.1.5	Curve fitting algorithm for ferrous debris density curves	157
5.4.1.6	Solution to the overdetermined system of equations	158
5.5	Validation of the methodology	158
5.5.1	Experimental set-up	158
5.5.2	Results	160
5.5.3	Discussion	162
5.5.4	Conclusions	165
5.6	Future research direction in particle size distribution modelling	165
6	<i>Discussion and conclusions</i>	169
6.1	Contributions to knowledge	169
6.2	Final remarks	170
6.3	Recommendations for further work	173

References	175
Appendices	185
A Wear	185
A.1 Wear mechanisms	185
A.1.1 Abrasive Wear	185
A.1.2 Adhesive Wear	186
A.1.3 Cavitation Wear	187
A.1.4 Corrosive Wear	187
A.1.5 Contact Stress Fatigue Wear	187
A.2 Contact pressure, Hertz theory	189
A.2.1 Simplifying Assumptions to Hertz's Theory	189
A.2.2 Geometry of Contacting Elastic Bodies	190
A.2.2.1 Two Elastic Bodies with Convex Surfaces in Contact	190
A.2.2.2 Convention	191
A.2.3 Two Elastic Bodies One Convex and One Flat Surface in Contact	192
A.2.4 Two Elastic Bodies with One Convex and One Concave Surface in Contact	193
A.2.4.1 Contact area, pressure	194
A.2.4.2 Contact between two spheres (point contact)	194
A.2.5 Contact between a sphere and a plane	196
A.2.6 Elliptical Contact between Two Elastic Bodies, General Case	196
A.2.7 Sliding speed and coefficient of friction for sliding four-ball and pin / ball-on-disc configurations	200
A.2.7.1 Four Ball Sliding Wear Test	200
A.2.7.2 Pin and ball-on-disc	203
B Lubrication theory	204
B.1 Viscosity classification	204
B.1.1 ISO Viscosity Grades	204
B.1.2 Hydraulic Fluid Classification	205
B.1.3 Viscosity temperature relationships	206
B.1.3.1 Viscosity-Temperature Equations	206
B.1.3.2 Effect of temperature	207
B.1.3.3 Viscosity index	207
B.1.3.4 Viscosity pressure relationship	208
B.1.4 Load-Bearing Capacity and Antiwear Properties	209

B.1.5	Setting Point/Pour Point _____	209
B.1.6	Gas Solubility _____	210
B.1.7	Ageing Behaviour _____	211
B.1.7.1	Detergent and Dispersant _____	211
B.2	Film thickness formulae for EHL contacts _____	212
B.2.1	Surface temperature at the conjunction between contacting solids ____	222
B.2.1.1	Flash Temperature in Circular Contacts _____	227
B.2.1.2	True Flash Temperature Rise _____	230
C	<i>Maintenance strategies</i> _____	231
D	<i>Topographic measurement considerations for sphere samples with optical profilometers</i> _____	237
D.1	The effect of horizontal errors in ball measurements _____	237

List of Figures

Figure 1.1. Machine failures using the particle size distribution (Baur 1982; Lockwood and Dalley 1992).....	2
Figure 1.2. Progress of wear.	3
Figure 2.1. Generalised Stribeck curve (Schipper 1988).....	9
Figure 2.2. Wear scar diameter (WSD) versus testing time for both hydraulic oils (392N)	18
Figure 2.3. Wear scar diameter (WSD) versus testing time for both hydraulic oils (147N)	18
Figure 2.4. Wear volume versus testing time for both hydraulic oils (392N)	19
Figure 2.5. Wear volume versus testing time for both hydraulic oils (147N)	19
Figure 2.6. Results of one of the replicas under ASTM D4172 test conditions at 60 min (images from the optical profilometer, wear volume calculated by the proposed procedure).....	20
Figure 2.7. Percentage of variation of the WSD and wear volume between Mobil DTE 13M and Esso Nuto H46 samples	21
Figure 2.8. Percentage of variation of the WSD and wear volume between Mobil DTE 13M and Esso Nuto H46 samples	21
Figure 2.9. Wear scar (3D images) 147N	22
Figure 2.10. Wear scar (3D images) 392N	23
Figure 2.11. Wear volume versus wear scar diameter for both hydraulic oils (392N)	25
Figure 2.12. Wear volume versus wear scar diameter for both hydraulic oils (147N)	26
Figure 3.1. Block diagram to calculate minimum film thickness	31
Figure 3.2. Block diagram to calculate the lubrication regime and different probabilities (Stolarski 1996).....	32

Figure 3.3. Velocity vectors at point P within the pin with imposed rotation ω_p in the pin-on-disc configuration. The system of reference (X', Y') is fixed to the disc36

Figure 3.4. Normalised x-component of the friction force (3D). $\alpha \in [10^{-2}, 10^3]$, $r_p/R \in [0.009, 0.9]$40

Figure 3.5. Normalised x-component of the friction force (2D). $\alpha \in [10^{-2}, 10^3]$, $r_p/R \in [0.009, 0.9]$41

Figure 3.6. Normalised friction torque (3D). $\alpha \in [10^{-2}, 10^3], r_p/R \in [0.009, 0.9]$41

Figure 3.7. Normalised friction torque (2D). $\alpha \in [10^{-2}, 10^3], r_p/R \in [0.009, 0.9]$42

Figure 3.8. Normalised x-component of the friction force (3D). $\alpha \in [-10^3, -10^{-2}]$, $r_p/R \in [0.009, 0.9]$42

Figure 3.9. Normalised x-component of the friction force (2D). $\alpha \in [-10^3, -10^{-2}]$, $r_p/R \in [0.009, 0.9]$43

Figure 3.10. Normalised friction torque (3D). $\alpha \in [-10^3, -10^{-2}], r_p/R \in [0.009, 0.9]$43

Figure 3.11. Normalised friction torque (2D). $\alpha \in [-10^3, -10^{-2}], r_p/R \in [0.009, 0.9]$44

Figure 3.12. Friction force F_{Rx} 45

Figure 3.14. Pin friction torque T_F 46

Figure 3.15. Friction force F_{Rx} acting on the pin. Positive alpha (α)48

Figure 3.16. Friction torque acting on the pin. Positive alpha (α)49

Figure 3.17. Friction force F_{Rx} acting on the pin. Negative alpha (α)49

Figure 3.18. Friction torque acting on the pin. Negative alpha (α).....50

Figure 3.19. Probability of asperity contact (PAC) is a function of α . Obtained from the experimental x-component of the friction force acting on the pin.
Positive α values.51

Figure 3.20. Friction torque predictions if the PAC varies with α .
Positive α values.52

Figure 3.21. Probability of asperity contact (PAC) is a function of α . Obtained from the experimental x-component of the friction force acting on the pin.
Negative α values.....52

Figure 3.22. Friction torque predictions if the PAC varies with alpha (α).
 Negative α values.....53

Figure 3.23. Elastic model of the pin. Principle of Superposition.
 L is the applied load.55

Figure 3.24. Equivalent force system for the determination of k_155

Figure 3.25. Maximum value of the probability of asperity contact (PAC) is a
 function of α . Obtained from the experimental x-component of the friction force
 acting on the pin. Positive α values.....60

Figure 3.26. Predictions of the y-component of the frictional force.
 Positive α values.60

Figure 3.27. Friction torque predictions of the proposed model. Positive α values. .61

Figure 3.28. Maximum value of the probability of asperity contact (PAC) is a
 function of α . Obtained from the experimental x-component of the friction force
 acting on the pin. Negative α values.61

Figure 3.29. Predictions of the y-component of the frictional force.
 Negative α values.....62

Figure 3.30. Friction torque predictions of the proposed model. Negative α values.62

Figure 3.31. PAC predictions for PP for different ϕ . Positive α values.....63

Figure 3.32. PAC predictions for LDPE for different ϕ . Positive α values.64

Figure 3.33. PAC predictions for PMMA for different ϕ . Positive α values.65

Figure 3.34. PAC predictions for PP for different ϕ . Negative α values.66

Figure 3.35. PAC predictions for LDPE for different ϕ . Negative α values.....67

Figure 3.36. PAC predictions for PMMA for different ϕ . Negative α values.68

Figure 4.1. Overview of common maintenance strategies and methods
 (Al-Shayea 2010)72

Figure 4.2. Typical connection types for oil sampling and analysis.....76

Figure 4.3. The two-plate model for the definition of the rheological parameters. ...77

Figure 4.4. Simplified classifications of fluids based on the characteristic of the viscosity curve (top) and the flow curve (bottom)..... 78

Figure 4.5. Conceptual diagram of AST100 viscometer..... 82

Figure 4.6. Phase shift vs. excitation frequency for two different damping values... 83

Figure 4.7. Viscosity results for different water in oil concentrations..... 86

Figure 4.8. Flow diagram for the measurement and characterisation of a material-electrode system (Macdonald 1992). 93

Figure 4.9. Bridge with voltage excitation..... 97

Figure 4.10. $\tan\delta$ sensors with cylindrical electrodes. 99

Figure 4.11. Definition of complex relative permittivity..... 100

Figure 4.12. Equivalent circuit for cylindrical electrodes ($\tan\delta$ sensor's head)..... 100

Figure 4.13. Operational range of impedance for Kittiwake's cylindrical sensor head. 101

Figure 4.14. $\tan\delta$ frequency sweep for a new (2A) and used oil (2B) (In green frequency candidates for the implementation of the sensor) (Murphy and Kent 2004)..... 102

Figure 4.15. $\tan\delta$ sensor (HW implementation)..... 103

Figure 4.16. Analysis of the circuits, for temperature and signal conversion ($V_{fd}\approx 0.6V$)..... 104

Figure 4.17. Oscillator circuit with the equivalent circuit for the sensor head, R_3 range represents the dielectric losses 105

Figure 4.18. Proposed system for IS studies using an improved Howland current source and GIC converter..... 107

Figure 4.19. Block diagram of the dual coil magnetometer (Hutchings 2000) 111

Figure 4.20. Implementation of the sensing part of Kittiwake's sensor, dashed lines are according to aforementioned block diagram (Hutchings 2000)..... 113

Figure 4.21. Output voltage waveforms from the sensing coils 114

Figure 4.22. Ferrous debris density sensor (Murphy and Kent 2004) 115

Figure 4.23. Kittiwake’s particle counter (Murphy and Kent 2004)..... 116

Figure 4.24. Typical waveforms for ferromagnetic and non-ferromagnetic conductive particles (Murphy and Kent 2004). 117

Figure 4.25. Particle counter based on a differential transformer..... 118

Figure 4.26. Particle counter based on a voltage control oscillator and PLL (Whittington et al. 1992). 120

Figure 4.27. Ferrous Debris Density Sensor, Kittiwake Sensor’s Suite and experimental rig. 122

Figure 4.28. Normalised ferrous debris density 124

Figure 4.29. Evolution of wear and roughness of pin counterface (Mobil DTE 13M) 126

Figure 4.30. Topographical images of the evolution of wear 127

Figure 4.31. Ferrogram analyses of used oil 127

Figure 4.32. Functional sketch of the oil condition monitoring rig 132

Figure 4.33. Oil condition monitoring rig and the pin-on-disc assembly for lubricated test 133

Figure 5.1. Single particle diameters used in microscope analysis (Rhodes 2008). 137

Figure 5.2. Particle travelling through the fluid under the action of three forces. ... 138

Figure 5.3. Particle travelling through the fluid (Stokes region). 143

Figure 5.4. Particle travelling through the fluid (Stokes region) 143

Figure 5.5. Particle travelling through the fluid (Stokes region) 144

Figure 4.7. Particle travelling through the fluid (Assuming Stokes region) 148

Figure 5.8. Example of settling speeds for different lubricant temperatures 150

Figure 5.9. Flowchart for the determination of particle size distribution from particle density measurements 151

Figure 5.10. Example of drag coefficients for different spherical particle diameters 153

Figure 5.11. Example of Reynolds Numbers for different spherical particle diameters 153

Figure 5.12. Elapsed time for a particle to travel a distance of 10cm..... 154

Figure 5.13. Elapsed time for a particle to settle down (approach to the problem). 155

Figure 5.14. Graphical representation of the M matrix (The size is 200 rows and 10^4 columns)..... 157

Figure 5.15. Experimental rig for ferrous debris density measurements 159

Figure 5.16. Deposition of ferrous debris at different temperatures 160

Figure 5.17. Debris density distribution (blue) and cumulative debris density distribution (red) for each radius of particles for a fluid temperature of 20°C..... 163

Figure 5.18. Debris density distribution (blue) and cumulative debris density distribution (red) for each radius of particles for a fluid temperature of 27°C..... 163

Figure 5.19. Debris density distribution (blue) and cumulative debris density distribution (red) for each radius of particles for a fluid temperature of 45°C..... 164

Figure 5.20. Debris density distribution (blue) and cumulative debris density distribution (red) for each radius of particles for a fluid temperature of 75°C..... 164

Figure 5.21. Centrifugal acceleration in ‘g’ for a radius up to 20cm and rpm up to 6000 166

Figure 5.22. Sketch of a sensor device under “centrifugal gravity”. 168

Figure A.1. Convex surfaces in contact (Stachowiak and Batchelor 2001)..... 190

Figure A.2. One convex and one flat surface in contact (Stachowiak and Batchelor 2001) 192

Figure A.3. One convex and one concave surface in contact (Stachowiak and Batchelor 2001)..... 193

Figure A.4. Contact between two spheres (Stachowiak and Batchelor 2001)..... 194

Figure A.5. Contact between a sphere and a plane (Stachowiak and Batchelor 2001) 196

Figure A.6. k_1, k_2 dependence with k_0 (Stachowiak and Batchelor 2001) 198

Figure A.7. Sliding four-ball test, lower balls and upper view200

Figure A.8. Upper Ball and one of the lower balls, sectional view.201

Figure B.1. Viscosity System for Industrial Fluid Lubricants204
 (Theo Mang 2007)204

Figure B.2. Low temperature viscosity grades for hydraulic fluid classification
 (ASTM 6080) (Nadkarni 2000)205

Figure B.3. Film thickness chart using Hamrock and Dowson Formulae (3D).....219

Figure B.4. Lubrication regime map using Hamrock and Dowson Formulae (2D) 220

Figure B.5. Central film thickness map using Hamrock and Dowson
 Formulae (2D).....220

Figure B.6. Central film thickness map using Nijenbanning Formulae (2D)221

Figure B.7. Different contact areas for surface contact temperature calculations ...225

Figure B.8. Dimensionless Peclet number227

Figure D.1. Differences in heights (e_v) versus the horizontal error (e_h) for the
 constrained sphere fitting algorithm, $R=6.35\text{mm}$ 239

Figure D.2. Differences in heights (e_v) versus the horizontal error (e_h) for the
 constrained sphere fitting algorithm when the peak value is not located
 in the centre239

Figure D.3. Errors in z (heights), after subtracting the sphere using the
 unconstrained algorithm.....240

Figure D.4. Errors in z (heights), calibration procedure242

List of Tables

Table 1.1. Tribological stresses in hydraulic machinery (Totten 1999) (A) Axial piston machine (R) Radial piston machine (E) extremely important with respect to failure behaviour	6
Table 2.1. Main properties of the lubricants (2011a, 2011b).....	16
Table 2.2. Film thickness and lubrication regime using the Hamrock and Dowson formulae.	16
Table 3.1. Experimental data obtained from (Stolarski 1989a)	47
Table 3.2. Best values of p in least mean square sense (LMSE) for $\phi \in [10^{-2}, 10]$	48
Table 3.3. Best values of p in least mean square sense (LMSE) for $\phi \in [-10, -10^{-2}]$...	48
Table 4.1. Summary of various condition monitoring methods (Appleby 2010)	74
Table 4.2. AST100 data frame transmitted every second (RS232).....	84
Table 4.4. Geometrical dimensions of the electrodes of commercial $\tan\delta$ sensors.	101
Table 4.4. Coefficient of friction	124
Table 4.5. Roughness evolution of the pin. The lubricant is Mobil DTE 13M	126
Table 4.6. Process viscometer selection according to quality versus price	129
Table 5.1. Viscosity of the lubricant under test	161
Table 5.2. Prediction model parameters.....	161
Table 5.3. Input parameters for the tests results.....	161
Table A.1. Differentiation of surface contact fatigue from other wear processes (Abrasion, Adhesion and Corrosion)	188
Table A.2. Contact between two spheres	195
Table A.3. Contact between two elastic bodies, general case.....	197
Table A.4. Contact between two elastic bodies, general case, Hamrock and Dowson simplification	199
Table B.1. Viscosity-temperature equations	206

Table B.2. Dimensionless groups (H,U,W,G,k) used to defined the film thickness formulae valid for isothermal, fully flooded conjunctions (Stachowiak and Batchelor 2001)	216
Table B.3. Film thickness formulae in EHD regime valid for isothermal, fully flooded conjunctions (Stachowiak and Batchelor 2001)	217
Table B.4. Relationship between dimensionless groups introduced by Moes and dimensionless groups of Hamrock and Dowson.....	217
Table B.5. Central film thickness (Nijenbanning et al. 1994)	217
Table B.6. Average and maximum flash temperature formulae for circular contacts	228

List of Equations

Equation (3-1). Archard adhesive wear law for dry contacts.....	29
Equation (3-2). Fractional film defect β	33
Equation (3-3). Volume of adhesive wear in lubricated contacts	33
Equation (3-4). Velocities of a point P within the pin	36
Equation (3-5). Velocities of a point P within the pin in polar coordinates.....	37
Equation (3-6). First definition of the probability of asperity contact	37
Equation (3-7). Friction forces F_{Rx} , F_{Ry}	38
Equation (3-8). Friction forces F_{Rx} , F_{Ry} according to α	38
Equation (3-9). Differential of friction force	39
Equation (3-10). The differential torque	39
Equation (3-12). Field of velocities (odd function)	39
Equation (3-13). A bar under a bending couple.....	54
Equation (3-14). Moment created by the distribution of σ_z	54
Equation (3-15). Determination of k_1	55
Equation (3-16). Contact pressure σ_z (compression and bending)	56
Equation (3-17). k_0, k_1 values in the contact pressure expression σ_z	56
Equation (3-18). Moment created by the distribution of σ_z	56
Equation (3-19). Condition of contacting surfaces separation.....	56
Equation (3-20). Stress distribution on the counterface of the pin	57
Equation (2-21). Friction forces and friction pin torque for the proposed model.....	57
Equation (3-22). Probability of asperity contact for the proposed model.....	58
Equation (3-23). Conductance according to (Greenwood and Williamson 1966)....	59
Equation (4-1). Shear rate or velocity gradient.....	77

Equation (4-2). Shear stress 77

Equation (4-3). Viscosity 77

Equation (4-4). Newtonian fluids (constant μ) 78

Equation (4-5). Non-Newtonian fluids. Power Law Fluids (PLF)..... 79

Equation (4-6). Non-Newtonian fluids. Herschel-Bulkley 79

Equation (4-7). Casson model..... 79

Equation (4-8). Viscosity relationship for the rod-type viscometer..... 83

Equation (4-9). Differential voltage in a bridge with voltage excitation 97

Equation (4-10). Null type impedance bridge..... 97

Equation (4-11). Differential voltage reading for a bridge circuit as reported in
(Allen 2006) 98

Equation (4-12). Bridge with the resonance condition in one of the arms 98

Equation (4-13). Bridge with a relaxed impedance condition 99

Equation (5-1). Gravitational force acting on the particle 138

Equation (5-2). The buoyant force acting on the particle 139

Equation (5-3). The drag force acting on the particle (Stokes region)..... 139

Equation (5-4). The drag force acting on the particle 139

Equation (5-5). The drag force acting on a spherical particle..... 139

Equation (5-6). Differential equation representing the motion of a particle through a
fluid 140

Equation (5-7). Differential equation representing the motion of a particle through
a fluid (General case) 140

Equation (5-8). Motion of a particle (Stokes region)..... 141

Equation (5-9). Position of the particle (Stokes region) 141

Equation (5-10). Velocity of the particle (Stokes region)..... 142

Equation (5-11). Velocity of the particle (Stokes region)..... 144

Equation (5-12). Velocity of the particle (Stokes region)..... 145

Equation (5-13). Drag coefficient for non-spherical particles
(Haider and Levenspiel 1989)..... 145

Equation (5-14). Drag coefficient for non-spherical particles (Ganser) 146

Equation (5-15). Dimensionless number to determine C_D from the particle
diameter D_p 149

Equation (5-16). Dimensionless number to determine C_D from the terminal
velocity V_T 149

Equation (5-17). Dimensionless number to determine C_D from the particle
diameter D_p (logarithm scale)..... 152

Equation (5-18). Elapsed time to settle down for each particle diameter 154

Equation (5-19). Overdetermined system of equations in matrix form 156

Equation (5-20). Notation for particle density samples over time 157

Equation (5-21). Solution to the problem in the least mean squares sense..... 158

Equation (5-22). Centrifugal acceleration..... 166

Equation (5-23). Terminal velocity of the particle (Stokes region)..... 166

Equation (5-24). Velocity of the particle (Stokes region)..... 167

Equation (A-1). Reduced radius of curvature of two elastic bodies with convex
surfaces in contacts 191

Equation (A-2). Reduced radius of curvature along ‘x’ and ‘y’ axis..... 191

Equation (A-3). Convection to locate ‘x’ and ‘y’ coordinates..... 191

Equation (A-4). Reduced radius of curvature of two elastic bodies one convex
and one flat surface in contact..... 192

Equation (A-5). Reduced radius of curvature of two elastic bodies one convex and
one concave surface in contact..... 193

Equation (A-6). Reduced radius of curvature of two spheres in contact 194

Equation (A-7). Reduced Young’s modulus of two bodies 195

Equation (A-8). Reduced radius of curvature between a sphere and a plane 196

Equation (A-9). Reference distance to calculate α201

Equation (A-10). Angle between the upper ball and the lower balls201

Equation (A-11). Angle between the upper ball and the lower balls when $h=0$ 201

Equation (A-12). Reference distance to calculate the sliding speed.....201

Equation (A-13). Sliding speed in the contact between the upper and one of the lower balls202

Equation (A-14). Friction Force Four-Ball Sliding Test.....202

Equation (A-15). Friction Torque in Four-Ball Sliding Test configuration.....202

Equation (A-16). Friction Torque T and the force F_c measured with the load cell 202

Equation (A-17). Coefficient of friction for 4 Ball Sliding Test Configuration203

Equation (A-18). The linear velocity, V , at the centre of the pin O_p 203

Equation (A-19). Coefficient of friction Pin and Ball-On-Disc.....203

Equation (B-1). Walther’s equation207

Equation (B-2). Viscosity index.....208

Equation (B-3). Barus Law208

Equation (B-4). Hamrock and Dowson central film thickness213

Equation (B-5). Hamrock and Dowson minimum film thickness.....213

Equation (B-6). Definition of the dimensionless film parameter λ218

Equation (B-7). Surface contact temperature.....223

Equation (B-8). Peclet number.....224

Equation (B-9). True flash temperature rise230

Nomenclature

V	The predicted volume of adhesive wear
L	The sliding distance or the load applied to the pin
k	The wear coefficient for non welded junctions
W	Applied load
H	Hardness of the worn surface
Ψ	Plasticity index
δ_p	Plastic deformation
W	Dimensionless Load Parameter
G	Dimensionless Material Parameter
U	Dimensionless Speed Parameter
K	Dimensionless Ellipticity Parameter
M	Dimensionless parameter (Moes expressions)
L	Dimensionless parameter (Moes expressions)
h_0	Minimum film thickness
h_C	Central film thickness
β	Fractional Film Defect
k_e	The wear coefficient for non welded junctions
k_p	The wear coefficient for welded junctions
A_r	The real area of contact
A_e	The area of contact for elastic asperity contacts
A_p	The area of contact for plastic asperity contacts
R	The distance from the centre of the pin O_p to the centre of the disc O_d
r	The distance of a point P on the pin from its centre, O_p
N	Normal Force
F_R	Resultant Friction Force
F	Force acting on the pin

τ	Interfacial Shear Stress
ω_d	The angular velocity of the disc about its centre, O_d
ω_p	The angular velocity of the pin about its centre, O_p
α	The ratio between angular velocity of the pin and angular velocity of the disc
r_p	Pin radius
T_F	Friction torque required to rotate the pin
T	Torque required to rotate the pin
ϕ	Dimensionless number defined as $2 \cdot \alpha \cdot r_p / R$
H	Hardness of the material
σ	Stress
M	Moment distribution
ρ	Resistivity of the material
G	Conductance
k_0, k_1	Constants for the stress distribution
h	Height of the pin
Θ_R	Angle between the x-axis and the resultant friction force
$\dot{\gamma}$	Shear rate
η	Viscosity
e	Error
e_h	Horizontal error
e_v	Vertical error
a	Radius of the contact area or semimajor axis of the contact ellipse
R_x	Reduced radius of curvature in the 'x' direction
R_y	Reduced radius of curvature in the 'y' direction
R_{ax}	Reduced radius of curvature of body 'A' in the 'x' direction
R_{ay}	Reduced radius of curvature of body 'A' in the 'y' direction

R_{bx}	Reduced radius of curvature of body 'B' in the 'x' direction
R_{by}	Reduced radius of curvature of body 'B' in the 'y' direction
ν_A, ν_B	Poisson's ratios of the contacting bodies 'A' and 'B' respectively
E_A, E_B	Young's moduli of the contacting bodies 'A' and 'B' respectively
E'	The reduced Young's modulus
R'	The reduced radius of curvature
P	The contact pressure (Hertzian stress)
W	The normal load
k_0, k_1, k_2	Contact coefficients, elliptical contacts
ϕ	Angle between the plane containing the minimum principal radius of curvature of body 'A' and the plane containing the minimum principal radius of curvature of body 'B'
$\bar{\varepsilon}, \bar{\xi}$	Simplified elliptic integrals
k	Contact coefficients, elliptical contacts
\bar{k}	Simplified ellipticity parameter
η_p	The lubricant viscosity a pressure 'p' and temperature 'θ'
η_o	The viscosity at atmospheric pressure and temperature 'θ'
α	The pressure-viscosity coefficient
h_c	The central film thickness
h_o	The minimum film thickness
U	U is the entraining surface velocity, i.e. $U = (U_A + U_B)/2$
W	The contact load
k	The ellipticity parameter $k = a/b$
W	Dimensionless load parameter
G	Dimensionless materials parameter

U	Dimensionless speed parameter
H	Dimensionless film parameter
g_e, g_v	Dimensionless groups used for the lubricant regime determination
λ	Dimensionless film parameter λ
$R_{q,a}, R_{q,b}$	Root mean square surface finishes of solids a and b .
T_c	Maximum surface contact temperature
T_b	Bulk temperature of the contacting solids before entering the contact
$T_{f\max}$	The maximum flash temperature
L	Peclet number
χ	Thermal diffusivity [m ² /s], i.e. $\chi = K/\rho\sigma$
K	Thermal conductivity
ρ	Density
σ	Specific heat
p_y	Flow or yield stress of the material
β	Fractional film defect
k_e	The wear factor characteristic of non-welded junctions
k_p	The wear factor characteristic for welded junctions
A_e	Elastic area of contact
A_p	Plastic area of contact
μ	Coefficient of friction
h	Distance separating the balls. Pure sliding $h=0$
d	Diameter of each ball
r	Radius of each ball.
g	Reference distance to calculate α
α	Angle between the upper ball and the lower balls

e	Reference distance to calculate the sliding speed
v	Sliding speed
T	Friction Torque
F'	Friction Force
ω_d	Angular velocity of the disc
ω_p	Angular velocity of the pin
V	Linear velocity
O_p	Centre of the pin
O_d	Centre of the disc
R	Distance from O_d to O_p

Abbreviations and Acronyms

ASTM	American Society for Testing and Materials
BL	Boundary Lubrication
CBM	Condition Based Maintenance
CBPM	Condition Based Predictive Maintenance
CFA	Current Feedback Amplifier
DDS	Digital Direct Synthesis
ECR	Electrical contact resistance
EIS	Electrochemical Impedance Spectroscopy
FBRM	Focused Beam Reflectance Method
FT-IR	Fourier Transform Infrared Spectroscopy
EHL	Full Film Lubrication (Elastohydrodynamic)
HL	Full Film Lubrication (Hydrodynamic)
GIC	General Impedance Converter
IS	Impedance Spectroscopy
KF	Karl Fisher
LDPE	Low-density polyethylene
λ	Lubrication Regime Parameter
MRI	Magnetic Resonance Imaging
ML	Mixed Lubrication
NCC	Negative Capacitive Circuits
NMR	Nuclear Magnetic Resonance
ODM	Oil Debris Monitoring
PQ	Particle Quantity
PSD	Particle Size Distribution
PLL	Phase Locked Loop

PMMA	Polymethyl metacrylate
PP	Polypropylene
PLF	Power Law Fluid
PdM	Predictive maintenance
PM	Preventive Maintenance
PAC	Probability of Asperity Contact
PPEC	Probability of Elastic Asperity Contact
PPAC	Probability of Plastic Asperity Contact
PID	Proportional–Integral–Derivative controller
PWM	Pulse Width Modulation
RCFA	Root Cause Failure Analysis
RMS	Root Mean Square
RNLI	Royal National Lifeboat Institution
RTF	Run to Failure Maintenance
SAW	Surface Acoustic Wave
TSM	Thickness Shear Mode
TAN	Total Acid Number
TBN	Total Base Number
TAN	Tribological Aspect Number
WSD	Wear Scar Diameter
YSF	Yield Stress Fluid

1 Introduction

This thesis is in line with the initial research proposal, which aims to seek further enhancements in the maintenance of the RNLI's hydraulic equipment required for launching and recovering lifeboats. The current maintenance plan consists of replacing the oil on a regular basis without taking into account whether the lubrication oil is still in an operating condition. In some cases, lubricant samples were taken for further off-line analysis but at times, results obtained from laboratory analysis confirmed they were non-representative of the status of health of the machine. Therefore, this type of maintenance is expensive and not effective for failure prediction. As a result, the RNLI requires the identification of a non-invasive condition monitoring approach suitable for their mode of operations and maintenance environment. This new maintenance strategy aims to achieve the goal of notifying in real time the condition of the oil and the onset of failures. This ideal solution will lead to a better utilisation of the lubrication oil with the associated cut down costs and the potential of preventing excessive wear rates in the RNLI's launch and recovery lifeboat equipment. It should be highlighted that the launch and recovery systems are exposed to a marine environment and they work under "start-stop" conditions with a very low utilisation factor (RNLI has 235 lifeboat stations and an average of 9000 launches per annum which results on average 40 launches per station and annum (Boyce 2010)). Under such conditions, the wear rate is much higher than the expected wear in normal operation. Although this abnormal wear rate can be assumed in many scenarios, from the RNLI's viewpoint whose main concern is to save lives at sea, all these small factors must be considered in order to achieve the highest success ratio on the launch and recovery of lifeboats, resulting in rescue systems operative all time avoiding undesirable and catastrophic situations.

Therefore, at the beginning of this research the initial goal was the assessment and characterisation of commercial oil condition sensor technology by means of the simulation of the typical failure processes that occur in real machinery. Aiming to simulate the widest range of failures optimising the costs (in terms of time and money), it was decided to use lubrication theory to select the operation conditions of experiments performed under standard bench test configurations. The main reason

for approaching the failure generation problem under this perspective is because machine components are designed to operate under certain operating conditions and, indirectly under a defined range of lubrication regimes. The lubrication regime is defined as the ratio of the squeeze film (oil film) thickness to the surface roughness. Therefore, machine failures and accelerated wear rates in lubricated systems are related to machine operation outside the range of lubrication regimes considered during the design stage.

During normal machine operation small wear debris particles of the order of sub-micrometer to 10 microns are generated. When abnormal wear begins, large debris particles in the range of 10 to 150 microns are produced. The particle size and concentration will increase gradually until machine failure, Figure 1.1. Besides this, during prolonged or extreme operating conditions several forms of contamination and additive depletion can begin to degrade the physical properties of the lubricant to an unsatisfactory level. From a theoretical approach, changes of the size and shape of debris (indirectly related to different wear mechanisms) and changes in oil quality are related to transitions of lubrication regimes (λ). Therefore, failures could be simulated from a theoretical approach if the lubrication regime could be controlled under experimental conditions, Figure 1.2.

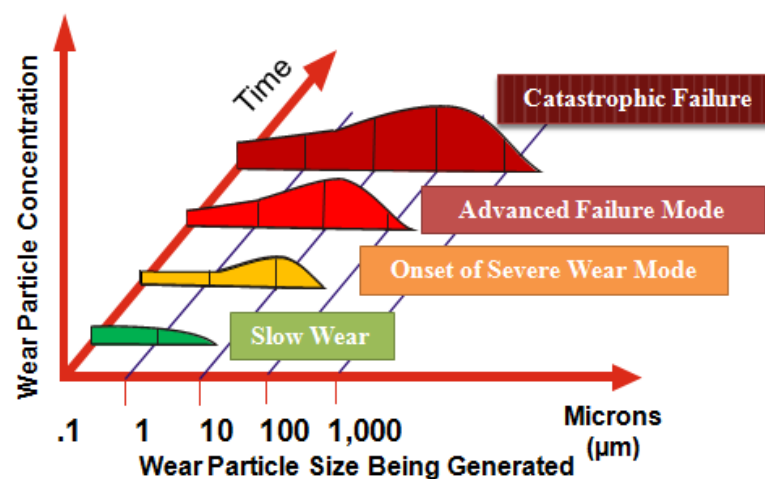
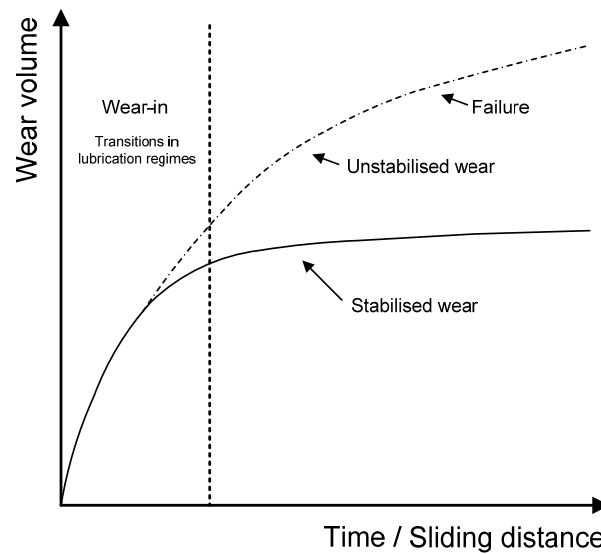
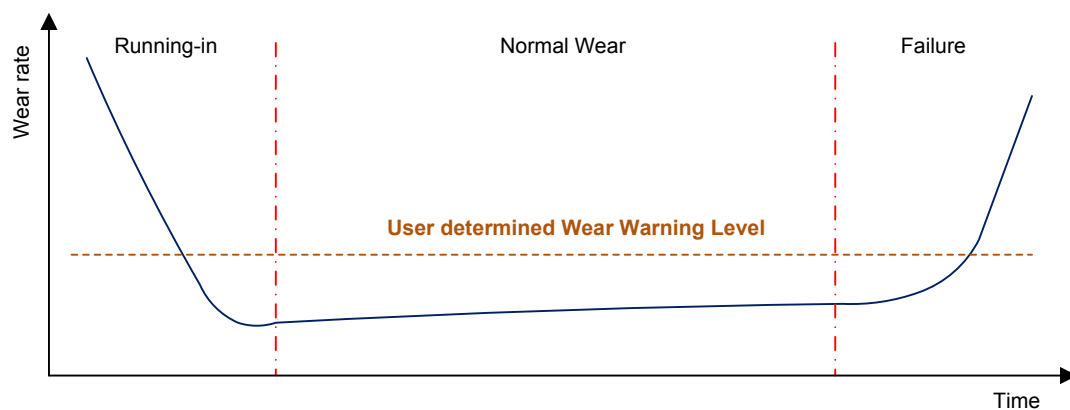


Figure 1.1. Machine failures using the particle size distribution
(Baur 1982; Lockwood and Dalley 1992)



a) Wear volume versus sliding distance → total amount of wear in the system



b) Bath curve (wear rate → debris density approach)
(Kent 2009; Roylance and Hunt 1999)

Figure 1.2. Progress of wear.

Although it seems reasonable from a theoretical perspective to simulate failures by means of controlling the lubrication regime, there is a long way to go in research in order to characterise the sensors in such way. Throughout this research, two fundamental limitations were found that make it impossible to assess oil condition sensors. The first one is the inability to control the lubrication regime using standard bench test configurations and the second one is related to the

technical limitations of current oil condition monitoring sensors. The first difficulty arises because the lubrication regime is a very complex time-dependant problem and hence, it depends on many factors; some of them impossible to control and very difficult to model, such as the additive behaviour under boundary and mixed lubrication regimes that is responsible for the tribofilm formation process and the changes of surface statistical parameters (such as roughness and surface texture orientations) over time. The second limitation appears because current sensor technology is not sufficiently sensitive to detect the onset of failures. As the sensor sensibility to measure the characteristic parameters of the lubrication oil increases, the detection time for failures or accelerated wear rates decreases.

In an attempt to solve the initial problem, this thesis aims to explain the physical operation and the limitations of commercial technology under controlled experimental conditions. The limitations are explained and several enhancements are proposed for future sensor development.

The type of tribological tests used for the assessment of the oil condition sensors were chosen considering the most frequent type of motion and contact geometry of tribological pairs that appear in hydraulic systems. As a result, this research is focused on sliding contacts as they are the most common type of motion in hydraulic machinery as can be seen from Table 1.1. There are several standard tests for sliding contacts depending on the contact geometry. Considering the lab facilities, point contacts were evaluated under ball-on-disc and four-ball sliding wear test configurations while area contacts were assessed under pin-on-disc. The type of test and experimental conditions were constrained by the sensor specifications.

Although accelerated wear testing does not exactly replicate real system behaviour, and failures cannot be accurately controlled, it gives the capability of a better understanding of how these sensors will behave in real machinery. Finally, this research was carried out under a multidisciplinary approach with the aim of achieving a significant improvement in actual trends of oil condition based maintenance.

	Type of motion			Contact geometry			Operational sequence of motions		
	Rolling	Sliding	Impacting	Point contact	Line contact	Area contact	Continuous	Oscillating	
Vane pumps/motors									
Vane-ring	(E)		(E)		(E)		(E)	(E)	
Vane-slot		(E)			(E)			(E)	
Vane-sides		(E)				(E)	(E)		
Gear pumps/motors									
Gear wheel-sides		(E)				(E)	(E)		
Gear wheel-work	(E)				(E)			(E)	
Piston pumps/motors									
Piston pin/ bushing (A/R)		(E)				(E)		(E)	
Sliding block / plate (A)		(E)				(E)	(E)		
Sliding block / plate (R)		(E)				(E)	(E)		
Sliding block / piston head (A)		(E)				(E)			
Sliding block / piston head (R)		(E)				(E)			

Valve plate / piston barrel (A)		(E)						(E)		
Support for Captive C-ring (A)	(E)						(E)			(E)
Axial distributor (R)		(E)						(E)		(E)
Rotary slide valve		(E)						(E)		(E)
Retainer plate- guide ball (A)		(E)					(E)			(E)
Sliding block-retainer plate (A)		(E)						(E)		(E)
Cylinders										
Piston – cylinder tube		(E)						(E)		(E)
Rod seal		(E)						(E)		(E)
Others										
Roller bearings	(E)							(E)		(E)
Shaft-radial		(E)						(E)		(E)
Shaft sealing ring										
Serrated shaft							(E)			(E)
Plain bearing		(E)						(E)		(E)

Table 1.1. Tribological stresses in hydraulic machinery (Totten 1999)
 (A) Axial piston machine (R) Radial piston machine (E) extremely important with respect to failure behaviour

1.1 Research Objectives and Scope

The aim of this research is to provide a detailed study of oil condition sensor technology linked with lubrication theory and wear modelling. This understanding assists the RNLI to improve oil replacement maintenance schedules associated with launch and recovery of marine equipment.

The main objectives set out in this thesis are:

- 1) Analyse the commercial oil condition monitoring technology and the products suitable to introduce a predictive maintenance plan for the RNLI hydraulic equipment.
- 2) Understand and explain the underlying physical principles of commercial sensor technology and their design limitations.
- 3) Contribute to the characterisation of ODM and oil quality sensors under controlled experimental conditions using wear prediction models and wear measurement techniques.
- 4) Design and develop a specific test rig to perform tribology tests using oil condition monitoring sensors highlighting the problems associated with controlled generation of failures under accelerated wear test configurations.
- 5) Contribute to new sensor technology knowledge to achieve a more effective indication of the oil condition and progression of wear within machinery.

1.2 Outline of the Report

This report is presented in 5 chapters. First chapter sets the project aim and goals. The second chapter presents a review of wear in lubricated contacts and discusses the reliability of bench tests under boundary or mixed lubrication regimes. Chapter 3 presents a model for the prediction of adhesive wear in lubricated contacts which is built upon the probabilistic approach of wear prediction. Chapter 4 describes the typical maintenance techniques, introduces the operational principles of commercial oil condition monitoring sensors and discusses their limitations. Chapter 5 contains a new methodology for the determination of the particle size distribution from ferrous debris density measurements. Chapter 6 sets out discussions, conclusions and recommendations for future work.

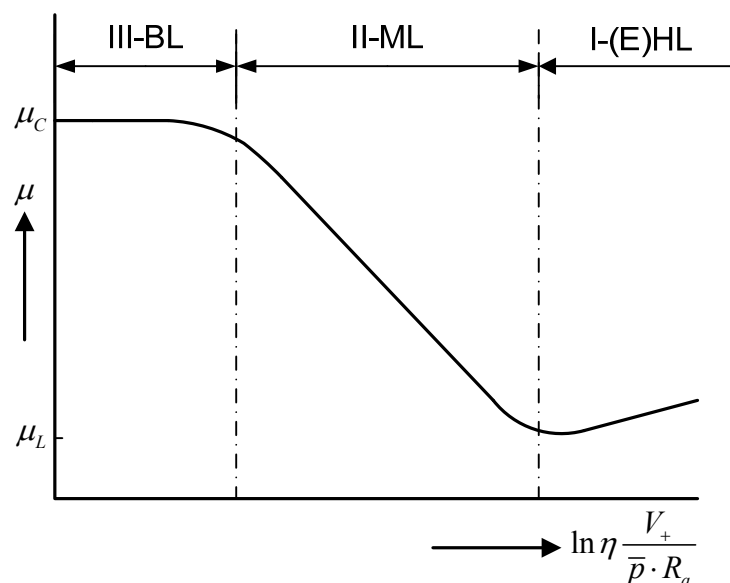
Finally, a summary of the theory used in this research is covered in the appendices.

2 Theoretical background

2.1 Review of lubrication theory

Improvements in maintenance of hydraulic systems are clearly of great interest to industry. For most hydraulic machinery the quality of the oil must be ensured for reliability (Totten 2006; Totten et al. 2001).

Machine components operate under different lubrication regimes. A good classification of the different regimes was found by Stribeck when he plotted the coefficient friction between two rough surfaces in relative motion against velocity. Stribeck presented a paper in 1902 describing the influence of the velocity of the interacting surfaces and the load on the coefficient of friction for plain journal bearings and for roller bearings (Stribeck 1902). A generalised Stribeck curve is obtained when the coefficient of friction is plotted against a lubrication number, Figure 2.1 (Schipper 1988).



I – Full film lubrication (E)HL

II – Mixed lubrication ML

III – Boundary lubrication BL

(η =dynamic inlet viscosity, V_+ = effective sum velocity, p =the average Hertz contact pressure, R_a combined roughness of the two surfaces)

Figure 2.1. Generalised Stribeck curve (Schipper 1988)

Three lubrication regimes can be distinguished in this curve, the boundary lubrication, the mixed lubrication and elastohydrodynamic lubrication regime. When the lubrication number is high the film thickness is greater than the surface roughness and no asperity contact occurs. The surfaces are fully separated by a fluid film that carries the entire load. As the lubrication number decreases asperity contact occurs and the mixed lubrication regime starts. The load is partly carried by the asperities and the remaining part by the fluid. If the lubrication number further decreases more asperity contact increases until the load is carried by the asperities. This is defined as boundary lubrication regime.

In industrial applications power and motion are transmitted in most cases through either lubricated or dry surface contacts between mechanical components. The type of contacts and lubrication characteristics define the performance and life of those components. In Tribology, lubrication models are used for better understanding of contact mechanisms, to predict lubrication characteristics and to improve component design, performance and reliability. These models are used to provide the following parameters: pressure due to surface contact and pressure due to hydrodynamic action, lubricant film thickness or gap, surface deformation and contact stiffness, friction/traction force, heat generation and surface temperature rise in the contact area, subsurface stresses. In lubricated contacts, film thickness is the most important parameter when designing tribological systems (Zhu and Hu 2001).

Due to increasing technical demands (more power through smaller volumes, i.e. the continuous optimisation of Formula 1 gearboxes) many applications involving non-conformal contacts transmit forces through a thin film of lubricant, which separates the two solid mating components. This thin film under severe operating conditions is not sufficient to completely separate the surfaces. At this point, the influence of surface roughness in lubrication is important as the film thickness is usually in the same order of magnitude as surface roughness. This is the regime of mixed elastohydrodynamic lubrication in which the applied load is shared by the asperities and the lubricant film. There are many factors to consider in this regime, such as the topography of the surfaces, the temperature distribution within the contact, the load sharing between the fluid and the asperities. As an example, surfaces may have different roughness features depending on the machining process.

Surface textures orientation can be of a longitudinal, transverse or oblique orientation. The effects of surface texture in EHL contacts has been studied by many researchers (Ai and Cheng 1996; Venner and Napel 1992). All these factors complicate the prediction of the friction under this regime.

The performance of components operating in a mixed EHL regime is of significant importance because asperity contacts give rise to high local pressures, which can reduce the fatigue life of components. At the same time, the asperity contact friction and lubricant shearing in mixed lubricated contacts generate heat, which results in extreme local temperature rise (Hu et al. 1999).

The conventional lubrication theory has been developed since the second half of the last century. The first attempt at solving analytically the Reynolds equation under EHL conditions is the Ertel solution (Ertel 1939), published by Gubrin in 1949 (Gubrin 1949). Petrusевич in 1951 (Petrusevich 1951) determined the first numerical solution and predicted the constriction near the outlet and the pressure spike. Dowson and Higginson (Dowson and Higginson 1966b) contributed to the numerical solution of line contacts in 1966, their expressions are widely used today which highlights the remarkable importance of their results. For elliptical contacts the work of Hamrock and Dowson (Hamrock 1984) was the first landmark. Hamrock and Jacobson, Chittenden et al. (Chittenden et al. 1985) also developed general predictive film thickness formulae, each having its own region of validity. The complete film thickness map was firstly obtained by Moes for line contacts using the numerical results obtained from the multigrid solver developed by Lubrecht and Venner (Venner 1991). Nijebanning et al. (Nijebanning et al. 1994) generalised the complete film thickness map for elliptical contacts. Once the determination of the central film thickness was established, the focus shifted towards surface roughness.

In an attempt to model the real rough surface contact, two approaches have been employed to establish the relationship between the average surface gap and normal applied load. The first approach is based on stochastic models developed by (Christensen 1969; Tonder 1977; Zhu and Cheng 1988). The probabilistic approach to wear prediction developed by Stolarski (Stolarski 1980, 1981a, 1981b, 1989b, 1990b, 1990c) is an important contribution within the context of a stochastic approach. Although the latest trends prefer to tackle real friction problems using the

second approach, the wear prediction based on probabilistic models is a fast and useful tool for the estimation of the average evolution of contact pressures and surface statistics within the contact, the wear volume and the friction coefficient. The accuracy of prediction is subject to the matching ratio of the friction model to the experimental data. There are not many probabilistic models available and most of them are developed considering the classical adhesive wear models and the theory of elastic and plastic asperity contact developed by Greenwood and Williamson (Greenwood and Williamson 1966). Within this scope, the next chapter introduces a novel model of friction for a generalised pin-on-disc setup that follows the statistical approach. It should be noted the probabilistic approach does not provide the detailed information that is critical for the study of lubrication breakdown and surface failure mechanism and it does not provide information about local pressure peaks, local film thickness fluctuations and asperity deformation.

The second approach provides detailed information; it is classified by deterministic models. This type of analysis predicts local pressures and film thickness fluctuations within the contacting surfaces. It is based on solving a multiphysics problem which in its simple form requires solving the Reynolds equation, the film thickness equation and the load balance equation for smooth contacting surfaces. Most recent solvers can include thermal effects, the compressibility of lubricants, starved lubrication conditions, rough surfaces, the Non-Newtonian behaviour of the lubricants and solve all the equations as a time-dependent problem making possible the simulation of the running-in period of different types of contacts. Although this way of tackling the problem is very promising, many aspects need to be improved to use this approach as an effective simulation tool for friction and wear prediction in real conditions. Implementation of real problems faces challenges such as computing power and solver convergence issues that make difficult the validation process of the models. However, this approach is useful to understand and solve some mixed lubrication problems and it seems to be a useful engineering tool to optimize component design, failure analysis and to improve the life of machine components. Most of these numerical techniques also calculate the influence of temperature rise within the contact (Deolalikar et al. 2008; Hu and Zhu 2000; Zhu and Hu 2001).

2.2 Review of flash temperature in sliding contacts

The prediction of the surface temperature at a sliding contact interface has great importance in Tribology. The temperature increase due to friction affects the viscosity of the lubricant, the materials in the tribological contact, the reactivity of the additives of the lubricant, the oxidation process of the surface and the geometry of the contact. As viscosity changes with temperature it also affects the lubrication regime (Dowson 1998; Stachowiak and Batchelor 2001).

The first study investigating frictional heating and contact temperatures was carried out by Blok (Blok 1937) who determined the maximum temperature rise for a semi-infinite body subject to a uniform square heat source at high Peclet numbers. Blok studied the quasi steady state temperature rise due to band-shape heat sources for high Peclet numbers. The Peclet number is a measure for the relative importance of convection and lateral conduction. Blok's models did not deal with situations where lateral conduction becomes a heat transfer mechanism.

Jaeger (Jaeger 1942) and Carslaw and Jaeger (Carslaw and Jaeger 1959) extended Blok's model taking into account situations where lateral conduction becomes an important heat transfer mechanism. Jaeger formalised the mathematical models for rectangular heat sources.

Archard (Archard 1959) introduced an approximation of circular contacts using the Jaeger solution for rectangular contacts. It is important to indicate most practical contacts are elliptic. Approximate solutions for elliptical contacts were introduced by (Kuhlmann-Wilsdorf 1986, 1987). Tian and Kennedy presented analytical and approximate solutions of maximum and average surface temperatures for the entire range of Peclet numbers each case of different shaped heat sources (Tian and Kennedy 1994).

Practical applications involve rough contacts, much effort has been expended trying to predict surface temperature rise in rough surfaces. In computational terms, the solution of this problem is very time-consuming. The solution to this problem allows the characterisation of the temperature rise of the solid to solid contact that might happen in the elastohydrodynamic lubrication (Zhao and Sadeghi 2001).

Different approaches to speed up the calculation of this problem has recently appeared (Deolalikar et al. 2008; Gao et al. 2000; Zhao and Sadeghi 2001).

2.3 Reliability and reproducibility of bench tests

The tribological performance of a machine is a complex problem as it depends on many factors. There are many well-classified ways to evaluate the tribological performance using bench tests. A specific bench test can be selected according to experience, standardised tests, or limitations in available equipment. The Tribological Aspect Number (TAN) was developed to assist the test designer in making the transition from the field problem to bench test (Anderson 2000). However, tests under real conditions are the most expensive and time consuming but the most effective in terms of results. The advantage of the bench tests is the capability of modifying each tribological system parameter such as velocity, load, temperature and configuration. In most cases, especially in lubricated systems, the conclusions drawn from bench testing do not match the real application results. In the literature, this topic is well-discussed comparing bench test results with real examples, a typical example is the reliability assessment of bench tests for performing lubricant performance comparisons in hydraulic pumps (Totten et al. 1997).

2.4 Bench test results analysis using lubrication theory

Lubricants interact with the contacting surfaces in a complex way, hence the characterisation of the lubrication regime considering the contact pressure, the velocity, the fluid temperature and the roughness of the contacting surfaces is insufficient for explaining the friction and wear behaviour of different lubricants. Although the λ parameter which is obtained from the aforementioned parameters is effective in describing the transaction of lubrication regimes and predicting the fatigue life of rolling elements, it is not so useful in sliding contacts. The main problems that λ parameter encounters are not only the difficulty in calculating the fluid film thickness accurately but also calculating the roughness of the contacting surfaces along the time. In rolling applications as the contact pressure is very high

most of the asperities are flattened and λ seemed to be a good indicator of the fatigue life (Skurka 1970).

An accurate lubrication model that describes the wear according to factors that can be known and controlled can relate conclusions drawn between bench test and real conditions. However, under some bench test conditions, for example mixed and boundary lubrication regimes there are factors that are not considered by models and conclusions could lead to wrong predictions.

The next section shows the variability of results that can be concluded by tests performed under a typical bench test configuration using the ASTM 4172D conditions. The aim of this example is to show that a rational way of extrapolating results using classical lubrication theory can lead in some cases to contradictory predictions when compared to real results. In some cases, the simplistic approach to problems using models which only consider a few parameters could lead to contradictory results. In Tribology, the final state is often the one described because not all the factors can be controlled under bench test conditions. As a result, this explains why the tendency of conclusions drawn from bench tests is sometimes non-representative of the real system behaviour.

2.4.1 A real example under ASTM D4172

The ASTM D4172 standard (2004) covers a procedure for making a preliminary evaluation of the antiwear properties of fluid lubricants in sliding contacts by means of the Four-Ball Wear Test Machine. This test method can be used to determine the relative wear preventive properties of lubricating fluids in sliding contacts under the prescribed test conditions. The antiwear properties of the lubricant fluids are measured through the wear scar diameters (dimensional measurements) on the three lower balls using a microscope. The ASTM D4172 measurement method can be improved by using the optical profilometry in order to measure wear volume instead of wear scar diameter measurements. The WSD method gives no information about the distribution of the wear within the worn surface and data can be misleading due to displaced or transferred material and/or irregular shape of the wear surfaces.

2.4.1.1 Experimental details

Two formulated hydraulic oils from different manufacturers were used. The main properties of the lubricants are listed in Table 2.1.

Table 2.1. Main properties of the lubricants (2011a, 2011b).

Oil properties	Hydraulic oils	
	Mobil DTE 13M	ESSO NUTO H46
ISO Viscosity Grade	32	46
Density (g/cm ³ @ 15°C), ASTM D1298	0.874	0.876
Viscosity (cSt @ 40°C), ASTM D445	32	46
Viscosity (cSt @ 75°C), ASTM D445	10.7	12.8
Viscosity (cSt @ 100°C), ASTM D445	6.10	6.70
Viscosity Index, ASTM D2270	141	104

All test-section components were cleaned ultrasonically with heptane for 10 minutes and dried with hot air before and after tests. Wear tests were performed on a TE92HS four-ball machine according to the ASTM D4172 standard conditions (speed: 1200 rpm, loads: 392 N / 147 N, testing time: 60 min, temperature: 75°C). Additional wear tests were performed for the same speed, load and temperature but different testing times (5, 15, 30 and 45 minutes). Three repetitions were made for each sample and test condition in order to estimate the deviations.

The balls used in this study are 12.7 mm diameter test balls with a roughness of $R_a < 0.020 \mu\text{m}$ manufactured from AISI 52100 steel with a hardness of 65 RC. Experiments were performed under boundary lubrication regime, according to the λ ratio (the ratio between film thickness and combined surface roughness) calculated using the Hamrock and Dowson formulae (Hamrock et al. 2004), Table 2.2.

Table 2.2. Film thickness and lubrication regime using the Hamrock and Dowson formulae.

Lubricant	Oil temperature (°C)	Load (N)	h_0 (nm)	h_c (nm)	$\lambda = \frac{h_0}{\sqrt{R_{q_1} + R_{q_2}}}$ *	Lubrication regime
ESSO NUTO H46	75	392	22	40	0.85	Boundary
Mobil DTE 13M	75	392	19	35	0.76	Boundary

(*) Average and RMS roughness of AISI 52100 steel balls, $R_a < 20 \text{ nm}$, $R_q < 18 \text{ nm}$
 Minimum (h_0) and central (h_c) film thicknesses were determined using the H&D formulae (Hamrock et al. 2004)

A white-light interferometer (Zygo New View 5000) was utilised for measuring the wear scar diameter and wear volume on the three lower balls. The interferometer has a vertical resolution up to 0.1 nm and a horizontal resolution up to 500 nm. Each lower ball was located in a special holder where the wear scar could be aligned and placed at an adequate working distance under the objective.

Special attention was paid to ensure the horizontal axes (x,y) of the apparatus were properly calibrated before performing any measurement. If the reference surface is a sphere, the horizontal errors should be kept as low as possible to have utmost accuracy when processing measurements. This is especially critical when measuring very small wear volumes. A detailed study of the horizontal calibration of an optical profilometer is found in Appendix D.

The wear volume and the wear scar diameter were measured after subtracting the reference sphere to the worn surface. The best reference sphere was obtained by removing the best fit sphere with a constrained radius of 6.35 mm. After removing the reference sphere, the resulting geometry is almost a flat surface where the worn areas can be easily identified. The wear volume and the transversal and longitudinal wear scar diameters were measured from this surface.

2.4.1.2 Results and discussion

The evolution of the wear scar diameter (WSD) versus the test times for both hydraulic oils is shown in Figure 2.2 for a load of 392 N and Figure 2.3 for a load of 147 N.

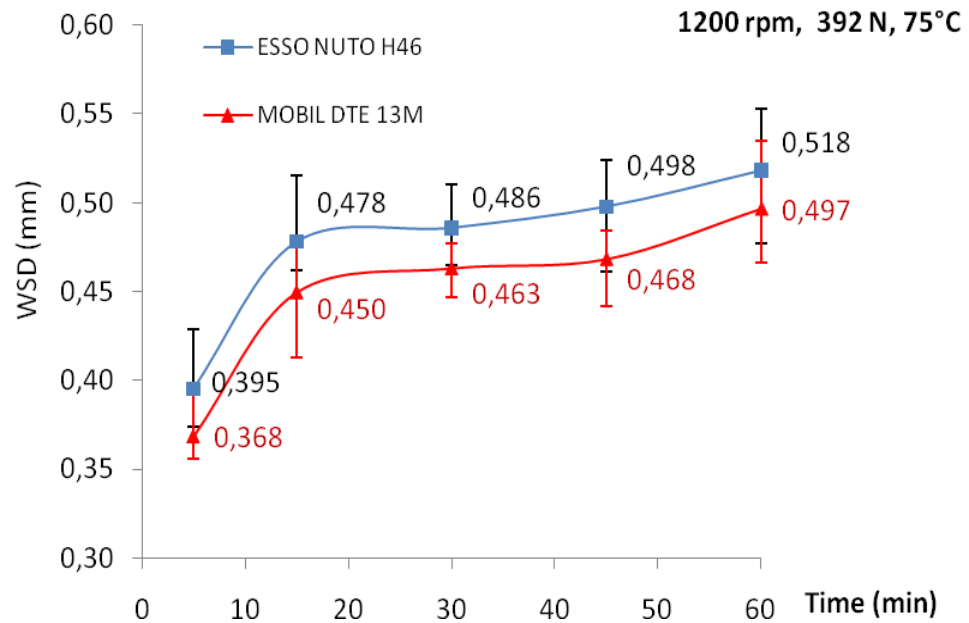


Figure 2.2. Wear scar diameter (WSD) versus testing time for both hydraulic oils (392N)

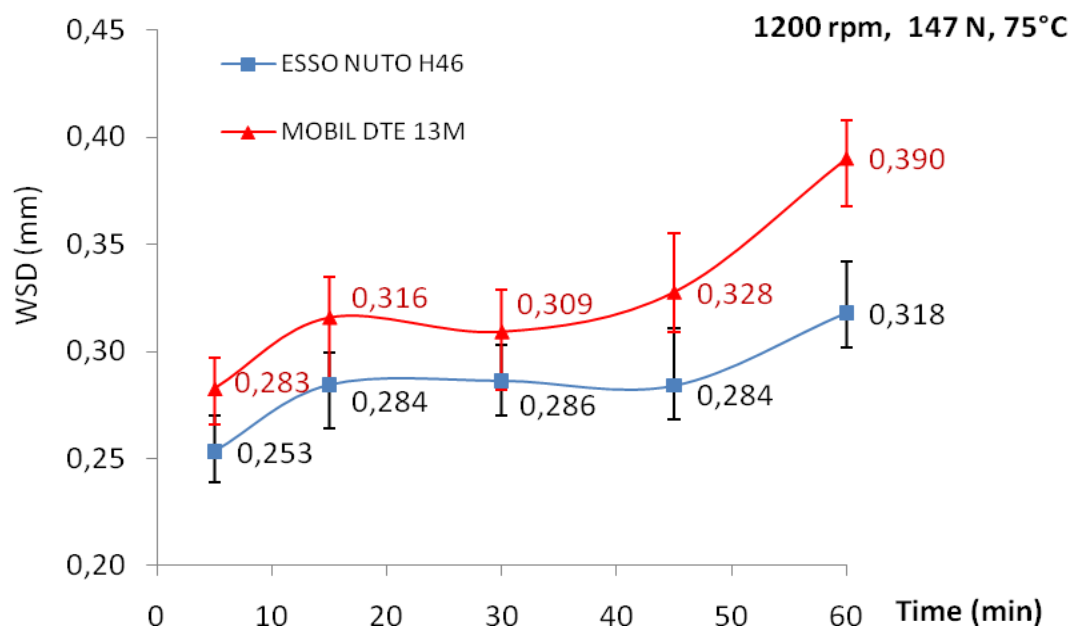


Figure 2.3. Wear scar diameter (WSD) versus testing time for both hydraulic oils (147N)

All measurements of WSD are larger than the theoretical Hertz diameter which is 0.2988 mm and 0.2148 mm for a load of 392 N and 147 N respectively. For the first 15 minutes wear increases sharply because of the running-in process. From 15

minutes onwards, the wear rate decreases, this behaviour is influenced by the action of the antiwear additives in the lubricants. The maximum and minimum differences among tests are shown with error bars, these deviations depend on the irregularities of the wear scars which affect the sample measurement process.

The wear volume and the maximum and minimum differences among tests are shown in Figure 2.4 for a load of 392 N and in Figure 2.5 for a load of 147 N.

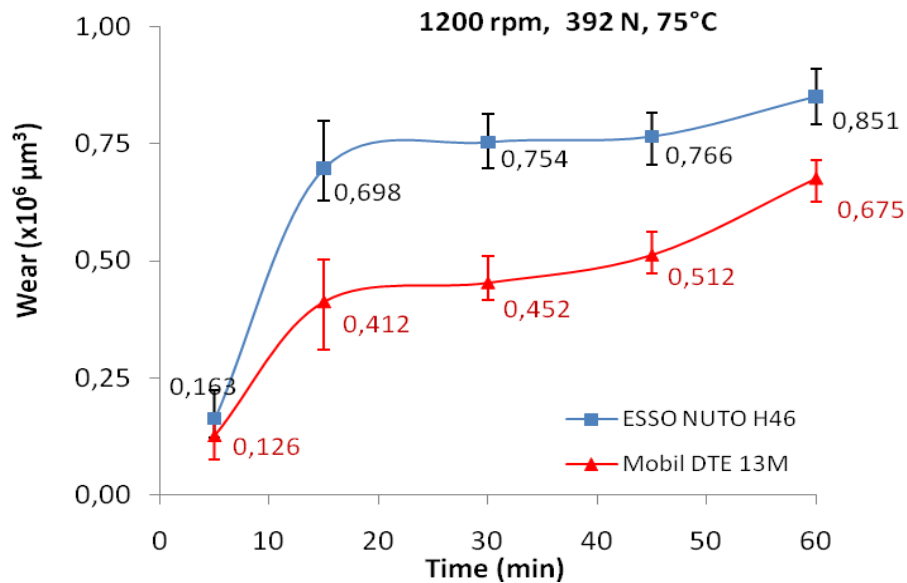


Figure 2.4. Wear volume versus testing time for both hydraulic oils (392N)

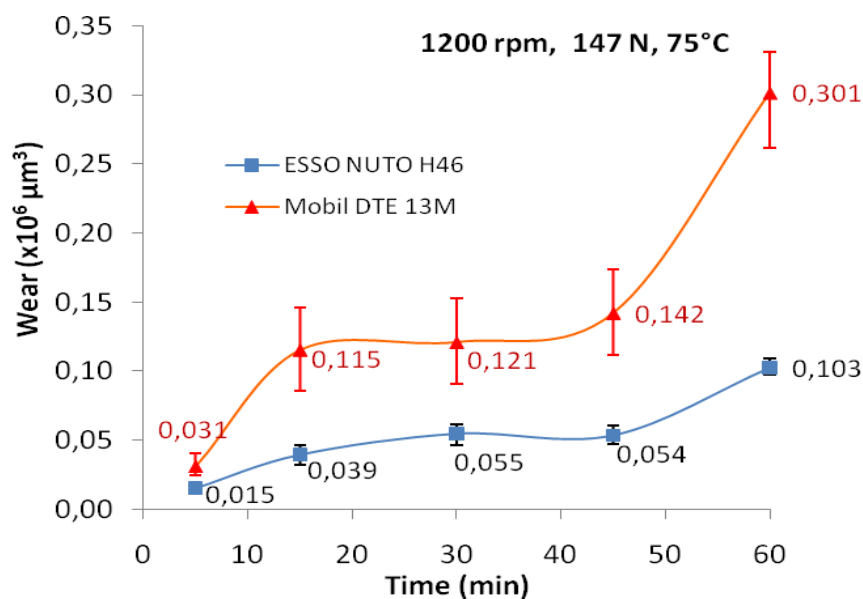


Figure 2.5. Wear volume versus testing time for both hydraulic oils (147N)

An important feature obtained from Figures 2.2, 2.3 when results are compared to wear volumes in Figure 2.4, 2.5 is the big overlap in WSD of both oils. This overlap does not occur in the wear volume measurements. This is an important argument for selecting the wear volume criteria instead of the WSD. An especial case could occur if only one experiment was run. This test could lead to different conclusions depending on the type of measurement criteria. As an example, considering the result of one replica of Figure 2.6, the WSD diameter for the 1 hour test could result in a WSD for the lowest viscosity oil larger than the WSD of the highest viscosity oil. According to the wear volume criteria, the largest viscosity oil always shows a higher wear volume than the low viscosity one. This situation leads to totally different results depending on the selected criteria, Figure 2.6.

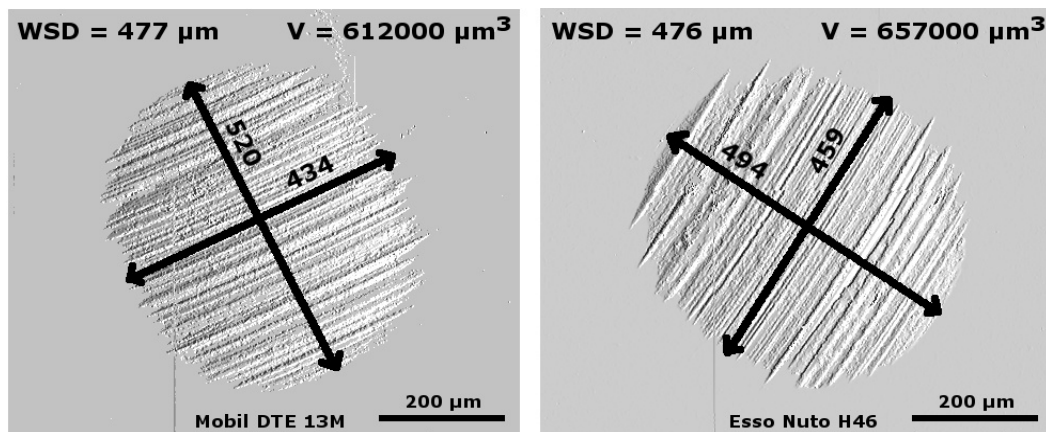


Figure 2.6. Results of one of the replicas under ASTM D4172 test conditions at 60 min (images from the optical profilometer, wear volume calculated by the proposed procedure).

The variations of wear volume measurements among tests for both hydraulic oils with similar testing times are larger than the differences in WSD, Figures 2.7, 2.8. This shows that the measurement of wear volume is more sensitive method than the WSD method. This fact is more important when the wear surface has an irregular shape which is typically found in experimental samples, Figures 2.9, 2.10. Considering that the purpose of the ASTM D4172 standard is the comparison of the antiwear properties among lubricants, the wear volume parameter would be a better choice. The variations in measurements using the wear volume criterion clearly decrease with time, whilst the variation in WSD is not so pronounced.

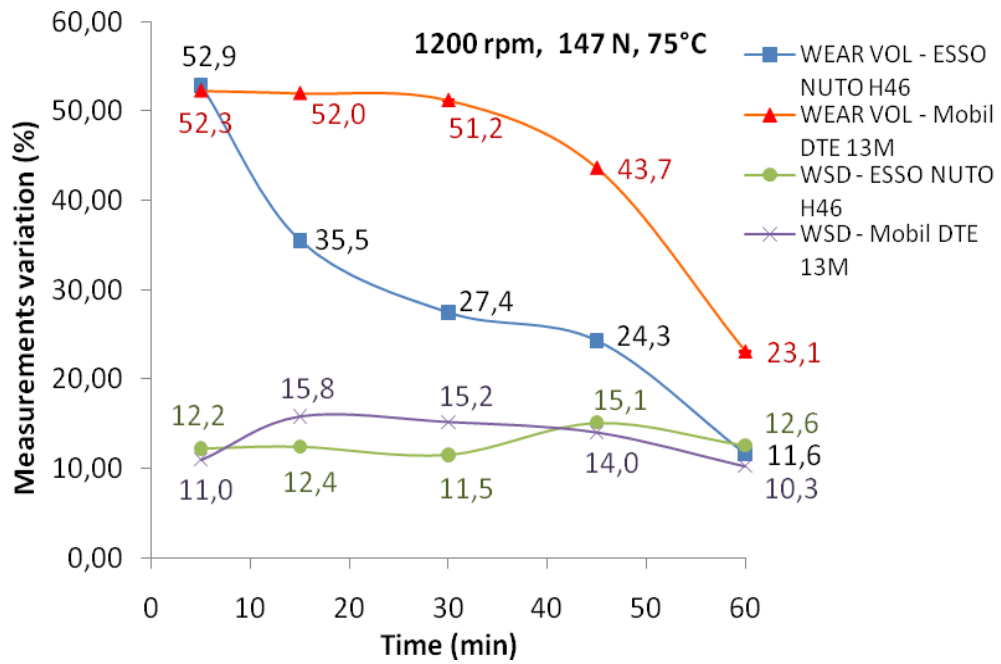


Figure 2.7. Percentage of variation of the WSD and wear volume between Mobil DTE 13M and Esso Nuto H46 samples

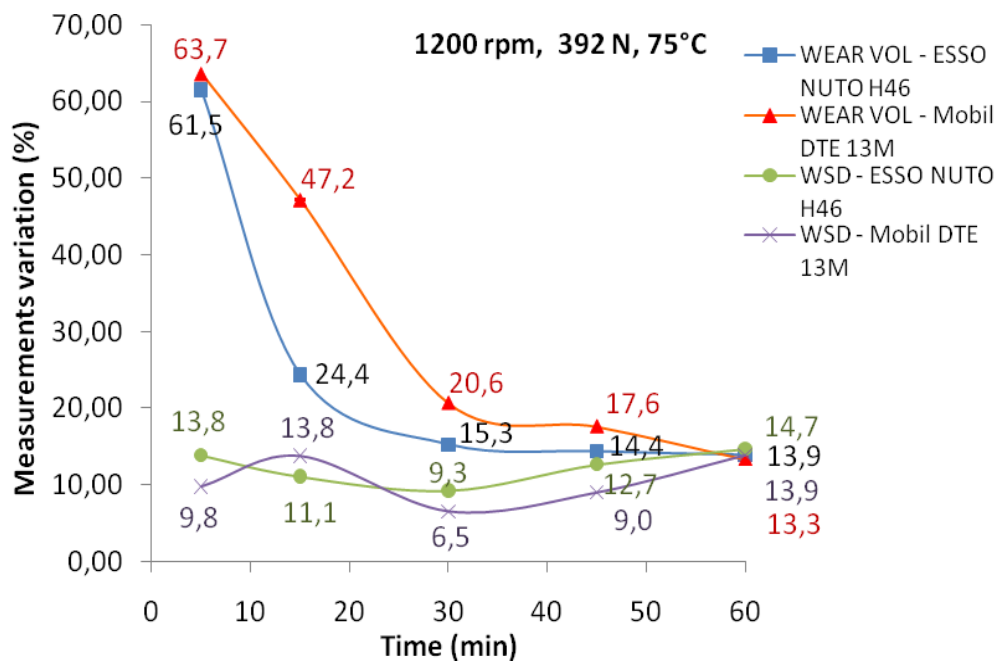
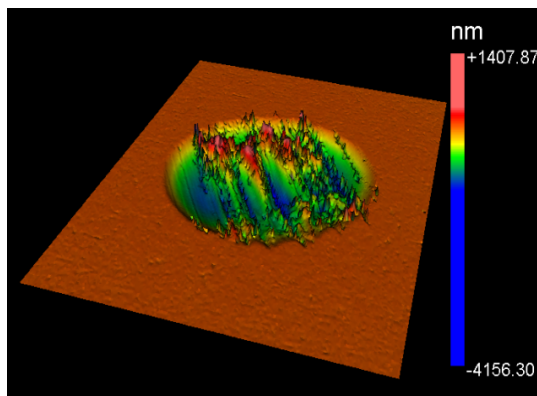
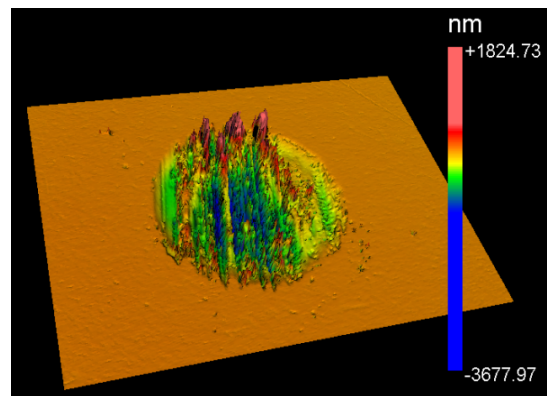


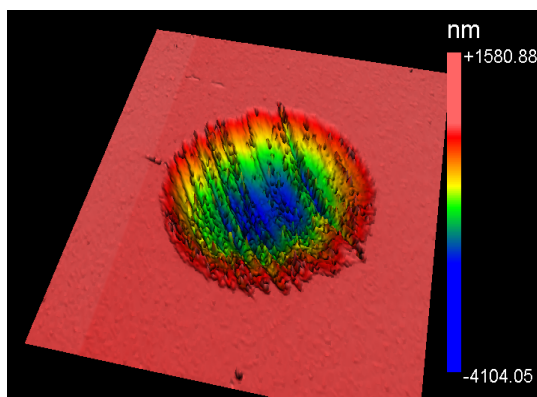
Figure 2.8. Percentage of variation of the WSD and wear volume between Mobil DTE 13M and Esso Nuto H46 samples



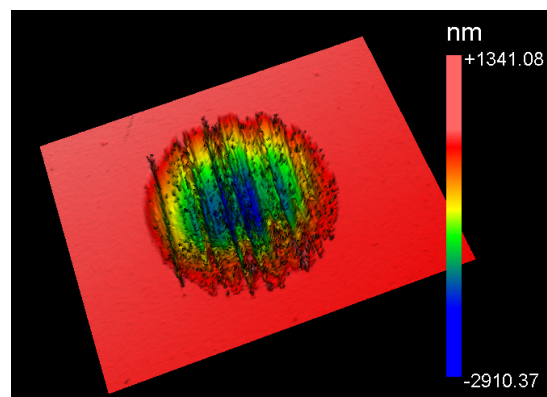
Mobil DTE 13M – 5 minutes, 147 N



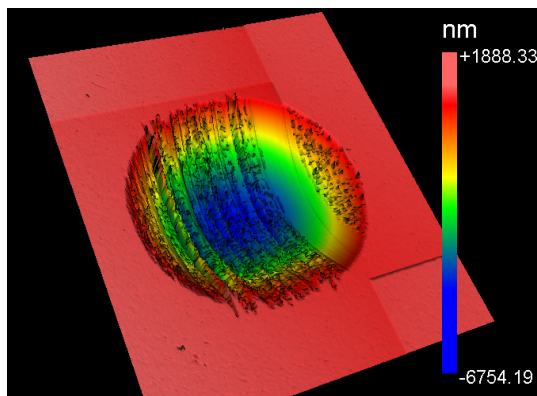
ESSO NUTO H46 – 5 minutes, 147 N



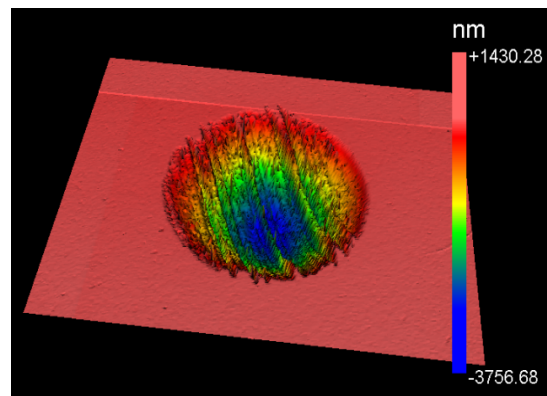
Mobil DTE 13M – 30 minutes, 147 N



ESSO NUTO H46 – 30 minutes, 147 N



Mobil DTE 13M – 60 minutes, 147 N



ESSO NUTO H46 – 60 minutes, 147 N

Figure 2.9. Wear scar (3D images) 147N

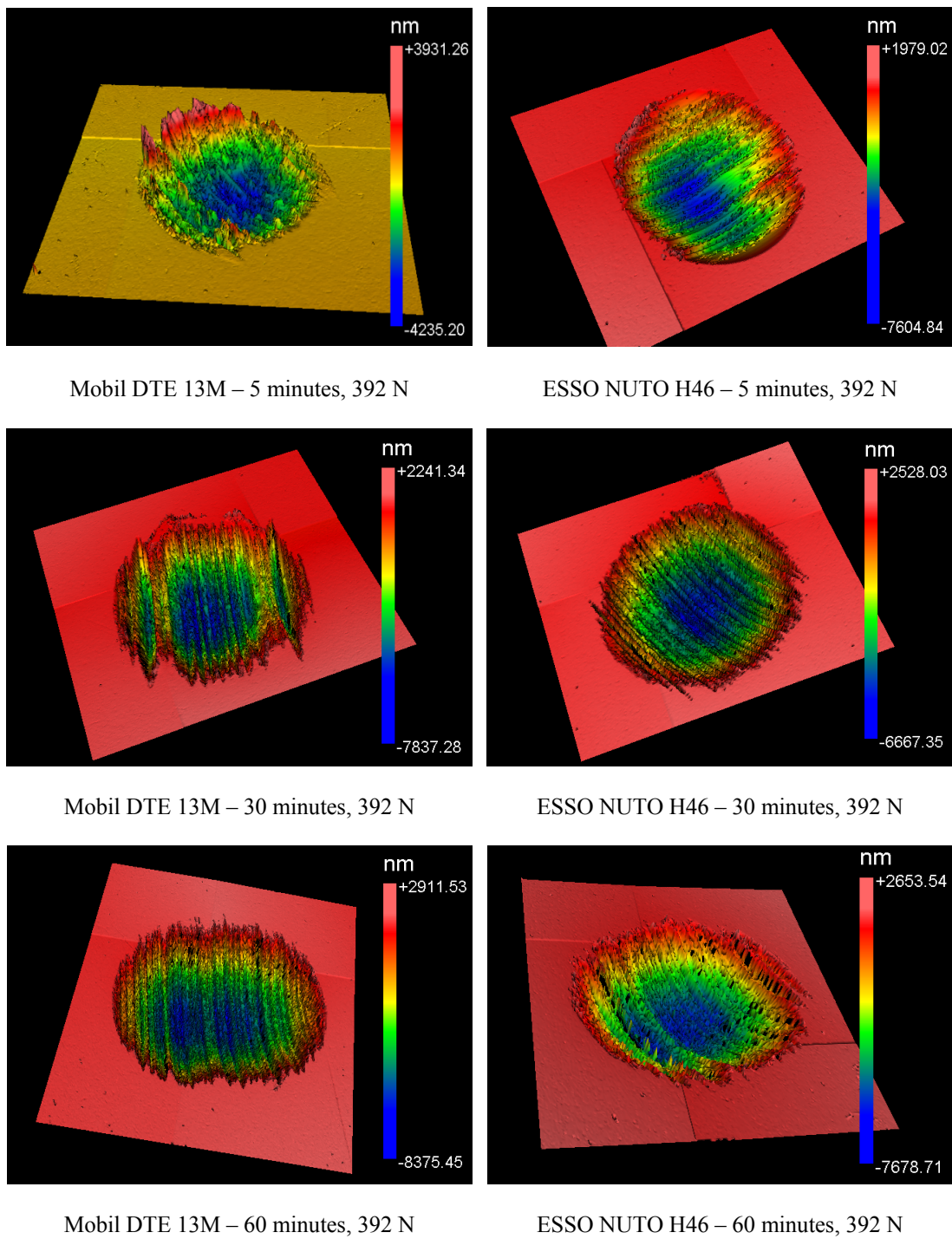


Figure 2.10. Wear scar (3D images) 392N

Another interesting feature is that similar wear scar diameters can correspond to large differences in wear volumes. This finding clearly supports the idea that wear volume measurements using the proposed procedure would perform with increased accuracy compared with typical WSD measurements.

Another factor to consider is shortening the test time. Although there is enough accuracy in wear volume and WSD measurements to shorten the test time, a practical limitation exists which limits the minimum test time to achieve representative results. It should be noted the accuracy in wear volume measurements is dependent on the length of the experiment and the type of objective and size of acquisition domain used in the measurement process. The accuracy for wear volume measurements can be higher than 90% for tests of duration 60 minutes. There is not an easy criterion to select the minimum time, as the type of wear depends on the contact pressure and flash temperature within the contact which is a time-dependant random process difficult to characterise. An attempt to explain this concept appeared in lubricated adhesive wear models introducing the fractional film defect (Stolarski 1987), equation (3.2). The fractional film represents the probability that an asperity encounters as it passes over the mating surface, a site unoccupied by adsorbed lubricant molecules.

The parameter β will vary to a high degree depending on the additive's behaviour during each test. This can be seen for test duration times equal or less than 30 minutes, where the wear volume variations are greater than 25%; Figures 2.7, 2.8. The parameters A_e , A_p are highly dependent on the contact pressures and the material properties within the contact which are influenced by the flash temperature. Then, for short periods, times as the wear rate is pronounced, the values of β , A_e , A_p will significantly change. The complex behaviour of the additives leads to non-expected results. As it can be seen from Figures 2.7, 2.8, oil with higher viscosity performs better at higher loads but its wear prevention characteristics at lower contact pressures are worse than the low viscosity oil. A rationale approach to this problem would expect an increased performance for the high viscosity oil at the lowest load, as the film thickness should be higher than the film thickness of the lowest viscosity oil. Therefore, the probability of asperity contact and wear volume should be lower than with the lowest viscosity oil. The last result should be considered when these

kinds of tests are used as the only criterion to select lubrication oils for real applications.

Long time duration tests cause small changes of the wear scar diameters whilst there are large wear volume variations, Figures 2.11, 2.12. Therefore, the last findings make it difficult to establish a reliable comparison of the wear-preventive properties of additivised lubricants using only wear scar diameters. As an example, Figure 2.11 shows that a change of the WSD of 25%, from 0.4mm to 0.5mm approximately corresponds to a change of volume of around 300%, from $0.2 \cdot 10^6 \mu\text{m}^3$ to $0.7 \cdot 10^6 \mu\text{m}^3$. Wear scars with diameters between 0.45mm to 0.5mm have volumes that vary from $0.41 \cdot 10^6 \mu\text{m}^3$ to $0.77 \cdot 10^6 \mu\text{m}^3$. An 11% diameter variation corresponds to 88% of volume variation. A similar situation and discussion can be obtained from Figure 2.12.

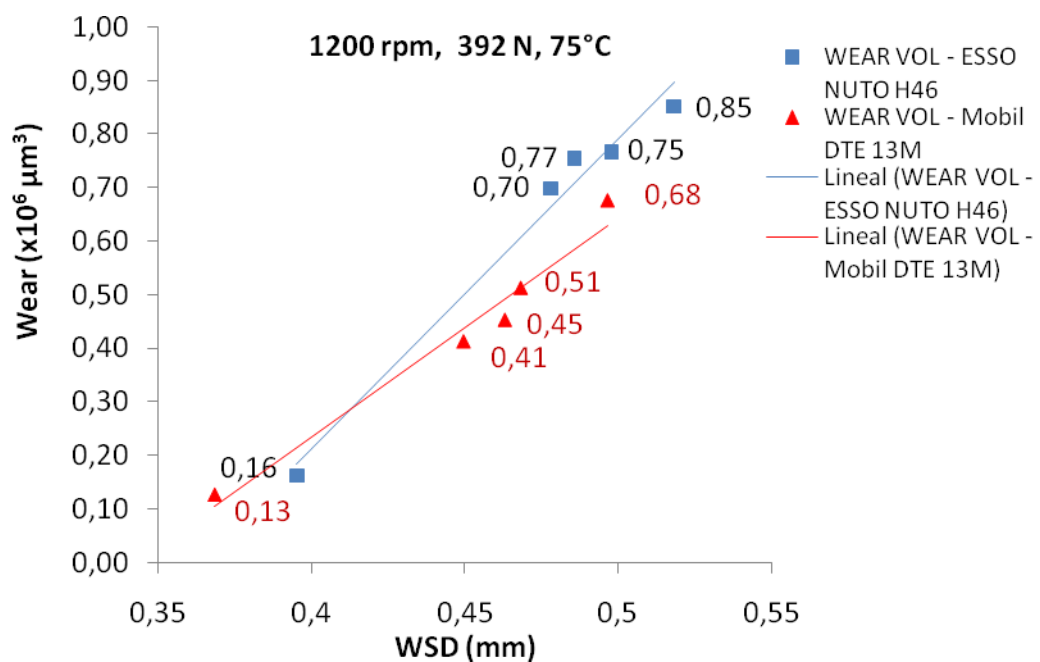


Figure 2.11. Wear volume versus wear scar diameter for both hydraulic oils (392N)

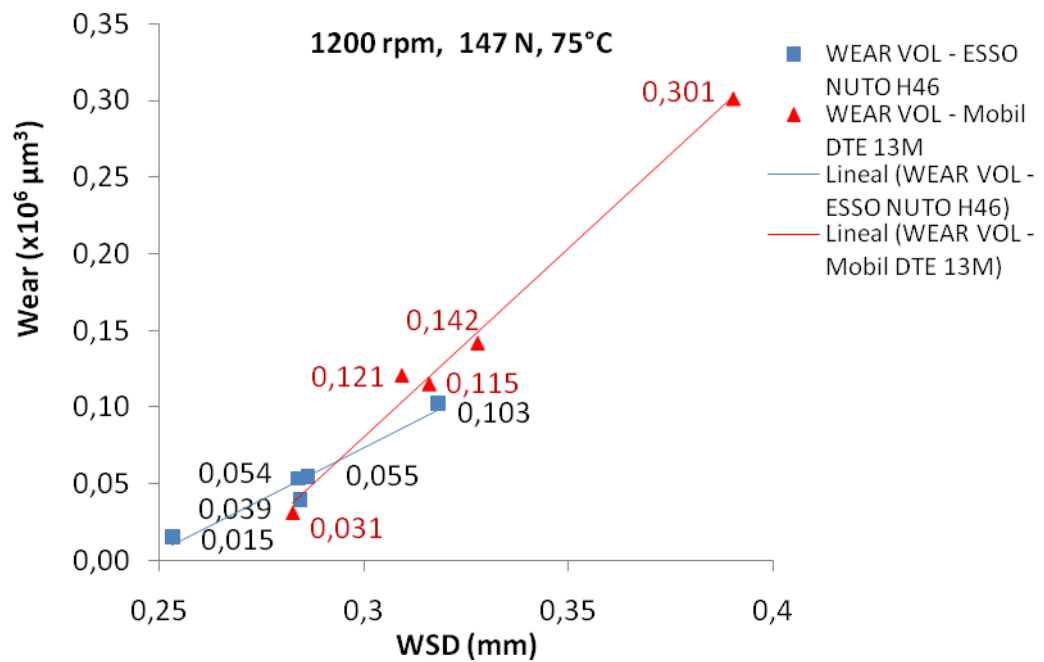


Figure 2.12. Wear volume versus wear scar diameter for both hydraulic oils (147N)

2.4.1.3 Conclusions of the example

Results show it is not easy to draw conclusions of lubricant performance under boundary or mixed lubrication regimes. According to ASTM 4172D, the simple modification of the load leads to a totally different selection of lubricants in terms of wear and the coefficient of friction.

Boundary lubrication is a difficult problem and for some tribologists is one the most complex aspects of the subject of friction and wear prevention. The complexity arises from the large number of variables. Larsen and Perry (Larsen 1950) listed 29 variables, and this list is probably not complete.

In boundary and mixed lubrication, the physical and chemical interaction of the lubricant with the solid body controls friction and wear. The most important aspect of boundary lubrication is the formation of surface films that will protect the contacting surfaces. The boundary films are formed by physical adsorption, chemical adsorption and chemical reaction (Godfrey 1968). The friction of the boundary

layers is strongly dependent on the operating factors such as contact pressure, velocity, temperature and atmosphere.

Some factors of the boundary film formation process are documented (Faraon 2005). The friction of the boundary layer can increase or decrease with the sliding velocity depending on the type of lubricant. The shear stress of the boundary layer decreases with increasing temperature (except when the temperature exceeds its melting point, the boundary layer breaks down in the contact). In the case of steel the existence of oxygen in the contact provides a reduction in friction due to adsorption or chemical reaction of the oxygen with the clean steel surface. The presence of water vapour in the contact leads to a higher coefficient of friction (a higher humidity results in a higher coefficient of friction). It is shown that in general an increase in pressure leads to an increase in shear stress but the relation between shear stress and pressure depends on the type of boundary layer.

As a result, there is a high degree of uncertainty in the boundary film formation process. Sometimes, all these factors, difficult or impossible to control, lead to bench test results that are not representative of the lubricant's performance in real machinery.

3 Wear prediction under the statistical approach

This chapter introduces the fundamentals of the statistical approach to wear prediction to devise a novel friction model for a pin-on-disc machine. The pin-on-disc is a standard bench test configuration that fulfils the initial requirements of this research because it can be used for sliding contacts, such as point and area contacts. The importance of this model is that it can predict the progression of wear for dry and lubricated sliding contacts. Therefore, this model is a step towards the real time prediction of wear volume distribution within the contact area that helps to characterise the performance of ODM sensors under controlled experimental conditions.

The model is developed considering the Coulomb friction model, a probabilistic approach of wear prediction, the kinematics of the pin-on-disc configuration and the elastic theory of bending. The model estimates the magnitude and direction of the frictional force, the pin torque, the probability of asperity contact and the real area of contact distinguishing between the parts due to elastic and plastic asperity contacts respectively. Therefore, the proposed model is suitable for the prediction of adhesive wear. It can be applied to metal contacts for conductance characterisation through the plastically deformed asperities which is of great interest for electrical contact resistance studies.

3.1 Real Area of Contact, an important concept for devising adhesive wear models

Surfaces consist of hills (asperities) of varying heights and spacing and valleys of varying depths and spacing. When two nominally flat surfaces are placed in contact, due to surface roughness physical contact takes place at localised spots within the area that is defined by the macro-geometry. The points at which the actual contact occurs are referred to as junctions. The sum of the areas of all the junctions constitutes the real area of contact. For most materials with applied load, this will be only a small fraction of the apparent area of contact that is determined through macro-geometry.

The real area of contact is a function of the surface texture, material properties and interfacial loading conditions (Bushan 1999).

No general equations exist for characterising all types of wear; this is due to the complex nature of wear. However, for adhesive and abrasive wear a relationship similar to Archard's model (Archard 1953) was proved to characterise these two main wear mechanisms in dry contacts, equation (3-1).

$$V = KA_r L \quad A_r \propto \frac{W}{H} \quad (3-1)$$

Equation (3-1). Archard adhesive wear law for dry contacts

Where

V is the predicted volume of adhesive wear.

L is the sliding distance.

K is the wear coefficient for non-welded junctions

A_r is the real area of contact

W is the applied load

H is the hardness of the worn surface

Adhesive wear occurs when two nominally flat solid bodies are in sliding contact, whether lubricated or not. It is formed due to localised bonding between contacting solid surfaces leading to material transfer between the two surfaces or loss from either surface.

Abrasive wear, on the other hand, is formed due to hard particles or hard protuberances forced against and moving along a solid surface. Material removal from a surface via plastic deformation can occur by several deformation modes which include ploughing, wedge formation and cutting.

3.2 Probability of Asperity Contact

The calculation of the probability of asperity contact can be achieved using two different approaches. The first uses the film thickness and surface statistical parameters for the estimation of the plastic and elastic asperity contact. The second is

based on fitting a frictional model to experimental data such as the friction force and torque readings. The second approach is an inverse problem where the probability of asperity contacts and the elastic and plastic area of contact are determined adjusting the input parameter of the proposed model until its output fits with a good degree of accuracy the experimental data. An example of this approach is shown in section 3.4.

The following block diagrams summarise the procedure to predict the asperity contact and plastic asperity contact using the first approach, Figures 3.1, 3.2. Figure 3.1 shows the procedure to apply the proper film thickness relationships depending on the lubrication regime. The most extended film thickness formulae in engineering for different lubrication regimes are the ones proposed by Hamrock&Downson and Moes&Venner. Figure 3.2 shows the procedure to calculate the probability of asperity contact after the minimum film thickness and lubrication regime is determined.

The input parameters are explained in appendices A and B. The expressions for PAC, PPAC and PPEC are found in (Stolarski 1996). The criterion for detectable plastic flow (plastic deformation δ_p) and the plasticity index $\psi=(\delta_p/\sigma)^{-1/2}$ are determined according to (Greenwood and Williamson 1966). A detailed study of film thickness formulae is covered in Appendix B.

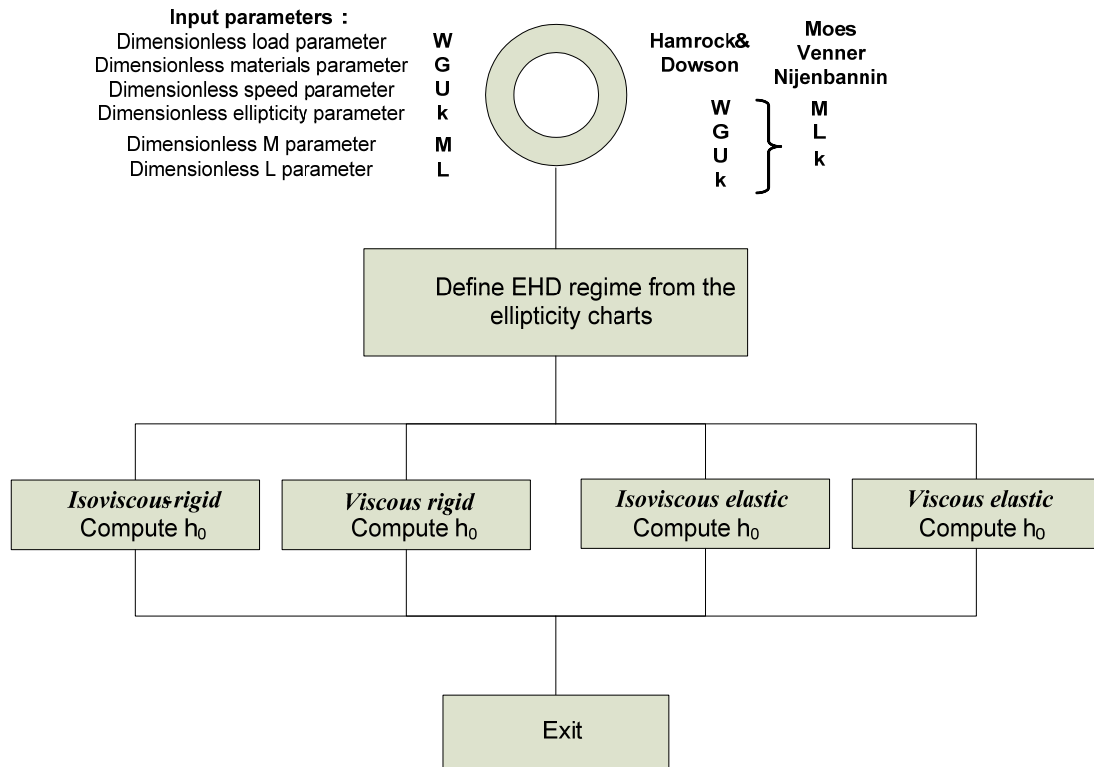


Figure 3.1. Block diagram to calculate minimum film thickness

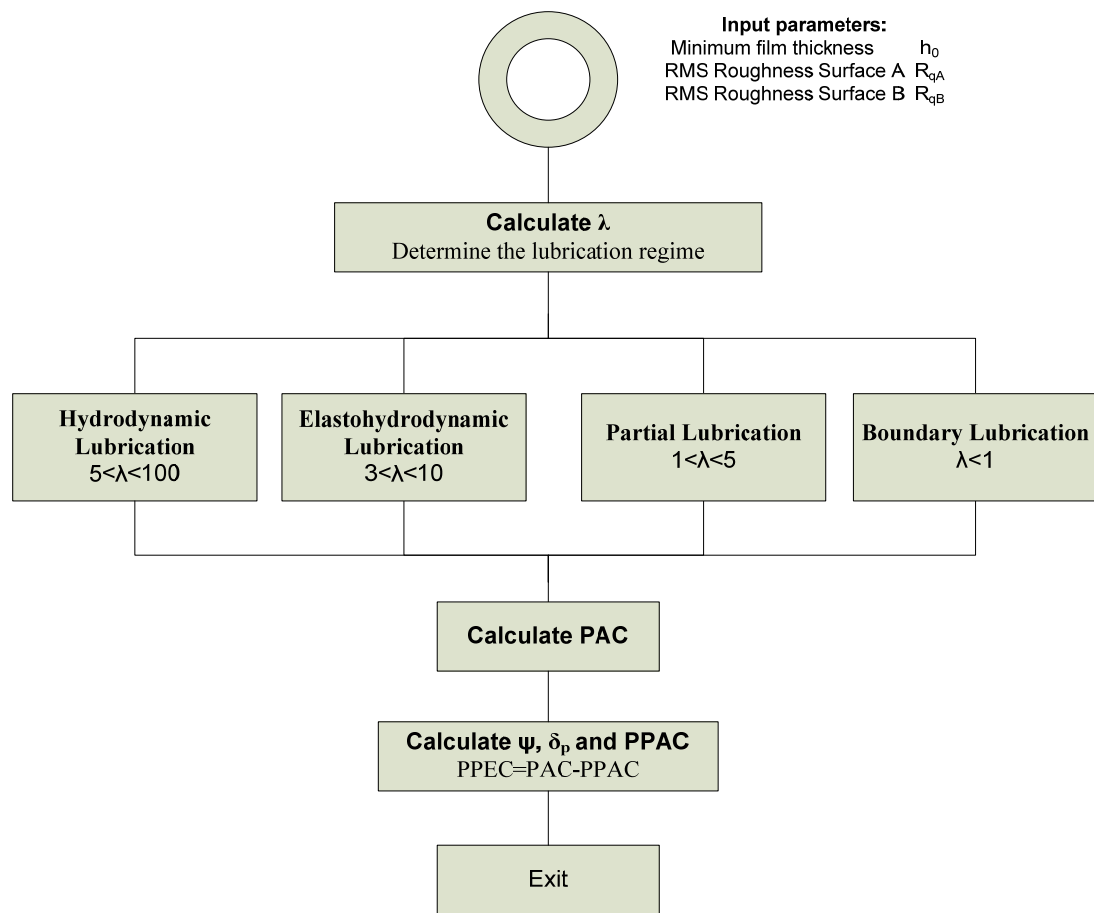


Figure 3.2. Block diagram to calculate the lubrication regime and different probabilities (Stolarski 1996).

3.3 Volume of Adhesive Wear in Lubricated Contacts

The volume of adhesive wear in lubricated contacts can be characterised by means of the fractional film defect parameter, equation (3-2). This represents the probability that an asperity encounters, as it passes over the mating surface, a site unoccupied by adsorbed lubricant molecules (Stolarski 1987, 1990a).

Under conditions of thin film lubrication, the area of direct metal-metal contact, A_m , is a proportion of the true area of contact, A_r .

$$\beta = \frac{A_m}{A_r} \quad (3-2)$$

Equation (3-2). Fractional film defect β

The fractional film defect is difficult to characterise. When it is known, the volume of adhesive wear in a lubricated contact can be estimated using equation (3-3). This equation was derived from the classical adhesive wear model (Archard 1953) and it considers the existence of plastic and elastic contacts within the true area of contact.

$$V = \sqrt{1 + 3\mu^2} \cdot \beta \cdot (k_e A_e + k_p A_p) L \quad (3-3)$$

Equation (3-3). Volume of adhesive wear in lubricated contacts

Where

V is the predicted volume of adhesive wear.

μ is the coefficient of friction.

β is the fractional film defect.

L is the sliding distance.

k_e is the wear coefficient for non-welded junctions

k_p is the wear coefficient for welded junctions

A_e is the area of contact for elastic asperity contacts.

A_p is the area of contact for plastic asperity contacts.

3.4 Model of friction for a pin-on-disc configuration with imposed pin rotation.

3.4.1 Introduction

The pin-on-disc configuration is a common test for the study of sliding wear. The pin-on-disc tester measures the friction and sliding wear properties of dry or lubricated surfaces of a variety of bulk materials and coatings (Bhushan 2001). For polymer testing, special pin-on-disc machines are used to report the wear rates produced using different pin-on-disc configurations (Briscoe and Stolarski 1985b). Two basic pin-on-disc configurations are typically found depending of the loading of the pin along its major axis, which can be either in a direction normal to (horizontal configuration) or parallel with (vertical configuration) the axis of rotation of the disc (Briscoe and Stolarski 1985a). If the rotation of the pin is controlled by an actuator two new categories appear “horizontal configuration with imposed pin rotation” and a “vertical configuration with imposed pin rotation” respectively. The interest of the former machines is that kinematically they are generalisation versions of the usual pin-on-disc machine where the pin is fixed. The kinematic study of this machine was reported in (Stolarski 1989a) and that is the fundamental pillar for devising the probabilistic friction and adhesive wear prediction models for the pin-on-disc apparatus.

Adhesive wear mechanisms always occur in sliding contacts (Lisowski 1981). Friction causes the asperities on one surface to become cold welded to the other surface. The volume of material transferred for adhesive wear is proportional to the real area of contact and the sliding distance. Several probabilistic models based on the Archard Adhesive Wear Law (Archard 1953) for dry and lubricated contacts are found in the literature (Stolarski 1979). A complete review of the probabilistic approach for wear prediction in lubricated sliding contacts can be found in (Stolarski 1996).

An initial study to model the friction is shown in (Stolarski 1989a). Improvements with respect to the model reported in (Stolarski 1989a) are achieved by proposing a novel frictional model based on the following assumptions:

- The pin is an elastic deformable body.
- The real area of contact changes with the ratio between the angular velocity of the pin and the angular velocity of the disc.
- The real area of contact depends on the compression and bending stresses acting on the pin.
- The interfacial shear strength, τ , is independent of the local velocity (Stolarski 1989a).
- The interfacial shear strength is not a function of the magnitude of the local relative displacement (Stolarski 1989a).

A sequential methodology is followed which discusses the steps to be taken in order to develop the proposed model.

3.4.2 Kinematic analysis of the pin-on-disc configuration with imposed rotation of the pin.

The kinematic analysis of the pin-on-disc setup with imposed rotation of the pin was developed in (Stolarski 1989a). As a result, the pin-on-disc configuration with imposed rotation of the pin is kinematically equivalent to the pin-on-plate configuration assuming a pin sliding along a straight line with velocity $V = \omega_d \cdot r$ and simultaneously rotating about its centre, O_p with angular velocity ω_p . The same conclusions can be drawn by analysing the kinematics of the velocity vectors \vec{u} , \vec{s} in one instant of time according to the system of reference $X'Y'$ of Figure 2.3.

The system of reference $X'Y'$ is fixed to the disc. The velocity field is analysed in the instant of time when the centre of the pin is aligned to the axis Y' . XY is a local system of reference centred at $(X', Y') = (0, R)$ for convenience. The disc rotates with an angular velocity ω_d and the pin rotates with an angular velocity ω_p according to Figure 3.3.

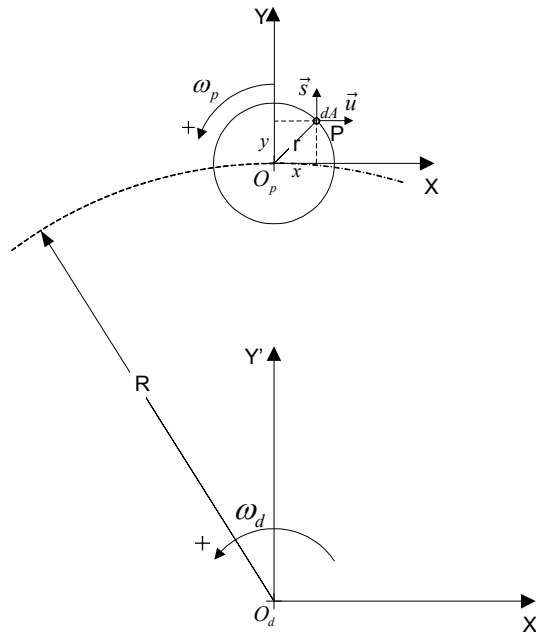


Figure 3.3. Velocity vectors at point P within the pin with imposed rotation ω_p in the pin-on-disc configuration. The system of reference (X', Y') is fixed to the disc .

From Figure 3.3, the field of velocities represented by the components \vec{u} , \vec{s} of a point P within the pin for the instant of time when the centre of the pin is in $(X', Y')=(0, R)$ are shown in equation (3-4).

$$\begin{aligned} u &= \omega_d R + (\omega_d - \omega_p) y \\ s &= (-\omega_d + \omega_p) x \end{aligned} \quad (3-4)$$

Equation (3-4). Velocities of a point P within the pin

Equation 3.4 can be expressed in polar coordinates as shown in equation (3-5).

$$\begin{aligned}
 x &= r \cos(\theta) \\
 y &= r \sin(\theta) \\
 u &= \omega_d R + (\omega_d - \omega_p) r \sin(\theta) \\
 s &= (-\omega_d + \omega_p) r \cos(\theta)
 \end{aligned} \tag{3-5}$$

Equation (3-5). Velocities of a point P within the pin in polar coordinates

Where

R is the distance from the centre of the pin O_p to the centre of the disc O_d (m).

r is the distance of a point P on the pin from its centre, O_p (m).

ω_d is the angular velocity of the disc about its centre, O_d (rad/s).

ω_p is the angular velocity of the pin about its centre, O_p (rad/s).

3.4.3 A probabilistic friction force model for pin-on-disc configuration.

The first step to devise the final model is to assume the following: the real contact area (A_r) is proportional to the contact area (A), the real area of contact (A_r) does not change with α , and it is uniformly distributed within the contact area (A). Under such conditions, the probability of asperity contact can be defined with a new parameter $p = A_r/A$. The real contact area and the probability of asperity contact are proportional to the local contact pressure. This is a result of the combination of the Coulomb friction model and adhesive models as shown in equation (3-6).

$$F_R = \mu N \quad ; \quad F_R = A_r \cdot \tau = \frac{A_r}{A} \tau \cdot A \quad \Rightarrow \quad p = \frac{A_r}{A} = \frac{\mu \cdot N}{\tau \cdot A} \propto \frac{N}{A} \propto \sigma_z \tag{3-6}$$

Equation (3-6). First definition of the probability of asperity contact

where 'p' is the local relation between the real contact area and the contact area. This 'p' is a parameter that can be experimentally obtained.

According to the adhesion model of friction, each differential of the contact area (dA) contributes to the force acting on the pin by the contribution of the differential real contact area (p·dA) and the interfacial shear stress, (τ). The sense of the friction force is opposite to the resultant force acting on the pin, where Ψ is the angle of the resultant force acting on the pin (Stolarski 1989a). The friction forces are shown in equation (3-7).

$$\begin{aligned}
 F_{R_x} &= -\int_A p \cdot \tau \cos(\psi) dA = -\int_0^{r_p} \int_0^{2\pi} p \cdot \tau \cos(\psi) r d\theta dr \\
 F_{R_y} &= -\int_A p \cdot \tau \sin(\psi) dA = -\int_0^{r_p} \int_0^{2\pi} p \cdot \tau \sin(\psi) r d\theta dr \\
 \sin(\psi) &= \frac{s}{\sqrt{(u^2 + s^2)}} = \frac{(\omega_p - \omega_d) r \cos(\theta)}{\sqrt{[\omega_d^2 R^2 + 2Rr\omega_d(\omega_d - \omega_p)\sin(\theta) + (\omega_d - \omega_p)^2 r^2]}} \\
 \cos(\psi) &= \frac{u}{\sqrt{(u^2 + s^2)}} = \frac{\omega_d R + (\omega_d - \omega_p) r \sin(\theta)}{\sqrt{[\omega_d^2 R^2 + 2Rr\omega_d(\omega_d - \omega_p)\sin(\theta) + (\omega_d - \omega_p)^2 r^2]}}
 \end{aligned} \tag{3-7}$$

Equation (3-7). Friction forces F_{R_x} , F_{R_y}

The previous expressions can be written in more elegant form using the parameter $\alpha = \omega_p / \omega_d$, equation (3-8). Alpha (α) is the ratio between the angular velocity of the pin and the angular velocity of the disc. According to Figure 3.3 if the pin and the disc are rotating in the same sense α is positive, otherwise is negative.

$$\begin{aligned}
 F_{R_x} &= -\int_0^{r_p} \int_0^{2\pi} p \cdot \tau \frac{[R + (1 - \alpha) r \sin(\theta)] r}{\sqrt{[R^2 + 2Rr(1 - \alpha)\sin(\theta) + (1 - \alpha)^2 r^2]}} dr d\theta \\
 F_{R_y} &= -\int_0^{r_p} \int_0^{2\pi} p \cdot \tau \frac{(\alpha - 1) r^2 \cos(\theta)}{\sqrt{[R^2 + 2Rr(1 - \alpha)\sin(\theta) + (1 - \alpha)^2 r^2]}} dr d\theta
 \end{aligned} \tag{3-8}$$

Equation (3-8). Friction forces F_{R_x} , F_{R_y} according to α

The differential friction force which contributes to the differential friction torque is shown in equation (3-9).

$$d\vec{F}_R = dF_{R_x}\hat{x} + dF_{R_y}\hat{y} \quad (3-9)$$

Equation (3-9). Differential of friction force

The differential torque is shown in equation (3-10).

$$\begin{aligned} d\vec{T}_F &= \vec{r} \times d\vec{F}_R \\ d\vec{T}_F &= r(-dF_{R_x} \sin(\theta) + dF_{R_y} \cos(\theta))\hat{z} \end{aligned} \quad (3-10)$$

Equation (3-10). The differential torque

The friction torque is shown in equation (3-11).

$$\begin{aligned} T_F &= \int_0^{r_p} \int_0^{2\pi} p \cdot \tau \frac{[R + (1-\alpha)r \sin(\theta)]r^2 \sin(\theta)}{\sqrt{[R^2 + 2Rr(1-\alpha)\sin(\theta) + (1-\alpha)^2 r^2]}} dr d\theta \dots \\ &- \int_0^{r_p} \int_0^{2\pi} p \cdot \tau \frac{(\alpha-1)r^3 \cos^2(\theta)}{\sqrt{[R^2 + 2Rr(1-\alpha)\sin(\theta) + (1-\alpha)^2 r^2]}} dr d\theta \end{aligned} \quad (3-11)$$

Equation (3-11). The friction torque

Assuming a uniform distribution of the real contact area, the y-component of the frictional force is 0. This conclusion can be drawn visualising that the sum of all velocities along the 'y' direction is 0. This is because the field of velocities along 'y' is an odd function in respect to the 'x' component which fulfils equation (3-12).

$$s(x, y) = -s(-x, y) \quad (3-12)$$

Equation (3-12). Field of velocities (odd function)

As a result, the 'y' component of the force acting on the pin is 0 according to this model.

3.4.4 Normalised results of the preliminary probabilistic friction model.

The integrals of equations (3-8), (3-11) are graphically determined for a wide range of input conditions using a suitable normalisation criterion. This normalisation criterion consists of dividing the equations (3-8), (3-11) by the pin area multiplied by the probability of asperity contact (p) and by the interfacial shear stress (τ). Results are shown in figures 3.4-3.7 for a distance R from the centre of the disc to the centre of the pin of 25mm and positive values of α . Results for negative values of α are shown in Figures 3.8-3.11.

When α is positive, the disc and the pin angular velocities have the same sense otherwise when α is negative, the angular velocities of the disc have opposite senses.

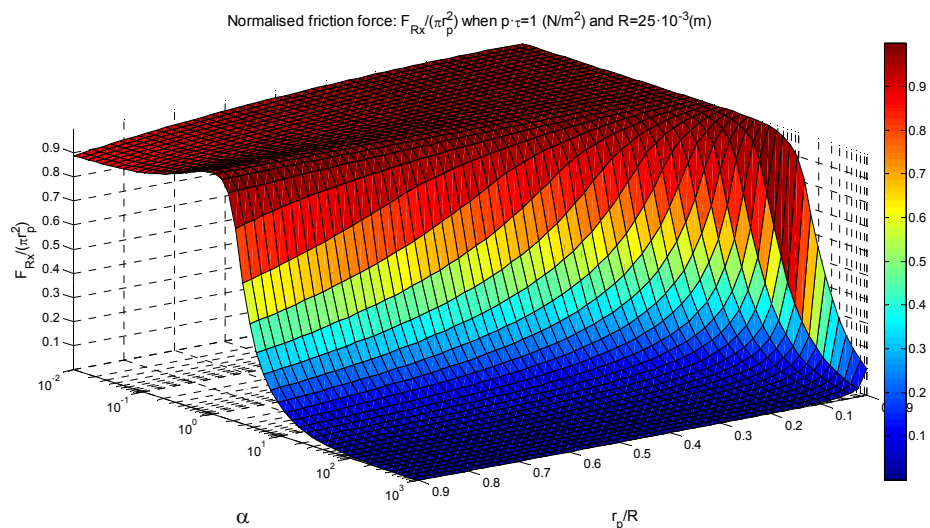


Figure 3.4. Normalised x-component of the friction force (3D). $\alpha \in [10^{-2}, 10^3]$, $r_p/R \in [0.009, 0.9]$

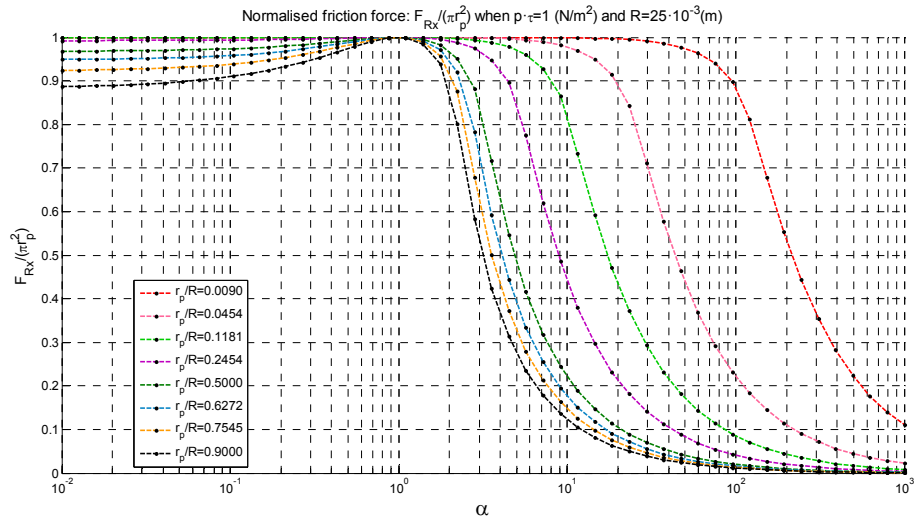


Figure 3.5. Normalised x-component of the friction force (2D). $\alpha \in [10^{-2}, 10^3]$, $r_p/R \in [0.009, 0.9]$

When $\alpha > 0$:

- F_x is proportional to the contact area of the pin
- F_x tends to be maximum when $\alpha < 1$ or $|\omega_d| \gg |\omega_p|$
- F_x tends to be 0 when $|\omega_p| \gg |\omega_d|$

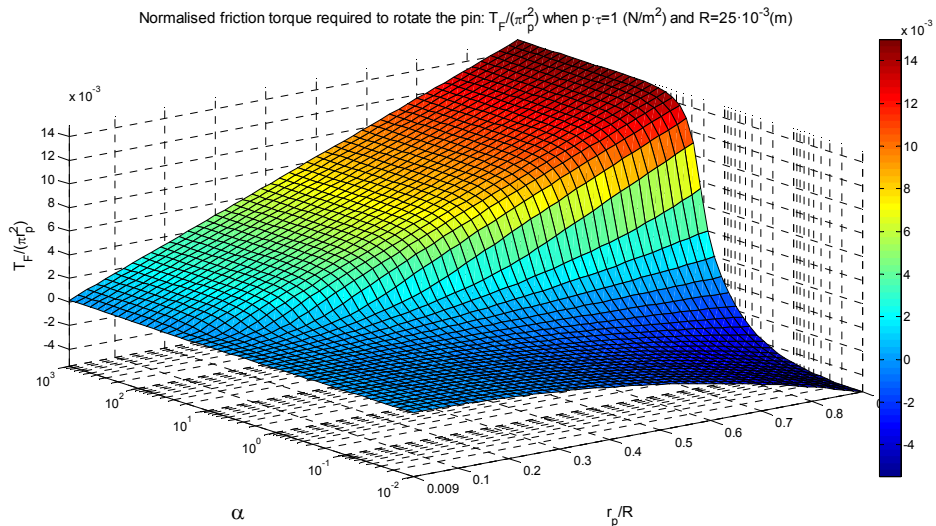


Figure 3.6. Normalised friction torque (3D). $\alpha \in [10^{-2}, 10^3]$, $r_p/R \in [0.009, 0.9]$

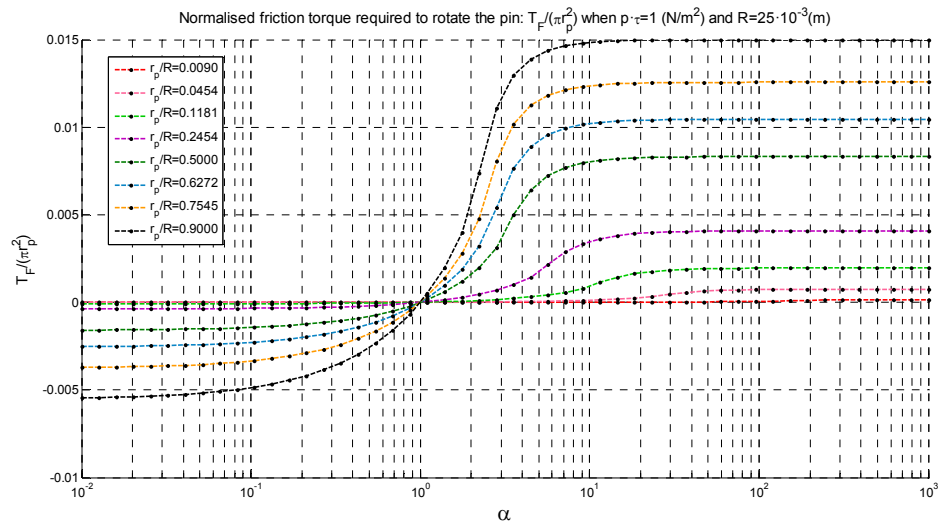


Figure 3.7. Normalised friction torque (2D). $\alpha \in [10^{-2}, 10^3]$, $r_p/R \in [0.009, 0.9]$

When $\alpha > 0$

- The frictional torque is 0 for $\alpha=1$ (The pin does not rotate in the pin-on-plate equivalent model)
- Frictional torque is negative for $\alpha < 1$
- Frictional torque is positive for $\alpha > 1$
- The frictional torque T is proportional to the pin radius.

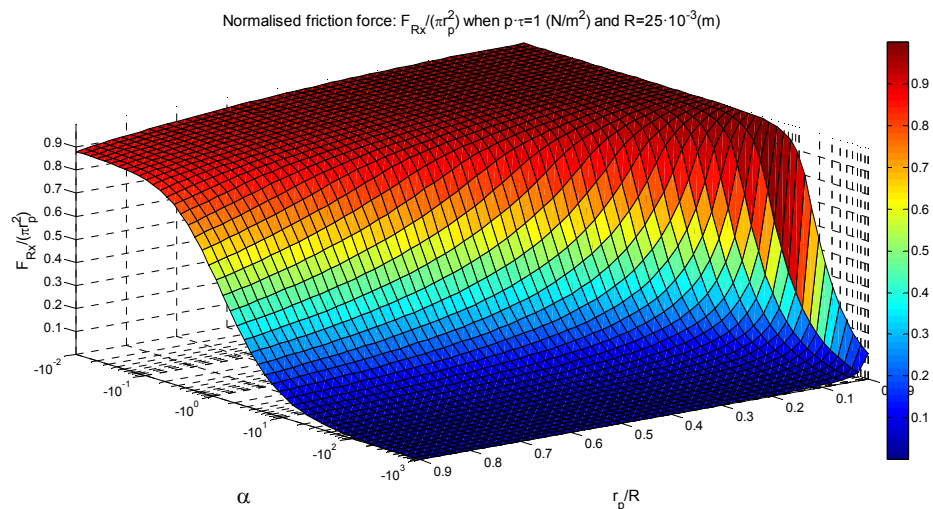


Figure 3.8. Normalised x-component of the friction force (3D). $\alpha \in [10^{-3}, 10^{-2}]$, $r_p/R \in [0.009, 0.9]$

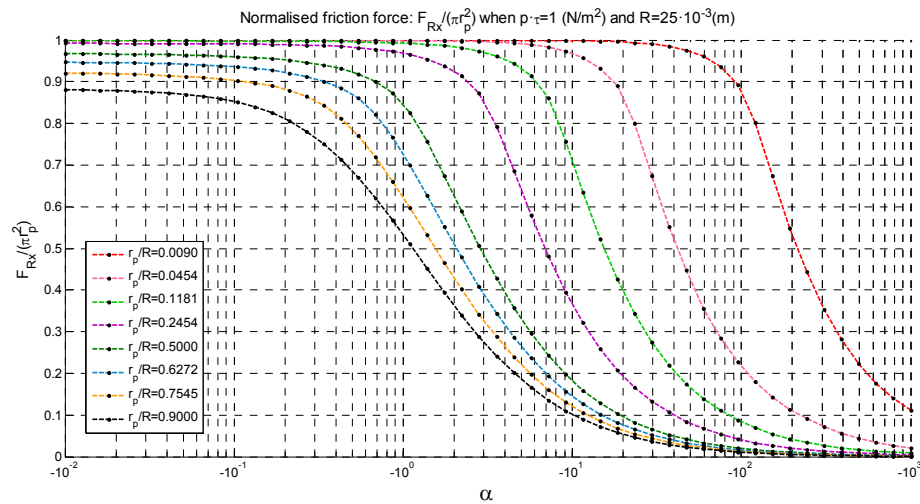


Figure 3.9. Normalised x-component of the friction force (2D). $\alpha \in [-10^3, -10^{-2}]$, $r_p/R \in [0.009, 0.9]$

When $\alpha < 0$:

- F_x is proportional to the contact area of the pin
- F_x tends to be maximum when $|\omega_d| \gg |\omega_p|$
- F_x tends to be 0 when $|\omega_p| \gg |\omega_d|$

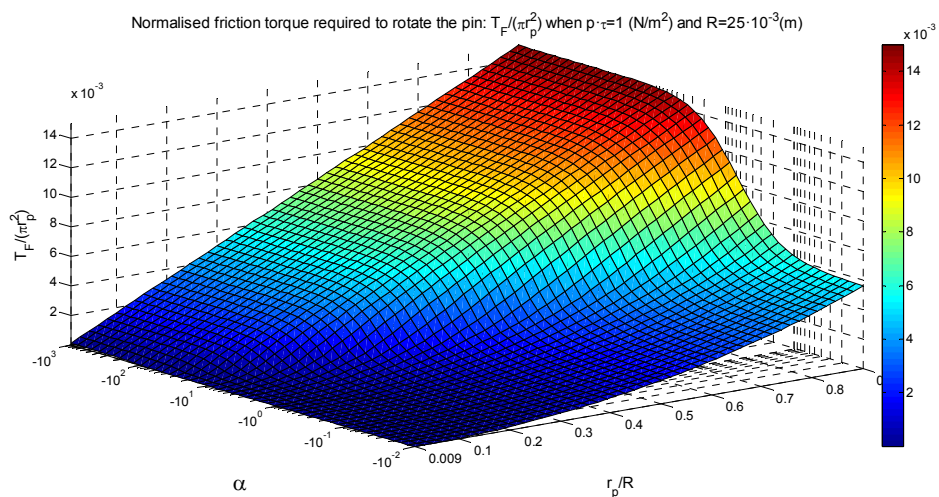


Figure 3.10. Normalised friction torque (3D). $\alpha \in [-10^3, -10^{-2}]$, $r_p/R \in [0.009, 0.9]$

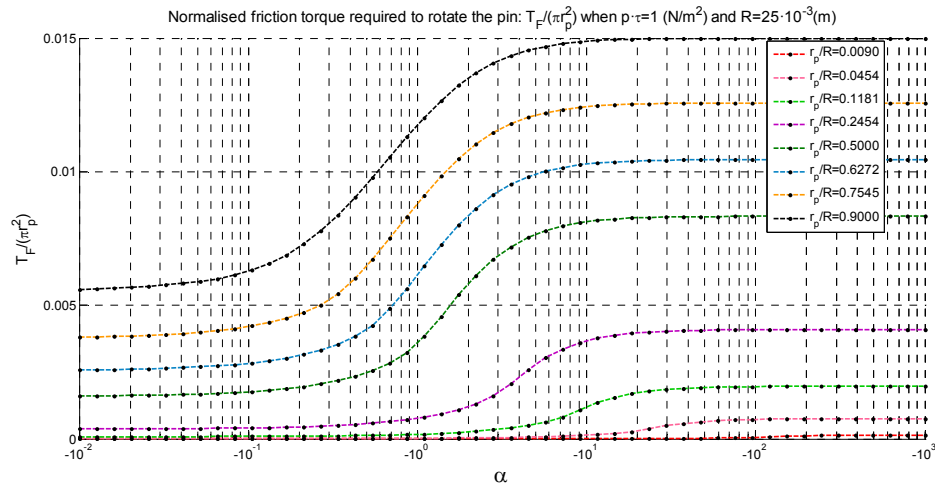


Figure 3.11. Normalised friction torque (2D). $\alpha \in [-10^3, -10^{-2}]$, $r_p/R \in [0.009, 0.9]$

When $\alpha < 0$:

- The sense of the frictional torque T_F does not change for all values of α .
- The frictional torque is maximum if $|\omega_p| \gg |\omega_d|$
- The frictional torque T_F is proportional to the pin radius.

3.4.5 Validation of the preliminary model with experimental data.

The model was devised without considering the type of material and in principle it can be applied to any material. The validation of the model is performed using polymer pins because no other set of experimental data is available. However, the use of polymers is interesting within the context of the thesis. Polymers are widely used in marine applications because they resist the marine environment for long periods of time. As a result, the study of polymer wear is of interest to assess their long term durability.

The experimental data is obtained from (Stolarski 1989a) and it is a curve fitted using different types of regression. The data is presented using a dimensionless variable ϕ introduced on the previous research (Stolarski 1989a). ϕ was initially defined as two times the pin radius (r_p) divided by the product of α by R . Although ϕ is reported in (Stolarski 1989a) as inversely proportional to α , the experimental data must be directly proportional to α to match the model tendencies according to Figures 3.12-3.14. The experimental points and the types of regression used in MATLAB are represented in Table 3.1.

No discussion of the sign of alpha is reported in the previous work (Stolarski 1989a). Therefore, the model predictions are plotted for positive and negative α .

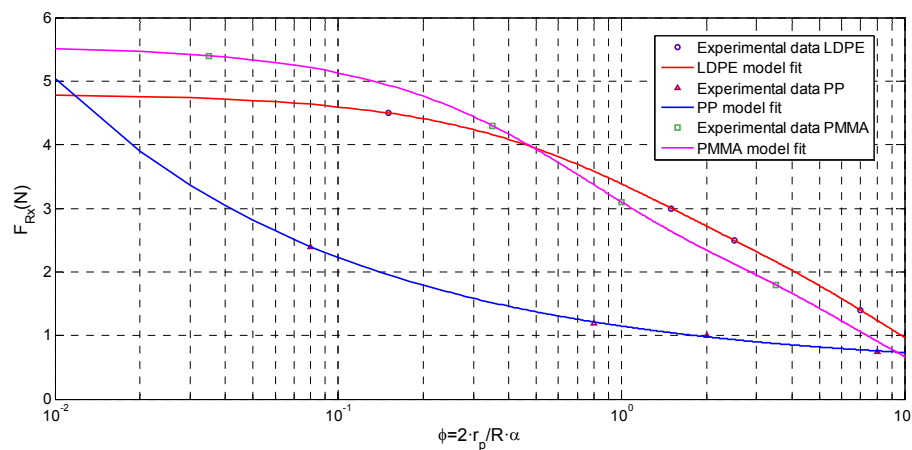


Figure 3.12. Friction force F_{Rx}

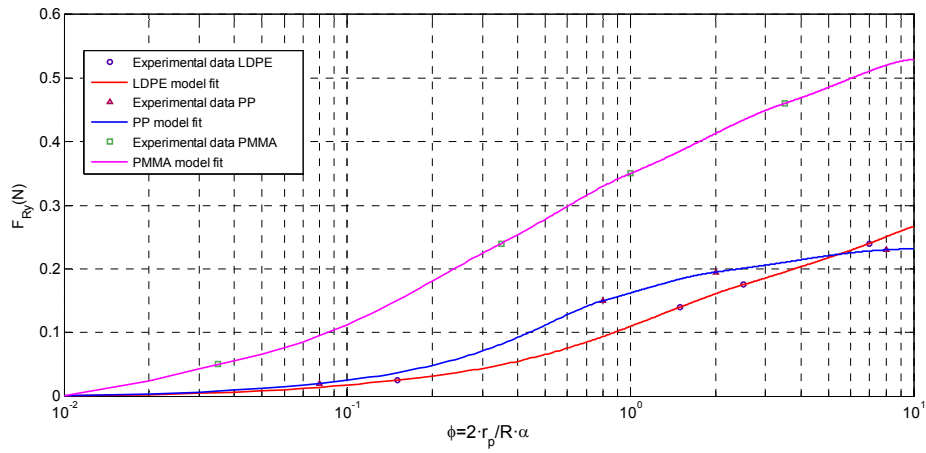
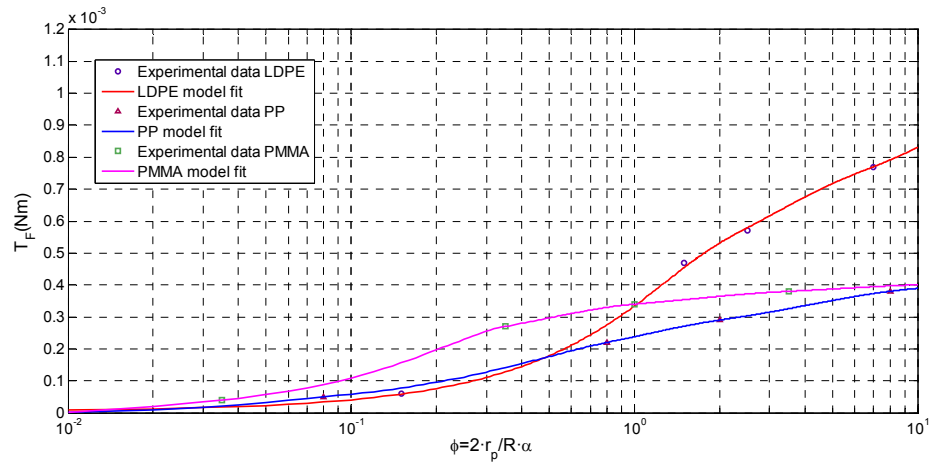
Figure 3.13. Friction force F_{Ry} Figure 3.14. Pin friction torque T_F

Table 3.1. Experimental data obtained from (Stolarski 1989a)

Experimental conditions		PP (Polypropylene) Normal Load=10N , R=25·10 ⁻³ (m) , τ=13·10 ⁶ (N/m ²)			
	φ ₁ =0.08	φ ₂ =0.8	φ ₃ =2	φ ₄ =8	Type of regression
F _{Rx} (N)	2.4	1.2	1	0.75	'power2'
F _{Ry} (N)	0.02	0.15	0.20	0.23	'pchipinterp'
T _F (Nm)	5·10 ⁻⁵	22·10 ⁻⁵	29·10 ⁻⁵	38·10 ⁻⁵	'pchipinterp'
Experimental conditions		LDPE (Low density polyethylene) Normal Load=10N , R=25·10 ⁻³ (m) , τ=8.1·10 ⁶ (N/m ²)			
	φ ₁ =0.15	φ ₂ =1.5	φ ₃ =2.5	φ ₄ =7	Type of regression
F _{Rx} (N)	4.5	3	2.5	1.4	'exp2'
F _{Ry} (N)	0.03	0.14	0.18	0.24	'pchipinterp'
T _F (Nm)	6·10 ⁻⁵	47·10 ⁻⁵	57·10 ⁻⁵	77·10 ⁻⁵	'smoothingspline'
Experimental conditions		PMMA (polymethyl metacrylate) Normal Load=10N , R=25·10 ⁻³ (m) , τ=1.95·10 ⁸ (N/m ²)			
	φ ₁ =0.035	φ ₂ =0.35	φ ₃ =1	φ ₄ =3.5	Type of regression
F _{Rx} (N)	5.4	4.3	3.1	1.8	'exp2'
F _{Ry} (N)	0.05	0.24	0.35	0.46	'pchipinterp'
T _F (Nm)	4·10 ⁻⁵	27·10 ⁻⁵	34·10 ⁻⁵	38·10 ⁻⁵	'pchipinterp'

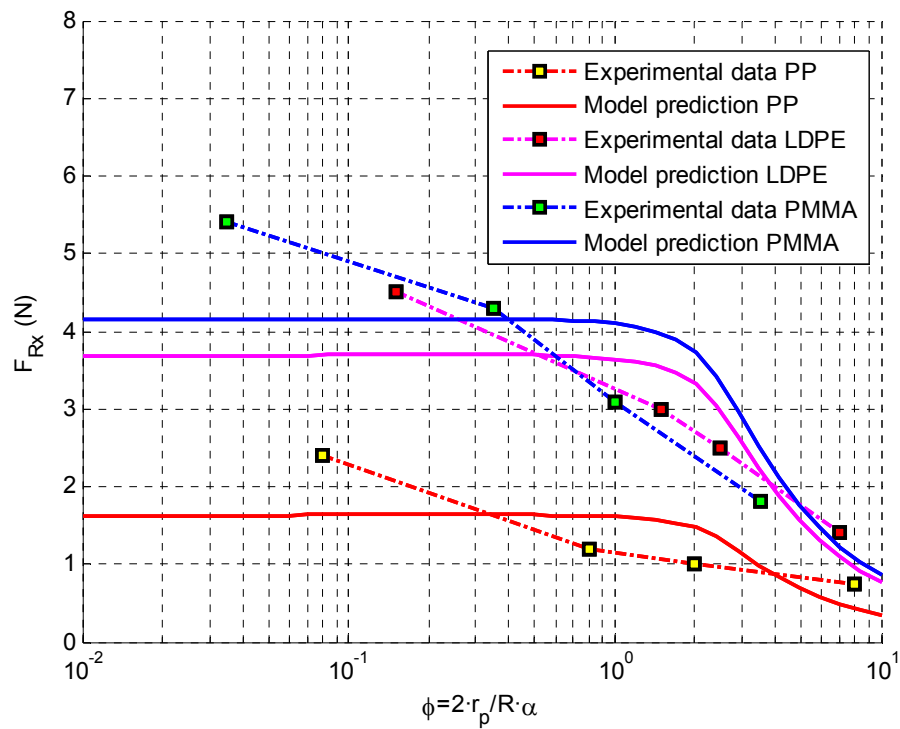
Using the experimental data provided in (Stolarski 1989a), the best value of p according to the x component of the frictional can be obtained by optimizing the best value in the least mean squares sense for the whole range of ϕ . The model predictions are shown in Figures 3.15, 3.16 for a pin of 8mm diameter and positive values of α and ϕ . The probabilities in such cases are shown in Table 3.2. The model predictions are shown in Figures 3.17, 3.18 for a pin of 8mm diameter and negative values of α and ϕ . The probabilities in such cases are shown in Table 3.3.

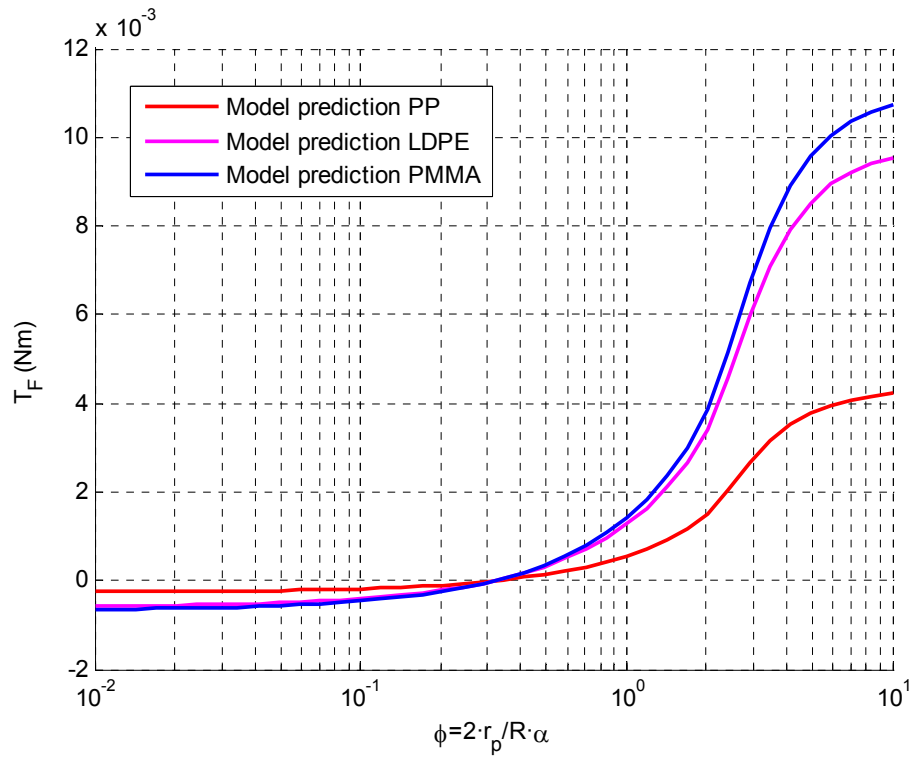
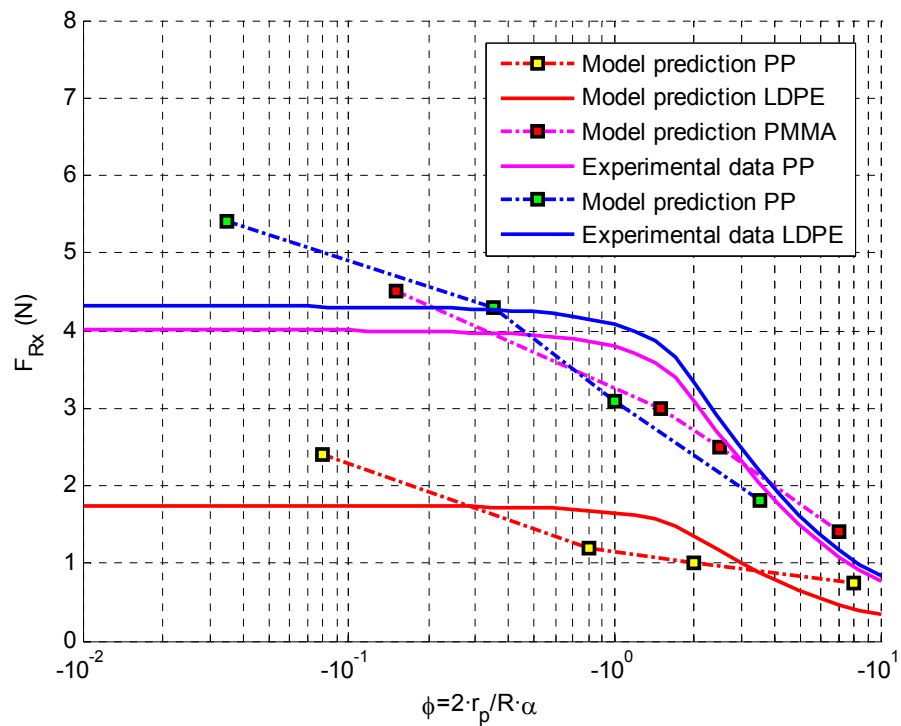
Table 3.2. Best values of p in least mean square sense (LMSE) for $\phi \in [10^{-2}, 10]$

Optimisation LMSE	PP	LDPE	PMMA
p	$2.5042 \cdot 10^{-3}$	$9.0846 \cdot 10^{-3}$	$4.2435 \cdot 10^{-4}$

Table 3.3. Best values of p in least mean square sense (LMSE) for $\phi \in [-10, -10^{-2}]$

Optimisation LMSE	PP	LDPE	PMMA
p	$2.6813 \cdot 10^{-3}$	$9.0881 \cdot 10^{-3}$	$4.4164 \cdot 10^{-4}$

Figure 3.15. Friction force F_{Rx} acting on the pin. Positive alpha (α)

Figure 3.16. Friction torque acting on the pin. Positive alpha (α)Figure 3.17. Friction force F_{Rx} acting on the pin. Negative alpha (α)

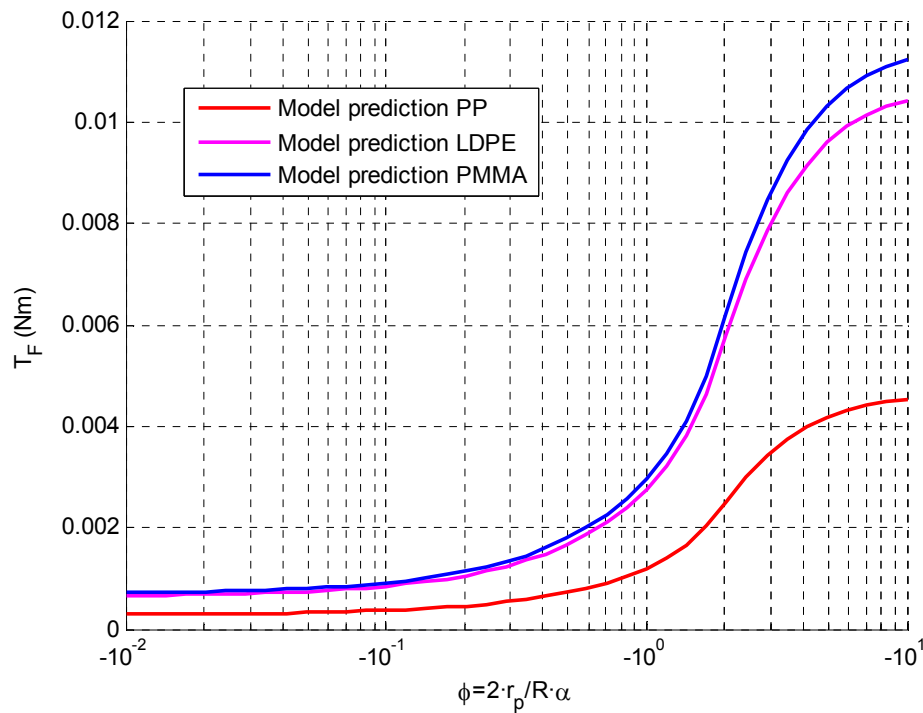


Figure 3.18. Friction torque acting on the pin. Negative alpha (α)

The friction torque predictions for positive and negative α are two orders of magnitude larger than the one reported in (Stolarski 1989a), but the scale of predictions is within the expected range according to the values of F_x and a pin radius of 4mm. The torque tendencies are not exactly the same as the experimental ones.

The next step is the assumption of a real area of contact uniformly distributed within the contact area that can vary with the ratio of the angular velocity of the pin and the angular velocity of the disc (α). This assumption is reasonable as the coefficient of friction could change with the relative velocities within the contacting surfaces (Bayer 2004). Therefore, the probability of asperity contact (p) is expressed as a function of (α).

The values of p can be obtained using the experimental 'x' component of the total frictional force. The model predictions for the 'y' component are still nil, so the experimental values cannot be used to determine p . According to this model, the real area of contact and the contact pressure change with α . The proposed model has the limitation of a real area of contact totally plastic or elastic depending on the contact

pressure. A suitable criterion to classify the type of contact area consists of selecting an elastic asperity contact if the contact pressure is less than $0.6H$, H is the hardness and plastic asperity contact (Greenwood and Williamson 1966).

The PAC and torque predictions for this model are shown in Figures 3.18, 3.19 for positive α values and Figures 3.21, 3.22 for negative α values respectively. The torque predictions have the same tendencies as the experimental data.

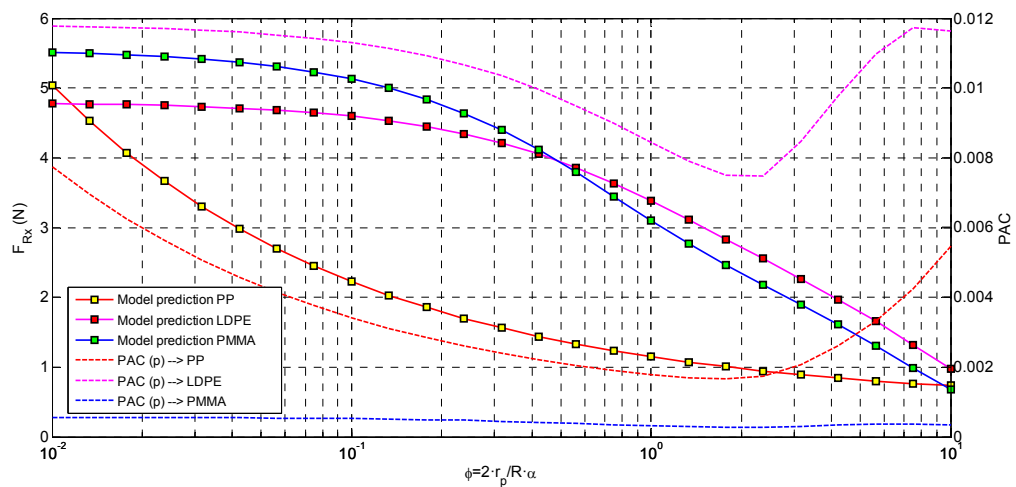


Figure 3.19. Probability of asperity contact (PAC) is a function of α . Obtained from the experimental x-component of the friction force acting on the pin. Positive α values.

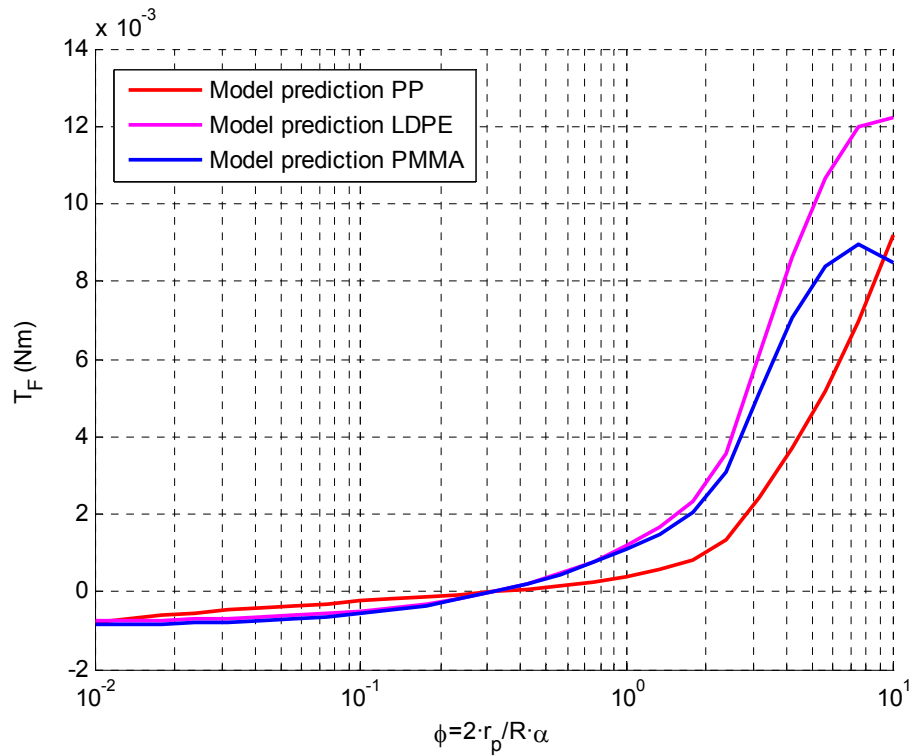


Figure 3.20. Friction torque predictions if the PAC varies with α . Positive α values.

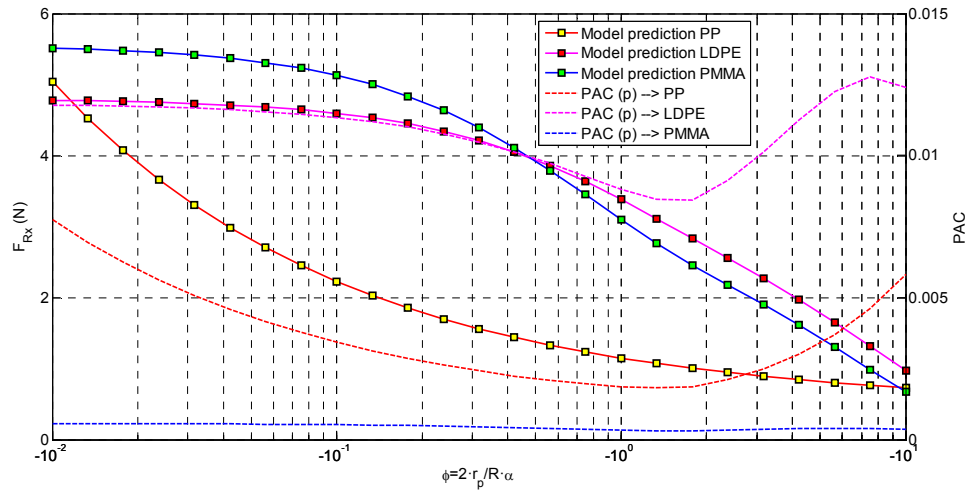


Figure 3.21. Probability of asperity contact (PAC) is a function of α . Obtained from the experimental x-component of the friction force acting on the pin. Negative α values.

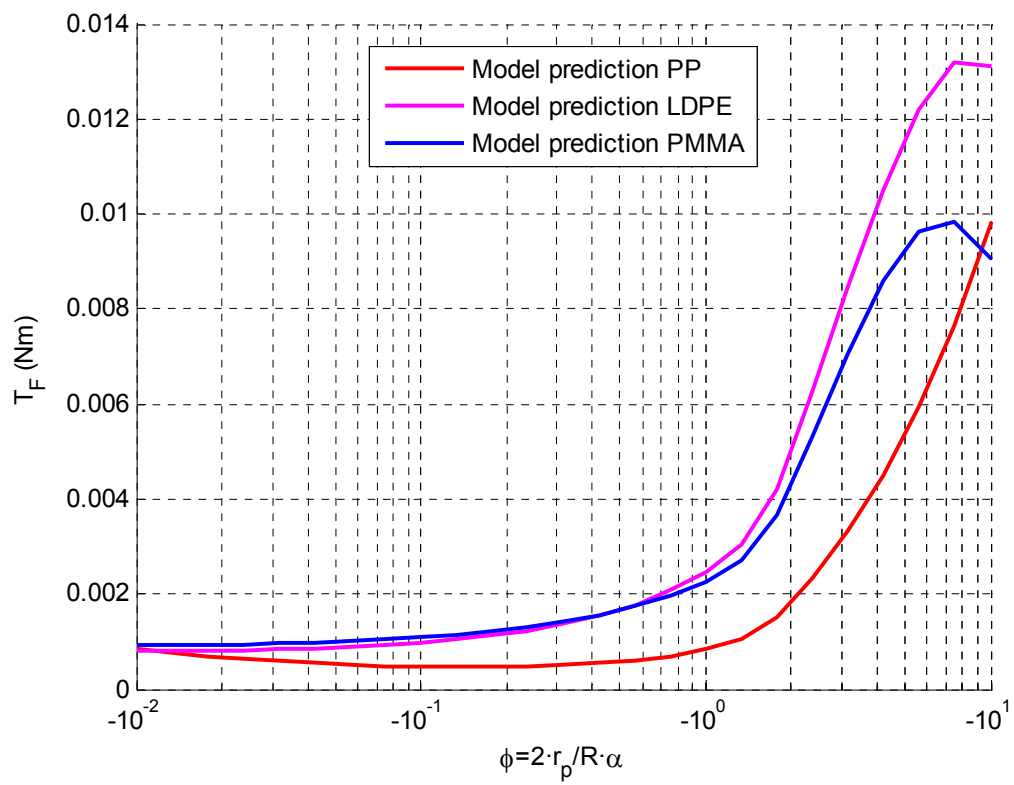


Figure 3.22. Friction torque predictions if the PAC varies with alpha (α). Negative α values.

3.4.6 Improved probabilistic model for pin-on-disc configuration.

If the pin behaves as an elastic deformable body, the stresses acting on the pin can be analysed using the elementary elastic bending theory. The contact pressure along the pin is higher in the vicinities of the “leading” edge than in the “trailing” edge due to a slight bending produced by the friction force, Figure 3.23.

A non-uniform distribution of contact stress leads to a non-uniform distribution of probability of asperity contacts within the contact area. Therefore, this new distribution can produce a frictional force in the direction of the y-axis different to 0. The contact pressure is obtained using the theory of elastic bending (Timoshenko and Goodier 1982). A bar under a bending couple has the stresses shown in equation (3-13).

$$\sigma_z = \frac{E}{R}x \quad \sigma_y = \sigma_x = \tau_{xy} = \tau_{xz} = \tau_{yz} = 0 \quad (3-13)$$

Equation (3-13). A bar under a bending couple

where R is the radius of curvature and E is the Young’s modulus. The value of σ_z is a plane that contains the ‘y’ axis.

The moment created by the distribution of σ_z is given by equation (3-14).

$$M_y = \int \sigma_z x dA \quad (3-14)$$

Equation (3-14). Moment created by the distribution of σ_z

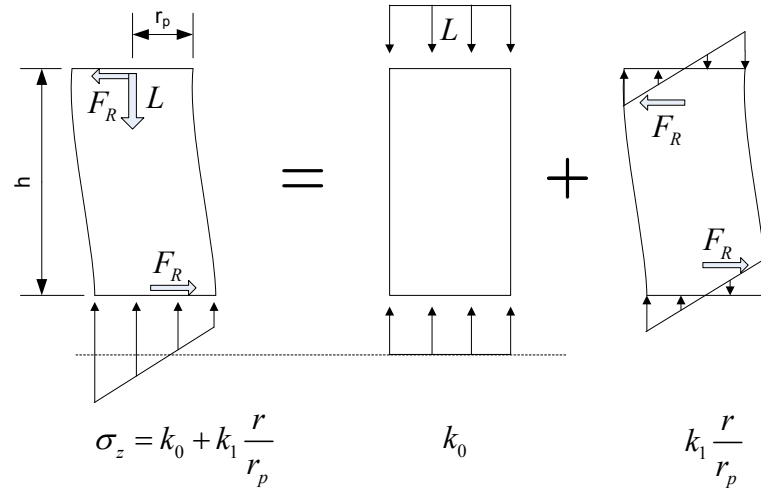


Figure 3.23. Elastic model of the pin. Principle of Superposition. L is the applied load.

The determination of 'k₁' is shown in equation (3-15). Figure 3.24 helps to visualise the equivalent force system.

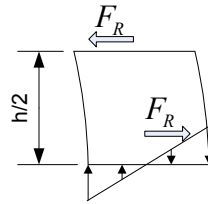


Figure 3.24. Equivalent force system for the determination of k_1 .

$$\sum M = 0$$

$$+F_R \frac{h}{2} = -\int \sigma_z x dA = -\int \frac{k_1}{r_p} x^2 dA = -\frac{k_1}{r_p} \int_0^{2\pi} \int_0^{r_p} (r \cos(\theta))^2 r dr d\theta = -\frac{k_1}{r_p} \frac{\pi}{4} r_p^4 \quad (3-15)$$

Equation (3-15). Determination of k_1

Although the contact pressure in contact theory is normally expressed as a positive, in elasticity a compressive stress is expressed as a negative. In this study the contact pressure is positive.

Using the Principle of Superposition it is possible to divide the stress distribution into two different problems areas: the compression and bending, Figure 21. As a result, the contact pressure can be expressed as shown in equation (3-16).

$$\sigma_z = k_0 + \frac{k_1}{r_p} x \quad (3-16)$$

Equation (3-16). Contact pressure σ_z (compression and bending)

and k_0 , k_1 expressions are shown in equation (3-17).

$$k_0 = \frac{L}{\pi r_p^2} \quad k_1 = \frac{-2F_R h}{\pi r_p^3} \quad (3-17)$$

Equation (3-17). k_0, k_1 values in the contact pressure expression σ_z

where L is the load, h is the height of the pin and F_R is the resultant friction force.

It is possible to calculate a limit of height or a limit of friction force to ensure the whole pin surface is making contact with the disc. This limitation is expressed in equation (3-18).

$$0 \leq k_0 + \frac{k_1}{r_p} x \quad (3-18)$$

Equation (3-18). Moment created by the distribution of σ_z

The condition of contacting surfaces separation occurs at the point where x is equal to the pin radius (r_p), equation (3-19).

$$k_0 \geq k_1 \rightarrow \frac{L}{\pi r_p^2} \geq \frac{2F_R h}{\pi r_p^3} \rightarrow \frac{L r_p}{2} \geq F_R h \quad (3-19)$$

Equation (3-19). Condition of contacting surfaces separation

This limitation implies a pin can be separated from the disc if its height or the friction force is high enough.

The stress distribution on the counterface of the pin can be generalised for a resultant friction force that it is not parallel to the x-axis. θ_R is the angle between the x-axis and the result of the experimental friction force, hence the contact stress distribution is as described by equation (3-20).

$$\sigma_z = k_0 + \frac{k_1}{r_p} \cos(\theta - \theta_R) \cdot r \quad (3-20)$$

Equation (3-20). Stress distribution on the counterface of the pin

The friction forces and friction pin torque for this new model can be expressed as shown in equation (3-21).

$$\begin{aligned} F_{R_x} &= - \int_0^{r_p} \int_0^{2\pi} PAC \cdot \tau \frac{[R + (1-\alpha)r \sin(\theta)]r}{\sqrt{[R^2 + 2Rr(1-\alpha)\sin(\theta) + (1-\alpha)^2 r^2]}} dr d\theta \\ F_{R_y} &= - \int_0^{r_p} \int_0^{2\pi} PAC \cdot \tau \frac{(\alpha-1)r^2 \cos(\theta)}{\sqrt{[R^2 + 2Rr(1-\alpha)\sin(\theta) + (1-\alpha)^2 r^2]}} dr d\theta \\ T_F &= \int_0^{r_p} \int_0^{2\pi} PAC \cdot \tau \frac{[R + (1-\alpha)r \sin(\theta)]r^2 \sin(\theta)}{\sqrt{[R^2 + 2Rr(1-\alpha)\sin(\theta) + (1-\alpha)^2 r^2]}} dr d\theta \dots \\ &\quad - \int_0^{r_p} \int_0^{2\pi} PAC \cdot \tau \frac{(\alpha-1)r^3 \cos^2(\theta)}{\sqrt{[R^2 + 2Rr(1-\alpha)\sin(\theta) + (1-\alpha)^2 r^2]}} dr d\theta \end{aligned} \quad (3-21)$$

$$\theta_R = \text{arctg}\left(\frac{F_y}{F_x}\right)$$

Equation (3-21). Friction forces and friction pin torque for the proposed model

The probability of asperity contact for this model is shown in equation (3-22).

$$PAC \equiv \begin{cases} p \left(k_0 + \frac{k_1}{r_p} r \cos(\theta - \theta_R) \right) \\ 0 \quad \text{if} \left(k_0 + \frac{k_1}{r_p} r \cos(\theta - \theta_R) \right) \leq 0 \end{cases} \quad (3-22)$$

$$r \in (0, r_p) \quad \theta \in (0, 2\pi)$$

Equation (3-22). Probability of asperity contact for the proposed model

The main improvements of this model in respect to the previous one are the predictions of:

- The “y” component of the friction force.
- The contact stress distribution within the pin.
- The probability of asperity contact.
- The real contact area due to elastic and plastic asperity contacts according to the mentioned criteria of $0.6H$.

3.4.7 Possible applications of the improved model

Several studies of adhesive wear in dry and lubricated contacts are reported in the literature (Stolarski 1979, 1996). The probabilistic models are based on Archard Wear Law. Among them, the most complete in the literature predicts adhesive wear for lubricated contacts (Stolarski 1979) and is shown in equation (3-3).

As the model predicts the contact stress distribution and the real contact area distinguishing between elastic and plastic asperity contacts according to Tabor criterion ($0.6 \cdot H$), the model can be used to predict wear volumes in dry and lubricated contacts. Another application of the proposed model is the study of the electrical contact resistance (ECR). This is of special interest for the assessment of the boundary layer formation under mixed and boundary lubrication regimes. The study can be carried out assuming only the plastic asperity contacts are the main

contributors for the final conductance (Greenwood and Williamson 1966) as shown in equation (3-23).

$$G_p = \frac{A_p}{\rho} \quad (3-23)$$

Equation (3-23). Conductance according to (Greenwood and Williamson 1966)

Where :

ρ → the electrical resistivity of the material.

3.4.8 Validation of the preliminary model with experimental data.

The model predictions for a pin diameter of 8mm, a pin height of 10mm, a load of 10N, a distance from the centre of the pin to the centre of the disc of 25mm and positive α values are shown in Figures 3.25-3.27. Results for same input parameters and negative α values are shown in Figures 3.28-3.30. The model predicts a y-component different than 0 but there is certain disagreement with the experimental data.

Figures 3.31-3.36 show the PAC estimates within the polymer pin for different values of ϕ (ϕ is directly related to α). These graphs according to the proposed contact stresses derived from elastic bending theory can explain the ‘edge loading effect’. The areas where there is a highest PAC implies there is more asperity contact and this will result in a higher wear rate according to the adhesive wear model. The PAC values within the pin are shown in Figures 3.31-3.33 for positive α values and Figures 3.34-3.36 for negative values of α .

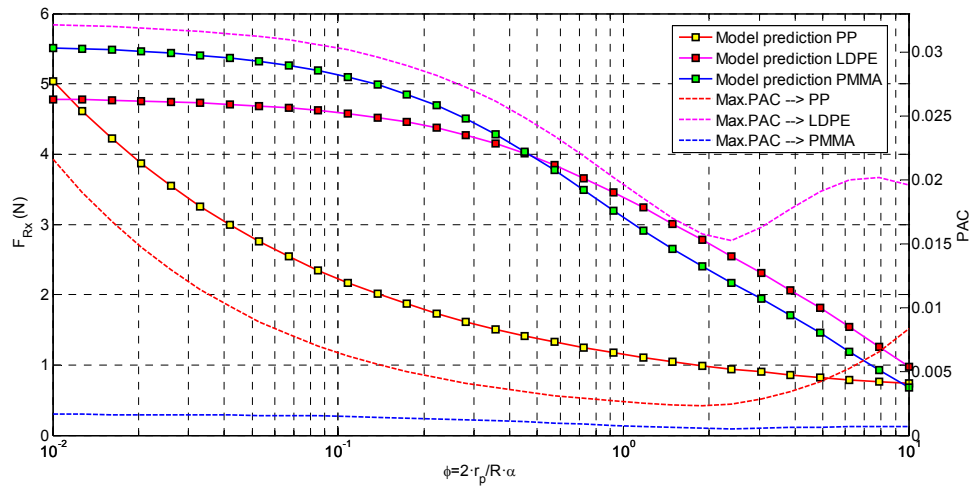


Figure 3.25. Maximum value of the probability of asperity contact (PAC) is a function of α . Obtained from the experimental x-component of the friction force acting on the pin. Positive α values.

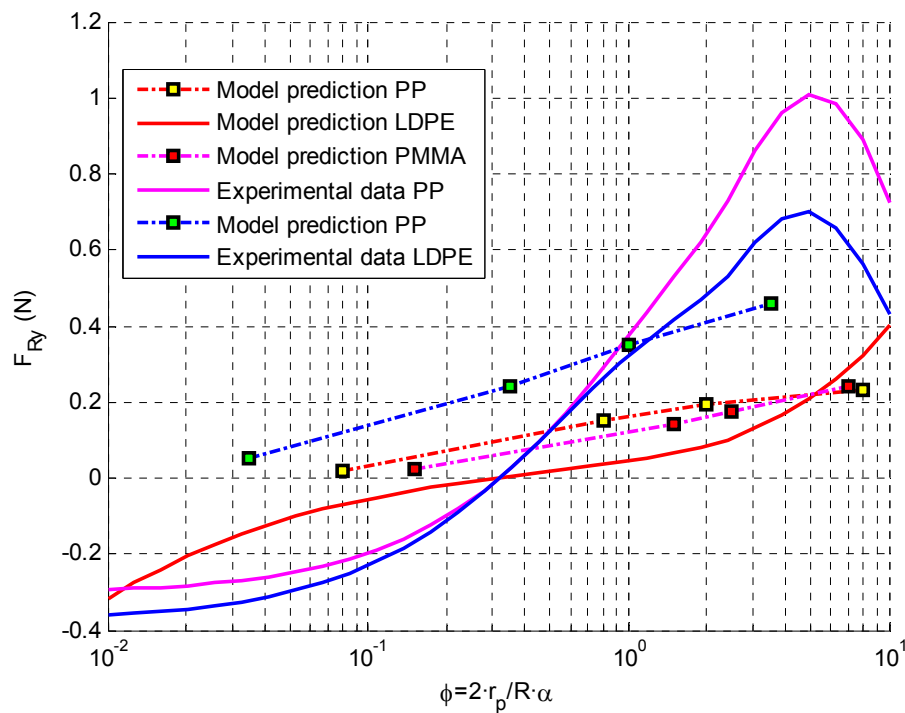


Figure 3.26. Predictions of the y-component of the frictional force. Positive α values.

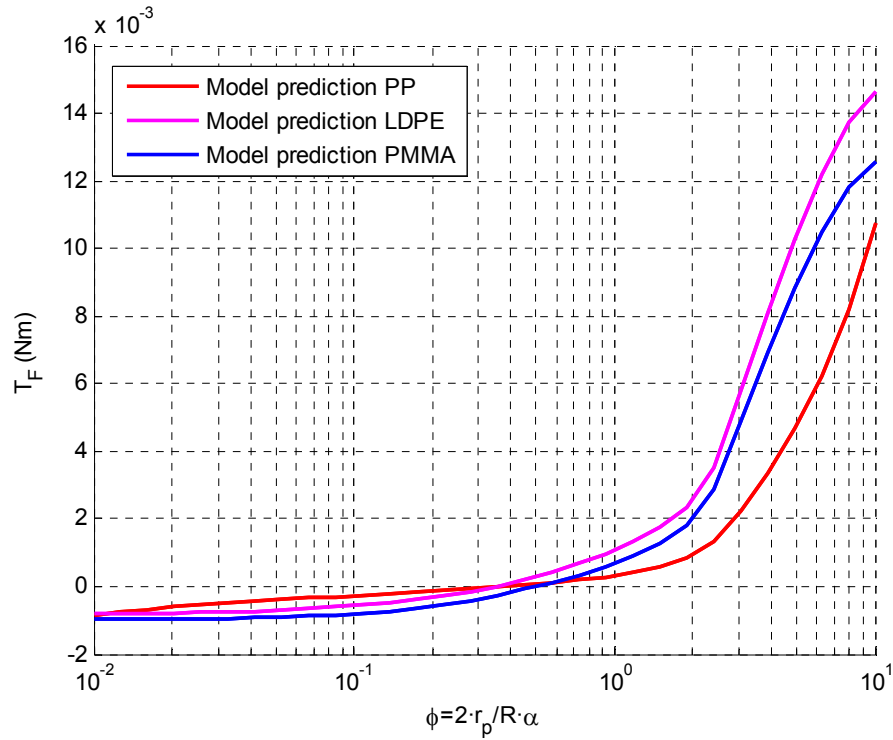


Figure 3.27. Friction torque predictions of the proposed model. Positive α values.

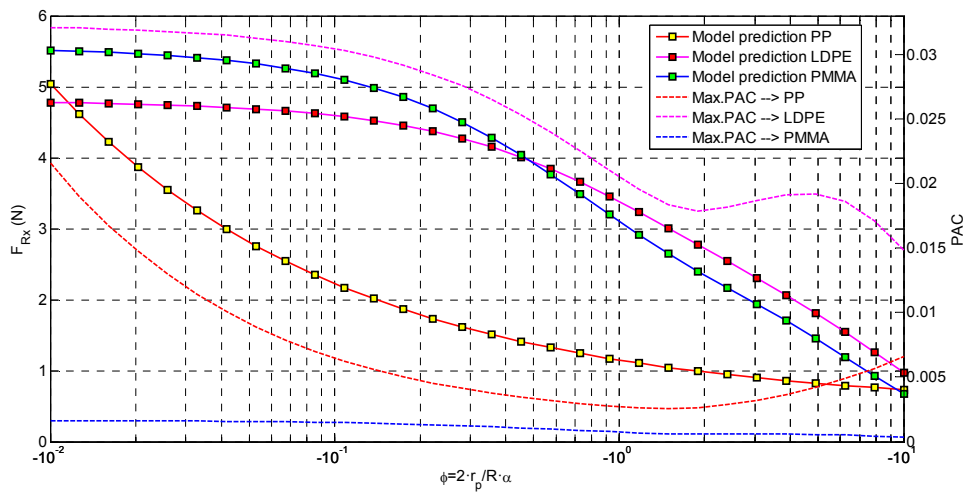


Figure 3.28. Maximum value of the probability of asperity contact (PAC) is a function of α .

Obtained from the experimental x-component of the friction force acting on the pin.

Negative α values.

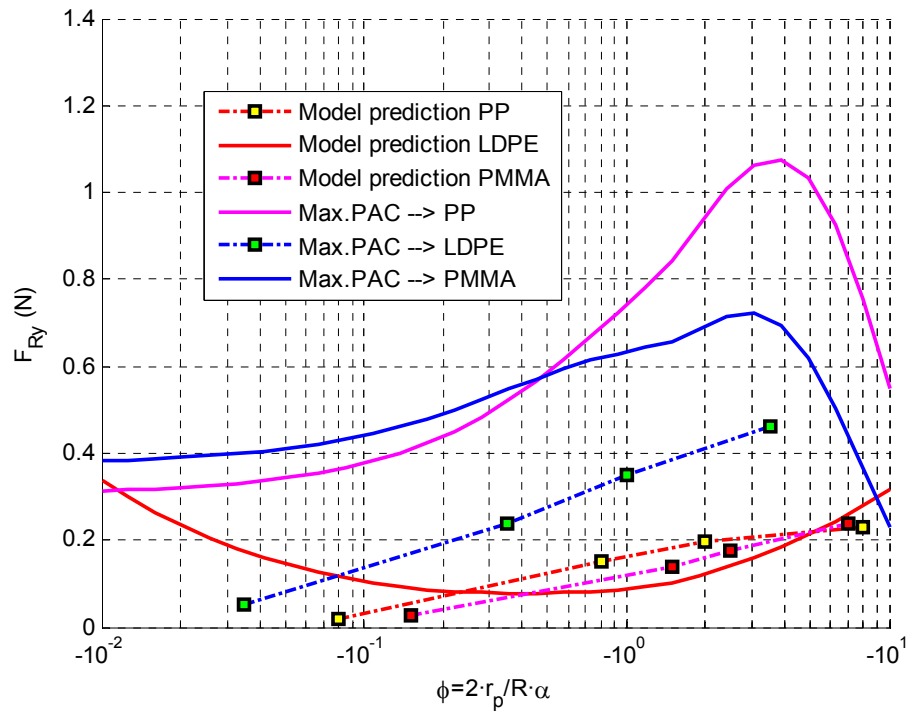


Figure 3.29. Predictions of the y-component of the frictional force. Negative α values.

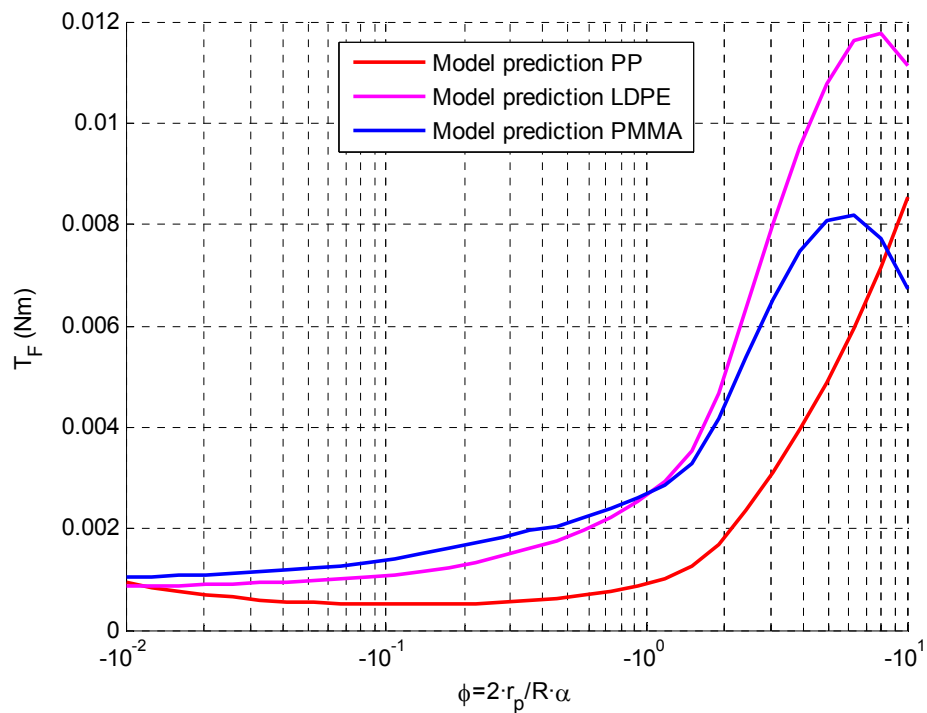
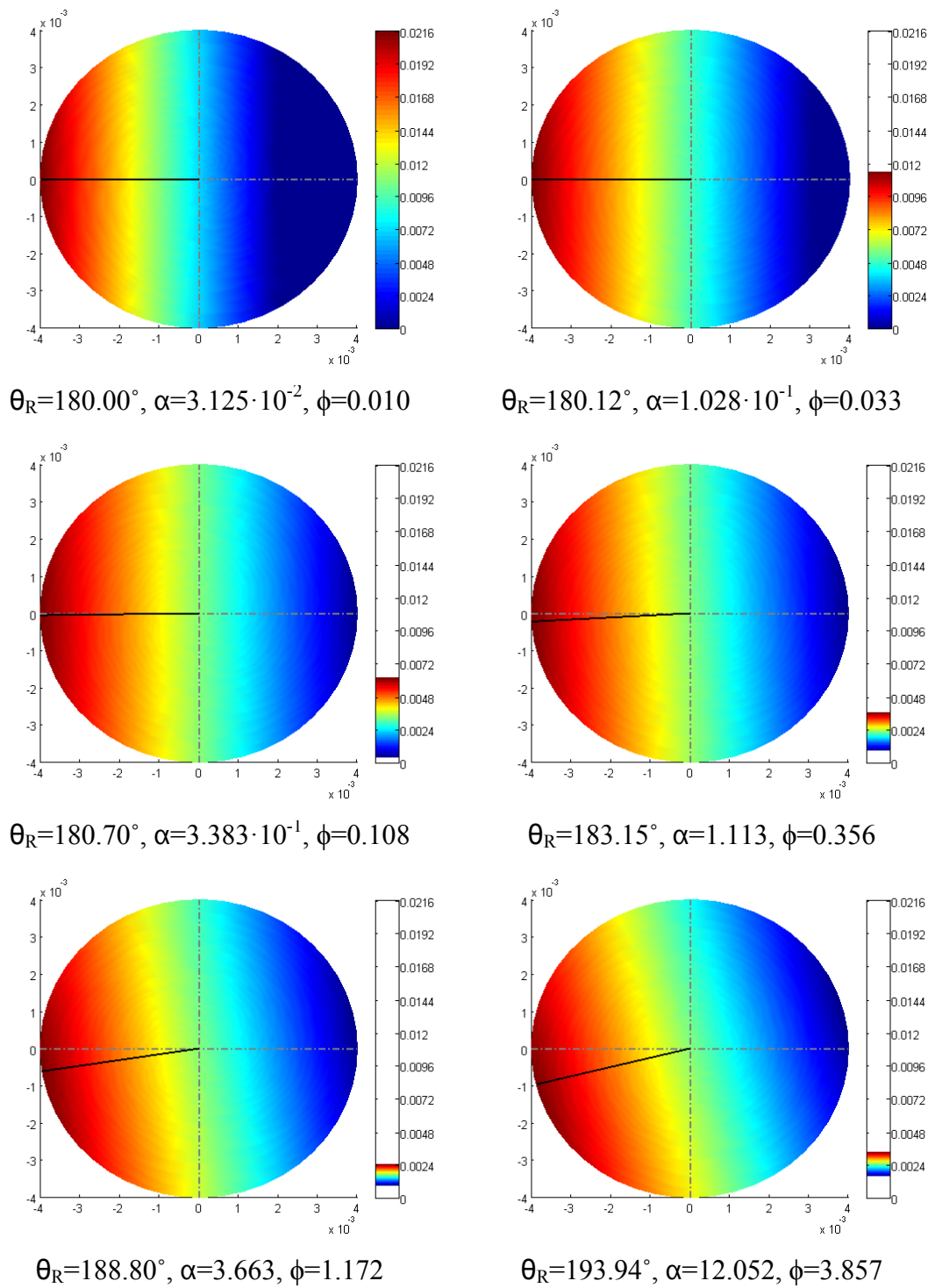
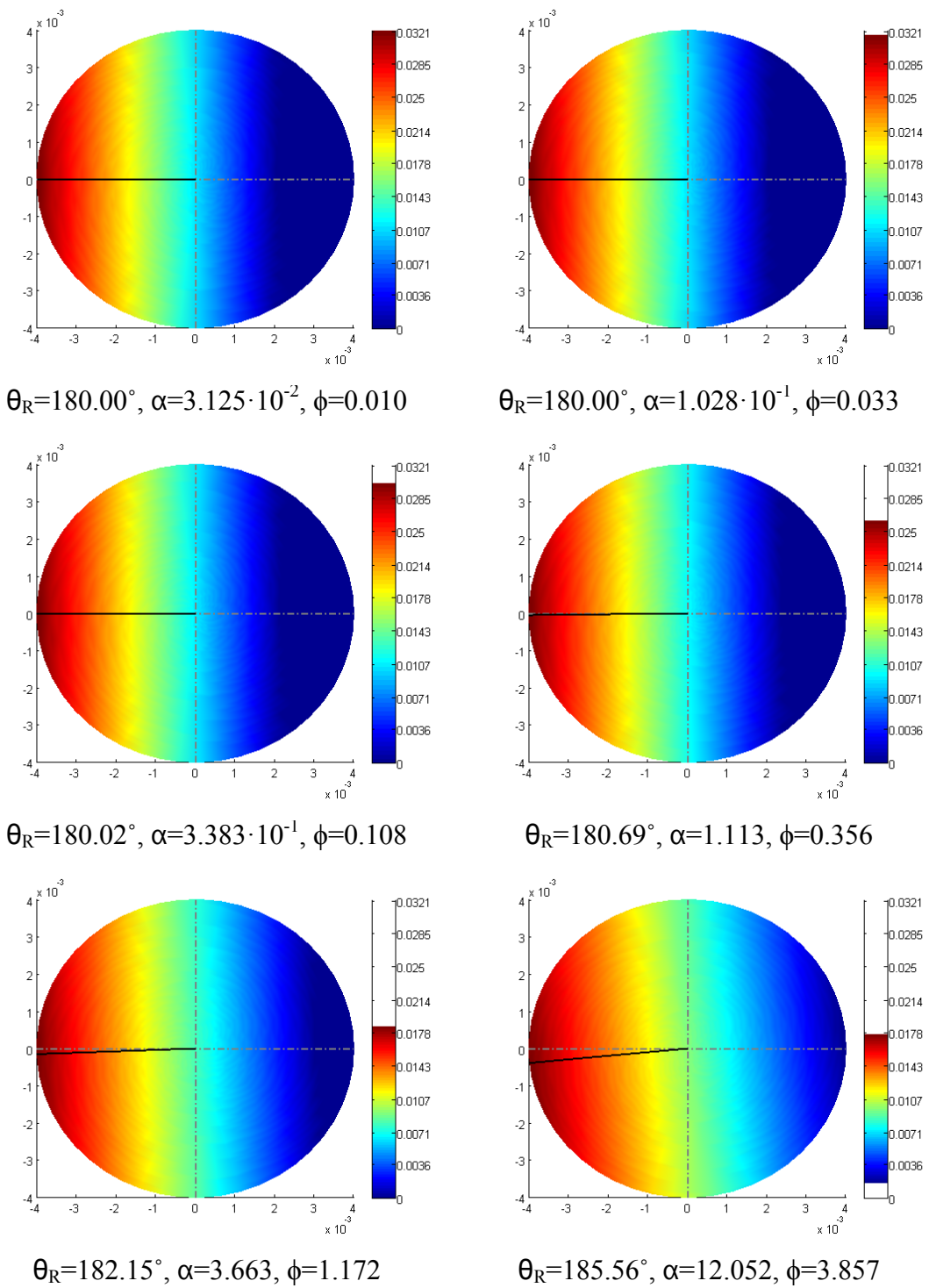
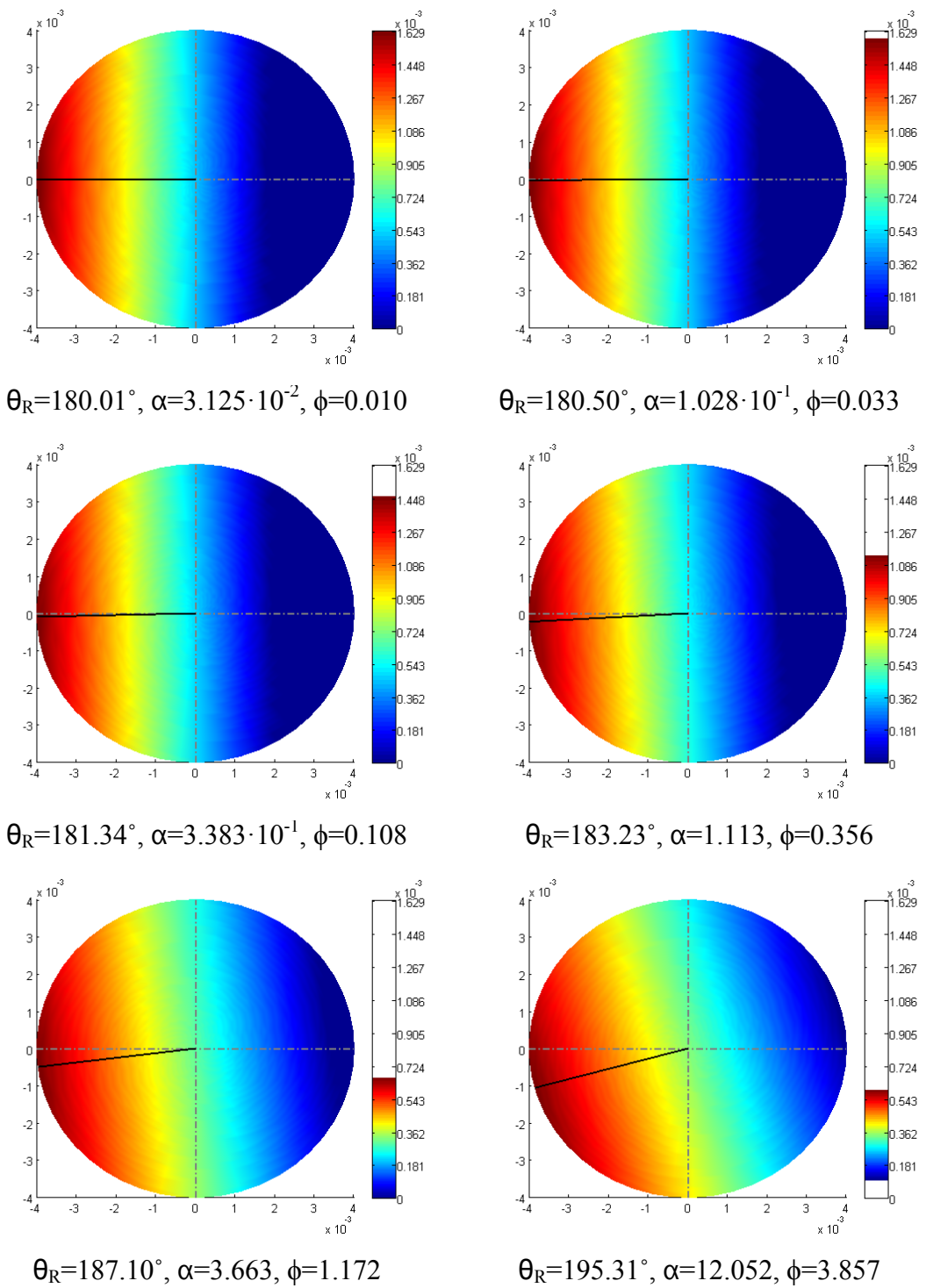
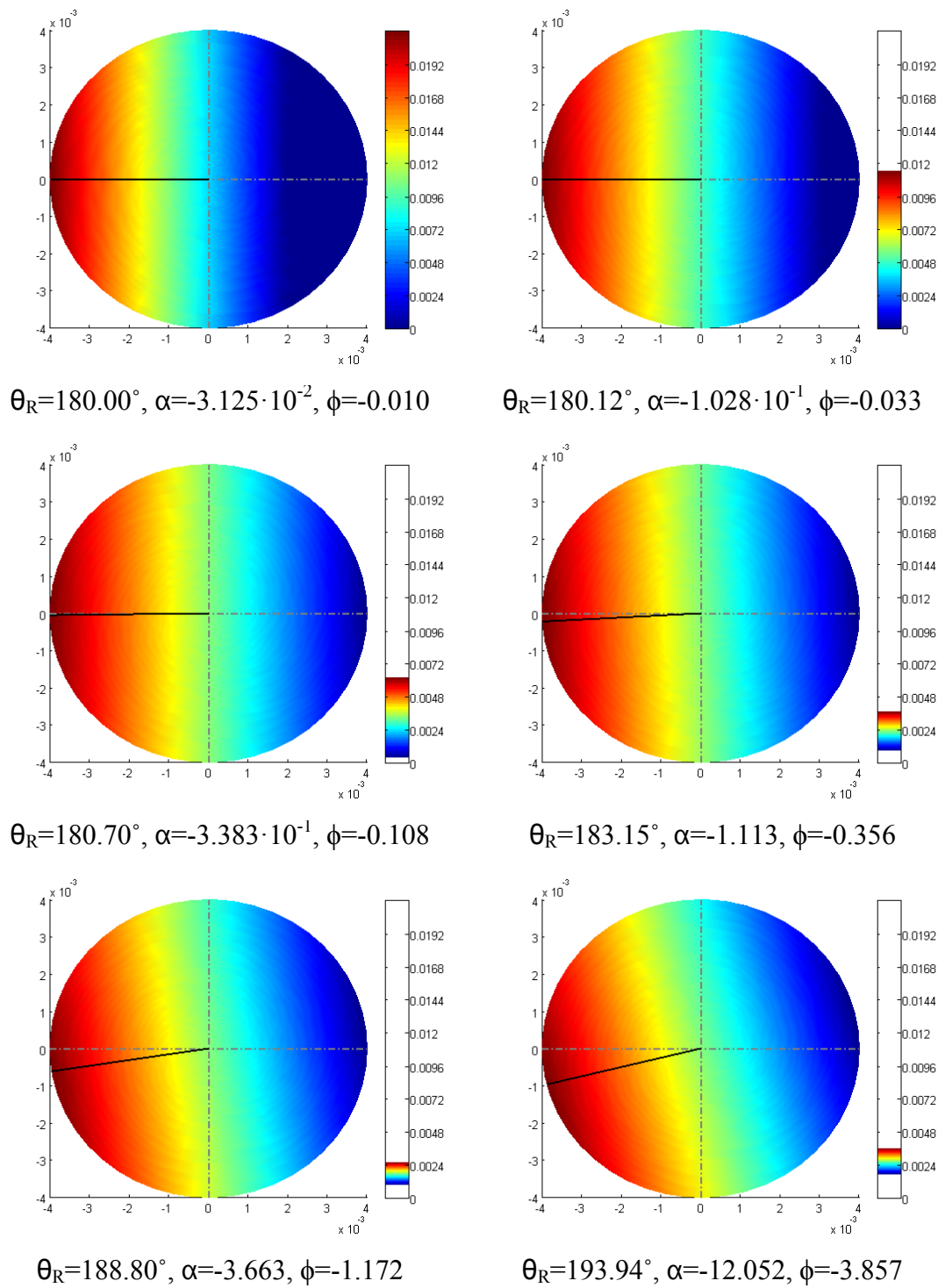


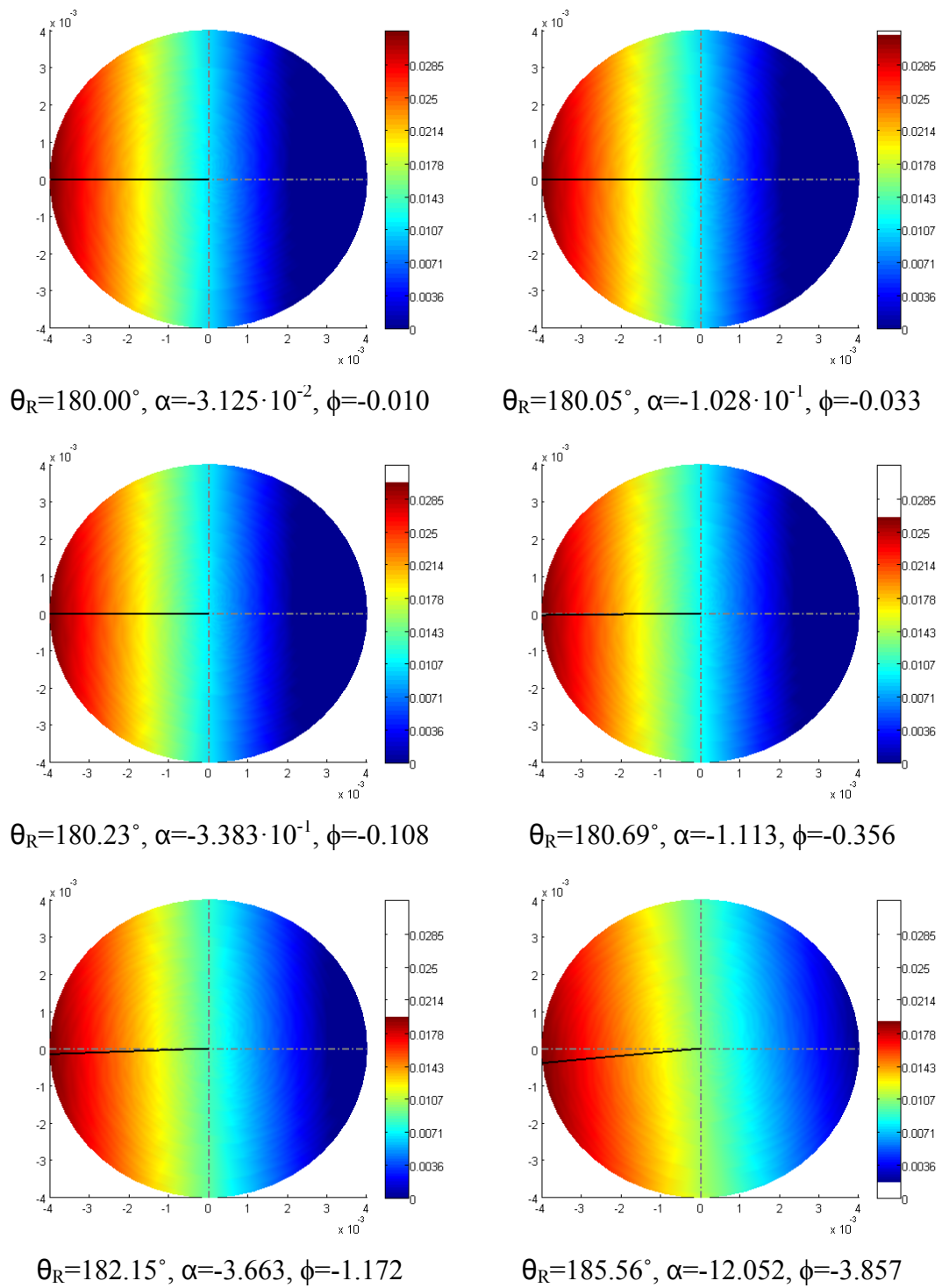
Figure 3.30. Friction torque predictions of the proposed model. Negative α values.

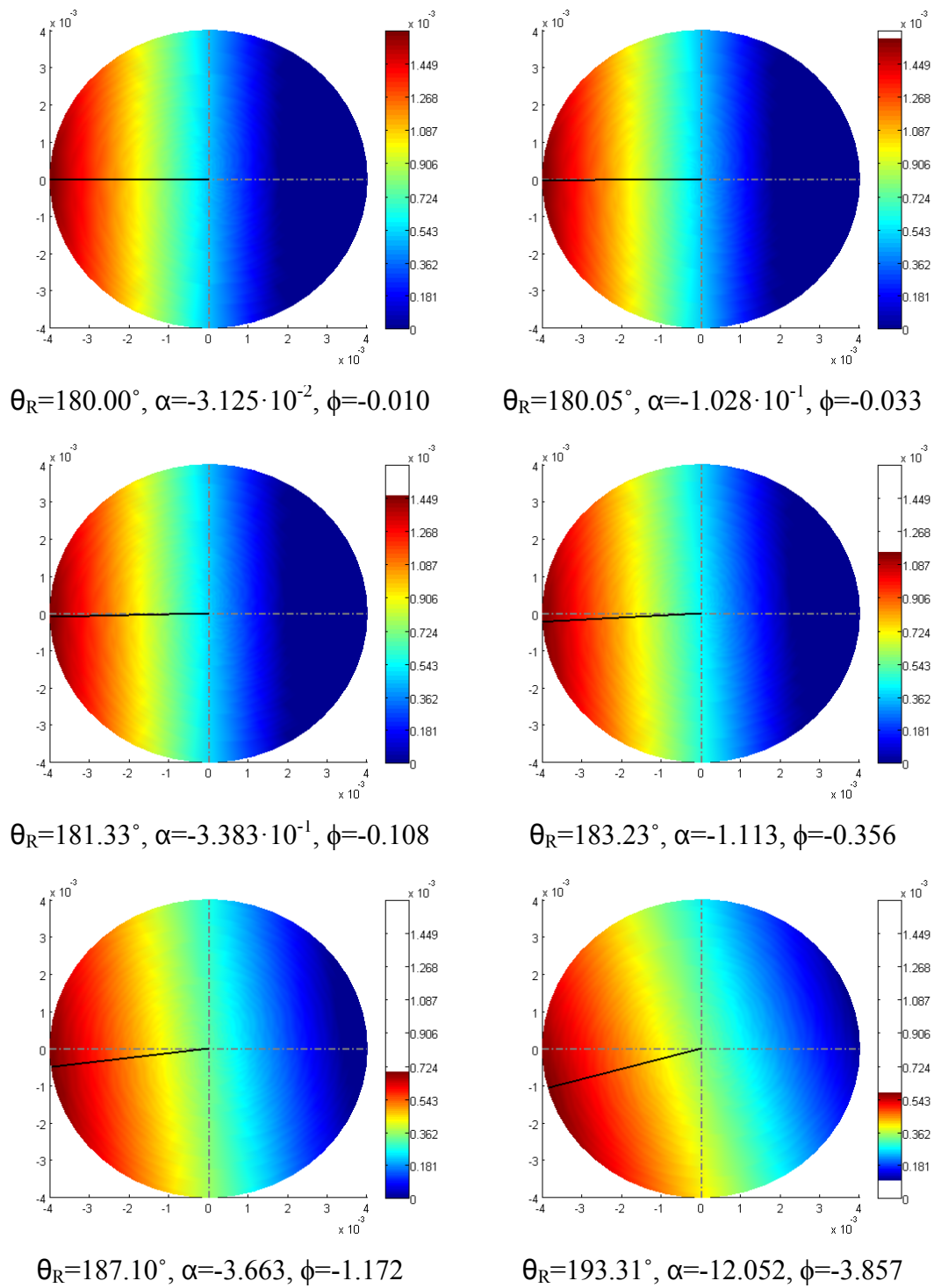
Figure 3.31. PAC predictions for PP for different ϕ . Positive α values.

Figure 3.32. PAC predictions for LDPE for different ϕ . Positive α values.

Figure 3.33. PAC predictions for PMMA for different ϕ . Positive α values.

Figure 3.34. PAC predictions for PP for different ϕ . Negative α values.

Figure 3.35. PAC predictions for LDPE for different ϕ . Negative α values.

Figure 3.36. PAC predictions for PMMA for different ϕ . Negative α values.

The proposed model results are in better agreement with the experimental data than the predictions of the previous models. The combination of a variable PAC within the area of contact for different experimental conditions ($\omega_d, \omega_p, r_p, R, h, \Theta_R$) make this model a useful tool for the evaluation of adhesive wear models.

3.4.9 Conclusions

A new model has been developed using the kinematics of the pin-on-disc apparatus, the use of the probabilistic approach for wear prediction and elastic bending theory which improves the previous study (Stolarski 1989a). This new model predicts the different components of the resultant force acting on the pin, the pin torque, the stress distribution along the contact area, the probability of asperity contact and the real area of contact distinguishing between elastic and plastic area of contact. It can also be used in electrical contact resistance (ECR) studies because it can predict the conductance using the predicted real contact area due to plastic asperity contacts.

The proposed model helps to estimate the PAC within the contact area of the pin fitting model predictions to known experimental data. This approach to the problem is classified as an inverse engineering problem. Basically, when a physical model fits experimental data, the parameters of the model can be obtained using an optimisation algorithm, i.e. fitting the model to experimental data in least mean squares sense. This strategy for solving problems is the one used for the validation of the models.

The validation process was performed using an existent experimental dataset due to the unavailability of a pin on disc machine with imposed pin rotation. This experimental dataset is valid for polymer pins and it is used to establish a comparison between the model developed by Stolarski and the proposed model. Obtained results for the proposed model are in better agreement with the experimental dataset. However, there is a considerable mismatch between experimental data and model predictions. The model can be further refined if experiments with different materials are available. New sets of experiments would help to assess the asperity contact theory and adhesion models to finally devise a

reliable physical model that can provide the pressure distribution and the plastic and elastic real area distribution within the pin contact area. Such model is of great interest to predict parameters impossible or very hard to measure that play a fundamental role in the prediction of the lubrication regime such as roughness evolution of the contacting surfaces in a pin-on-disc setup and the on-line prediction of the volume of wear and wear mechanism. Therefore, this novel model should be seen as a basic pillar for future model development that could match real results with higher accuracy.

4 Oil condition monitoring techniques

4.1 Maintenance strategies and techniques

According to the ASTM Dictionary of Engineering Science & Technology maintenance is “the act of maintaining by keeping in an existing state and preserving from failure or decline”. Maintenance is defined by the British Standard Glossary of terms (3811:1993) as “the combination of all technical and administrative actions, including supervision actions, intended to retain an item in, or restore it to, a state in which it can perform a required function”. The primary focus of maintenance is to restore an item to its original operation after it has failed or to prolong the life of equipment by not allowing it to deteriorate in the first place. From an economic point of view, maintenance is a set of organised activities that are carried out in order to keep an item in its best operational condition with minimum cost acquired (Al-Shayea 2010).

There are two basic categories of maintenance: 1) unplanned maintenance, and 2) planned maintenance. Unplanned maintenance is a reactionary measure applied to something that has already failed, whereas planned maintenance aims to avoid faults and machine failures.

Figure 4.1 shows a detailed block diagram of the maintenance techniques, a further explanation can be found in Appendix C.

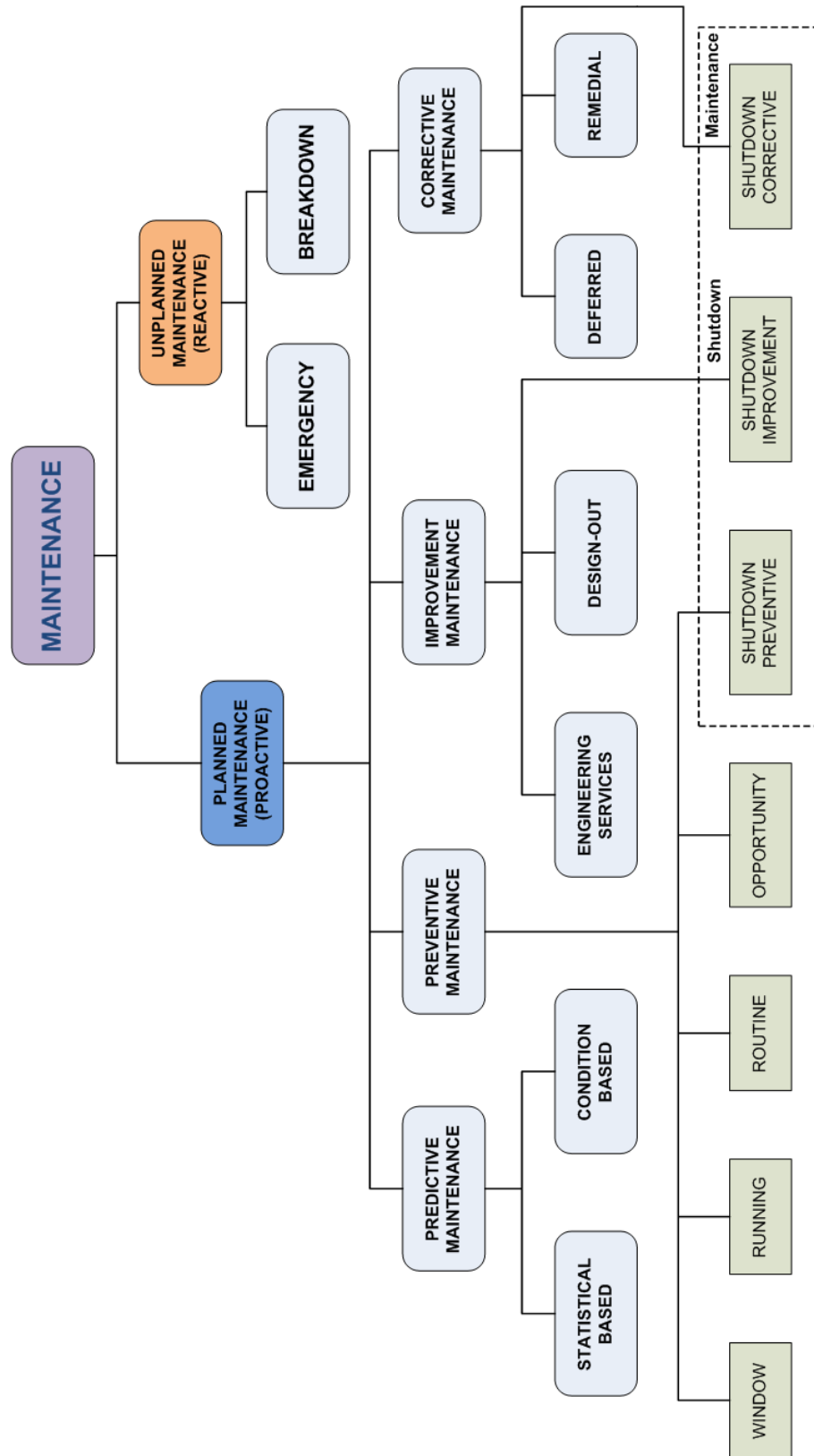


Figure 4.1. Overview of common maintenance strategies and methods (Al-Shayea 2010)

4.2 Condition-Based Predictive Maintenance

The term Condition Monitoring term is typically used in Condition-Based Predictive Maintenance (CBPM). CBPM is a condition-driven preventive maintenance programme, usually referred to by the more general term CBM (Condition Based Maintenance).

CBM provides the potential for reductions in operational costs, increases in operational performance and improvements in machine safety. Offline laboratory analysis at regular sampling intervals has traditionally provided the data for CBM programmes. Current maintenance trends are shifting towards the use of sensors to monitor the machine condition. In-situ sensors address concerns of sampling issues, time-lags from sampling to report delivery and analysis errors, particularly infield testing. Currently, considerable activity is in progress to develop simple, robust, cost effective sensors and systems for monitoring important oil condition parameters.

The main difference between CBM and preventive maintenance is that the time to failure can be estimated by monitoring mechanical conditions representative of machine status and it does not need to rely on time-driven average machine life statistics (e.g. the mean-time-to-failure or 'bathtub' curve) to schedule maintenance activities. Therefore, the main advantage when compared to preventive maintenance is that it monitors machine health and indicates the variable service intervals. Knowledge about the condition of machinery is build-identification and selecting a list of parameters that are indicative of machine condition deterioration. Depending on the application and the importance of the chosen parameter to machine health, these parameters can be measured continuously, periodically, in-line, on-line or off-line. The data gathering process starts by collecting physical data on the performance or condition of the equipment in order to optimise maintenance scheduling and to predict a failure before it occurs. The more serious the failure the more quickly the machine can be out of service for longer terms. Therefore, costly repairs and unscheduled shutdowns can be avoided if the potential problem is detected at an early stage. There are a number of condition monitoring and diagnostic techniques that have been developed over the years. While in the past all sampling and analysis had to be done off-line and in laboratory settings by highly skilled technicians, the

current trend in condition monitoring systems is to move towards automation. This includes the use of on-line transducers and instrumentation that can perform self-tests and calibrations, as well as software-based decision making tools and pattern recognition systems. Many modern systems are capable of routine monitoring, trending and evaluation of the condition of all mechanical equipment in a typical plant. These programmes can then be used to schedule maintenance on rotating, reciprocating and most continuous-process mechanical equipment.

Table 4.1 shows a classification of the different strategies or condition monitoring techniques that can be used in a CBM program.

Table 4.1. Summary of various condition monitoring methods (Appleby 2010)	
Method	Data Collected / Measurement Technique
Oil Analysis	<ul style="list-style-type: none"> ➤ Viscosity ➤ Density ➤ Complex permittivity ➤ Acid Content <ul style="list-style-type: none"> – TAN – TBN – pH ➤ Contamination <ul style="list-style-type: none"> – Metal Particles – Soot – Water – Fuel – Coolant
Wear Debris Monitoring/Analysis	<ul style="list-style-type: none"> ➤ Spectrography ➤ Ferrography ➤ Magnetic Chip ➤ Acoustic / Ultrasound ➤ Capacitive ➤ Inductive
Vibration Monitoring	<ul style="list-style-type: none"> ➤ Displacement Transducers ➤ Velocity Transducers ➤ Acceleration Transducers
Visual Inspection	<ul style="list-style-type: none"> ➤ Borescope ➤ Fiberscope ➤ Liquid Penetrates ➤ Radiography
Performance Monitoring	<ul style="list-style-type: none"> ➤ Pressures ➤ Flow Rates ➤ Fuel Consumption
Thermography	<ul style="list-style-type: none"> ➤ Contact Measurements ➤ Non-Contact Measurements (Infrared)

4.3 Oil condition monitoring

In-situ monitoring of lubricant oil quality and wear debris are important maintenance techniques within military, transportation and manufacturing industries. Costs are potentially reduced as these techniques typically allow the effective scheduling of maintenance downtime and oil changes. Although frequent oil changes are an effective way to protect machinery they may add operational expense in terms of unnecessary costs in labour, material and the impact on environmental pollution. Therefore, in most industrial machinery a well-planned maintenance strategy is critical in order to reduce equipment failure rates and costs.

It is standard practice with lubricant manufacturers to add to the base oil active extreme pressure and/or anti-corrosive additives which tend to inhibit the formation and build-up of contaminant materials. As the machines are operated, the concentration of the additives is depleted to the point where they fail to perform their inhibitory function, thus resulting in important increases in the amount of contaminant material existing within the lubricant. Furthermore, these additives are organic and organo-metallic chemical compounds which due to the operating environment and conditions of the machines may degrade into acid or basic components. Therefore, the degradation process of lubrication oil is influenced by the accumulation of metal wear particles and other contaminants like mineral acids such as sulphur, nitrogen and hydrophalic based acids, soot and water amongst others. The final result is oil whose acidity, corrosivity and viscosity have substantially increased as it reaches the end of its useful life which can result in disastrous consequences. As a result, it seems reasonable to monitor the oil quality and wear debris (Fitch 2001). They offer an important insight into the appropriate time when oil should be changed and also when mechanical repairs might be necessary (Evans and Hunt 2000; Roylance and Hunt 1999).

The analysis of oil can be performed in-line, on-line or off-line, figure 4.2 (Evans and Hunt 2000). Off-line analysis allows an accurate characterisation of the sample but it has the drawback of analysing misrepresentative samples. Real time oil monitoring is based on in-line and on-line analysis. Real time results are not as

detailed and accurate as off-line results but they reduce the problem of misrepresentative samples.

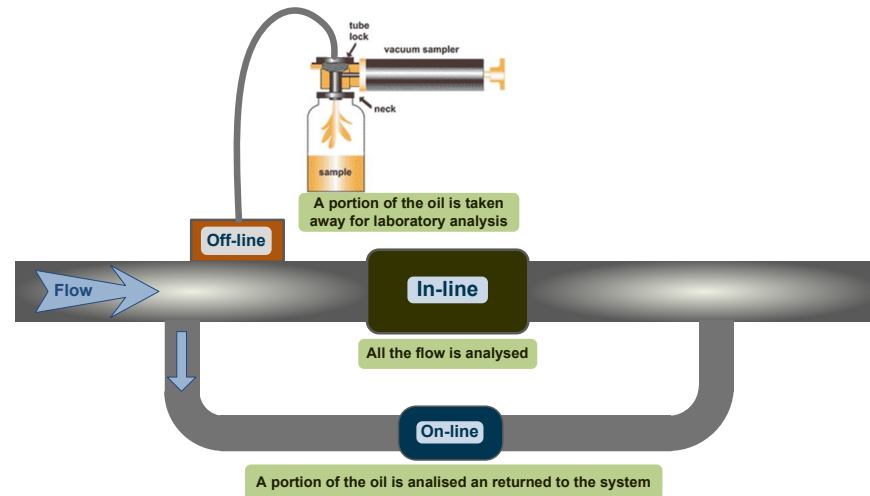


Figure 4.2. Typical connection types for oil sampling and analysis

4.3.1 Process viscometers

Viscosity is an important parameter to be monitored because it is an indicator of lubricant degradation. Viscosity changes abruptly when there is a lubricant breakdown. Therefore, measuring viscosity is very useful for characterising a new lubricant as well as following its overall performance while in service.

4.3.1.1 Rheology

Rheology studies the deformation and flow of materials considering both liquids and solids. The flow behaviour of a given fluid is expressed either by the flow curve, which is the relationship between the shear rate $\dot{\gamma}$ and the shear stress τ , or by the viscosity curve η , which is the relationship between the shear rate $\dot{\gamma}$ and the viscosity η . The two-plate model in Figure 4.3 helps to visualise these parameters $\dot{\gamma}$, τ , η .

The fluid is contained between two plates which are separated by a small gap H . The bottom plate is at rest. The upper plate with the surface area A is moved with the velocity v by applying the force F . Accordingly the fluid in between is sheared and it is required that the fluid adheres to the plates surface (no slip) and the shear flow is a laminar layer flow not a turbulent flow. If these two conditions are fulfilled

plus a small gap H , the velocity distribution within the fluid is linear and the rheological parameters are defined as shown in Equations (4-1), (4-2), (4-3).

$$\dot{\gamma} = \frac{v}{H} \quad (4-1)$$

Equation (4-1). Shear rate or velocity gradient

$$\tau = \frac{F}{A} \quad (4-2)$$

Equation (4-2). Shear stress

$$\eta = \frac{\tau}{\dot{\gamma}} \quad (4-3)$$

Equation (4-3). Viscosity

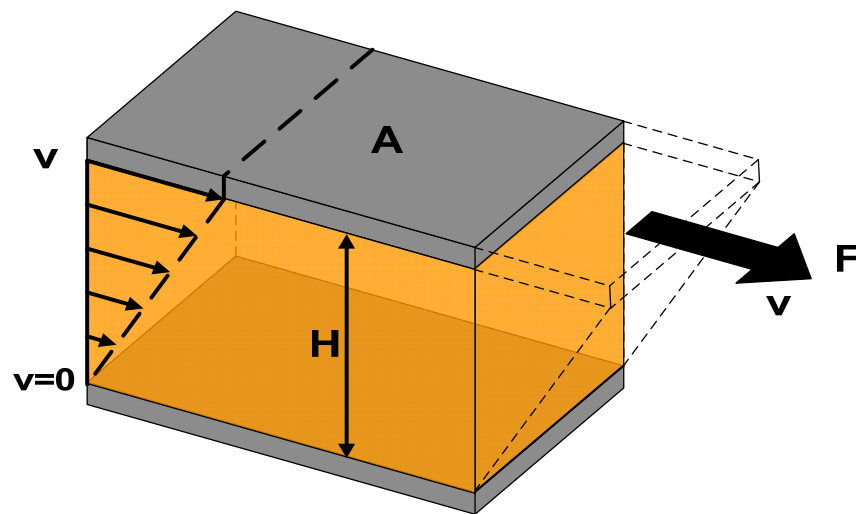
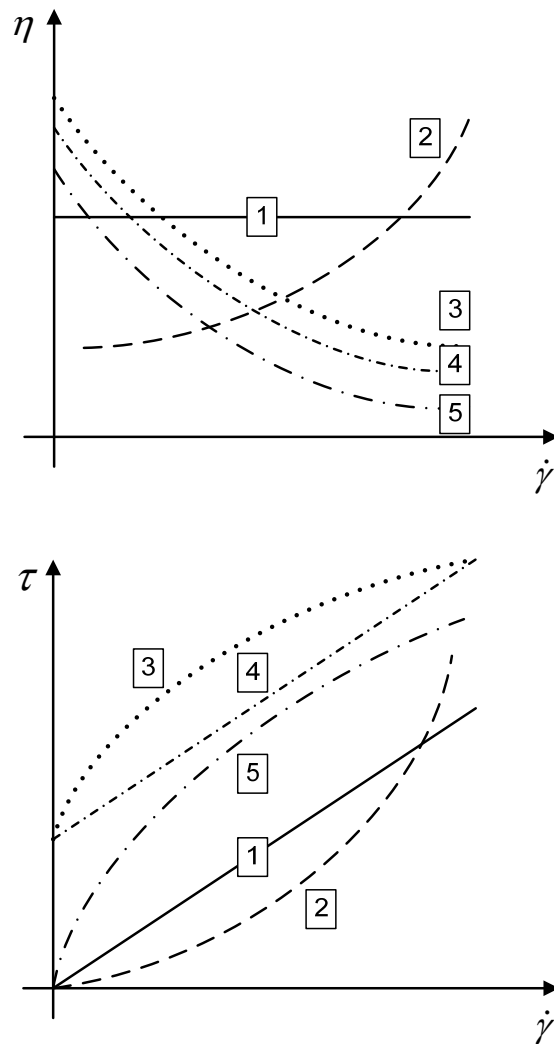


Figure 4.3. The two-plate model for the definition of the rheological parameters.

Assuming a time independent viscosity, different behavioural patterns can be obtained from the shear stress and viscosity curves depending on the range of shear rate considered and the fluid type, Figure 4.4.



1	Newtonian fluid $\eta = \text{constant}$
Non-Newtonian fluids $\eta \neq \text{constant}$	
2	Power Law Fluid (PLF) with shear-thickening behaviour
3	Yield Stress Fluid (YSF) with shear-thinning behaviour
4	Bingham Fluid: Simplified Yield Stress Fluid (YSF)
5	Power Law Fluid (PLF) with shear-thinning behaviour

Figure 4.4. Simplified classifications of fluids based on the characteristic of the viscosity curve (top) and the flow curve (bottom).

For **Newtonian fluids** the flow curve is described by a linear function, equation (4-4).

$$\tau = \mu \cdot \dot{\gamma} \quad (4-4)$$

Equation (4-4). Newtonian fluids (constant μ)

where μ is the Newtonian viscosity. Lubrication oils are an example of Newtonian liquids at low shear rates; at very high shear rates (within the range 10^6 - 10^7s^{-1} , typically found in EHL contacts) additivised oils tend to behave as Non-Newtonian liquids.

Depending on the flow curve, Non-Newtonian fluids are classified into Power Law Fluids (PLF) and Yield Stress Fluids (YSF). The PLF relationship is shown in equation (4-5),

$$\tau = m \cdot \dot{\gamma}^n \quad (4-5)$$

Equation (4-5). Non-Newtonian fluids. Power Law Fluids (PLF)

where m is the power-law consistency coefficient and n is the power-law index. If n is less than one the power-law fluid behaves as a shear-thinning fluid and if n is greater than one the power-law fluid behaves as a shear thickening fluid.

Yield Stress Fluids (YSF) are characterised by the presence of a yield stress τ_y , which expresses the stress that must be overcome to set a fluid into motion. If the applied stress is lower than τ_y no fluid motion takes place. The flow curve of a YSF is usually described either by a Herschel-Bulkley or a Casson model.

The Herschel-Bulkley model function is shown in equation (4-6).

$$\tau = \tau_y + m \cdot \dot{\gamma}^n \quad (4-6)$$

Equation (4-6). Non-Newtonian fluids. Herschel-Bulkley

where m is the Herschel-Bulkley consistency coefficient and n is the Herschel-Bulkley index. A simplification of the Herschel-Bulkley model is the Bingham model. This continuous until n is one and m reduces to $m = \mu_B$, where μ_B is the Bingham viscosity.

The Casson model function is shown in equation (4-7).

$$(\tau)^{0.5} = (\tau_y)^{0.5} + (\mu_c \cdot \dot{\gamma})^n \quad (4-7)$$

Equation (4-7). Casson model

where μ_c is the Casson viscosity parameter.

YSF typically described by the Herschel-Bulkley model are sediment water mixtures rich in fine material. YSF typically described by a Casson model are yoghurt, tomato purée, molten chocolate or blood.

4.3.1.2 Devices to measure viscosity

There are many different devices to measure viscosity. Below is a list of the most common ones (Dinser 2009):

1. *Capillary rheometers* that are primarily used for Newtonian fluids. They are also called U-tube viscometers or Ostwald viscometers.
2. *Falling sphere viscometers* that are based on Stokes Law.
3. *Piston viscometers*. There are two types: the falling piston invented by Austin Norcross and the oscillating piston or electromagnetic Viscometer invented by Cambridge Viscosity Corporation.
4. *Rotational rheometers* with various geometries.
5. *Resonator techniques*, amplitude or phase measurement.
6. *Impedance measurement* of a mechanical system in contact with a fluid. Mechanical impedance is a measure of how much a structure resists motion when subjected to a given force. The transfer function of a mechanical system is measured without and with contact with a fluid. Their comparison together with an impedance model allows the determination of the viscosity.

From the list of rheometers shown above two are of special interest, rotational rheometers (typically available in laboratory facilities) and resonator-based rheometers the most common type in process monitoring.

Rotational rheometers are used for Non-Newtonian and visco-elastic fluids. They basically consist of a vertical shaft driven by an electro motor. The fluid is placed between a resting bottom plate and a measuring geometry fixed to the shaft, e.g. a plate. A rotation of the shaft creates a shear field in the fluid which is then used for viscosity determination. For higher accuracy, a bearing with minimal friction is

chosen. In most cases, this is an air bearing. With the help of different geometries - plate-plate, cone-plate, cylinder-cylinder the fluid is sheared in a small gap. The goal is to have a constant shear rate throughout the whole fluid sample. This permits the measurement of the shear rate dependent viscosity or the frequency dependent complex viscosity (Dinser 2009).

Resonator based rheometers are based on the measurement principle of resonators which are based on changes of two characteristics parameters, the resonant frequency and the damping or Q factor. If the mechanical structure of the resonator is brought into contact with a fluid or solid medium both resonance frequency and damping are changed. Both depend on the viscosity and the elasticity of the fluid. For a fluid the motion of the resonator induces a sinusoidal shear in the fluid, therefore the damping is increased due to the viscous shear stress acting on the resonator. The characteristics can be observed by different methods. The amplitude or the phase between excitation and response are measured in the vicinity of the resonant frequency. Resonators have been used primarily in process monitoring.

Finally, a recent review of methods for on-line monitoring of viscosity of Lubrication Oils is reported by (Markova et al. 2010). The integration and miniaturisation of current sensor technology is an active field in research, and several viscometers exist based on microstructures such as passive capacitors and SAW or TSM resonators (Farone et al. 2002). An interesting sensor for oil condition monitoring of engines is based on a basic tuning fork resonator by Hella KGaA Huck. Hella's sensor is not in production but it claims to provide all variables of interest to determine the condition of the oil: viscosity, density, temperature, and permittivity and conductivity in a highly integrated and robust package, a more detailed description can be found in the following patents (Kolosov et al. 2010; Leonid Matsiev 2004). A brief comparison between the last sensor and all miniaturised oil quality sensors can be found in (Buhrdorf et al. 2005).

4.3.1.3 Principles of measurement of AST100

The viscometer use in this research is the Brookfield AST100. This viscometer is a rod-type resonator intended for applications in the field of industrial fluid process control and monitoring (Dual et al. 1990), Figure 4.5.

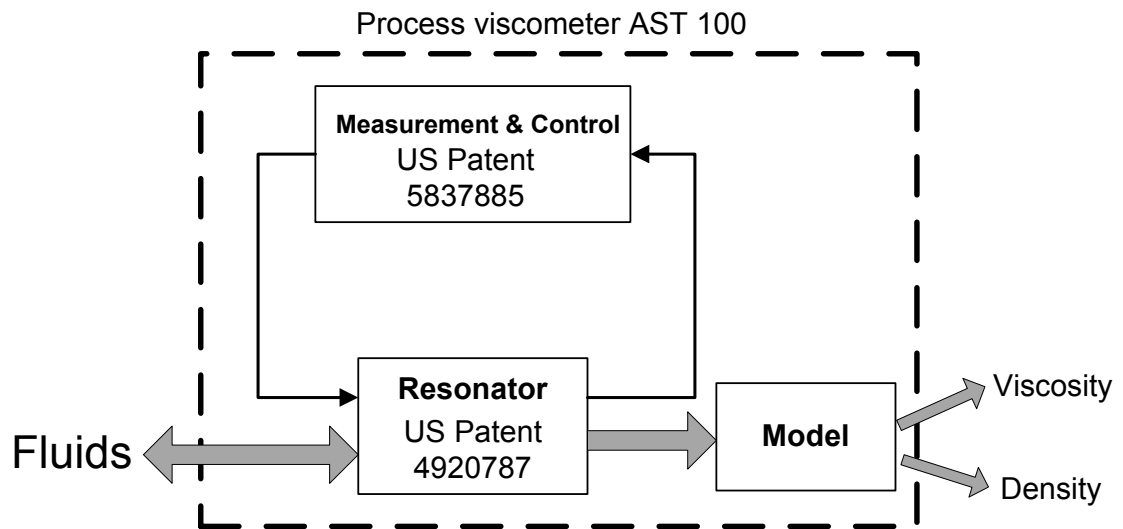


Figure 4.5. Conceptual diagram of AST100 viscometer

A rod type resonator consists of a rod which is excited at one end electromagnetically in torsion. The other end of the rod is attached to a tube exciting it in a torsional mode. This assembly is plunged into the fluid sample.

The mechanical resonator is characterised by its resonance frequency and the damping or Q-factor. If a resonator is brought into contact with a fluid or solid medium both resonance frequency and damping are changed.

The motion of the resonator induces a sinusoidal shear in the fluid; therefore the damping is increased due to the viscous shear stress acting on the resonator.

The damping effect of the mechanical resonator can be determined from the slope of the phase curve near the resonance with great accuracy and stability (Goodbread et al. 1998). The viscosity of the fluid is obtained from the damping of the resonator by modelling the fluid motion in the vicinity of the resonator (Dual 1989; Valtorta 2007). The slope $\Delta f/\Delta\Phi$ of the phase curve near resonance is a measure of damping, Figure 4.6. The theoretical model relates the damping to the viscosity of the fluid following equation (4-8).

$$\eta = K \cdot \Delta f^2 \quad (4-8)$$

K is a constant determined in the calibration process

Equation (4-8). Viscosity relationship for the rod-type viscometer

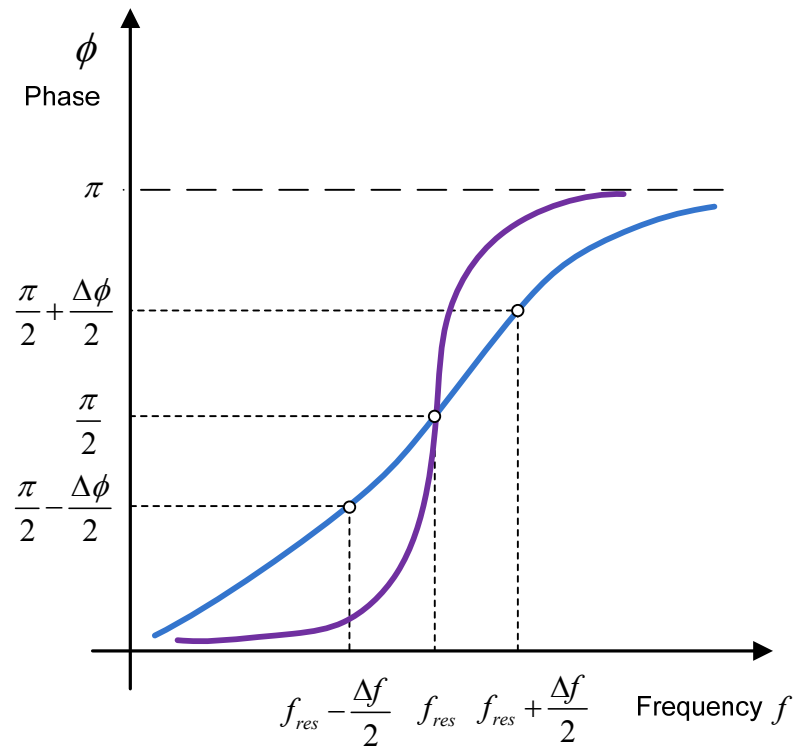


Figure 4.6. Phase shift vs. excitation frequency for two different damping values

The AST100 viscometer has an RS232 output for data gathering and every second the following parameters are transmitted, Table 4.2.

Table 4.2. AST100 data frame transmitted every second (RS232)

Order	Name	Data length (bytes)	Name of the variable	Intended use
1	V	13	Viscosity (cSt)	Reading
2	T	11	Temperature (°C)	Reading
3	cF	11	Central frequency	Calibration and Diagnosis
4	dF	9	Differential frequency	Calibration and Diagnosis

4.3.1.4 Commercial alternatives to AST100 viscometer

There are many types of process viscometers based on resonators. Promising process viscometers for oil condition monitoring for its stiffness and immunity to vibrations are acoustic viscometers based on piezoelectric transducers (SAW or TSM) (Agoston et al. 2005). ViSmart products from Vectron International are leading products for engine oils and are based on quartz resonators. However, for Non-Newtonian fluids, as the shear rate where viscosity is measured is very high (depends on the excitation frequency of the piezoelectric transducer) results of viscosity could be non representative of viscosity readings in most applications (at low shear rates). As a result, to achieve the best accuracy there is the need to correlate acoustic viscosity readings with real viscosity ones using a Non-Newtonian model. The final accuracy depends upon the reliability of the Non-Newtonian model which limits the practical viscosity detection tolerance. For example, ViSmart products provide unit AW (acoustic wave units). These units are quite accurate at the measuring shear rate, and they are proportional to kinematic viscosity at the shear rate of the transducer (excitation frequency). The basic calibration process for this sensor consists of providing the kinematic viscosity of the oil at a certain temperature and the density for absolute viscosity readings. However, the main problem arises if the fluid being tested behaves as a Non-Newtonian fluid at the transducer shear rates (for example, an excitation frequency of 1MHz implies a shear rate 10^6s^{-1}) which makes it difficult to extrapolate acoustic wave units to viscosity

because viscosity would be a function depending on the shear rate. A custom calibration depending on the working fluid with a suitable Non-Newtonian model must be the best way of sensing viscosity using this type of sensor. For mineral oils that are Newtonian, these sensors work very well. However, it should be noted that the viscosity behaviour at shear rates within the range 10^6 - 10^7 s^{-1} (transducer excitation frequencies 1MHz to 10MHz) for most types of lubrication oils based on mineral based oil with additives presents a Non-Newtonian behaviour (Torres 2010).

Kittiwake is developing a process viscosity sensor which is a resonator based on a smart and quite old patent (Roth and R.Rich 1953, 1958a, 1958b). A block diagram and operation of the Kittiwake sensor can be found in (Baldwin and Lunt 2010). The main advantage of this sensor when compared to the rod-type resonator is that it is a more cost effective solution but the main disadvantage is the immunity to vibrations that is not as good as rod-type resonators such as the Brookfield AST100 process viscometer. This last shortcoming can limit its application in some machinery. For laboratory facilities, this type of viscometer can be a low cost alternative to other types with the advantage of minor maintenance and the possibility of using different probes to characterise viscosity at different shear rates.

Before selecting the AST100, another two process viscometers SPC/L372(J) Flow-Through, SPC/L571 OILSENSE™ Miniature from Cambridge Viscosity were shortlisted because the sensing principle is based on a simple and reliable electromagnetic concept. A piston is magnetically driven back and forth by two coils at a constant force. From the piston's two-way travel time the absolute viscosity can be obtained. The shear rate is not controlled by this way of sensing but the shear is low enough to monitor lubricant at low shear rates.

4.3.1.5 The effectiveness of process control viscometers in hydraulic oils contaminated with water

Oil samples were prepared blending 0.5ml of water and 1l of oil (500ppm in volume) and 2ml of water and 1l of oil (2000ppm in volume). These were stored in a lab bottle. The type of oil is ESSO NUTO H46, which is a commercial hydraulic oil of viscosity grade ISO VG46. Although the bottle was properly stirred for a long period of time the emulsion was not stable at room temperature (the solubility of the

oil is less than 100ppm at room temperature) and free water in the oil must be present. The oil was pumped to the measurement chamber of the AST100 using a peristaltic pump and the fluid temperature was controlled using a thermoelectric cooling plate. According to the specifications of AST100, a reproducibility unit to unit of $\pm 5\%$ of the reading implies that the viscosity readings between experiments under the same conditions can vary by $\pm 5\%$ of the real value. Reproducibility should not be confused with the stability of the reading that means the output value of the sensor during one experiment can vary within $\pm 1\%$ range.

Figure 4.7 shows the viscosity of the samples at three different temperatures and the theoretical boundaries of viscosity values according to the reproducibility of the AST100. The real viscosity curve is determined using the Andrade's viscosity-temperature relationship. According to this result, viscosity monitoring is not effective for small concentrations of water in this particular type of hydraulic oil.

The readings of the AST100 were validated with a rotational viscometer available in the lab facilities. Considering the oil has a constant density of 876g/m^3 the kinematic viscosity for the oil at 25°C was 91.5 cSt according to the Brookfield DVII pro viscometer. The difference between measurements of these two viscometers is approximately 1 cSt .

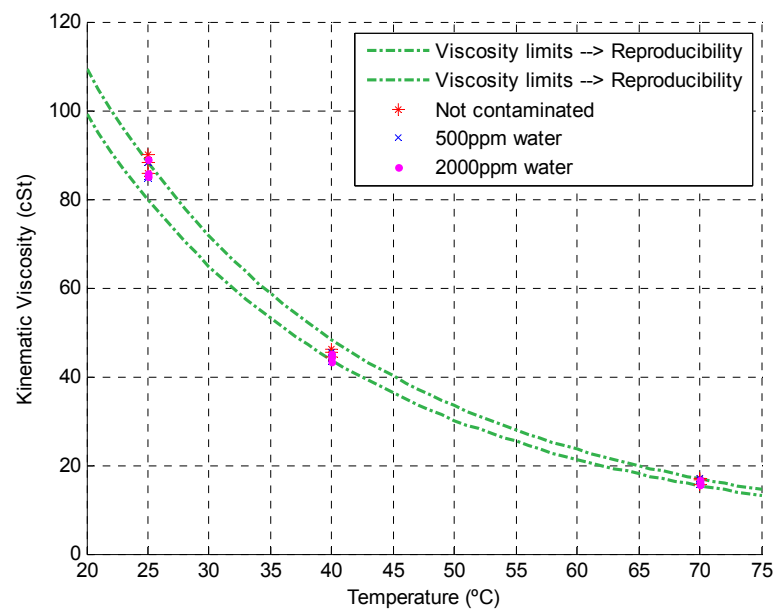


Figure 4.7. Viscosity results for different water in oil concentrations

4.3.2 Sensors to monitor the water content in oil

Water in lubrication oils is presented in two states: dispersed (water content below the saturation point) and free water. As long as water is dispersed within the oil water is not harmful to machinery but free water is not desirable because it causes corrosion, absorbs or consumes additives in the oil and shortens bearing life and the service life of the entire machine. Free water promotes oxidation of the lubricant's base oil and washes out some additives which are attracted to water. Later, water will typically separate to the bottom of the sump. It hydrolyses (chemically attacks) additives, which compromises their performance, and in some cases, produces highly-corrosive by-products. A water-degraded lubricant cannot fully lubricate and protect the machine, which leads to excessive wear and failure. Water also attacks the machine directly. The following is a summary of common water-induced wear mechanisms: Rust and Corrosion, Vaporous Cavitation, Film strength loss (Troyer 2001).

Every fluid has the ability to hold a certain amount of dissolved water. The maximum amount of water that a given fluid can contain in solution is referred to as its saturation point. Once the fluid has reached its saturation point, any additional water introduced will separate out as free water by forming a distinct layer. Since most oils are less dense than water, the water layer will usually settle below the oil. An oil's saturation point is a function of many different factors such as the composition of its base stock (mineral or synthetic) as well as the type of additives, emulsifiers, and oxidising agents present. Aside from these initial composition differences, the saturation point of an oil will vary over its lifetime as a working fluid. Two major factors that influence an oil's saturation point as it ages are fluctuations in temperature as well as changes in chemical make-up due to the formation of new substances produced as by-products of chemical reactions taking place within a dynamic oil system (Jiroutek 2007).

Typical methods to monitor water contamination in the lubrication industry are:

- Karl-Fisher titration method. Laboratory based method.
- Infrared spectroscopy (FTIR-Fourier transform infrared spectroscopy). Laboratory based measurements and on-line measurements of total water in oil.
- Water activity sensors (aw). They provide an output relative to the saturation point. When the medium is air, they are also called relative hygrometers. On-line measurements of the relative percentage to saturation point.
- Dielectric measurements.

Karl-Fisher titration method is a standard method for detecting water in lubrication oils (ASTM D6304). For the contamination levels of water in oils, coulometry based methods with generator electrodes with diaphragm are used. This is because of the accuracy of the “electronic buret” based on Faraday’s law so the current applied with the time (charge) releases the stoichiometrically corresponding amount of iodine from the iodine-containing KF reagent by electrolysis. From the Karl Fisher reaction in methanolic solution, 1 mol of iodine reacts with 1 mol of water. When the end point of the titration is detected, the amount of water can be determined. The end point of the titration is detected by bivoltametric indication. For a further explanation of the Karl Fisher Method, and considerations of how to apply it to petrochemical products the reader is referred to (Bruttel and Schlink 2006).

Karl Fisher works very well in theory but in reality it requires a lot of expertise and proper equipment to measure water in oil with an accuracy of +/-50ppm. This is because it is hard to detect the end point of the reaction and moist contamination should be removed from the titration cell which requires extreme care and time.

Infrared spectroscopy methods for water content monitoring are based on the principle that water absorbs strongly in the IR portion of the spectrum due to the O-H stretching and bending vibrations. However, it is difficult to quantify because of the spectral interferences from other OH containing constituents. In new and used

oils, many additives, breakdown products and contaminants containing OH groups are commonly present, making it difficult to predict moisture especially at moisture levels less than 100ppm. The solubility curve of oils depends on temperature and lubricant degradation. For new mineral based oils the solubility of water can be around 100ppm at ambient temperature. A preliminary test to enable a guess to be made of the solubility levels of water in hydraulic oils was performed in University of Oviedo using a FTIR spectrometer. The first step was the characterisation of the minimum detection of water by the FTIR spectrometer. The characterisation is performed introducing known amounts of water to the base oil and ensuring the miscibility of the emulsion with a Vortex. It has been found that the minimum detection capability was between 100 and 200 ppm. After characterising the detection capability of the equipment, a saturated water in oil sample was prepared mixing oil and water at 50 % in volume. By centrifugal action the water was separated from the oil and the oil in the top was extracted using a pipette. This oil was saturated with water at an ambient temperature 20°C. The spectrum of the oil was measured and the absorption peak was slightly less than the one obtained with the known contamination of 150ppm. Therefore, the solubility level of the test oil was in the range of 100ppm at an ambient temperature of 20°C.

Alternative approaches to the direct determination of moisture in oils is the extraction of the moisture into a suitable dry solvent or reaction with the water with an appropriate reagent followed by IR quantitation of the reaction products. However, attempts to use absolute Methanol to extract water from mineral oils proved unworkable. Other authors propose monitoring the products of stoichiometric reactions as an alternative technique. The underlying idea is to react 2,2-dimethoxypropane (DMP) with water under acidic conditions to produce acetone. Then the absorbance is measured by IR spectroscopy at 1703 cm^{-1} . Dong studied the water content in oil using the reaction of DMP with water (Dong 2000) and explained the calibration method and using Partial Least Squares (PLS) regression. He concluded that it is possible to measure water content with a higher precision than KF volumetric and coulometric methods. An uncertainty of +/-35ppm can be reached using Dong's method but the sample preparation is quite time consuming. Process equipment based on IR absorption monitors the total water content with the

minimum detection capability of several hundreds of ppm. Kittiwake is developing a total water content in oil sensor (WaterScan) based on two infrared detectors (one monitoring the water absorption at the wavelength 3400 cm^{-1} , and the other the soot content at the wavelength 2000 cm^{-1}) (Baldwin and Lunt 2010; Robinson 2000). There are other commercial units for on-line monitoring of water levels based on the Michelson Interferometer (FT-IR) such as the PAL series of portable Fourier Transform Infrared manufactured by A2 Technologies.

Water activity sensors (aw). Moisture Related to Saturation

Water activity is derived from fundamental principles of thermodynamics and physical chemistry. Several requirements must be met to define water activity. These requirements are: pure water ($aw=1.0$) is the standard state, the system is in equilibrium, and the temperature is defined. It is usually assumed that under normal working conditions of ambient temperature and atmospheric pressure, gas phases behave ideally and so the ratio of fugacities can be taken as the ratio of partial pressures (Reid 1987). Therefore, water activity is the amount of water in a substance relative to the total amount of water it can hold. It is defined as the ratio between the partial pressure of water in the oil and the saturated vapour pressure of pure water at the same temperature.

Capacitive probes consist of two electrodes with a dielectric made by a hygroscopic polymer. Water molecules penetrate into the polymer depending on the water content relative to saturation level of the oil. Other molecules than water may likewise diffuse into the polymer and thus change the capacity, leading to a measurement error. This influence is considered to be small, since e.g. ageing by products have a low permittivity compared to water. Careful considerations must be taken if some material is condensed on the polymer because it may inhibit the water molecules diffusing in the polymer.

Water penetration changes the capacity C because of the high relative permittivity of water $\epsilon_r = 80$. Two types of circuits are typically used with this type of sensor depending on the type of output: frequency or voltage. A frequency output circuit is typically based on an astable multivibrator where the sensing element is a variable capacitor (the humidity sensor). A voltage output is obtained using an

astable multivibrator (pulse generator) that continuously triggers a monostable. The output of the monostable is a pulse which is proportional to the sensor capacity. The generated PWM signal is later low-pass-filtered to provide a voltage output.

There are several commercial humidity sensors based on capacitive type polymers. Some of the sensors provide all the signal conditioning circuitry in a very small package with direct digital interface (Humidity Sensors manufactured by Sensirion), others provide a voltage output (Honeywell Sensors), and another needs a small circuit for signal conditioning such as the previously mentioned. (Measurement Specialist, Honeywell).

Research with thermoset polymer-based capacitive sensors has also demonstrated that the RH calibration in air applies to relative saturation measurements in oil within 0.3% (2010). Kittiwake's water in oil sensor is based on a SHT15 manufactured by Sensirion. Water in oil sensors based on capacitive probes has been extensively used to monitor the water in oil levels in transformer oils.

Dielectric measurements. Oil and water have very dissimilar dielectric properties, the typical value of oil is between 1.7 and 2.1 and water has a typical value of 80. So water can store much more energy per volume than average oil. Any device based on monitoring the changes of the dielectric constant can detect the presence of water in oils. The simplest and cheapest circuit that detects changes of capacitance is based on using an RC circuit (astable multivibrators). A further explanation of different strategies to monitor impedance is covered in the next section. One commercial sensor for total water in oil measurements based on this principle is EASZ-1 (EESIFLO).

4.3.3 Tan Delta and Impedance Spectroscopy sensors

4.3.3.1 Introduction to Impedance Spectroscopy (IS) in lubrication oils

The complex permittivity of lubrication oils changes with use mainly because of the process of oxidation and degradation of additives which is affected by the presence of contaminants such as water, soot particles, acid combustion products, glycol, ferrous and non-ferrous metallic particles (Collister 1998).

The degradation of most oils implies the generation of molecules that are generally more polar than the previous ones. The base oil consists of large hydrocarbon molecules that are generally weakly polar, so the presence of most contaminants results in an increase of one or both parts of the oil's complex permittivity. Depending on the geometry of the electrodes, the permittivity is directly related to complex impedance. As a result, any measurement of complex impedance is also an indicator of the degradation of the oils.

In analytical chemistry the technique where the electrode impedance is assessed as a function of frequency is commonly referred to as electrochemical impedance spectroscopy (EIS). The underlying idea of Impedance spectroscopy is the measurement and characterisation of a material-electrode system.

Figure 4.8 shows a flow diagram for a complete impedance spectroscopy (IS) study whose goal is the characterisation of important properties of the material-electrode system from its impedance vs frequency response. The experimental data is denoted by $Z_d(\omega)$, the impedance predicted by a theoretical fitting model by $Z_m(\omega)$, and that of a possible electrical equivalent circuit by $Z_{ec}(\omega)$, where $\omega = 2\pi \cdot f$ and f is frequency.

A complete IS analysis involves more than a single set of measurements of impedance vs. frequency. Frequently, full characterisation requires that such sets of measurements be carried out over a range of temperatures and/or other externally controlled experimental variables. IS characterisation may be used to yield basic scientific and/or engineering information on a wide variety of materials and devices, ranging from solid and liquid electrolytes to dielectrics and semiconductors, to

electrical and structural ceramics, to magnetic ferrites, to polymers and protective paint films, and to secondary batteries and fuel cells.

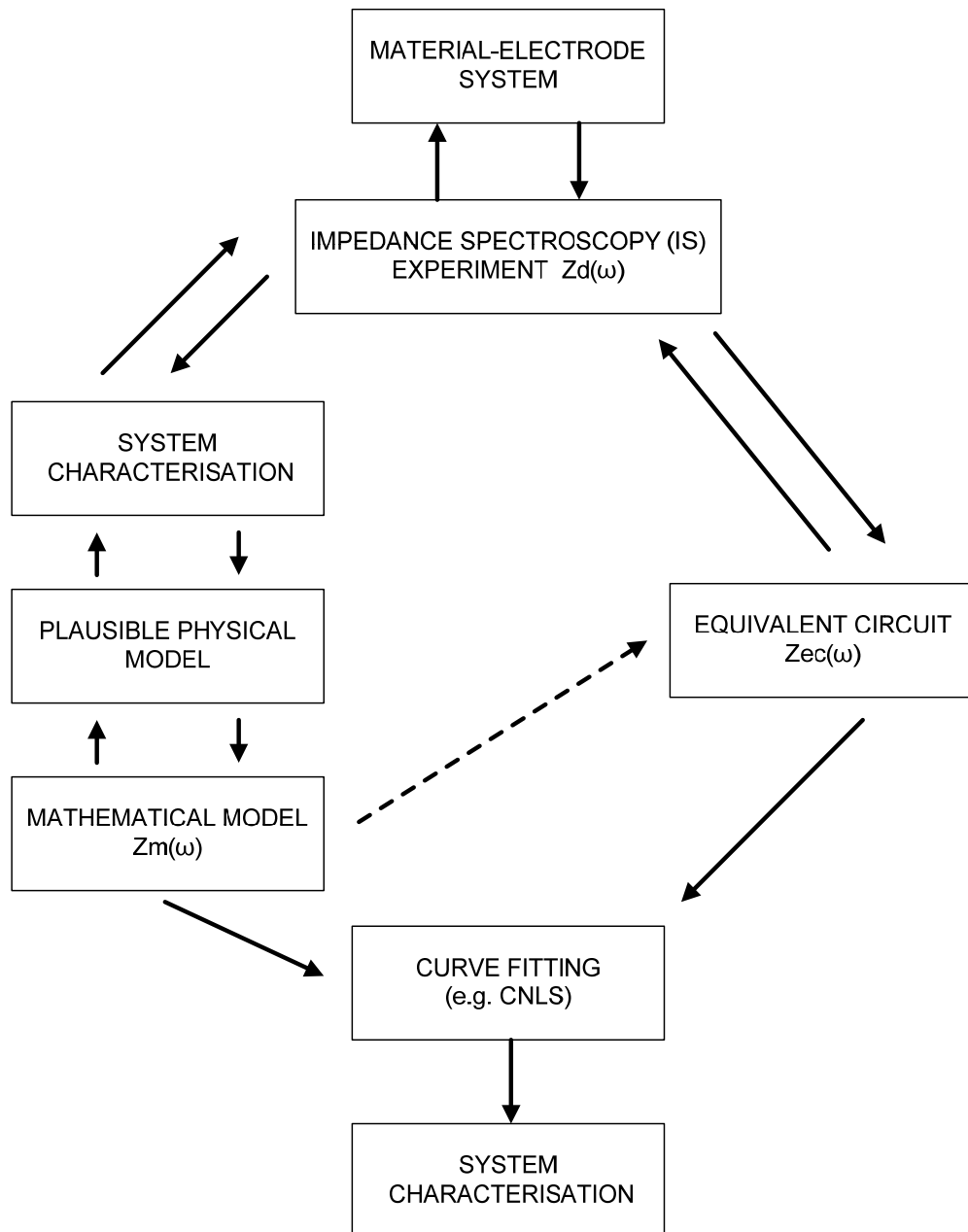


Figure 4.8. Flow diagram for the measurement and characterisation of a material-electrode system (Macdonald 1992).

During the last decade several research studies considering the application of electrochemical impedance spectroscopy in lubrication have appeared (Boyle and Lvovich 2007, 2008; Lvovich 2007, 2008; Lvovich et al. 2004; Lvovich and Smiechowski 2006, 2007, 2008; Smiechowski and Lvovich 2005; Ulrich et al.

2007). Important conclusions can be drawn from these studies; the impedance response is dependent on the electrode geometry and their contact with the medium.

From electrochemical studies and patents of impedance measurements of lubrication oils mentioned, it can be concluded from the Nyquist plots that low frequencies are the most sensitive to variations in impedance due to changes of the dielectric of the fluid under test with the shortcoming that small deviations of the measurement frequency lead also to great changes in impedance. However, high frequencies provide similar readings of impedance in the vicinities of the nominal measurement frequency but lower sensitivity to detect changes in the dielectric.

From an engineering point of view, the design criteria should consist of optimising the final cost and functionality of an impedance monitoring sensor within an adequate range of frequencies. The final operational frequency range depends on the electrode-lubricant interface, and therefore, the previous information should be considered for the selection of the best strategy (electronic topology and electrode system) that minimises the cost of electronics and maximises the sensitivity of the sensor. Therefore, oscillator based methods should not be used at low frequencies because small drifts of the measurement frequency can imply a misleading reading. Frequency drifts at high frequencies (i.e, $1\text{MHz} < f < 100\text{MHz}$) are not important as impedance does not vary significantly with respect to the reference frequency of oscillation. $\text{Tan}\delta$ sensors use this feature to provide an indicator of the degradation of the oil at an affordable cost. At lower frequencies, the influence of frequency drifts is greatly noticeable and makes the oscillator-based method not suitable for impedance monitoring. For lower frequencies other strategies can be considered to monitor impedance (bridge methods, autobalancing bridge methods), all of them based on a very stable excitation source in frequency terms. There are other patents such as (Philips et al. 2006) that specifically exploit the idea of measuring at lower frequencies and hence they propose a sensor which monitors the lubricant impedance at lower frequencies ($<100\text{Hz}$).

New products for oil condition sensing are currently appearing in the market. The most promising one claims to be an IS sensor manufactured by Impact Technologies (Wooton et al. 2006). However, there are two limitations that make this type of sensors not as competitive as alternative well-established technology.

The first limitation is the complexity of devising an accurate IS model that can isolate and indicate each type of contaminant in the oil such as water, soot, glycol by means of using impedance measurements. The second is the final high cost of the electronics when compared to single frequency measurements which makes this kind of sensor less attractive than established technology. An interesting research would be the validation of the accuracy of the Impact IS sensor for different service oils exposed to different water contamination levels and the understanding of the behaviour of their electrochemical model with their type of geometry of electrodes. From a research point of view, impedance monitoring of oils is still an active field in research and far from a mature discipline. It has the potential to be a very useful tool from a custom based calibration approach. The development of a low-cost IS sensor that could be calibrated for specific types of oils with a reliable IS model is definitively the best alternative compared with the other commercial solutions based on conductance or complex dielectric measurements at one frequency.

4.3.3.2 Current commercial technology

In the market there are many types of oil quality sensors based on conductivity and permittivity measurements at one frequency. The conductivity ones are based on potentiostat measurements. The electrodes can be based on a polymeric bead matrix structure in which the detection principles are based on changes to the resistance of the polymer that depend on oxidation products and free water (Voelker and Hedges 1998), or electrodes made with dissimilar metals where the potential difference between the sensitive and reference electrodes can be detected (pH probe) (Kuroyanagi et al. 1996) or detecting the point when the lubricant starts to conduct applying a triangular waveform to the electrodes and using current to voltage converters (current follower) (Schoes 2004; Wang et al. 1992). The sensors based on permittivity measurements are classified in two types depending on output. The first ones only monitor changes in the real part of the permittivity and the output of the second ones is related to the complex permittivity. The ones that monitor the real part of the permittivity measure changes in the capacitance of the electrodes whilst the sensors that monitor the complex permittivity provide output related to the capacitance and dielectric losses. The parameter that relates these two quantities of the complex permittivity is the dissipation factor (D or $\tan\delta$) which is the ratio

between the imaginary part and the real part, hence the reason for their name “tan δ sensors”.

The real part of the permittivity can be measured using low cost electrodes (Raadnui and Kleesuwa 2005) and very simple circuits such as bridges (Hopkins and K.Irwin 1965; Hopkins and Wedel 1977), resonant circuits (Appleby 2010), astable multivibrators (Liu et al. 2000).

The main shortcoming of these sensors when compared to the ones that monitor complex permittivity is that their sensitivity is much lower. In other words, a circuit that monitors dielectric losses is more effective than ones that monitor the energy stored in the dielectric. According to Chris Collister (Collister 1998), the imaginary part of the permittivity is two orders of magnitude (100 times) more sensitive than the real part of the permittivity.

All measurement techniques for impedance spectroscopy are suitable to characterise the dissipation factor for a broad range of frequencies. For a detailed description of impedance measurement methods the reader is referred to (Okada and Sekino 2003).

If the dissipation factor is to be measured at one or small frequency ranges, the following three alternatives are the most cost effective methods for lubrication oils. For high frequencies, i.e greater than 1MHz, the best strategy is the use of a Colpitts oscillator, and for lower frequencies the use of a voltage bridge feed by a stable oscillator or the use of an autobalancing bridge method as it is shown in patent (Lin et al. 2010).

For high frequencies, the impedance of the oils does not vary significantly at frequencies close to the reference or default oscillation frequency. In other words, small drifts on the oscillator frequency (i.e $\Delta f < \pm 5\%$) do not influence the impedance reading (i.e $\Delta Z < \pm 1\%$) which makes this technique a cheap and interesting alternative. Therefore, the optimum design criterion for oscillator based methods is to find the lowest frequency where the impedance deviations due to a maximum drift of frequency are less than the permissible variation in impedance reading.

4.3.3.3 Single frequency circuits based on bridges and Tan δ sensors

Bridge methods can be a cost effective method to monitor the dissipation factor (Allen 2006; Okada and Sekino 2003).

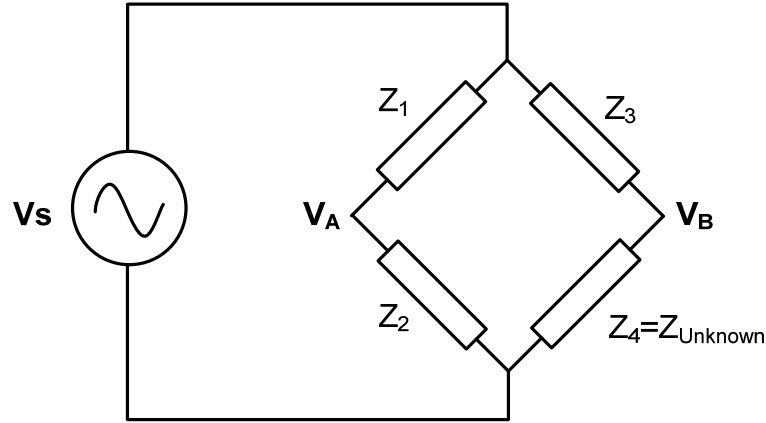


Figure 4.9. Bridge with voltage excitation

The working principles of the bridge shown in figure 4.9 are based on the calculation of the differential voltage $V_A - V_B$, equation (4-9).

$$V_A = V_s \frac{Z_2}{Z_1 + Z_2} \quad V_B = V_s \frac{Z_4}{Z_3 + Z_4}$$

$$V_A - V_B = V_s \left(\frac{Z_2}{Z_1 + Z_2} - \frac{Z_4}{Z_3 + Z_4} \right) = V_s \left(\frac{Z_2 \cdot (Z_3 + Z_4) - Z_4 \cdot (Z_1 + Z_2)}{(Z_1 + Z_2) \cdot (Z_3 + Z_4)} \right) \quad (4-9)$$

Equation (4-9). Differential voltage in a bridge with voltage excitation

The typical condition for finding the value of the unknown impedance is to modify one of the arms of the bridge until the differential voltage is 0, equation (4-10). This is not the best strategy for a low cost implementation of a sensor as it needs a circuit capable of self-adjusting the impedance of the arms.

$$V_A - V_B = 0 \Rightarrow \frac{Z_2}{Z_1 + Z_2} = \frac{Z_4}{Z_3 + Z_4} \Rightarrow Z_4 = \frac{Z_2}{Z_1} Z_3 \quad (4-10)$$

Equation (4-10). Null type impedance bridge

A circuit that imposes a condition as shown in equation (4-11) has been claimed to monitor the dissipation factor (Allen 2006).

$$\text{Im}(Z_2) \cong \text{Im}(Z_4) \quad \& \quad Z_1 = Z_3$$

$$V_A - V_B = V_s \left(\frac{\text{Re}(Z_2) + \text{Im}(Z_2)}{Z_1 + \text{Re}(Z_2) + \text{Im}(Z_2)} - \frac{\text{Re}(Z_4) + \text{Im}(Z_2)}{Z_1 + \text{Re}(Z_4) + \text{Im}(Z_2)} \right) \quad (4-11)$$

Equation (4-11). Differential voltage reading for a bridge circuit as reported in (Allen 2006)

This condition imposes a differential voltage that hardly changes with the real part of the unknown impedance as $\text{Re}(Z_2)$ and $\text{Re}(Z_4)$ appears in both numerators and denominators. As a result, this circuit is not recommended and the sensitivity to monitor the $\text{Re}(Z_4)$ is questionable.

Imposing the resonance condition in the arm of the bridge with the unknown load, the differential voltage can be obtained as shown in equation (4-12). This condition can be achieved for a broad range of frequencies using a general impedance converter (GIC). The phase angle of Z_4 can be monitored using a low cost phase detector, such as two zero cross detectors (one connected to V_A and the other to V_B) implemented with high speed operational amplifiers (comparators) and the output of the zero cross detectors connected to an XOR gate. The output of the XOR gate is a PWM signal that can be filtered to obtain a voltage proportional to the phase of Z_4 . The main limitation of this circuit is that at low frequencies the oscillation condition (LC circuit) can impose large values of inductance $L=1/(C \cdot \omega^2)$.

$$\text{Im}(Z_3) = -\text{Im}(Z_4) \Rightarrow V_A - V_B = V_s \left(\frac{Z_2}{Z_1 + Z_2} - \frac{Z_4}{\text{Re}(Z_3) + \text{Re}(Z_4)} \right) \quad (4-12)$$

$$Z_1 = Z_2 \quad \& \quad \text{Re}(Z_3) \ll \text{Re}(Z_4) \Rightarrow V_A - V_B \approx V_s \left(\frac{1}{2} - \frac{Z_4}{\text{Re}(Z_4)} \right)$$

Equation (4-12). Bridge with the resonance condition in one of the arms

If it is possible to impose the condition such as the total impedance of one of the arms of the bridge as the inverse of the total impedance of the other arm,

considering that Z_3 is less than Z_4 then the differential voltage is proportional to the unknown impedance, equation (4-13).

$$Z_1 + Z_2 = \frac{1}{Z_3 + Z_4} \Rightarrow V_A - V_B = V_S (Z_2 (Z_3 + Z_4) - Z_4 (Z_1 + Z_2))$$

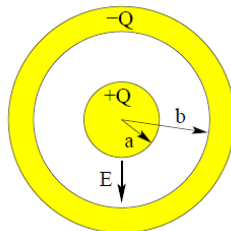
$$Z_2 = \frac{1 - Z_1 (Z_3 + Z_4)}{Z_3 + Z_4} \Rightarrow V_A - V_B = V_S \left(1 - Z_1 (Z_3 + Z_4) - \frac{Z_4}{(Z_3 + Z_4)} \right) \quad (4-13)$$

$$Z_3 \ll Z_4 \Rightarrow V_A - V_B \approx V_S (-Z_1 \cdot Z_4)$$

Equation (4-13). Bridge with a relaxed impedance condition

As shown above, bridge methods are a cost effective alternative because the phase detector circuit is cheap and nowadays programmable sine wave generator can be implemented at low cost and high accuracy with Digital Direct Synthesis (DDS) chips.

Tan δ sensors are based on oscillators and the sensing unit is a cylindrical capacitor as shown in Figure 4.10. For the circuit design of an impedance sensor, the first step is to characterise the impedance range of the sensing element. Some circuit theory concepts are needed in the first instance, Figures 4.10-4.12.



Capacitance for cylindrical geometry

$$C = \frac{Q}{V} = \frac{2\pi\epsilon_0\epsilon_r L}{\ln\left(\frac{b}{a}\right)}$$

Figure 4.10. Tan δ sensors with cylindrical electrodes.

The losses of the dielectric can be represented with a complex permittivity as shown in Figure 4.11. The equivalent circuit for a cylindrical capacitor is found in Figure 4.12.

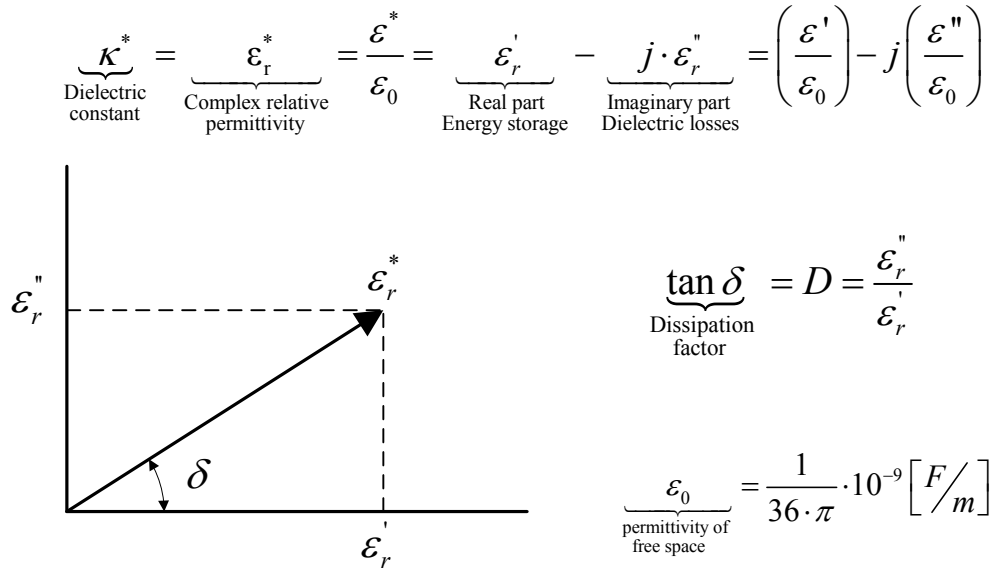


Figure 4.11. Definition of complex relative permittivity

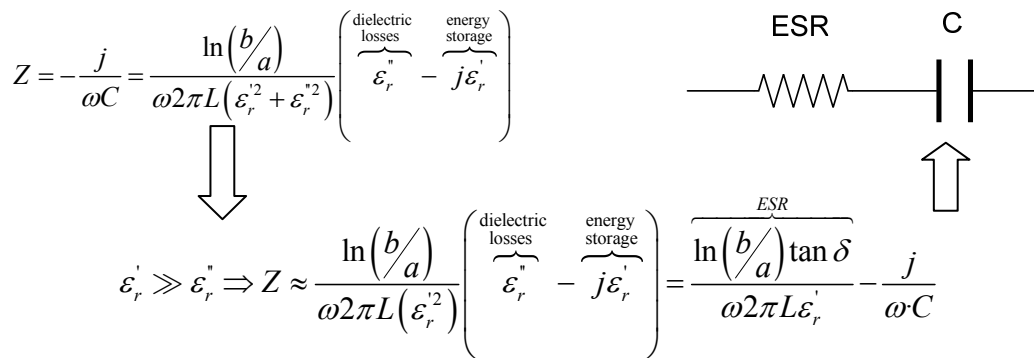


Figure 4.12. Equivalent circuit for cylindrical electrodes (tanδ sensor's head)

The range of complex permittivity can be obtained using an impedance analyser and impedance frequency sweeps for different lubricant degradation stages. The conclusion of these impedance frequency sweeps should be similar to the data appearing in (Collister 1998). According to (Collister 1998), the values of the dissipation factor (tanδ) are between 0.001 and 0.1 within 10-300MHz (0.001 for an unused mineral oil and 0.1 for a heavily contaminated sample) and the real permittivity is typically 2.25 (unused oil) to 2.45 respectively. Using the range of

permittivity and the geometrical dimensions of both cylindrical heads, Table 4.4, the working range of impedance is obtained, Figure 4.13.

Table 4.4. Geometrical dimensions of the electrodes of commercial $\tan\delta$ sensors.

Sensor	L (mm)	b (mm)	a (mm)	C_0 (pF)@ ϵ_0	C (pF) @($\epsilon_r=2.25$)	C (pF) @($\epsilon_r=2.45$)
Kittiwake	12	8	7	50	112	122
OMS	23	5.6	3.5	27	61	66

From the impedance charts, it is easy to realise that the real part of the impedance varies in a larger range (2 decades) than the imaginary part (low range of values). Any circuit which monitors the real part of the impedance will be more sensitive than a circuit which only measures the imaginary part. This explains why the methods to monitor the degradation of oils using capacitive measurements capacitance (imaginary part of the impedance) are not as effective as methods that monitor the dielectric losses.

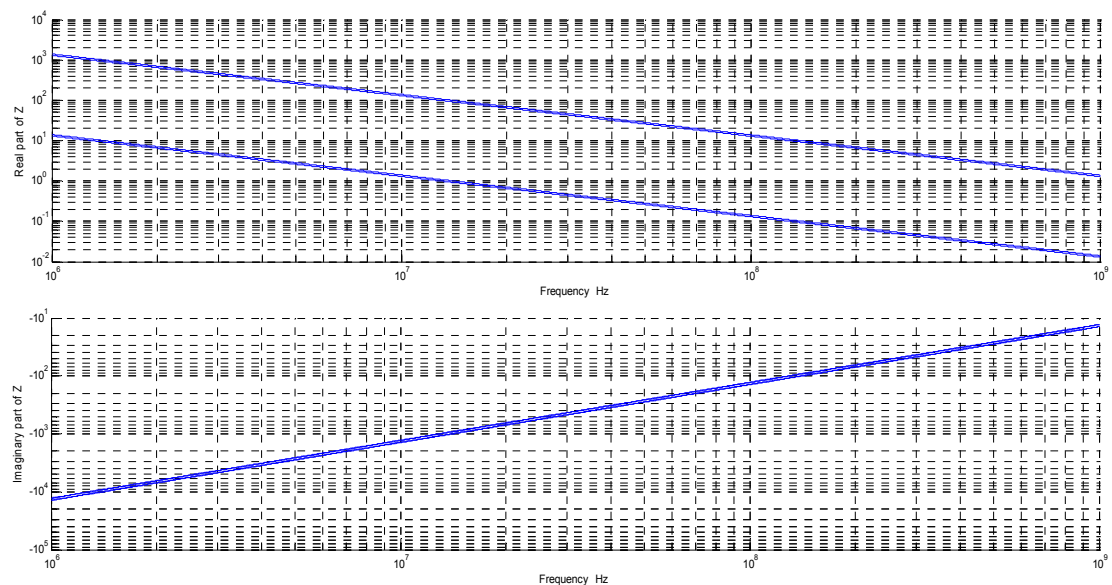


Figure 4.13. Operational range of impedance for Kittiwake's cylindrical sensor head.

When the range of impedance of the sensing element is known, the optimum design should be chosen to achieve the lowest measurement frequency (to gain more sensitivity) and should fulfill the requirement that impedance reading variations due to oscillator frequency drifts should be within a specified tolerance (e.g. $\Delta Z < 1\%$). The optimum frequency selection can be graphically chosen from the frequency sweeps of the dissipation factor bearing in mind that $\tan\delta$ is directly related to the

complex impedance. For example, Figure 4.14 shows to frequency sweeps of the dissipation factor for two mineral oils, new and used oils respectively. Possible candidates for the best operational frequencies seem to be around 1MHz where $\tan\delta$ has almost the same value within a 1 MHz span and 12MHz where $\tan\delta$ is almost constant for oscillator frequency drifts. The lower frequency is the best one because $\tan\delta$ sensitivity to oil degradation is better than the 12MHz region. As a result, considering these two observations the best frequency to perform measurements is around the MHz region. For the optimum frequency selection dissipation factor plots for new and used oils at several temperatures should be utilized for selecting the best frequency that fulfill the design requirements.

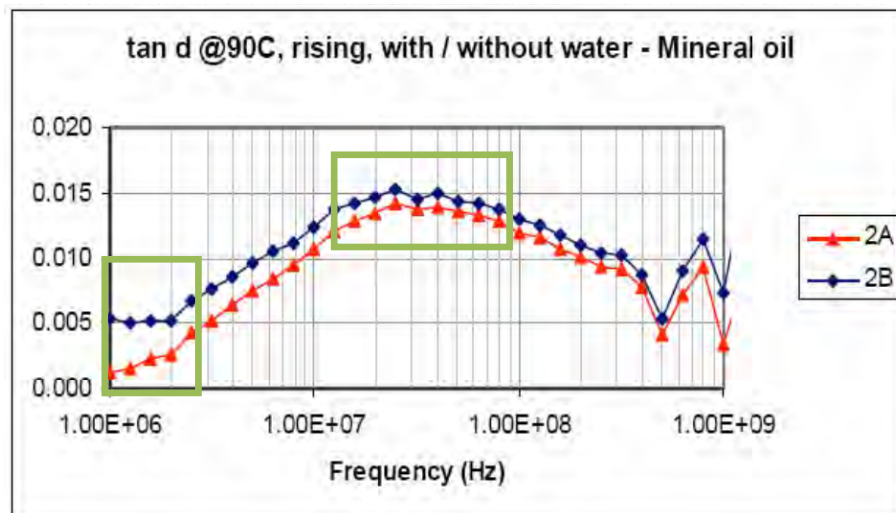


Figure 4.14. $\tan\delta$ frequency sweep for a new (2A) and used oil (2B) (In green frequency candidates for the implementation of the sensor) (Murphy and Kent 2004)

High frequency measurements can be measured using an oscillator-based method. A Colpitts RF oscillator is used to monitor the dielectric losses. The implementation of this oscillator can be performed using one transistor or with a marginal oscillator implemented using a CFA (current feedback amplifier), Figures 4.15, 4.16. The amplitude of oscillation is inversely proportional to the ESR of the sensing capacitor. A low ESR implies high amplitude of oscillation. After signal conditioning the output voltage is proportional to the ESR of the sensing capacitor.

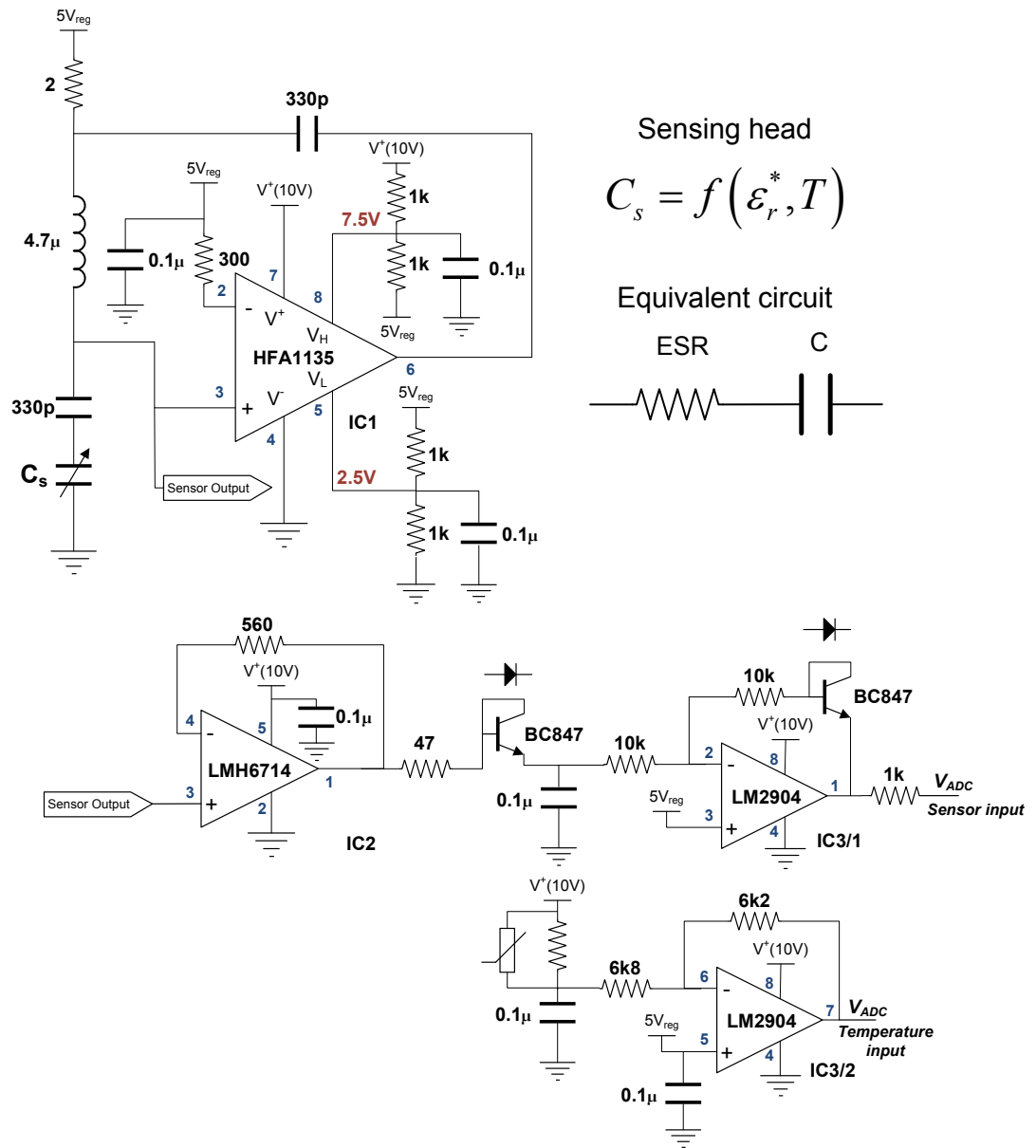


Figure 4.15. $\tan\delta$ sensor (HW implementation)

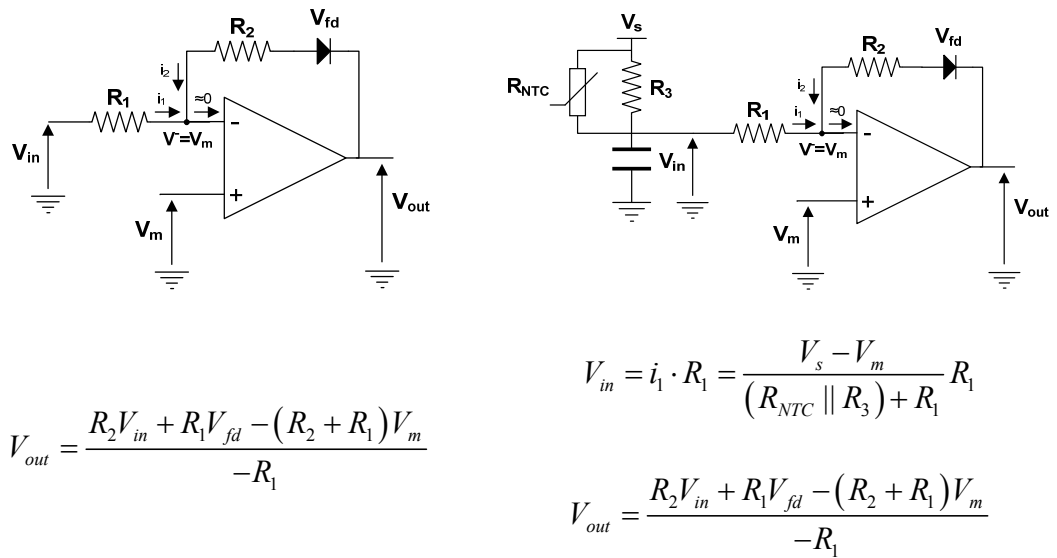


Figure 4.16. Analysis of the circuits, for temperature and signal conversion ($V_{fd} \approx 0.6V$)

The circuit was simulated using ISIS v7.6 distributed by Labcenter Electronics, Figure 4.17. The models of the CFA were substituted for an equivalent AD8001A for the oscillator part and MAX4113ESA for the non-inverting buffer. The inductor was simulated according to a real ceramic inductor with high Q, the equivalent series resistor for the inductor is 5Ω .

A functional design of the sensor needs to have temperature compensation. The calibration is based on recording the output voltage of the lubricant under different temperatures. When the temperature increases the ESR tends to decrease (in lubrication oils this always occurs, please see the Nyquist plots of other EIS studies (Lvovich 2008)) and the voltage in the (sensor output port) increases. After signal conditioning the final voltage output decreases. Recording the output voltage at different temperatures allows the selection of a suitable temperature compensation relationship.

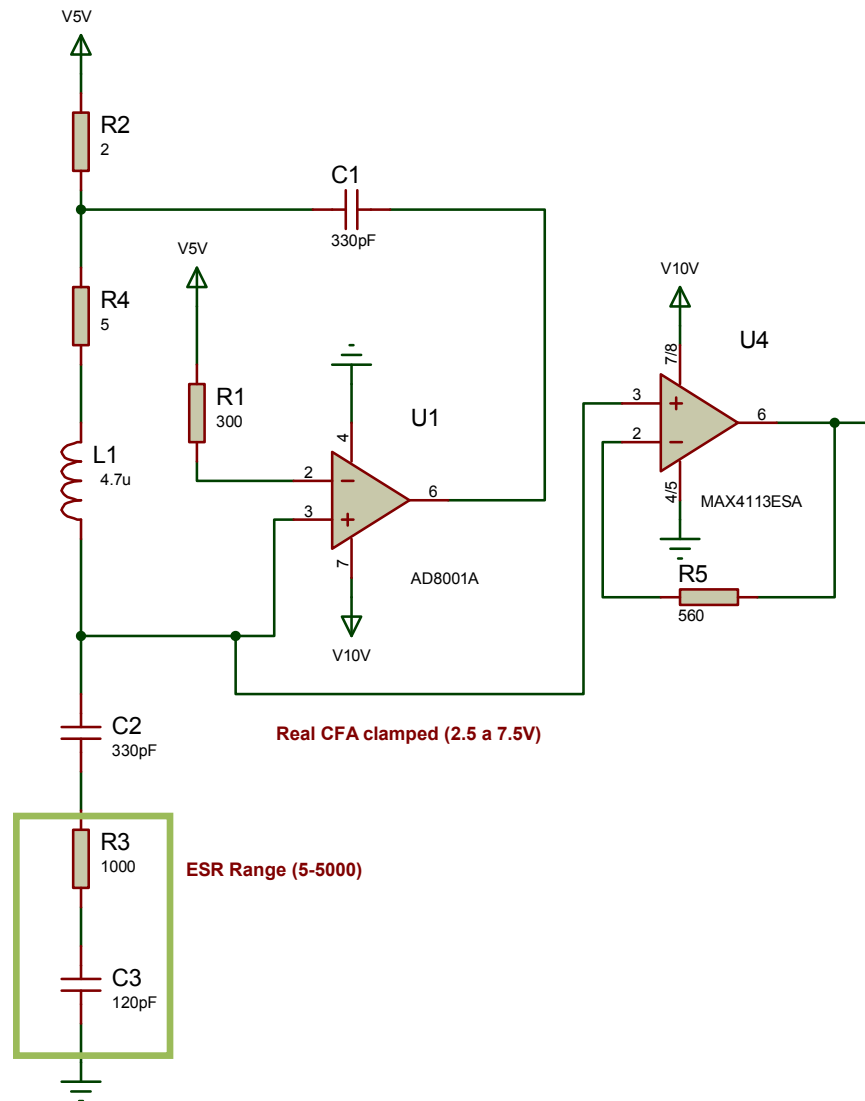


Figure 4.17. Oscillator circuit with the equivalent circuit for the sensor head, R_3 range represents the dielectric losses

4.3.3.4 New ideas for the development of a low-cost and highly-integrated oil condition IS sensor

The last trends in oil quality monitoring are shifting towards IS studies rather than monitoring impedance at a single frequency (Lvovich 2007, 2008; Lvovich et al. 2004; Lvovich and Smiechowski 2006, 2007, 2008; Ulrich et al. 2007). The need to measure the impedance in a broad range of frequencies requires the selection of the most cost effective solution for a final and portable implementation (Wooton et al. 2006). Typical strategies are the use of autobalancing bridge methods (for non-

grounded loads), or current-voltage (V-I) methods (grounded and non-grounded loads) (Okada and Sekino 2003). Another interesting approach to monitor impedance is the use of AC current sources. The underlying principle is that a known current waveform can be pumped to the unknown load. The voltage drop across the unknown load can be measured and therefore, the unknown load can be determined from the known current and voltage waveforms. This method for measuring impedance is widely used in biomedical studies (Frounchi et al. 2009).

The main considerations during the design process of an AC current source for an IS application is the bandwidth and the output impedance. A good current source design should provide accurate current over a wide frequency range and accommodate large load impedance variations. For IS studies up to 1MHz there are several amplifiers that have enough bandwidth, but the output impedance is a big limitation specially at low frequencies for lubricant condition monitoring. The best way to improve the output impedance of the Howland current source for a broad range of frequencies is the use of Generalised Impedance Converters that neutralize the parasitic capacitance at the output node (including stray parasitic) but GIC has a narrow band-pass frequency response, and there is the need to tune it at different frequency ranges. Another approach to neutralising the output capacitance is the use of negative capacitive circuits (NCC) that show a wide response in frequency. However, NCC circuits should have other limitations as they are not used in Electrical Impedance Tomography. The approach of using an improved Howland Current Source and GIC can lead to a system that can monitor impedances over the frequency range of 100Hz to 1MHz considering a range of output impedances in the order of hundreds of $M\Omega$ as the output impedance of the current source is within the $G\Omega$ order as shown in patent (Ross and Saulnier 2006). This system can be an effective alternative for portable IS studies.

During the last 5 years new ICs are appearing on the market making impedance monitoring of lubricants viable at an affordable price. A popular IC is the AD5933. A variety of designs using this chip for impedance measurements depending on the application are available. For example, AD5933 is becoming popular in non destructive testing such as corrosion monitoring where impedance measurements are based on the V-I method with two AD5833 (Hoja and Lentka

2010). Possible implementations that suggest the use of AD5833 of hand-held devices that perform IS measurements based on the V-I method are reported as industrial research (Koehler et al. 2008; Tomlinson 2009). However, amongst the circuits used for IS based on AD5933; a system that performs impedance spectroscopy measurements for lubrication oils using AC current sources has not been yet reported.

A prototype (concept design) of an impedance monitoring system based on AD5933 that can be used for EIS studies of oils is shown in Figure 3.18. The circuit is based on an improved Howland AC current source and a GIC. The AD5933 is used for characterising the frequency response of the whole system. Finally, it is important to note that the use of an AC current source for impedance monitoring of lubrication oils has not been reported in the literature and it is not patented for this intended use. Therefore, this prototype is a new contribution to knowledge.

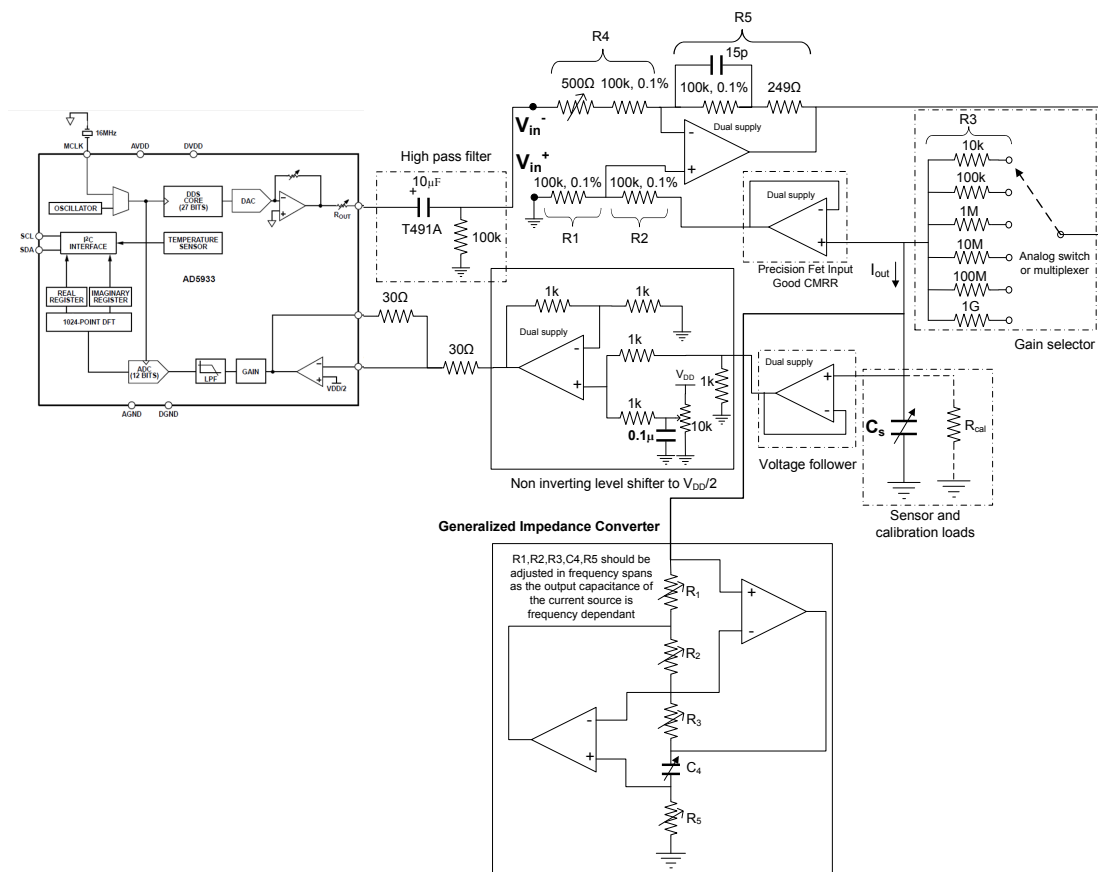


Figure 4.18. Proposed system for IS studies using an improved Howland current source and GIC converter.

4.3.4 Particle counters and ferrous debris density sensors

Industrial lubricants lose their functionality through high temperature-induced oxidative degradation and/or the increased presence of major contaminants such as water, glycol, fuel, soot and wear metals. As a result, performance deteriorates and industrial equipment eventually fails. The condition of critical components can be evaluated by the collection and analysis of wear debris being generated at interacting surfaces and transported by the fluid medium to a suitable sampling position. The analysis of the debris relates to various aspects such as quantity (concentration), size distribution, composition and morphology. Measurement of concentration and size distribution provides quantitative (objective) information on the progression of wear while the morphology and composition of the particles provide interpretive (subjective) information of the wear as a whole. The current sensor technology for on-line monitoring only monitors quantitative information. Wear debris sensors provide information on the **progression of wear**.

Currently, wear monitoring techniques provide the particle size distribution using particle counters and the density of the particles using particle density sensors.

As is shown in Figure 1.1, the density of particle size distribution is a strong indicator of the onset of faults. The values should not be considered as a guideline for all types of machinery. This graph should be read in terms of trends. As soon as there is an onset of wear the particle sizes tend to increase. For example, this kind of graph under controlled experimental conditions (sliding four-ball tests, pin/ball-on - disc) does not match the experimental results.

The bath curve is related to the wear rate considering all possible particle sizes. It is another indicator of fault detection but not as strong as the previous one.

Nowadays technology is limited to particle size detection; hence there are two types of sensors to tackle the progression of wear, each of them with their own strengths and limitations.

Commercial ODM (Oil Debris Monitoring) particle counters and ferrous debris density sensors rely on the principles of magnetometry. Their output when compared to other particle counting technologies is not significantly affected by air

bubbles and water droplets in the fluid stream. It should be noted that optical particle counters based on light scattering are the most accurate ones but their main shortcoming when applied in the lubrication industry is that air bubbles or other gasses (depends on the solubility of gases in liquids) that are transported by lubrication oils and also water droplets are counted at the end of the process providing misleading results. These final effects are really difficult to compensate using light scattering principles as they depend on many external conditions such as flow rate, type of lubricant, pressures, vibration, etc. Other types of particle counters based on ultrasonic, capacitive coulter counter, and capacitance measurements are reported in the literature (Murali et al. 2009; Murali 2008; Nemarich et al. 1988; Sarangi 2007; Xia 2009; Zhe et al. 2007; Zhe et al. 2010). During the last decade, several publications related to wear debris monitoring based on electrostatic sensors have appeared (Craig et al. 2009a; Craig et al. 2009b; Harvey 2002, 2007; Powrie 2001; Wang et al. 2007). Electrostatic monitoring in lubrication oils is an interesting field of research but the charging process of the particles is not fully understood as it depends on many factors, such as the chemical composition of the lubricants, the temperature, the type of materials, the type of flow, the geometry of the electrodes, the deposition of the particles... As a result, it is not easy to visualise a future commercialisation for these sensors. Perhaps the detection capabilities of electrostatic sensors could be improved combining these sensors with other physical sensing principles. Currently, none of these research alternatives can compete with ODM sensors based on magnetometry.

The best particle counters count the particles in size ranges of 40-60 μm , 60-100 μm , 100-200 μm , 200-300 μm , >300 μm so are limited by the minimum size detection and resolution. However, ferrous debris sensors can measure all types of particles as they measure the bulk amount of debris contained in a known volume, the only limitation is the minimum detectable mass, i.e Kittiwake's sensor limit is approximately 4 μg for steel particles. The final limit number is obtained from the debris density sensor sensitivity (1 ppm in mass) and the volume of the sensing chamber 4ml. From the specifications of commercial sensors the minimum size detection and current size limitation of magnetometry devices can be obtained by referring to the spherical particles, Table 4.3. As can be seen, an increment in

particle radius is related to a cubic increment of particle volumes. That means the sensitivity of current devices needs to be improved 1000 times for a reduction of ten times the particle size diameter. This is a considerable challenge and it can be assumed that current detection schemes, that are discussed later, present a constraint of 30 μm in particle size detection. There is therefore the need for another type of strategy that can overcome this size limitation with current magnetometry sensing techniques. The next chapter discusses a very interesting methodology to overcome the size limitation boundary.

Table 4.3. Mass and particle size limitation of particle counters and ferrous debris density sensors based on magnetometry.

Spherical particle (μm) radius	Estimated volume (μm^3) $4/3 \cdot \pi \cdot r^3$	Estimated mass (μg) (steel = 7850 Kg/m^3)
1	4,189	$3,288 \cdot 10^{-5}$
10	$4,189 \cdot 10^3$	$3,288 \cdot 10^{-2}$
20	$3,351 \cdot 10^4$	$2,631 \cdot 10^{-1}$
30	$1,131 \cdot 10^5$	$8,878 \cdot 10^{-1}$
40	$2,681 \cdot 10^5$	2,104
50	$5,236 \cdot 10^5$	4,110
60	$9,048 \cdot 10^5$	7,102
70	$1,438 \cdot 10^6$	11,279
80	$2,145 \cdot 10^6$	16,836
90	$3,053 \cdot 10^6$	23,971
100	$4,189 \cdot 10^6$	32,882
In red is the technological barrier / In orange is the best current technology / In blue the standard size detection		

4.3.4.1 Ferrous debris density technology

The physical and operational principles of the ferrous debris density sensor developed by Kittiwake can be found in (Hutchings 2000). The sensor has the same technology as the PQ (particle quantity) index and is a dual-coil magnetometer. It has two coils arranged so that the sample coil (sensor) and the reference coil are in balance when there is no sample on the sensor. The system is carefully designed to ensure that both coils respond equally to changes in ambient temperature. When a sample of oil containing ferromagnetic debris is placed on the sensor (sample coil), the balance between the coils is altered. The resulting out-of-balance signal is

amplified, filtered and displayed as a "PQ Index", which relates directly to the mass of ferromagnetic debris in the sample.

A very stable electronic oscillator is used to excite an inductive bridge. The sensing circuit is a balanced inductive bridge, Figures 4.19, 4.20. Any imbalance in the sensing coils is related to a change of impedance that is related to a change in the complex permeability of the medium.

Ideal voltage waveforms of the imbalance of the bridge are shown in Figure 4.21. Ferromagnetic particles are in phase with the reference oscillator and the amount is proportional to the amplitude whilst non-ferromagnetic particles introduce a phase shift that is proportional to the amount of non-ferromagnetic particles.

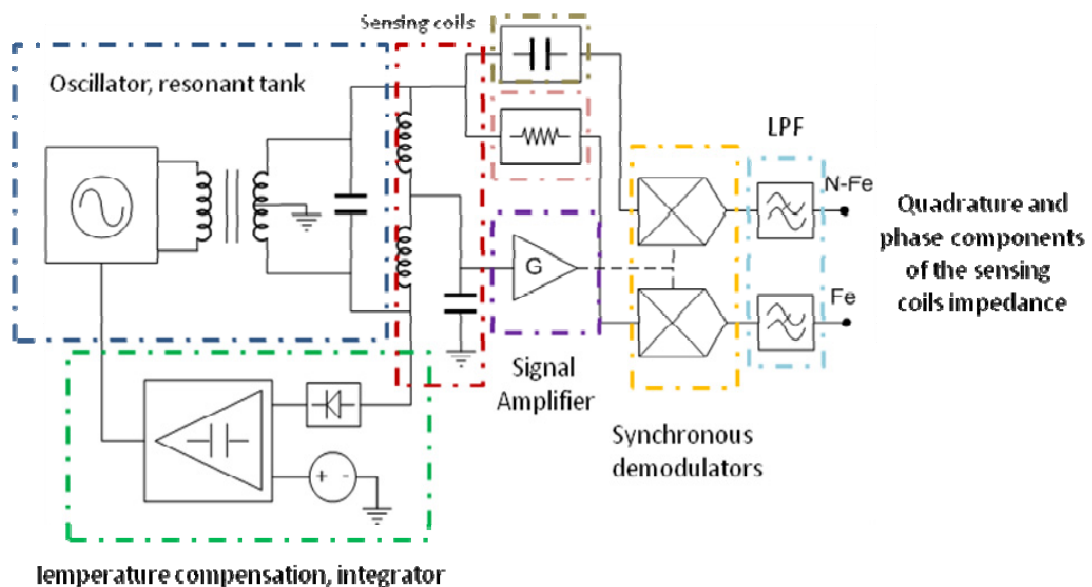
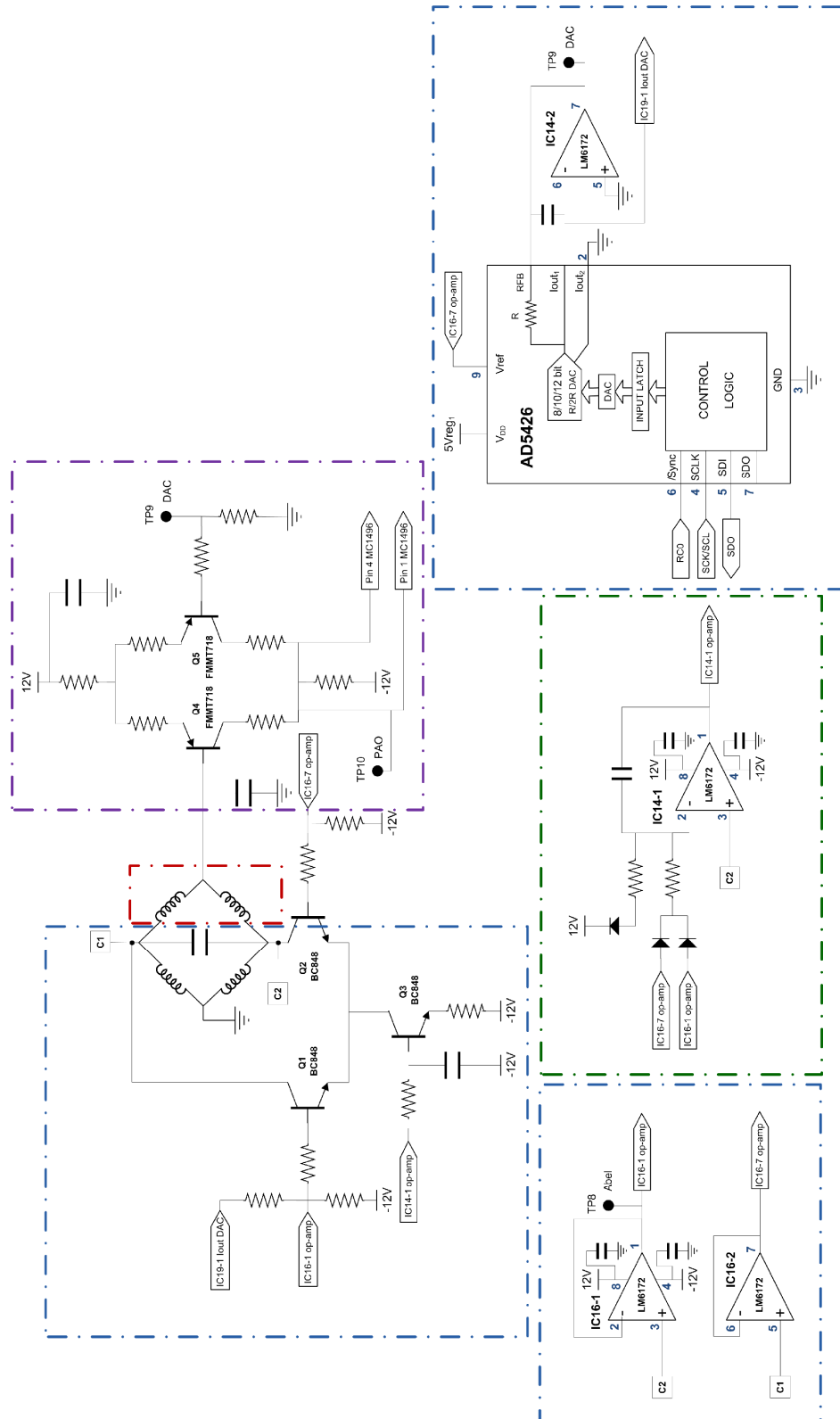


Figure 4.19. Block diagram of the dual coil magnetometer (Hutchings 2000)



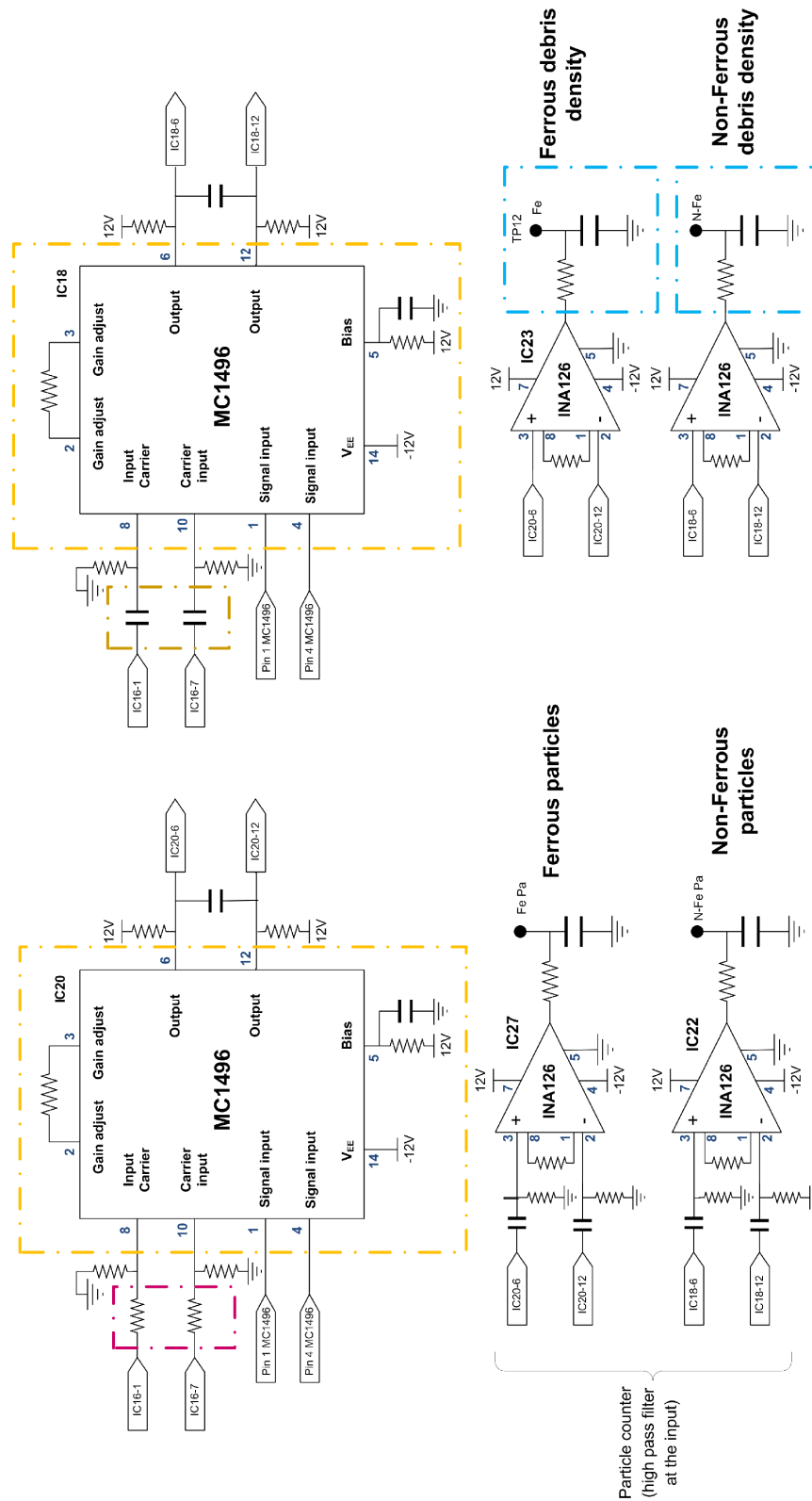


Figure 4.20. Implementation of the sensing part of Kittiwake’s sensor, dashed lines are according to aforementioned block diagram (Hutchings 2000)

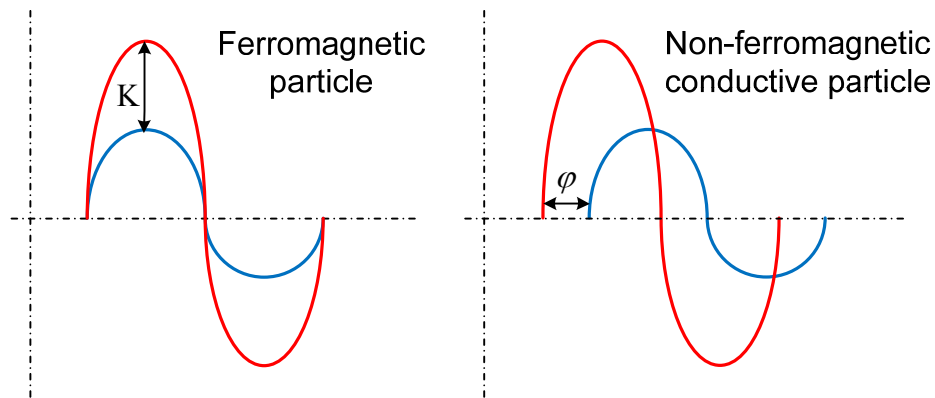


Figure 4.21. Output voltage waveforms from the sensing coils

The real part of the permeability is related to the energy stored and the imaginary part to the losses (for example eddy currents in particles of conductive non-ferromagnetic materials). The imbalance generated is amplified and the carrier signal is synchronously demodulated using the in phase (resistor network) carrier and the 90 degrees shifted (capacitor network) carrier. The output is low-pass filtered and the output voltage is proportional to the real and imaginary part of the coil's impedance (which is directly related to changes of the permeability and therefore to conductive ferromagnetic and non-ferromagnetic particles respectively).

The ferrous debris density sensor has the coils arranged in parallel (Kent 2004; Leigh-Jones et al. 2004) as shown in Figure 4.22.

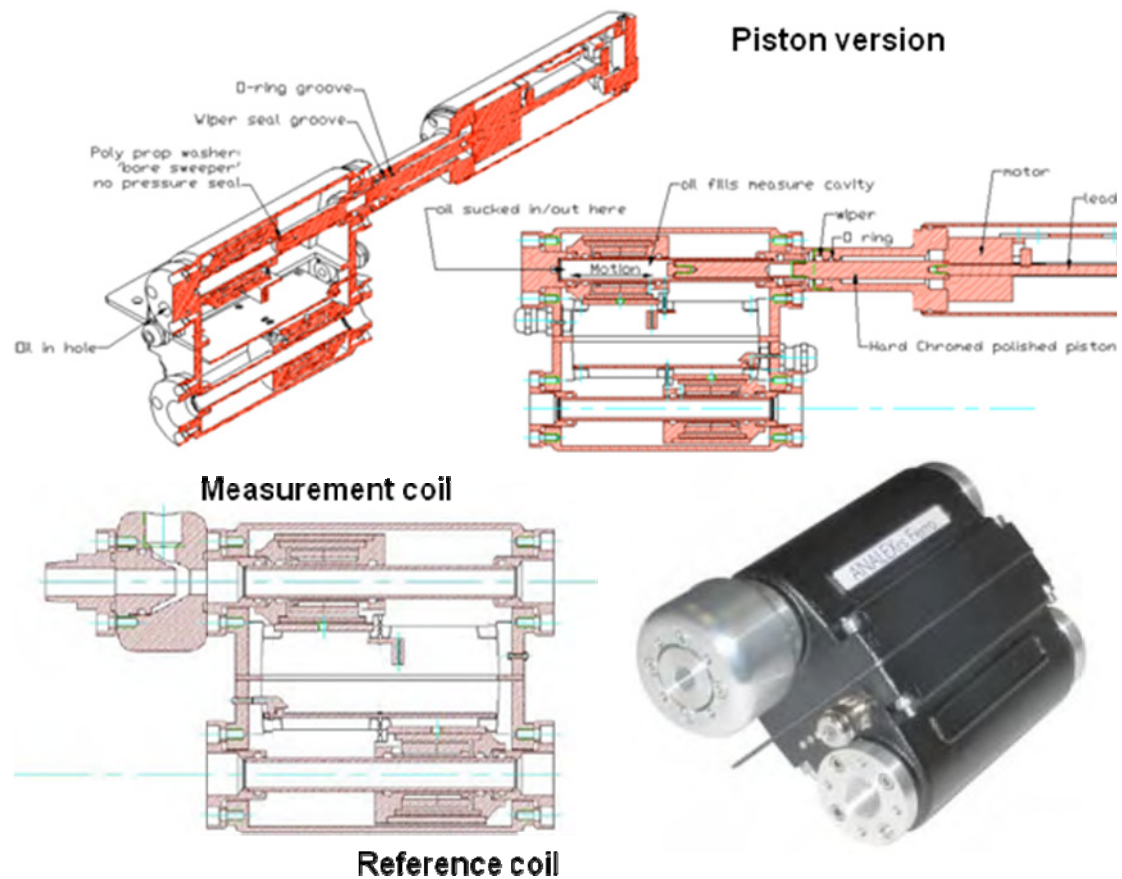


Figure 4.22. Ferrous debris density sensor (Murphy and Kent 2004)

4.3.4.2 Particle counters technology

The particle counter developed by Kittiwake uses the same electronics as the ferrous debris density sensor. The only difference is the location of the coils. For particle counting and classification of the type of debris (metallic conductive or metallic non-conductive particles) the coils need to be arranged in series, Figure 4.23. Since ferromagnetic and non-ferromagnetic conductive particles interact with the field via permeability and eddy current effects, the transient phase signature of a particle has a characteristic lag-lead or lead-lag sequence depending on the nature of the particle. Any imbalance of the bridge with this coil configuration allows the classification of the type of debris because the output voltage of the bridge for non-ferrous conductive particles is shifted 180° with respect to ferrous debris, Figure 4.24. The peak of the signal or the RMS (root mean square) value of the

voltage waveforms are indicators of the size of the debris and the initial phase of the type of debris. The amplitude of metallic conductive particles (i.e iron, steels with magnetic properties) leads to peak voltages higher than non-metallic conductive particles (i.e aluminium). This is due to the effect of eddy currents generated in aluminium (losses) which result in a final decrease of permeability which is less noticeable than the increase of permeability due to ferromagnetic particles. The calibration of sensor technology is dependent on the type of material of the particles, therefore in systems with many types of particles different than steel, this issue should be considered. Some questions could arise regarding the effects of the shape of the particles in the detection process, but considering the coarse range of size classification all these questions are not so relevant and it can be assumed the detected particle is equivalent to a spherical particle within the classified detection size. The improbable event that metallic and non-metallic particles pass through the coil at the same time can cause cancellation effects and lead to incorrect results. This is more likely to occur if the flow rate of the oil within the system is slow and not constant.

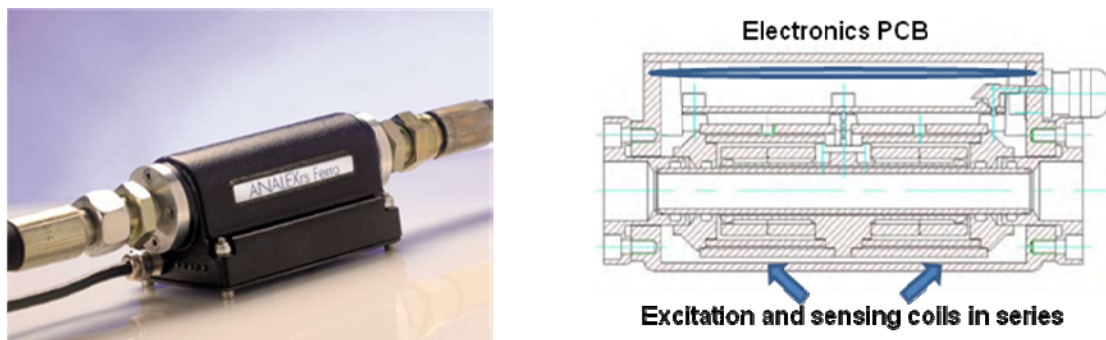


Figure 4.23. Kittiwake's particle counter (Murphy and Kent 2004).

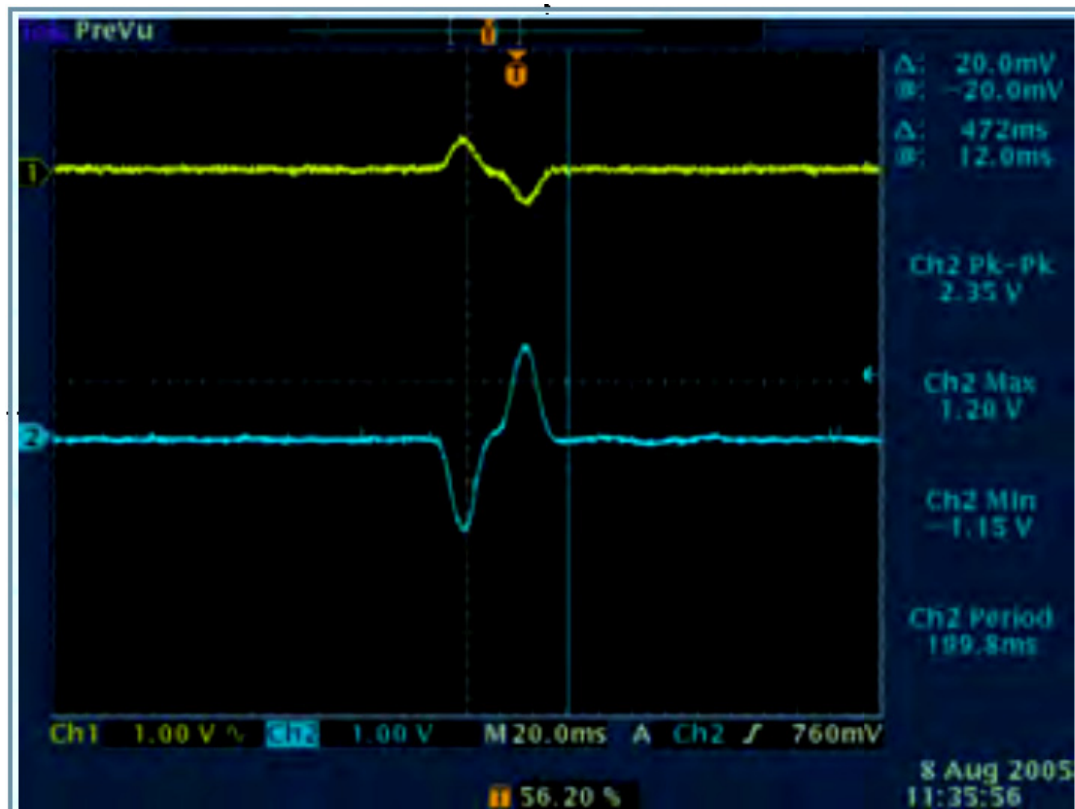


Figure 4.24. Typical waveforms for ferromagnetic and non-ferromagnetic conductive particles (Murphy and Kent 2004).

Another type of particle counter technique which is a simpler configuration and presents a similar operation to the bridge used by Kittiwake is the MetalScan particle counter (Kempster and George 1993; Miller and Kitaljevich 2000). The sensing unit is based on a differential transformer. Two field coils are wound in opposite directions to each other cancelling the magnetic flux so there is no magnetic coupling with the sensing coil. When a metallic particle passes through field and sensing coils, depending on its position, there is magnetic coupling between one of the field coils and the sensing coil and subsequently a voltage waveform proportional to the size of the particle with a characteristic phase depending on the type of particles appears in the sensing coil. The phase of voltage waveform in the sensing coil allows the determination of the type of particle and the peak amplitude detects the size in the same fashion as explained before for the bridge method. A schematic of this type of sensor is shown in Figure 4.25.

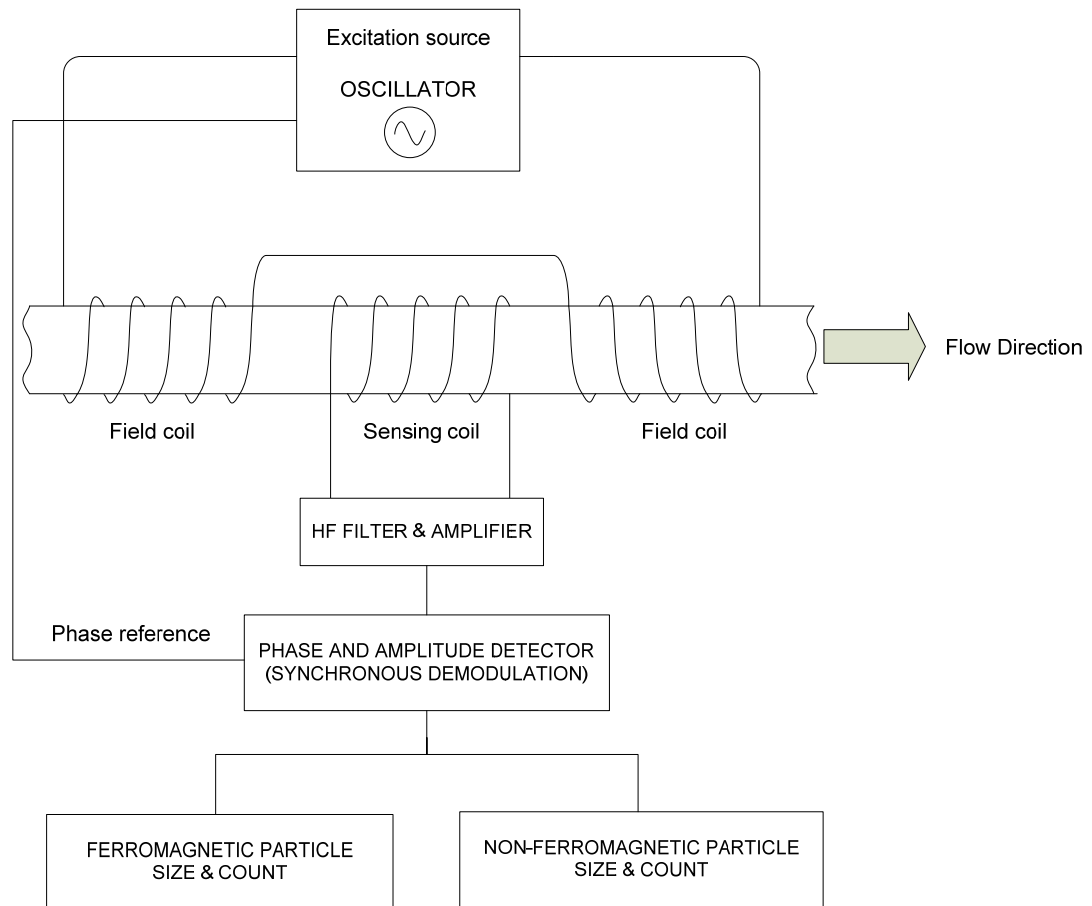


Figure 4.25. Particle counter based on a differential transformer.

Another very sensitive technique to monitor changes in permeability because of the travelling of debris through a coil is to use oscillators. Macom TechAlert units are based on marginal oscillators (Flanagan et al. 1989; Whittington et al. 1989). This way of detection has proven to be very sensitive as it is widely used in the field of Nuclear Magnetic Resonance (NMR). The Robinson marginal oscillator which was invented by Neville Robinson (an English physicist) forms the underlying basis of Magnetic Resonance Imaging (MRI) systems used in many hospitals.

The underlying idea of a marginal oscillator is that the resonant tank imposes the amplitude and frequency of oscillation. When the oscillation reaches the steady state signal level, the voltage in the resonant tank V is obtained when the amplifier transconductance or gain equals the conductance G of the tuned circuit. As a result, the passage of a metallic particle through the coil changes its inductance and

increases its conductance, changing the frequency and reducing the amplitude of the oscillation. By frequency and amplitude demodulating the signal from the oscillator, one can obtain two pieces of information on each particle which passes through the coil (Whittington et al. 1992).

Temperature compensation should be done carefully with this type of oscillator. It should be noted that oil quality sensors based on Colpitts oscillators ($\text{Tan}\delta$ sensors) have the same behaviour as marginal oscillators, in their case the sensing unit is a capacitor and the losses of the capacitor (real part of the impedance) increase the oscillation frequency and decrease the amplitude of the oscillation, which is another way of understanding the circuit shown in Figure 4.26. In $\text{Tan}\delta$ sensors the only parameter of interest is the amplitude as it is related to the losses of the capacitor (the imaginary part of the permittivity) .

Another approach to count particles using oscillators to cause changes in the real part of the permittivity (inductance of the sensing coil) is the use of a phase-locked loop (PLL) to detect the shifts in frequency and to correct for the effects of slow frequency drifts such as those caused by temperature changes (Whittington and Flynn 1998; Whittington et al. 1992) , Figure 4-26. Basically, when a particle passes through the pipe, an increase in the permeability or the inductance of the coil would occur if it was a ferrous debris particle causing the frequency to decrease with respect to the default oscillation frequency. The control loop with the PLL will compensate for the decrease of frequency with a constant voltage applied to the voltage control oscillator. If the particle is non-ferrous the opposite effect occurs. A more detailed description of noise sources and how to maximize the sensitivity can be found in (Whittington and Flynn 1998; Whittington et al. 1992).

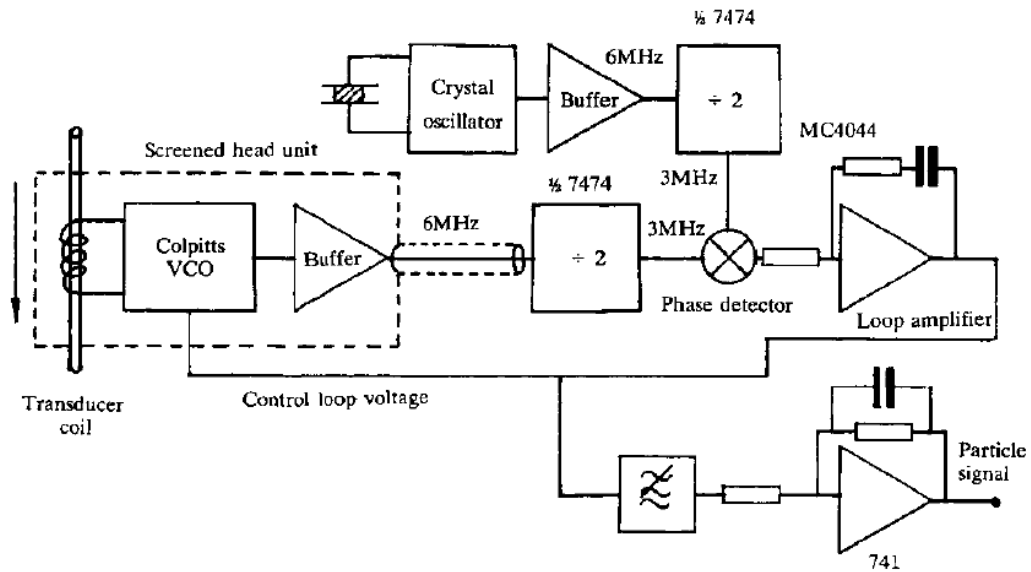


Figure 4.26. Particle counter based on a voltage control oscillator and PLL
(Whittington et al. 1992).

4.3.4.3 The influence of particle deposition in ferrous debris density measurements

This section introduces the evaluation of the ferrous debris density sensor according to the progression of wear obtained from pin-on-disc configuration. The sensor samples the oil directly from the lubrication bath where the pin-on-disc test is performed. Results reveal important considerations to take into account when using these sensors.

As explained, Kittiwake's ferrous debris density sensor relies on the principles of Magnetometry and it is based on two fundamental physical effects (Kittiwake 2007):

- The change of inductance due to the presence of a magnetic material (changes in permeability).
- The change of energy losses due to electromagnetic induction in any conductive material (Eddy currents losses).

The changes of permeability and the losses due to eddy currents are measured by means of two symmetrical measurement chambers: measure and reference halves as shown in Figure 4.27. The reference half remains empty, while the other chamber is filled with the fluid sample. The sensor measures the difference between the halves. As a result, its output is a signal proportional to ferrous debris density. The sensor provides the ferrous density by weight ($\text{ppm} = \text{g/m}^3$ or $\mu\text{g/ml}$).

This method of sensing is independent of the particle size. The sensor has a repeatability of 5 ppm if the difference of temperature between the two halves is less than 10°C . The volume of the oil inside the sensing chamber is approximately 4ml. The sensor is connected to a PLINT TE92HS rotary tribometer Figure 4.27. When the test is finished the used oil is stored in a laboratory bottle for further characterisation of the deposition of ferrous debris particles under different temperatures. Oil temperature is controlled by a heater block underneath the lubricant bath. Accurate regulation of temperature during the experiment is possible with a K-type thermocouple and PID control.

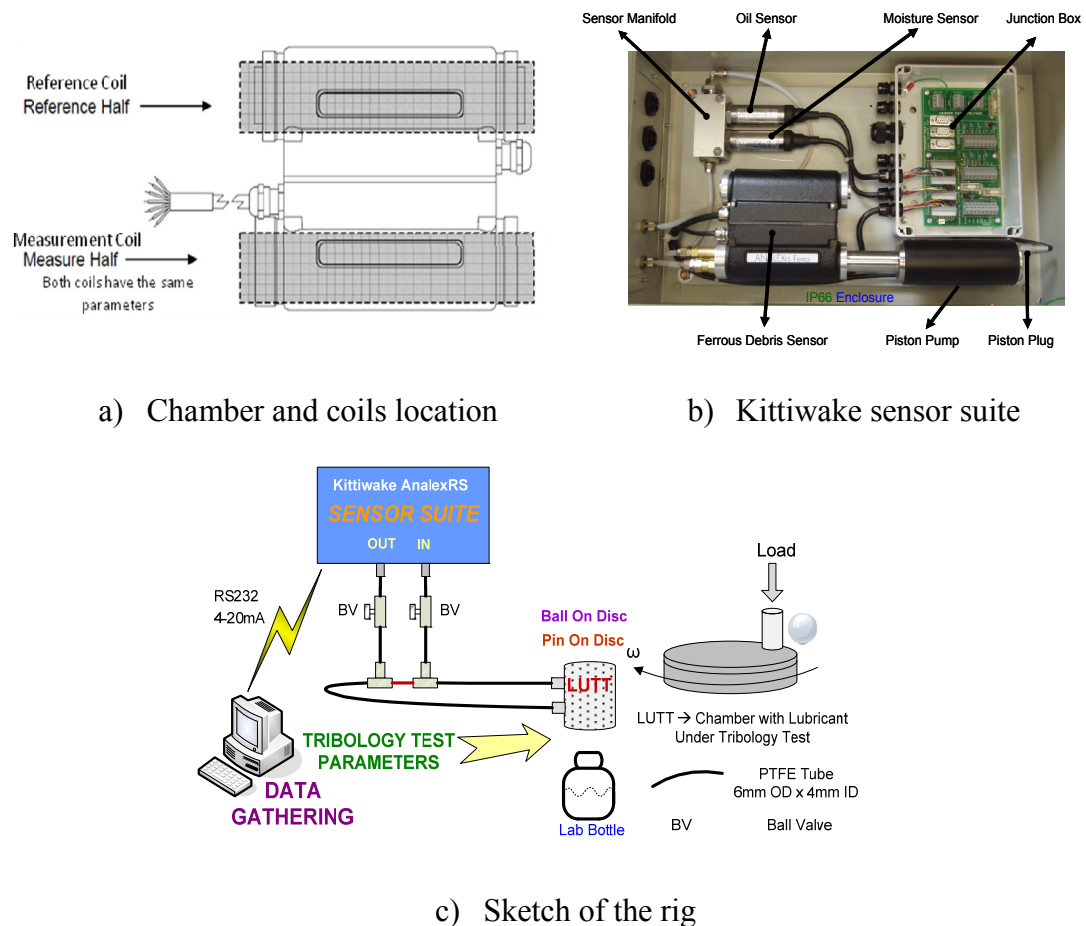


Figure 4.27. Ferrous Debris Density Sensor, Kittiwake Sensor's Suite and experimental rig.

4.3.4.3.1 Test Methodology

When the goal is the characterisation of the output of real time sensors, the selection of the bench test configuration is constrained by the inherent limitations of the sensors. The most important parameters to be taken into account are the sensitivity and repeatability. The sensitivity of the sensor implies a limitation in typical tribological test configurations. For the ferrous debris sensor, the sensitivity is 1ppm but it should be borne in mind that 1ppm for an oil chamber of 4ml corresponds to a total weight of ferrous debris of $4\mu\text{g}$. The sensor has a repeatability of 5ppm thus the minimum discernible mass is $20\mu\text{g}$ (Repeatability is the variation in measurements under the same conditions). This wear mass is equivalent to a wear volume loss in steel (density of steel is 7850 Kg/m^3) of $2.5478 \cdot 10^{-3} \text{ mm}^3$. This amount is above the typical wear volumes in concentrated contact experiments. For example, wear volume results using 4 ball sliding wear test are below this limit as

shown in figures 2.9 and 2.10. For 4 ball sliding wear test arrangement, a specific chamber was used. This chamber allows lubricant circulation and enables the connection of the sensors. Apart from the calculations, experimental results confirm the sensors could not measure any variation in ferrous debris density during the tests under this configuration. Performance of the ferrous debris sensor was also assessed under ball on disc configuration. Samples consisted on a chromium steel ball (EN31) of grade 5 and a carbon steel disc (EN24T). Maximum debris density peaks were around 30ppm under very high loads (150N). Ball-on-disc was discarded as the contact pressures were very high and wear rates not enough to be accurately measured with the sensor. Besides, in some cases the ball rolled inside the holder before reaching the welding point making results misrepresentative.

As a result, a line or area contact is the best option to achieve wear rates measurable by the sensor. Therefore, the selected type of test is pin-on-disc. The disc is EN24T steel hardened to 500-550HV and surface finish less than $R_a=0.2\mu\text{m}$ while the pin is mild steel 150-200HV with a surface finish of $R_a=5\mu\text{m}$, figure 4.29. Within this tribological pair severe wear in the pin is expected. Tests were developed under fully-flooded oil conditions at 60°C (typical operating temperature of hydraulic machinery), normal load of 300N (this corresponds to mean contact pressure assuming an ideal flat surface of 5.97MPa) and a spindle speed of 100rpm (contact speed is 0.2618m/s as the distance from the centre of the pin to spindle centre is 2.5cm). Two types of oils were considered: a pure mineral oil without additives of grade ISO VG46 provided by REPSOL and a commercial additivised hydraulic oil ESSO MOBIL DTE 13M of grade ISO VG32. The quantity of oil per test is 125ml and it is pumped with a flow rate of 4ml/min.

The contact area was operated outside the visco-elastic range so there were only two possibilities: either a lubricant film separated the mating surfaces or not. The pin was machined with a flat profile without chamfer in order to avoid a converging wedge inlet which could generate a lubricant film. The reading of high values of coefficient of friction ensured that the experiment run in a boundary lubrication regime. Several parameters were collected during the test: the applied load, the speed, the temperature, the coefficient of friction and the ferrous debris density.

4.3.4.3.2 Results

Ferrous debris density measurements for the two reference oils are shown in Figure 4.28 with an approximate value of friction coefficient shown in Table 4.4. During the running-in process the coefficient of friction decreased until it stabilised. At this point, the waviness and roughness of the two contact surfaces were similar with a high degree of conformity.

The comparison of ferrous density results was evaluated by means of normalising the ferrous debris density using the peak value. As was expected the wear was less when the lubricant had an additive package. Important trends are highlighted in Figure 4.28.

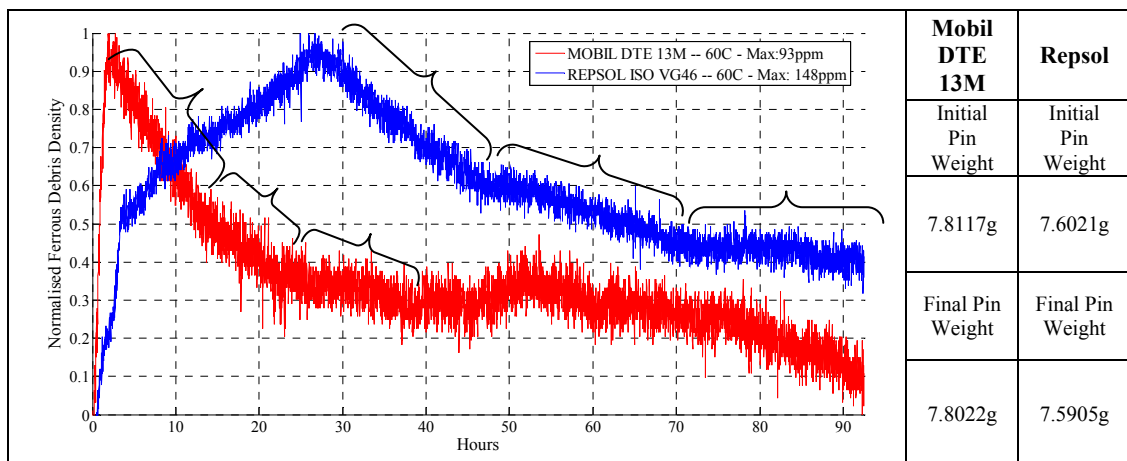


Figure 4.28. Normalised ferrous debris density

Table 4.4. Coefficient of friction

	0 min	15 min	30min	45min	60min	90min	120min	48hrs
Repsol	0.08	0.08	0.08	0.08	0.07	0.05	0.04	0.01
Mobil DTE 13M	0.08	0.07	0.04	0.04	0.04	0.04	0.04	0.01

The evolution of wear is shown in Figure 4.29. This graph is obtained by doing experiments with the same pin running 12 repetitions at regular intervals of 10min. The pin was properly aligned to the same position before each test. Figure 4.29 is especially helpful to identify the matching degree of sensor response and the real wear in the tribological pair. A representation of roughness evolution of the pin and the mass loss of the pin is shown. The roughness is measured using a white light interferometer after subtracting the reference surface (a plane), it is the average

roughness of the pin counterface. Some topographic images are shown in Figure 4.30 at different sampling times. As shown in Figure 3.28, after 2 hours the wear rate is almost in the stabilised region (this concept appears in Figure 1.2). According to Figure 4.28, the pin losses approximately 0.0095g after 4 days (Mobil DTE 13M results). Results after two hours, Table 4.5 confirm the weight loss is 0.006g which represents a 62% compared to the total weight loss after 4 days. Therefore, after two hours the wear rate is in the stabilised wear rate region (there is wear but at much lower rate than the running-in period). The sensor predicts this effect with a delay of almost 30 min. After the wear is stabilised, the influence of the particle deposition process plays an important role. The generated wear is negligible compared to the effect of particle deposition phenomena, thus the ferrous debris density trend drops quickly. From Figure 4.28 different slopes are easily identified. Each slope corresponds to a different range of particle sizes. Small particles are in suspension for a longer time than large ones. The type of particles is measured using ferrographic techniques (Poley 2001; Roylance 2005; Roylance and Hunt 1999; Roylance et al. 2004). The particles are trapped in a glass specimen using a variable magnetic field. Different particle sizes and clusters of particles are shown in Figure 4.31. One of the problems encountered performing linear ferrography (Roylance and Hunt 1999) is the formation of particle clusters due to magnetic field lines. This problem can be further reduced using rotary ferrography.

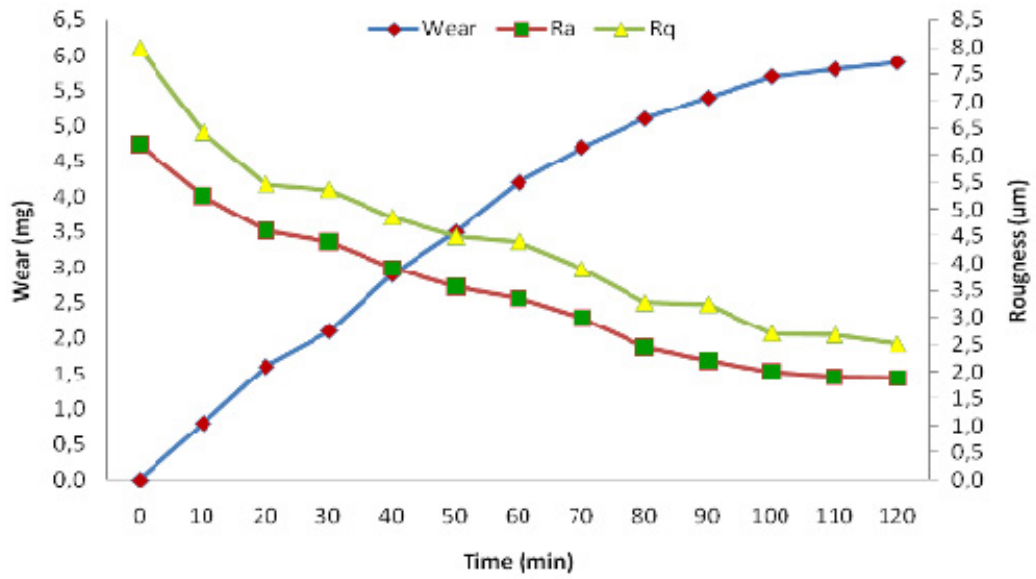


Figure 4.29. Evolution of wear and roughness of pin counterface (Mobil DTE 13M)

Table 4.5. Roughness evolution of the pin. The lubricant is Mobil DTE 13M

Time (min)	Weight (g)	Time (min)	Weight (g)	Time (min)	Weight (g)
0	7.7983	50	7.7948	100	7.7926
10	7.7975	60	7.7941	110	7.7925
20	7.7967	70	7.7936	120	7.7924
30	7.7962	80	7.7932		
40	7.7954	90	7.7929		

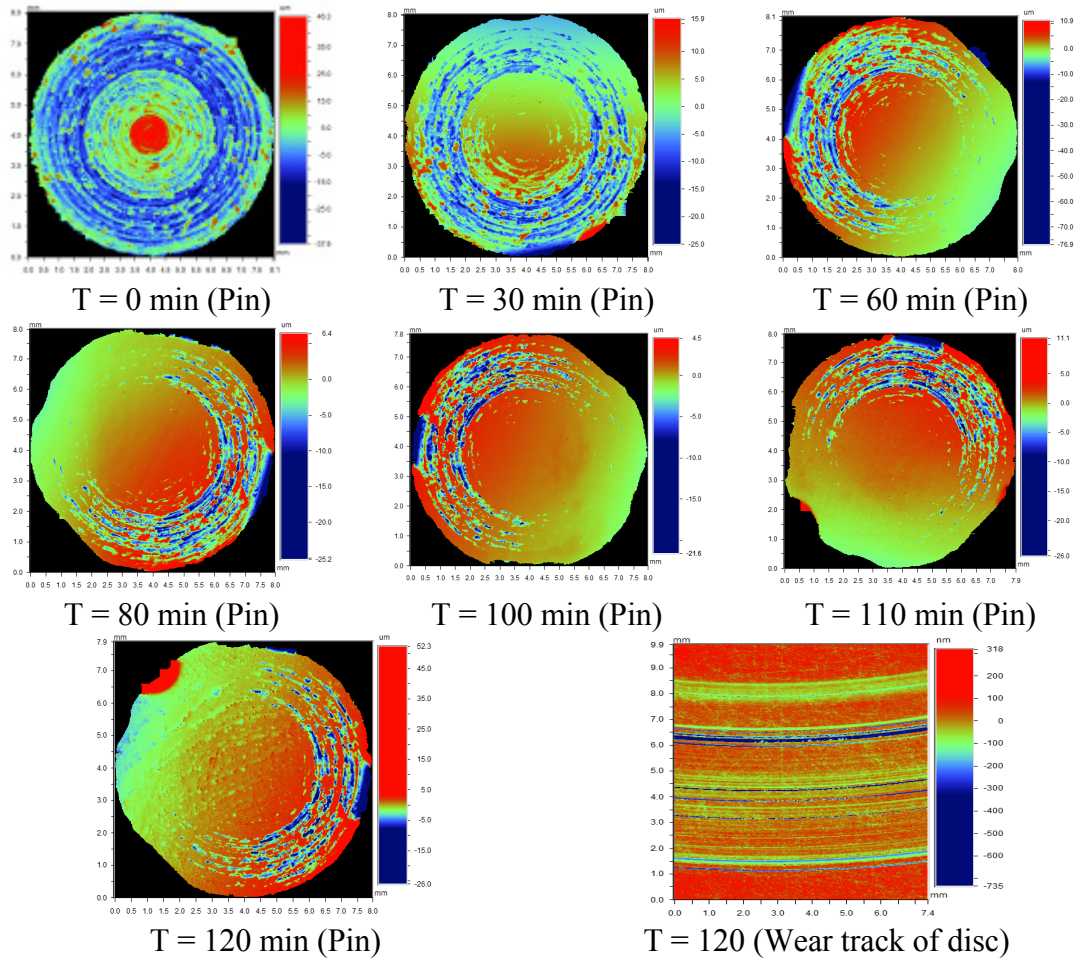


Figure 4.30. Topographical images of the evolution of wear

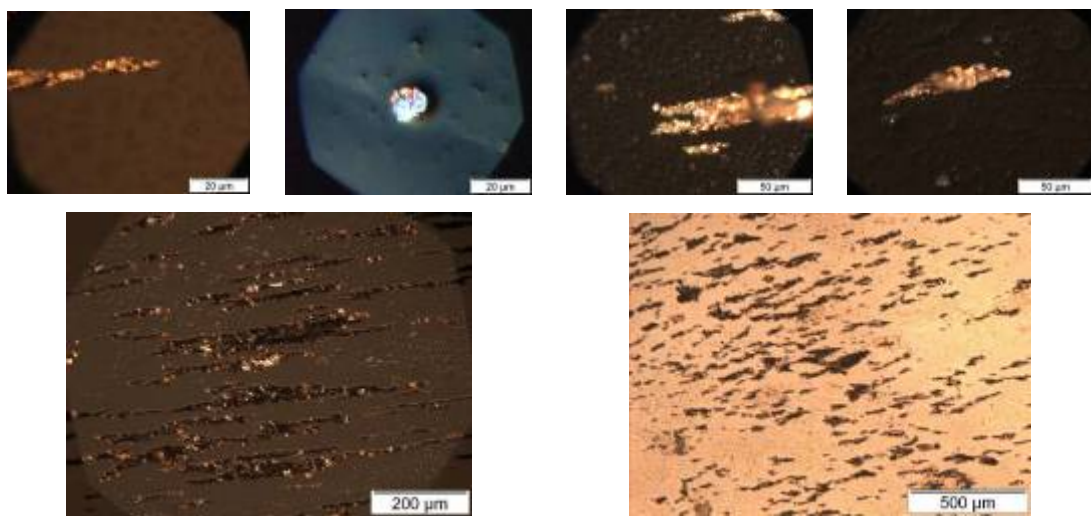


Figure 4.31. Ferrogram analyses of used oil

The characterisation of temperature influence within the particle deposition process is done by storing the used oil in a bottle. Before taking any measurements the lubricant is stirred and the temperature is stabilised using a hot-plate. Oil viscosity decreases as temperature increases. As a result, when the temperature increases, particles sink more quickly.

The experimental results confirm that ferrous debris density monitoring can be a late indicator of the incipient fault detection depending on the particle deposition process. It should be considered that the on-line detection of ferrous debris density has an inherent lag depending on the sampling point. Although this lag affects the response time in the fault detection process it can be determined if the flow and pipe distance from the sensor are known.

The effectiveness of ferrous debris density measurements depends on the uniform suspension of the debris. Therefore, this monitoring technique is reliable in systems with small sumps and rapid circulation of the lubricating oils. It can be concluded that ferrous debris density is suited to small-sump systems with a rapid oil circulation loop in which the major wearing components are ferrous-based, and in which there is no filtration or at least relatively coarse filtration. Good candidates include reciprocating engines, power train components, aviation turbine applications, engines and gear-heavy components, like most automotive and industrial gearboxes and transmissions. If this monitoring technique is applied to machinery with stationary lubrication oils and large sumps the effective fault detection cannot be guaranteed. An extreme case of misrepresentative results may arise when machinery operates under start-stop conditions with low utilisation factor. In this case, the incipient fault detection using ferrous debris density measurements can only be achieved if the particle deposition process can be determined within machinery.

Experimental results reveal that particle deposition is a complex physical phenomenon mainly dependent on temperature, particle size, oil viscosity and the capacity to stir the oil. Therefore, modelling the particle deposition process in real machinery implies a significant challenge.

4.4 Design of the oil condition monitoring rig

4.4.1 Selection criteria of oil condition sensors

A rig was built to facilitate the use of commercial oil condition sensors under accelerated bench test configurations. These sensors were selected according to the following criteria:

- Specifications.
 - *Viscosity determination.* After an exhaustive search of the different types of process viscometers (section 3.3.1.2) three candidates were shortlisted Brookfield AST100, Cambridge Viscosity SPC/L372(J) and Vectron ViSmart Low Shear Sensor. The final decision was supported in terms of quality versus price and technology considerations, Table 4.6.

Table 4.6. Process viscometer selection according to quality versus price

	Process viscometer		
	Cambridge Viscosity SPC/L372(J)	Brookfield AST100	SenGenuity Threaded-Bolt (Low Shear)
Technology	Electromagnetic piston	Resonator	Resonator (SAW)
Specifications	Excellent accuracy up to ± 0.1 cP	Good accuracy +1%	Lowest accuracy ($\pm 10\%$)
	Broad range of viscosities depending on the type of piston	Excellent working range of viscosities 2-3000cP	Difficult to find an accurate model to relate acoustic viscosity to real viscosity
Final price	≈£10000	≈£4000	≈£3000

- *Water content in the oil.* Four sensors were shortlisted: Macom, Vaisala Humicap, Kittiwake, EESIFLO EASZ-1. EASZ-1 was discarded because of its high cost and because it is technologically based on changes in the dielectric constant (same principle as the qualitative oil

condition sensor). The others are based on a polymer that responds to water vapour changing its dielectric constant. The Kittiwake unit was finally selected because its output was assessed against Vaisala Humicap sensors (the leader in moisture-in-oil devices) (Murphy and Kent 2004).

- *Ferrous debris density or particle counters.* Three companies are the leaders in the market: Kittiwake, Macom Technologies, GasTops MetalScan. Kittiwake was selected because it is the only company that provides a ferrous debris density sensor. For particle counters according to the specifications Macom TE-10 and Kittiwake are best.
- *Oil quality.* Lubrigard, Kittiwake and Voelker sensor. The first two sensors are based on the same principle whilst the last one was discarded because it is not clear how the ageing process of the polymer affects its sensitivity.
- Cost and origin of manufacture (giving priority to UK providers and manufacturers).
 - Kittiwake and Macom Technologies had priority because they are UK based companies.
- Delivery time.

Brookfield AST100 was selected because its design is used to monitor viscosity in many industrial processes and the quality versus price is best compared to the other alternatives. Kittiwake's sensors were selected because they provide a sensor suite with all sensors needed to perform the measurements which facilitates the rig construction and they also quote the items with an academic discount.

The final rig contains the last commercial oil condition monitoring technology and it is valid to establish comparison with future sensor development.

4.4.2 Test rig design

A block diagram of the rig with default connections is shown in Figure 4.32. The connections can be easily modified with standard pneumatic pipe adaptors. A

real picture of the rig is shown in Figure 4.33. For pin and ball-on-disc tests the lubricant chamber (lubricant bath) is shown in Figure 4.33.

The rig is intended to be used with low fluid pressures, hence the use of pneumatic fittings. The rig is based on a fast oil circulation loop with an intermittent flow driven by a peristaltic pump to avoid any contamination due to the pump operation. The thermoelectric cooling plate allows accurate thermal regulation of the fluid temperature. Kittiwake's on-line sensor suite was placed after the chamber where the tribological test was running to monitor the debris released during the experiment. The viscosity of the oil is preferably measured before the oil enters the chamber where the tribological test is running because it has an oil chamber with a geometry that can trap the debris.

The rig was designed to be easily adaptable to different bench test configurations which in the present research were used as follows:

For detecting water contamination in lubrication oils by means of using viscosity measurements (Section 4.3.1.5) it is used the peristaltic pump, the cooling plate and the viscometer.

To generate debris (experimental part considered in section 4.3.4.3), it is used the sensor suite and the lubricant bath for pin on disc tests.

The novel methodology based on particle deposition theory of particles due to gravity requires the sensor suite and a lab bottle that contains the contaminated oil.

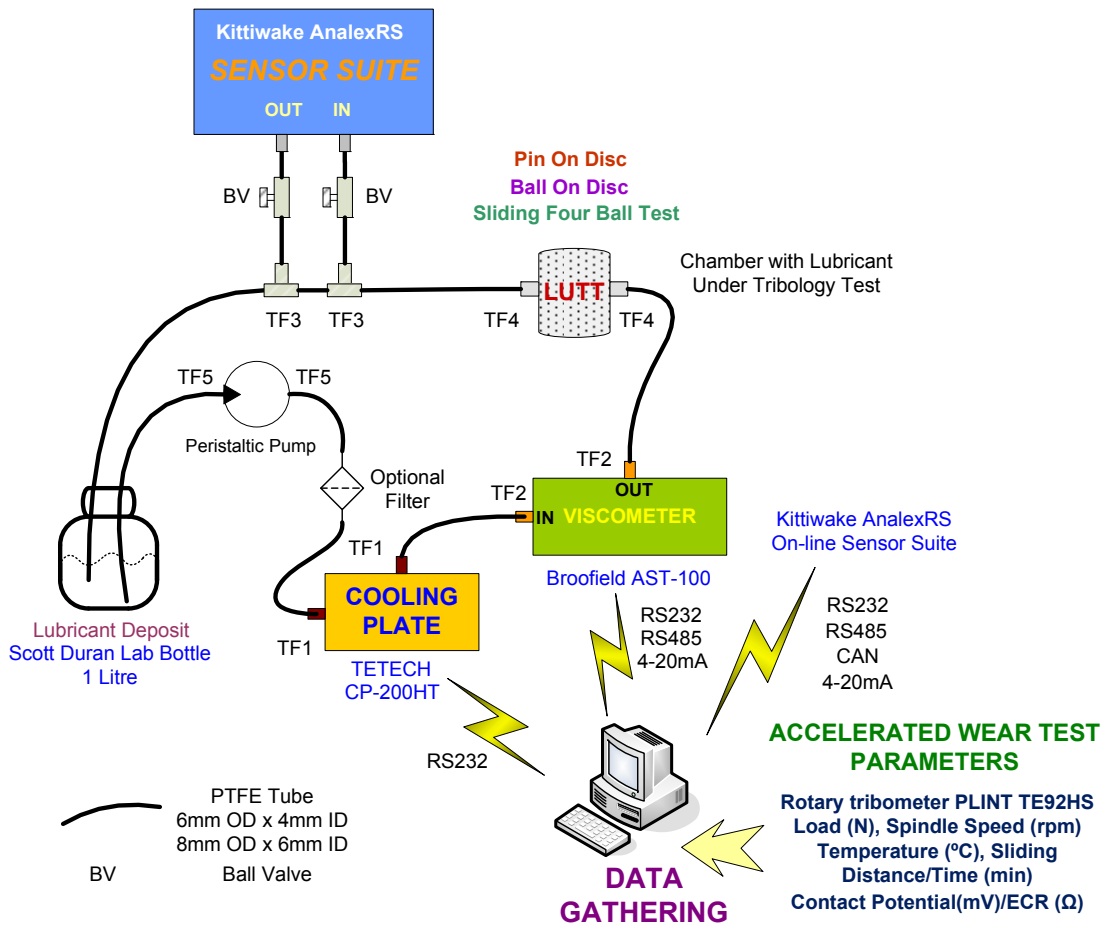


Figure 4.32. Functional sketch of the oil condition monitoring rig

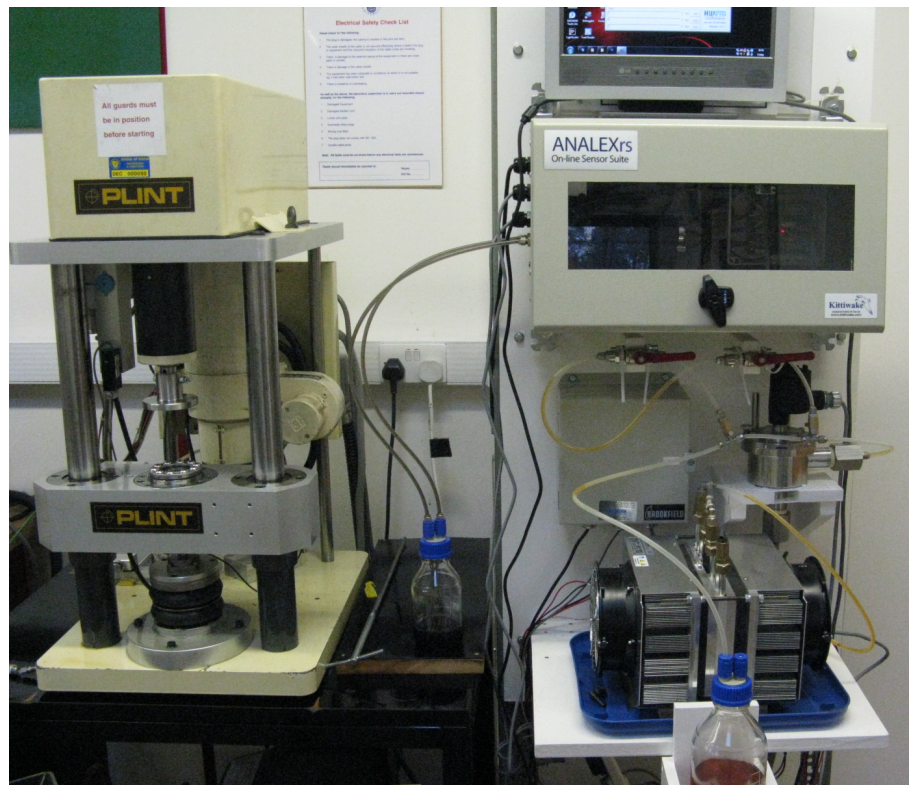


Figure 4.33. Oil condition monitoring rig and the pin-on-disc assembly for lubricated test

The rig has different elements that are shown below:

- Lubricant deposit. The deposit is a lab bottle with GL45 standard thread. The screw cap system of these bottles allows the hose connection of 6mm O.D PTFE tubing. The cap is available with two and three ports connections.
- Peristaltic pump. This kind of pump avoids fluid contamination and it is easy to maintain. The cooling capability of the rig is strongly dependant on flow.

Therefore, different dosing rates and pipe diameters are required for flow selection.

- Optional filter. Particle contamination can be removed using a filter.
- Thermoelectric cooling plate. This device allows accurate control of oil temperature. Two cooling plates are assembled together to duplicate the cooling power.
- Viscometer. This sensor provides the temperature and kinematic viscosity of the oil. The volume of oil in the viscometer chamber is approximately 100ml.
- Sensor suite. This suite has one ferrous debris sensor mounted with a piston pump that sucks the oil from the main stream (fast circulation loop), one oil condition sensor and one water-in-oil sensor. It pumps at a rate of 4ml per minute.
- Chamber with lubricant under tribology test. Different tribology experiments can be performed using different chambers for lubricant circulation. There are two chambers:
 - For pin-on-disc and ball-on-disc the lubricant bath designed by PLINT.
 - For Four-Ball Sliding Wear Test a custom pressurised chamber design can be used.

5 Determination of particle size distribution from ferrous debris density measurements

This chapter covers all the theory necessary for the determination of the particle size distribution from debris density measurements of solid particles falling under gravity through a fluid. A novel methodology is presented which solves this inverse problem. A promising application of this methodology is the development of low cost particle counters in the micrometer and submicrometer range for wear debris monitoring in industrial machinery. As a result, this methodology could lead to a considerable breakthrough in respect to the current technology which is based on direct particle detection using magnetometry techniques which presents serious size discrimination limitations.

The most accurate particle size distribution techniques are based on light scattering. Although dynamic and static laser light scattering techniques such as the Focused Beam Reflectance Method (FBRM) can be used in-situ, the main limitation encountered when used in the lubrication industry is the misleading results caused by the high difference in the refraction index of contaminants in oil. The measurement technique is dependent on the size, shape, environment and the complex refractive index of the particle. Air bubbles, dirt and water droplets (when the lubricant is emulsified) trapped in the oil result in misleading results. Optical counters can be reliably used in wear debris monitoring only if there is an adequate sample preparation for removing these contaminants from the oil. This last shortcoming makes optical counters inadequate for real time monitoring of particle size distribution in oils.

Finally, another useful application of the proposed methodology is the evaluation of the motion of particles in close proximity when the particle size distribution is known. Under such conditions, the particle density is very high and the use of this methodology could be interesting for the evaluation and determination of the effective drag coefficients which are of importance in sedimentation theory.

5.1 Classification of a single particle

The classification of particles according to their size depends on the shape. Regular-shaped particles can be accurately described by giving the shape and a number of dimensions. Some particle shapes and dimensions are shown in Table 5.1

Table 5.1. Examples of shape and dimensions of particles (Rhodes 2008)

Shape	Sphere	Cube	Cylinder	Cuboid	Cone
Dimensions	Radius	Side length	Radius and height	Three side lengths	Radius and height

However, the size of irregular-shaped particles cannot be described by a single physical dimension. Therefore, the dimension to be used is constrained by the measurement technique. For example, Figure 5.1 shows some common diameters used in microscope analysis e.g. statistical diameters such as Martin's diameter (length of the line which bisects the particle image), Feret's diameter (distance between two tangents on opposite sides of the particle) and equivalent circle diameters such as the projected area diameter (area of circle with same area as the projected area of the particle resting in a stable position) (Rhodes 2008). The orientation of the particle on the microscope slide will affect the projected image and consequently the measured equivalent sphere diameter. If a sieve is used to measure the particle size then an equivalent sphere diameter is defined, this is the diameter of a sphere passing through the same sieve aperture. If a sedimentation technique is used to measure particle size then this diameter is expressed as the diameter of a sphere having the same sedimentation velocity under the same conditions. As a result, it is important to use the method of size measurement which directly gives the particle size relevant to the situation or process of interest.

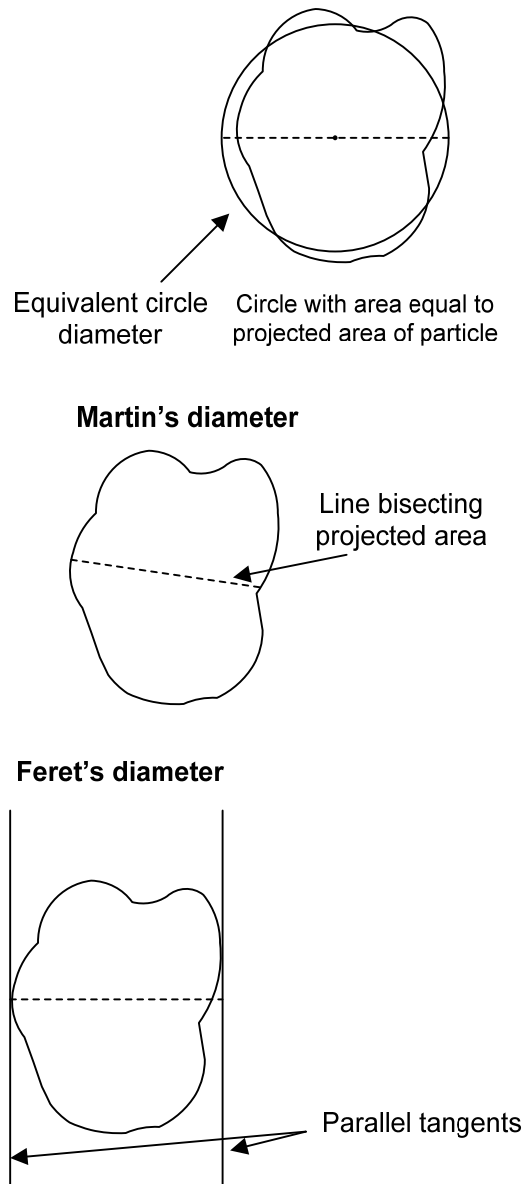


Figure 5.1. Single particle diameters used in microscope analysis (Rhodes 2008).

The shape of a non-spherical particle can be described for convenience by a single number. One simple approach is to describe the shape of a particle in terms of its sphericity, the ratio of the surface area of a sphere of volume equal to that of the particle to the surface area of the particle. For example, a cube of side one unit has a volume of 1 (cubic units) and a surface area of 6 (square units). A sphere of the same volume has a diameter of 1.24 units. The surface area of a sphere of diameter 1.24 units is 4.836 units. The sphericity of a cube is therefore 0.806 (4.836/6). The

sphericity is a suitable parameter to describe the drag coefficients of non-spherical particles.

5.2 Motion of particles under gravity through a fluid

The motion of particles under gravity through a fluid is analysed considering the forces applied to a single particle, Figure 5.2.

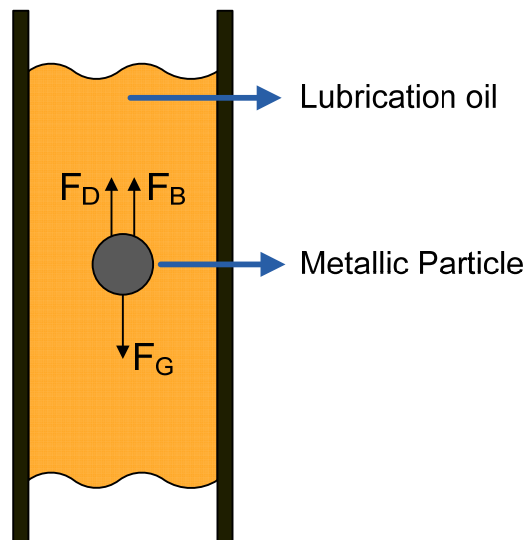


Figure 5.2. Particle travelling through the fluid under the action of three forces.

The particle of mass m is moving through a fluid under the action of an external force F_G . The velocity of the particle is U and the density of the particle is ρ_p . There are three forces acting on a particle moving through a fluid:

1. The external force, gravitational or centrifugal (F_G): where v is the volume of the particle, ρ_p is the particle density and g is the acceleration due to gravity, equation (5-1).

$$F_G = \rho_p v g \quad (5-1)$$

Equation (5-1). Gravitational force acting on the particle

2. The buoyant force, a force acting parallel with the external force but in an opposite direction (F_B) given by Archimedes principle: where ρ_f is the density of the fluid through which the particle is moving (lubrication oil), equation (5-2).

$$F_B = \rho_f v g \quad (5-2)$$

Equation (5-2). The buoyant force acting on the particle

3. The drag force due to the relative motion between the particle and the fluid (F_D) is given by Stokes law applied to particles with a Reynolds number less than 1: where μ is the dynamic viscosity of the fluid (lubrication oil), D_p is the diameter of the particle and U is the velocity of the particle in oil, equation (5-3).

$$F_D = 3 \cdot \mu \cdot U \cdot \pi \cdot D_p \quad (5-3)$$

Equation (5-3). The drag force acting on the particle (Stokes region)

The drag force can be expressed in terms of the drag coefficient. The drag force can be obtained for particles with a Reynolds number greater than one if the drag coefficient of the particle is known, equation (5-4).

$$F_D = \frac{1}{2} C_D \cdot \rho_f \cdot U^2 \cdot A \quad (5-4)$$

Equation (5-4). The drag force acting on the particle

Equation (5-5) is the particular case for spherical particles where the projected area is $A = (\pi/4) \cdot D_p^2$

$$F_D = \frac{1}{2} C_D \cdot \rho_f \left(\frac{dx}{dt} \right)^2 \frac{\pi}{4} D_p^2 = \frac{1}{8} \pi \cdot C_D \cdot \rho_f \left(\frac{dx}{dt} \right)^2 D_p^2 \quad (5-5)$$

Equation (5-5). The drag force acting on a spherical particle

When the particle is at rest, the velocity and the distance travelled by the particle is zero. Then, the initial conditions of the problem are assumed as $t=0$, $U = u = \frac{dx}{dt} = 0$ and $x=0$;

The differential equation and the initial conditions that represent the motion of the particle are obtained by applying Newton's 2nd law, where the resultant force applied to the particle is equal to the mass multiplied by the acceleration of the particle. In expression, v is the volume of the particle and ρ_p is the particle density, equation (5-6).

$$F_G - F_B - F_D = \rho_p \cdot v \frac{d^2x}{dt^2} \quad (5-6)$$

$$\text{Initial conditions: At } t=0, u = \frac{dx}{dt} = 0 \text{ and } x=0$$

Equation (5-6). Differential equation representing the motion of a particle through a fluid.

Substituting equations (5-1, 5-2, 5-5) into the previous equation lead to the mathematical representation of the motion of the particle, equation (5-7).

$$v \cdot g (\rho_p - \rho_f) - \frac{1}{8} \pi \cdot C_D \cdot \rho_f \left(\frac{dx}{dt} \right)^2 D_p^2 = \rho_p \cdot v \frac{d^2x}{dt^2} \quad (5-7)$$

$$\text{Initial conditions: At } t=0, u = \frac{dx}{dt} = 0 \text{ and } x=0$$

Equation (5-7). Differential equation representing the motion of a particle through a fluid (General case)

The analytical solution for the general equation (5-7) is a hard problem because the drag force is expressed in terms of the drag coefficient which is dependent on the Reynolds number of the particle. The Reynolds number depends on the particle velocity and it is stabilised to a certain value when the force equilibrium is reached (at this point the particle velocity is the terminal velocity). Therefore, this problem should be solved numerically considering the evolution of the drag coefficient with the time. For the Stokes region, the differential equation is not

dependant on the drag coefficient of the particle and therefore, there is a theoretical solution for equation (5-8).

$$C_D = \frac{24}{\text{Re}_p} \quad \text{Re}_p = \frac{D_p U \rho_f}{\mu} = \frac{D_p \left(\frac{dx}{dt} \right) \rho_f}{\mu}$$

$$v \cdot g (\rho_p - \rho_f) - 3 \cdot \mu \cdot \pi \cdot D_p \frac{dx}{dt} = \rho_p \cdot v \frac{d^2 x}{dt^2} \quad (5-8)$$

$$\text{Initial conditions: At } t=0, u = \frac{dx}{dt} = 0 \text{ and } x=0$$

Equation (5-8). Motion of a particle (Stokes region)

The theoretical solution of the differential equation (5-8) for the Stokes Law region is shown in equation (5-9).

$$\Delta \rho = (\rho_f - \rho_p)$$

$$x = \frac{\rho_p v^2 g \Delta \rho}{(3\pi\mu D_p)^2} e^{\left(-\frac{3\pi\mu D_p}{\rho_p v} \right) t} + \frac{vg \Delta \rho}{(3\pi\mu D_p)} t - \frac{\rho_p v^2 g \Delta \rho}{(3\pi\mu D_p)^2} \quad (5-9)$$

Equation (5-9). Position of the particle (Stokes region)

The terminal velocity from a mathematical point of view is the maximum value of the derivate with respect to the time of the former equation, and in physical terms this is obtained when the force equilibrium is reached (the particle is not subject to any acceleration), equation (5-10).

$$\Delta\rho = (\rho_f - \rho_p)$$

$$\frac{dx}{dt} = \frac{vg\Delta\rho}{3\pi\mu D_p} e^{\left(-\frac{3\pi\mu D_p}{\rho_p v}\right)t} - \frac{vg\Delta\rho}{3\pi\mu D_p}$$

$$V_T = \max\left(\frac{dx}{dt}\right) = -\frac{vg\Delta\rho}{3\pi\mu D_p} = \frac{D_p^2 g (\rho_p - \rho_f)}{18\mu} \quad (5-10)$$

Or physically when the force equilibrium is reached

$$v \cdot g (\rho_p - \rho_f) - 3 \cdot \mu \cdot \pi \cdot D_p V_T = 0$$

Equation (5-10). Velocity of the particle (Stokes region)

Figures 5.3, 5.4, 5.5 show the solution of the equation for different particle diameters and assuming the particles are in the Stokes Region. A complete sweep of particle diameters from 10 μ m to 1mm is used, to show graphically the degree of influence of the size of the particles to reach the particle terminal velocity. The experimental conditions consider steel particles $\rho_p = 7950 \text{ kg/m}^3$ and a hydraulic oil with a fluid density of $\rho_f = 876 \text{ kg/m}^3$ and kinematic viscosity 46 cSt@40°C.

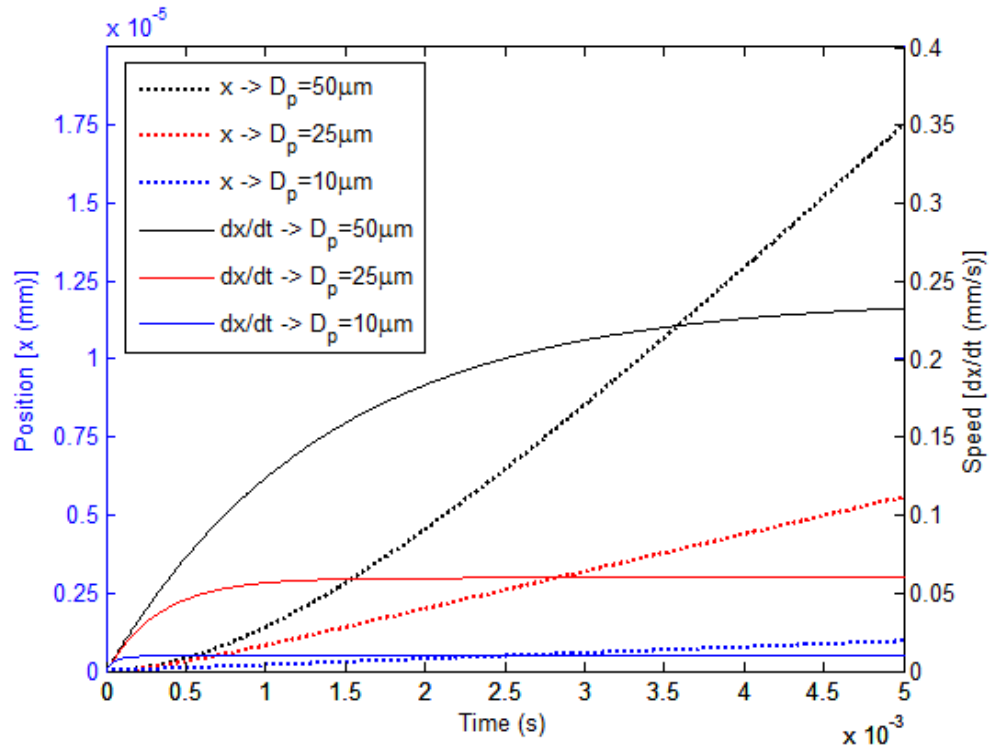


Figure 5.3. Particle travelling through the fluid (Stokes region).

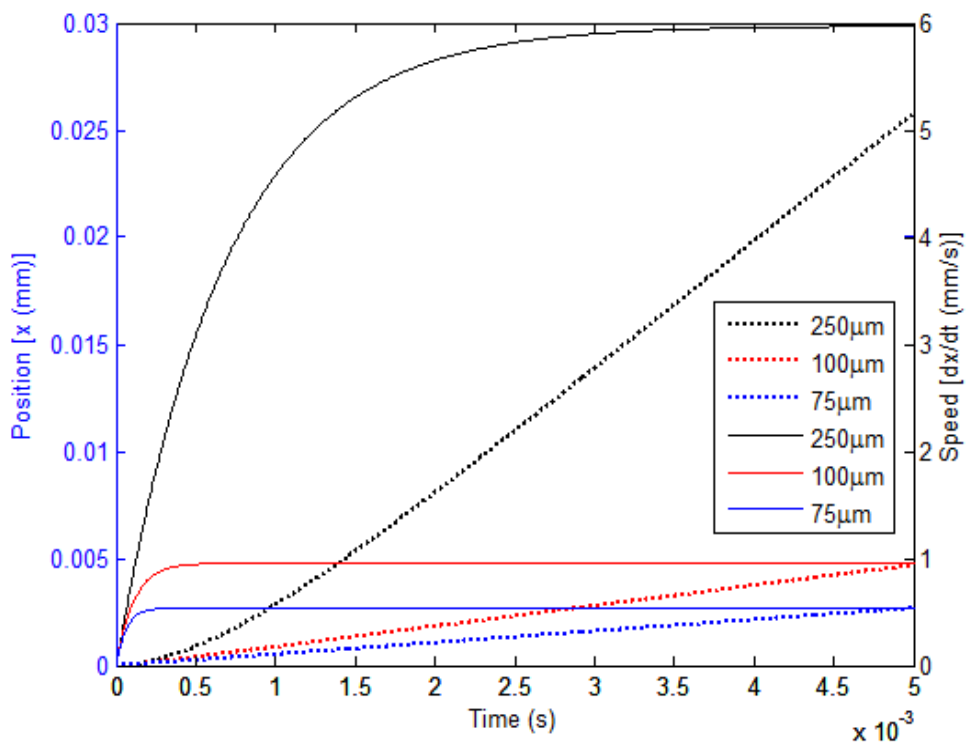


Figure 5.4. Particle travelling through the fluid (Stokes region)

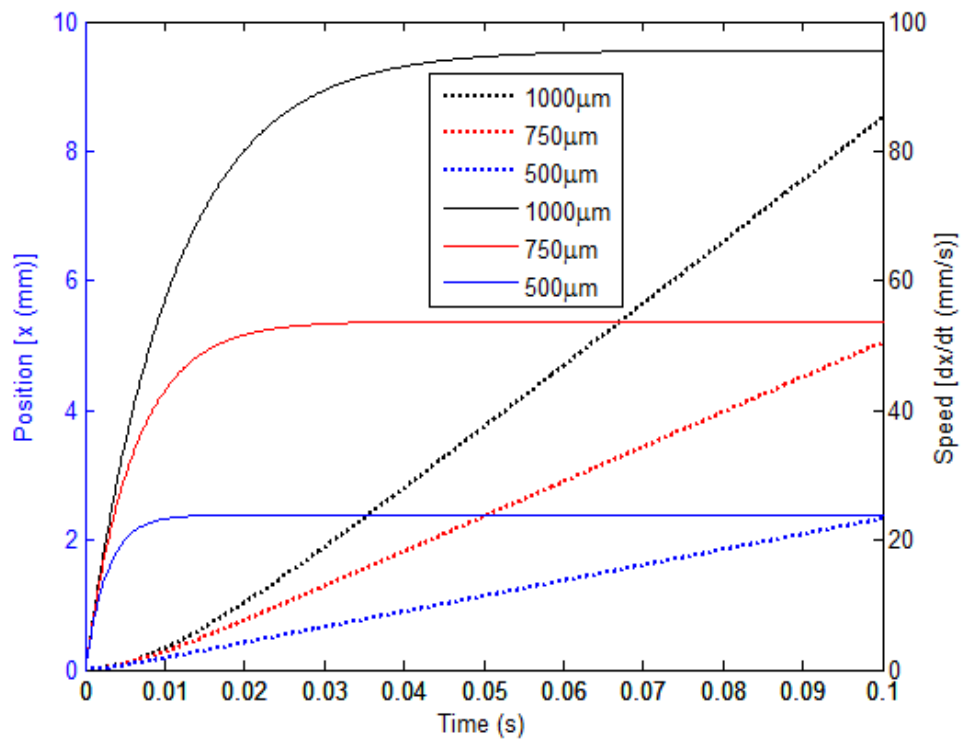


Figure 5.5. Particle travelling through the fluid (Stokes region)

In general, see equation (5-7), the terminal velocity is obtained when the force equilibrium is reached and the resultant force acting on the particle is nil, equation (5-11).

$$v \cdot g (\rho_p - \rho_f) - \frac{1}{8} \pi \cdot C_D \cdot \rho_f \cdot V_T^2 \cdot D_p^2 = 0 \quad (5-11)$$

$$V_T = \sqrt{\frac{4 \cdot g \cdot D_p}{3 \cdot C_D} \left[\frac{\rho_p - \rho_f}{\rho_f} \right]}$$

Equation (5-11). Velocity of the particle (Stokes region)

5.3 Determination of the drag coefficient for a given particle diameter

The drag coefficient for a single particle depends on the Reynolds Number of the particle. These relationships (C_D versus Re_p) are experimentally obtained and they can be found for spherical and non-spherical particles. For spherical particles, of all the relationships cited in the literature, the Haider and Levenspiel's expressions are the most referenced (Haider and Levenspiel), equation (5-12).

$$C_D = \frac{24}{Re_p} \left(1 + 0.1806 \cdot Re_p^{0.6459} \right) + \left(\frac{0.4251}{1 + \frac{6880.95}{Re_p}} \right) \quad (5-12)$$

$$Re_p < 2 \cdot 10^5$$

Equation (5-12). Velocity of the particle (Stokes region)

For non spherical particles, some of the drag coefficient correlations are expressed in terms of the sphericity such as the Haider, Levenspiel and Ganser (Ganser 1993; Haider and Levenspiel 1989). The Haider and Levenspiel correlation is shown in equation (5-13). There are two expressions for variables C and D.

$$C_D = \frac{24}{Re_p} \left(1 + A \cdot Re_p^B \right) + \left(\frac{C}{1 + \frac{D}{Re_p}} \right) \quad Re_p < 2 \cdot 10^5$$

$$A = e^{2.3288 - 6.4581 \cdot \phi + 2.4486 \cdot \phi^2} \quad (5-13)$$

$$B = 0.0964 + 0.5565 \cdot \phi$$

$$C = e^{4.905 - 13.8944 \cdot \phi + 18.4222 \cdot \phi^2 - 10.2599 \cdot \phi^3} \quad \text{or} \quad C = e^{-5.0748 \cdot \phi}$$

$$D = e^{1.4681 + 12.2584 \cdot \phi - 20.7322 \cdot \phi^2 + 15.8855 \cdot \phi^3} \quad \text{or} \quad D = 5.378 \cdot e^{6.2122 \cdot \phi}$$

Equation (5-13). Drag coefficient for non-spherical particles (Haider and Levenspiel 1989)

The coefficients depend on the sphericity ϕ , defined as the ratio between the surface of a volume equivalent spherical particle and the surface of the non-spherical particle.

Ganser correlation (Ganser 1993) distinguishes between isometric and non-isometric particles. For isometric particles, the expression is shown in equation (5-14).

$$C_D = \frac{24}{K_1 \cdot \text{Re}_p} \left(1 + 0.1118 \cdot (K_1 \cdot K_2 \cdot \text{Re}_p)^{0.6567} \right) + 0.4305 \cdot K_1 \cdot K_2^2 \cdot \left(\frac{\text{Re}_p}{3305 + K_1 \cdot K_2 \cdot \text{Re}_p} \right) \quad (5-14)$$

$$\text{Re}_p < 2 \cdot 10^5$$

$$K_1 = \frac{3}{1 + 2 \cdot \phi^{-0.5}} \quad K_2 = 10^{1.8148 \cdot (-\log \phi)^{0.5743}}$$

Equation (5-14). Drag coefficient for non-spherical particles (Ganser)

From the drag curves four regions are identified: the Stokes' law region; the Newton's law region in which drag coefficient is independent of the Reynolds number; an intermediate region between the Stokes and Newton regions; and the boundary layer separation region.

As shown in Figure 5.6, shape affects drag coefficient far more in the intermediate and Newton's law regions than in the Stokes' law region (Rhodes 2008).

Haider, Levenspiel and Ganser relationships are plotted in Figure 5.7.

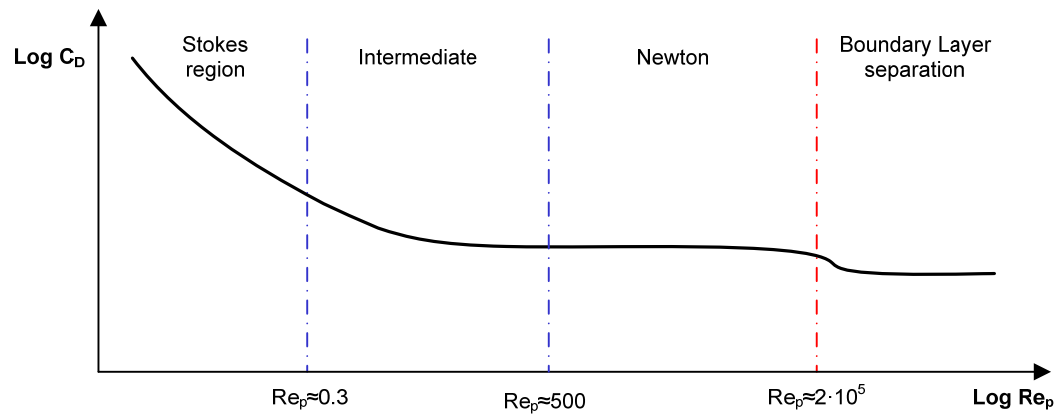


Figure 5.6. Different regions for the drag coefficient of a sphere in a fluid

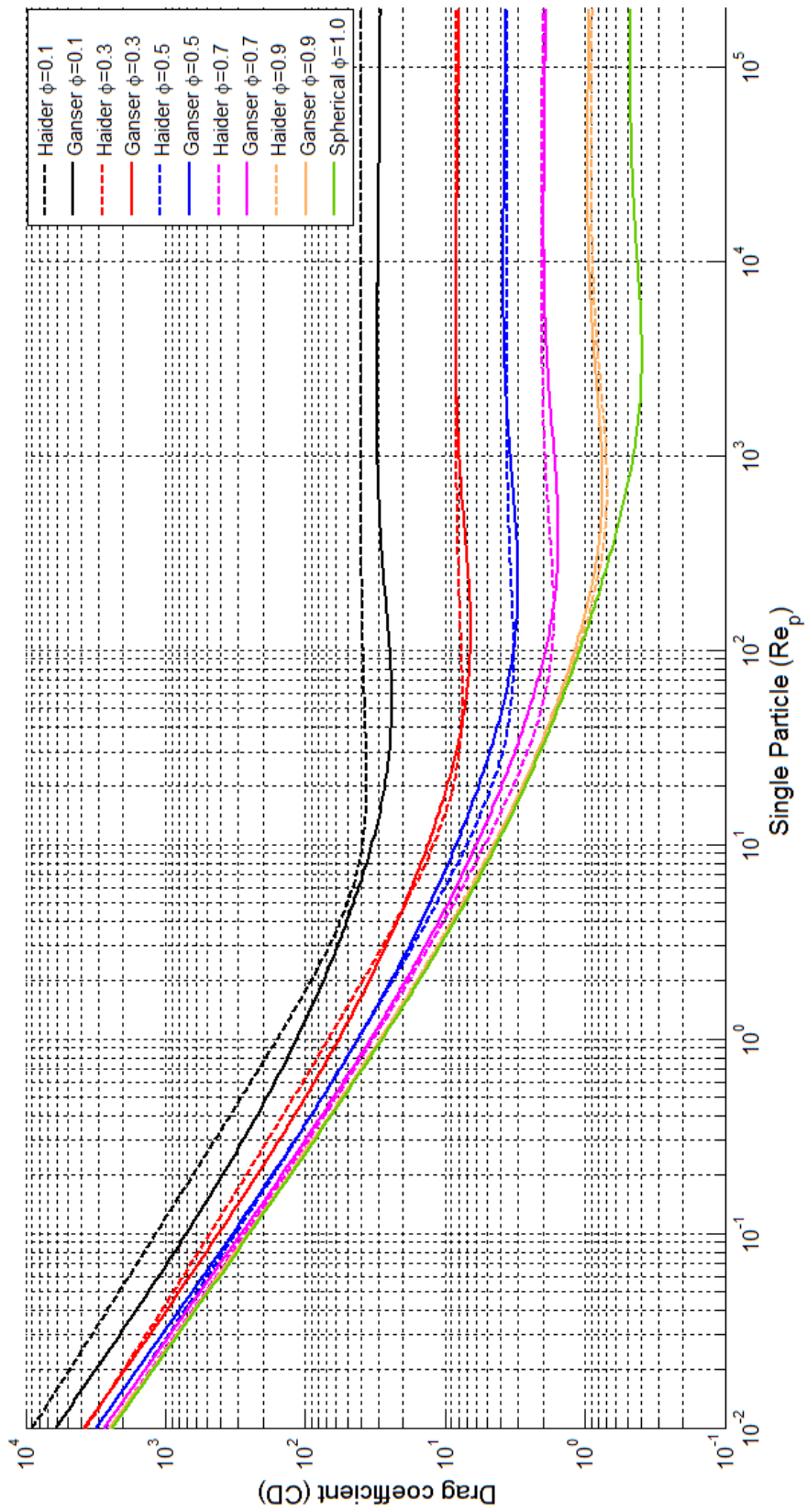


Figure 4.7. Particle travelling through the fluid (Assuming Stokes region)

The drag coefficient for a given particle diameter is obtained intersecting the experimental drag coefficient expressions with a dimensionless number which does not depend on the terminal velocity. This number is obtained from the equations of drag force, Reynolds number and buoyancy equations, equation (5-15).

$$C_D \cdot \text{Re}_p^2 = \frac{4 D_p^3 \cdot \rho_f (\rho_p - \rho_f) g}{3 \mu^2} = \frac{4}{3} \text{Ar} \quad (5-15)$$

Equation (5-15). Dimensionless number to determine C_D from the particle diameter D_p

The particle diameter can be calculated if the terminal velocity is known, intersecting the experimental drag coefficient expression and the dimensionless number of equation (5-16).

$$\frac{C_D}{\text{Re}_p} = \frac{4 g \cdot \mu (\rho_p - \rho_f)}{3 V_T^3 \cdot \rho_f^2} \quad (5-16)$$

Equation (5-16). Dimensionless number to determine C_D from the terminal velocity V_T

The intersection of these two curves is implemented numerically in two ways: using a graphical method developed by Douglas M. Schwarz which allows definition of the intersection of any curves or solution of the non-linear equation using an optimisation algorithm such as Gauss-Newton or Levenberg-Marquardt.

Figure 4.8 is an example of the settling velocities for a steel spherical particle ($\rho_p=7850 \text{ g/m}^3$) of several diameters travelling through a lubricant of kinematic viscosity of 78cSt@20°C, 55cSt@27°C, 26.5cSt@45°C, 10.7cSt@75°C and density of $\rho_f=874 \text{ Kg/m}^3$. The settling velocities are shown assuming the particles are in the Stokes Region and using the real terminal velocity using the procedure mentioned above.

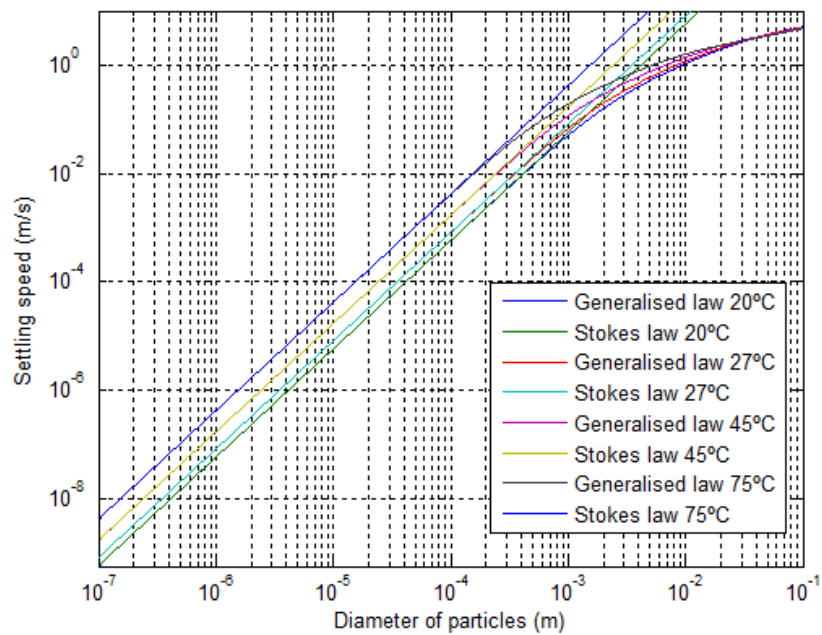


Figure 5.8. Example of settling speeds for different lubricant temperatures

5.4 Methodology for the determination of the particle size distribution from particle density measurements

The particle size distribution from particle density measurements is determined assuming that there is an effective height (H) of particles in uniform suspension inside a container. The experiment is simple; a container (i.e a lab bottle) is filled with known oil (contaminated with debris) up to a certain height H . The contaminated lubricant should be stirred to ensure uniform suspension before performing any particle density measurements. Ferrous debris density readings are taken for a certain period of time or until the particles settle down. Knowing the ferrous debris density evolution over time and applying the proposed methodology, the particle size distribution can be obtained.

5.4.1 Numerical model

The procedure for the calculation of the particle distribution is summarised in Figure 5.9. The flowchart is explained sequentially considering each block.

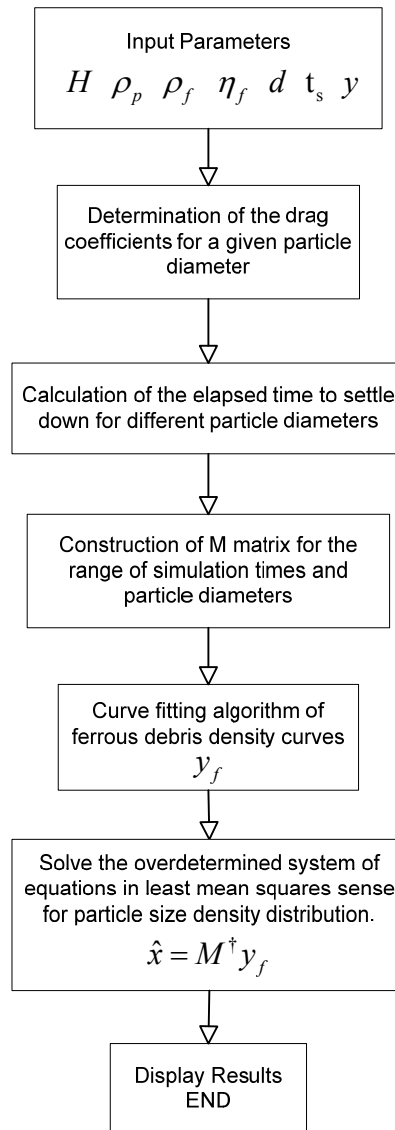


Figure 5.9. Flowchart for the determination of particle size distribution from particle density measurements

5.4.1.1 Input Parameters

The algorithm needs the following parameters:

- | | | |
|----------|---|---|
| H | → | The effective height. The height of the level of the fluid in the container (m) |
| ρ_p | → | particle material density (Kg/m ³) |
| ρ_f | → | fluid density (Kg/m ³) |
| η_f | → | absolute viscosity of the fluid (Pa·s) |

- d → logarithmically equally spaced array of particle diameters (m)
- t_s → array of linear equally distributed times (s)
- y → array of ferrous debris density evolution sampled at t_s (ppm)

5.4.1.2 Determination of the drag coefficients for a given particle diameter

The calculation of the terminal velocity for each particle diameter is based on the mentioned graphical method. This method consists of the intersection of the drag coefficient curve for spherical particles $\Phi=1$ and the curve defined by a dimensionless group which does not depend on the terminal velocity. In logarithmic scale the drag coefficient is determined as a result of the intersection of equation (5-12) with a line of slope -2 according to equation (5-17). In equation (5-17), $\log(A)$ is a different constant for each particle diameter. Although the shapes of wear particles are important to select the most adequate drag coefficient, results from this study are shown for sphericity values of one.

$$\log(C_D) = \log\left(\frac{4}{3} \frac{d^3 \rho_f (\rho_p - \rho_f) g}{\eta^2}\right) - 2 \log(\text{Re}_p) = \log(A) - 2 \log(\text{Re}_p) \quad (5-17)$$

Equation (5-17). Dimensionless number to determine C_D from the particle diameter D_p (logarithm scale)

After the application of this procedure, the Reynolds number and drag coefficients are determined for each particle diameter. The numerical implementation is based on a fast and robust curve intersection algorithm (Schwarz 2006). For the previous example, the relationship between the particle diameters and drag coefficients can be obtained for the four different viscosities as shown in Figure 5.10. It is easy to realise, the oil viscosity plays an important role in the Stokes region while is not as noticeable as the Reynolds number of the particle increases and the particle reaches the Newton region. It is important to consider a simple expression that could fit the data and saves computational time for processing other measurements. For the previous example, a two-exponential expression as shown in Table 5.2 fits well the data of Figures 5.10, 5.11.

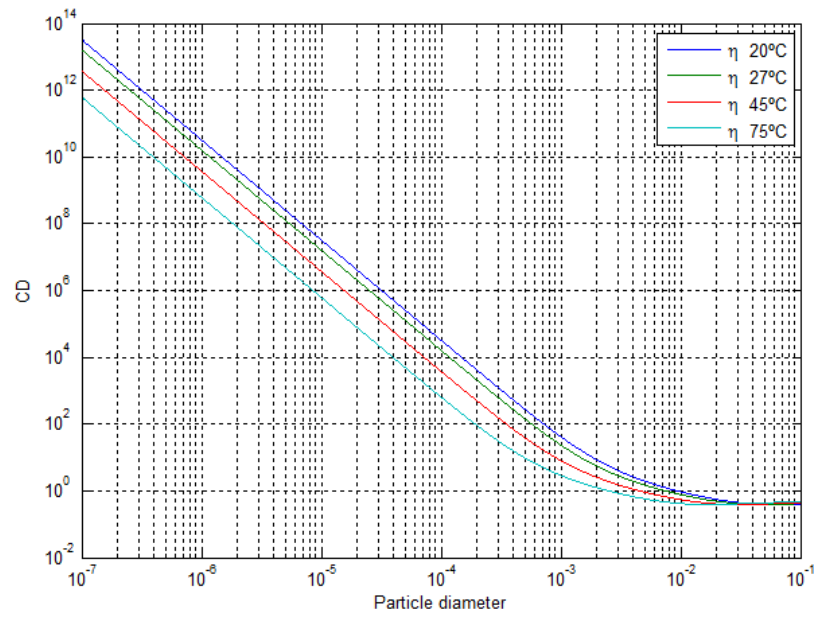


Figure 5.10. Example of drag coefficients for different spherical particle diameters

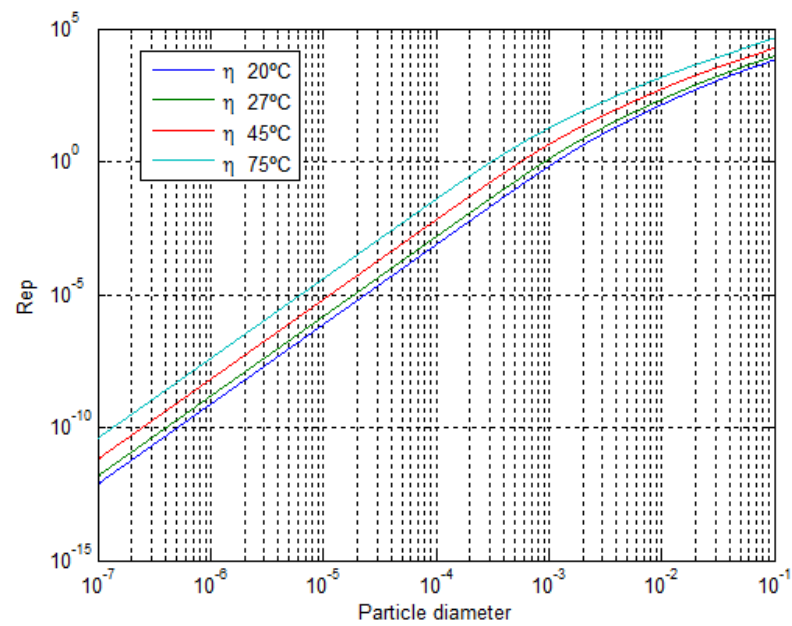


Figure 5.11. Example of Reynolds Numbers for different spherical particle diameters

5.4.1.3 Elapsed time to settle down for each particle diameter

When the drag coefficients are known for each particle diameter the elapsed time for a particle to travel a distance H is obtained from the terminal velocity, equation (5-18).

$$t = t_{\text{orig}} = H \sqrt{\frac{3C_D \rho_f}{4dg(\rho_p - \rho_f)}} \quad \text{Re}_p < 2 \cdot 10^5 \quad (5-18)$$

Equation (5-18). Elapsed time to settle down for each particle diameter

As an example, considering the terminal velocities of Figure 5.7, the time to travel a distance $H=0.1\text{m}$ is shown in Figure 5.12.

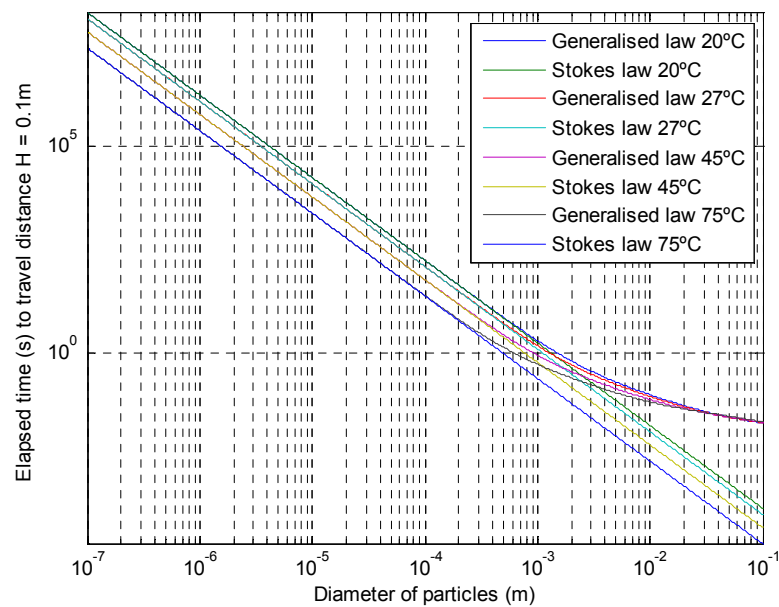


Figure 5.12. Elapsed time for a particle to travel a distance of 10cm

5.4.1.4 Construction of the M matrix for the range of simulating times and particle diameters

The linear system of equations is constructed considering the evolution of ferrous debris density over a period of time. A graphical representation of the construction of this linear system of equations is obtained representing the elapsed time for each particle until settled or travelling the distance H . Figure 5.13 corresponds to the black trace t_0 . Parallel traces to the black one (represented in dashed lines) show the remaining time for a particle before it settles down at different sampling times. Considering a small range of particle diameters, for the initial time t_0 , the amount of elapsed time until a particle settles down is P_0 , for the next instant of time, t_1 is P_1 and so on. The percentage of particles remaining in the oil for this small range of particle diameters is easily calculated at each instant of time, for example at the instant t_1 is the ratio $(P_1 - P_0) \cdot 100$, at the instant t_2 $(P_2 - P_0) \cdot 100$ and so on. As can be seen from Figure 5.13, the purple trace is below the horizontal axis, and physically this means a particle contained in this range of diameters settles down in a instant of time between t_3 and t_4 . Then, for times t_i greater than P_0 the percentage of particles in the selected interval remaining in the oil is 0. Having in mind all these simple concepts, a matrix that represents all percentages of particles remaining in the oil for each instant of time (sampling time) and particle diameters can be constructed.

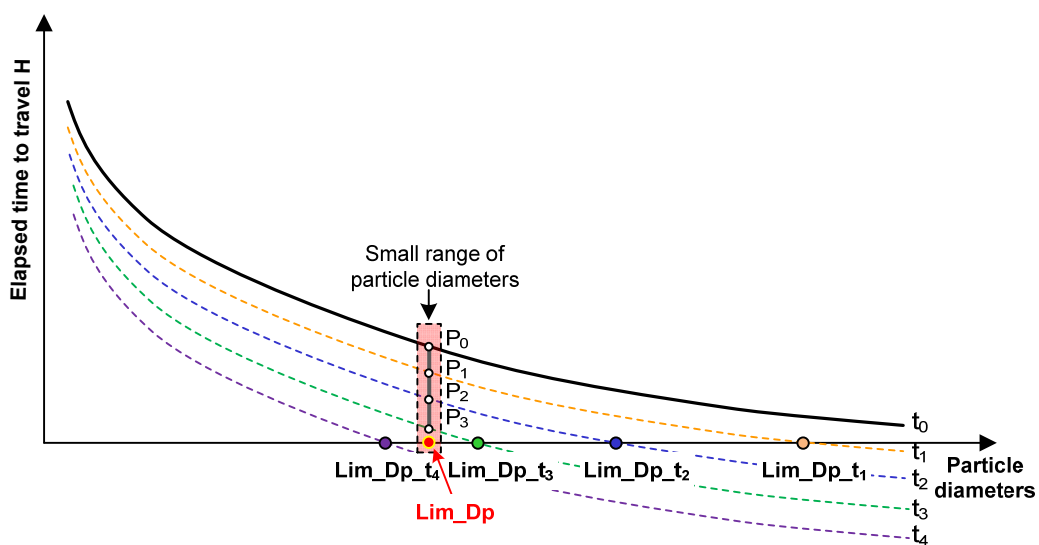


Figure 5.13. Elapsed time for a particle to settle down (approach to the problem)

The matrix M has n rows and m column dimensions and it is constructed as shown in equation (5-19).

$$\begin{aligned} t_{\text{orig}} &= (t_{\text{orig},1}, t_{\text{orig},2}, \dots, t_{\text{orig},m}) \\ t_s &= (t_{s,1}, t_{s,2}, \dots, t_{s,n}) \\ d &= (d_1, d_2, \dots, d_m) \\ d_{\text{lim}} &= (d_{\text{lim},1}, d_{\text{lim},2}, \dots, d_{\text{lim},n}) \end{aligned} \tag{5-19}$$

$$M \equiv (m_{i,j}) = \begin{cases} \frac{t_{\text{orig},j} - t_{s,i}}{t_{\text{orig},j}} & i = 1 \dots n \\ 0 & \text{if } d_j < d_{\text{lim},i} \\ & j = 1 \dots m \end{cases}$$

Equation (5-19). Overdetermined system of equations in matrix form

Where

- $t_{\text{orig},j}$ → The elapsed time for particle of diameter d_j to settle down (s).
- $t_{s,i}$ → Sampling times of curve-fitted ferrous debris density measurements (s)
- d → Array of particle diameters logarithmically equally spaced (m).
- $d_{\text{lim},i}$ → For each $t_{s,i}$ is the maximum diameter of the particles in suspension (m)

The values of m_{ij} are explained in physical terms as the percentage of particles for each diameter d_j and sampling time $t_{s,i}$ which are still suspended in the fluid. As an example, $m_{ij}=0.8$ means that 20% of the total particles of size d_j at time $t_{s,i}$ have already settled.

An example of the graphical representation of M matrix is shown in Figure 5.14.

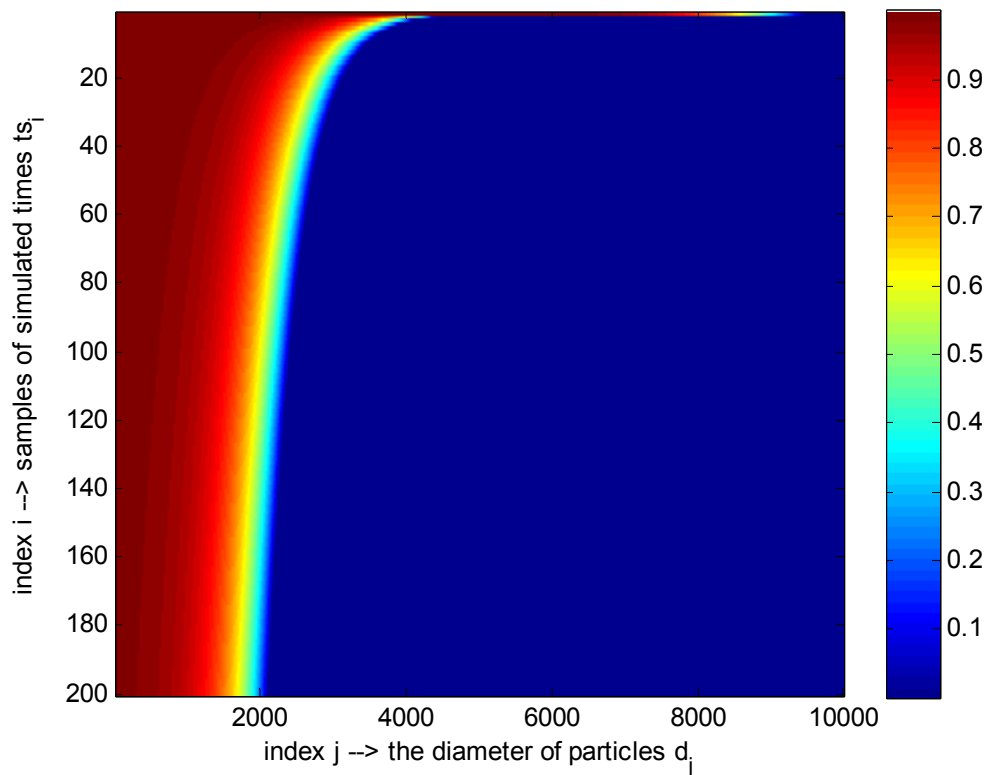


Figure 5.14. Graphical representation of the M matrix
(The size is 200 rows and 10^4 columns).

5.4.1.5 Curve fitting algorithm for ferrous debris density curves

The debris density samples (y) are normalised according to the maximum value and also curve fitted using robust regression. Equation (5-20) shows the array notation for each sample at $t_{s,i}$ of the curve fitted model.

$$y_f = (Pd_1, Pd_2, \dots, Pd_m) \quad (5-20)$$

Equation (5-20). Notation for particle density samples over time

Where:

$Pd_i \rightarrow$ Particle density sampled at $t_{s,i}$

5.4.1.6 Solution to the overdetermined system of equations

The solution of the overdetermined system is the particle size distribution. The linear system is solved in the least mean squares sense using the pseudoinverse matrix, as shown in equation (5-21).

$$\hat{x} = (\hat{x}_1, \hat{x}_2, \dots, \hat{x}_m) \quad (5-21)$$

$$\hat{x} = M^\dagger y_f$$

Equation (5-21). Solution to the problem in the least mean squares sense

Where:

$x_i \rightarrow$ Normalised particle size distribution within diameter range d_i

5.5 Validation of the methodology

5.5.1 Experimental set-up

Ferrous debris density measurements are performed using a procured ferrous debris density sensor. The sensor relies on the principals of magnetometry and is based on two fundamental physical effects (Kittiwake 2007).

- The change of inductance due to the presence of a magnetic material (changes in permeability).
- The change of energy losses due to electromagnetic induction in any conductive material (Eddy currents losses).

The changes of permeability and the losses due to eddy currents are measured by means of two symmetrical measurement chambers: measure and reference halves. The reference half remains empty, while the other chamber is filled with the fluid sample (Kent 2004). The sensor measures the difference between the two halves using an inductive bridge and synchronous demodulator technique (Hutchings 2000). As a result, its output is a signal proportional to ferrous debris density by weight. The sensitivity of the sensor is 1ppm (1 μ g/1ml) but the repeatability is 5 ppm if the

temperature between chambers is less than 10°C (Kittiwake 2007). The volume of the measurement chamber is 4ml, hence the minimum discernible mass of the sensor is 20 µg. The sensor is equipped with a piston which pumps the oil to the measure half at a rate of one pump per minute.

The experimental sample is lubricant contaminated with ferrous debris. The sample is stored in a lab bottle. The size distribution of the debris is not important because the unknown variable is the PSD. However, the particles should be ferromagnetic and preferably steel to achieve more accurate readings. Accuracy is dependent on the material used for the sensor's calibration process. The factory calibration process of the sensors consists of correlating the unbalance of permeability between the sensing and reference coils with a known amount of mass of steel in the sensing chamber. Therefore the particles stored in the bottle are made of steel and for convenience were originated in a pin on disc experiment. The experiment is simple, a bottle filled with known oil contaminated with debris is filled up to a certain height H . The oil is stirred to ensure uniform suspension of the particles, and after, ferrous debris density readings are taken until the particles settle down or for a certain amount of time. Knowing the ferrous debris density evolution over time and applying the proposed methodology the particle size distribution can be obtained. The temperature of the lubricant is controlled using a hot plate with a PID controller which indirectly controls viscosity due to temperature drifts. Figure 5.15 is a functional sketch of the rig for ferrous debris density measurements.

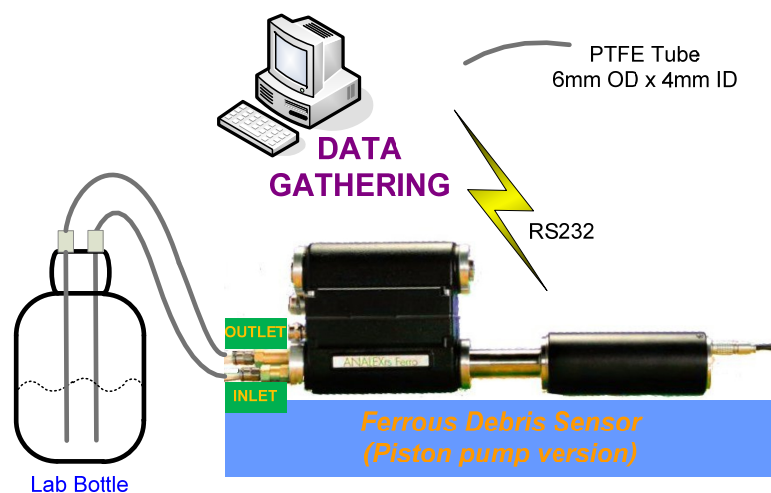
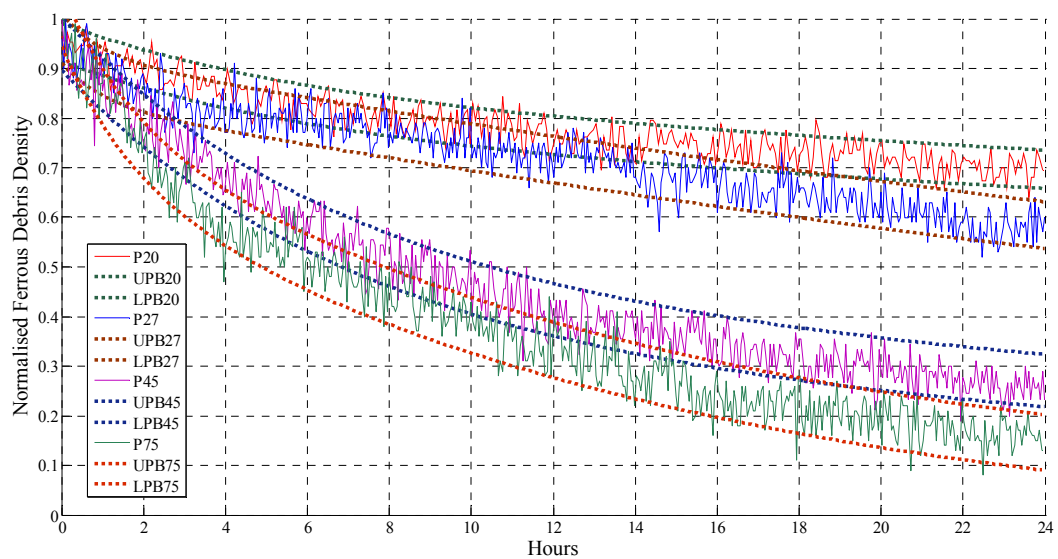


Figure 5.15. Experimental rig for ferrous debris density measurements

5.5.2 Results

The validation of the methodology was carried out monitoring the evolution of ferrous debris density for the same oil at four different lubricant temperatures (20°C, 27°C, 45°C, 75°C). Lubricant temperature affects viscosity which is shown in Table 4.1. The test procedure uses a container filled up with contaminated hydraulic oil of ISO grade 32 to the level of H=10cm. Before performing any measurement, the bottle was stirred to ensure uniform suspension of ferrous debris. Figure 4.16 shows the evolution of the normalised ferrous debris density and Table 4.2 contains the proposed curve fitted model for the experimental data. The input parameters for the example are summarised in Table 4.3. After the calculation of the drag coefficients of each particle diameter the terminal velocity and elapsed time were calculated as shown in Figures 4.7, 4.12 respectively. Results obtained after solving the linear system of equations are shown in Figures 4.17, 4.18, 4.19, 4.20.



P → ferrous debris density (ppm in weight)

UPB /LPB → upper / lower prediction bounds
for the 80% of confidence level

Figure 5.16. Deposition of ferrous debris at different temperatures

Table 5.1. Viscosity of the lubricant under test

	Temp: 20°C	Temp: 27°C	Temp: 45°C	Temp: 75°C
Kinematic Viscosity (cSt)	78 cSt	55 cSt	26.5 cSt	10.7 cSt
Absolute Viscosity (cP)	68.2 cP	48.1 cP	23.2 cP	9.4 cP

Table 5.2. Prediction model parameters

	Temp: 20°C	Temp: 27°C	Temp: 45°C	Temp: 75°C
Maximum Values (ppm)	107 ppm	102 ppm	93 ppm	100 ppm
	Proposed regression model for the normalised data $y_f = f(x) = ae^{bx} + ce^{dx}$			
Temp: 20 °C	a = 0.1519	b = -0.1696	c = 0.8005	d = -0.00593
Temp: 27 °C	a = 0.0829	b = -1.0390	c = 0.8788	d = -0.01703
Temp: 45 °C	a = 0.5368	b = -0.1509	c = 0.4142	d = -0.01996
Temp: 75 °C	a = 0.2439	b = -0.5979	c = 0.7595	d = -0.06879

Table 5.3. Input parameters for the tests results

H	→	Effective height	0.1m
ρ_p	→	Particle density	Steel 7850 Kg/m ³
ρ_f	→	Fluid density	874 Kg/m ³
η_f	→	Absolute viscosity	(68.2 · 10 ⁻³ Pa · s, 20 °C), (48.1 · 10 ⁻³ Pa · s, 27 °C) (23.2 · 10 ⁻³ Pa · s, 45 °C) (9.4 · 10 ⁻³ Pa · s, 75 °C)
d	→	Particle diameters	10 ⁻⁷ to 10 ⁻² m logarithmically equally space (10 ⁴ diameters)
t _s	→	Sampling times	0.1 to 20 · 10 ³ s linearly equally distributed (200 samples of time)
y	→	Array of ferrous debris density. Figure 5.16.	Evolution of ferrous debris density sampled at t _{sim} (ppm)

5.5.3 Discussion

The proposed methodology was applied using a lubricant which mainly contained rubbing wear particles from a pin-on-disc test. The particles were formed by the sliding of metal against metal. The underlying wear mechanism is abrasive two-body wear which is caused by hard particles or hard protuberances being forced against and moved along a solid surface (Pérez et al. 2010). After ferrogram analyses using linear ferrography the most predominant size of particles were identified as normal rubbing wear particles with a size less than $15\mu\text{m}$, the lubricant was also filtered with a filter of pore size $20\mu\text{m}$ before performing any measurement. Therefore, results of the prediction of particle distribution were in good agreement with measured particles.

Drag coefficients play an important role for the final prediction. As an example, if only Stokes region is considered, the final prediction of 4.17, 4.18, 4.19, 4.20 is slightly shifted to the right and ferrous debris density distribution for small radius particles ($<1\mu\text{m}$) is slightly higher than using higher Reynolds numbers.

Commercial particle counters give the number of particles for a range of sizes of particles. Typical ODM particle counters classify the range as $60\text{-}100\ \mu\text{m}$, $100\text{-}200\ \mu\text{m}$, $200\text{-}300\ \mu\text{m}$, $>300\ \mu\text{m}$. The cumulative distribution which appears in Figures 4.17, 4.18, 4.19, 4.20 helps to select different ranges of size of particles. The total number of particles can be determined using an effective radius for each selected range of size of particles and there is no size limitation apart from the initial size discretisation of the diameters.

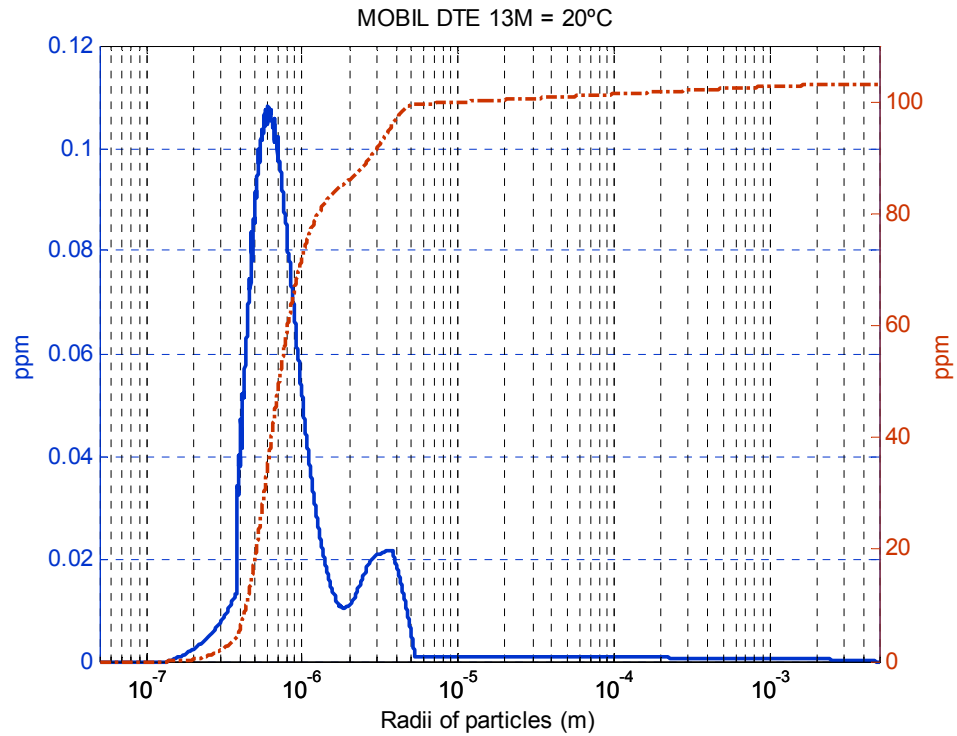


Figure 5.17. Debris density distribution (blue) and cumulative debris density distribution (red) for each radius of particles for a fluid temperature of 20°C

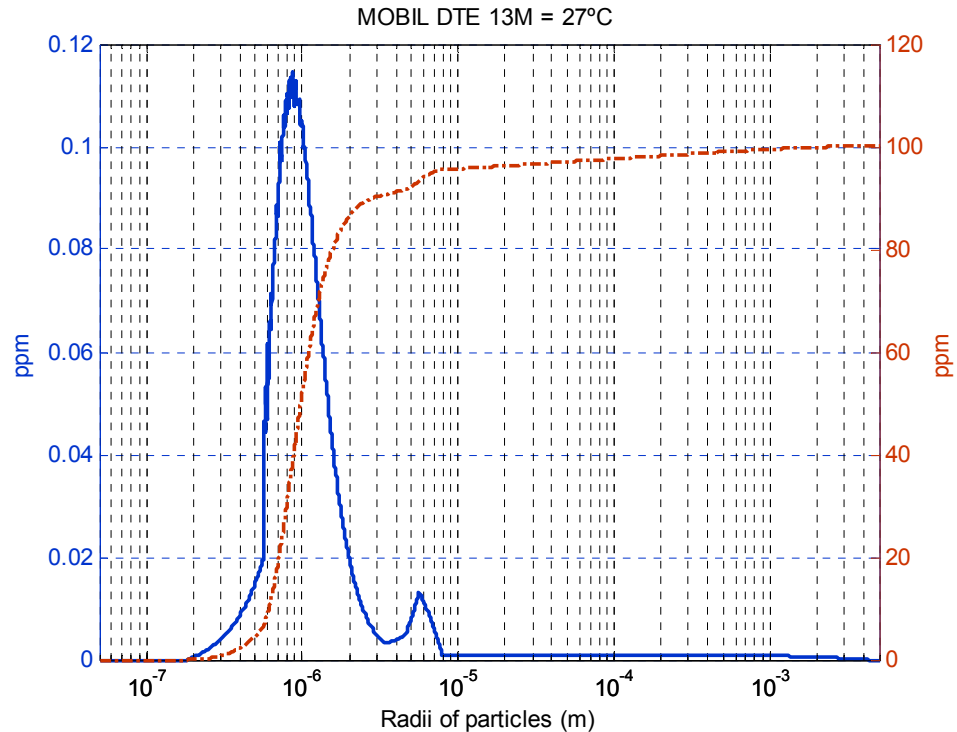


Figure 5.18. Debris density distribution (blue) and cumulative debris density distribution (red) for each radius of particles for a fluid temperature of 27°C

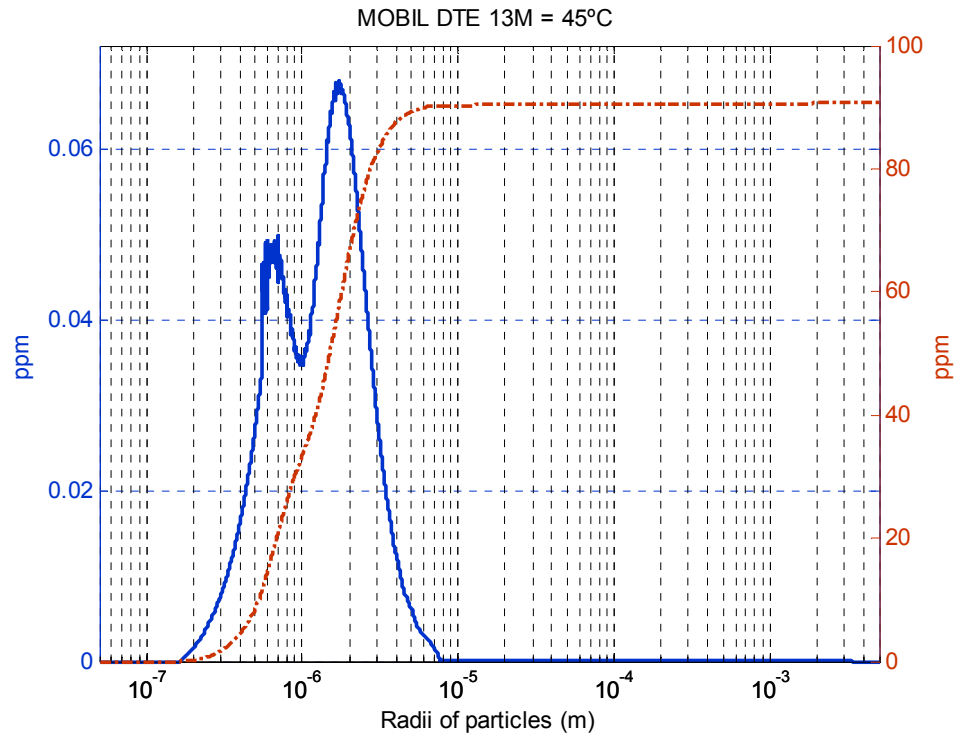


Figure 5.19. Debris density distribution (blue) and cumulative debris density distribution (red) for each radius of particles for a fluid temperature of 45°C

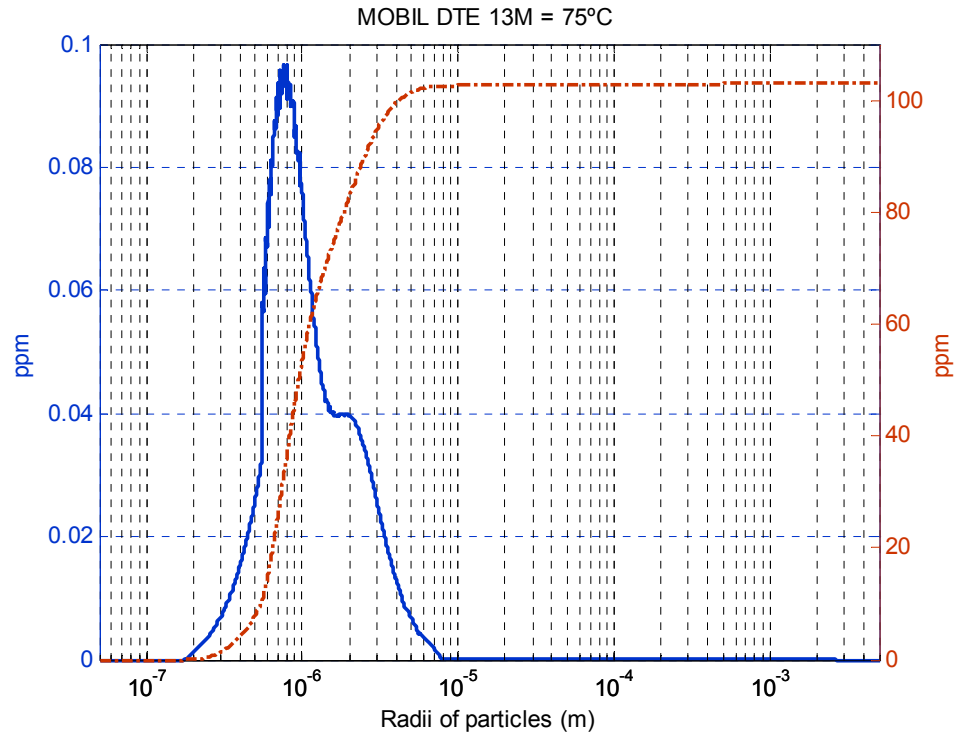


Figure 5.20. Debris density distribution (blue) and cumulative debris density distribution (red) for each radius of particles for a fluid temperature of 75°C

5.5.4 Conclusions

A novel methodology is developed to predict particle size distribution based on sedimentation theory using on/in-line ferrous debris density measurements. Although the type of sensor is an on-line ferrous debris density sensor, the proposed algorithm is not dependant on the particle density sensing method. Therefore, this technique can be extended for non-ferrous density measurements as for example density measurements based on changes in the dielectric constant of the fluid or methods based on light absorbance.

Results of particle size distribution are mainly affected by the selection of the drag coefficients. The present study only considers drag coefficients for spherical particles. This assumption could be improved if drag coefficients were selected according to the shape of the particles. This last consideration should be borne in mind as non-spherical particles used to happen in most types of wear mechanisms and industrial processes. Accuracy of results is also dependant on the repeatability and sensitivity of the particle density sensor and the stable control of viscosity which is achieved by means of precise fluid thermal regulation.

The application of this methodology has the potential to develop sensors which measure particle sizes in the micrometer and submicrometer range using low cost sensing technology as impedance spectroscopy or magnetometry techniques. Improvements of this methodology would allow further development of low cost particle counters in the micrometer and submicrometer range which can be widely applied in many industrial processes and scientific disciplines.

5.6 Future research direction in particle size distribution modelling

The process of particle deposition can be speeded up using a centrifugal force. The gravity is several times lower than the centrifugal acceleration equation (5-22), and then it can be neglected in the analysis. Figure 5.21 shows the typical values with respect to gravity 'g'.

$$G = \omega^2 \cdot r \quad (5-22)$$

Equation (5-22). Centrifugal acceleration

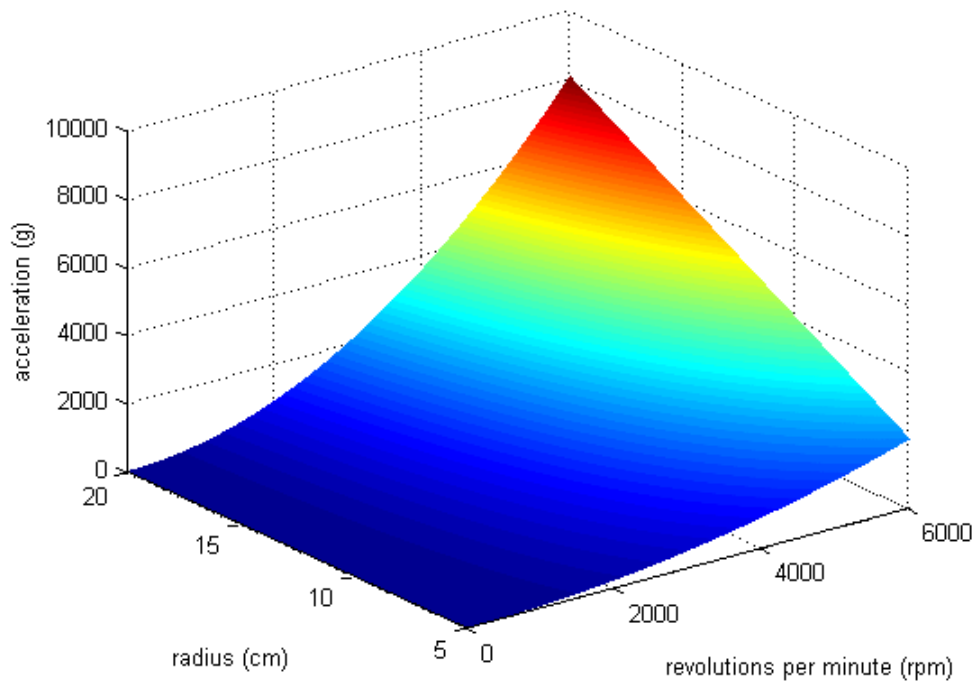


Figure 5.21. Centrifugal acceleration in ‘g’ for a radius up to 20cm and rpm up to 6000

In a rotating flow, Stokes Law is modified by the “centrifugal gravity” as shown in equation (5-23).

$$V_T = \frac{1}{18\mu} D_p^2 \cdot \omega^2 r \cdot (\rho_p - \rho_f) \quad (5-23)$$

Equation (5-23). Terminal velocity of the particle (Stokes region)

In general, the terminal velocity is modified by the “centrifugal gravity” as shown in equation (5-24).

$$V_T = \sqrt{\frac{4 \cdot \omega^2 r \cdot D_p}{3 \cdot C_D} \left[\frac{\rho_p - \rho_f}{\rho_f} \right]} \quad (5-24)$$

Equation (5-24). Velocity of the particle (Stokes region)

In a rotating flow, the determination of the drag coefficient depends on the particle diameter and the radial distance of the particle from the centre. Therefore, the terminal velocity is a function of the particle diameter and the drag coefficient which varies depending on the radial distance. The terminal velocity can be assumed to be constant in adequately spaced intervals of the radial distance and hence, the application of the methodology for the determination of particle size distribution from particle density measurements is still valid.

A sketch of a device under “centrifugal gravity” is shown in Figure 5.22. The whole assembly rotates supported by two bearings, it has an inlet port where the fluid enters and is stirred to ensure uniform suspension. Two valves are used for timing when the fluid enters the measurement chambers. Two coils are assembled at an adequate distance from the rotating axis to measure the evolution of the concentration of debris in the oil.

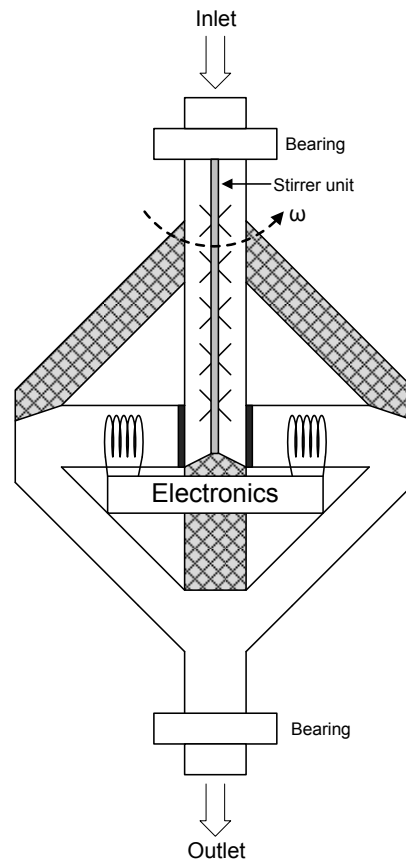


Figure 5.22. Sketch of a sensor device under “centrifugal gravity”.

6 Discussion and conclusions

6.1 Contributions to knowledge

The outcome of this research are seven important contributions to knowledge that are the following:

- A complete example under four ball sliding wear test is fully described and discussed which clarifies some of the problems that can be encountered using rational approaches based on lubrication models.
- A new model of friction for the pin-on-disc configuration based on the probabilistic approach of wear prediction is devised. This model is of relevant interest for predicting wear volume within the contact during test. The model is useful as it can predict local features within the contact area such as the edge-loading effect that happens in flat pins without chamfer.
- Based on the obtained experimental results, the effectiveness of process viscometers for detecting hydraulic oils contaminated with water is reported. In general terms viscosity measurements does not provide a clear indication of water contamination. It should be noted that water content is one of the most important contaminants in hydraulic oils.
- It is explained and experimentally proved the degree of influence of particle deposition process in wear debris monitoring in relation to the real progress of wear.
- A complete and novel review of the state of the art and future trends in oil quality sensors based on changes of the complex permittivity is presented. All possible types of electronic sensing topologies are discussed in terms of sensing performance and final implementation costs.

Based on the new trends of oil quality sensing based on IS, a new and portable implementation of an oil quality sensor based on AC current sources for low cost impedance spectroscopy measurements is presented.

- A novel methodology that can break the technological particle size limitation of particle counters based on magnetometry is presented in chapter 4. This methodology when the external force is the “centrifugal gravity” leads to a novel type of particle counters which can be of great interest for industrial and military machinery. This methodology is the most important contribution to knowledge of this research as it represents a novel type of particle counter that can compete and find a gap amongst all the other alternatives in the ODM market.

6.2 Final remarks

At the beginning of this thesis special effort was kept trying to explain oil condition sensor technology using lubrication theory and controlled experimental conditions. This work was based on the idea of characterising the output of the sensors according to the lubrication regime parameter (λ). The reason of selecting the lubrication regime parameter is because a fine control of λ implies an accurate control of the failure progress. From the theoretical viewpoint, the initial idea was a promising line of research mainly because no other research studies tackling this problem from this perspective were found. Under such an assumption, the response time of the sensors with respect to a controlled failure could be quantified and related to the accurate prediction of λ . This is basically an assessment of the sensors under controlled failure and hence, it provides a way of evaluating their limitations (failure prediction capabilities) in real applications, which is one of the interests of the RNLI. Unfortunately, the assumption that the lubrication regime modifying test conditions could be finely controlled was shown to be incorrect and not realisable.

During the course of this research, this promising idea was finally discarded because of the difficulty to control λ in accelerated wear test configurations. Needless to say, accelerated wear tests need to be operated in boundary or mixed lubrication regimes. Under such regimes, there are many factors that play a relevant role and many of them are not considered in current lubrication models. The difficulty to control and quantify them, makes it impossible to include these factors in lubrication models. As a result, it is hard to get a reliable estimation of λ which makes meaningless the assessment of the sensors because the failure process cannot

be controlled. In practical terms, λ parameter can only be used as a coarse indicator of the lubrication regime which is typically used to indicate transitions between lubrication regimes. However, researchers around the world are putting a lot of effort trying to improve lubrication models and more and more complicated multiphysics problems can be solved using efficient numerical solvers. Maybe in the future, the attempt to characterise the output of the sensors according to the lubricant regime could be realisable. The fine control of the lubrication regime (λ) mentioned, which means a fine control of the failure process, should not be mistaken with the easy task of the uncontrolled generation of failures. Faults or accelerated wear rates can be easily generated, but these tests are not representative of typical faults in machinery. For example, wear debris can be easily generated in a tribometer under pin-on-disc configuration by means of running experiments under boundary lubrication regime with a tribological pair composed of two materials with noticeable difference in hardness. From the readings of coefficient of friction and wear rate of a test performed in such an uncontrolled condition it can only be concluded if the experiment is performed under boundary lubrication regime or with a lubricating film separating the contacting surfaces. In last case, when the aim of the experiment is the generation of debris it can always be performed using a cost-effective alternative, e.g. a set of grinders. Aiming to predict the wear volume within the contact area during the experiments a novel probabilistic model of friction for a generalised pin on disc setup was devised. Due to the lack of this machine in the lab facilities, this model could not be validated using different tribological pairs. However, the model is an important contribution in wear modelling and it needs to be further assessed with other tribological pairs. When the model predictions matches experimental results, this model is of great interest for characterising the effectiveness of ODM sensors under controlled experimental conditions.

Manufacturers evaluate the sensors based on calibration samples and finally draw up their specifications according to “laboratory conditions”. Manufacturer’s specifications are ideal ones and do not consider problems associated with the real working environment. As an example, considering the ferrous wear debris density sensor, the manufacturer uses a known mass of steel in the sensing chamber and sensors are calibrated according to that. A real scenario for this sensor implies more

factors to consider such as the influence of the type of debris (magnetic properties), the shape and size of the debris, the problems of particle deposition within the system, the response time of the sensor etc. All these factors are not specified by manufacturers. Therefore, the most reliable way of determining the performance of the sensor in the operation environment is to compare its results against the current status of the machinery. This method of assessing the failure is far from an “ideal predictive maintenance” which aims to monitor the onset of failure. Basically, if the sensors work, they would provide an indicator of an existing and well-developed failure condition which is far beyond predicting the onset of failure. From this point of view, this research tries to explain typical commercial oil condition sensor technology from the operation viewpoint and it highlights possible factors that can limit the detection capability of the sensors. It was concluded that in-line viscosity sensors are not as effective in hydraulic systems monitoring as other types of sensors such as water in oil and oil impedance monitoring. This conclusion was supported because changes in oil viscosity are not so pronounced when water is dissolved in the oil and even when there is a water and oil emulsion within the range of high water content contamination levels (500-3000ppm, an oil with this level of water is considered unserviceable). The typical temperature operation in hydraulic systems is quite low (60-70°C), therefore there are not significant changes in oil viscosity due to thermal degradation of the oil. However, viscosity monitoring is more efficient for engine oils because of the higher operational temperatures. The theoretical and practical limitations of wear debris sensors based on magnetometry were also explained and reported not only discussing the minimum detection capability but also their effectiveness that depends on the uniform suspension of the debris within the oil circulation system. Therefore, it was suggested using this monitoring technique when it could be reliable, that means, in systems with small sumps and rapid circulation of the lubricating oils. Examples of these systems are reciprocating engines, power train components and aviation turbine applications. If this monitoring technique is applied to machinery with stationary lubrication oils and large sumps, the effective fault detection cannot be guaranteed. An extreme case of misrepresentative results may arise when machinery operates under start-stop conditions with a low utilisation factor. In this last case, the incipient fault detection using ferrous debris density measurements can only be achieved if the particle

deposition process is determined within machinery, which is almost an impossible task as the particle deposition process is a complex physical phenomenon mainly dependent on temperature, particle size, oil viscosity and the capacity to stir the oil (vibrations in machinery).

Finally, this research also proposes novel ideas for future sensor development.

6.3 Recommendations for further work

From a tribological point of view, the model of friction for a pin-on-disc setup needs to be validated with new experimental datasets and different types of tribological pairs. Any modification to the proposed model that could improve the matching ratio of its prediction to experimental data should be considered as an important contribution in wear modelling under pin-on disc configuration.

The most noticeable contribution of this research should be seen from an industrial maintenance perspective. Any improvement in oil quality or wear debris detection is a considerable breakthrough and it has a direct impact on machine maintenance and failure prediction. Within this scope, this work introduces two new lines of research for further sensor development.

The first line of research contemplates the possibility of developing an oil debris sensor that monitors the particle size distribution with greater accuracy than current particle counters based on magnetometry. This is a considerable breakthrough as it allows an earlier detection of the onset of failures. This research proposes the development of a particle counter based on centrifugal acceleration with a new methodology that determines the particle size distribution from ferrous debris density measurements that is not limited by the size of the debris.

The second line of research is related to oil quality sensing based on impedance spectroscopy. Impedance monitoring of oils provides a qualitative indication of the degradation of the oils. Typical commercial sensors monitor the impedance of the oils in a reduced range of frequencies. Commercial $\text{Tan}\delta$ sensors monitor the impedance of the oils in the MHz range while other patented sensor technology and research in electrochemical studies has proved the improved sensibility of the sensors at lower frequencies within the range of frequencies

between (10mHz to 100kHz). This research, considering all this information, introduces a new prototype for performing low-cost electrochemical impedance studies within the range of 100Hz-100kHz using the impedance analyser chip (AD5933) and significantly improving its impedance measurement capabilities to enable it to be used for grounded loads by means of an AC Impedance current source built by combining an Improved Howland Current Source with a Generalised Impedance Converter. The proposed impedance monitoring sensor is not available in the lubrication industry and it has other potential applications in other non-destructive testing such as corrosion and health structure monitoring.

References

- Agoston, A., Ötsch, C., and Jakoby, B., 2005. Viscosity sensors for engine oil condition monitoring-application and interpretation of results. *Sensors and Actuators A*, 121, 327-332
- Ai, X., and Cheng, H. S., 1996. The effects of surface texture on EHL point contacts. *ASME Journal of Tribology*, 118, 59-66
- Al-Shayea, A. M. IE 434: Reliability and Maintenance Engineering (Lecture I, maintenance definition). Course presentation. Available from: [Accessed June 2011]. URL: <http://faculty.ksu.edu.sa/Alshayea/IE%20434%20Lectures/LEC3-%20Reliability%20Engineering%20and%20Maintenance.ppt>
- Allen, K. J. K., 2004. Design and Fabrication of a Laboratory Test Unit to Demonstrate the Characterization and Collection of Data from Condition Monitoring Sensors. Defense R&D Canada (DRDC) Atlantic Contract Report No. CR 2005-243.
- Anderson, M., 2000. The use of Tribological Aspect Numbers in Bench Test Selection – A Review Update. In: Symposium on Bench Testing of Lubrication and Wear Properties of Industrial Fluids Used in Machinery Applications, Seattle: ASTM STP1404, ISBN 0-8031-2867-3
- Anon. 2004. Standard Test Method for Wear Preventive Characteristics of Lubricating Fluid (Four-Ball Method). ASTM (D4172-94).
- Anon. 2007. Application-Sheet: Thermoset polymer-based capacitive sensors Honeywell. Available from: <http://honeywell.com/sensing/> [Accessed: June 2011].
- Anon. 2011. Exxo Nuto Product description Available from: http://xom.ee/files/20_euxxenindesnuto_h.pdf [Accessed: June 2011].
- Anon. 2011. Product description for mobil dte 10m series. Available from: http://www.gwinc.com/downloads/mobil/glxxenindmomobil_dte_10m.pdf [Accessed: June 2011].
- Appleby, M.P., 2010. *Wear debris detection and oil analysis using ultrasonic and capacitance measurements*. Thesis (MSc). University of Akron.
- Archard, J. F., 1953. Contact and rubbing of flat surfaces. *Journal of Applied Physics*, 24 (8), 981-988.
- Archard, J. F., 1959. The temperature of rubbing surfaces. *Wear*, 2, 438-355.
- Baldwin, A., and Lunt, S., 2010. *Latest developments in online oil condition monitoring sensors*. Paper presented at The Seventh International Conference on Condition Monitoring and Machinery Failure Prevention Technologies, Stratford-Upon-Avon, UK. ISBN no.: 978-1-901892-33-8.
- Baur, P. S., 1982. Ferrography: Machinery-wear analysis with predictable future. *Power Magazine*, 126, 114-117.
- Bayer, R. G., 2004. *Mechanical wear fundamentals and testing*. New York: Marcel Dekker.
- Bhushan, B. 2001. *Modern tribology handbook*. New York : CRC Press.

- Blok, H., 1937. *Theoretical study of temperature rise at surfaces of actual contact under oiliness lubricating conditions*. Proc. Inst. Mech. Eng. General Discuss. Lubr. 2, 222 – 235
- Boyce, L. RNLI annual report and accounts. Available from: <http://www.rnli.org.uk> [Accessed: June 2011].
- Boyle, F. P., and Lvovich, V. F., 2007. *Method for on-line monitoring of condition of non-aqueous fluids*. The Lubrizol Corporation. US Patent Number US2007/0151806A1.
- Boyle, F. P., and Lvovich, V. F., 2008. *Method for on-line monitoring of condition of non-aqueous fluids*. The Lubrizol Corporation. US Patent Number US7355415.
- Briscoe, B. J., and Stolarski, T. A., 1985a. Transfer wear of polymers during combined linear motion and load axis spin. *Wear*, 104 (2), 121-137.
- Briscoe, B. J., and Stolarski, T. A., 1985b. Wear of polymers in the pin-on-disk configuration. *American Chemical Society (ACS) Symposium Series*, 303-313.
- Bruttel, P., and Schlink, R., 2006. *Water determination by Karl Fisher Titration*. Metrohm Monograph (article number 8.026.5013), 80 pages.
- Buhrdorf, A., Dobrinski, H., Lüdtkke, O., Bennett, J. W., Matsiev, L. F., Uhrich, M., and Kolosov, O., 2005. *Multiparametric oil condition sensor based on the tuning fork technology for automotive applications*. Technical presentation given in the Forum on Advanced Microsystems for Automotive Applications (AMAA 2005 Yearbook).
Available from:
http://www.amaa.de/previous/amaa_2005/presentations2005/01-01-03_Buhrdorf.pdf [Accessed: June 2011]. (v)
- Bushan, B. 1999. *Principles and applications of tribology*. New York : John Wiley.
- Carslaw, H. S., and Jaeger, J. C., 1959. *Conduction of heat in solids*. Second ed.: Oxford University Press.
- Collister, C. J., 1998. *Electrical measurement of oil quality*. WO Patent Number WO/1998/050790.
- Craig, M., Harvey, T. J., Wood, R. J. K., Masuda, K., Kawabata, M., and Powrie, H. E. G., 2009b. Advanced condition monitoring of tapered roller bearings, Part 1. *Tribology International*, 42 (11-12), 1846-1856.
- Chittenden, R. J., Dowson, D., Dunn, J. F., and Taylor, C. M., 1985. I. A Theoretical Analysis of the Isothermal Elastohydrodynamic Lubrication of Concentrated Contacts. II. General Case, with Lubricant Entrainment along Either Principal Axis of the Hertzian Contact Ellipse or at Some Intermediate Angle. *Proceedings of the Royal Society of London, Series A*, 397, 245-270.
- Christensen, H., 1969. Stochastic models for hydrodynamic lubrication of rough surfaces. In: *Proc. Inst.Mech.Engrs. (London)*, 185, (Pt.1 No.55), 1013-1026
- Deolalikar, N., Sadeghi, F., and Marble, S., 2008. Numerical modelling of mixed lubrication and flash temperature in EHL elliptical contacts. *ASME Journal of Tribology*, 130, (1), Art. No. 011004
- Dinser, S. M., 2009. *Resonator based superposition rheometry*. Thesis (PhD). ETH Zürich.

- Dong, J., 2000. *Quantitative condition monitoring of lubricating oils by Fourier Transform Infrared (FTIR) Spectroscopy*. Thesis (PhD). McGill, Montreal.
- Dowson, D., 1998. *History of tribology*. Second ed.: John Wiley & Sons
- Dowson, D., and Higginson, G. R. 1966. *Elastohydrodynamic lubrication, the fundamentals of roller and gear lubrication*. Oxford: Pergamon Press. ISBN: 978-0080114729
- Dual, J., 1989. *Experimental methods in wave propagation in solids and dynamic viscometry*. Thesis (PhD) Diss. ETH No. 8659, Zürich.
- Dual, J., Sayir, M., and Goodbread, J., 1990. *Viscometer*. US Patent US4920787.
- Ertel, A. M., 1939. Hydrodynamic theory of lubrication based on new principles. *In: Acad. Nauk SSSR, Prikadnaya Matematika i Mekhanika*. 3, (2), 41-52.
- Evans, S., and Hunt, T., 2000. *The oil analysis handbook*. Oxford: Coxmoor Publishing company.
- Faraon, I. C., 2005. *Mixed lubricated line contacts*. Thesis (PhD). University of Twente.
- Farone, W. A., Sacher, R. F., and Fleck, C., 2002. *Acoustic viscometer and method of determining kinematic viscosity and intrinsic viscosity by propagation of shear waves*. US Patent US6439034.
- Fitch, J. C., 2001. *Advancements in fluid analysis technologies and strategies for hydraulic systems condition-based maintenance*. Conference paper presented at the Symposium on Hydraulic failure analysis: fluids, components and system effects, Nevada. Available in "Hydraulic Failure Analysis: Fluids, Components and System Effects", ASTM Stock No: STP1339. ISBN: 0-8031-2883-5
- Flanagan, I. M., Jordan, J. R., and Whittington, H. W., 1989. An inductive method for estimating the composition and size of metal particles. *Measurement Science and Technology*, 1 (90), 381-384.
- Frounchi, J., Dekhoda, F., and Zarifi, M. H., 2009. A low-distortion wideband integrated current source for tomography applications. *European Journal Of Scientific Research*, 27 (1), 56-65.
- Ganser, G. H., 1993. A rational approach to drag prediction of spherical and non-spherical particles. *Power Technology and Engineering*, 77, 143-152.
- Gao, J., Lee, S. C., Ai, X., and Nixon, H., 2000. An fft-based transient flash temperature model for general three-dimensional rough surface contacts. *ASME Journal of Tribology*, 123, 75-82.
- Garland, N., 2004. *Sustainable design of hydrocarbon refrigerants applied to the hermetic compressor*. Thesis. Bournemouth Universit.
- Godfrey, D., 1968. *Boundary lubrication*. Paper presented at the Interdisciplinary Approach to Friction and Wear, P.M. Ku, NASA Special Publication SP-181, 335-353.
- Goodbread, J., Sayir, M., Hausler, K., and Dual, J., 1998. *Method and device for measuring the characteristics of an oscillating system*. US Patent Number US5837885.
- Greenwood, J. A., and Williamson, J. B. P., 1966. Contact of nominally flat surfaces. *Proceedings of the Royal Society of London A.*, 295, 300-319.

- Gubrin, A. N., 1949. Fundamentals of the hydrodynamic theory of lubrication of heavily loaded cylindrical surfaces. In: *Symp: Investigation of the Contact of Machine Components, Central Scientific Research Institute for Technology and Mechanical Engineering*, 30, Moscow, Kh. ft Ketova (ed.), D.S.L.R. London Translations No. 337, pp 115-166.
- Haider, A., and Levenspiel, O., 1989. Drag coefficient and terminal velocity of spherical and non-spherical particles. *Power Technology and Engineering*, 58, 63-70.
- Hamrock, B. J., Schmid, S. R., and Jacobson, B. O., 2004. *Fundamentals of fluid film lubrication*. Second ed.: Marcel Dekker.
- Hamrock, B. J. , Jacobson, B. O. 1984. Elastohydrodynamic lubrication of line contacts. *ASLE Transactions*, 27, (4) 275–287.
- Harvey, T. J., Morris, S., Wang, L., Wood, R.J.K. And Powrie, H.E.G. 2007. Real-time monitoring of wear debris using electrostatic sensing techniques . *Proceedings of the Institution of Mechanical Engineers, Part J: Journal of Engineering Tribology*, 221 (1), 27-40.
- Harvey, T. J., Wood, R.J.K., Denuault, G. And Powrie, H.E.G., 2002. Investigation of electrostatic charging mechanisms in oil lubricated tribo-contacts. *Tribology International*, 35 (9), 605-614.
- Hoja, J., and Lentka, G., 2010. Interface circuit for impedance sensors using two specialized single-chip microsystems. *Sensors and Actuators A*, 163, 191-197.
- Hopkins, E. L., and K. Irwin, L., 1965. *Instrument for capacitively testing the condition of the lubricating oil*. US Patent Number US3182255.
- Hopkins, E. L., and Wedel, J. L., 1977. *Fluid condition monitoring system*. US Patent Number US4064455.
- Hu, Y. Z., Barber, G. C. , and Zhu, D., 1999. Numerical analysis for the elastic contact of real rough surfaces. *Tribology Transactions*, 42 (3), 443-452.
- Hu, Y. Z., and Zhu, D., 2000. A full numerical solution to the mixed lubrication in point contacts. *Journal of Tribology*, 122, 1-9.
- Hutchings, I. M., 1992. *Tribology: Friction and wear of engineering materials*. CRC Press
- Hutchings, M. J., 2000. *System and method for measuring the magnetic conductive properties of a test fluid*. US Patent Number US6051970.
- Jaeger, J. C., 1942. *Moving sources of heat and the temperature of sliding contacts*. Paper presented at the Proc. Roy. Soc, New South Wales, 76, (3), 203-224.
- Jiroutek, S. 2007. Understanding Water Activity (aw). Vaisala's Application Note: Available from <http://www.vaisala.com/> [Accessed: June 2011].
- Johnson, K. L., 1985. *Contact mechanics*. Cambridge University Press.
- Keith, P., and Hodges, B., 1996. *Hydraulic fluids*. John Wiley & Sons.
- Kempster, R. W., and George, D. B., 1993. *Method and apparatus for detecting particles in a fluid having coils isolated from external vibrations*. US Patent Number US5444367.

- Kent, T., 2009. When, What and How to monitor wind turbine gearboxes, using experience from other industries. Presentation given in the IMeche, London.
Available from: <http://events.imeche.org> [Accessed: June 2011].
- Kent, T., Leigh-Jones, C., 2004. On-line lube oil sensing for tomorrows engines. Presentation given in: *Lubrication Excellence*, Nashville, Tennessee, USA.
Available from: www.kittiwake.com/webshare/Papers/lubexcellence.pdf [Accessed: June 2011].
- Kittiwake., 2007. *Online sensor suite user guide*.
Available from: http://www.kittiwake.com/webshare/instruction_pdfs/MA-K16637-KW.pdf [Accessed: June 2011].
- Koehler, C., Seitz, M., Hirthe, R., and Wooton, D. L., 2008. *Impedance spectroscopy (IS) methods and systems for characterizing fuel*. US2008/0167823A1.
- Kolosov, O., Matsiev, L., Varni, J. F. , Dales, C., Ludtke, O., Wullner, D., Buhrdorf, A., and Dobrinski, H., 2010. *Resonator sensor assembly*. US Patent Number US7721590B2.
- Kuhlmann-Wilsdorf, D., 1986. Sample calculations of flash temperatures at a silver-graphite electric contact sliding over copper. *Wear*, 107 (1), 71-90.
- Kuhlmann-Wilsdorf, D., 1987. Temperature at interfacial contact spots: Dependence on velocity and on role reversal of two materials in sliding contact. *ASME Journal of Tribology*, 109, (2), 321-329.
- Kuroyanagi, S., Fujii, T., Okada, K., Nozawa, M., Yamagucho, S., and Naito, K., 1996. *Oil deterioration detector*. US Patent Number US5523692.
- Larsen, R. G. , Perry, G.L. , 1950. *Chemical aspects of wear and friction*. Paper presented at Mechanical Wear, Cleveland, American Society for Metals, 73-94.
- Leigh-Jones, C., Kent, T., M.Hutchings, Matsuda, M. 2004. Development of a range of on-line lubricant and machinery condition sensors for engine monitoring. *In: CIMAC Congress*, Tokio.
Available from: http://www.kittiwake.com/webshare/instruction_pdfs/MA-K16637-KW.pdf Accessed: June 2011]
- Leonid Matsiev, J. B., Eric Mcfarland. 2004. *Method and apparatus for characterizing materials by using mechanical resonator*. US Patent Number US2004/0074302A1.
- Lin, Y., Wang, S., and Resendiz, H., 2010. *Liquid properties sensor circuit*. US Patent Number US7659731B2.
- Lisowski, Z., Stolarski, T.A., 1981. A modified theory of adhesive wear in lubricated contacts. *Wear*, 68 (3), 333-345.
- Liu, Y., Liu, Z., Xie, Y., and Yao, Z., 2000. Research on an on-line wear condition monitoring system for marine diesel engine. *Tribology International*, 33 (12), 829-835.
- Lockwood, F. E., and Dalley, R. 1992. Lubricant analysis. In Henry, S. D. et al. (Ed.), *ASM Handbook: Friction, Lubrication, and Wear Technology*. Metals Park: The Materials Information Society Vol. 18, 299-312.
- Lvovich, V. F., 2007. Monitoring industrial-machinery lubricants. *Sensors Magazine*.
Available from: [<http://www.sensorsmag.com>]

- Lvovich, V. F. 2008. Impedance characterisation of time dependent degradation of industrial lubricants: The Lubrizol Corporation. The Electrochemical Society. Available from: [<http://www.electrochem.org/dl/ma/203/pdfs/2900.pdf>]
- Lvovich, V. F., Skursha, D. B. , and Boyle, F. P., 2004. *Method for on-line monitoring of quality and condition of non-aqueous fluids*. US Patent Number US2004/0075448A1.
- Lvovich, V. F., and Smiechowski, M. F., 2006. Impedance characterization of industrial lubricants. *Electrochimica Acta*, 51 (8-9), 1487-1496.
- Lvovich, V. F., and Smiechowski, M. F., 2008. Non-linear impedance analysis of industrial lubricants. *Electrochimica Acta*, 53 (25), 7375-7385.
- Lvovich, V. F., and Smiechowski, M. F., 2009. AC impedance investigation of conductivity of automotive lubricants using two and four electrode electrochemical cells. [Journal of Applied Electrochemistry](#) 39 (12), 2439-2452
- Macdonald, J. R., 1992. Impedance spectroscopy. *Annals of Biomedical Engineering*, 20, 289-305.
- Markova, L. V., Makarenko, V. M., Semenyuk, M. S., and Zozulya, A. P., 2010. On-line monitoring of viscosity of lubrication oils. *Journal of Friction and Wear*, 31 (6), 433-442.
- Miller, J. L., and Kitaljevich, D., 2000. *In-line oil debris monitor for aircraft engine condition assessment*. Paper presented at the IEEE Aerospace Conference Proceedings. 6, 49-56 .
- Murali, S., Xia, X., Jagtiani, A. V., Carletta, J., and Zhe, J., 2009. Capacitive coulter counting: Detection of metal wear particles in lubricant using a microfluidic device. *Smart Materials and Structures*, 18, 03700
- Murali, S. V., 2008. *A microfluidic coulter counting device for metal wear detection in lubrication oil*. Thesis (MSc). The University of Akron.
- Murphy, M., and Kent, T. S., 2004. On-line sensors for condition monitoring. Presentation given in: *61st STLE Annual Meeting*, Calgary, Canada. Available from: www.conditionmonitoringintl.com/resources/sensors_case_studies.pdf [Accessed: June 2011].
- Nadkarni, R. A. K., 2000. *Guide to ASTM Test Methods for the Analysis of Petroleum Products and Lubricants: Second edition*. ASTM International.
- Nemarich, C. P., Whitesel, H. K., and Sarkady, A. 1988. On-line wear particle monitoring based on ultrasonic detection and discrimination: David Taylor Research Center. Document Reference PAS-89-7.
- Nijenbanning, G., Venner, C. H., and Moes, H., 1994. Film thickness in elastohydrodynamically lubricated elliptic contacts. *Wear*, 176 (2), 217-229.
- Okada, K., and Sekino, T., 2003. *Impedance measurement handbook*. (Agilent Technologies), Available from: cp.literature.agilent.com/litweb/pdf/5950-3000.pdf [Accessed: June 2011]
- Perez, A. T. 2010. Characterisation of the rheological properties of lubricants for EHL film thickness prediction: TribouK VII, Imperial College. Available from: [<http://www3.imperial.ac.uk/tribology/events/tribouk>]

- Pérez, A. T., Hadfield, M., and Austen, S., 2010. On-line ferrous debris density monitoring in sliding area contacts under boundary lubrication regime. Paper presented at The Seventh International Conference on Condition Monitoring and Machinery Failure Prevention Technologies, Stratford-Upon-Avon, UK. ISBN no.: 978-1-901892-33-8.
- Petrusevich, A. I., 1951. *Fundamental conclusions from the hydrodynamic theory of lubrication*. Paper presented at the Izv. Akad. Nauk. SSSR (OTN), 1951, 2, 209–223.
- Philips, A. D., Eggers, W. J., Rodgers, R. S., and Pappas, D., 2006. *On-line oil condition sensor system for rotating and reciprocating machinery*. US Patent Number US7043402B2.
- Poley, J., 2001. Wear particle analysis. In: STP 1339, A. ed. *Symposium on Hydraulic failure analysis: fluids, components and system effects*, Nevada.
- Powrie, H. E. G., Wood, R. J. K., Harvey, T. J. And Morris, S., 2001. Re-analysis of electrostatic wear-site sensor data from FZG gear scuffing tests. International Conference on Condition Monitoring, University of Wales, Swansea, 177, 6-12.
- Raadnui, S., and Kleesuwa, S., 2005. Low-cost condition monitoring sensor for used oil analysis. *Wear*, 259 (7-12), 1502–1506.
- Reid, R. C., Prausnitz, J.M., Poling, B.E. 1987. *The properties of gases & liquids*. Fourth Editio, New York: McGraw-Hill.
- Rhodes, M., 2008. *Introduction to particle technology*. John Wiley and Sons. Second Edition, Chichester: John Wiley & Sons, Ltd,
- Robinson, N., 2000. Monitoring oil degradation with infrared spectroscopy. *WearCheck Tecncial Bulletin* (18). Available from: <http://www.wearcheck.com>. [Accessed: June 2011].
- Rodríguez, J. L. V., 2007. *Rheology, wettability and tribology of mixtures, nanoparticles - polyalphaolefin oils*. Thesis (PhD) . University of Oviedo.
- Roelands, C. J. A., 1966. *Correlation aspects of the viscosity-temperature-pressure relationship of lubrication oils*. Thesis (PhD). The University of Delft, The Netherlands.
- Ross, A. S., and Saulnier, G. J., 2006. *High output impedance current source*. US Patent Number US71161157/B2.
- Roth, W., and Rich, S. R., 1953. A new method for continous viscosity measurement. General theory of the ultra-viscoson. *Journal of Applied Physics*, 24 (7), 940-950.
- Roth, W., and Rich, S. R., 1958a. *Method and apparatus for measuring viscosity, etc., of fluid like materials*. US Patent Number US2839915.
- Roth, W., and Rich, S., 1958b. *Temperature compensated viscometer*. US Patent Number US2837913.
- Roylance, B. J., 2005. Ferrography - Then and now. *Tribology International*, 38 (10), 857-862.
- Roylance, B. J., and Hunt, T. M.. 1999. *Wear debris analysis* (First ed.). Oxford: Coxmoor Publishing company.
- Roylance, B. J., Sperring, T. P., and Barraclough, T. G., 2004. *Bench test determination of wear modes to classify morphological attributes of wear*

- debris*. In: Symposium on Bench Testing of Lubrication and Wear Properties of Industrial Fluids Used in Machinery Applications, Seattle: ASTM Special Technical Publication (STP1404)
- Sarangi, M. A., 2007. *Oil debris detection using capacitance and ultrasonic measurements*. Thesis (MSc). The University of Akron.
- Schipper, D. J., 1988. *Transitions in the lubrication of concentrated contacts*. Thesis (PhD). University of Twente.
- Schoes, J. N., 2004. *Oil quality sensor system, method and apparatus*. US Patent Number US6718819B2.
- Schwarz, D. M. 2006. Fast and robust curve intersections: File ID: 11837. Available from: <http://www.mathworks.com/matlabcentral/fileexchange/> [Accessed: June 2011].
- Skurka, J. C., 1970. Elastohydrodynamic lubrication of roller bearings. *Journal of Lubrication Technology*, 92, 281.
- Smiechowski, M. F., and Lvovich, V. F., 2005. Characterization of non-aqueous dispersions of carbon black nanoparticles by electrochemical impedance spectroscopy. *Journal of Electroanalytical Chemistry*, 577 (1), 67-78.
- Stachowiak, G. W., 2005. *Wear-materials, mechanisms and practice*. John Wiley & Sons.
- Stachowiak, G. W., and Batchelor, A. W., 2001. *Engineering tribology*. Second ed.: Butterworth-Heinemann.
- Stolarski, T. A., 1979. Adhesive wear of lubricated contacts. *Tribology International*, 12 (4), 169-179.
- Stolarski, T. A., 1980. A contribution to the theory of lubricated wear. *Wear*, 59 (2), 309-322.
- Stolarski, T. A., 1981a. A model for lubricated wear in a pin-on-disc configuration. *Wear*, 68 (2), 141-150.
- Stolarski, T. A., 1981b. A modified theory of adhesive wear in lubricated contact. *Wear*, 68 (3), 333-345.
- Stolarski, T. A., 1987. Evaluation of the heat of adsorption of a boundary lubricant. *ASLE Transactions*, 30 (2), 472-478.
- Stolarski, T. A., 1989a. Friction in a pin-on-disc configuration. *Mechanism and Machine Theory* 24 (5), 373-381.
- Stolarski, T. A., 1989b. Probability of scuffing in lubricated contacts. *Proceedings of the Institution of Mechanical Engineers* 203 (6), 361-369.
- Stolarski, T. A., 1990a. Fractional film defect as measure of lubricant effectiveness. *Tribology Transactions* 33 (1), 21-24.
- Stolarski, T. A., 1990b. A probabilistic approach to wear prediction. *Journal of Physics D: Applied Physics*, 23 (9), 1143-1149.
- Stolarski, T. A., 1990c. *Tribology in machine design*. Butterworth Heinemann.
- Stolarski, T. A., 1996. A system for wear prediction in lubricated sliding contacts. *Lubrication Science*. [8 \(4\)](#), 315-351.
- Stribeck, R., 1902. Diw wesentlichen eigenschaften der gleit und rollwlnlager. *VDI-Zeitschrift, Z. Verein. Deut. Ing.* 46, (38), 1341-1348.
- Theo Mang, W. D., 2007. *Lubricants and lubrication*. Second ed.: Wiley-VCH.

- Tian, X., and Kennedy, F. E., 1994. Maximum and average flash temperatures in sliding contacts. *ASME Journal of Tribology*, 116, 167-173.
- Timoshenko, S. P., and Goodier, J. N., 1982. *Theory of elasticity*. London: McGraw Hill.
- Tonder, K., 1977. Mathematical verification of the applicability of modified Reynolds equation to striated rough surfaces. *Wear*, 44, 329-343.
- Torres, A., 2010. Characterisation of the rheological properties of lubricants for EHL film thickness prediction. *Presentatio in: TriboUK VII*, Imperial College. London.
- Totten, G. E., 1999. *Handbook of hydraulic fluid technology*. Mechanical Engineering Series: CRC Press.
- Totten, G. E., 2006. *Handbook of Lubrication and Tribology*. Second ed. Vol. I: CRC Taylor & Francis.
- Totten, G. E., Kling, G. H., and Smolenski, D. J., 1997. *Tribology of hydraulic pump testing*. ASTM Special Technical Publication (STP1310).
- Totten, G. E., Wills, D. K., and Feldmann, D. G., 2001. *Hydraulic failure analysis*. ASTM Special Technical Publication (STP1339).
- Troyer, D., 2001. Removing Water Contamination from oil. *Machinery Lubrication*. On-line article available from: <http://www.machinerylubrication.com/Read/177/removing-water-oil-contamination> [Accessed: June 2011].
- Ulrich, C., Petersson, H., Sundgren, H., Björefors, F., and Krantz-Rülcker, C., 2007. Simultaneous estimation of soot and diesel contamination in engine oil using electrochemical impedance spectroscopy. *Sensors and Actuators B: Chemical*, 127 (2), 613-618.
- Valtorta, D., 2007. *Dynamic torsion test for the mechanical characterization of soft biological tissues*. Thesis (PhD). Politecnico di Milano.
- Venner, C. H., 1991. *Multilevel solutions of the EHL line and point contact problems*. Thesis (PhD). University of Twente.
- Venner, C. H., and Napel, W. E., 1992. Surface roughness effects in an EHL line contact. *ASME Journal of Tribology*, 114, 329-343.
- Voelker, P. J., and Hedges, J. D., 1998. *Oil quality sensor for use in a motor*. US Patent Number US5789665.
- Wang, L., Wood, R. J. K., Harvey, T. J., and Powrie, H. E. G., 2007. Electrostatic charge – a novel condition monitoring technology. *Condition Monitoring*, 248, 5-9.
- Wang, S.-C. S., Lee, H.-S., Mcgrath, P. B., and Staley, D. R., 1992. *Oil sensor systems and methods of qualitatively determining oil type and condition*. US Patent Number US5274335.
- Whittington, H. W., and Flynn, B. W., 1998. *Debris monitoring*. US Patent Number US5811664.
- Whittington, H. W., Flynn, B. W., and Mills, G., 1992. An on-line wear debris monitor. *Measurement Science and Technology*, 3, 656-661.
- Whittington, H. W., Jordan, J.R., and Flanagan, I., 1989. *Method and apparatus for identifying the metal included in a metallic item*. US Patent Number US4837511.

-
- Wooton, D. L., Hirthe, R. W., and Seiz, M. A., 2006. *Fluid quality control using broad spectrum impedance spectroscopy*. US Patent Number US7199595B2.
- Xia, X., 2009. *Modeling a microfluidic capacitive sensor for metal wear debris detection in lubrication oil*. Thesis (MSc). University of Akron.
- Zhao, J., and Sadeghi, F., 2001. Analysis of EHL circular contact start up: Part II - Surface temperature rise model and results. *ASME Journal of Tribology*, 123, (1), 75-82.
- Zhe, J., Choy, F. K., Murali, S. V., Sarangi, M. A., and Wilfong, R., 2007. Oil debris detection using capacitance and ultrasonic measurements. In: *ASME/STLE proceedings of the STLE/ASME, International Joint Tribology Conference (IJTC2007)*, San Diego, California, 113-115
- Zhe, J., Du, L., Carletta, J. E., and Veillete, R. J., 2010. *Metal wear detection apparatus and method employing microfluidic electronic device*. US Patent Number US2010/0109686A1.
- Zhu, D., and Cheng, H. S., 1988. Effect of surface roughness on the point contact EHL. *ASME Journal of Tribology*, 110, 32-37.
- Zhu, D., and Hu, Y.Z., 2001. A computer program package for the prediction of EHL and mixed lubrication characteristics, friction, subsurface stresses and flash temperature based on measured 3-D surface roughness. *Tribology Transactions*, 44, 383-390.

Appendices

A Wear

Wear is the major cause of material wastage and loss of mechanical performance. Any reduction in wear can result in considerable savings. Friction is a principal cause of wear and energy dissipation. Considerable savings can be made by improved friction control. It is estimated that one third of the world's energy resources in present use is needed to overcome friction in one form or another (Dowson 1998; Garland 2004; Hutchings 1992; Stachowiak 2005; Totten 1999). Lubrication is an effective means of controlling wear and reducing friction. Tribology is a field of science which applies an operational analysis to problems of great economic significance such as reliability, maintenance and wear of technical equipment ranging from household appliances to spacecraft (Totten 1999).

A.1 Wear mechanisms

To conduct successful failure analysis, it is necessary to understand and recognise common wear mechanisms upon inspection. The wear processes commonly encountered are abrasive, adhesive, corrosive and contact stress fatigue wear (Stachowiak and Batchelor 2001; Totten 1999).

A.1.1 Abrasive Wear

Abrasive wear refers to the cutting of a metal by hard particles or a rough surface by ploughing or microcutting mechanisms. When abrasion wear is caused by a hard particle between two surfaces, it is termed three-body wear. Hard particles causing three-body wear can be introduced into the system from the component manufacturing process, generated internally as wear debris, ingested through a breather or seals, or it may be added as a contaminant in the fluid upon addition to the system.

Abrasive wear is dependent on particle size distribution, shape, toughness, and hardness.

Two body wear, is caused by a harder surface with asperity dimensions sufficiently large to penetrate the lubricant oil film causing ploughing or microcutting action on the other softer surface, which is in relative motion. This form of wear is also known as cutting, ploughing, gouging, lapping, grinding or broaching wear.

Abrasive wear may be minimised by the following:

1. Removing potential residual manufacturing debris by proper draining and flushing procedures before starting.
2. Using wear-resistant materials of construction
3. Minimising ingested particles by the use of proper breather filters and by keeping the system tight.
4. Using fine filtration to minimise particulate contamination of the hydraulic fluid.
5. Performing proper fluid maintenance and periodic fluid analysis.

A.1.2 Adhesive Wear

Adhesive wear occurs when surface asperities come into sliding contact under a load. If sufficient heat is generated, microwelding of the asperity with subsequent shearing and material transfer of the contact will be observed. This process will continue until larger surfaces are in contact and macrowelding or seizure occurs. Generation of adhesive wear debris may then cause abrasion wear.

Adhesive wear may be reduced by increasing the material hardness, surface hardening, avoiding metallurgically similar material pairs, avoiding highly soluble materials, using weld-resistant materials, and providing a soft layer containing compounds of sulphur and phosphorus.

Adhesive wear may occur in areas where the friction contacts are greatest due to tight fits, misalignment, high loading, poor lubrication and high temperature.

A.1.3 Cavitation Wear

Cavitation arises when there is a sudden collapse (implosion) of a gas bubble within the hydraulic fluid. The bubble implosion process forms a high velocity microjet capable of high-impact energy. After repeated bubble implosions, material fatigue damage results in surface damage with the formation of pitting and larger cavities.

A.1.4 Corrosive Wear

Corrosive wear is surface damage related to electrochemical attack on the metallic component. Corrosion may be avoided by using an appropriate design methodology. Some guidelines for proper design include the following:

- Avoid galvanic couples by insulating dissimilar metals or increase electrical resistance.
- Avoid small anodic areas (the more active metal is the anode) relative to the cathode (the less active metal), such as fasteners, by using alloys that are cathodic with respect to the main structure.
- Protect from corrosive environments.
- Avoid sections with high tensile stress.
- Use protective coatings where possible.

A.1.5 Contact Stress Fatigue Wear

Contact-surface fatigue is favoured by small contact areas, high loadings, and repeated flexing action under cyclic rolling or reciprocal sliding loads, although each occurs under different conditions. The applied stresses are less than the material yield stress, and the process is often accompanied by frictional heat and plastic flow. Subsurface structural changes are also observed metallographically. Contact-stress fatigue wear processes are distinguishable as abrasion, adhesion and corrosion.

Mechanism	Surface-contact fatigue	Other wear mechanisms
Nature of contact	Counter formal	Conformal
Stress system	Hertzian and alternating	Dispersed and continuous
Lubrication and oil film	Partial to full EHD	Partial hydrodynamic to starvation
Relative motion to mating surface	Usually rolling	Usually sliding
Wear process	Crack initiation and propagation or material subsidence	Material adhesion, ploughing, corrosion, etc
Wear particle characteristics	Lamina, sphere, and spall particles	Normal/severe sliding wear particles, cutting, wear particles, etc
<p>Table A.1. Differentiation of surface contact fatigue from other wear processes (Abrasion, Adhesion and Corrosion) (Totten 1999)</p>		

A.2 Contact pressure, Hertz theory

From elementary mechanics it is known that two contacting surfaces under load will deform. The deformation may be either plastic or elastic depending on the magnitude of the applied load and the material's hardness. In many engineering applications, for example, rolling contact bearings, gears, cams, seals, etc., the contacting surfaces are non-conformal hence the resulting contact areas are very small and the resulting pressures are very high. From the viewpoint of machine design it is essential to know the values of stresses acting in such contacts. These stresses can be determined from the analytical formulae, based on the theory of elasticity, developed by Hertz in 1881.

A.2.1 Simplifying Assumptions to Hertz's Theory

Hertz's model of contact stress is based on the following simplifying assumptions (Johnson 1985; Stachowiak and Batchelor 2001):

- The materials in contact are homogeneous and the yield stress is not exceeded.
- Contact stress is caused by the load which is normal to the contact tangent plane which effectively means that there are no tangential forces acting between the solids.
- The contact area is very small compared with the dimensions of the contacting solids.
- The contacting solids are at rest and in equilibrium.
- The effect of surface roughness is negligible.

Subsequent refinements of Hertz's model by later researcher have removed most of these assumptions, and Hertz's theory forms the basis of the model of elastohydrodynamic lubrication.

A.2.2 Geometry of Contacting Elastic Bodies

The shape of the contact area depends on the shape (curvature) of the contacting bodies. For example, point contacts occur between two balls, line contacts occur between two parallel cylinders and elliptical contacts, which are most frequently found in many practical engineering applications, occur when two cylinders are crossed, or a moving ball is in contact with the inner ring of a bearing, or two gear teeth are in contact. The curvature of the bodies can be convex, flat or concave. It is defined by convention that convex surfaces possess a ‘positive curvature’ and concave surfaces have a ‘negative curvature’. The following general rule can be applied to distinguish between these surfaces: if the centre of curvature lies within the solid then the curvature is positive, if it lies outside the solid then the curvature is negative. This distinction is critical in defining the parameter characterising the contact geometry which is known as the reduced radius of curvature.

A.2.2.1 Two Elastic Bodies with Convex Surfaces in Contact

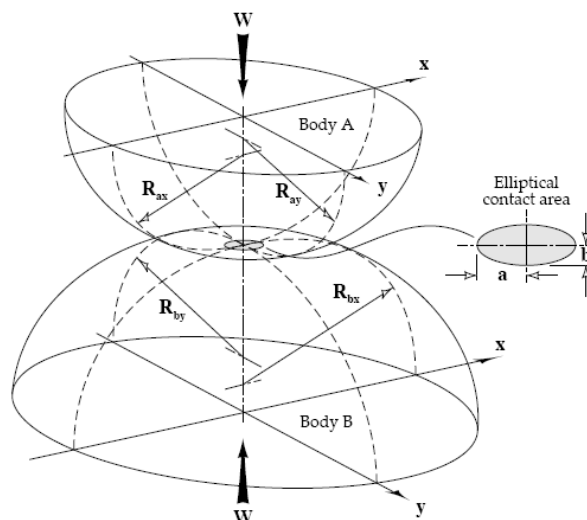


Figure A.1. Convex surfaces in contact (Stachowiak and Batchelor 2001)

$$\frac{1}{R'} = \frac{1}{R_x} + \frac{1}{R_y} = \frac{1}{R_{ax}} + \frac{1}{R_{bx}} + \frac{1}{R_{ay}} + \frac{1}{R_{by}} \quad (\text{A-1})$$

Equation (A-1). Reduced radius of curvature of two elastic bodies with convex surfaces in contacts

$$\frac{1}{R_x} = \frac{1}{R_{ax}} + \frac{1}{R_{bx}} \quad \frac{1}{R_y} = \frac{1}{R_{ay}} + \frac{1}{R_{by}} \quad (\text{A-2})$$

Equation (A-2). Reduced radius of curvature along 'x' and 'y' axis

Where:

R_x is the reduced radius of curvature in the 'x' direction [m]

R_y is the reduced radius of curvature in the 'y' direction [m]

R_{ax} is the reduced radius of curvature of body 'A' in the 'x' direction [m]

R_{ay} is the reduced radius of curvature of body 'A' in the 'y' direction [m]

R_{bx} is the reduced radius of curvature of body 'B' in the 'x' direction [m]

R_{by} is the reduced radius of curvature of body 'B' in the 'y' direction [m]

A.2.2.2 Convention

The method of arranging the 'x' and 'y' coordinates plays an important role in the calculation of contact parameters. It is important to locate the 'x' and 'y' coordinates so that the following condition is fulfilled, equation (A-3).

$$\frac{1}{R_x} \geq \frac{1}{R_y} \quad (\text{A-3})$$

Equation (A-3). Convection to locate 'x' and 'y' coordinates

When this rule is applied, the coordinate 'x' determines the direction of the semiminor axis of the contact and the coordinate 'y' determines the direction of the semimajor axis. If $1/R_x = 1/R_y$ then there is a circular contact and when $1/R_x < 1/R_y$ then it is necessary to transpose the directions of the coordinates, i.e. 'x' becomes 'y' and vice versa and 'Rx' becomes 'Ry'.

A.2.3 Two Elastic Bodies One Convex and One Flat Surface in Contact

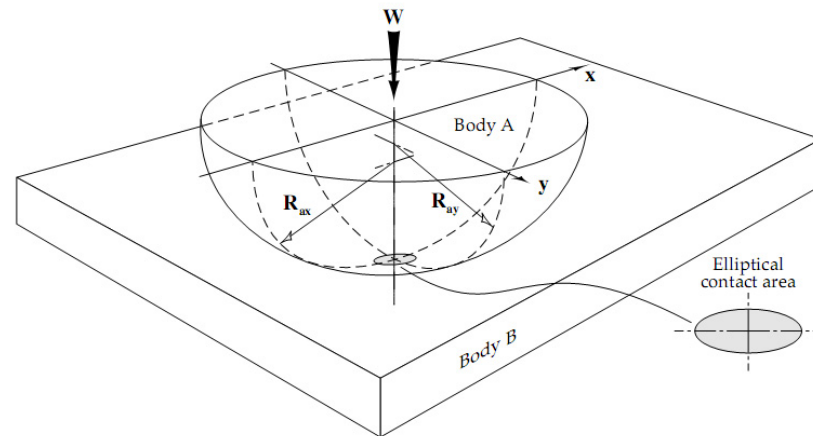


Figure A.2. One convex and one flat surface in contact (Stachowiak and Batchelor 2001)

Since one of the contacting surfaces is a plane then it has infinite radii of curvature, Equation (A-4).

$$\frac{1}{R'} = \frac{1}{R_x} + \frac{1}{R_y} = \frac{1}{R_{ax}} + \frac{1}{R_{ay}} \quad (\text{A-4})$$

$$R_{bx}} = R_{by} = \infty$$

Equation (A-4). Reduced radius of curvature of two elastic bodies one convex and one flat surface in contact

A.2.4 Two Elastic Bodies with One Convex and One Concave Surface in Contact

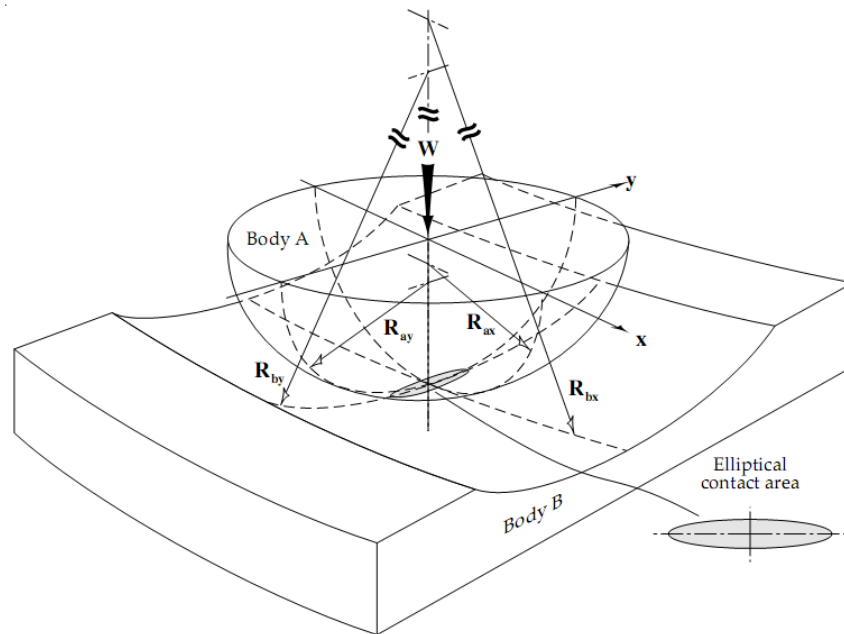


Figure A.3. One convex and one concave surface in contact (Stachowiak and Batchelor 2001)

Body ‘B’ has a concave surface and according to convention its curvature is negative, i.e. ‘ R_{bx} ’ and ‘ R_{by} ’ are negative.

$$\frac{1}{R'} = \frac{1}{R_x} + \frac{1}{R_y} = \left(\frac{1}{R_{ax}} + \frac{1}{R_{ay}} \right) - \left(\frac{1}{R_{bx}} + \frac{1}{R_{by}} \right) \quad (A-5)$$

Equation (A-5). Reduced radius of curvature of two elastic bodies one convex and one concave surface in contact

A.2.4.1 Contact area, pressure

The evaluation of contact parameters is essential in many practical engineering applications.

The most frequently used contact parameters are:

- The contact area dimensions.
- The maximum contact pressure, often called the Hertzian stress.
- The position of the maximum shear stress under the surface.

A.2.4.2 Contact between two spheres (point contact)

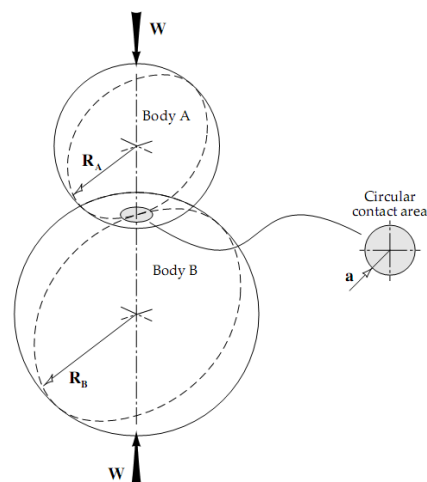


Figure A.4. Contact between two spheres (Stachowiak and Batchelor 2001)

$$\frac{1}{R'} = 2 \left[\frac{1}{R_A} + \frac{1}{R_B} \right] \quad (\text{A-6})$$

Equation (A-6). Reduced radius of curvature of two spheres in contact

$$\frac{1}{E'} = \frac{1}{2} \left[\frac{1-\nu_A^2}{E_A} + \frac{1-\nu_B^2}{E_B} \right] \quad (\text{A-7})$$

Equation (A-7). Reduced Young's modulus of two bodies

ν_A , $\nu_B \rightarrow$ are the Poisson's ratios of the contacting bodies 'A' and 'B' respectively

E_A , $E_B \rightarrow$ are the Young's moduli of the contacting bodies 'A' and 'B' respectively

$E' \rightarrow$ is the reduced Young's modulus [Pa]

$R' \rightarrow$ is the reduced radius of curvature [m]

For example, reduced Young's modulus for contact between steel spheres of $\nu_{Steel} = 0.3$ $E_{Steel} = 2.1 \cdot 10^{11}$ [Pa] thus $E' = 2.308 \cdot 10^{11}$ [Pa]

Contact area dimensions	Maximum contact pressure	Average contact pressure
$a = \left(\frac{3WR'}{E'} \right)^{1/3}$	$P_{\max} = \frac{3W}{2\pi a^2}$	$P_{\text{average}} = \frac{W}{\pi a^2}$
Table A.2. Contact between two spheres		

$a \rightarrow$ is the radius of the contact area [m]

$W \rightarrow$ is the normal load [N]

$P \rightarrow$ is the contact pressure (Hertzian stress) [Pa]

A.2.5 Contact between a sphere and a plane

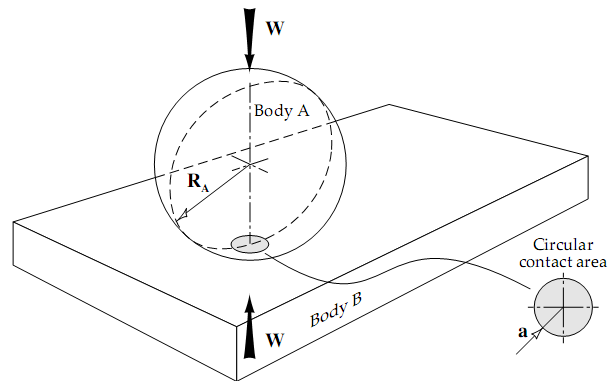


Figure A.5. Contact between a sphere and a plane (Stachowiak and Batchelor 2001)

The contact area between a sphere and a plane surface is also circular.

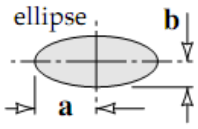
The radii of curvature of a plane's surface are infinite and symmetry of the sphere applies so that $R_{bx} = R_{by} = \infty$ and $R_{ax} = R_{ay} = R_A$. The reduced radius of curvature is:

$$\frac{1}{R'} = \frac{2}{R_A} \quad (\text{A-8})$$

Equation (A-8). Reduced radius of curvature between a sphere and a plane

A.2.6 Elliptical Contact between Two Elastic Bodies, General Case

Elliptical contacts are found between solid bodies which have different principal relative radii of curvature in orthogonal planes. Examples of this are encountered in spherical balls.

Contact area dimensions	Maximum contact pressure	Average contact pressure
$a = k_1 \left(\frac{3WR'}{E'} \right)^{1/3}$ $b = k_2 \left(\frac{3WR'}{E'} \right)^{1/3}$ 	$p_{\max} = \frac{3W}{2\pi ab}$ <p>Elliptical pressure distribution</p>	$p_{\text{average}} = \frac{W}{\pi ab}$
<p>Table A.3. Contact between two elastic bodies, general case (Stachowiak and Batchelor 2001)</p>		

$a \rightarrow$ is the semimajor axis of the contact ellipse [m]

$b \rightarrow$ is the semiminor axis of the contact ellipse [m]

$W \rightarrow$ is the normal load [N]

$P \rightarrow$ is the contact pressure (Hertzian stress) [Pa]

$E' \rightarrow$ is the reduced Young's modulus [Pa]

$R' \rightarrow$ is the reduced radius of curvature [m]

$k_1, k_2 \rightarrow$ are the contact coefficients

Contact coefficients can be found from the following charts:

$$k_0 = \frac{\left[\left(\frac{1}{R_{ax}} - \frac{1}{R_{ay}} \right)^2 + \left(\frac{1}{R_{bx}} - \frac{1}{R_{by}} \right)^2 + 2 \left(\frac{1}{R_{ax}} - \frac{1}{R_{ay}} \right) \left(\frac{1}{R_{bx}} - \frac{1}{R_{by}} \right) \cos 2\phi \right]^{1/2}}{\left(\frac{1}{R_{ax}} + \frac{1}{R_{ay}} + \frac{1}{R_{bx}} + \frac{1}{R_{by}} \right)}$$

$\phi \rightarrow$ is the angle between the plane containing the minimum principal radius of curvature of body 'A' and the plane containing the minimum principal radius of curvature of body 'B'.

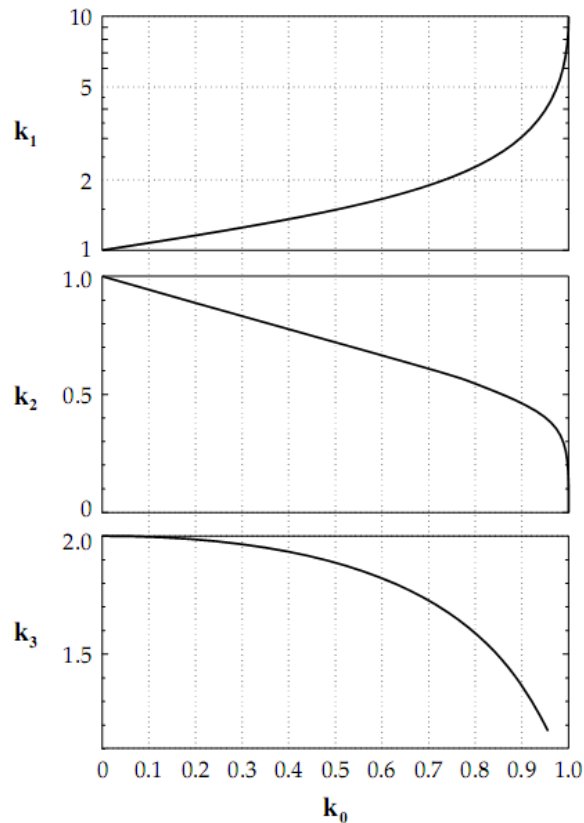
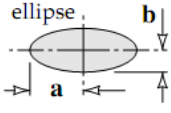


Figure A.6. k_1, k_2 dependence with k_0 (Stachowiak and Batchelor 2001)

A very useful development in the evaluation of contact parameters is due to Hamrock and Dowson. The method of linear regression by the least squares method has been applied to derive simplified expressions for the elliptic integrals required for the stress and deflection calculations in Hertzian contacts. The derived formulae apply to any contact and eliminate the need to use numerical methods or charts.

The formulae are shown below. Although they are only approximations, the differences between the calculated values and the exact predictions from the Hertzian analysis are very small.

Contact area dimensions		Maximum contact pressure	Average contact pressure
$a = \left(\frac{6\bar{k}^2 \bar{\varepsilon} W R'}{\pi E'} \right)^{1/3}$ $b = \left(\frac{6\bar{\varepsilon} W R'}{\pi \bar{k} E'} \right)^{1/3}$		$p_{\max} = \frac{3W}{2\pi ab}$	$p_{\text{average}} = \frac{W}{\pi ab}$
Maximum shear stress		Ellipticity parameter	
$\bar{\varepsilon} = 1.0003 + \frac{0.5968 R_x}{R_y}$ $\bar{\xi} = 1.5277 + 0.6023 \ln \left(\frac{R_y}{R_x} \right)$		$\bar{k} = 1.0339 \left(\frac{R_y}{R_x} \right)^{0.636}$	
<p>Table A.4. Contact between two elastic bodies, general case, Hamrock and Dowson simplification (Stachowiak and Batchelor 2001)</p>			

where

$\bar{\varepsilon}$ and $\bar{\xi}$ are the simplified elliptic integrals;

\bar{k} is the simplified ellipticity parameter. The exact value of the ellipticity parameter is defined as the ratio of the semiaxis of the contact ellipse in the transverse direction to the semiaxis in the direction of motion, i.e. $k = a/b$

A.2.7 Sliding speed and coefficient of friction for sliding four-ball and pin / ball-on-disc configurations

This section explains how to calculate the sliding speed, effective load and friction torque for four-ball sliding and pin/ball-on-disc test configurations. These expressions are important to calculate the effective load applied in the contact between balls for further estimation of the Hertz contact pressure needed for film thickness predictions.

A.2.7.1 Four Ball Sliding Wear Test

In this test the balls are placed according to Figure A.7.

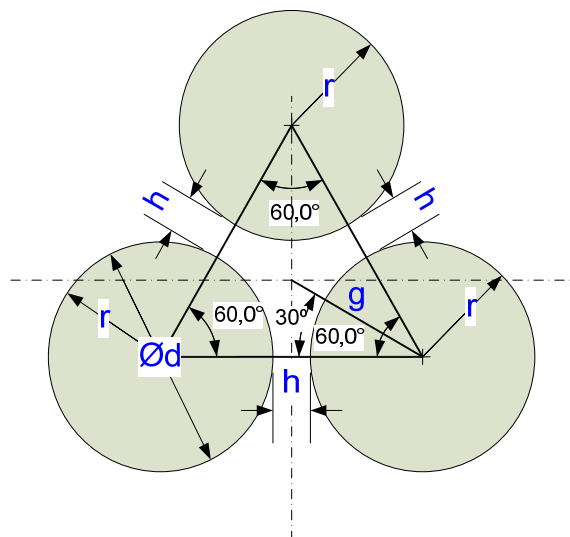


Figure A.7. Sliding four-ball test, lower balls and upper view

Where: **h** is the distance separating the balls. In pure sliding the balls are in contact, therefore $h=0$, **d** is the diameter of each ball, **r** is the radius of each ball.

With this configuration, the centres of the balls form an equilateral triangle.

g can be easily calculated (this relation is straightforward from the triangle forming a 30° angle)

$$g = \frac{\frac{d+h}{2}}{\cos(30^\circ)} = \frac{d+h}{\sqrt{3}} \quad (\text{A-9})$$

Equation (A-9). Reference distance to calculate α

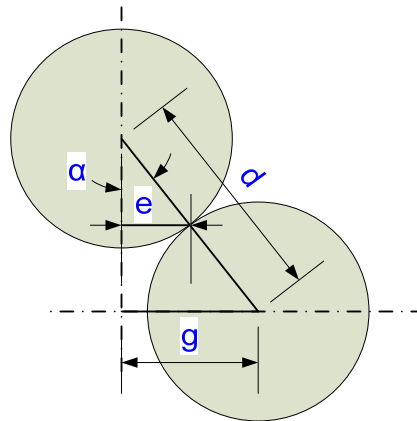


Figure A.8. Upper Ball and one of the lower balls, sectional view.

$$\alpha = \arcsin\left(\frac{g}{d}\right) = \arcsin\left(\frac{d+h}{\sqrt{3}d}\right) \quad (\text{A-10})$$

Equation (A-10). Angle between the upper ball and the lower balls

If $h=0$ (pure sliding ball test) $\rightarrow \alpha = \arcsin\left(\frac{1}{\sqrt{3}}\right) = 35.2644^\circ$ (A-11)

Equation (A-11). Angle between the upper ball and the lower balls when $h=0$

$$e = r \sin(\alpha) = \frac{r}{\sqrt{3}} \quad (\text{A-12})$$

Equation (A-12). Reference distance to calculate the sliding speed

$$v = \omega \cdot e = \omega \frac{r}{\sqrt{3}} \quad (\text{A-13})$$

Equation (A-13). Sliding speed in the contact between the upper and one of the lower balls

Where:

ω is the angular speed of the spindle.

The friction force according to the Coulomb friction model for each ball is defined as

$$F' = \mu \cdot N = \mu \cdot \frac{W}{3 \cdot \cos(\alpha)} = \mu \cdot \frac{W}{\sqrt{6}} \quad (\text{A-14})$$

Equation (A-14). Friction Force Four-Ball Sliding Test

Where:

W is the applied load.

μ is the coefficient of friction.

N is the normal load.

The friction force F' between the upper ball and the three lower balls produces a friction torque.

$$T = 3 \cdot F' \cdot e = 3 \cdot F' \cdot \frac{r}{\sqrt{3}} = 3 \cdot \mu \frac{W}{\sqrt{6}} \cdot \frac{r}{\sqrt{3}} = \frac{1}{\sqrt{2}} \mu \cdot W \cdot r \quad (\text{A-15})$$

Equation (A-15). Friction Torque in Four-Ball Sliding Test configuration

The torque T is obtained with equation 7 from the applied force to the torque load cell positioned at a distance q from the axis of rotation.

$$T = F_c \cdot q \quad (\text{A-16})$$

Equation (A-16). Friction Torque T and the force F_c measured with the load cell

F_c is the force measured with the torque transducer (load cell),

q is the horizontal distance from the centre of the upper ball (axis of rotation) to the point of contact on the load cell.

Once the torque is determined, the coefficient of friction is obtained from equation (A-17).

$$\mu = \frac{\sqrt{2} \cdot T}{W \cdot r} = \frac{\sqrt{2} \cdot F_c \cdot q}{W \cdot r} \quad (\text{A-17})$$

Equation (A-17). Coefficient of friction for 4 Ball Sliding Test Configuration

A.2.7.2 Pin and ball-on-disc

The expressions for pin and ball-on-disc assemblies are simpler than the previous ones. The sliding speed in the contact between the disc and the pin is the angular speed of the spindle by the distance from the pin to the axis of rotation.

The friction force assuming the Coulomb friction model is the normal force (the applied load) by the coefficient of friction.

The coefficient of friction is shown in Equation (A-19).

$$V = \omega_d \cdot R = \left(\frac{n_d}{60} \cdot \frac{2\pi}{1} \right) \cdot R = \frac{n_d \cdot \pi}{30} \cdot R \quad (\text{A-18})$$

Equation (A-18). The linear velocity, V , at the centre of the pin O_p

n_d angular speed of the spindle in revolutions per minute (rpm)

R is the horizontal distance from the centre of pin to the axis of rotation.

$$\mu = \frac{T}{W \cdot R} = \frac{F_c \cdot q}{W \cdot R} \quad (\text{A-19})$$

Equation (A-19). Coefficient of friction Pin and Ball-On-Disc

B Lubrication theory

B.1 Viscosity classification

B.1.1 ISO Viscosity Grades

Hydraulic fluid viscosity is classified according to ASTM D2422 (Industrial Fluid Lubricants by Viscosity System) commonly referred to as ISO viscosity grades. The ISO system classifies fluids solely on kinematic viscosity measured at 40°C. The choice of 40°C as the reference temperature is a compromise between maximum operating and ambient temperatures. However, it is also convenient because it is a reference temperature used for determination of the viscosity index, a commonly reported property of hydraulic fluids.

TABLE 1 Viscosity System for Industrial Fluid Lubricants^A

Viscosity System Grade Identification	Mid-Point Viscosity, cSt (mm ² /s) at 40.0°C	Kinematic Viscosity Limits, cSt (mm ² /s) at 40.0°C ^{B,C}	
		min	max
ISO VG 2	2.2	1.98	2.4
ISO VG 3	3.2	2.88	3.52
ISO VG 5	4.6	4.14	5.06
ISO VG 7	6.8	6.12	7.48
ISO VG 10	10	9.00	11.0
ISO VG 15	15	13.5	16.5
ISO VG 22	22	19.8	24.2
ISO VG 32	32	28.8	35.2
ISO VG 46	46	41.4	50.6
ISO VG 68	68	61.2	74.8
ISO VG 100	100	90.0	110
ISO VG 150	150	135	165
ISO VG 220	220	198	242
ISO VG 320	320	288	352
ISO VG 460	460	414	506
ISO VG 680	680	612	748
ISO VG 1000	1000	900	1100
ISO VG 1500	1500	1350	1650
ISO VG 2200	2200	1980	2420
ISO VG 3200	3200	2880	3520

^AThis system implies no evaluation of quality.

^BThis system is used in ISO 3448.

^CIf 40°C is not the temperature used when determining the viscosity (as is sometimes the case with very viscous fluids) then the related viscosity at 40°C shall be established by using Viscosity Temperature Charts D341.

Figure B.1. Viscosity System for Industrial Fluid Lubricants

(Theo Mang 2007)

B.1.2 Hydraulic Fluid Classification

For hydraulic fluid classification, ASTM D6080 (Standard Practice for Defining the Viscosity Characteristics of Hydraulic Fluids) builds upon the current ISO VG classification in order to describe rheological properties.

This standard addresses the following:

- Cold temperature requirements.
- Multigraded (HV) oils.
 - Description
 - Viscosity after shearing.
- Viscosity at operating temperature (viscosity index).
- Classification of fluids outside ISO ranges.

Viscosity Grade	Temperature, ° C, for Brookfield Viscosity of 750 mPa·s ^A	
	min	max
L5	...	-50
L7	-49	-42
L10	-41	-33
L15	-32	-23
L22	-22	-15
L32	-14	- 8
L46	- 7	- 2
L68	- 1	4
L100	5	10
L150	11	16

^A The temperature range for a given L-grade is approximately equivalent to that for an ISO grade of the same numerical designation and having a viscosity index of 100, that is, the temperature range for the L10 grade is approximately the same as that for an ISO VG 10 grade with a viscosity index of 100.

Figure B.2. Low temperature viscosity grades for hydraulic fluid classification (ASTM 6080) (Nadkarni 2000)

The cold-temperature grade is an appropriate ISO number preceded by “L” (for low temperature) with measurements by a Brookfield viscometer (ASTM D 2983) and a limit of 750 cP maximum. This limit was chosen because it represents a relatively severe case for allowable viscosity at pump start-up.

B.1.3 Viscosity temperature relationships

The viscosity of lubricating oils is extremely sensitive to the operating temperature (Keith and Hodges 1996; Totten 1999; Totten et al. 1997; Totten et al. 2001)

B.1.3.1 Viscosity-Temperature Equations

There are several viscosity-temperature equations available, some of them are purely empirical whereas others are derived from theoretical models. The most commonly used equations are summarised in Table B.1. (Totten 1999)

Name	Equation	Comments
Reynolds	$\eta = b \cdot e^{-aT}$	Early equation; accurate only for a very limited temperature range
Slotte	$\eta = \frac{a}{(b + T)^c}$	Reasonable; useful in numerical analysis
Walther	$(\nu + a) = b \cdot d^{1/T^c}$	Forms the basis of the ASTM viscosity-temperature chart
Vogel	$\eta = a \cdot e^{b/(T-c)}$	Most accurate; very useful in engineering calculations
Table B.1. Viscosity-temperature equations (Stachowiak and Batchelor 2001)		

where:

a, b, c, d are constants;

ν is the kinematic viscosity [m^2/s];

T is the absolute temperature [K].

Among them the most accurate is the Vogel equation. Three viscosity measurements at different temperatures for specific oil are needed in order to determine the three constants in this equation. A numerical approach to fit some of these models can be found in (Perez 2010).

B.1.3.2 Effect of temperature

The most widely used chart is the ASTM (American Society for Testing Materials) Viscosity-Temperature chart (ASTM D341) which is entirely empirical and is based on Walther's Equation first described in 1921 (Theo Mang 2007; Totten 1999).

$$\log \log(\nu + 0.7) = A - B \cdot \log T \quad (\text{B-1})$$

Equation (B-1). Walther's equation

where

ν is the kinematic viscosity

T is the absolute temperature

A, B are constants for a given liquid.

The negative slope B indicates that higher temperatures result in lower viscosities. The slope $B = (1/\nu)(\delta\nu/\delta T)$, is usually referred to as the viscosity-temperature coefficient.

B.1.3.3 Viscosity index

The extent of decreasing viscosity with temperature is described by viscosity index (VI), a dimensionless number calculated from measured viscosities using ASTM D2270 (Calculating Viscosity Index from Kinematic Viscosity at 40° and 100°C). This arbitrary VI scale was originally devised by Dean and Davis. The viscosity index is an entirely empirical parameter which compares the kinematic viscosity of the oil of interest to the viscosities of two reference oils which have a considerable difference in sensitivity of viscosity to temperature.

$$\text{Viscosity index} = \left(\frac{L-U}{L-H} \right) \cdot 100 \quad (\text{B-2})$$

Equation (B-2). Viscosity index

Firstly the kinematic viscosity of the oil of interest is measured at 40°C ('U') and at 100°C. Then from Table 2.2 (ASTM D2270), looking at the viscosity at 100°C of the oil of interest, the corresponding values of the reference oils, 'L' and 'H' are read. Substituting the obtained values of 'U', 'L' and 'H' into the above equation yields the viscosity index.

The viscosity index is an inverse measure of the decline in oil viscosity with temperature. High values indicate that the oil shows less relative decline in viscosity with -temperature. The viscosity index of most of the refined mineral oils available on the market is about 100, whereas multigrade and synthetic oils have higher viscosity indices of about 150.

B.1.3.4 Viscosity pressure relationship

Lubricant viscosity increases with pressure. The best known equation to calculate the viscosity of a lubricant at moderate pressures (close to atmospheric) is the Barus equation. The application of this equation to pressures above 0.5 [GPa] can, however, lead to serious errors. The equation becomes even more unreliable if the ambient temperature is high (Theo Mang 2007; Totten 1999). The Barus equation is of the form, Equation (B-3)

$$\eta_p = \eta_0 e^{\alpha p} \quad (\text{B-3})$$

Equation (B-3). Barus Law

Where:

η_p is the lubricant viscosity a pressure 'p' and temperature 'θ' [Pa·s];

η_0 is the viscosity at atmospheric pressure and temperature 'θ' [Pa·s];

α is the pressure-viscosity coefficient [m^2/N];

p is the pressure of concern [Pa].

For hydraulic fluids, the approximate range of the viscosity pressure coefficient is $(1.5 - 5.0) \times 10^{-8} \text{Pa}^{-1}$. At pressures less than 280 MPa the difference among viscosity-pressure coefficients of different fluids is relatively small, but at high pressures, the differences can be quite marked. A more accurate relationship for high pressures is the Roelands equation (Roelands 1966), temperature effects are also included in Roelands formula. Density variations due to high pressures are typically modelled using the Dowson and Higginson relationship (Dowson and Higginson 1966) which is typically used for film thickness prediction in EHL contacts.

Finally, Barus Law can be used within a narrow gap of high pressures for EHL film thickness prediction if an effective pressure-viscosity coefficient can be determined. A suitable experimental procedure to obtain this effective α can be found in (Perez 2010).

B.1.4 Load-Bearing Capacity and Antiwear Properties

A high load-bearing capacity is one of the most important requirements to be met by a hydraulic fluid, and this also implies good antiwear properties. The dynamic viscosity is the most important parameter for the antiwear properties for hydrodynamic lubrication. If the forces acting at low sliding velocities in the mixed friction zone are not sufficient to separate the mating frictional surfaces completely, then the antiwear properties of a fluid are determined by its ability to wet a metal surface and form friction-reducing reaction layers on the mating faces. The wetting capability is also referred to as “oiliness.” The load-bearing capacity of a hydraulic fluid in the mixed friction zone can be improved by means of antiwear additives and substances which reduce the coefficient of friction (Totten 1999).

B.1.5 Setting Point/Pour Point

The setting point of a fluid is determined by the temperature at which the medium just ceases to flow under certain testing conditions. By comparison, the

pour-point corresponds to the temperature at which the medium just continues to flow. This is around 6 °C to 8 °C higher than the setting point. Determination of the pour-point alone is not admissible for an evaluation of the low-temperature characteristics of ester-based hydraulic media. The slow crystallisation processes occurring here are such that the time dependence must be taken into consideration as well as just the temperature, and this is determined by means of special test procedures (Totten 1999).

B.1.6 Gas Solubility

All hydraulic fluids are capable of dissolving a certain proportion of gas. This gas solubility is proportional to the pressure up to around 300 bar (Henry's Law "Solubility of gases in fluids").

Under normal circumstances, dissolved air does not exert any influence on the properties of the hydraulic fluid. It may bleed out of the fluid, however, if static pressure is applied locally, particularly if the fluid is simultaneously subjected to shearing stresses. The process is referred to as cavitation.

Cavitation and cavitation erosion, which presents a serious problem, particularly with respect to water-based fluids, can be reduced by suitable design measures, such as the selection of special materials for surfaces which are susceptible to erosion, guiding the cavitation stream into uncritical areas away from the walls, or diminishing the differential pressure at one resistor by fitting several resistors, one behind the other (Totten 1999).

B.1.7 Ageing Behaviour

The term “ageing” includes changes that take place in the composition and chemical structure of a hydraulic fluid. Ageing is brought about by such chemical reactions as oxidation, hydrolysis, polymerization, and thermal decomposition, or by mechanical influences, such as shearing action. Oxidation refers to the reaction with O_2 producing left-over acids. Polymerisation refers to the enlargement of hydrocarbons resulting from the formation of side chains or macromolecules. This process produces waste products such as sludge or resin-like coatings on components. Hydrolysis refers to the cracking of esters when they come into contact with water. The ageing process breaks down or destroys the additives and, at the same time, changes the molecules of the base fluid. It is accelerated by high operating temperatures and contamination through extraneous air, water, and metallic catalysts, predominantly copper, copper alloys and iron. One example of a measure for the ageing condition of a fluid is the acid number (AN). The acid value indicates the acid content of a fluid by defining how many milligrams of caustic potash solution would be required to neutralise 1 gram of a sample (Totten 1999)

The ageing stability of mineral oils can be improved by means of certain additives.

B.1.7.1 Detergent and Dispersant

The separation of water and solids in the tank offers a satisfactory means of keeping the fluid clean, at least for low circulation rates. Mineral oils and HFD fluids usually have a good separating capacity, which is referred to as their detergent properties.

Owing to their large-scale use in engine oils, the surface-active detergents and dispersants have become the most significant types of additives, accounting for approximately 50% of the market share. Their function is to keep oil-insoluble substances, resinous and bituminous oxidation products and water in suspension, or to accelerate their sedimentation in order to prevent deposits forming on metal surfaces, thickening of the fluid, precipitation of sludge and corrosion. The

dispersant or detergent effects of many of these additives depend on the respective concentration (Totten 1999).

B.2 Film thickness formulae for EHL contacts

Elastohydrodynamic lubrication which is commonly abbreviated to EHL or EHD is a form of hydrodynamic lubrication where the elastic deformation of the contacting bodies and the changes of viscosity with pressure play fundamental roles.

The lubricating films are very thin, in the range of 0.1 to 1 [μm], but manage to separate the interacting surfaces, resulting in a significant reduction of wear and friction. Although this regime generally operates between non-conforming surfaces, it can also occur under certain circumstances in the contacts classified as conformal such as highly loaded journal and pad bearings which have a significant component of contact and bending deformation.

The term elastohydrodynamic lubricating film refers to the lubricating oil which separates the opposing surfaces of a concentrated contact. The applied load is concentrated in a reduced area, leading to very high pressure values. This pressure has two beneficial consequences in the lubrication regime: the elastic deformation of the surfaces and the viscosity increment of the lubricant. According to these effects, the film thickness in EHD is much higher than the film thickness predicted by the theory of hydrodynamic lubrication. The EHD theory takes into account the combination of the analytical solution of Reynolds equation, the elasticity Hertz theory and the pressure-viscosity coefficient.

The viscosity-pressure relationship is usually described by a mathematically convenient but approximate equation known as the Barus law. A method to determine the pressure-viscosity coefficient is found in (Perez 2010).

The exact analysis of elastohydrodynamic lubrication by Hamrock and Dowson provided the most important information about EHL. The results of this analysis are the formulae for the calculation of the minimum film thickness in elastohydrodynamic contacts.

The formulae derived by Hamrock and Dowson apply to any contact, such as point, linear or elliptical, and are now routinely used in EHL film thickness calculations. They can be used with confidence for many material combinations including steel on steel even up to maximum pressures of 3-4 [GPa]. The numerically derived formulae for the central and minimum film thicknesses are in the following form (Hamrock et al. 2004; Rodríguez 2007).

$$\frac{h_c}{R'} = 2.69 \left(\frac{U\eta_0}{E'R'} \right)^{0.67} (\alpha E')^{0.53} \left(\frac{W}{E'R'^2} \right)^{-0.067} (1 - 0.61e^{-0.73k}) \quad (\text{B-4})$$

Equation (B-4). Hamrock and Dowson central film thickness

$$\frac{h_o}{R'} = 3.63 \left(\frac{U\eta_0}{E'R'} \right)^{0.68} (\alpha E')^{0.49} \left(\frac{W}{E'R'^2} \right)^{-0.073} (1 - e^{-0.68k}) \quad (\text{B-5})$$

Equation (B-5). Hamrock and Dowson minimum film thickness

h_c is the central film thickness [m]

h_o is the minimum film thickness [m]

U is the entraining surface velocity [m/s], i.e. $U = (U_A + U_B)/2$, where the subscripts

‘A’ and ‘B’ refer to the velocities of bodies ‘A’ and ‘B’ respectively;

η_0 is the viscosity at atmospheric pressure of the lubricant [Pa·s]

E' is the reduced Young's modulus [Pa]

R' is the reduced radius of curvature [m]

α is the pressure-viscosity coefficient [m^2/N]

W is the contact load [N]

k is the ellipticity parameter defined as: $k = a/b$, where ‘a’ is the semiaxis of the contact ellipse in the transverse direction [m] and ‘b’ is the semiaxis in the direction of motion [m]. It can be calculated with sufficient accuracy from:

$$\bar{k} = 1.0339 \left(\frac{R_y}{R_x} \right)^{0.636} \quad k = 1 \text{ for point contact } k = \infty \text{ for line contact}$$

R_x, R_y are the reduced radii of curvature in the 'x' and 'y' directions respectively.

The above EHL film thickness equations are applicable for 'k' values between 0.1 and ∞ .

Although the EHL film thickness equations (B.14-15) apply to most of the elastohydrodynamic contacts, there may be some practical engineering applications where more precise formulae can be used. For example, in heavily-loaded contacts where the elastic deformations and changes in viscosity with pressure are significant, equations (B.14-15) give accurate film thickness predictions. However, there are other engineering applications, such as very lightly-loaded rolling bearings where the elastic and viscosity effects are small, yet the contacts are classified as elastohydrodynamic. The magnitude of elastic deformation and changes in lubricant viscosity depend mostly on the applied load and the Young's modulus of the material. Depending on the values of load and material properties, the changes in film geometry and lubricant viscosity can be either more or less pronounced. In general, four well-defined lubrication regimes are distinguished in full-film elastohydrodynamics. Each of these regimes is characterised by the operating conditions and the properties of the material. Accurate equations for minimum film thickness have been developed for each of these regimes which are:

- isoviscous-rigid body (comparable to classical hydrodynamics),
- piezoviscous-rigid body.
- isoviscous-elastic body.
- piezoviscous-elastic body.

These four regimes are defined as follows:

Isoviscous-rigid: In the isoviscous-rigid regime, elastic deformations are small and can be neglected. The maximum film pressure is too low to significantly increase the lubricant viscosity. This form of lubrication is typically encountered in circular-arc thrust bearing pads and in industrial coating processes in which paint, emulsion, or protective coatings are applied to sheet or film materials passing between rollers.

Viscous-rigid: If the contact pressure is sufficiently high to significantly increase the fluid viscosity within the conjunction, it may be necessary to consider pressure-viscosity characteristics of the lubricant while assuming that the solids remain rigid. For the latter part of this assumption to be valid the deformation of the surfaces must remain insignificant relative to the fluid film thickness. This form of lubrication may be encountered on roller end-guide flanges, in contacts in moderately loaded cylindrical tapered rollers, and between some piston rings and cylinder liners.

Isoviscous-elastic: In this regime the elastic deformation of the solids is significant relative to the thickness of the fluid film separating them, but the contact pressure is quite low and insufficient to cause any substantial increase in viscosity. This situation arises with low-elastic modulus materials (soft EHL) and may be encountered in seals, human joints, tyres, and elastomeric-material machine elements.

Viscous-elastic: In fully-developed elastohydrodynamic lubrication the elastic deformation of the solids is often significant relative to the thickness of the fluid film separating them, and the contact pressure is high enough to significantly increase the lubricant viscosity within the conjunction. This form of lubrication is typically encountered in ball and roller bearings, gears and cams.

Using a convenient set of dimensionless numbers, Table B., Hamrock and Dowson proposed a suitable film thickness formula for each lubrication regime assuming isothermal and fully-flooded conjunctions, Table B.3. Moes & Venner derived a most accurate set of film thickness expressions for line and point contacts

(Venner 1991; Venner and Napel 1992), and Nijenbanning (Nijenbanning et al. 1994) extended the formulae to elliptical contacts, Table B.4.

W	Dimensionless load parameter	$W = \left(\frac{W}{E' R^2} \right)$
G	Dimensionless materials parameter	$G = (\alpha E')$
U	Dimensionless speed parameter	$U = \left(\frac{U \eta_0}{E' R'} \right)$
H	Dimensionless film parameter	$H = \frac{h}{R'}$
k	Dimensionless ellipticity parameter	$k = \frac{a}{b} = \frac{R_x}{R_y}$
$g_V = \frac{GW^3}{U^2}$		$g_E = \frac{W^{8/3}}{U^2}$
Table B.2. Dimensionless groups (H,U,W,G,k) used to defined the film thickness formulae valid for isothermal, fully flooded conjunctions (Stachowiak and Batchelor 2001)		

EHD regime	Equations
Isoviscous-rigid Rigid Isoviscous asymptote	$\left(\hat{H}_{\min} \right)_{IR} = \left(\hat{H}_c \right)_{IR} = 128 \alpha_r \left[0.131 \tan^{-1} \left(\frac{\alpha_a}{2} \right) + 1.683 \right]^2 \varphi^2$ $\alpha_r = \frac{R_y}{R_x} = k^{\pi/2}$ $\varphi = \left(1 + \frac{2}{3 \alpha_r} \right)^{-1}$
Viscous-rigid Rigid Piezoviscous asymptote	$\left(\hat{H}_{\min} \right)_{VR} = \left(\hat{H}_c \right)_{VR} = 1.66 g_V^{0.667} \left(1 - e^{-0.68 \alpha_r^{2/\pi}} \right)$
Isoviscous-elastic Elastic Isoviscous asymptote	$\left(\hat{H}_{\min} \right)_{IE} = 8.70 g_E^{0.67} \left(1 - 0.85 e^{-0.31 k} \right)$ $\left(\hat{H}_c \right)_{IE} = 11.15 g_E^{0.67} \left(1 - 0.72 e^{-0.28 k} \right)$

Viscous-elastic	$\left(\hat{H}_{\min}\right)_{VE} = 3.42g_V^{0.49}g_E^{0.17}\left(1 - e^{-0.68k}\right)$
Elastic Piezoviscous asymptote	$\left(\hat{H}_c\right)_{VE} = 3.61g_V^{0.53}g_E^{0.13}\left(1 - 0.61e^{-0.73k}\right)$
Table B.3. Film thickness formulae in EHD regime valid for isothermal, fully flooded conjunctions (Stachowiak and Batchelor 2001)	

$M = W \cdot (2 \cdot U)^{\frac{3}{4}}$	$L = G \cdot (2 \cdot U)^{\frac{1}{4}}$
Table B.4. Relationship between dimensionless groups introduced by Moes and dimensionless groups of Hamrock and Dowson.	

$H_c = \left[\left(H_{RI}^{3/2} + (H_{EI}^{-4} + H_{00}^{-4})^{-3/8} \right)^{2s/3} + (H_{RP}^{-8} + H_{EP}^{-8})^{-s/8} \right]^{1/s}$
$s = 1.5 \left(1 + e^{-1.2 \frac{H_{EI}}{H_{RI}}} \right)$
$H_{00} = 1.8 \cdot k^{-1}$
$H_{RI} \approx 145 \left(1 + 0.796 \cdot k^{14/15} \right)^{-15/7} k^{-1} M^{-2}$
$H_{EI} \approx 3.18 \left(1 + 0.006 \cdot k + 0.63 \cdot k^{4/7} \right)^{-14/25} k^{-1/15} M^{-2/15}$
$H_{RP} \approx 1.29 \left(1 + 0.691 \cdot k \right)^{-2/3} L^{2/3}$
$H_{EP} \approx 1.48 \left(1 + 0.006k + 0.63 \cdot k^{4/7} \right)^{-7/20} k^{-1/24} M^{-1/12} L^{3/4}$
Table B.5. Central film thickness (Nijenbanning et al. 1994)

In engineering calculations it is important first to assess which EHL regime applies to the contact or mechanical component under study and then apply the appropriate equation to determine the minimum film thickness.

In the evaluation of EHL film thickness it has been assumed that the contacting surfaces lubricated by elastohydrodynamic films are flat. In practice the surfaces are rough.

The relationship between the dimensionless film parameter λ and the minimum film thickness h_{\min} is shown in Equation (B-6).

$$\lambda = \frac{h_{\min}}{R_q} \quad R_q = \sqrt{(R_{q,a}^2 + R_{q,b}^2)} \quad (\text{B-6})$$

$R_{q,a}$, $R_{q,b}$ are the root mean square surface finishes of solids **a** and **b** respectively

Equation (B-6). Definition of the dimensionless film parameter λ

The film parameter is used to define the four important lubrication regimes.

The range of λ for these four regimes is

1. Hydrodynamic lubrication, $5 < \lambda < 100$
2. Elastohydrodynamic lubrication, $3 < \lambda < 10$
3. Partial lubrication, $1 < \lambda < 5$
4. Boundary lubrication, $\lambda < 1$

These values are rough estimates. The great differences in geometric conformity in going from hydrodynamically-lubricated conjunctions to elastohydrodynamically-lubricated conjunctions make it difficult for clear distinctions to be made.

It is important to consider that running-in is a process that affects the film parameter. This process allows wear to occur so that the mating surfaces can adjust to each other to provide for smooth running. This type of wear may be viewed as beneficial. The film parameter will increase with running-in, since the composite surface roughness will decrease. Running-in also has a significant effect on the shape

of the asperities that is not captured by the composite surface roughness. With running-in, the peaks of the asperities in contact become flattened.

Figures B.3-6 show an application example of film thickness formulae under Four-Ball Sliding Wear test configuration when the applied load is constant (147N), the spindle speed is within 100 to 2000 rpm and the temperature varies from 40 to 100°C. The balls are chromium steel balls of grade 10 ($R_a=20\text{nm}$ and R_q assuming sinusoidal profiles is $R_a/1.11$) and radius 6.35mm. The lubricant properties are 46cSt@40°C, 5cSt@100°C, density 874kg/m^3 , pressure-viscosity coefficient $2.3 \cdot 10^8 \text{ m}^2/\text{N}$. Similar charts can be represented assuming a constant temperature and varying the load and speed. The importance of these charts is the graphical and intuitive idea of the degree of influence of each parameter in the final prediction of the lubrication regime.

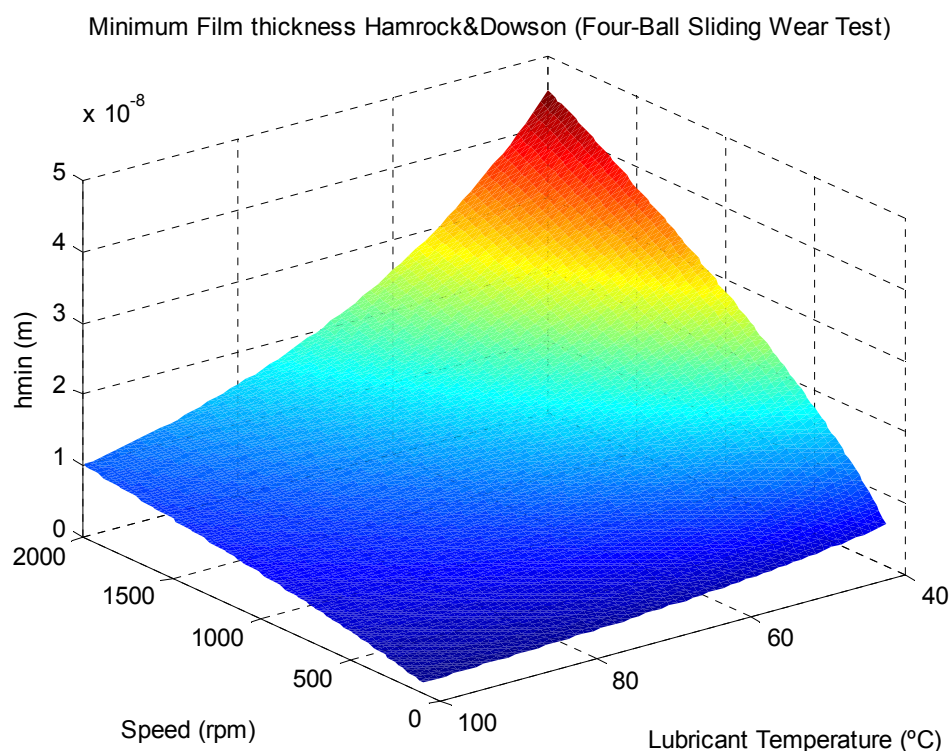


Figure B.3. Film thickness chart using Hamrock and Dowson Formulae (3D)

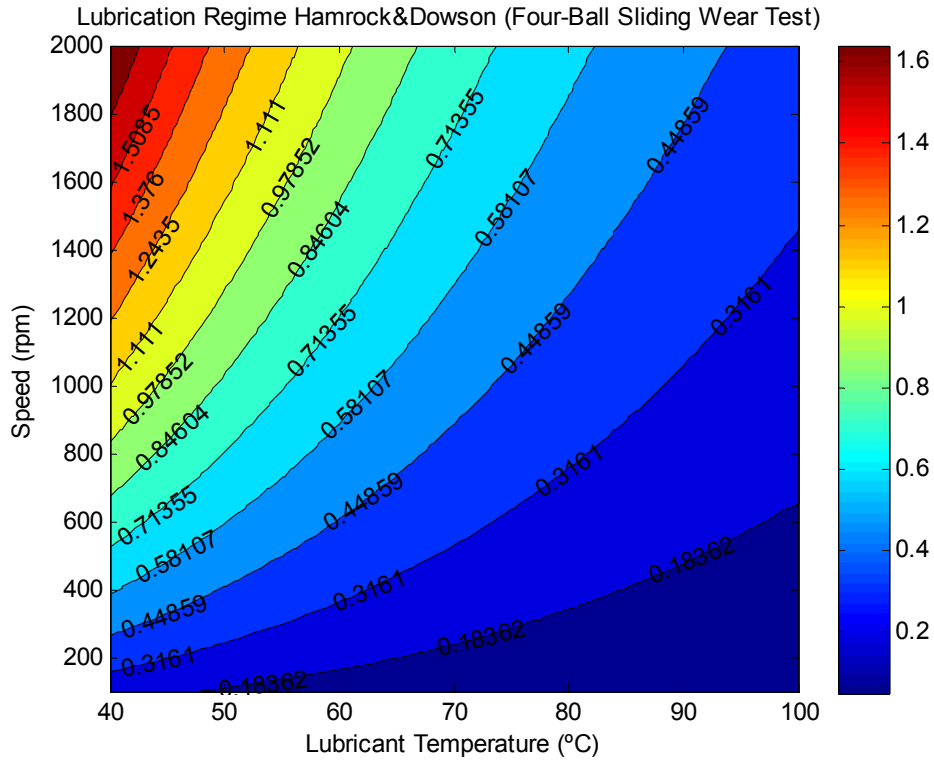


Figure B.4. Lubrication regime map using Hamrock and Dowson Formulae (2D)

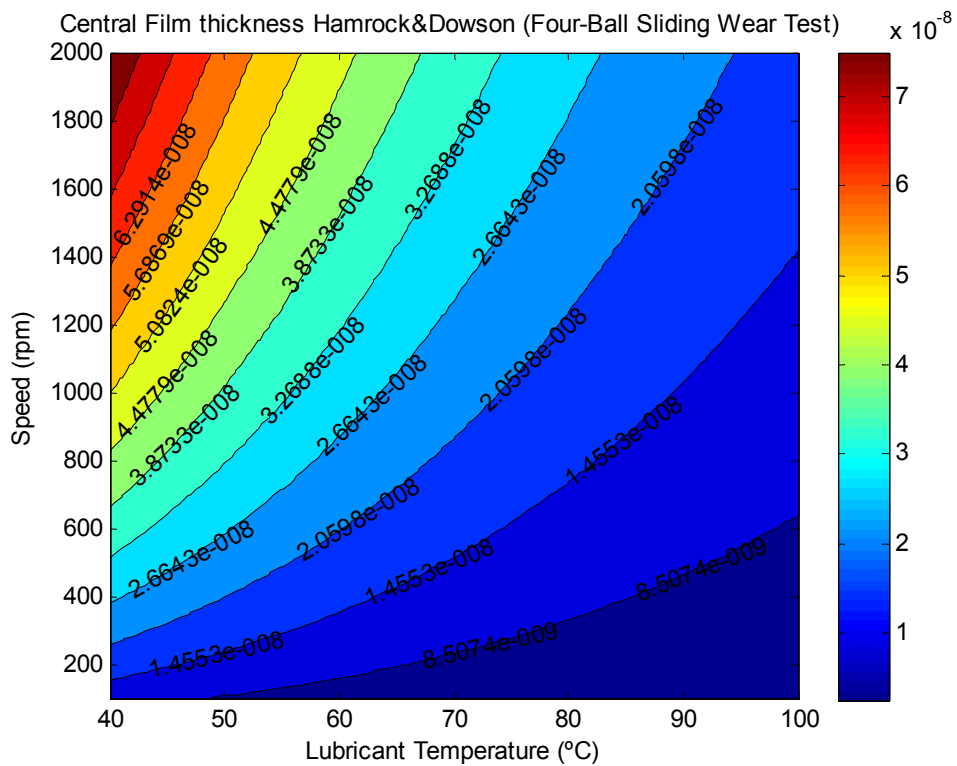


Figure B.5. Central film thickness map using Hamrock and Dowson Formulae (2D)

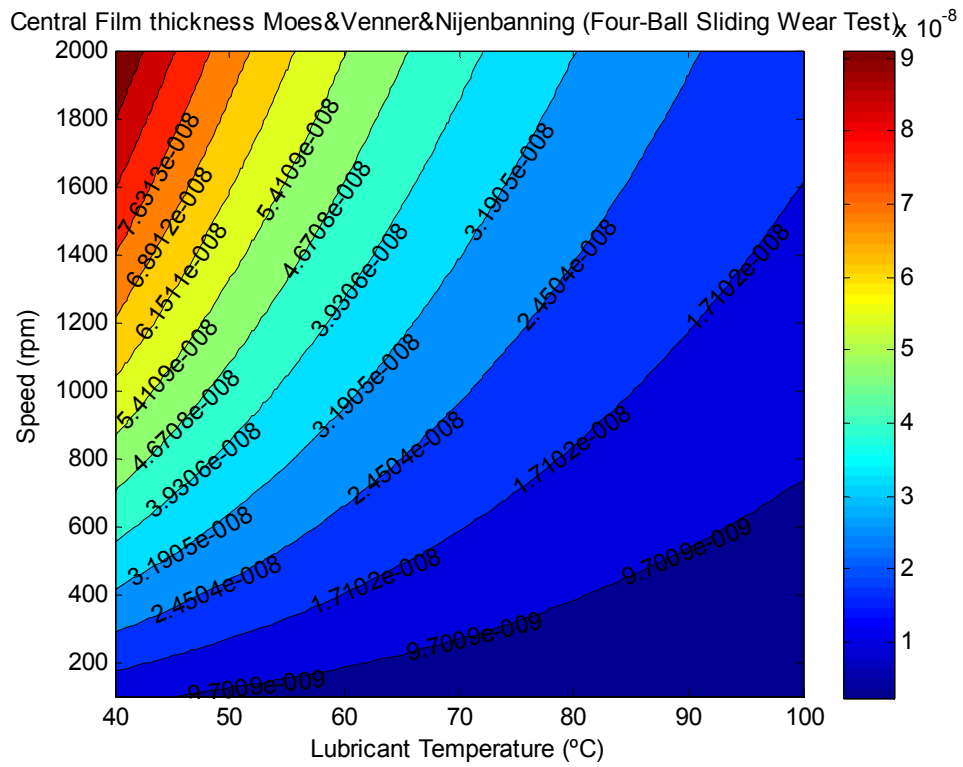


Figure B.6. Central film thickness map using Nijebanning Formulae (2D)

B.2.1 Surface temperature at the conjunction between contacting solids

Surface temperature has a strong effect on EHL, as is the case with hydrodynamic lubrication. Elevated temperatures lower the lubricating oil viscosity and usually decrease the pressure-viscosity coefficient ' α '. A reduction in either of these parameters will reduce the EHL film thickness which may cause lubricant failure. Excessively high temperatures may also interfere with some auxiliary mechanisms of lubrication necessary for the stable functioning of partial EHL (Rodríguez 2007; Stachowiak and Batchelor 2001).

Calculation of Surface Conjunction Temperature EHL is almost always found in concentrated contacts and in order to estimate the temperature rise during sliding contact, it is convenient to model the contact as a point or localised source of heat as a first approximation. Since the intense release of frictional heat occurs over the small area of a concentrated contact, the resulting frictional temperatures within the contact are high, even when outside temperatures are close to ambient. The temperatures at the interface between contacting and mutually sliding solids is known as the 'surface conjunction temperature' (Stachowiak and Batchelor 2001)

For any specific part of the sliding surface, frictional temperature rises are of very short duration and the temperatures generated are called 'flash temperatures'.

From the engineering viewpoint it is important to know the expected values of these temperatures since they can severely affect not only EHL but also wear and dry friction through the formation of oxides, production of metallurgically-transformed surface layers, alteration of local geometry caused by thermal expansion effects, or even surface melting. As well as the transient 'flash temperatures' there is also a steady state 'flash temperature rise' at the sliding contact. When the contact is efficiently lubricated, the transient flash temperatures are relatively small and are superimposed on a large, steady-state temperature peak. In dry friction, or where lubrication failure is imminent, the transient flash temperatures may become larger than the steady-state component (Stachowiak and Batchelor 2001).

The flash temperature theory was originally formulated by Blok in 1937 and developed further by Jaeger in 1944 and Archard in 1958. The theory provides a set of formulae for the calculation of flash temperature for various velocity ranges and contact geometries. According to Blok, Jaeger and Archard's theory, the flash temperature is the temperature rise above the temperature of the solids entering the contact which is called the 'bulk temperature'. The maximum contact temperature has therefore two components: the bulk temperature of the contacting solids and the maximum flash temperature rise, Equation (B-7).

$$T_c = T_b + T_{f_{\max}} \quad (\text{B-7})$$

Equation (B-7). Surface contact temperature

$T_c \rightarrow$ is the maximum surface contact temperature [$^{\circ}\text{C}$]

$T_b \rightarrow$ is the bulk temperature of the contacting solids before entering the contact [$^{\circ}\text{C}$]

$T_{f_{\max}} \rightarrow$ is the maximum flash temperature [$^{\circ}\text{C}$]

Evaluation of the flash temperature is basically a heat transfer problem where the frictional heat generated in the contact is modelled as a heat source moving over the surface. The following simplifying assumptions are made for the analysis (Stachowiak and Batchelor 2001):

- thermal properties of the contacting bodies are independent of temperature,
- the single area of contact is regarded as a plane source of heat,
- frictional heat is uniformly generated at the area of the contact,
- all heat produced is conducted into the contacting solids,
- the coefficient of friction between the contacting solids is known and attains some steady value.
- a steady state condition (i.e. $\partial T/\partial t = 0$, the temperature is steady over time) is attained.

Some of these assumptions appear to be dubious. For example, the presence of the lubricant in the contact will affect the heat transfer characteristics. Although most of the heat produced will be conducted into the solids, a portion of it will be convected away by the lubricant resulting in cooling of the surfaces. An accurate value of the coefficient of friction is very difficult, if not impossible, to obtain. The friction coefficient is dependent on the level of the heat generated as well as many other variables such as the nature of the contacting surfaces, the lubricant used and the lubrication mechanism acting. Even when an experimental measurement of the friction coefficient is available, in many cases the friction coefficient continually varies over a wide range. It is therefore necessary to calculate temperatures using a minimum and maximum value of friction coefficient. Temperatures at the beginning of sliding movement should also be considered since flash temperatures do not form instantaneously. Flash temperatures tend to stabilise within a very short sliding distance but the gradual accumulation of heat in the surrounding material and consequent slow rise in bulk temperature should not be overlooked. Notwithstanding these assumptions, the analysis gives temperature predictions which, although not very precise, are a good indication of the temperatures that might be expected between the operating surfaces.

A non-dimensional measure of the speed at which the ‘heat source’ moves across the surface called the ‘Peclet number’ has been introduced as a criterion allowing the differentiation between various speed regimes.

The Peclet number is defined in Equation (B-8).

$$L = \frac{Ua}{2\chi} \quad (\text{B-8})$$

Equation (B-8). Peclet number

where

L is the Peclet number;

U is the velocity of a solid (‘A’ or ‘B’) [m/s];

a is the contact dimension [m], (i.e. contact radius for circular contacts, half width of the contact square for square contacts and the half width of the rectangle for linear contacts)

χ is the thermal diffusivity [m²/s], i.e. $\chi = K/\rho\sigma$ where:

K is the thermal conductivity [W/mK];

ρ is the density [kg/m³];

σ is the specific heat [J/kgK].

The Peclet number is an indicator of the heat penetration into the bulk of the contacting solid, i.e. it describes whether there is sufficient time for the surface temperature distribution of the contact to diffuse into the stationary solid. A higher Peclet number indicates a higher surface velocity for constant material characteristics.

Since all frictional heat is generated in the contact, the contact is modelled and treated as a heat source in the analysis. Flash temperature equations are derived, based on the assumption that the contact area moves with some velocity 'U' over the flat surface of a body 'B' as shown in Figure B.7.

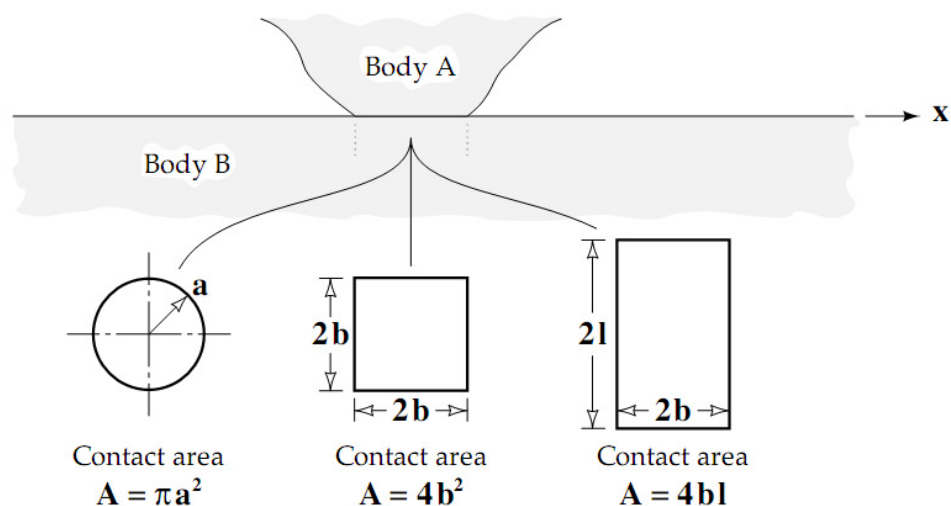


Figure B.7. Different contact areas for surface contact temperature calculations

(Stachowiak 2005; Stachowiak and Batchelor 2001)

The heat transfer effects vary with the Peclet number as shown schematically in Figure B.8. The following velocity ranges, defined by their Peclet number, are considered in flash temperature analysis:

$L < 0.1$ one surface moves very slowly with respect to the other. There is enough time for the temperature distribution of the contact to be established in the stationary body. In this case, the situation closely approximates to steady state conduction.

$0.1 < L < 5$ intermediate region. One surface moves faster with respect to the other and a slowly moving heat source model is assumed, $L > 5$ one surface moves fast with respect to the other and is modelled by a fast moving heat source. There is insufficient time for the temperature distribution of the contact to be established in the stationary body and the equations of linear heat diffusion normal to the surface apply. The depth to which the heat penetrates into the stationary body is very small compared to the contact dimensions.

Flash temperature equations are given in terms of the heat supply over the contact area, the velocity and the thermal properties of the material. They are derived based on the assumption that the proportion of the total heat flowing into each contacting body is such that the average temperature over the contact area is the same for both bodies. The flash temperature equations were developed by Blok and Jaeger for linear and square contacts and by Archard for circular contacts .

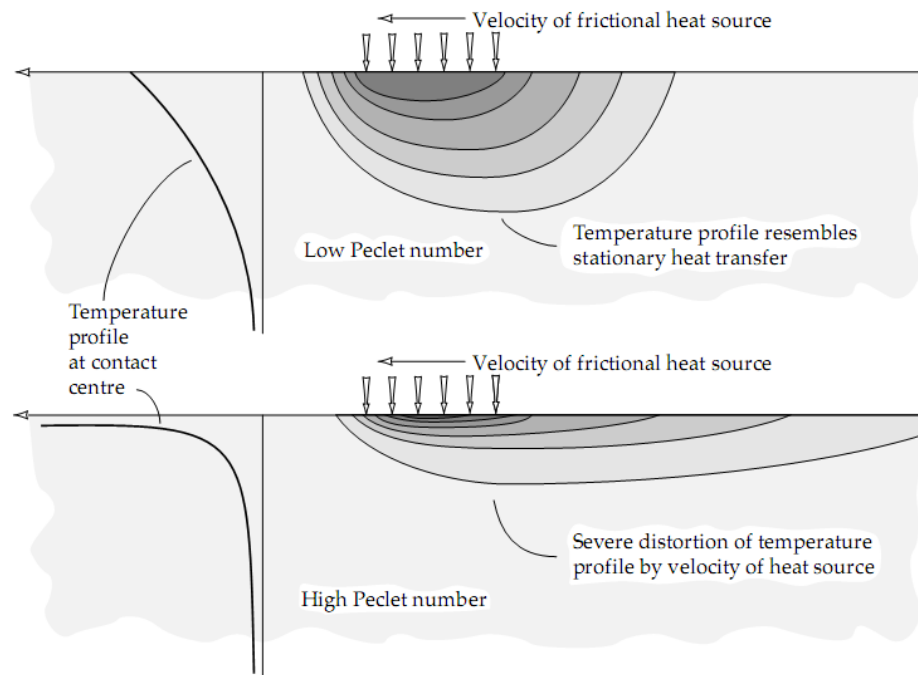


Figure B.8. Dimensionless Peclet number

(Stachowiak and Batchelor 2001)

B.2.1.1 Flash Temperature in Circular Contacts

In developing the flash temperature formulae for a circular contact, it was assumed that the portion of the surface in contact was of approximately equal height to the contact radius 'a'. The temperature at the distance 'a' from the surface was considered as a bulk temperature 'T_b' of the body. This can be visualised as a cylinder of height equal to its radius with one end in contact and the other end maintained at the bulk temperature of the body.

Average and maximum flash temperature formulae for circular contacts and various velocity ranges are summarised in Table B.6. The average flash temperature corresponds to the steady-state component of flash temperature, while the maximum value includes the transient component. The maximum flash temperature occurs when the maximum load is concentrated at the smallest possible area, i.e. when the load is carried by a plastically deformed contact.

Peclet Number	Average flash temperature	Maximum flash temperature
$L < 0.1$	$T_{f_a} = 0.5NL = \frac{\pi qa}{4K}$ or $T_{f_a} = 0.25 \frac{\mu W U_A - U_B }{Ka}$	$T_{f_{max}} = 0.25N'L'$ or $T_{f_{max}} = 0.222 \frac{\mu U}{K} (p_y W)^{0.5}$
$0.1 < L < 5$	$T_{f_a} = 0.5\alpha NL = \alpha \frac{\pi qa}{4K}$ or $T_{f_a} = 0.25\alpha \frac{\mu W U_A - U_B }{Ka}$ <p>α ranges from 0.85 at $L = 0.1$ to 0.35 at $L = 5$</p>	$T_{f_{max}} = 0.25\beta N'L'$ or $T_{f_{max}} = 0.222\beta \frac{\mu U}{K} (p_y W)^{0.5}$ <p>β ranges from 0.95 at $L = 0.1$ to 0.50 at $L = 5$</p>
$L > 5$	$T_{f_a} = 0.435NL^{0.5} = \frac{\pi q}{3.251K} \left(\frac{\chi a}{U} \right)^{0.5}$ or $T_{f_a} = 0.308 \frac{\mu W U_A - U_B }{Ka} \left(\frac{\chi}{Ua} \right)^{0.5}$	$T_{f_{max}} = 0.435\gamma N'L^{0.5}$ or $T_{f_{max}} = 0.726\gamma\mu p_y \left(\frac{U}{K\rho\sigma} \sqrt{\frac{W}{p_y}} \right)^{0.5}$ <p>or in general $T_{f_{max}} = 1.64 \cdot T_{f_a}$ γ ranges from 0.72 at $L = 5$ to 0.92 at $L = 100$. For $L > 100$, $\gamma = 1$ or in general</p>
Table B.6. Average and maximum flash temperature formulae for circular contacts (Stachowiak and Batchelor 2001)		

where:

T_{f_a} is the average flash temperature [°C];

$T_{f_{max}}$ is the maximum flash temperature [°C];

μ is the coefficient of friction;

W is the normal load [N];

p_y is the flow or yield stress of the material [Pa];

U_A, U_B are the surface velocities of solid 'A' and solid 'B' respectively [m/s];

U is the velocity of solid 'A' or 'B';

a is the radius of the contact circle [m] (Figure 7.28);

χ is the thermal diffusivity, $\chi = \frac{K}{\rho\sigma}$, [m²/s];

K is the thermal conductivity [W/mK];

ρ is the density [kg/m³];

σ is the specific heat [J/kgK];

α, β, γ are coefficients;

L is the Peclet number; $L = \frac{Ua}{2\chi} = \frac{Ua\rho\sigma}{2K}$

N is the variable [°C], defined as: $N = \pi \frac{q}{\rho} \sigma U$ where:

$$q = \frac{Q}{\pi a^2} = \mu W \frac{|U_A - U_B|}{\pi a^2}$$

is the rate of heat supply per unit area (circular) [W/m²];

L' is the variable defined as: $L' = \frac{U}{2\chi} \left(\frac{W}{\pi p_y} \right)^{0.5}$

N' is the variable [°C], defined as: $N' = \pi \mu \frac{p_y}{\rho} \sigma$

Average and maximum flash temperature formulae for rectangular and line contacts can be found in (Stachowiak and Batchelor 2001)

B.2.1.2 True Flash Temperature Rise

The heat generated in frictional contacts is divided between the contacting solids. The proportion of the total heat flowing to each body is determined on the basis that the average surface temperature is the same for both bodies. A simple way of estimating the true temperature rise in the contact is to assume that initially all the heat generated is supplied to body 'A'. The appropriate flash temperature equation for a given speed and contact geometry conditions is then selected and a flash temperature ' T_{fA} ' calculated. The next step in the calculation procedure is to assume that all the heat generated is transferred to body 'B'. The appropriate flash temperature equation for this second model is then selected and the corresponding flash temperature ' T_{fB} ' calculated. The true flash temperature rise must be the same for both solids in contact and is calculated from (Stachowiak and Batchelor 2001):

$$\frac{1}{T_f} = \frac{1}{T_{fA}} + \frac{1}{T_{fB}} \quad (\text{B-9})$$

Equation (B-9). True flash temperature rise

C Maintenance strategies

Maintenance is divided into two classes: Unplanned maintenance which is a reactionary measure applied to something that has already failed and planned maintenance which aims to avoid faults and machine failures.

Unplanned maintenance is also called **Run to Failure Maintenance (RTF)**. This is the oldest type of maintenance. RTF consists of repair, replacement, or restorative action performed on a machine or a facility after the occurrence of a failure in order to bring this machine or facility to at least its minimum acceptable condition. It is subdivided into two types (Al-Shayea 2010).

- Emergency maintenance: carried out as fast as possible in order to bring a failed machine or facility to a safe and operationally efficient condition.
- Breakdown maintenance: performed after the occurrence of an advanced considered failure for which advanced provision has been made in the form of repair method, spares, materials, labour and equipment.

RTF can be useful in the following situations:

- The failure of a component in a system is unpredictable.
- The cost of performing run to failure maintenance activities is lower than performing other activities or other types of maintenance.
- The equipment failure priority is too low to include the activities of preventing it within the planned maintenance budget.
- This type of maintenance has serious disadvantages:
- Its activities are expensive in terms of both direct and indirect cost.
- Using this type of maintenance, the occurrence of a failure in a component can cause failures in other components in the same equipment, which leads to low production availability.

- Its activities are very difficult to plan and schedule in advance.

Planned maintenance is any variety of scheduled maintenance to an object or item of equipment to ensure that an item of equipment is operating correctly to avoid any unscheduled breakdown and downtime. It is subdivided into four types of maintenance: Preventive, Predictive, Improvement and Corrective (Al-Shayea 2010).

Preventive Maintenance (PM) is a set of activities that are performed on plant equipment, machinery and systems before the occurrence of a failure in order to protect them and to prevent or eliminate any degradation in their operating conditions. The British Standard 3811:1993 Glossary of Terms defined preventive maintenance as: “the maintenance carried out at predetermined intervals or according to prescribed criteria and intended to reduce the probability of failure or the degradation of the functioning and the effects limited”. Although it satisfies most maintenance objectives there are factors that affect the efficiency of this type of maintenance:

1. The need for an adequate number of staff in the maintenance department in order to perform this type of maintenance.
2. The right choice of production equipment and machinery that is suitable for the working environment and that can tolerate the workload of this environment.
3. The required staff qualifications and skills, which can be gained through training.
4. The support and commitment from executive management to the PM programme.
5. The proper planning and scheduling of the PM programme.
6. The ability to apply the PM programme properly.

Summarising, this type of maintenance is good for those machines and facilities where failure would cause serious production losses. Its aim is to maintain machines and facilities in such a condition that breakdowns and emergency repairs are minimised. Its activities include replacements, adjustments, major overhauls, inspections and regular lubrications.

Preventive maintenance can be further subdivided into different kinds according to the nature of activities:

- *Routine maintenance* which includes those maintenance activities that are repetitive and periodic in nature such as lubrication, cleaning and minor adjustments.
- *Running maintenance* which includes those maintenance activities that are carried out while the machine or equipment is running. They represent those activities that are performed before the actual preventive maintenance activities take place.
- *Opportunity maintenance* which is a set of maintenance activities that are performed on a machine or a facility when an unplanned opportunity exists during the period of performing planned maintenance activities to other machines or facilities.
- *Window maintenance* which is a set of activities that are carried out when a machine or equipment is not required for a definite period of time.
- *Shutdown preventive maintenance*, which is a set of preventive maintenance activities that are carried out when the production line is in total stoppage situation.

Corrective maintenance corresponds to actions such as repair, replacement or restore which will be carried out after the occurrence of a failure in order to eliminate the source of the failure or reduce the frequency of its occurrence. In the British Standard 3811:1993 Glossary of Terms, corrective maintenance is defined as: “the maintenance carried out after recognition and intended to put an item into a state in which it can perform a required function”.

This type of maintenance is subdivided into three types:

- *Remedial maintenance*, which is a set of activities that are performed to eliminate the source of failure without interrupting the continuity of the production process. The way to carry out this type of corrective maintenance is by taking the item to be corrected out of the production line and replacing it with a reconditioned item or transferring its workload to its redundancy.

- *Deferred maintenance*, which is a set of corrective maintenance activities that are not immediately initiated after the occurrence of a failure but are delayed in such a way that will not affect the production process.
- *Shutdown corrective maintenance*, which is a set of corrective maintenance activities that are performed when the production line is in a total stoppage situation.

The main objectives of corrective maintenance are the maximisation of the effectiveness of all critical plant systems, the elimination of breakdowns, the elimination of unnecessary repair and the reduction of the deviations from optimum operating conditions.

The difference between corrective maintenance and preventive maintenance is that for corrective maintenance, the failure should occur before any corrective action is taken.

Corrective maintenance is different from run to failure maintenance in that its activities are planned and regularly implemented to keep plant's machines and equipment in optimum operating condition.

The way to perform corrective maintenance activities is by conducting four important steps:

1. Fault detection.
2. Fault isolation.
3. Fault elimination.
4. Verification of fault elimination.

In the fault elimination step several actions could be taken such as adjusting, aligning, calibrating, reworking, removing, replacing or renovation.

Corrective maintenance has several prerequisites in order to be carried out effectively:

1. Accurate identification of incipient problems.

2. Effective planning which depends on the skills of the planners, the availability of a well-developed maintenance database including standard time to repair, complete repair procedures and the required labour skills, specific tools, parts and equipment.
3. Proper repair procedures.
4. Adequate time to repair.
5. Verification of repair.

Improvement maintenance aims to reduce or eliminate entirely the need for maintenance.

This type of maintenance is subdivided into three types as follows:

1. *Design-out maintenance* which is a set of activities that are used to eliminate the cause of maintenance, simplify maintenance tasks or raise machine performance from the maintenance point of view by redesigning those machines and facilities which are vulnerable to frequent occurrence of failure and their long term repair or replacement cost is very expensive.

2. *Engineering services* which includes construction and construction modification, removal and installation and rearrangement of facilities.

3. *Shutdown improvement maintenance*, which is a set of improvement maintenance activities that are performed while the production line is in a complete stoppage situation.

Predictive maintenance is a set of activities that detect changes in the physical condition of equipment (signs of failure) in order to carry out the appropriate maintenance work for maximising the service life of equipment without increasing the risk of failure.

It is classified into two kinds according to the methods of detecting the signs of failure:

- *Condition-based predictive maintenance* **depends on** continuous or periodic condition monitoring equipment to detect the signs of failure.

-
- *Statistical-based predictive maintenance* **depends on** statistical data from the meticulous recording of the stoppages of the in-plant items and components in order to develop models for predicting failures.

The drawback of predictive maintenance is that it depends heavily on information and the correct interpretation of the information.

Some researchers classified predictive maintenance as a type of preventive maintenance.

The **main difference** between preventive maintenance and predictive maintenance is that *predictive maintenance monitors the condition of machines or equipment to determine the actual mean time to failure* whereas *preventive maintenance depends on industrial average life statistics* (Al-Shayea 2010).

D Topographic measurement considerations for sphere samples with optical profilometers

Optical profilometers achieve the greatest precision in height measurement when there is no curvature or tilt in the samples (flat surfaces). The dimensions of the acquisition domain when using spherical samples should be taken bearing in mind that the curvature of the sphere introduces a tilt that should be within the limits of the measurement range of the interferometer in order to ensure reliable data.

The measuring process of balls in interferometers requires the location of the ball in a special holder where the wear scar could be aligned and placed at an adequate working distance under the objective. Special attention needs to be paid to ensure the horizontal axes (x,y) of the apparatus are properly calibrated before performing any topographic measurement. If the reference surface is a sphere, the horizontal errors should be kept as low as possible to have utmost accuracy when processing measurements. The next section discusses in detail the errors depending on the type of algorithm used.

D.1 The effect of horizontal errors in ball measurements

Horizontal errors in interferometers should be carefully considered in the case of spherical surfaces. Small horizontal errors could result after removing the reference sphere, in an error distribution of heights which alters the value of the wear volume predictions. A realistic way of assessing these errors consists of establishing a comparison between the typical types of sphere fitting algorithm under ideal conditions. For the numerical test conditions, a convex section of the sphere of a radius of 6.35mm is considered which is centred in (0,0,0) and corresponds to a hypothetical measurement within $x,y \in (-0.5,0.5)$ mm. The range of measurement in the z axis for the mentioned x,y is 39.49 μ m, which is within the normal measurement range of the apparatus.

This sphere is distorted scaling the mesh of points in x and y directions using a scale factor within the range of -5% to 5%. The differences in the z axis (heights) after removing the reference sphere are the indicator used to compare the

performance between several sphere fitting algorithms. Two algorithms were considered which are typically implemented in commercial software packages. The first algorithm is constrained providing only the centre as the radius is fixed to 6.35mm while the second one is unconstrained providing the radius and the centre of the sphere. The first one is a non-linear least squares algorithm which minimises the orthogonal distances using the radius as a constraint and the second one, is a fast linear least squares algorithm which minimises the orthogonal distances from the data points to the reference sphere using the pseudoinverse.

The results of differences of heights (e_v) after subtracting the original sphere to the calculated reference sphere are shown versus the horizontal errors (e_h) for each type of algorithm. Figures D.1-2 are for the first algorithm and Figure D.3 is for the second algorithm. Although this subtraction is chosen because it represents the real error, the measurement apparatus subtracts the calculated sphere from the distorted measured sphere, therefore the display surface is always distorted. If the highest point of the ball is located in the middle of the acquisition domain, the application of the constrained algorithm leads to a uniform error distribution of heights within the domain, Figure D.1. This property is exploited for the calibration of the interferometer. If the highest point is not located in the centre, the height distributions are no longer uniform and correspond to a plane within the acquisition domain, Figure D.2.

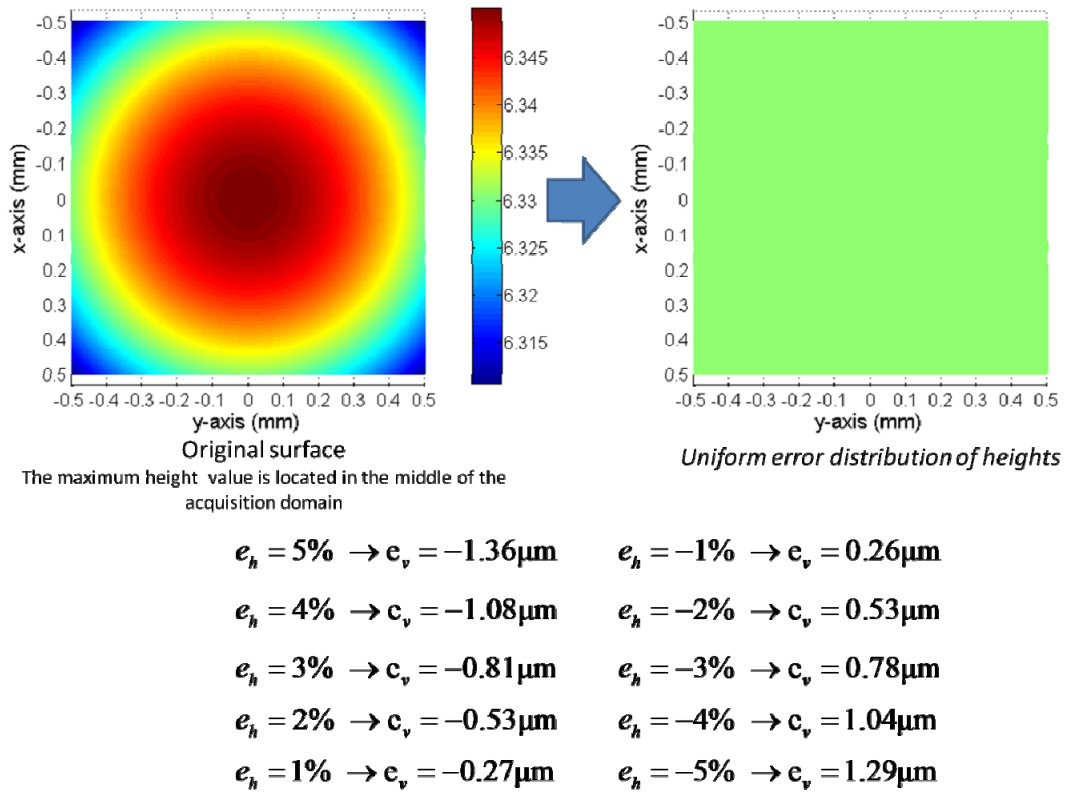


Figure D.1. Differences in heights (e_v) versus the horizontal error (e_h) for the constrained sphere fitting algorithm, $R=6.35\text{mm}$

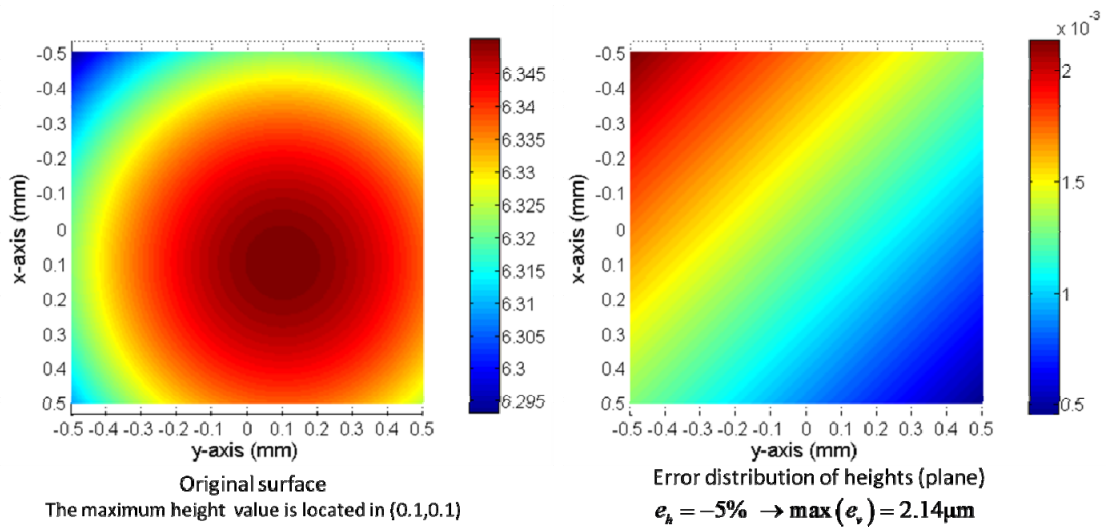


Figure D.2. Differences in heights (e_v) versus the horizontal error (e_h) for the constrained sphere fitting algorithm when the peak value is not located in the centre

The second algorithm leads to a non-uniform distribution of errors of height which increases in the radial direction, Figure D.3. Furthermore, the height errors are always higher than the constrained algorithm. As a result, the first algorithm should be finally chosen because it performs better than the others.

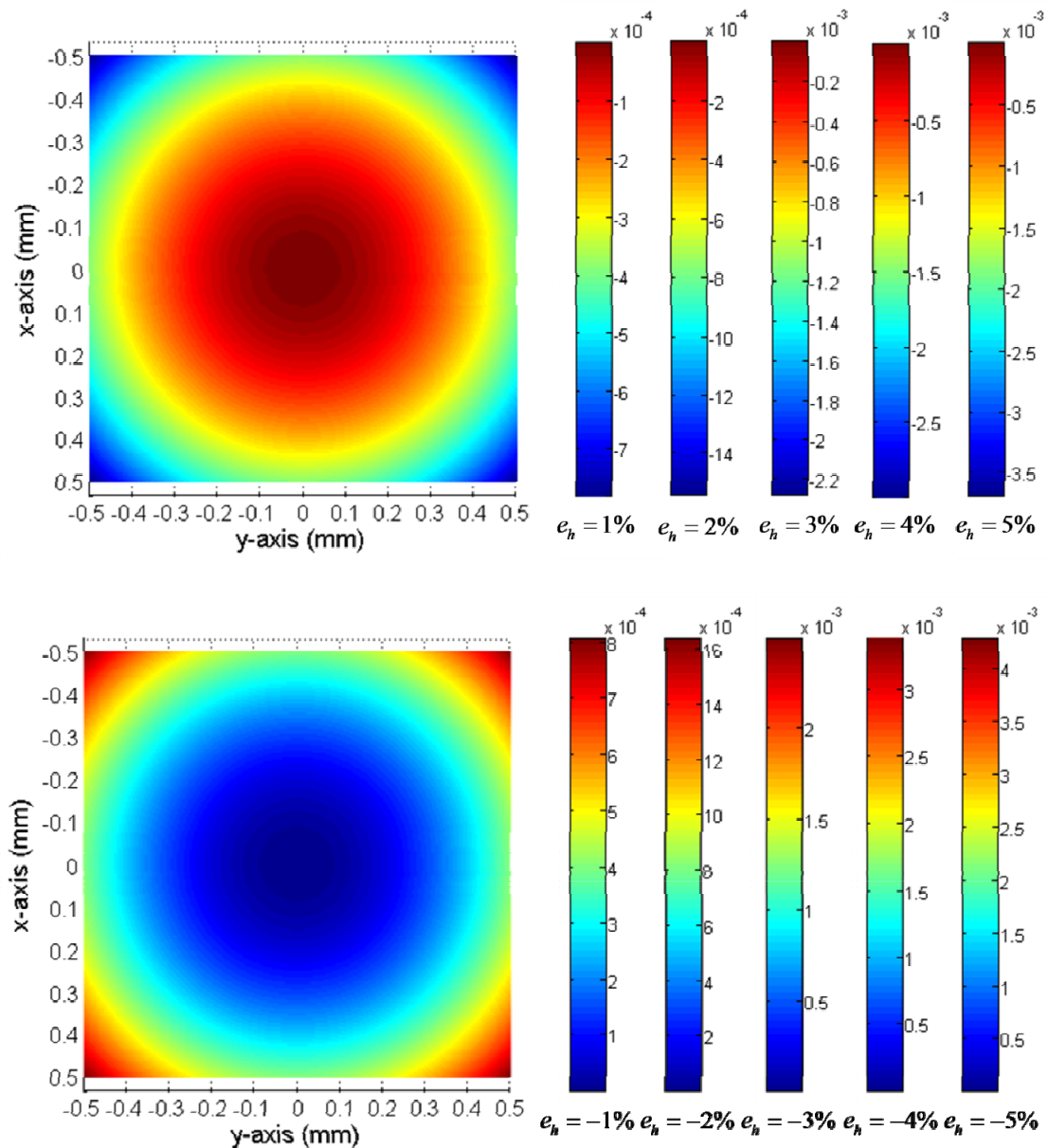


Figure D.3. Errors in z (heights), after subtracting the sphere using the unconstrained algorithm

As shown above, a way of minimising the horizontal errors is to perform measurements on calibrated apparatus using the most adequate sphere fitting algorithm. In the market, metrology manufacturers provide horizontal calibration samples with known dimensions that are more expensive than a steel ball bearing of

grade 5. A fast and cheap procedure for the horizontal calibration of the interferometer for ball measurements consists of performing several measurements of an unworn steel ball of known diameter and grade 5. The highest point of the ball should be placed in the centre of the acquired data to ensure the final height errors are uniformly distributed within the acquisition domain, Figure D.1. On each measurement, the horizontal magnification scale is adjusted until the processed data (after removing the sphere with the fixed radius) is within the range of the roughness of the original steel ball. It should be noted that the horizontal calibration is unique and it depends on the zoom and objective used. A change in the mechanical zoom of the interferometer may need a check in the horizontal calibration.

A graphical example of the calibration procedure is shown in Figure D.4, which shows three real measurements of the same surface with different horizontal calibration factors. The roughness of the original ball is of the order of 3nm, which is shown in Figure D.4(b) when the horizontal error is almost zero. The horizontal magnification of Figure D.4(a) was too high during measurement while in Figure D.4(c) it was too small. The interferometer is considered properly-calibrated when the processed measurement is within the range of the original roughness, Figure D.4(b).

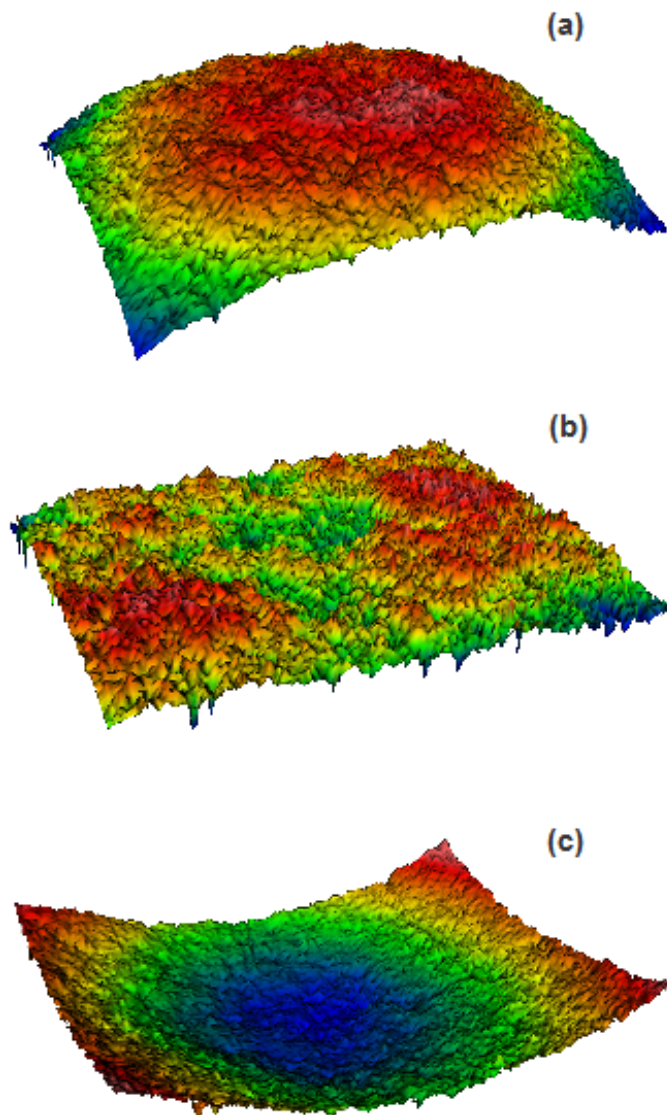


Figure D.4. Errors in z (heights), calibration procedure