

الجمهورية الجزائرية الديمقراطية الشعبية

République Algérienne Démocratique Et Populaire

وزارة التعليم العالي والبحث العلمي

Ministère de L'Enseignement Supérieur et de La Recherche Scientifique

جامعة فرحات عباس- سطيف 1

Université Ferhat Abbas - Sétif 1

THÈSE

Présentée à l'Institut d'Optique et de Mécanique de Précision

pour l'obtention du Diplôme de

DOCTORAT 3^{ème} Cycle LMD

Domaine : Sciences et Techniques

Filière : Optique et mécanique de Précision

Spécialité : Mécanique Appliquée.

Par

BENMAHDI Dalila

THÈME

*Contribution à l'étude du diagnostic des défauts mécaniques par
classification non supervisée*

Soutenue, le: 06/01/2020

Devant le jury composé de:

Président du Jury	HARRAG Abdelghani	Prof.	UFA Sétif 1
Directeur de thèse	FELKAOUI Ahmed	Prof.	UFA Sétif1
Co-directeur de thèse	RASOLOFONDRAIBE Lanto	MC-PHD	URCA Reims
Examineur	LADJAL Mohamed	MCA	UMB M'sila
Examineur	MAHGOUN Hafidha	MCA	UFA Sétif 1

Remerciements

Je tiens à remercier tout d'abord mon directeur de thèse, le professeur Ahmed FELKAOUI, pour m'avoir accueilli au sein du laboratoire LMPA, encadré et soutenu tout au long des années de mon doctorat à l'institut d'optique et mécanique de Précision. Je le remercie de m'avoir apporté toute son expérience et ses compétences sans lesquelles ce travail n'aurait pas pu aboutir.

Je tiens à exprimer ma gratitude et reconnaissance envers mon Co-directeur de thèse, Lanto RASOLOFONDRAIBE, maître de conférence PhD à l'université Champagne Ardenne Reims France. Merci de m'avoir accepté au sein du laboratoire Crestic. Merci d'avoir m'aidé à traduire mes idées, merci pour ta compréhension et ta disponibilité et le temps que tu as consacré pour que je puisse accomplir ma recherche et ma thèse. Merci d'avoir cru en moi, Merci pour tes encouragements, ta bienveillance et surtout pour les conditions que tu as assuré pour mon bien être,..., Merci pour tout.

J'adresse mes sincères remerciements à Monsieur Xavier CHIEMENTIN, maître de conférences PHD à l'université Champagne Ardenne Reims France, de m'avoir accueilli à son laboratoire GRESPI/MAN et qui grâce à son soutien moral et précieux ce projet est rendu possible. Grâce à ton esprit tu m'as laissé travailler à l'aise, et pour ton humilité tu m'as considéré comme collègue. Merci pour le temps que tu as consacré malgré ton programme chargé pour que je puisse bien comprendre le domaine de la maintenance industriel ainsi que pour toutes les corrections et conseils. Merci d'avoir m'accepter comme ta doctorante. Et je souhaite continuer mes recherches avec vous deux Merci pour tout.

Je voudrais adresser mes sincères remerciements à monsieur Abdelghani HARRAG, Professeur à l'Université Sétif 1, d'avoir accepté de présider l'honorable jury de cette thèse.

Je remercie vivement Monsieur Mohamed LADJAL, maître de Conférences à l'Université Mohamed Boudiaf UMB M'sila, et à madame Hafida MAHGOUN, maître de Conférences à l'institut d'optique et de mécanique de précision Sétif 1, de m'avoir fait l'honneur d'accepter d'être les Examineurs de cette thèse.

Un merci spécial au professeur Abdellatif HAMOUDA que je considère mon exemple, grâce à lui j'ai pu effectuer mon séjour. Je vous remercie pour tous vos conseils que vous m'avez donnés tout au long de mes années universitaires.

Je remercie tous les membres et Ex-membres de laboratoire Crestic et Grespi, en particulier Monsieur *Samuel* CREQUY pour l'aide qu'il m'a apportée dans la résolution de certains problèmes techniques et scientifiques rencontrés dans mes travaux de thèse et aussi pour les essais qu'il a effectués.

Je remercie également Tarek NOUIOUA de l'université de Tébessa pour ces conseils et son aide ainsi que pour ces corrections, merci.

Je remercie tous les membres du LMPA et en particulier Mr. Semcheddine FEDALA et Mr. Ridha ZIANI pour leurs encouragements et leurs aides, ainsi que tous mes amis.

Enfin,

Les mots les plus simples étant les plus forts, j'adresse toute mon affection à ma famille, en particulier à mes parents qui m'ont soutenus tout au long de ma thèse moralement et physiquement et d'avoir toujours été fiers de moi, merci à mes frères et sœurs merci pour votre amour et vos encouragements. Merci d'avoir fait de moi ce que je suis aujourd'hui.

A mon mari Salaheddine, tu étais toujours là pour moi, tu m'as écouté quand je te raconte mes soucis, tu m'as remonté le moral quand je suis fatiguée, tu m'as encouragé quand j'ai baissé les bras, et tu m'as consolé quand j'ai subi un échec. Tu as fait tout ça pour moi. Merci beaucoup.

A mes enfants, Mariane et Mohamed Ali, vous êtes ma plus belle histoire d'amour. Chaque fois que je pose le regard sur vous, je sais pourquoi j'existe...

Je vous adore.

Dalila

Résumé

La complexité des installations actuelles nécessite d'avoir recours à des techniques de surveillance sophistiquées et efficaces. La technique la plus utilisée reste l'analyse vibratoire. En effet l'état de chaque composant peut être caractérisé par un ou plusieurs indicateurs de défaut issus de l'analyse vibratoire. Le suivi de ces indicateurs permet de détecter le défaut. Cependant, l'évolution de ces indicateurs peut être influencée par d'autres paramètres. Cela peut provoquer des fausses alarmes et remettre en question la fiabilité du diagnostic. Pour cette raison la communauté scientifique s'intéresse largement à combiner ces méthodes avec des méthodes d'apprentissage pour renforcer et améliorer le diagnostic. Ces méthodes permettent de rendre automatique et fiable la détection et le diagnostic. Cette thèse a pour objectif de combiner l'analyse vibratoire avec la méthode de reconnaissance des formes afin d'une part d'améliorer la détection de défaut de roulement et d'autre part de mieux suivre l'évolution de la dégradation dans le temps. Pour cela nous avons proposé une nouvelle approche de reconnaissance des formes non supervisée dans le cadre de la détection et le diagnostic dynamique des défauts de roulements, RT-OPTICS (Real Time Ordering Points to Identify the Clustering Structure). Cette approche cible deux étapes de l'endommagement : la détection de défaut par classification et le suivi de la nouvelle classe qui représente l'état dégradé du roulement. Ces deux étapes sont réalisées par une méthode bidimensionnelle impliquant des indicateurs scalaires : le Kurtosis et la valeur efficace. Ces deux indicateurs apportent des informations complémentaires quant à la présence de défauts dans le roulement. La première étape déploie la méthode de classification non supervisée OPTICS (Ordering Points To Identify Clustering Structure) pour détecter l'apparition de défaut sur la bague intérieure et/ou extérieure de roulement. L'étape suivante consiste à suivre l'évolution de la dégradation en utilisant cinq paramètres de la nouvelle classe : le saut du centre de classe en terme de distance euclidienne, la densité, la distance interclasse maximale, le contour et la surface de cette classe. Après une mise en évidence des possibilités de la méthode proposée sur des signaux simulés dont les variations de paramètres sont testées, cette approche fut validée expérimentalement sur un banc d'essais composés de roulements N.206.E.G15. Des variations de charge et/ou de vitesses sont appliquées. Une étude comparative est réalisée entre l'approche proposée et un suivi d'indicateurs scalaires et fréquentiels dans le temps a été également réalisé.

Table des matières

Introduction générale.....	2
1 Généralité sur la surveillance et le diagnostic industriel	8
1.1 Introduction.....	8
1.2 Différents niveaux de maintenance.....	8
1.2.1 Maintenance corrective	9
1.2.2 Maintenance préventive	9
1.3 L’analyse vibratoire en tant qu’outil de diagnostic.....	11
1.4 Image vibratoire des défauts de roulement	11
1.4.1 Constitution des roulements.....	11
1.4.2 Modes de dégradation des roulements	12
1.4.3 Fréquences caractéristiques de défauts	13
1.4.4 Indicateurs scalaires	16
1.4.4.1 Kurtosis.....	16
1.4.4.2 Skewness	18
1.4.4.3 Facteur crête	18
1.4.4.4 Facteur K	19
1.4.4.5 Analyse cepstrale	19
1.4.4.6 Analyse spectrale	20
1.4.5 Analyse d’enveloppe.....	22
1.4.6 Analyse temps-fréquence.....	23
1.4.7 Analyse temps-échelle:	24
1.4.8 Transformation en paquets d’ondelettes	26
1.4.9 Décomposition en mode empirique (EMD).....	27
Cyclo-stationnarité.....	29
1.5 Conclusion	30
2 Classification non supervisée dans le cadre de diagnostic des machines tournantes ...	33
2.1 Introduction.....	33
2.2 Systèmes de clustering.....	34
2.2.1 Objectif des méthodes de classification.....	34
2.2.1.1 Clusters	34

2.2.1.2	Propriétés d'un cluster	35
2.2.1.3	Différentes étapes d'un système de clustering	35
2.2.1.3.1	Représentation des données : extraction et réduction de dimension.....	36
2.3	Classification.....	37
2.3.1	Propriétés des techniques de clustering	38
2.3.2	Evaluation de la qualité d'une classification non supervisée.....	39
2.3.3	Différentes représentation des résultats de la classification	44
2.3.3.1	Classification dure, douce et floue	44
2.3.3.2	Classification hiérarchique	46
2.4	Différentes méthodes de Partitionnement ou de clustering	47
2.4.1	Méthodes basées sur une distance.....	47
2.4.1.1	Méthodes basées sur des prototypes	47
2.4.1.2	Les K plus proches Voisins (K- Nearest Neighbors)	50
2.4.2	Méthodes neuronales	50
2.4.3	Approches basées sur les Séparateurs à Vaste Marge (SVMs).....	53
2.4.4	Méthodes basées sur la densité	55
2.4.4.1	Algorithme DBSCAN.....	55
2.4.4.2	Algorithme OPTICS	59
2.4.5	Méthodes basées sur une grille	65
2.4.6	Méthodes probabilistes	65
2.4.7	Méthodes hiérarchiques	66
2.4.7.1	Classification ascendante hiérarchique.....	66
2.4.7.2	Arbre de décision.....	67
2.5	Conclusion	68
3	Contribution à l'étude de diagnostic de défauts par la méthode de classification dynamique	71
3.1	Introduction.....	71
3.2	Classification dynamique.....	71
3.2.1	Phénomènes dynamiques en environnement non stationnaire.....	71
3.2.2	Apparition de nouveautés dans un espace de classification.....	72
3.2.3	Evolution de la partition dynamique.....	74
3.2.4	Fusion d'informations mutuelles	76

3.2.5	Séparation d'informations exclusives	77
3.2.6	Suppression de connaissances obsolètes ou parasites.....	77
3.3	Système du diagnostic vibratoire	78
3.3.1	Extraction des caractéristiques.....	78
3.3.2	Détection de défaut par la méthode de classification dynamique.....	79
3.3.3	Adaptation de la méthode RT-OPTICS	79
3.3.3.1	Présentation de la méthode RT-OPTICS.....	80
3.3.3.2	Nouvelle méthode pour le calcul des clusters	81
3.3.3.3	Détection du défaut.....	83
3.3.4	Suivi des caractéristiques géométrique	84
3.3.4.1	Enveloppe convexe (contour).....	84
3.3.4.2	Densité.....	86
3.3.4.3	Saut du centre	86
3.3.4.4	Distance maximale d_{max}	87
3.3.4.5	La surface	87
3.3.5	Limitations de la classification par RT-OPTICS	87
3.4	DDBSCAN	88
3.5	Conclusion	88
4	Validation numérique et expérimentale de la méthode RT-OPTICS.....	90
4.1	Introduction.....	90
4.2	Validation numérique de la méthode RT-OPTICS	90
4.2.1	Création du modèle mathématique	90
4.2.2	Rapport signal sur bruit RSB	93
4.2.3	Validation numériques	94
4.2.3.1	Calcul d'indicateurs scalaires	96
4.2.3.2	Influence de la variation de l'amplitude	97
4.2.3.3	Variation de la vitesse.....	98
4.2.3.4	Variation du taux d'amortissement.....	99
4.2.4	Détection du défaut	99
4.2.4.1	Suivi de la classe C2.....	101
4.3	Validation expérimentale et étude comparative.....	102

4.3.1	Description du banc expérimental	102
4.3.2	Création des défauts sur la bague extérieure.....	103
4.3.3	Calcul des indicateurs	104
4.3.3.1	Indicateurs scalaires.....	105
4.3.3.1.1	Effet de la vitesse de rotation sur les indicateurs scalaires	107
4.3.3.1.2	Effet de la charge sur les indicateurs scalaires.....	108
4.3.3.2	Calcul des indicateurs fréquentiels	110
4.3.3.2.1	Effet de la vitesse de rotation sur les indicateurs fréquentiels	112
4.3.3.2.2	Effet de la charge sur les indicateurs fréquentiels.....	113
4.3.3.3	Analyse par les ondelettes	114
4.3.3.3.1	Effet de la vitesse	116
4.3.3.3.2	Effet de la charge.....	116
4.3.4	Validation de la méthode proposée.....	117
4.3.4.1	Cas d'une vitesse et d'une charge constante	117
4.3.4.1.1	Détection du défaut avec la méthode adaptée	118
4.3.4.1.2	Suivi de la classe de roulement défectueux.....	119
4.3.4.2	Cas d'une vitesse variable	120
4.3.4.2.1	Détection du défaut avec la méthode adaptée	121
4.3.4.2.2	Suivi de la classe de roulement défectueux.....	122
4.3.4.3	Cas d'une charge variable (vitesse constante).....	124
4.3.4.3.1	Détection du défaut avec la méthode adaptée	124
4.3.4.3.2	Suivi de la classe de roulement défectueux.....	125
4.3.4.4	Cas d'une vitesse et d'une charge variables	127
4.3.4.4.1	Détection du défaut avec la méthode adaptée	127
4.3.4.4.2	Suivi de la classe de roulement défectueux.....	128
4.3.5	Comparaison entre la méthode RT-OPTICS et DDBSCAN	129
4.4	Conclusion	131
	Conclusion générale & perspectives.....	134
	Références bibliographiques	137

Liste des tableaux

Tableau 1-1: Modes défailants sur les machines tournantes	15
Tableau 4-1: Dimensions de roulement SKF6206.....	95
Tableau 4-2 : Comparaison entre la méthode classique et la méthode RT-OPTICS.....	101
Tableau 4-3: Dimensions des défauts créés sur la bague extérieure de roulement.....	104
Tableau 4-4: Comparaison de la détection de défaut entre DBSCAN et RT-OPTICS. Oui: détection, Non: non détecté.....	130

Liste des figures

Figure 1-1: Les stratégies de maintenance (Larose, 2012)	9
Figure 1-2: Géométrie d'un roulement (Morel, 2005).....	14
Figure 1-3: Vue éclatée d'un roulement	14
Figure 1-4: Représentation schématique de la distribution d'une variable aléatoire correspondant à trois valeurs du Kurtosis	17
Figure 1-5: Arbre de décomposition en paquets d'ondelettes pour 3 trois niveaux de décomposition (Kerroumi, 2016).....	27
Figure 2-1: (a) Classes réelles et (b) classe instrumentale (Blansché, 2006)	34
Figure 2-2: Structure générale de la méthode de reconnaissance de formes	36
Figure 2-3: Classification par partitionnement (Blansché, 2006)	44
Figure 2-4: Classification hiérarchique (Blansché, 2006).	46
Figure 2-5: Algorithme des k-moyennes (Fouchal, 2011).....	48
Figure 2-6: L'architecture d'un réseau de neurones (Kerroumi, 2016)	50
Figure 2-7: Schéma des hyperplans avec une forme sphérique pour trois classes (Wang et al., 2012)	54
Figure 2-8: Algorithme DBSCAN (Ester et al., 1996)	56
Figure 2-9: Notions de points dense-accessibles et de dense-connectivité (Ester et al., 1996) ..	57
Figure 2-10: Noyau(q) et point limite (p) ; p est dd-a à partir de q ; q n'est pas dd-a à partir de p (Ester et al., 1996).....	58
Figure 2-11: p est dd-a à partir de q ; q n'est pas dd-a à partir de p (Ester et al., 1996).....	59
Figure 2-12: Création des clusters par DBSCAN ; densité-connectée (Ester et al., 1996).....	59
Figure 2-13: densité-accessibilité et connectée (Ankerst et al., 1999).....	61
Figure 2-14: Distance-noyau(o), distance-accessibilité.....	62
Figure 2-15: L'algorithme OPTICS (Ankerst et al., 1999).....	63
Figure 2-16: La procédure EtendreClusterOrdre (Ankerst et al., 1999).	64
Figure 2-17: La méthode OrderSeeds ::mise-à-jour (Ankerst et al., 1999).	64
Figure 3-1: (a) Apparition de classe dans un espace de classification « vide », (b) Apparition d'une nouvelle classe après un saut brut et (c) Apparition d'une nouvelle classe après une dérive rapide (Amadou, 2002)	73
Figure 3-2: (a) Développement de classe, (b) Modification quelconque de classe et (c) Evolution de classe suite à une dérive de données non-stationnaires (Amadou, 2010)	75

Figure 3-3: (a) Fusion de classes partageant des informations mutuelles (b) Scission de classes après séparation des informations partagées et (c) Elimination de classes obsolètes et parasites.....	78
Figure 3-4: Extraction d'un cluster (Ankerst et al., 1999).....	82
Figure 3-5: Algorithme principal de RT-OPTICS (Benmahdi et al. 2017)	82
Figure 3-6: Organigramme de l'algorithme principal.....	83
Figure 3-7: Algorithme de l'enveloppe convexe (contour)	85
Figure 3-8: les différentes étapes de l'algorithme de l'enveloppe convexe	86
Figure 4-1: Création de la base de données	94
Figure 4-2: Signaux simulés (a) bague intérieure ; (b) bague extérieure, pour une amplitude $A=2$	96
Figure 4-3: Évolution de la valeur efficace en fonction du kurtosis lors d'une augmentation de l'amplitude pour un défaut de bague extérieure.....	97
Figure 4-4: Signaux temporels pour (a) $A = 0$ (RSB = -1.60dB), (b) $A = 4 \cdot 10^{-2}$ (RSB = -0.23dB), (c) $A = 5 \cdot 10^{-2}$ (RSB = 2.45dB)	98
Figure 4-5: Évolution de RMS en fonction du kurtosis pour une amplitude $A=0.1$	98
Figure 4-6: Évolution de la valeur efficace en fonction du kurtosis pour une amplitude $A=0.1$	99
Figure 4-7: Dynamique de la détection de défaut par la méthode OPTICS. (a) roulement sain ($A = 0$). (b) roulement sain ($A = 3 \cdot 10^{-2}$, RSB = -3.28dB). (c) roulement défectueux ($A = 4 \cdot 10^{-2}$ et RSB = -0.23 dB).	100
Figure 4-8: (a) Evolution de centre de la classe C1, (b) Evolution de contour de la classe C2, (c) Evolution de la densité de la classe C2	101
Figure 4-9: (a) Banc expérimental. (b) Roulement N.206.E.G15. (c) Exemple de défaut créé	102
Figure 4-10: (a) défaut crée, (b) mesure de la taille de défaut	103
Figure 4-11: Signal temporel, défauts (sain, défaut 4 et défaut9).....	105
Figure 4-12: Evolution de la valeur RMS en fonction du kurtosis pour toutes les tailles de défauts et dans le cas de tous les paramètres (9 défauts \times 10 vitesses \times 6 charges \times 10 tranches = 5400points).....	106
Figure 4-13: Evolution des indicateurs en fonction de la taille de défauts, vitesse et charge constantes.....	106
Figure 4-14: influence de la vitesse de rotation sur l'évolution des indicateurs scalaires	107
Figure 4-15: Évolution de la valeur efficace en fonction du kurtosis pour une taille de défaut de 0.21 mm, charge= 100 daN et une vitesse variable.....	108
Figure 4-16: Influence de la charge sur l'évolution des indicateurs scalaires chaque point est une moyenne de dix valeurs)	109
Figure 4-17: Évolution de la valeur efficace en fonction du kurtosis pour taille de défaut de 0.21 mm, vitesse= 1405 tr/mn et une charge variable	110
Figure 4-18: (a) spectre de puissance du signal vibratoire pour une vitesse de 1430 rpm et (b) Les indicateurs fréquentiels calculés à partir de la puissance de l'enveloppe du spectre.....	110
Figure 4-19: (a) Spectre de roulement sans défaut pour une vitesse de 1405 rpm, (b) Spectre de roulement avec défaut ($R_a = 0.33 \mu\text{m}$) pour une vitesse de 1435 rpm	111
Figure 4-20: Les indicateurs calculés à partir du spectre d'enveloppe	112
Figure 4-21: Evolution des indicateurs fréquentiels en cas de présence de défaut, et en cas de différentes vitesses	113

Figure 4-22: Evolution des indicateurs fréquentiels en cas de présence de défaut et en cas de différentes charges	114
Figure 4-23: Procédure de l'extraction de ces deux indicateurs W_{rms} et P_{CWT}	115
Figure 4-24: Extractions des indicateurs : (a) W_{rms} et (b) P_{cwt} A partir de la transformée en ondelettes continues	115
Figure 4-25: Influence de la vitesse sur les indicateurs : (a) W_{rms} et (b) P_{cwt}	116
Figure 4-26: Influence de la charge sur les indicateurs : (a) W_{rms} et (b) W_{Pcwt}	117
Figure 4-27: Evolution de la valeur efficace RMS en fonction du kurtosis des 9 défauts....	118
Figure 4-28: (a), (b) dynamique de la détection de défaut et (c) ordre de la distance-accessibilité.....	119
Figure 4-29: Evolution des caractéristiques géométriques: (a) surface, (b) Distance maximale, (c) densité, (d) saut de centre.	120
Figure 4-30: Evolution de la valeur efficace RMS en fonction du kurtosis pour différentes vitesses et une charge constante.....	121
Figure 4-31: (a), (b) et (c); dynamique de la détection de défaut.	122
Figure 4-32: Evolution des caractéristiques géométriques: (a) surface, (b) Distance maximale, (c) densité, (d) saut de centre et (e) le contour.....	123
Figure 4-33: Evolution de la valeur efficace RMS en fonction du kurtosis pour différentes charges et une vitesse constante.....	124
Figure 4-34: (a), (b) et (c); dynamique de la détection de défaut.	125
Figure 4-35: Evolution des caractéristiques géométriques: (a) surface, (b) Distance maximale, (c) densité, (d) saut de centre et (e) le contour.....	126
Figure 4-36: Dynamique de la méthode adaptée. (a) Signal 2 : roulement sain ; vitesse =1405 tr/mn, charge= 130 daN. (b) Signal 8 : roulement sain ; vitesse =1405 tr/mn, charge= 130 daN. (c) Signal 60 : roulement sain ; vitesse =1560 tr/mn, charge= 250 daN(=) (d) Signal 61 : Roulement défectueux, deux classes détectées.....	127
Figure 4-37: Evolution des caractéristiques géométriques: (a) surface, (b) Distance maximale, (c) densité.....	129
Figure 4-38: Comparaison de la dynamique de détection	131

Abréviations et notations

φ	Angle de contact
$\psi(t)$	ondelette
$C_{a,b}$	Coefficient d'ondelette
ξ_i	facteur d'amortissement
$\varepsilon - \text{voisinage}$	Rayon de voisinage
μ	Degré d'apprentissage
Γ	Indice de Hubert
θ	Ampleur angulaire de la zone de charge
A	Amplitude du signal
$br(t)$	bruit de fond additif
c	contour
C	classe
Cep	cepstre
D	Ensemble de points
d	distance
DB	Indice de Davies et Bouldin
$DBSCAN$	
Den	Densité
$Dist$	Saut du centre d'une classe
d_{max}	Distance maximale
DPO	décomposition en paquets d'ondelettes

<i>DWV</i>	Distribution wigner ville
<i>E(t)</i>	Energie instantanée
<i>Eps</i>	Rayon de voisinage (dbscan)
<i>f_{bpf_i}</i>	Fréquence caractéristique du défaut sur la bague intérieure
<i>f_{bpf_o}</i>	Fréquence caractéristique du défaut sur la bague extérieure
<i>FC</i>	centre de fréquence
<i>f_m</i>	fréquence de la m ^{ème} ligne
<i>f_{ni}</i>	Fréquence de résonance
<i>f_r</i>	Fréquence de rotation
<i>f_{rms}</i>	spectre de puissance de l'enveloppe du signal
<i>inter</i>	Critère inter-classe
<i>intra</i>	Critère intra-classe
<i>K</i>	kurtosis
<i>L</i>	Ensemble d'objets d'apprentissage
<i>MinPts</i>	Minimum nombre de points
<i>n</i>	Nombre d'éléments roulant de roulement
<i>N</i>	Nombre de points du signal
<i>P_o</i>	objets
<i>P_{CWT}</i>	valeur moyenne des amplitudes du spectre d'enveloppe
<i>Peak</i>	peak
<i>q(t)</i>	Répétition de la charge autour des éléments roulant du roulement
<i>Ra</i>	Rugosité arithmétique
<i>RMS</i>	Root Mean Square

rms_f	RMS fréquentielle
RSB	Rapport signal sur bruit
$RT-OPTICS$	Real Time Ordering Points to Identify the Clustering Structure
S	surface
$S(C)$	silhouette
$skweness$	skweness
$SPRI$	Pique du spectre de la bague intérieure
$SPRO$	Pique du spectre de la bague extérieure
std_f	déviatiion standard des fréquences
$STFT$	Transformée de fourier en court terme
T_d	Période de répétition des impultions
TOD	Décomposition en ondelettes discrètes
V	Voisinage d'un objet
W_{rms}	valeur efficace des fréquences
X	Matrice d'indicateurs
$x(t)$	Signal vibratoire temporel
σ	variance
τ	Variable aléatoire

Introduction générale

Introduction générale

Dans de nombreux secteurs industriels, la sûreté de fonctionnement est un enjeu majeur pour assurer une compétitivité optimale de l'outil de production. Les gains de productivité représentent un souci quotidien pour les dirigeants des entreprises. La compétitivité de ces dernières passe par la maîtrise de la disponibilité de l'outil de production, l'amélioration de son fonctionnement, la sécurité de ses utilisateurs et la réduction des coûts d'entretien. Il existe plusieurs notions permettant une optimisation des systèmes industriels et qui s'intègrent à la discipline de la sécurité de fonctionnement : fiabilité, maintenabilité, sécurité et maîtrise des risques. Le diagnostic des défaillances de systèmes industriels, lorsqu'il est réalisé avec efficacité et s'il permet de détecter de façon précoce une dégradation, représente un des moyens pour contribuer à obtenir un meilleur gain de productivité. Sa vocation première est de détecter et de localiser une possible défaillance des matériels. Le diagnostic industriel est devenu une discipline incontournable dans le domaine de la sûreté de fonctionnement.

L'identification d'anomalies à un stade précoce, le suivi de leurs évolutions et le diagnostic de l'état d'une machine ne sont possibles que si, préalablement, l'on connaît la cinématique de la machine et les manifestations dynamiques ou symptômes associés aux différents défauts potentiels tels que les défauts de roulements, défauts d'engrenage, défaut de moteur, etc. Pour ces types de défauts, l'analyse vibratoire s'est avérée comme outil très apprécié dans l'industrie ces dernières décennies. Elle consiste en l'extraction des indicateurs capables d'identifier l'état du (ou des) composant(s) à partir des signaux temporels (valeur efficace, facteur crête, Kurtosis, etc.) ou fréquentiels (amplitude des fréquences caractéristiques de défaut, Kurtosis fréquentiel, centre de fréquence, etc.) dans des conditions normales et défectueuses et de les faire suivre dans le temps, afin de définir les indicateurs de surveillance et pour formuler un diagnostic sur son état. L'extraction de ces paramètres repose sur des techniques de traitement du signal classiques ou des techniques récentes. Parmi les techniques de traitement du signal classiques on peut citer l'analyse temporelle, l'analyse spectrale, la démodulation ou l'analyse Temps-Fréquence. Dans le cas des machines complexes ces techniques ne permettent pas un diagnostic fiable. Cependant, les techniques récentes de traitement du signal telles que la décomposition empirique modale (EMD), la transformée en ondelettes, la cyclo-stationnarité, les méthodes temporelles et les méthodes issues de l'intelligence artificielle permettent de mieux diagnostiquer les machines tournantes. Néanmoins, l'efficacité de toutes ces techniques dépendent essentiellement des conditions de fonctionnement de la machine ainsi que le bruit environnant. Par conséquent, les indicateurs

de défauts d'un composant (roulement, engrenages, courroie, etc.) dépendent de plusieurs facteurs comme la cinématique de la machine, des défauts de montage/démontage de la machine tournante, etc. Ces facteurs rendent le diagnostic plus difficile car l'évolution des indicateurs de défauts ne signifie pas nécessairement l'apparition d'un défaut mécanique ou son évolution vers un état critique.

Parallèlement, en raison de variabilité permanente et irrégulière des paramètres de fonctionnement des machines tournantes en régime de fonctionnement non stationnaire, le diagnostic par ces indicateurs devient insuffisant voire inutile. Par conséquent, un nombre impressionnant de méthodes ont été présentées pour prendre des décisions automatiques, fiables et rapides sur la santé des machines en cours d'exécution. En outre, le défi consiste à développer des méthodes de diagnostic capables de détecter de façon précoces l'apparition des défauts avec plus de précision que celles qui existent actuellement. Parmi les méthodes existantes, les plus appropriées pour le diagnostic des systèmes complexes sont celles basées sur l'approche par Reconnaissance de Formes (RdF), car elles ne nécessitent pas de connaissance à priori du système.

La RdF regroupe l'ensemble des méthodes permettant la classification des formes (vecteurs de caractéristiques de l'état de fonctionnement du système) dans des classes. Chaque classe est associée à un mode de fonctionnement du système et elle est représentée par un modèle. Les modèles de classes sont utilisés pour la classification d'une nouvelle forme à une des classes existantes. Il existe plusieurs approches de reconnaissance de formes (RdF): La comparaison à des motifs ou à des modèles, l'approche statistique, l'approche syntaxique ou structurelle et les réseaux de neurones. Chaque approche peut être effectuée au moyen des divers algorithmes d'apprentissage automatique.

Dans l'apprentissage statique des données, les paramètres du modèle de classification demeurent inchangés au cours du temps. Pour ce type de données, l'enrichissement de la base de connaissance se traduit par une déformation locale du contour ou de la forme des classes sans mettre en cause les informations acquises précédemment. Toutefois, lorsque les connaissances évoluent dans le temps, il est important que le système d'apprentissage automatique soit capable de s'adapter en prenant en compte ces évolutions. Dans ce cas, les classes et/ou les formes sont dynamiques. Cette dynamique se traduit par le déplacement, l'élimination, la fusion, la scission, ..., des classes. Parmi les applications réelles nécessitant une classification dynamique, on peut citer le diagnostic industriel (évolution des modes de

fonctionnement), le diagnostic médical (expansion de maladie tel qu'un cancer), la surveillance vidéo (mouvement des cibles), etc.

Les algorithmes connus de clustering nécessitent des paramètres d'entrées qui sont difficiles à déterminer, mais qui ont une influence significative sur le résultat de clustering. La détermination des paramètres d'entrées pour les algorithmes de partitionnement (le nombre de clusters k) nécessitent une certaine connaissance du domaine qui n'est malheureusement pas disponible pour de nombreuses applications. Contrairement aux algorithmes de partitionnement, les algorithmes hiérarchiques n'ont pas besoin du nombre de clusters k comme une entrée. Toutefois, une condition de terminaison doit être définie en indiquant si le processus de fusion ou de division doit être arrêté. D'autres algorithmes de clustering tentent de remédier aux problèmes de définition des paramètres d'entrées, les méthodes de clustering basées sur la densité.

Une méthode de classification dynamique basée sur le principe DBSCAN a été utilisée dans le cadre de suivi de l'état de dégradation de roulement. Cette méthode reste sensible au choix de paramètre de densité. Une manière de remédier à ce problème est de définir un algorithme qui permet de prendre en considération toutes les valeurs possibles pour ce paramètre de densité. L'algorithme OPTICS qui est une extension de DBSCAN a été proposé. Il fournit un ordre de tous les objets de la base de données à partir duquel nous pouvons extraire tous les niveaux de clustering où chaque niveau est caractérisé par un paramètre de densité différent. Nous avons opté pour l'utilisation de l'algorithme OPTICS dans notre approche de détection de défaut, car, l'ordre fourni est plutôt insensible aux paramètres d'entrées. Cette approche permet juste la visualisation de la distribution des régions de différentes densités à un instant t donné. Plus particulièrement, ce travail de thèse concerne le développement d'une méthode de classification dynamique basée sur l'algorithme OPTICS en associant le paramètre temps afin d'analyser au fur et à mesure les signaux enregistrés.

Cette méthode devrait être capable d'identifier l'origine du changement observée sur les signaux vibratoire, elle doit permettre également de différencier l'évolution d'un défaut à l'évolution de la vitesse de rotation ou à l'évolution de la charge.

Dans ce contexte, la thèse comporte quatre chapitres. Le premier chapitre présente les différentes défaillances qui peuvent se produire sur une machine tournante ainsi que l'occurrence de chacun de ces défauts. Ce chapitre se décompose en deux grandes parties. La première partie fait un état de l'art sur les différentes défaillances que peut subir une machine tournante. Quant à la seconde partie, elle présente les techniques d'analyse vibratoire et

d'extraction d'indicateurs dans la littérature. Cet état de l'art, a pour but de mettre en évidence les avantages et les inconvénients de différentes techniques utilisées dans l'analyse vibratoire.

Le deuxième chapitre est consacré à une description de la structure générale des méthodes de reconnaissance des formes: prétraitement, extraction des caractéristiques, sélection, classification, post traitement et décision. Dans ce chapitre, un intérêt particulier est accordé à l'étape "classification" que nous considérons comme objectif principal de notre étude. Nous rappelons dans une synthèse bibliographique les différentes approches de classifications déjà employées dans le diagnostic des machines tournantes. Enfin nous introduirons, dans ce chapitre, la limitation des méthodes statiques souvent utilisés dans le diagnostic et l'apport de la classification dynamique par rapport la classification statique.

Dans le troisième chapitre, nous présentons en premier lieu le processus de conception d'une méthode de classification (reconnaissance des formes) développée pour le diagnostic et le suivi de défaut de roulements, ensuite nous détaillerons l'approche de diagnostic proposée qui est réalisée en deux étapes: la détection et le suivi. La première étape déploie la détection de défaut par la méthode adaptée. La deuxième étape permet de discriminer la nature de la dérive par le suivi de quelques caractéristiques géométriques du nuage de points dans le temps.

Le quatrième chapitre est consacré à l'application du nouveau processus de diagnostic des défauts de roulements à bille et des roulements à galet. Une description du banc d'essais ainsi que la démarche expérimentale sont présentées. Nous avons d'abord simulé numériquement le défaut du roulement à l'aide d'un modèle qui existe déjà dans la littérature (BPFO, BPFI, BSF). Nous avons ensuite créé le défaut sur la bague extérieure d'un roulement à galet. Les essais (numériques et expérimentaux) ont été réalisés pour différents niveaux de charges et pour différentes valeurs de la vitesse de rotation du moteur. La première partie de ce chapitre est consacrée à l'extraction des différents indicateurs et à l'étude de l'influence de la variation de la charge et de la vitesse sur l'évolution de la gravité du défaut. La normalisation des données a été aussi réalisée. Celle-ci a pour but d'éliminer la dispersion d'un mode de fonctionnement due à la variation de charge. De ce fait, la discrimination des différents modes ainsi que des défaillances possibles est facilitée. La deuxième partie est consacrée à la mise en œuvre de l'algorithme proposé, RT-OPTICS présenté au chapitre III sur des signaux simulés et sur des signaux réels (issus d'un banc expérimental). Aussi, nous avons utilisé des caractéristiques géométriques des classes détectées pour le suivi de défaut.

Une étude comparative entre la méthode proposée et la méthode dynamique DBSCAN a été réalisée, suivi d'une discussion autour de ces deux méthodes.

Enfin, ces travaux donneront lieu à une conclusion générale et des perspectives.

Chapitre 1 :

Généralité sur la surveillance et le diagnostic
industriel

1 Généralité sur la surveillance et le diagnostic industriel

1.1 Introduction

Le premier pas vers la réduction des arrêts non planifiés réside dans la mise en œuvre d'une bonne stratégie de maintenance. Depuis plus de vingtaine d'années, la maintenance des installations industrielles fait appel à des techniques de diagnostic de plus en plus sophistiquées. Ainsi la communauté scientifique s'intéresse particulièrement à la détection et aussi à la surveillance des éléments tournants sans avoir arrêté ou perturber le fonctionnement des machines. Parmi les techniques utilisées pour détecter et suivre les défauts des éléments mécaniques d'une machine tournante, l'analyse vibratoire s'est avérée comme outil très apprécié dans l'industrie depuis plusieurs dernières décennies. Dans ce chapitre, nous allons présenter les différentes stratégies de maintenance ainsi que les différentes techniques de traitement du signal les plus utilisées en analyse vibratoire.

1.2 Différents niveaux de maintenance

Une détérioration ou une défaillance des installations dans le temps est un souci permanent. Il est alors légitime de maîtriser les risques inhérents pour une bonne utilisation des machines. L'enjeu ici, est d'assurer la sécurité des installations dont le matériel fait partie et au-delà, de la sécurité de l'ensemble des ressources nécessaires à leur exploitation (humaines et énergétiques). En outre, les pannes coutent chères en termes d'arrêt de production et de remise en état. Tous ces problèmes nous contraignent à une meilleure maîtrise de la durabilité des matériels de production et ainsi qu'une amélioration permanente de la productivité.

Il est nécessaire pour assurer une exploitation correcte de maintenir cette installation en bon état. La maintenance a un coût, et il est important de trouver le niveau optimum qui maintiendra au moindre coût l'installation à niveau sans mettre en danger la sécurité des personnes.

Selon la norme NF-X60-010 (AFNOR, 1990) « *La maintenance est définie comme étant un ensemble d'activités destinées à maintenir ou rétablir un bien dans un état ou dans des conditions données de sûreté de fonctionnement pour accomplir une fonction requise* » (Monchy, 1996). Les activités de maintenance permettent une diminution des couts de production et l'amélioration de la qualité des produits. Il existe deux stratégies de maintenance : corrective ou préventive (ou conditionnel). Chacune d'entre elles correspondant

à un concept particulier et répondant à un besoin particulier. La figure 1-1 résume les types de maintenance rencontrés.

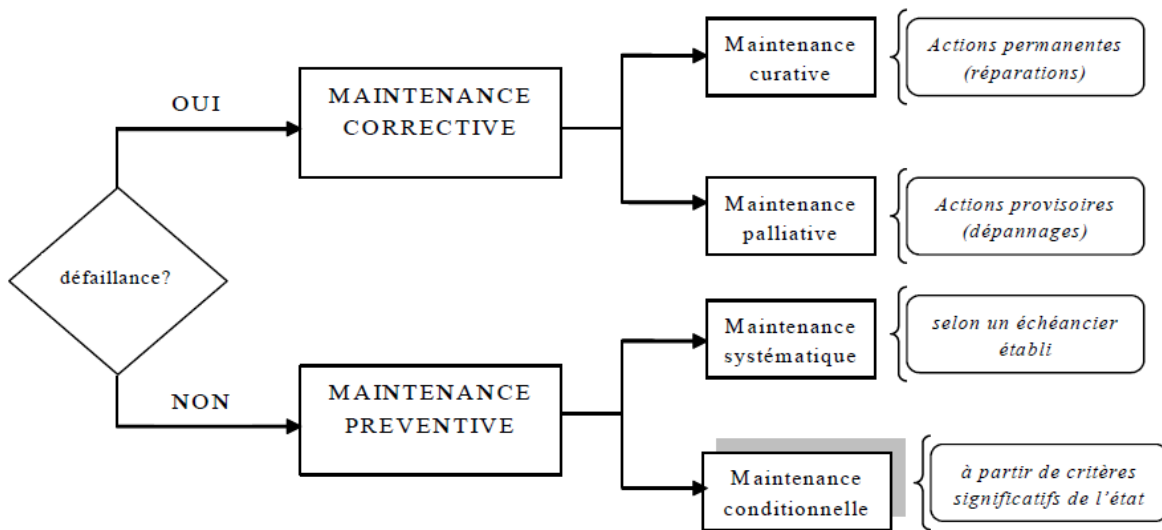


Figure 1-1: Les stratégies de maintenance (Larose, 2012)

1.2.1 Maintenance corrective

La maintenance corrective est une forme délibérée d'entretien qui consiste, une fois la défaillance survenue, à en éliminer les effets et les causes. On distingue généralement (AFNOR, 2002) :

- **Maintenance palliative** : effectue une remise en état provisoire du matériel ayant subi un dysfonctionnement : c'est un dépannage. Cette maintenance est le plus souvent associée à des systèmes ne présentant pas d'impératif de sécurité.
- **Maintenance curative** : correspond à la remise en état de l'équipement et revêt un caractère définitif. L'équipement concerné par cette maintenance retrouve, après intervention, les caractéristiques qu'il avait avant l'apparition du problème.
- **Maintenance corrective d'amélioration** : représente une forme dérivée de la maintenance curative. Elle ne se limite pas à la remise en état du système mais implique en plus une modification de ses caractéristiques afin d'éviter la répétition du problème.

1.2.2 Maintenance préventive

La maintenance préventive vise à réduire la probabilité de défaillance d'une unité fonctionnelle. Les actions peuvent être déclenchées selon un échéancier établi à partir d'un

nombre prédéterminé d'unités d'usage (maintenance systématique) et/ou à partir de critères prédéterminés significatifs de l'état de dégradation (seuils de maintenance conditionnelle) (Monchy, 1996). On peut distinguer deux types essentiels de maintenance préventive:

- **Maintenance systématique** : elle est caractérisée par la connaissance des dates de visite, l'intervalle d'inspection étant déterminé arbitrairement ou en fonction des lois de comportement du système.

- **Maintenance conditionnelle** : ce type de maintenance comprend toutes les tâches de restauration de matériels ou de composants non défaillants, entreprise en application d'une évaluation d'état et de la comparaison avec un critère d'acceptation préétabli (défaillance potentielle).

Contrairement à la précédente, celle-ci n'est pas basée sur la statistique, mais sur l'induction scientifique. Par cet aspect, elle rappelle le principe fondamental des sciences expérimentales, l'interaction permanente entre l'observation des faits réels et l'élaboration de théories. Cette forme de maintenance possède aussi des objectifs bien plus ambitieux. Pour une machine, elle permet de déterminer d'une part, quel organe défaillant devra être remplacé et, d'autre part, la date à laquelle imposer l'intervention. Bien évidemment pour faire de telles prédictions, il est nécessaire d'accéder à une bonne compréhension du comportement interne de la machine, et par exemple à une bonne interprétation des phénomènes vibratoires.

En effet, il n'est pas possible de prédire l'avenir d'une machine sans connaître son état de santé au temps précis par rapport à un état de référence initial. Les avantages de ce type de maintenance sont :

- élimination d'arrêts imprévus (fiabilité et productivité accrues).
- élimination des dommages secondaires (un roulement défectueux non détecté à temps peut détruire une boîte de vitesse ou provoquer un incendie).
- élimination du gaspillage des pièces de rechange (pas de remplacement inutiles).
- réduction du stock de pièces de rechange (l'alarme est donnée suffisamment tôt pour que l'on puisse commander les pièces nécessaires).
- réduction du nombre des arrêts (en profitant d'un arrêt et production planifiée pour effectuer la réparation, en intervenant sur plusieurs machines lors d'un seul arrêt).
- amélioration de la sécurité des personnels, de la qualité du service.

Les avantages d'une maintenance conditionnelle bien conçue sont tellement évidents qu'elle devrait être implantée dans la plupart des entreprises. Pour y parvenir, plusieurs méthodes peuvent être utilisées telles que la thermographie infrarouge, l'analyse des huiles, l'analyse des signaux électriques, l'analyse vibratoire, etc. Parmi les techniques existantes, l'analyse vibratoire apparaît comme étant prometteur et sont en plein essor au cours des vingt dernières années. Par ailleurs, plusieurs techniques de traitement des signaux vibratoires ont été développées pour détecter et diagnostiquer les différents défauts mécaniques dans les machines tournantes. Ces techniques sont basées sur l'extraction des paramètres significatifs de l'état du (ou des) des composant(s) mécaniques soit par l'analyse temporelle du signal, soit par l'analyse fréquentielle ou l'analyse temps-fréquence soit par l'exploitation des caractéristiques du signal comme la cyclo stationnarité.

1.3 L'analyse vibratoire en tant qu'outil de diagnostic

La détection d'un défaut nécessite d'une part une prise de mesure vibratoire puis une exploitation du signal recueilli. Nous présentons, au cours de ce chapitre, les méthodes de traitement de signal les plus utilisées.

Le traitement des données vibratoires et leurs analyses subséquentes dépendent bien entendu de la nature du signal. Certaines techniques de traitement de fonctions bidimensionnelles peuvent aussi produire des indicateurs graphiques, à l'exemple de l'analyse en temps-fréquence ou en temps-échelle. C'est pourquoi il est courant de catégoriser les analyses en trois catégories selon les domaines de représentation : en temps, en fréquence et en Temps-Fréquence.

Ce chapitre décrit les principes et les applications de ces techniques d'analyse vibratoire dans le cadre de diagnostic automatisé des machines tournantes. Nous nous intéressons, en particulier, à celles que nous allons utiliser dans le chapitre développement expérimentale.

1.4 Image vibratoire des défauts de roulement

1.4.1 Constitution des roulements

De manière générale, les roulements sont constitués de quatre composants essentiels. Une bague extérieure, une bague intérieure, des éléments roulants et une cage. En fonction de l'application envisagée, les éléments roulants peuvent être des billes, des rouleaux cylindriques ou des aiguilles (rouleaux cylindriques très longs devant leur diamètre). Chaque type de roulement possède des caractéristiques spécifiques en termes de charges radiales et axiales supportées, de vitesse limite ou encore de défaut d'alignement admissible.

1.4.2 Modes de dégradation des roulements

Les roulements sont parmi les composants les plus sollicités des machines et représentent une source de panne fréquente. Les défauts que l'on peut y rencontrer sont les suivants : écaillage, grippage, corrosion (qui entraîne l'écaillage), faux effet Brinell, etc... Tous ces défauts se traduisent par une perte de métal et provoquent des chocs répétés des billes sur la cage de roulement (Kerroumi, 2016).

Par définition, la durée de vie d'un roulement est une donnée statistique caractérisant le nombre de cycles, c'est-à-dire de tours qu'il peut effectuer en étant soumis à une contrainte spécifique, avant de présenter des défauts d'écaillage des surfaces de contact. Une courbe de fiabilité des roulements est obtenue analytiquement à partir de la durée de fonctionnement atteinte par 90% des roulements d'un même type. Cependant, cette évaluation de la durée de vie des roulements ne prend en compte que les phénomènes de fatigue des matériaux.

La dégradation du roulement peut être le résultat des causes naturelles (la fatigue) ou le résultat d'une mauvaise utilisation ou mauvais montage du roulement. Les défauts que l'on peut y rencontrer sont les suivants (Chimentin, 2007) :

- Le grippage : dû à l'absence de lubrification, à une vitesse excessive ou un mauvais choix du type de roulement. Ceci se manifeste par un transfert de matière arrachée sur les surfaces et redéposée par microsoudure.
- Les empreintes par déformation, dues à des traces de coups, des fissures ou des cassures.
- L'incrustation de particules étrangères, due à un manque de propreté au montage ou de l'entrée accidentelle d'impuretés.
- La corrosion, due à un mauvais choix du lubrifiant, surtout quand les roulements viennent d'être nettoyés et sont contaminés par la transpiration des mains.
- La corrosion de contact, due au mauvais choix d'ajustements entre les bagues et les logements ou les arbres.
- Les criques, fissures étroites ou autres amorces de cassures dues aux contraintes exagérées au montage ou au démontage.
- L'usure par abrasion, due à une mauvaise lubrification. L'usure par abrasion donne aux roulements un aspect gris, givré.

1.4.3 Fréquences caractéristiques de défauts

Les pistes des roulements et les billes sont chargées cycliquement, ce qui engendre à la surface une dégradation par fatigue qui se présente sous la forme de fissures conduisant à l'écaillage (Chimentin, 2007) puis à la ruine du roulement. Ces fissures peuvent être d'origine superficielle ou provenir de la dégradation de la sous couche du matériau. Cette ruine peut être détectée sur un spectre à partir de l'identification des fréquences caractéristiques du roulement. Celles-ci correspondent aux fréquences des impacts lorsqu'un élément roulant rencontre un défaut. Elles sont données en écrivant que les vitesses aux points de contact entre billes et bague sont nulles (Morel, 2005). Elles dépendent du diamètre des éléments roulants (d), du diamètre moyen du roulement (D), du nombre d'éléments roulants (n_b), de l'angle de contact (φ) et la fréquence de rotation relative entre les bagues intérieure et extérieure (f_r) (voir figure 1-2 et figure 1-3). Elles sont données par les formules suivantes :

- Fréquence d'un défaut localisé sur la bague externe du roulement :

$$f_{be} = \left(\frac{1}{2}\right) f_r n_b \left[1 - \left(\frac{d}{D}\right) \cos \varphi \right] \quad (1.1)$$

- Fréquence d'un défaut localisé sur la bague interne du roulement :

$$f_{bi} = \left(\frac{1}{2}\right) f_r n_b \left[1 + \left(\frac{d}{D}\right) \cos \varphi \right] \quad (1.2)$$

- Fréquence d'un défaut localisé sur l'élément roulant :

$$f_b = \left(\frac{D}{d}\right) f_r \left[1 - \left(\left(\frac{d}{D}\right) \cos \varphi \right)^2 \right] \quad (1.3)$$

- Fréquence d'un défaut de cage :

$$f_c = \left(\frac{1}{2}\right) f_r \left[1 - \left(\frac{d}{D}\right) \cos \varphi \right] \quad (1.4)$$

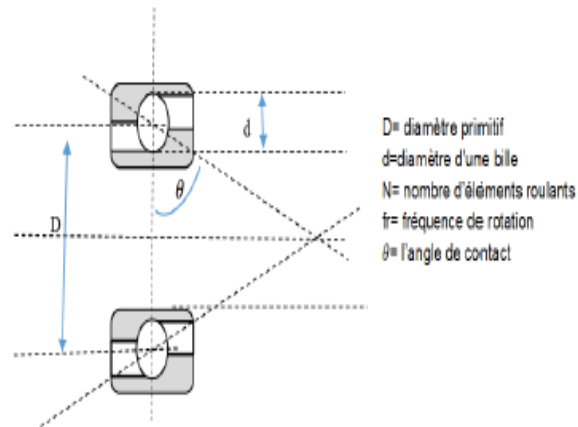


Figure 1-2: Géométrie d'un roulement (Morel, 2005)

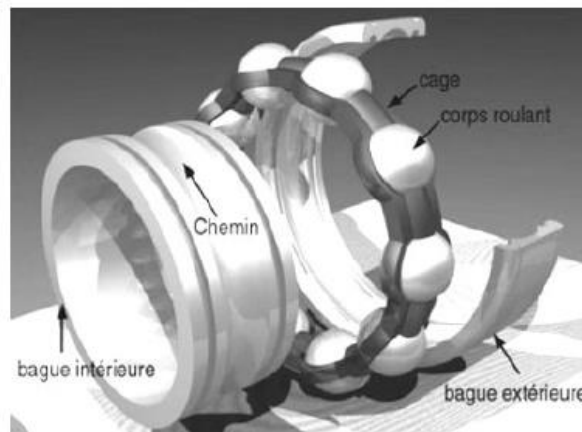


Figure 1-3: Vue éclatée d'un roulement

Ces fréquences peuvent être rencontrées pour un roulement en fonctionnement normal (du fait du non linéarité de la raideur de contact sur un roulement en fonctionnement), mais surtout lors de l'apparition de défauts sur les éléments constituant qui vont engendrer des impulsions produites à chaque fois qu'un élément roulant rencontre un écaillage sur son chemin.

Il existe d'autres types de défauts qui manifestent lors de fonctionnement des machines tournante. Ces défauts sont résumés dans le tableau (1-1).

Défaut	Fréquence à laquelle le défaut est attendu se manifester (f_r : fréquence de rotation)
Déséquilibre (balourd)	f_r (manifestation radiale)
désalignement	f_r : désalignement angulaire, manifestation axiale $2f_r$: désalignement parallèle, manifestation radiale
Jeux mécanique, courroie	$0.5 f_r$, f_r et harmonique d'ordre supérieur
Instabilités hydrodynamiques dans les paliers	40-50 % f_r
Défaut de palier à éléments roulants	<p>Défaut de bague externe : $f_{be} = \left(\frac{1}{2}\right) f_r n_b \left[1 - \left(\frac{d}{D}\right) \cos \varphi\right]$</p> <p>Défaut de bague interne : $f_{bi} = \left(\frac{1}{2}\right) f_r n_b \left[1 + \left(\frac{d}{D}\right) \cos \varphi\right]$</p> <p>Défaut d'élément roulant : $f_b = \left(\frac{D}{d}\right) f_r \left[1 - \left(\left(\frac{d}{D}\right) \cos \varphi\right)^2\right]$</p> <p>Défaut de cage : $f_c = \left(\frac{1}{2}\right) f_r \left[1 - \left(\frac{d}{D}\right) \cos \varphi\right]$</p> <p>(+ : dans le cas où la bague externe est tournante ; - dans le cas où la bague interne est tournante)</p> <p>f_r : fréquence de rotation de l'arbre D : diamètre de l'élément roulant α : angle de contact ; n : nombre d'éléments roulants.</p>
Défaut d'engrènement	Modulation d'amplitude et de phase à la fréquence d'engrènement $z f_r$; composantes satellites à $f_e \pm k f_r$ (z : nombre de dents, k : entier)
Défaut de ventilation	PPMC (bras, pales) f_r
Défaut de pompage	PPMC (aubes, 2) f_r si on suppose une admission et un refoulement
Défaut de cavitation	f_r ; $N f_r$ et spectre étendu aux fréquences sup. à 2000Hz
Défauts électriques (Didier, 2004)	<p>f_r et harmonique d'ordre élevé : balourd magnétique</p> <p>$f_{el} 2f_{el}$: défauts de champ fixe</p> <p>$f_{r \pm 2g} f_{el} ; f_r \left(k \left(\frac{1-g}{p} \right) \pm g \right)$: défaut de champ tournant ;</p> <p>g : glissement</p> <p>f_{el} : fréquence d'alimentation électrique</p> <p>$k/p=1, 5, 7...$ en fonction de bobinage moteur</p>

Tableau 1-1: Modes défailants sur les machines tournantes

1.4.4 Indicateurs scalaires

Dans la pratique, la présence des défauts aux fréquences caractéristiques n'est pas toujours facile à déceler. Pour les mettre en évidence, plusieurs techniques d'analyse peuvent être mises en œuvre. La première technique est l'analyse temporelle. Ce type d'analyse est basés directement sur l'extraction des paramètres dits « indicateurs de défauts » à partir d'un signal vibratoire, les plus simples étant les indicateurs scalaires. Un indicateur scalaire peut correspondre à une grandeur caractéristique : (i) de son amplitude (valeur efficace, amplitude crête, taux de modulation...), (ii) de sa distribution d'amplitude (facteur de crête, kurtosis) ou (iii) de sa composition spectrale (amplitude d'une composante spectrale, valeur efficace d'une famille de composantes, taux d'harmoniques...).

Ces indicateurs permettent de caractériser l'état de fonctionnement global des équipements mais ne localisent pas le défaut. Leur utilisation très répandue s'explique aisément par leur facilité d'utilisation : ils se réduisent à un nombre, se prêtent facilement à l'automatisation de leur gestion (archivage, courbes d'évolution, comparaison à des seuils).

De nombreux indicateurs ont été développés dans la littérature. Il existe également des indicateurs obtenus par une combinaison de plusieurs indicateurs existants (Zhang et al., 2013a) ont proposé un ensemble d'indicateurs scalaires pour le diagnostic automatisé des défauts de roulements.

1.4.4.1 Kurtosis

Le Kurtosis (mot grecque signifiant « bosse ») est défini comme étant le quatrième moment normé d'une distribution dite gaussienne: il traduit la largeur de la courbe de densité de probabilité du signal. Le Kurtosis fournit une mesure de la taille de la queue de la distribution et est utilisé comme indicateur de pics principaux dans un ensemble de données. Pour une distribution Gaussienne de référence, le Kurtosis vaut 3. Lorsque les valeurs de la variable sont fortement concentrées autour de la moyenne, l'écart type est faible et le Kurtosis excède 3 (figure 1-4). Dans le cas contraire, la courbe est nettement plus étalée et le Kurtosis est inférieur à 3. L'équation d'aplatissement est donné par:

$$K = \frac{\sum_{i=1}^N [x(i) - \bar{x}]^4}{N \cdot (\sigma^2)^2} \quad (1.5)$$

Où $x(i)$ est la série temporelle au point n , \bar{x} est la moyenne des données, σ^2 est la variance des données, et N est le nombre total de points de données.

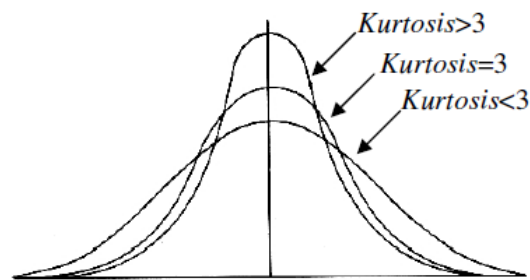


Figure 1-4: Représentation schématique de la distribution d'une variable aléatoire correspondant à trois valeurs du Kurtosis

Les travaux de Dyer & Stewart (1978) constituent la première application de ce paramètre à la détection de défauts dans les roulements. En effet, un roulement en bon état génère un signal vibratoire proche d'une distribution Gaussienne. Lorsqu'il vieillit, les traces d'usures, telles qu'un écaillage d'un chemin des éléments roulants de roulement, vont générer des chocs plus ou moins périodiques. La densité de probabilité se concentre autour de la valeur particulière associée et le Kurtosis dépasse la valeur 3, c'est le cas en présence de chocs répétés. Par contre, lorsque l'endommagement devient trop sévère (généralisé), le signal retrouve une composante aléatoire plus marquée et le Kurtosis décroît en tendant à nouveau vers 3. Il est aussi sensible à certains paramètres comme la vitesse de rotation, les bandes de fréquences étudiées ou le niveau du bruit de fond. Cela explique en partie son succès relativement limité auprès des industriels. Dans plusieurs travaux, le kurtosis s'est montré plus sensible que les autres indicateurs scalaires, ce qui le place en tant qu'indicateur privilégié lorsqu'il s'agit de défauts de type chocs, notamment ceux des roulements et des engrenages (Pachaud, 1997). Néanmoins, il a été utilisé par les chercheurs comme base de référence afin d'évaluer les possibilités de nouveaux indicateurs (Honarvar & Martin, 1997 ; Li et al., 2013a ; Samanta & Nataraj, 2009 ; Zhang et al., 2013a ; Li et al., 2013b). Il a aussi servi d'application à de nouveaux développements en particulier concernant l'extraction du bruit de fond. Dron et al. (2004) ont proposé la soustraction spectrale comme méthode de débruitage de la partie aléatoire du signal pour accroître les sensibilités du Kurtosis et du facteur de crête.

1.4.4.2 Skewness

Le coefficient de dissymétrie (skewness) est défini comme étant le troisième moment statistique centré. Il correspond à une mesure de l'asymétrie de la distribution d'une variable aléatoire réelle.

$$\text{skewness} = \frac{\sum_{i=1}^N (x(i) - \bar{x})^3}{(N-1)\sigma^3} \quad (1.6)$$

Avec : $x(i)$ est la série temporelle au point n ; \bar{x} est la moyenne des données et σ est la valeur efficace.

La valeur efficace ou *RMS* (Root Mean Square), notée *RMS*, est l'indicateur scalaire le plus couramment utilisé vu sa relation directe avec l'énergie contenue dans le signal vibratoire. La *RMS* est la racine carré de la moyenne quadratique du signal vibratoire temporel discrétisé $x(i)$ de longueur N et de moyenne empirique \bar{x} . Appelé également niveau vibratoire global, il est représentatif de l'énergie de la vibration basses fréquences.

$$RMS = \sqrt{\frac{\sum_{i=1}^N (x(i))^2}{N}} \quad (1.7)$$

La méthode la plus simple pour mesurer les défauts dans le domaine temporel est en l'utilisation de la valeur efficace (RMS). Cependant, le niveau *RMS* ne peut pas montrer des changements appréciables dans les premiers stades de défaillances des roulements. Cet indicateur ne fournit donc pas une information sur la naissance du défaut. Il est utile pour le suivi du niveau global. La valeur *RMS* est utilisée souvent pour détecter des dissipations d'énergie anormalement élevées, elle peut être utilisée aussi pour la surveillance acoustique.

1.4.4.3 Facteur crête

Cet indicateur correspond au rapport entre la valeur crête (valeur absolue maximale atteinte par la fonction représentative du signal pendant la période de temps considérée) et la valeur efficace du signal.

$$FC = \frac{\sup|x(i)|}{\sqrt{\frac{1}{N} \sum_{i=1}^N [x(i)]^2}} \quad (1.8)$$

Avec N nombre d'échantillons prélevé dans le signal, $x(i)$ signal temporel.

Le facteur de crête "*FC*" est l'indicateur le mieux adapté pour détecter des changements dans le modèle de signal dû à des sources de vibrations impulsives comme un défaut d'écaillage sur la bague extérieure d'un roulement. En effet la valeur du facteur crête d'un signal dont la distribution des amplitudes est gaussienne est principalement comprise entre 3 et 6, cet indicateur est donc adapté à la détection de ce type de défauts car sa valeur est au moins égale à 6 en présence du défaut (Campagna-Varenne, 1985)

Le facteur crête s'interprète par son évolution au fur et à mesure de la dégradation du roulement. La valeur de *FC* augmente avec l'usure du roulement. Il donne une information précoce sur la dégradation du roulement tout en étant indépendant des caractéristiques de fonctionnement des machines (vitesse et/ou charge) ou bien les dimensions du roulement. L'inconvénient majeur de cet indicateur réside dans sa décroissance lors d'un développement important du défaut. D'après Boulenger & Pachaud (1995), une valeur faible du facteur de crête, associée à un niveau bas de la valeur efficace et de la valeur crête, correspond à un roulement en bon état. Lorsque la dégradation s'amorce, les chocs répétés vont amplifier la valeur de crête et donc le facteur de crête. Ce n'est que lorsque le roulement est vraiment dégradé que le *RMS* du signal augmente sensiblement, réduisant le facteur de crête.

Il existe encore de nombreux indicateurs dérivés du *FC* (Valeur crête à crête, Valeur crête mini, facteur de forme,...) mais ils présentent tous en quelque sorte des avantages et des faiblesses comparables, qui ont conduit à une utilisation relativement limitée.

1.4.4.4 Facteur K

Le facteur K d'un signal est défini comme étant le produit entre la valeur crête et la valeur efficace.

$$K = peak \times RMS = peak \times \sqrt{\frac{1}{N} \sum_{i=1}^N x(i)^2} \quad (1.9)$$

L'interprétation du facteur *K* se fait au travers de son évolution au fur et à mesure de la dégradation du roulement. La valeur du facteur *K* augmente avec l'usure du roulement.

1.4.4.5 Analyse cepstrale

Les signaux issus des machines tournantes sont typiquement non-stationnaires. La technique d'analyse de Fourier suppose que les signaux soient stationnaires, elle ne permet donc pas d'associer à une signature fréquentielle l'intervalle temporel où elle se réalise

(Rubini & Meneghetti, 2001). Une autre méthode a été utilisée pour remédier ce problème, le cepstre.

Le cepstre est un outil mathématique qui permet la mise en évidence des périodicités dans un spectre. Le cepstre d'énergie d'un signal est défini par :

$$Cep(t) = TF^{-1}\left(\ln(|S(f)|^2)\right) \quad (1.10)$$

Il résulte de la transformée de Fourier inverse (TF^{-1}) du logarithme d'un spectre de puissance ($|S(f)|^2$). Le cepstre associe à une famille de raies harmoniques ou un ensemble de bandes latérales une raie unique dans sa représentation graphique. Chaque raie s'appelle le "*rhamonique*" et leur position sur l'axe des abscisses s'appelle le "quefrenc". Le cepstre est utilisé pour le diagnostic des phénomènes de chocs périodiques (desserrages, défauts de dentures, écaillage de roulements) et des phénomènes de modulation en fréquence ou en amplitude (Badaoui, 1999).

1.4.4.6 Analyse spectrale

L'analyse fréquentielle ou l'analyse spectrale établit la signature de la machine et révèle les fréquences auxquelles se produisent des variations d'amplitude significatives (thomas, 2007).

L'utilisation directe des composantes de la transformée de Fourier comme indicateurs, d'une manière quantitative, n'est pas pratique en diagnostic des défauts due à la quantité massive de l'information contenue dans cette représentation et à la résolution fréquentielle adoptée. Pour cette raison, plusieurs indicateurs ont été proposés dans la littérature. Dans Zhao et al. (2013), les auteurs proposent une série d'indicateurs afin de détecter les défauts d'engrenages d'une boîte de vitesse. Ces indicateurs ont été utilisés aussi dans le cadre du diagnostic des défauts de roulement par Kerroumi et al. (2013). Parmi les caractéristiques statistiques calculé dans la domaine fréquentielle, on trouve le centre de fréquence FC , le RMS fréquentielle ou la valeur efficace des pics rms_f , le spectre de puissance de l'enveloppe du signal $frmset$ la déviation standard des fréquences std_f :

$$FC_{env} = \frac{\sum_{m=1}^L f_m(H(f_m))}{\sum_{m=1}^L H(f_m)} \quad (1.11)$$

$$rmsf_{env} = \sqrt{\frac{\sum_{m=1}^L f_m^2 (H(f_m))}{\sum_{m=1}^L H(f_m)}} \quad (1.12)$$

$$frms_{env} = \frac{\sum_{m=1}^L (H(f_m))}{L} \quad (1.13)$$

$$stdf_{env} = \sqrt{\frac{\sum_{m=1}^L (f_m - FC_{env})^2 (H(f_m))}{\sum_{m=1}^L H(f_m)}} \quad (1.14)$$

Avec $H(f_m)$ est le spectre de puissance du signal $x(i)$, pour $m = 1, 2, \dots, L$. L est le nombre de lignes spectrales et f_m est la fréquence de la $m^{ème}$ ligne.

Les fréquences fondamentales dépendent de la géométrie du composant et de sa vitesse de rotation. Cependant les fréquences de défauts calculées ne correspondent pas souvent aux fréquences réelles qui apparaissent dans le spectre. Ce qui peut être interprété par un dépassement de charges admissibles supportée par le roulement conduisant à un fonctionnement avec un angle de contact différent de celui qui le caractérise ou à une variation de la vitesse de rotation qui provoque son échauffement.

En plus des trois premiers indicateurs, deux autres ont été proposés dans Zhang et al. (2013a) pour identifier la présence des pics aux alentours des fréquences de défauts de la bague intérieure et extérieure d'un roulement à rouleaux : on peut citer

Pique du spectre de la bague intérieure (Spectrum peak ratio inner)

$$SPRI = \frac{k \sum_{h=1}^H PI(h)}{\sum_{k=1}^K S(f_k)} \quad (1.15)$$

Pique du spectre de la bague extérieure (Spectrum peak ratio outer)

$$SPRO = \frac{k \sum_{h=1}^H PO(h)}{\sum_{k=1}^K S(f_k)} \quad (1.16)$$

$PO(h)$ et $PI(h)$ représentent les amplitudes de la $h^{ème}$ harmonique des fréquences caractéristiques de la bague extérieure BPFO et de la bague intérieure BPFI et $h = 1, 2, \dots, H$. H est le nombre d'harmoniques.

1.4.5 Analyse d'enveloppe

L'analyse d'enveloppe est une technique très utilisée pour la détection des défauts se manifestant dans les hautes fréquences. Ces défauts sont de faible énergie. Ils peuvent s'apparenter à de petits chocs qui excitent la fréquence de résonance d'une structure de la machine. Ces fréquences se situent généralement entre 1 et 15 kHz.

Il est à noter que cette méthode nécessite de connaître une résonance de structure en hautes fréquences et elle est inefficace devant un bruit trop élevé (Randall, 2001). En effet, le choc généré par l'élément roulant sur le défaut excite la structure sur toutes ses fréquences de résonance. A chaque fois que ce phénomène se produit, il génère une vibration. Ainsi l'amplitude de la vibration à la fréquence de résonance varie avec une période égale à la période de répétition de chocs caractéristiques du défaut. L'amplitude du signal vibratoire est donc modulée (Randall, 2011). La transformée de Hilbert permet d'extraire l'enveloppe d'un signal $x(t)$ (Decoulon, 1984 ; Safizadeh, 1999). En effet, la technique de détection d'enveloppe (HFRT, High frequency resonance technique) est basée sur la transformée de Hilbert et se décompose en plusieurs étapes (Huang et al., 1998). Dans la première étape, on réalise le filtrage passe-bande du signal $x(t)$ autour d'une fréquence particulière (en générale la fréquence de résonance), on calcule ensuite la transformée de Hilbert pour extraire l'enveloppe du signal modulé en amplitude (Decoulon, 1984). La transformée de Fourier de l'enveloppe permet ainsi de retrouver la fréquence caractéristique de défaut et les harmoniques de cette fréquence.

McFadden & Smith (1985) a étudié la façon dont le signal de défaut de roulement (représenté par un déplacement plutôt que d'une accélération) se manifeste par le spectre d'enveloppe. Les auteurs ont développé un modèle mathématique, pour des vibrations à haute fréquence produite par un défaut sur la bague intérieure d'un roulement, sous une charge radiale. Le modèle intègre les effets de la géométrie des roulements, de la vitesse, de la répartition de la charge, de la fonction de transfert et de la décroissance des vibrations. Les spectres de réponse vibratoires ont été obtenus en utilisant la technique de détection d'enveloppe. Une comparaison des spectres prédits et mesurés pour un roulement avec deux défauts ponctuels confirme les performances satisfaisantes du modèle. Ce modèle a été étendu par Su & Lin (1992) pour caractériser les vibrations des roulements qui ont été soumis à une charge variable.

D'autres applications de l'analyse d'enveloppe en diagnostic des défauts de roulements, peuvent être consultées dans les références (Randall, 2001 ; Stepanic et al., 2009 ; Li et al., 2012 ; Pan & Tsao, 2013).

L'analyse par la détection d'enveloppe d'un signal est un outil particulièrement efficace pour la recherche des informations fréquentielles contenu dans un signal périodique complexe rencontré dans la surveillance des machines tournantes. Elle est utilisée pour la détection de chocs cycliques de roulement en particulier (démodulation d'enveloppe) ou pour la caractérisation d'oscillation de rotation (engrenage, variations de couple dans des moteurs électriques,...).

1.4.6 Analyse temps-fréquence

Les techniques qui utilisent la transformée de Fourier ne sont pas suffisantes pour analyser la plupart de signaux vibratoires qui sont généralement non stationnaires. Les transformations temps-fréquences donnent un cadre mathématique optimal pour l'analyse des signaux non stationnaires (Boashash, 1990 ; Flandrin & Hermes, 1993).

L'analyse Temps-Fréquence traditionnelle utilise des distributions qui représentent l'énergie ou puissance des signaux en deux dimensions de temps et fréquence pour révéler l'évolution des fréquences caractéristiques de défauts dans le temps. La transformée de Fourier à court terme (*STFT*) ou spectrogramme (Jardine et al., 2006) et la distribution Wigner-Ville (*DWV*), sont les méthodes temps-fréquence les plus utilisées.

La transformation de Wigner-Ville est définie par :

$$W_x(t, f) = \int_{-\infty}^{+\infty} x\left(t + \frac{\tau}{2}\right) x^*\left(t - \frac{\tau}{2}\right) e^{-2\pi j f \tau} d\tau \quad (1.17)$$

Avec : $x(t)$ signal complexe quelconque.

L'énergie instantanée par unité de temps est définie par :

$$E(t) = \int_{f_1}^{f_2} |W_{xx}(t, f)| df \quad (1.18)$$

L'énergie instantanée peut être utilisée comme indicateur de défaut. La transformée de Wigner-Ville (plus généralement la distribution de Cohen) (Raad, 2003) est caractérisée par une bonne localisation précise de l'énergie dans le plan temps-fréquence. Cependant, le temps

de calcul d'une telle représentation peut être prohibitif et l'interprétation de l'image résultante est souvent difficile, ce qui rend la détection de défaut complexe (Didier, 2004)

Par ailleurs, le diagnostic de défaut de rotor est l'un des grands axes de plusieurs recherches. La transformée de Wigner Ville a été utilisée par Penga et al. (2003) dans le but de détecter les défauts provoqués par le frottement entre le rotor et stator avec une présentation de la méthode de scalogramme (conventionnelle).

1.4.7 Analyse temps-échelle:

Par rapport à l'analyse temps-fréquence, l'analyse temps-échelle consiste à adapter la taille de la fenêtre analysante $\Psi(t)$ suivant les caractéristiques du signal. Cette fenêtre analysante s'appelle « ondelette ». Des fenêtres courtes permettront une bonne localisation temporelle; des fenêtres plus longues fourniront une bonne résolution fréquentielle en basse fréquence. Plus précisément, Une ondelette désigne une fonction qui oscille sur un intervalle de longueur fini (un temps donné si la variable est de type temporels). L'ondelette, notée $\Psi(t)$ est une fonction continue, a des moments nuls, et est nulle au-delà d'un segment de \mathcal{R} . La condition d'admissibilité pour Ψ est définie par l'équation :

$$\int_{\mathcal{R}^-} \frac{|\Psi(S)|^2}{|S|} ds < +\infty \quad (1.19)$$

La forme de l'ondelette $\Psi(t)$ étant fixée, on modifie sa longueur par un facteur de dilatation ou de contraction pour changer de résolution. De ce fait, on examine le signal à différentes échelles et on réalise alors sa transformation en ondelette continue.

$$C_{a,b} = \int_{-\infty}^{+\infty} x(t)\psi_{a,b}(t)dt \quad (1.20)$$

Avec $\psi_{a,b}(t) = \frac{1}{\sqrt{a}}\psi\left(\frac{t-b}{a}\right)$ représente l'ondelette mère. $x(t)$ étant le signal temporel, ' b ' est un paramètre de translation (position temporelle) et ' a ' est un paramètre d'échelle qui mesure la dilatation ou la contraction de l'ondelette (notion de fréquence). Les coefficients d'ondelette $C_{a,b}$ sont obtenus par projection de $x(t)$ sur les bases $\Psi_{a,b}(t)$.

La transformée en ondelettes permet d'étudier la variation des caractéristiques fréquentielles dans le temps. Elle consiste à décomposer le signal en une somme d'ondelettes dilatées ou non et localisées temporellement. Notons que les ondelettes sont utilisées soit pour

réaliser un dé-bruitage du signal (Abbasion et al., 2007) ou bien pour réaliser un diagnostic en analyse vibratoire (Li & Ma, 1997). Le choix de la meilleure 'ondelette' en fonction du problème considéré est important dans un souci d'interprétation physique de la décomposition.

La démodulation d'amplitude à partir des ondelettes est efficace pour extraire une fréquence caractéristique du signal basse fréquence. L'utilisation de cette méthode nécessite dans un premier temps de choisir l'ondelette mère, sa fonction et ses paramètres (décroissance et fréquence). Ensuite, on doit faire varier le paramètre a sur une plage de valeurs qui est donné de façon arbitraire. La décomposition en ondelettes du signal génère une matrice W de coefficients $C_{a,b}$. Chaque colonne b correspond à un moment t et chaque ligne a correspond à une fréquence f_i . Sur la carte en temps échelle de $C_{a,b}$, que l'on appelle aussi "Scalogramme", les grandes valeurs de la matrice représentent les impulsions dues aux chocs dans le signal. L'enveloppe, $x_{max}(t)$, est formée en conservant la valeur maximale de chaque colonne de la matrice W . Le spectre du signal enveloppe est obtenu à partir de la transformée de Fourier de $x_{max}(t)$.

Deux indicateurs peuvent être extraits d'un spectre d'enveloppe : le W_{rms} ou la valeur efficace des fréquences et le P_{CWT} qui représente la valeur moyenne des amplitudes du spectre d'enveloppe.

$$W_{rms} = \sqrt{\frac{\sum_{j=1}^k f_j^2 (W_S(f_j))}{\sum_{j=1}^k W_S(f_j)}} \quad (1.21)$$

$$P_{CWT} = \frac{\sum_{j=1}^k W_S(f_j)}{k-1} \quad (1.22)$$

$W_S(j)$ est le spectre des coefficients max de la transformée en ondelettes continue pour $j = 1, 2, \dots, k$. k est le nombre de raies spectrales, f_j est la valeur de fréquence de la $j^{ème}$ ligne de spectre.

La première application de cette transformée, dans le domaine de la détection et du diagnostic des roulements, a été réalisé par Mori et al. (1996). Ensuite, de nombreuses études ont amélioré son utilisation et ont étendu le nombre d'ondelette mère (Misiti, 2003). Une revue bibliographique sur l'application de la transformation en ondelettes dans le domaine de diagnostic des machines tournantes a été réalisée par Peng et Chu (2004). Cette transformée a

été appliquée pour détecter les variations très rapides dans des signaux physiques, comme dans le cas d'un défaut de bague extérieure d'un roulement. Les impulsions dans le signal peuvent être détectées en hautes fréquences avec une bonne résolution. Par ailleurs, l'ondelette doit être choisie en fonction du signal à analyser et les objectifs à atteindre (débruitage ou détection) pour avoir des résultats satisfaisants.

Rubini & Meneghettien (2001) ont étudié la dégradation de la surface des éléments de roulement par l'utilisation de deux méthodes : l'analyse d'enveloppe et l'analyse par la transformation d'ondelettes. L'analyse par la transformée d'ondelettes est une méthode très sensible aux phénomènes de passage de bille, elle permet l'identification ou la détermination de la périodicité du choc dans des signaux de défauts de bague extérieure et intérieure et le défaut de bille. Ce travail s'intéresse aussi à l'étude de l'évolution de la fissure sur la bague externe du roulement.

Le choix de l'ondelette est un aspect primordial quant à la qualité des résultats. Chimentin (2007) a développé une ondelette adaptée pour des signaux de types réponses aux chocs avec une méthodologie adéquate réduisant considérablement le temps de calcul. Les résultats d'analyse par cette ondelette ont été comparés avec les résultats obtenus par l'ondelette de Morlet. Les formes des ondelettes sont très différentes. L'ondelette de Morlet est très compressée dû à une haute décroissance. Cette propriété montre qu'elle est très sensible aux singularités de la réponse aux chocs alors que l'ondelette adaptée est plus sensible à la réponse au choc.

La transformée en ondelettes trouve également son application dans le cadre de la démodulation du signal vibratoire et donne des résultats satisfaisants dans la détection précoce des défauts de roulements (Rubini & Meneghetti, 2001 ; Nikolaou & Antoniadis, 2002a ; Nikolaou & Antoniadis, 2002b).

La combinaison de différentes techniques s'avère souvent pertinente. Wang (2003) a conjugué avantageusement les transformations de Fourier et en ondelettes pour caractériser des défauts naissants et bien localisés sur les paliers à roulement.

1.4.8 Transformation en paquets d'ondelettes

La décomposition en paquets d'ondelettes (DPO) est similaire à la transformation en ondelettes discrète. Les deux sont des méthodes d'analyse multi-résolution. La DPO décompose simultanément les versions des détails (D_i) et les approximations (A_i) (figure 1.5), tandis que la Transformée en Ondelettes Discrète (TOD) décompose uniquement les

approximations. Aussi, la DPO a la même largeur de bande fréquentielle dans chaque résolution, propriété qu'on ne trouve pas dans la transformée en ondelettes discrète.

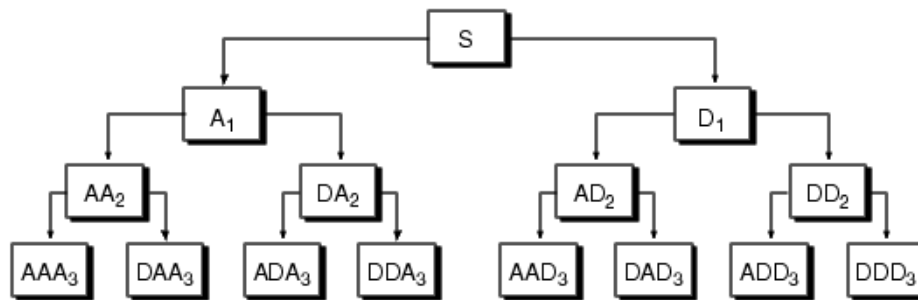


Figure 1-5: Arbre de décomposition en paquets d'ondelettes pour 3 trois niveaux de décomposition (Kerroumi, 2016).

La décomposition en paquets d'ondelettes a été utilisée par Li et al. (2013b) pour la détection des défauts multiples d'une boîte de vitesse. Dans cette étude, le signal est décomposé en 2^j coefficients (j est le niveau de décomposition). Puis, l'énergie, l'entropie, le skewness, et le kurtosis de chaque coefficient obtenu par DPO, ont été utilisés comme indicateurs pour la détection des défauts d'engrenages. Enfin une méthode de sélection appelée (locally linear embedding LLE) a été appliquée pour sélectionner les indicateurs les plus pertinents.

Le choix de niveau de décomposition en paquet d'ondelettes constitue un paramètre important dans la phase d'extraction d'indicateurs. Dans (Shen et al., 2013), il a été montré que seulement 3 niveaux de décomposition sont suffisants pour cette tâche.

1.4.9 Décomposition en mode empirique (EMD)

La décomposition en mode empirique, EMD (Empirical Mode Decomposition) a été proposée par Huang et al. (1998) pour l'analyse des signaux non stationnaires. Cette méthode est une technique adaptative, sans surveillance, et par laquelle tout ensemble de données peut être décomposé en un nombre fini de fonctions de modes intrinsèques dans le domaine temporel. L'extraction des composantes oscillantes appelées *modes empiriques* est non-linéaire, mais leur recombinaison est linéaire.

Cette méthode est une décomposition adaptée au signal, ne nécessitant pas d'informations a priori sur ce dernier. L'EMD n'a pas de formulation analytique, elle est définie par un algorithme et par un processus appelé *tamissage* (Mahgoun et al., 2011),

permettant de décomposer le signal en modes empiriques ou IMFs (Intrinsic mode functions). La décomposition est locale, itérative, séquentielle et entièrement pilotée par les données.

Pour calculer les IMFs, la procédure est comme suit :

- 1) Extraire les maxima et minima locaux du signal.
- 2) Calculer les enveloppes supérieure et inférieure par interpolation,
- 3) Calculer l'enveloppe moyenne locale $m(t)$ à partir des enveloppes supérieure et inférieure
- 4) soustraire l'enveloppe moyenne du signal d'entrée $h(t) = x(t) - m(t)$
- 5) Si $h(t)$ est un IMF, le résidu est $r(t) = x(t) - h(t)$ et le nouveau signal sera $x(t) = h(t)$,
- 6) Si $h(t)$ n'est pas un IMF le nouveau signal sera $x(t) = r(t)$,

Plusieurs applications de l'*EMD* ont été déjà réalisées dans le cadre du diagnostic de défauts des machines tournantes depuis plus d'une dizaine d'années (Yang et al., 2006 ; Dybala & Zimroz, 2012 ; Parey et al., 2006 ; Li et al., 2007).

De résultats encourageants ont été obtenus en diagnostic de palier à roulement. La propriété de l'*EMD* comme étant un filtre passe bande a été exploité par Pan & Tsao (2013) pour la détermination des IMFs contenant les fréquences de résonance avec précision avant d'appliquer l'analyse d'enveloppe. Une étude comparative entre cette méthode et l'analyse classique de l'enveloppe a été réalisée pour la détection des défauts multiples de roulements.

Dans Grasso et al. (2014), les auteurs proposent une approche basée sur le couplage de l'*EMD* avec l'Analyse en Composante Principale (*ACP*) pour la maintenance conditionnelle d'une machine de découpage par jet d'eau.

On retrouve une autre application de cette méthode dans le cadre du diagnostic des anomalies différentes du désalignement ou du balourd qui ne sont souvent ni linéaires ni stationnaires (Wu & Qu, 2008).

Une autre étude sur les roulements a été menée par Yang et al. (2007). Dans cette étude, le signal vibratoire a été décomposé en plusieurs fonctions de mode intrinsèque (en anglais : Intrinsic Mode Function ; IMF) en utilisant la décomposition en modes empirique (Empirical Mode Décomposition ; EMD), puis le spectre d'enveloppe a été calculé pour certains IMFs. Les amplitudes des pics autour des fréquences caractéristiques des défauts, extraits du signal d'enveloppe ont été utilisées pour construire le vecteur forme nécessaire à la tâche de classification.

Cyclo-stationnarité

L'analyse des signaux cyclostationnaires a pour origine le domaine des télécommunications où l'on peut transmettre des signaux dont les paramètres statistiques d'ordre 1 (moyenne, variance) et d'ordre 2 (fonction de corrélation) qui le gouvernement varient périodiquement.

D'un point de vue mathématique, la stationnarité ou la cyclostationnarité d'un signal peut être identifiée à partir de l'état de son autocorrélation (Breneur, 2002). L'autocorrélation d'un signal s'exprime de la façon suivante :

$$R_{xx}(t, \tau) = E\left[x(t + \tau/2) \cdot x^*(t - \tau/2)\right] \quad (1.23)$$

Le signal $x(t)$ est stationnaire à l'ordre 2 si son autocorrélation est indépendante de t .

Le signal $x(t)$ est purement cyclostationnaire si son autocorrélation dépend de t de façon périodique tel que : $R_{xx}(t, \tau) = R_{xx}(t + T, \tau)$. La fréquence $1/T$ alors associée est la fréquence de cyclostationnarité.

Cette théorie connaît beaucoup de développements et d'applications dans le domaine du diagnostic par analyse acoustique et vibratoire des machines tournantes (Lee & Lee, 2003 ; Randall, 2011 ; Randall et al., 2001 ; Bonnardot, 2004).

La corrélation spectrale permet d'étudier le lien existant entre les fréquences caractéristiques du système telles que la fréquence de rotation et la fréquence d'engrènement et ces harmoniques (Breneur, 2002 ; Capdessus, 1992 ; Bonnardot, 2005).

Une étude menée par Bouillaut & Sidahmed (2001) montre que l'analyse de cyclostationnarité, comparée au bispectre, présente bien des intérêts d'un point de vue 'souplesse' d'estimation et temps de calcul. Elle fournissait de meilleurs résultats pour le diagnostic de défauts sur les systèmes d'engrenage, et ce, quelle que soit la nature du défaut étudié. Capdessus et al. (2000) ont développé l'application de la théorie de la cyclostationnarité dans le but de diagnostic de défaut des dents d'engrenage dans le cas des signaux non stationnaires. Ils définissent les propriétés principales du processus cyclostationnaire comprenant l'évolution de la fonction de corrélation de la densité spectrale.

Dans Capdessus et al. (1995), il a été également montré que les liens qui apparaissent dans la corrélation spectrale, outil cyclostationnaire, permettent de détecter de manière plus précoce un défaut d'engrenage (les défauts apparaissant à certaines fréquences cycliques et rendent le signal de plus en plus cyclostationnaire).

1.5 Conclusion

Nous avons présenté dans ce chapitre la majorité des travaux touchant à l'analyse des signaux vibratoires par extraction d'indicateurs en vue de diagnostic de défauts dans les machines tournantes.

Les indicateurs temporels tels que la valeur efficace, le Kurtosis, le facteur crête, le facteur d'impulsion, etc ... permettent une première appréciation de l'état vibratoire d'une machine ou d'une installation. Malheureusement, ils trouvent leur limite pour un diagnostic précis quand il s'agit d'identifier la nature du défaut et surtout de localiser l'élément défaillant dans la machine. Les méthodes avancées de diagnostics telles que l'analyse d'enveloppe, donnent des résultats encourageantes quand il s'agit de localiser le défaut dans les machines tournantes.

La fiabilité du diagnostic par l'utilisation d'un ou des indicateurs (temporels et fréquentiels) seuls dépend essentiellement de la sensibilité de ces derniers face aux bruits et au changement du mode de fonctionnement de la machine comme la variation de la vitesse de rotation ou la variation de charge. C'est pourquoi, de plus en plus de travaux de recherche s'orientent vers l'utilisation simultanée d'un ensemble d'indicateurs par la méthode de classification de données.

La méthode de classification de données et plus généralement la méthode de reconnaissance des formes est une discipline issue de différents domaines telles que les mathématiques, les sciences de l'ingénieur, l'informatique et l'intelligence artificielle (Tan et al., 2005).

En dépit des efforts considérables consentis pour maîtriser le domaine de diagnostic des machines, les méthodes classiques de traitement de signal présentent encore des limitations. Ces limitations sont mises en évidence quand la machine observée est par exemple d'une cinématique complexe ou quand elle est sous des conditions de fonctionnement non stationnaires. De ce fait, il est primordial d'aller vers des nouveaux outils (Dong & Luo, 2013). Moosavian et al. (2012) qui permettent d'établir un diagnostic fiable quand c'est difficile autrement, et sans l'intervention d'un expert. Les méthodes de reconnaissance de formes furent des solutions acclamées dans la littérature de diagnostic de par leur autonomie (le diagnostic peut être établie sans l'intervention d'un agent), leur fiabilité, leur facilité d'interprétation...

Le chapitre suivant introduit les méthodes de reconnaissance des formes, leurs applications, leurs avantages et inconvénients, et des exemples pertinents des leurs mises en œuvres dans le diagnostic des machines tournantes.

Chapitre 2 :

Classification non supervisée dans le cadre de
diagnostic des machines tournantes

2 Classification non supervisée dans le cadre de diagnostic des machines tournantes

2.1 Introduction

Les techniques de classification font partie de la statistique exploratoire multidimensionnelle. Elles ont pour but d'expliquer la structure d'un ensemble de données importante, permettent ainsi de formuler des hypothèses à vérifier dans une étape ultérieure. Elles sont à distinguer des méthodes de classement qui ont un but explicatif ou prédictif.

La classification est une partie importante pour l'analyse des données. Elle consiste à regrouper les objets d'un ensemble de données en classes homogènes. Il existe deux types d'approches : la classification supervisée et la classification non supervisée. Ces deux se différencient par leurs méthodes et leurs buts. La classification supervisée (ang. Classification) est basée sur un ensemble d'objets L (appelé ensemble d'apprentissage) de classes connues. Le but étant de découvrir la structure des classes à partir de l'ensemble L afin de pouvoir généraliser cette structure sur un ensemble de données plus large. La classification non supervisée (ang. Clustering) consiste à diviser un ensemble de données D en sous-ensembles, appelés classes (ang. Clusters), de sorte que les objets soient les plus homogènes possible suivant un critère défini. Les critères les plus couramment utilisés sont la similarité entre les objets, la densité entre les classes et les mesures probabilistes. Les objets sont regroupés selon le critère qu'utilise la méthode de classification employée.

Dans le cadre de cette thèse, nous ne nous intéresserons qu'à la classification non supervisée. Le but de ce chapitre est d'introduire les notions et les concepts de base sur lesquels s'appuiera la suite de cette thèse, et de mettre en évidence la diversité qu'il existe parmi les différentes méthodes de classification non supervisées. Nous insistons particulièrement sur le problème de la définition d'un indice de ressemblance dans le but de pouvoir classer des individus et de les suivre dans le temps. La partie principale de ce chapitre sera consacrée à une synthèse bibliographique des différentes méthodes classiques de classification fournies dans la littérature. L'objectif recherché est d'introduire et de positionner nos contributions au regard de l'existant. Pour pouvoir apprécier la justesse des résultats obtenus par la classification, nous exposons les principaux critères d'évaluation des résultats d'un algorithme de classification non supervisée.

2.2 Systèmes de clustering

2.2.1 Objectif des méthodes de classification

Il s'agit de consulter des groupes, c'est-à-dire une partition de l'ensemble étudié: objets, variables, instances, cas, etc.

L'objectif général de la classification est donc d'explorer des données, de les réduire éventuellement, de les valider en infirmant ou en confirmant une hypothèse soutenue, d'agir sur des individus suivant la nature du groupe où ils se trouvent. Le regroupement d'objets ou d'individus en classes peut générer des hypothèses à tester dans une étape ultérieure.

La partition fournie par une méthode de classification peut être constituée soit de classes réelles (données hétérogènes, figure 2-1.a), soit de classes dites instrumentales (données homogènes, figure 2-1.b), qui sont obtenues par découpage d'un ensemble continu : il s'agit, dans ce dernier cas, d'une méthode généralisant la construction d'histogrammes en statistique unidimensionnelle.

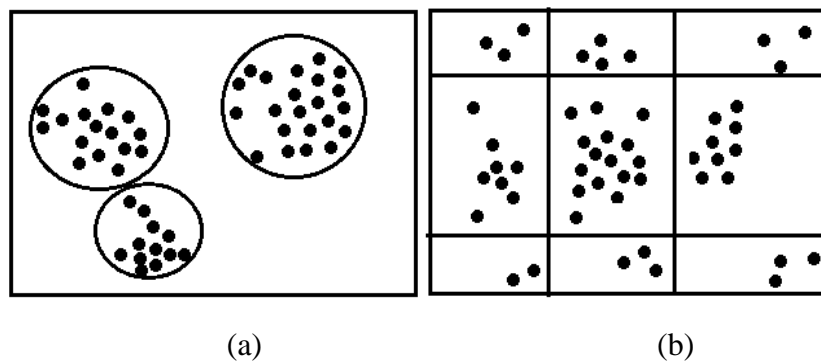


Figure 2-1:(a) Classes réelles et (b) classe instrumentale (Blansch , 2006)

2.2.1.1 Clusters

A la base, un cluster est un ensemble d' l ments. Cet ensemble est distinct des autres. Donc chaque  l ment d'un cluster a de fortes ressemblances avec les autres  l ments de ce m me cluster, et doit  tre diff rent des  l ments des autres clusters. C'est ce que l'on appelle : la forte similarit  intra-classe, et la faible similarit  inter-classe. Il y a donc une id e de recherche des groupes distincts. Les m thodes d'analyse de clusters sont des algorithmes non-supervis s, ils permettent de g n rer et de trouver des classes naturelles. Par exemple ce genre de m thodes d'analyse de clusters est utilis  dans le marketing pour d couvrir le profil de certains groupes de clients, et ainsi s'adapter   un march . Une m thode d'analyse de clusters doit se montrer fiable, donc elle doit pouvoir cr er des clusters bien distincts,  tre faiblement

sensible au bruit, mettre à jour des classes cachées, être insensible aux paramètres d'entrées. Il existe deux types de clusters (Blanché, 2006):

- les clusters durs qui sont totalement distincts les uns des autres, ainsi un élément d'un cluster n'est pas du tout dans un autre.
- les clusters mous comportent des éléments dont l'appartenance est pondérée, donc un élément peut être distribué parmi plusieurs clusters.

2.2.1.2 Propriétés d'un cluster

Les deux propriétés importantes définissant un cluster pertinent sont :

- sa cohésion interne (que les objets appartenant à ce cluster soient les plus similaires possibles)
- son isolation externe (que les objets appartenant aux autres clusters soient les plus éloignés possible).

Pour observer cela, plusieurs mesures sont associées à un cluster :

- sa densité (la masse d'objets par unité volumique)
- sa variance (le degré de dispersion des objets dans l'espace depuis le centre du cluster)
- sa dimension (typiquement son radius ou son diamètre)
- sa forme (hypersphérique/allongée/concave/convexe,...)
- sa séparation (par rapport aux autres clusters).

Un cluster forme un ensemble et cet ensemble occupe donc un espace. Pour pouvoir mesurer l'appartenance d'un élément à un cluster et pouvoir prendre une décision, il nous faut une fonction de mesure. Dans la littérature, on utilise beaucoup la distance de Minkowski (eq 2.1) (Nakache & Confais, 2004):

$$d(i, j) = \sqrt[q]{|x_{i1} - x_{j1}|^q + |x_{i2} - x_{j2}|^q + \dots + |x_{in} - x_{jn}|^q} \quad (2.1)$$

Plusieurs variantes sont utilisées, notamment avec $q = 1$ ou $q = 2$. D'autres distances sont utilisées parfois, telles que la paramétrique de Pearson (Hastie et al., 2009).

2.2.1.3 Différentes étapes d'un système de clustering

Typiquement, les systèmes de clustering se différencient par une fonction choisie pour évaluer la qualité du clustering, et par la stratégie de contrôle pour parcourir l'espace des

clusters possibles. Mais tous ces systèmes suivent le principe général traditionnel en clustering qui consiste à maximiser la similarité des observations à l'intérieur d'un cluster, et minimiser la similarité des observations entre clusters pour arriver à une partition de la base aussi pertinente que possible. Les différentes étapes d'une tâche de clustering sont données par la figure 2-2 :

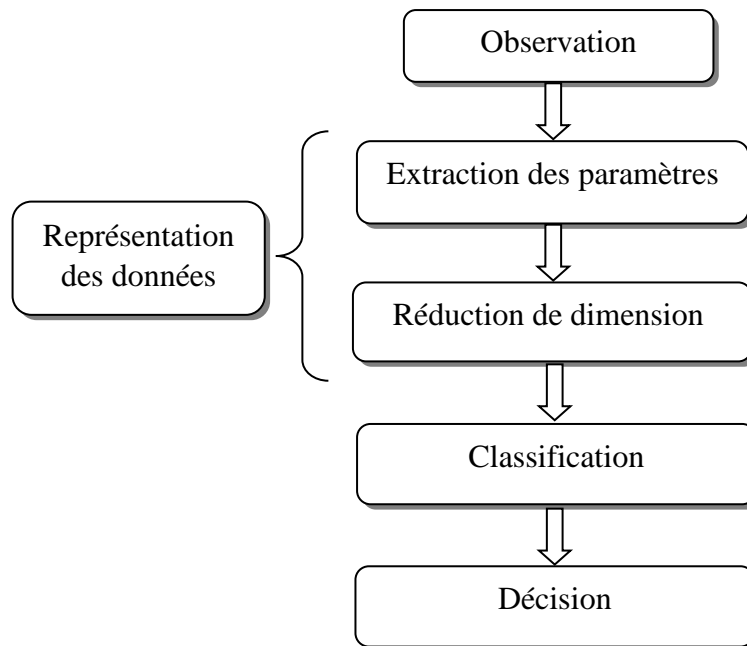


Figure 2-2: Structure générale de la méthode de reconnaissance de formes

2.2.1.3.1 Représentation des données : extraction et réduction de dimension

Lorsque les données issues de l'observation du fonctionnement d'un système sont représentées par des paramètres statistiques que l'on appelle aussi « indicateurs » ou « attributs », elles sont transformées en formes, c'est-à-dire un ensemble de points dans l'espace de représentation. Les groupes de formes similaires sont appelés « classes ». Si les indicateurs sont bien choisis, les classes sont bien discriminées et elles sont situées dans différentes régions de l'espace de représentation.

En règle générale, plus les paramètres du système de clustering permettent de discriminer les formes, meilleurs sont les résultats de la classification.

La première étape d'une méthode de clustering consiste donc à extraire ces paramètres, caractéristiques informatives de l'observation. Ces caractéristiques doivent être robustes et insensibles au bruit. L'une des difficultés principales liées à la représentation des données est sa dimension qui peut être très importante. Dans le cadre du diagnostic des défauts des

machines tournantes, l'extraction des caractéristiques fait appel à des techniques de traitement du signal classique ou avancées (voir chapitre 1).

Le problème de la dimension des données concerne le nombre et la qualité des variables descriptives caractérisant chacun des individus (des données issues de l'observation ou des paramètres). En effet, plusieurs paramètres peuvent être redondants ou non pertinents de l'ensemble de données et cela risque de conduire à la réduction de la précision de la classification et à une augmentation inutile des coûts de calcul.

La réduction de dimension est une étape importante dans le système de clustering et plus généralement, dans la méthode de reconnaissance de formes. Cette réduction de dimension repose sur deux types de méthode : la sélection ou la projection.

La sélection repose sur le choix des caractéristiques les plus discriminantes de l'ensemble des caractéristiques disponibles (Janecek et al., 2008), c'est-à-dire ceux qui permettront de discerner au mieux les classes tout en représentant la totalité ou le maximum d'informations véhiculés par les caractéristiques disponibles. Les méthodes de sélection choisissent le sous-ensemble de paramètres les plus informatifs (Janecek et al., 2008). Ce sous-ensemble doit bien évidemment remplacer l'ensemble des caractéristiques extraites sans engendrer une perte d'information, et permet d'atteindre les résultats avec une très bonne précision.

Quant à la projection, elle consiste à représenter les données dans un espace à dimension réduite. Les nouvelles données ainsi obtenues sont donc une combinaison linéaire ou non des données initiales. Parmi les techniques les plus utilisées dans la méthode de reconnaissance de formes, on peut citer l'analyse en composante principale (PCA) (Kambhatla & Leen, 1997), l'analyse en composantes principale à noyau (Kernel Principal Component Analysis KPCA) (Li et al., 2005), etc.

2.3 Classification

La méthode de classification consiste à regrouper les individus (données sélectionnée ou projetée) dans différentes classes ou catégories. Dans le cadre de la surveillance des machines tournantes, chaque classe caractérise un état de fonctionnement du système. Ce regroupement sera fait sur la base de la similitude des caractéristiques ou des indicateurs (Wang et al., 2012 ; Zhang et al., 2013 ; Zhang & Zhou, 2013). La classification est effectuée en sorte que deux contraintes soient respectées : (1) maximisation de la variance au sein du même groupe ou classe et (2) minimisation de la variance entre les différents groupes/classes. Chaque classe a

ses propres caractéristiques et elle peut être définie par un modèle qui est représenté par une fonction d'appartenance. Chaque individu est associé à une classe. Cette association ou attribution à la bonne classe se fait en mesurant la similarité ou la différence entre les caractéristiques de chaque individu et les caractéristiques des individus déjà classifiés dans ces classes. La similarité ou la différence est souvent mesurée par une distance ou par un degré d'appartenance. Ces mesures sont souvent dérivées de certaines fonctions discriminantes dans la reconnaissance des formes statistique. Les mesures de distance couramment utilisées sont la distance euclidienne (Ghaemi et al., 2014.), la distance de Mahalanobis (Staszewski et al., 2002.), la distance bayésienne (Goumas et al., 2001.). Plusieurs méthodes de classification ont été utilisées dans le cadre de diagnostic de défaut des machines tournantes. On peut citer K-means pour classifier des signaux enregistrés pour des roulements avec différents défauts (Yiakopoulos et al., 2011.).

2.3.1 Propriétés des techniques de clustering

Plusieurs propriétés peuvent être associées aux différentes techniques de clustering :

➤ **Ascendant versus descendant :**

Une méthode de classification ascendante commence avec autant de clusters que d'objets, puis va concaténer successivement les clusters jusqu'à ce qu'un critère d'arrêt soit satisfait. A l'inverse, une méthode de classification descendante va démarrer avec un cluster réunissant tous les objets, puis va diviser les clusters jusqu'à ce qu'un critère d'arrêt soit satisfait (Blanché, 2006).

➤ **Déterministe versus stochastique :**

Avec les mêmes données en entrée, un algorithme déterministe exécutera toujours la même suite d'opérations et fournira donc toujours le même résultat. A l'inverse, une méthode stochastique pourra donner des résultats différents pour des données en entrée identiques, car elle permet l'exécution d'opérations aléatoires.

Les algorithmes stochastiques sont donc moins précis mais moins coûteux. C'est pourquoi ils sont utilisés lorsqu'on a à faire face à de larges bases de données.

➤ **Incrémental versus non-incrémental :**

Une méthode incrémentale va être exécutée de façon continue, et va intégrer les données au fur et à mesure de leur arrivée dans l'algorithme (Blanché, 2006).

A l'inverse, une méthode non-incrémentale va considérer un ensemble de données fournies en entrée, et sera exécutée sur cet ensemble de données. Si, par la suite, une nouvelle donnée devait être fournie en entrée de l'algorithme, celui-ci devrait être relancé à nouveau.

➤ **Hard versus Fuzzy :(ou dure versus floue)**

Une méthode Hard associe à chaque objet un unique cluster alors qu'une méthode Fuzzy associe à chaque objet un degré d'appartenance à chaque cluster.

A noter qu'un Fuzzy clustering peut être converti en un Hard clustering en assignant chaque donnée au cluster dont la mesure d'appartenance est la plus forte (Blansché, 2006).

2.3.2 Evaluation de la qualité d'une classification non supervisée

Avant de décrire les principales méthodes, il est nécessaire de décrire les différents critères d'évaluation d'une classification non supervisée. En effet, quelque soit le type du résultat que l'on souhaite construire, il existe de nombreuses classifications possibles pour un ensemble de données. Par exemple, dans le cas d'une classification dure d'un ensemble D et N objets en K classes, il existe K^N classifications possibles. Or, sélectionner la « meilleure » classification parmi ces K^N possibilités n'a de sens que si l'on dispose d'un critère d'évaluation de celles-ci. Malheureusement, contrairement à la classification supervisée, il n'existe pas de critère naturel. En effet, dans le cas des algorithmes de classification supervisée, il est assez aisé de vérifier la validité des règles de classification obtenues, étant donné que la classe de chaque objet de l'ensemble d'apprentissage est connue. En revanche, un algorithme non supervisée n'a pas cette information à disposition. Des critères statistiques, évaluant la cohérence des classes ont dû être définis. Ces critères évaluent la pertinence de la classe par rapport au paradigme utilisé, par exemple si les objets d'une classe sont similaires entre eux et différents des objets des autres classes (Blansché, 2006).

De fait, l'évaluation d'une classification réalisée par un algorithme non supervisé a engendré un grand nombre de travaux (Zhu et al., 2009 ; Bezdek & Pal., 1998 ; Halkidi et al., 2001 ; Chavent & Lechevallier, 2002). Nous allons présenter ici une partie des critères qui ont été proposés dans la littérature. Ces critères d'évaluation se basent sur différentes notions. On distinguera principalement les critères basés sur une mesure de distance, les critères spécifiques aux classifications floues, les critères spécifiques aux méthodes de classification probabiliste et les critères basés sur une mesure de rééchantillonnage des données.

Définition 2.1 (erreur au carré)

L'erreur au carré d'une classification non supervisée est définie dans Nakache & Confais (2004) par :

$$\text{cost}_{km}(c, C) = \sum_{1 \leq k \leq K} \sum_{o \in C_k} d(o, c_k)^2 \quad (2.2)$$

Avec $c = \{c_1, \dots, c_n\}$ les centres des classes et $C = \{C_1, \dots, C_n\}$ les classes et o est l'objet de la k -ième classe.

L'erreur au carré (ang. square error) est une des mesures de compacité les plus couramment utilisés dans le cas des classifications dures; une faible valeur indique une forte compacité des classes.

Cette définition peut être étendue à une classification floue en pondérant les distances par le degré d'appartenance aux classes. Une faible valeur indique toujours une forte compacité des classes (Nakache & Confais, 2004).

Définition 2.2 (erreur au carré (floue))

L'erreur au carré d'une classification non supervisée floue est définie dans par :

$$\text{cost}_{fcm}(c, \mu) = \sum_{1 \leq k \leq K} \sum_{o \in C_k} (\mu_k(o))^f d(o, c_k)^2 \quad (2.3)$$

Avec $C = \{C_1, \dots, C_K\}$ les classes, $c = \{c_1, \dots, c_K\}$ leurs centres et $\mu = \{\mu_1, \dots, \mu_K\}$, où $\mu_k(o)$ le degré d'appartenance de l'objet o à la k -ième classe et f un réel tel que $f > 1$.

Ce critère tient compte à la fois de la compacité et de la séparabilité des classes dans le cas d'une classification dure. Ce critère est cependant rarement utilisé car il est coûteux en temps de calcul. Une forte valeur indique une forte compacité des classes et une forte séparation des classes.

Définition 2.3 (indice de Dunn)

L'indice de Dunn (1974b) d'une classification en K classes est défini par :

$$D = \frac{\min_{k, k' \in [1; K]} \left\{ \min_{o \in C_k, o' \in C_{k'}} d(o, o') \right\}}{\max_{k'' \in [1; K]} \left\{ \max_{o, o' \in C_{k''}} d(o, o') \right\}} \quad (2.4)$$

Avec $d(o,o')$ est la distance entre les classes C_k et $C_{k'}$. o et o' des points appartenant respectivement aux classes C_k et $C_{k'}$.

Ce critère tient compte à la fois de la compacité et de la séparabilité des classes dans le cas d'une classification dure. Une faible valeur indique une forte compacité des classes et une forte séparation des classes.

Définition 2.4 (indice de Davies et Bouldin)

L'indice de Davies et Bouldin (Davies & Dunn, 1979) est défini par :

$$DB = \frac{1}{K} \sum_{k=1}^K \max_{k' \in \{1; K\} \setminus \{k\}} \left\{ \frac{S(C_k) + S(C_{k'})}{d(c_k, c_{k'})} \right\} \quad (2.5)$$

$$\text{Avec } S(C_k) = \frac{1}{|C_k|} \sum_{o \in C_k} d(o, c_k) \quad (2.6)$$

$S(C_k)$ est la distance intra-classe de la classe K.

Ce critère mesure la séparabilité des classes dans le cas d'une classification dure. Une valeur élevée sur cet indice indique une forte séparation des classes (Elghazel, 2007).

Définition 1.5 (Statistique de Hubert Modifiée)

La statistique de Hubert Modifiée est présentée dans Hubert et al. (1985) par :

$$\Gamma = \sum_{o, o' \in D} d(o, o') S(o, o') \quad (2.7)$$

Avec $S(o, o') = 1$ si o et o' sont dans la même classe et $S(o, o') = 0$. Sinon, il est possible de normaliser cet indice pour qu'il prenne ses valeurs entre -1 et 1 :

$$\bar{\Gamma} = \frac{1}{M} \times \frac{1}{\sigma_d \sigma_s} \times \sum_{o, o' \in D} (d(o, o') - \mu_d)(S(o, o') - \mu_s) \quad (2.8)$$

Où $M = \frac{N(N-1)}{2}$, et μ_d et μ_s (respectivement σ_d et σ_s) représentent la moyenne (respectivement la variance) de d et de S .

Définition 2.6 (indice de Xie et Beni)

L'indice de Xie et Beni (1991) d'une classification floue en K classes est défini par :

$$XB = \frac{\sum_{k=1}^K \sum_{o \in D} \mu_k(o)^2 \times d(o, c_k)^2}{n \times \min_{o, o' \in D} d(o, o')^2} \quad (2.9)$$

Avec $\sum_{k=1}^K \mu_k(o) = 1, \forall o \in D$.

μ_d représentent la moyenne de d.

L'indice de *Xie* et *Beni* peut également être défini en fonction du critère de qualité de *fuzzy-C-means* par :

$$XB = \frac{\text{cost}_{cm}}{n \times d_{\min}^2} \quad (2.10)$$

Avec $d_{\min} = \min_{o, o' \in D} d(o, o')$.

L'indice de compacité de *Xie* et *Beni* peut enfin être étendu à une classification dure par :

$$XB = \frac{\text{cost}_{kn}}{n \times d_{\min}^2} \quad (2.11)$$

Définition 2.7 (Distance inter et intra classe)

De nombreux critères de qualité sont basés sur une mesure de distance entre les objets. Ils utilisent des notions de *compacité* des classes (eq. 2.12, critères intra-classes)

$$\text{Intra} = \frac{1}{n} \sum_{C=1}^p \sum_{i=1}^{n_c} d^2(X_i, C_C) \quad (2.12)$$

Ou de *séparabilité des classes* (eq. 2.13 critère inter-classes), ou les deux à la fois.

$$\text{Inter}(C_i) = \frac{1}{n} \sum_{i=1}^p n_c d^2(C_c, c_G) \quad (2.13)$$

La compacité est définie en fonction de la distance entre les objets d'une même classe ou en fonction de la distance entre les objets d'une classe et le centre de la classe. La séparabilité des classes est définie en fonction de la distance entre objets de classe différentes ou en fonction de la distance entre les centres des classes (Ghribi et al., 2010 ; Lebart et al., 1979).

Définition 2.8 (Indice de silhouette)

L'indice de silhouette (Rousseeuw, 1987) est défini pour tout individu X_i de l'ensemble X par la formule suivante

$$\forall X_i \in X, S(X_i) = \frac{b(X_i) - a(X_i)}{\max(a(X_i), b(X_i))} \quad (2.14)$$

Avec $a(X_i)$ est la dissimilarité moyenne entre l'individu X_i et tous les autres individus de la classe à laquelle il appartient $C(X_i)$

$$a(X_i) = \frac{1}{n_{c_i}} \sum_{X_j \in C(X_i), X_j \neq X_i} d(X_j, X_i) \quad (2.15)$$

Et $b(X_i)$ est le minimum des dissimilarités moyennes entre l'individu X_i et tous les autres individus des classes de la partition P .

$$\begin{aligned} b(X_i) &= \min_{C \in P, C \neq C(X_i)} d(X_i, C) \\ &= \frac{\min_{1 \leq i < j \leq n} \Delta_{i,j}}{\max_{1 \leq k \leq n} \Delta'(k)} \end{aligned} \quad (2.16)$$

Cet indice travaille à l'échelle microscopique, c'est à dire qu'il s'intéresse aux objets en particulier et non pas aux classes. Le but de l'indice de Silhouette est de vérifier si chaque objet a été bien classé. Pour cela, et pour chaque élément i de la partition, l'indice de silhouette est calculé. S'il est proche de 1, cela signifie que l'objet est bien classé : la distance qui le sépare de la classe la plus proche est très supérieure à celle qui le sépare de sa classe. Par contre, si $S(X_i)$ est proche de -1, cela veut dire que l'objet est mal classé. Mais si $S(X_i)$ est proche de 0 alors il pourrait également être classé dans la classe la plus proche (Nakache & Confais, 2004 ; Lebart et al., 1979).

Synthèse

La définition d'un critère d'évaluation de la qualité d'une classification non supervisée dépend de la définition que l'on se fait de la structure d'une classe. En effet, certains critères sont spécifiques à une définition précise des classes, comme par exemple l'erreur au carré qui implique que les classes sont spécifiques dans l'espace de données. De nombreuses méthodes de classification sont basées sur l'optimisation de l'un ou de l'autre de ces critères d'évaluation, ce qui les limite à une certaines définition des classes.

On remarque également que de nombreux critères ne sont pas normalisés, l'ordre de grandeur pouvant varier considérablement d'un jeu de données à l'autre, ce qui rend plus difficile leur interprétation.

2.3.3 Différentes représentation des résultats de la classification

Il existe différentes façons de représenter le résultat d'une classification selon qu'il y ait des chevauchements entre les classes ou non (on distingue les classifications dures, douces et floues), et selon qu'il y ait des objets non classés. Par ailleurs, le résultat d'une classification peut également être représenté sous la forme d'une structure plate ou d'une hiérarchie de classes.

2.3.3.1 Classification dure, douce et floue

La plus simple façon de représenter les résultats d'un algorithme de classification non supervisée est une *classification dure* (ang. hard clustering). Dans une classification dure, chaque objet appartient à une et seule classe (Blansché, 2006). L'ensemble de données D , c'est-à-dire que $\bigcup_{k=1}^K C_k = D$ et $C_k \cap C_{k'} = \emptyset$ pour $k \neq k'$. On notera $C(o)=k$ le fait que $o \in C_k$.

Les méthodes de classification non supervisée produisant ce type de résultat sont appelées méthodes de partitionnement.

Exemple :

Si l'on considère l'ensemble de données présenté sur la figure 2-3, un algorithme de partitionnement en quatre classes produira les classes présentées sur la figure 2-3.a. Un algorithme de partitionnement en trois classes regroupera les classes 1 et 2 et produira les classes présentées sur la figure 2-3.b.

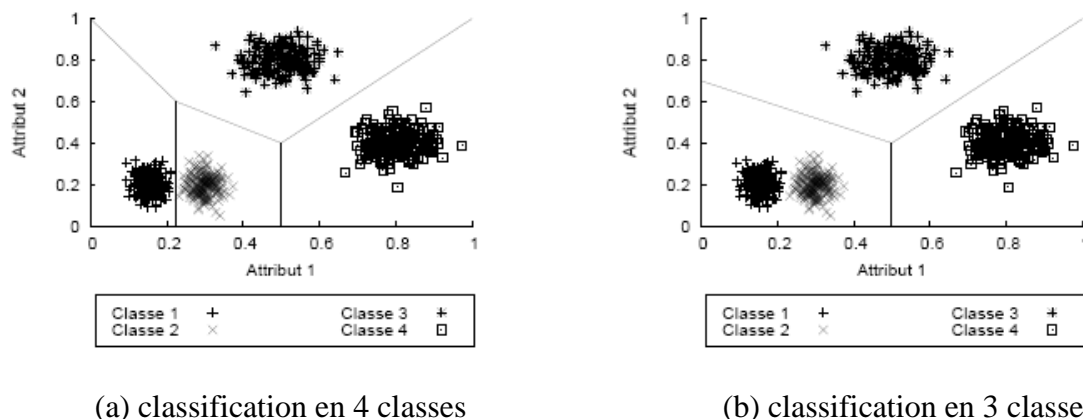


Figure 2-3: Classification par partitionnement (Blansché, 2006)

Ce type de résultats à l'avantage d'être facilement interprétable par un utilisateur, mais il est parfois nécessaire de donner plus de flexibilité à la définition des classes. En effet, il se peut que des objets se distinguent trop des autres. Ainsi, dans une *classification dure partielle* (ang. partial hard clustering), chaque objet appartient à au plus une classe (Nakache & Confais, 2004). Une telle classification en K classes $C = \{C_1, \dots, C_K\}$ vérifie néanmoins $C_k \cap C_{k'} = \emptyset$. On appelle objet atypique (ang. outlier) un objet o tel que $o \notin C_k, \forall C_k \in C$.

Dans une *classification douce* (ang. soft clustering), un objet peut appartenir à une ou plusieurs classes. L'ensemble de données D est donc divisé en un ensemble de classes $C = \{C_1, \dots, C_K\}$ de telle sorte que $\bigcup_{k=1}^K C_k = D$

Une *classification douce partielle* (ang. partial soft clustering) peut alors être définie par une classification dans laquelle un objet peut appartenir à une ou plusieurs classes ou à aucune d'entre elles. L'ensemble de données D est donc divisé en un ensemble de classes $C = \{C_1, \dots, C_K\}$ sans autre contraire que celle du critère de regroupement utilisé par la méthode de classification.

Dans une *classification floue* (ang. fuzzy clustering), chaque objet o appartient à chacune des classes avec un certain degré d'appartenance $\mu_k(o)$ (chaque classe est définie par une fonction d'appartenance μ_k). On note l'ensemble des fonctions d'appartenance $\mu = \{\mu_1, \dots, \mu_K\}$. Ces méthodes ont l'avantage de tenir d'avantage compte des objets stéréotypiques d'une classes par rapport aux objets atypiques ou à cheval entre deux classes. De nombreuses méthodes imposent $\sum_{k=1}^K \mu_k(o) = 1$.

Une classification floue reste cependant difficilement interprétable par un utilisateur et est souvent transformée en classification dure en affectant chaque objet o à la classe C_k telle que $\mu_k(o) = \max_{\mu_h \in \mu} \mu_h(o)$ (Nakache & Confais, 2004).

2.3.3.2 Classification hiérarchique

La plupart des méthodes représentent les résultats de classification sous forme d'une structure plate.

Il est cependant naturel de représenter la connaissance sous forme de hiérarchie de classes ou de concepts. Par exemple, une classe de végétation découverte dans une image de télédétection peut-être divisée en différentes sous classes selon le type de végétation. Dans une *classification hiérarchique*, une classe peut être divisée en plusieurs sous-classes, l'ensemble des classes formant alors une hiérarchie (représentée par un arbre). Un objet appartient généralement à une et une seule classe-feuille mais aussi à la classe mère de celui-ci, etc. les algorithmes de classification hiérarchique permettent de construire ce type de résultats. On distingue deux types d'approches : les méthodes ascendantes ou agglomératives et les méthodes descendantes ou divisives. Les méthodes agglomératives partent d'un grand nombre de classes (éventuellement une classe par objet) et fusionne les classes similaires entre elles. Les méthodes divisives partent de l'ensemble de données et le divisent en classes qui sont alors ensuite divisées récursivement (Nakache & Confais, 2004).

Exemple :

Si l'on considère l'ensemble de données présenté sur la figure 2-3, un algorithme de classification hiérarchique produira la hiérarchie de classes présentées sur la figure 2-4 la racine de l'arbre correspond à l'ensemble de données à classifier.

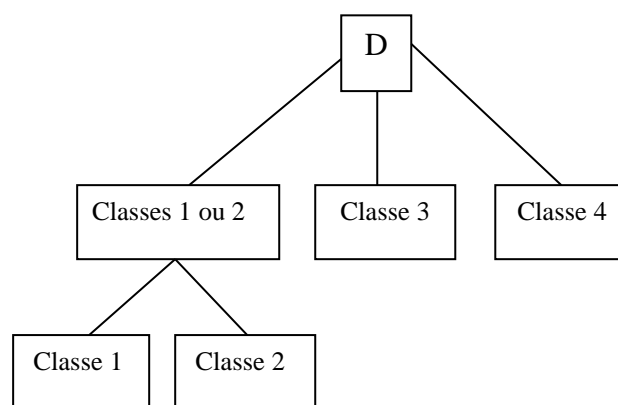


Figure 2-4: Classification hiérarchique (Blansché, 2006).

Synthèse

Nous avons vu que la classification non supervisée renvoie à un grand nombre d'approches très différentes les unes des autres, dans la façon de représenter les résultats et dans la définition même de ce qu'est une classe.

Le résultat d'une classification peut prendre la forme d'une structure plate ou d'une hiérarchie de classe. Un objet peut appartenir à plusieurs classes à la fois (éventuellement avec un certain degré d'appartenance) ou à aucune d'entre elles, selon le type du résultat proposé par la méthode de classification.

Une classe peut être définie par des prototypes, c'est-à-dire des objets stéréotypiques de la classe, par une relation de voisinage, par une loi de probabilité ou par une description des objets qui la composent.

2.4 Différentes méthodes de Partitionnement ou de clustering

Après avoir présenté les différents types de résultats d'un algorithme de classification, nous allons présenter les principales familles de méthodes de regroupement des données dans des classes (Jain et al., 1999) :

- Les méthodes basées sur une distance
- Les méthodes basées sur une grille
- Les méthodes probabilistes
- Les méthodes hiérarchiques

2.4.1 Méthodes basées sur une distance

De nombreuses méthodes de regroupement de données se basent sur une mesure de distance ou de similarité entre les objets. Ces méthodes se basent sur l'idée que des objets proches appartiennent à une même classe. Trois types d'approches utilisent la notion de distance pour construire les classes : les méthodes basées sur un prototype, les méthodes neuronales, les méthodes basées sur les Séparateurs à Vaste Marge (SVMs) et les méthodes basées sur la densité.

2.4.1.1 Méthodes basées sur des prototypes

Les méthodes basées sur des prototypes (*ang. Prototype-based methods*) sont des méthodes qui définissent les classes en fonction d'un objet représentatif de la classe. Cet objet

représentatif est appelé centroïde s'il s'agit du centre de gravité des objets appartenant à la classe ou médoïde s'il s'agit d'un objet particulier de la classe (Fouchal, 2011).

La méthode K-means ou C-means (ou k-moyenne) et Fuzzy-C-means sont les algorithmes de classification les plus populaires vue leur simplicité d'implémentation ainsi que leur applicabilité à des données de grande taille, sans problème de mémoire et temps de calcul. Ces méthodes ne permettent généralement que de trouver des classes convexes dans l'espace de données considéré. Cependant, la méthode présente l'inconvénient du choix des paramètres initiaux parlant des centres initiaux et la connaissance à priori du nombre de classe. Nous présentons l'algorithme des k-moyennes en figure 2-5.

Algorithme K-means :

Entrées : k le nombre de clusters désirés, d une mesure de dissimilarité sur l'ensemble des objets à traiter X ;

Sorties : une partition $C = \{C_1, C_2, \dots, C_k\}$.

Etape 0 : 1. Initialisation par tirage aléatoire dans X , de k centres $x_{1,0}^*, \dots, x_{k,0}^*$

2. constitution d'une partition initiale $C_0 = \{C_1, C_2, \dots, C_k\}$ par allocation de chaque objet $x_i \in X$ au centre le plus proche :

$$C_1 = \left\{ x_i \in X \mid d(x_i, x_{1,0}^*) = \min_{h=1, \dots, k} d(x_i, x_{h,0}^*) \right\} \quad (2.17)$$

3. calcul des centroïdes des classes k obtenues $x_{1,1}^*, \dots, x_{k,1}^*$.

Etape t : 4. Constitution d'une nouvelle partition $C_t = \{C_1, C_2, \dots, C_k\}$ par allocation de chaque objet $x_i \in X$ au centre le plus proche :

$$C_t = \left\{ x_i \in X \mid d(x_i, x_{1,t}^*) = \min_{h=1, \dots, k} d(x_i, x_{h,t}^*) \right\} \quad (2.18)$$

5. calcul des centroïdes des classes obtenues $x_{1,t+1}^*, \dots, x_{k,t+1}^*$.

6. répéter les étapes 4 et 5 tant que des changements s'opèrent d'un schéma C_t à un schéma C_{t+1} ou jusqu'à un nombre T d'itérations.

7. retourner la partition finale C_{finale}

Figure 2-5: Algorithme des k-moyennes (Fouchal, 2011)

A partir d'un tirage aléatoire de k "centres" dans X , l'algorithme des k-moyennes procède par itérations de super-étapes d'allocations des objets x_1, \dots, x_{ng} aux centres

(initialement les graines), suivies du calcul de la position des nouveaux centres, dits "mobiles". On peut montrer que, par ce processus, le critère de variance intra-cluster converge vers un minimum local, dépendant de l'initialisation de l'algorithme. Rappelons que ce critère est défini par :

$$V(t) = \sum_{h=1}^k \sum_{i=1}^n d(x_i, x_{h,t}^*)^2 \quad (2.19)$$

où $x_{h,t}^*$ désigne le centroïde du cluster C_h à l'étape t .

Cette méthode a été utilisée dans le cadre du diagnostic des défauts mécaniques. C.T. Yiakopoulos et al. (2011) ont exploité la méthode k-means pour le diagnostic automatisé des roulements défectueux. Pour remédier à la sensibilité de la méthode au choix des paramètres initiaux, les centres de départ sont sélectionnés à l'aide des caractéristiques extraites à partir de signaux simulés, résultant d'un modèle bien établi pour le comportement dynamique des défauts des éléments roulant de roulement. Le diagnostic a été réalisé en deux étapes. En premier lieu la présence du défaut sur tous les éléments de roulement est envisagée. La deuxième étape consiste à l'identification du type de défaut (par exemple défauts sur les éléments roulant). La méthode a permis de classifier les données en quatre classes : sain, défaut bague intérieure, défaut bague extérieure et éléments roulants. La méthode a été testée avec des données enregistrées sur un banc expérimental et sur des données enregistrées sur un site industriel. Chaque cas comprend des mesures successives de différents types de défauts de roulements. Dans tous les cas, la méthode présente un pourcentage de succès de 100%.

Dans Jiang et al. (2013), une projection des données dans un nouvel espace par une fonction noyau a été utilisée avant d'appliquer la méthode k-means. Cela a permis de réduire le nombre d'itérations et d'améliorer légèrement la précision de classification de 91% à 95%.

L'hybridation avec les algorithmes génétiques (AG) est une direction intéressante pour surmonter l'inconvénient de la classification par k-means de converger vers des optimums locaux (Ghaemi et al., 2014). L'algorithme génétique est appliqué pour le prétraitement des données avant de les classifier. Cette combinaison a permis d'augmenter la précision de la classification en quatre classes : sain, défaut bague intérieur, défaut bague extérieure, défaut élément roulant. Cette hybridation a donné une amélioration de la précision de classification de (64% à 97%) à (91% à 100%).

2.4.1.2 Les K plus proches Voisins (K- Nearest Neighbors)

C'est la première méthode de classification qui ne suppose pas un type de distribution des données et qui propose une analyse de similitude entre des données en utilisant la distance entre elles (Fix, 1951). L'algorithme fait un calcul de distance entre tous les individus et chaque objet est classé dans le groupe où se trouvent ses K plus proches voisins.

Les k plus proches voisins sont des individus de l'ensemble des données d'entraînement. Habituellement, la distance employée pour l'analyse est la distance euclidienne. Normalement le nombre des K voisins est un ou trois, mais cette valeur dépend de chaque application (Duda et al., 2001).

2.4.2 Méthodes neuronales

Les réseaux de neurones s'inspirent des neurones biologiques et les systèmes neuronaux, les réseaux de neurones sont aujourd'hui très utilisés dans les problèmes d'apprentissage.

Les réseaux neuronaux artificiels (Artificial neural network ANN) ont été initialement développés selon le principe élémentaire de l'opération du système neuronal (humain). Depuis lors, une très grande variété de réseaux a été construite. Tous les réseaux neuronaux artificiels sont composés d'unités simples (neurones), et des connexions reliant les unités entre elles, l'ensemble (les unités et leurs connexions) détermine le comportement du réseau. L'ANN peut être vue comme un groupe interconnecté des unités simples qui utilisent un modèle mathématique ou informatique pour le traitement de l'information (Kerroumi, 2016). L'ANN est un système adaptatif qui modifie sa structure basée sur l'information qui circule à travers le réseau (voir figure 2-6).

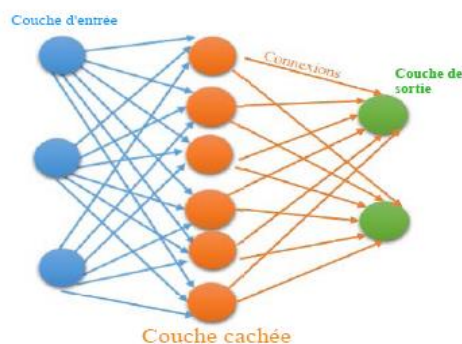


Figure 2-6: L'architecture d'un réseau de neurones (Kerroumi, 2016)

Les réseaux de neurones artificiels sont construits sur une architecture semblable, en première approximation, à celle du cerveau humain. Le réseau reçoit les informations sur une

couche réceptrice de « neurones », traite ces informations avec ou sans l'aide d'une ou plusieurs couches « cachées » contenant un ou plusieurs neurones et produit un signal (ou plusieurs) de sortie. Chaque neurone, qu'il appartienne à la première couche (réceptrice), aux couches cachées ou à la couche de sortie, est lié aux autres neurones par des connexions (similaires aux synapses du cerveau) auxquelles sont affectés des poids (eux même assimilables aux potentiels synaptiques). Les réseaux de neurones peuvent employer les deux types d'apprentissage : supervisé et non supervisé (exemple la carte auto-adaptive). Dans le cas des réseaux de neurones avec apprentissage supervisé (Perceptron, Adaline, etc.), on présente au réseau des entrées et en même temps les sorties que l'on désirerait pour cette entrée. Le réseau doit alors se reconfigurer, c'est-à-dire calculer ses poids afin que la sortie qu'il donne corresponde bien à la sortie désirée. Ils peuvent être considérés comme des systèmes informatiques massivement parallèles constitués d'un très grand nombre de processeurs simples (ou unité de calcul) avec de nombreuses interconnexions (Ghahramani, 2003). Les avantages des réseaux de neurones sont leurs apprentissages adaptifs, leurs auto-organisations et leurs capacités de tolérance aux pannes (Jain et al., 2000).

Dans les méthodes d'apprentissage supervisé, les réseaux de neurones artificiels (ANN) ont été largement appliqués pour la détection automatisée et le diagnostic des conditions de la machine (Dellomo, 1999 ; McCormick & Nandi, 1997 ; Nandi, 2000 ; Samanta et al., 2001 ; Jack & Nandi, 2002 ; Samanta et al., 2003 ; Rafiee et al., 2007). Cependant, les approches traditionnelles de réseau de neurones ont des limites sur la généralisation donnant naissance à des modèles qui peuvent sur-apprendre aux données de formation. Cette lacune est due aux algorithmes d'optimisation utilisés dans ANNs pour la sélection des paramètres d'entrée du réseau de neurones. Un système expert basé sur le réseau de neurones a été appliqué dans (Yang et al., 2012) pour les défauts de boîte de vitesses dans la turbine éolienne. Une autre application de cette méthode a été présentée dans (Qian et al, 2008). Dans Li et al. (2013b), un classificateur k-plus proche voisin (FKNN) a été proposé pour l'identification d'un motif de défaut d'une boîte de vitesses. Pour évaluer l'efficacité de l'approche diagnostique proposée, deux cas d'études ont été réalisés dont l'un porte sur le diagnostic des défauts d'engins, et l'autre sur le diagnostic des défauts de roulements de la boîte de vitesses.

Kumar et al. (2013) ont choisi d'utiliser le réseau de neurones de type perceptron multicouches dans le but de classifier des caractéristiques vibratoires en trois classes : sain, défaut bague intérieure, défaut bague extérieure. Le réseau de neurones a été testé avec trois types d'algorithme d'apprentissage : *trainbfg*, *trainscg* et *trainlm*. La comparaison des trois

algorithmes a montré que *trainbfg* est supérieur aux deux autres avec un taux de reconnaissance de 86%, ce taux augmente par l'utilisation des données débruités. Ce réseau de neurones a été combiné avec la méthode des ondelettes.

Dans Yu et al. (2006), un réseau de neurones de type perceptron multicouches a été employé et combiné avec la méthode *EMD* où l'entropie de chaque *IMF* alimente les nœuds d'entrées. Le réseau a été formé pour classifier les données en trois classes : sain, défaut bague intérieure, défaut bague extérieure.

Une ondelette de Morlet modifiée associée à des réseaux de neurones (WNN) est présentée par Guo et al. (2005) pour la détection des défauts de roulement. La fonction d'ondelette de Morlet modifiée est utilisée comme fonction d'activation des nœuds d'ondelettes dans la couche cachée. Le spectre de fréquence des données est calculé et introduit dans une phase d'apprentissage, dans laquelle 6 défauts et 7 bandes de fréquences sont sélectionnées pour former un vecteur de caractéristiques. Les approches WNN ont été testées sur les machines tournantes et comparées aux techniques BP (algorithme back-propagation). Les résultats des tests montrent que l'approche proposée de Morlet WNN modifié nécessite beaucoup moins d'époques d'entraînement et présente un taux de convergence et une précision de diagnostic plus élevés que la méthode BP. Li et al. (2009) quant à lui, a utilisé le cepstre pour le diagnostic des défauts d'engrenages. Dans cette étude, cet indicateur est utilisé comme paramètre d'entrée aux Réseaux de Neurones Artificiels (RNA).

Wang et al. (2017) ont proposé une nouvelle méthode appelée réseau neuronal convolutif profond adaptatif (CNN) pour le diagnostic des défauts de roulement. Tout d'abord, pour se débarrasser de l'extraction manuelle des entités, le modèle CNN profond est initialisé pour l'apprentissage automatique des fonctionnalités. Deuxièmement, pour s'adapter aux différentes caractéristiques du signal, les principaux paramètres du modèle CNN profond sont déterminés à l'aide d'une méthode d'optimisation des essaims de particules. Troisièmement, pour évaluer la capacité d'apprentissage des caractéristiques de la méthode proposée, l'incorporation de t-distribués stochastiques voisins (t-SNE) est également adoptée pour visualiser le processus d'apprentissage des entités hiérarchiques. Les résultats confirment que la méthode proposée est plus efficace et robuste que les autres méthodes intelligentes.

Dans Tyagi (2008), l'auteur a utilisé plusieurs indicateurs statistiques comme entrées au RNA et SVM pour la détection automatique des défauts de roulements. Ces indicateurs ont été extraits des coefficients de la TOD (Transformée en Ondelettes Discrète) au sixième

niveau de décomposition. (Ziani et al, 2012) ont proposé un algorithme génétique pour la sélection des indicateurs les plus pertinents. Dans cette étude, l'AG a été combinée avec les réseaux de neurones artificiels pour un problème de classification des défauts de roulement. La fonction objective utilisée est le critère de trace de la matrice de dispersion intra-classe. Ce critère a été utilisé pour évaluer la pertinence de chaque sous ensemble généré par les chromosomes de l'AG. Des résultats prometteurs ont été obtenus avec cette approche.

2.4.3 Approches basées sur les Séparateurs à Vaste Marge (SVMs).

Les machines à vecteurs supports (SVM) ont été proposées par Vapnik (1998). L'idée des SVM a été initialement introduite pour la recherche d'hyperplans de séparation dans le cadre de la classification binaire. Le principal but de cette méthode est de chercher un hyperplan qui sépare parfaitement les données de d dimensions dans deux classes. La distance entre l'hyperplan et la donnée la plus proche détermine la marge du classificateur. La fonction de décision peut être spécifiée par un petit sous-ensemble de données qui définit la position de l'hyperplan dans l'espace. Ces points sont appelés les vecteurs supports (SVs). Le principe des SVMs repose sur deux notions fondamentales : la marge maximale et la fonction noyau. Grâce à elles, un problème de discrimination non-linéaire peut être reformulé en un problème d'optimisation quadratique pour lequel on connaît des méthodes de résolution (Hastie et al., 2009).

Les SVMs ont été proposés pour de nombreuses applications pratiques dans la surveillance de l'état de santé des machines de surveillance de rotation (Samanta et al., 2001; Yang et al., 2007; Kurek & Osowski, 2010; Konar & Chattopadhyay, 2011). Samanta (2004) a utilisé les caractéristiques extraites de signaux bruts et prétraitées comme entrées des ANNs et les supports vecteurs machines (SVM) pour la reconnaissance de deux classes. Il a utilisé Les algorithmes génétiques (GA) pour la sélection des caractéristiques d'entrée du domaine temporel et optimiser les paramètres des classificateurs. Pour la plupart des cas, la précision de la classification des SVM était mieux que celle des ANNs, sans GA. Le temps d'exécution des SVMs a été moins par rapport à ANNs. Enfin, le taux de réussite des deux classificateurs avec sélection basée-GA était proche de 100%. La principale différence entre ce travail et celui de Jack & Nandi (2002) est le processus d'extraction de caractéristiques à partir du signal dans le domaine temporel.

Plusieurs variantes de SVM ont récemment trouvé une application dans le diagnostic des machines tournantes. La plupart des techniques SVM développées pour la classification est basée sur l'approche de séparation. Wang et al. (2012) ont développé une variante de SVM

appelée « Hyper-sphere-structured multi-class SVM » dans le but de déterminer le degré de dégradation d'un roulement. Elle consiste en la classification des caractéristiques calculées à partir des signaux vibratoires, traités par plusieurs méthodes de traitement du signal en plusieurs groupes de données, en utilisant un séparateur ou un hyperplan qui a une forme sphérique (voir figure 2-7).

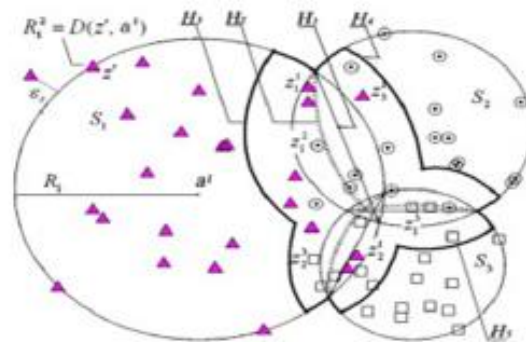


Figure 2-7: Schéma des hyperplans avec une forme sphérique pour trois classes (Wang et al., 2012)

Une version de SVM avec une fonction noyau de type « Radial Basis Function (RBF) » est utilisée par Zhang et al. (2013) appelée « Particle Swarm Optimization for SVM (SVM & PSO) ou Optimisation par essaims particulaires pour le SVM ». L'emploi de la méthode d'optimisation par essaims particulaires nécessite l'initialisation de plusieurs paramètres pour trouver les valeurs optimales pour ces paramètres. Un ensemble de données a été construit, avant la mise en route de la SVM-PSO, par l'extraction et la sélection des indicateurs de défauts de roulement. Des signaux vibratoires pour un roulement en différents état de fonctionnement ont été choisis pour en servir comme ensemble d'apprentissage de la méthode.

Une variante de l'optimisation multi-classes de SVM par la distance interclasse (Multi-class SVM optimized by inter-cluster distance) a été proposée par Zhang & Zhou (2013b) pour le diagnostic des roulements. Cette optimisation de SVM consiste en un choix aléatoire des valeurs d'initialisation des paramètres de la fonction noyau pour définir le degré de séparabilité des classes optimal. Le résultat est d'avoir un intervalle de valeur pour chaque paramètre pour une distance optimale entre les classes dans l'espace des caractéristiques. Cela se traduit par une meilleure séparabilité des classes et donc une meilleure immunisation contre les erreurs de classification.

Dans Samanta et al. (2003), les auteurs proposent un AG (algorithme génétique) pour la sélection d'un sous ensemble optimal d'indicateurs, dans le cadre de diagnostic automatisé des machines tournantes. Le sous ensemble sélectionné est utilisé pour entraîner deux classificateurs différents à savoir, les SVMs et les réseaux de neurones artificiels RNA. Afin de mettre en évidence l'effet de la sélection sur la performance de la classification, chaque classificateur a été entraîné dans deux cas différents : avec et sans sélection d'indicateurs.

La surveillance des conditions et le diagnostic des pannes dans les giravions ont un effet significatif sur l'amélioration du niveau de sécurité et la réduction des coûts d'exploitation et de maintenance.

2.4.4 Méthodes basées sur la densité

Les méthodes basées sur la densité consistent à définir la classe d'un objet en fonction de la classe des objets dans son voisinage le plus proche. Le voisinage d'un objet $o \in D$ peut être défini par $V(o) = \{o' \mid d(o, o') < \varepsilon\}$. Si $\text{card}(V(o))$ est suffisamment élevé, o est considéré comme étant dans une région dense de l'espace. Ainsi chaque zone dense dans l'espace des données sera vue comme une classe par l'algorithme. Contrairement aux méthodes basées sur des prototypes, ces méthodes permettent de découvrir des classes concaves dans l'espace de données. L'algorithme DBSCAN est un algorithme basé sur la densité.

2.4.4.1 Algorithme DBSCAN

DBSCAN (Density-Based Spatial Clustering of Applications with Noise) permet l'identification de classes, c'est à dire le regroupement des objets d'une base de données en sous-classes significatives.

Les applications de cette méthode aux bases de données spéciales conséquentes augmentent les exigences des algorithmes de clustering tel que :

- un minimum de connaissances sur les domaines afin de déterminer les paramètres d'entrée car les valeurs appropriées ne sont pas souvent connues à l'avance lorsque l'on travaille avec des bases de données importantes.
- la découverte de cluster de forme arbitraire car les formes de clusters dans les bases de données spatiales peuvent être sphérique, étiré, linéaire, allongé etc...
- une bonne efficacité sur les larges bases de données, c'est à dire celle contenant plus que quelques milliers d'objets.

L'algorithme DBSCAN (Density-Based Spatial Clustering of Applications with Noise) a pour but est de découvrir les clusters et le bruit dans une base de données spatiale (figure 2-8). Idéalement, il faudrait connaître les paramètres appropriés le rayon de voisinage Eps et le nombre minimum de points $MinPts$ de chaque cluster et un point de chacun des clusters respectifs. On peut ensuite retrouver tous les points de densité accessibles des points donnés à partir de ces paramètres corrects. Mais il n'est pas facile d'obtenir ces informations à l'avance pour chaque cluster de la base de données. Pourtant il existe une heuristique simple et efficace pour déterminer les paramètres Eps et $MinPts$ des clusters les plus minces. Ces paramètres de densité sont des bons candidats pour les paramètres globaux spécifiant des densités les plus basses qui ne sont pas considérés comme du bruit (Ester et al., 1996).

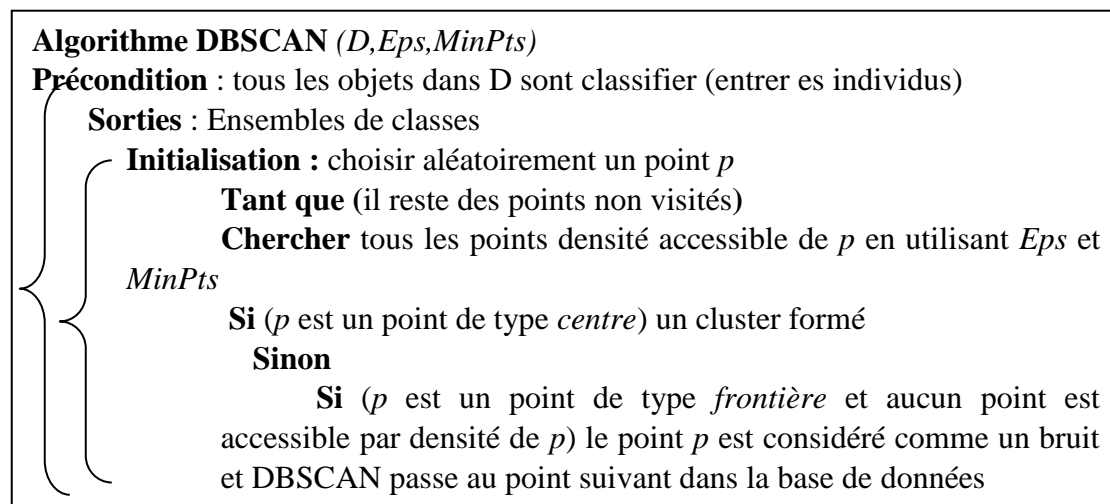


Figure 2-8: Algorithme DBSCAN (Ester et al., 1996)

Pour trouver un cluster, DBSCAN commence par un point arbitraire p et recherche tous les points de densité accessibles à partir de p . Si p est un point central, la procédure ajoute p au cluster. Si p est un point de bordure alors aucun point n'est atteignable à partir de p et DBSCAN visitera le prochain point de la base de données.

Grâce à l'utilisation des valeurs globale Eps et $MinPts$, DBSCAN peut fusionner 2 clusters dans le cas où 2 clusters de densité différente sont proches l'un de l'autre. Deux ensembles de points ayant au moins la densité la plus petite seront séparés l'un de l'autre si la distance entre les deux est plus large que Eps . En conséquence, un appel récursif de DBSCAN peut se révéler nécessaire pour les clusters détectés avec la plus haute valeur de $MinPts$. Cela n'est pas forcément un désavantage car l'application récursive de DBSCAN reste un algorithme basique, et n'est nécessaire que sous certaines conditions.

Ainsi, pour chaque objet que l'on ajoute, on a une zone de croissance qui va permettre d'étendre le cluster. Evidemment plus cette zone (une sphère) est grande et plus le cluster aura de chances de s'étendre. La notion de voisinage est la clé de cette méthode (figure 2-9). On forme donc le cluster de proche en proche. La difficulté que l'on peut rencontrer vient de la taille de la zone (rayon de la sphère) d'extension. DBSCAN a une complexité en $(n * \log n)$.

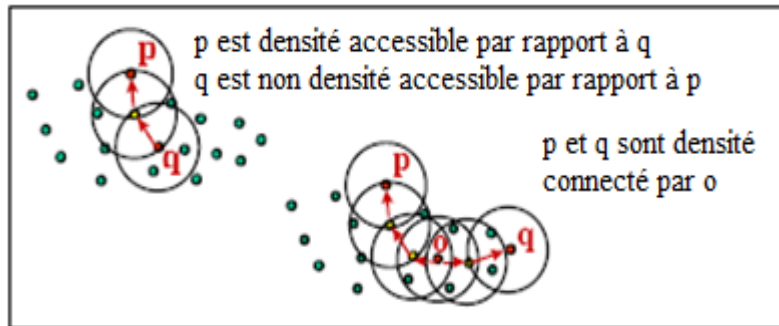


Figure 2-9: Notions de points dense-accessibles et de dense-connexité (Ester et al., 1996)

ϵ -voisinage d'un point p : $V_\epsilon(p)$

C'est l'ensemble des points de D qui sont à une distance inférieure ou égale à ϵ , soit :

$$V_\epsilon(p) = \{y \in D; d(p, y) \leq \epsilon\} \quad (2.20)$$

Une approche naïve consisterait à requérir, pour chaque point dans une classe, qu'il y ait au moins un nombre minimum ($MinPts$) dans son ϵ -voisinage. Cependant cette approche fait défaut car il y a deux sortes de points dans une classe : les points qui sont à l'intérieur de la classe (*points-noyau*) et ceux qui se trouvent en bordure de la classe (*points limites*). En général, un ϵ -voisinage d'un point limite contient significativement moins de points qu'un ϵ -voisinage d'un point-noyau. Et il faudrait donc fixer le nombre minimum de points à une valeur relativement basse pour inclure tous les points appartenant à la même classe. Par conséquent, on impose que tout point p d'une classe C , il existe un point y dans C tel que p soit à l'intérieur de l' ϵ -voisinage de y , et que $V_\epsilon(y)$ contienne au moins $MinPts$ points.

Points directement densité-accessible (dd-a)

Un point p de D est directement densité-accessible à partir d'un point q , ϵ et $MinPts$ étant fixé, si :

1. $p \in V_\epsilon(q)$ et

2. $|V_\epsilon(q)| \geq \text{MinPts}$ (condition noyau) où $||$ désigne le cardinal de l'ensemble.

La densité-accessibilité est évidemment une relation symétrique pour les paires de noyaux. Elle n'est cependant, en général, pas symétrique si un point de la paire est un noyau et l'autre un point-limite (figure 2-10).

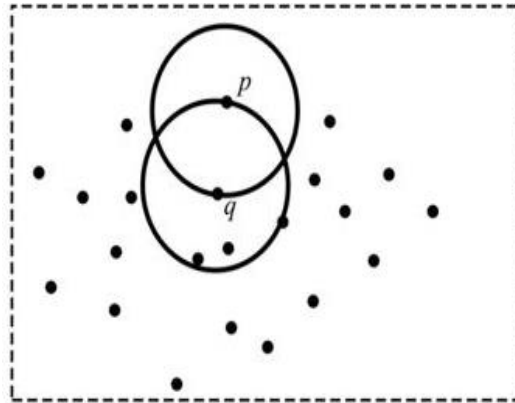


Figure 2-10: Noyau(q) et point limite (p) ; p est dd -a à partir de q ; q n'est pas dd -a à partir de p (Ester et al., 1996)

Point densité-accessible (d-a)

Un point p est densité-accessible à partir d'un point q , ϵ et MinPts étant fixé, s'il existe une chaîne de points $p_1, p_2, p_3, \dots, p_n$, avec $p_1=q$ et $p_n=p$, telle que p_{i+1} soit directement densité-accessible à partir de p_i .

La densité-accessibilité est une relation transitive mais en général non symétrique ; elle est évidemment symétrique pour les points noyaux.

Deux points-limite d'une même classe C peuvent ne pas être densité-accessibles à partir l'un de l'autre car la condition noyau pourrait ne pas être vérifiée pour chacun d'eux (figure 2-11). Cependant il doit y avoir un noyau de C à partir duquel les deux points-limite sont densité-accessibles.

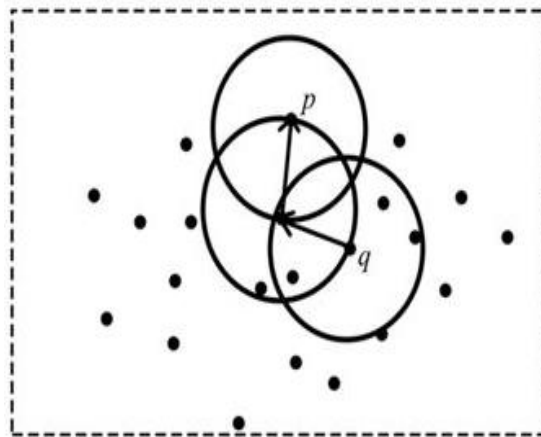


Figure 2-11: p est $dd-a$ à partir de q ; q n'est pas $dd-a$ à partir de p (Ester et al., 1996).

Point densité-connecté (d-c)

Un point p est densité-connecté à un point q , ϵ et $MinPts$ étant fixé, s'il existe un point o tel que les deux points p et q soient densité-accessibles à partir de o , ϵ et $MinPts$ étant fixé.

La densité-connecté est une relation symétrique. Pour les points densité-accessibles, la relation de densité-connexité est aussi réflexive (figure 2-12).

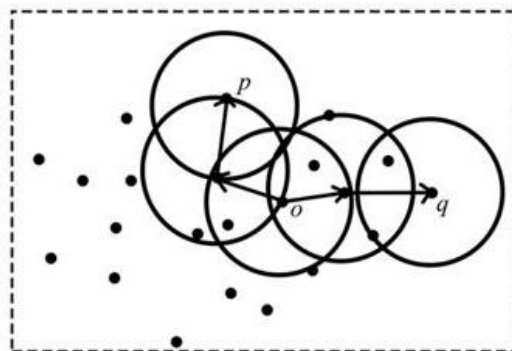


Figure 2-12: Création des clusters par DBSCAN ; densité-connectée (Ester et al., 1996).

2.4.4.2 Algorithme OPTICS

La présentation de cette méthode est largement inspirée de l'article de Ankerst et al. (1999).

OPTICS (acronyme de Ordering Point To Identify Clustering Structure en anglais) est un algorithme de partitionnement non supervisé de données. Il a été proposé par Ankerst et al. (1999). Il s'agit d'un algorithme basé sur la densité dont le principe est similaire à celui de

DBSCAN, son principal défaut est qu'il se base sur la visualisation des partitions de densités différentes.

Comme *DBSCAN*, *OPTICS* nécessite deux paramètres : ε , définissant un rayon de voisinage maximum, et *MinPts*, définissant un nombre minimum de points contenant dans son ε -voisinage. Ces 2 paramètres définissent donc une densité minimale pour constituer un cluster. Un point P appartient à un groupe si au moins *MinPts* points existent dans le voisinage d'un point. Par contre, à l'inverse de *DBSCAN*, le paramètre ε est optionnel. S'il est omis, il sera alors considéré comme infini. L'algorithme définit pour chaque point une distance, appelée distance-noyau, qui décrit la distance du *MinPts*^{ème} point le plus proche au point noyau.

OPTICS est un algorithme d'analyse de données qui ne produit pas des clusters réels mais il permet de distinguer et de visualiser la distribution des zones de haute densité par rapport à ceux de faibles densités en créant un ordre d'accessibilité de l'ensemble de données par rapport au point noyau. La méthode *OPTICS* est peu sensible au choix des deux paramètres cités. Ces propriétés mettent en évidence l'identification des clusters contenant dans un cluster principal.

L'idée clé de clustering basée sur la densité est que pour chaque objet, le voisinage d'un rayon (ε) doit contenir au moins un nombre minimum d'objets et la cardinalité de la zone voisinant doit dépasser un certain seuil. Les définitions formelles pour cette notion de clustering sont brièvement introduites dans la suite.

L'algorithme *OPTICS* crée un ordre de la base de données. En outre, il enregistre les deux grandeurs : le distance-noyau et la distance-accessibilité pour chaque objet. Cette information peut aider à extraire toutes les zones les plus denses :

Définition 1 : directement densité-accessible (dd-a)

L'objet p est directement densité-accessible à partir de l'objet q à partir de ε et *MinPts* dans un ensemble d'objets si :

- 1- $p \in N_\varepsilon(q)$ ($N_\varepsilon(q)$ est le sous ensemble de D contenu dans le voisinage de q .)
- 2- $\text{card}(N_\varepsilon(q)) \geq \text{MinPts}$ ($\text{card}(N_\varepsilon(q))$ désigne le cardinal de l'élément N .)

La condition $\text{card}(N_\varepsilon(q)) \geq \text{MinPts}$ est appelée "condition de distance-noyau". Si cette condition est valable pour un objet p , cet objet s'appellera *objet noyau*. Seulement à partir de cet objet, d'autres objets peuvent être directement densité-accessible.

Définition 2. : densité-accessible (d-a)

Un objet p est densité-accessible à partir de l'objet q à partir de ε et $MinPts$ dans l'ensemble des objets D s'il y a une chaîne d'objets $p_1, p_2, p_3, \dots, p_n$ $p_1 = q, p_n = p$ tel que $p_i = D$ et p_{i+1} est directement densité-accessible à partir de p_i à partir de ε et $MinPts$.

Cette relation n'est en général pas symétrique. Seuls les objets de bases peuvent être mutuellement densité-accessible.

Définition 3. : Densité-connecté (d-c)

L'objet p est densité-connecté à l'objet q à partir de ε et $MinPts$ dans l'ensemble d'objets D si il y a un objet $O \in D$ tel que p et q sont density-accessible depuis O à partir de ε et $MinPts$ dans D .

La densité-connecté est une relation symétrique. La figure (2-13) illustre les définitions sur un ensemble de données d'un échantillon de points en deux dimensions à partir d'un espace vectoriel. Notez que les définitions ci-dessus exigent seulement une mesure de distance et sera également applicable aux données à partir d'un espace métrique.

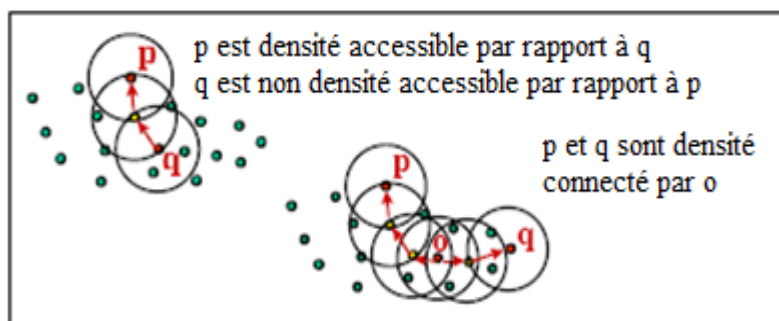


Figure 2-13: densité-accessibilité et connectée (Ankerst et al., 1999).

Un cluster basé sur la densité est désormais défini comme un ensemble d'objets de densité-connecté, qui est maximale à partir de densité-accessible. Un point est considéré comme bruit s'il n'est contenu dans n'importe quel cluster.

Définition 4 : (cluster et bruit)

Soit D un ensemble d'objets. Un cluster C à partir de ε et $MinPts$ dans D est un sous ensemble non vide de D satisfaisant aux conditions suivantes :

- 1- Maximalité : $\forall p, q \in D$: if $p \in C$ et q est densité-accessible de p à partir de ε et $MinPts$, alors aussi $q \in C$.

2- Connectivité : $\forall p, q \in D$ est densité-connecté à q à partir de ε et $MinPts$ dans D .

Chaque objet qui n'appartient à aucun cluster est un bruit.

Notez qu'un cluster contient non seulement des objets de base mais aussi des objets qui ne remplissent pas la condition de l'objet-noyau. Ces objets appelés *objets de bordure* du cluster sont, cependant, directement densité accessible depuis au moins un objet de base de la classe (contrairement aux objets bruit).

Définition 5: distance-noyau

P est un individu de l'ensemble de données D , ε est la valeur de la distance de voisinage, $N_\varepsilon(p)$ est le ε -voisinage de p , $MinPts$ est le nombre minimal de points dans le ε -voisinage de p et $MinPts$ -distance(p) est la distance de p à son $MinPts$ voisinage, le distance-noyau de p est défini comme :

$$distance - noyau_{\varepsilon, MinPts}(p) = \begin{cases} Ind\acute{e}fini, & \text{if } Card(N_\varepsilon(p)) < MinPts \\ MinPts - distance(p), & \text{otherwise} \end{cases} \quad (2.21)$$

Le distance-noyau de l'objet p est simplement la plus petite distance ε 'entre p et l'objet en son voisinage tel que p soit un distance-noyau à partir de ε 'si ce voisin est contenu dans $N_\varepsilon(p)$. Autrement le distance-noyau est indéfini.

Définition 6. : distance-accessibilité d'un objet

La distance-accessibilité de p à partir de o (figure 2-14) est définie comme :

$$distance - accessibl\acute{e}_{\varepsilon, MinPts}(p, o) = \begin{cases} Ind\acute{e}fini, & \text{if } |N_\varepsilon(o)| < MinPts \\ \max(distance - noyau(o), distance - noyau(p)) \end{cases} \quad (2.22)$$

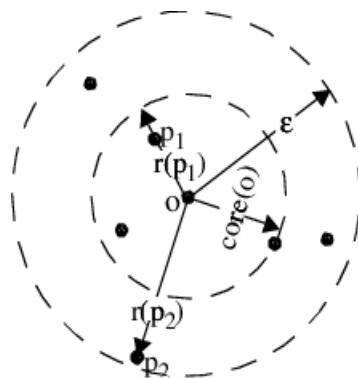


Figure 2-14: Distance-noyau(o), distance-accessibilité.

$r(p_1, o)$, $r(p_2, o)$ pour un $MinPts=4$ (Ankerst et al., 1999)

Intuitivement, la distance accessibilité de l'objet p à partir de l'autre objet o est la plus petite distance tel que p est directement densité accessible de o si o est un objet-noyau (figure 2-14). Dans ce cas, la distance accessibilité ne peut être que la distance-noyau de o parce que pour de plus petites distances aucun objet est directement densité accessible à partir de o . autrement, si on a pas un objet-noyau, même à la distance de génération ε , la distance accessibilité de p à partir de o est indéfinie. La distance accessibilité d'un objet p dépend de l'objet-noyau par rapport à laquelle il est calculé. Brièvement, la distance-noyau et la distance-accessibilité sont non définis si l'ensemble de points n'est pas suffisamment dense.

Algorithme de la méthode OPTICS

Pour chaque objet de la base de données, nous stockons la distance noyau et la distance d'accessibilité correspondante. Ces informations sont suffisantes pour extraire tous les clusters à base de densité. La Figure 2-15 illustre la boucle principale de l'algorithme OPTICS. Chaque objet de la base de données *EnsDeObjets* est pris en charge par une procédure *EtendreClusterOrdre* si l'objet n'est pas encore traité.

```

OPTICS (EnsDeObjets,  $\varepsilon$ , MinPts, FichierOrdre)
  FichierOrdre.ouvrir() ;
  Pour  $i$  de 1 à EnsDeObjets.taille faire
    Objet= EnsDeObjets.obtenir( $i$ ) ;
    Si non Objet.traite alors
      EtendreClusterOrdre(EnsDeObjets, Objets,  $\varepsilon$ , MinPts,
FichierOrdre)
  FichierOrdre.ferme() ;
Fin ; // OPTICS

```

Figure 2-15: L'algorithme OPTICS (Ankerst et al., 1999).

La procédure **EtendreClusterOrdre** est représentée dans la Figure 2.16. Premièrement, cette procédure recherche les ε -voisins de l'objet «**Objet**» passé dans la boucle principale OPTICS, place sa distance d'accessibilité à indéfinie et détermine sa distance noyau. L'objet est écrit dans **FichierOrdre**. La condition *Si* contrôle les propriétés noyau de l'objet, s'il n'est pas un noyau principal avec la distance ε , le contrôle est simplement retourné à la boucle principale de OPTICS qui sélectionne le prochain objet non traité dans la base de données. Autrement, si **Objet** est un objet noyau à une distance inférieure ou égale à ε , nous collectons itérativement les objets avec la densité d'accessibilité directe avec ε et *MinPts*. Les objets qui sont directement densités accessibles à partir de l'objet noyau courant sont insérés dans la liste **OrderSeeds**. Les objets contenus dans **OrderSeeds** sont triés par leurs distances

d'accessibilité à l'objet noyau le plus proche qui ont été directement densités accessibles. Dans chaque étape de la boucle *Tant que*, un objet **ObjetCourant** ayant la plus petite distance d'accessibilité dans la liste **OrdreSeeds** est sélectionné par la méthode **OrderSeeds.suivant**. Les e-voisinage de cet objet et sa distance noyau sont déterminés. L'objet est simplement écrit dans le fichier **FichierOrdre** avec sa distance noyau et sa distance d'accessibilité.

```

EtendreClusterOrdre (EnsDeObjets, Objets, ε, MinPts, FichierOrdre) ;
  Voisinage= EnsDeObjets.voisinage(Objets, ε) ;
  Objet.traite=vrai ;
  Objet.accessibilite_distance=INDEFINIE ;
  Objet.EnsDistanceNoyau (voisins, ε, MinPts) ;
  FichierOrdre.écrire(Objet) ;
  Si Objet.noyau_distance <> INDEFINIE alors
    OrdreSeeds.Mise-à-jours (voisins, Objets) ;
    Tant que non OrdreSeeds.vide ( ) faire
      ObjetCourant .traite=OrderSeeds.suivant( ) ;
      Voisins=EnsDeObjets.voisins (ObjetCourant, ε) ;
      ObjetCourant.traite=vrai
      ObjetCourant.EnsDistanceNoyau(voisins, ε, MinPts) ;
      FichierOrdre.écrire( ObjetCourant) ;
      Si ObjetCourant.noyau_distance <> INDEFINIE alors
        OrdreSeeds.Mise-à-jours(voisins, ObjetsCourant) ;

```

Figure 2-16: La procédure EtendreClusterOrdre (Ankerst et al., 1999).

```

OrdreSeeds ::Mise-à-jour(voisins, ObjetCentre) ;
  c_dist=ObjetCentre.distance_noyau ;
  Pour tous Objet DE voisinage FAIRE
    Si non Objet.traite alors
      nouveau_r_dist=max(c_dist, ObjetCentre.dist(Objet));
      Si Objet.distance_accessibilité=INDEFINIE alors
        Objet.distance_accessibilité=nouveau_r_dist ;
        Insérer (Objet, nouveau_r_dist)
      Sinon
        Si nouveau_r_dist < Objet.distance_accessibilité alors
          Objet.distance_accessibilité=nouveau_r_dist ;
          Décroître (Objet, nouveau_r_dist ) ;
Fin ; // OrdreSeeds ::Mise-à-jours

```

Figure 2-17: La méthode OrderSeeds ::mise-à-jour (Ankerst et al., 1999).

L'insertion dans la liste *OrderSeeds* et le traitement de la distance d'accessibilité sont gérés par la méthode *OrderSeeds* :: Mise-à-jour présenté dans la Figure 2-17. La distance d'accessibilité pour chaque objet dans l'ensemble des voisins est déterminée par rapport à l'objet *ObjectCentre*.

Les objets qui ne sont pas encore dans la liste *OrderSeeds* sont simplement insérés avec leurs distances d'accessibilité. Les distances d'accessibilité des objets qui sont dans la liste sont remplacées par leurs nouvelles distances d'accessibilités si ces nouvelles distances sont plus petites que leurs distances d'accessibilité précédentes.

2.4.5 Méthodes basées sur une grille

Les méthodes de classification basées sur une grille (*ang. grid-based clustering*) consiste à discrétiser l'espace des données (représentées par des attributs numériques), c'est-à-dire à découper chaque individus en intervalles de valeurs (pouvant être fixes ou dynamiques). Une *unité* se définit comme un point de cet espace discret ou comme un hyperrectangle construit par l'intersection d'un intervalle de chacune des dimensions de l'espace de données considéré. Les classes sont construites en utilisant un algorithme de type *seed fill* en propageant les classes tant que le nombre d'objets par unité est supérieur à un seuil. Parmi ces méthodes, on retrouve l'algorithme CLIQUE (Agrawal et al., 1998). On ne retrouve pas pour le moment des applications de cette approche dans le cadre du diagnostic de défauts des machines tournantes.

2.4.6 Méthodes probabilistes

Un problème de classification peut se traduire par un problème d'estimation des paramètres d'une loi de probabilité, définie par un modèle de mélange de lois (le plus souvent un mélange gaussien). Ces méthodes font l'hypothèse qu'à chaque classe C_k est associée une loi de probabilité $P(o, \theta_k)$ de paramètres θ_k qui permet de déterminer la probabilité d'appartenance d'un objet o à la classe C_k . on note π_k la proportion de la k -ième loi dans le mélange (Blansché, 2006).. Les paramètres du mélange sont alors notés $\Phi = (\pi_1, \dots, \pi_k, \theta_1, \dots, \theta_k)$ et la fonction de densité est donnée par :

$$P(o, \Phi) = \sum_{k=1}^K \pi_k P(o, \theta_k) \quad (2.23)$$

Les méthodes de classification basées sur ce principe, comme par exemple l'Algorithme d'Espérance-Maximisation EM, cherche à approximer les paramètres Φ du modèle. Pour

chaque objet, une probabilité d'appartenance à chacune des classes, pouvant être assimilée à un degré d'appartenance, est calculée

2.4.7 Méthodes hiérarchiques

A la différence des méthodes basées sur la distance présentées précédemment, les méthodes hiérarchiques produisent une séquence de partitions imbriquées appelée dendrogramme (Johnson, 1967). Ce dernier permet tout de même d'obtenir une partition lorsque l'on en considère une coupe transversale à un certain niveau de dissimilarité. Il existe traditionnellement deux types d'approches hiérarchiques :

- les approches ascendantes ou agglomératives qui partent d'un ensemble de n classes constituées d'un unique objet (appelées singletons) et qui vont être itérativement fusionnées de façon à n'obtenir qu'une seule classe à la fin,
- les approches descendantes ou divisives qui, inversement, partent d'une classe contenant toutes les données et s'arrêtent lorsque n singletons ont été créés.

2.4.7.1 Classification ascendante hiérarchique

La classification ascendante hiérarchique est dite ascendante car elle part d'une situation où tous les individus sont seuls dans une classe, puis sont rassemblés en classes de plus en plus grandes. Le principe de base de cette méthode est le suivant : initialement chaque individu est considéré comme une classe, soit n classes. On cherche à réduire le nombre de classes à $nb_{classes} < n$, ceci se fait itérativement. À chaque étape, on fusionne deux classes, réduisant ainsi le nombre de classes. Les deux classes choisies pour être fusionnées sont celles qui sont les plus « proches », en d'autres termes, celles dont la dissimilarité entre elle est minimale, cette valeur de dissimilarité est appelée indice d'agrégation. Comme on rassemble d'abord les individus les plus proches, la première itération a un indice d'agrégation faible, mais celui-ci va croître d'itération en itération. L'algorithme continue jusqu'à ce que l'ensemble des individus se retrouve dans une unique classe (Székely & Rizzo, 2005). Chaque classe d'une partition est incluse dans une classe de la partition suivante. Le principal problème de cette méthode hiérarchique est de définir le bon critère de regroupement de deux classes, c'est-à-dire trouver une bonne méthode de calcul des distances entre classes.

Les méthodes hiérarchiques sont déjà utilisées dans le diagnostic des machines tournantes dans différents objectifs, soit d'une manière directe, c'est à dire que l'on utilise directement pour la classification (Karabadjji et al., 2014 ; Amarnath et al., 2013) soit indirecte, c'est-à-dire que l'on combine cette méthode à d'autres méthodes de classification

automatique nécessitant la connaissance à priori du nombre de classe optimal ou bien comme outil d'apprentissage des méthodes supervisées (Saimurugana et al., 2011).

2.4.7.2 Arbre de décision

Un arbre de décision est une méthode hiérarchique descendante, composé de plusieurs nœuds internes reliés entre eux par des branches. Ces nœuds internes représentent les états (classes) intermédiaires lors de la construction de l'arbre tandis que les branches représentent le processus de construction. Il possède également un certain nombre de nœuds terminaux ou feuilles, correspondant aux classes finales, auxquels on associe un indicateur de classe (étiquette correspondant à son mode de fonctionnement), parmi les M classes du problème à traiter.

L'intérêt de la classification par arbres de décision est de diviser les données contenus dans un ensemble d'apprentissage, par des tests définis à l'aide d'attributs (calcul de distance) jusqu'à ce que l'on arrive à des sous-ensembles de points, obtenus en suivant l'évolution d'un critère, ne contenant presque que des données appartenant toutes à la même classe.

Plusieurs variantes de l'arbre de décision ont été utilisées, comme méthodes de classification des données, dans le cadre du diagnostic des roulements en utilisant des paramètres issus des signaux vibratoires et des signaux acoustiques. La différence principale entre ces variantes réside dans les méthodes utilisées pour résoudre le problème de sur-apprentissage de l'AD et le problème de l'expansion (la complexité de l'AD).

Amarnath et al. (2013) ont présenté un arbre de décision induit basé sur l'algorithme classique ID3, plus précisément la variante C4:5. Cet arbre emploie les valeurs des indicateurs de défauts pour décider l'état de roulement (sain, défaut de bague intérieure, défaut de bague extérieure).

Une étude comparative entre les différentes versions de l'AD a été présentée dans l'article (Karabadi et al. 2014) dans le domaine industriel. Pour surmonter les deux challenges de l'AD que sont la complexité et le sur/sous apprentissage et donc améliorer cette méthode en une version non élaguée (Improved Unpruned Decision Tree IU DT), Les auteurs ont proposé d'hybrider l'arbre de décision avec des méthodes de sélection des caractéristiques (Wrapper ou enveloppante pour choisir les attributs les plus pertinents) et une méthode de réduction de dimensions. Ceci convertit le problème de la construction de l'arbre de décision en une exploration combinatoire de l'espace de recherche graphique. Cette comparaison a été élaborée sur une base de 420 signaux vibratoires pour des roulements avec différents états

d'endommagement et différentes vitesses de rotation. Les résultats de cette comparaison ont montré que l'arbre de décision de type REPTree IUDT (avec une précision de 98% et la possibilité de retrouver les indicateurs les plus pertinents pour cette classification qui sont Skewness, RMS, Crestfactor, l'intervalle moyen entre les fréquences des quatre amplitudes maximales (par les ondelettes) est le meilleur choix de l'AD pour le diagnostic des roulements.

Loin de la classification, les auteurs dans l'article (Saimurugana et al., 2011) ont présenté une autre façon pour employer l'arbre de décision, comme outil de sélection des données. Les auteurs ont utilisé l'AD pour sélectionner les caractéristiques des vibrations les plus pertinentes afin d'améliorer la précision de la méthode de classification utilisée, SVM, et de réduire sa complexité de convergence vers un optimum de classification global.

2.5 Conclusion

La reconnaissance des formes peut être utilisée pour le diagnostic des machines tournantes en général ou des roulements en particulier. Dans la littérature du diagnostic, plusieurs méthodes de classification ont été employées. Ces méthodes ont été souvent combinées avec des techniques de traitement du signal afin d'élaborer un diagnostic automatique de la machine. Ces approches ont été appliquées soit sur une base des signaux enregistrés auparavant, ou bien par l'extraction des indicateurs soit du même roulement avec différents états de dégradation (le même défaut mais avec des tailles différentes) ou bien de différents roulements avec différents défauts (défaut bague extérieure, défaut bague intérieure, défaut de bille).

L'application de la méthode de classification dans le diagnostic des machines tournantes se fait soit en se basant sur la qualité de « séparation automatique » des données différentes ou pour trouver un « modèle » qui peut expliquer la relation reliant les données avec leurs classes. Ces méthodes peuvent représenter un choix intéressant dans le cas d'un apprentissage hors ligne d'un système de RdF utilisé par exemple pour la discrimination des signaux vibratoires des roulements avec différents états de dégradation.

En fait toutes ces méthodes s'appuient sur la construction d'un système d'apprentissage reposant sur la définition d'un ensemble fini d'observations. L'utilisation de ces méthodes présente toutefois un inconvénient dans le sens où elle suppose la connaissance a priori de toutes les observations nécessaires à la caractérisation des classes. Or durant le suivi d'un système mécanique qui naturellement passe d'un mode opératoire normal à un autre anormal

implique que les données collectées sur ce système vont connaître des évolutions dans le temps. La détection et l'analyse de ces évolutions pourra nous apporter des informations supplémentaires d'où la nécessité de mettre en place une méthode capable d'exploiter ce type de données. L'extraction de ces informations complémentaires nécessite un apprentissage continu (en ligne) du système ce qui peut nous permettre d'employer la RdF pour le diagnostic et la surveillance continue des roulements. L'apprentissage en ligne signifie que la RdF peut être utilisée pour classifier les caractéristiques des signaux qui arrivent au fur et à mesure. Cependant dans la surveillance d'un système industriel, certains phénomènes évolutifs peuvent se présenter, le passage d'un mode de fonctionnement vers un autre étant rarement instantané. Ces contraintes de représentation ou de modélisation sont solutionnées avec l'introduction de la théorie de la classification dynamique.

L'utilisation de la classification dynamique permet d'introduire la notion du temps, c'est à dire qu'on pourrait suivre l'évolution temporelle des formes, la quantifier et l'intégrer dans le processus de diagnostic. La différence entre la classification statique et la classification dynamique réside dans la représentation des classes et des objets. Dans la classification statique, les classes et les objets sont statiques c'est à dire que leurs caractéristiques ne changent pas dans le temps. Par contre, dans la classification dynamique, les classes et/ou les objets évoluent, ils changent au fil du temps (Mouchaweh et Lughofer, 2010), les classes peuvent être créées si nécessaire, ou supprimés si la nécessité a disparu, les classes peuvent se déplacer, ou pivoter, les classes semblables peuvent être fusionnées pour former de nouvelles classe, etc. Cette évolution des classes traduit un changement dans les données qui signifient un changement dans le système observé. La classification statique, omet cette évolution temporelle des données exposées par l'évolution des classes. En effet, une classification statique des données qui changent dans le temps ne peut pas être considérée comme une représentation complète d'un système réel qui change naturellement au cours du temps pour passer d'un fonctionnement normal à un fonctionnement anormal. Dans le cas de surveillance des roulements, le changement brutal est toujours associé à l'émergence d'un défaut.

En d'autres termes les méthodes de classification utilisées jusqu'à maintenant dans le diagnostic des machines tournantes n'ont pas été exploitées à leur maximum.

Le chapitre suivant décrit la méthode de classification dynamique, son intérêt dans le diagnostic des roulements. Ce chapitre développe une méthode de classification spécifique au suivi des roulements.

Chapitre 3 :

Contribution à l'étude de diagnostic de défauts par la
méthode de classification dynamique

3 Contribution à l'étude de diagnostic de défauts par la méthode de classification dynamique

3.1 Introduction

Dans ce chapitre, nous présentons les méthodes de classification dynamiques et la stratégie de détection de défaut développée durant cette thèse. Les méthodes de diagnostic par les méthodes de classification basées sur la notion de densité ont été privilégiées pour nos travaux, OPTICS et DBSCAN. Ce choix a été motivé d'une part par le fait que ces méthodes ne nécessitent pas trop d'informations pour les manipuler et qu'elles peuvent détecter des classes de formes convexes. Ces méthodes, en tant que telles restent des méthodes de classification statiques. Nous avons donc mis en place une architecture autour de ces méthodes pour les rendre dynamique.

Nous présentons, dans un premier temps, les méthodes de classification dynamique en général. Nous détaillerons ensuite notre contribution au tour de la méthode OPTICS.

3.2 Classification dynamique

3.2.1 Phénomènes dynamiques en environnement non stationnaire

Cette section détaille les mécanismes généraux engendrés par l'incorporation des nouvelles données dans la partition et les phénomènes dynamiques causés par la non-stationnarité des données. Avant d'exposer ces mécanismes, nous allons tout d'abord donner les définitions d'une classe et d'une partition dynamique dans le contexte de notre étude.

Définition 1: Une classe est définie comme un ensemble (ou regroupement homogène) de données similaires. La similarité est évaluée d'un point de vue de la proximité géométrique des données. Lorsque les données d'une classe évoluent dans le temps, cette classe est dite évolutive. Son modèle doit donc être adapté.

De cette définition, on déduit que deux classes différentes ne se chevauchent pas ou très peu, sinon, leurs données seront suffisamment proches pour constituer une seule et unique classe.

Définition 2: Une partition dynamique représente l'ensemble des regroupements homogènes des données non-stationnaires. Elle est associée à un modèle de connaissances (i.e. classes) susceptible d'évoluer dans le temps en fonction de l'incorporation de nouvelles données.

Dans le cadre de l'étude de la problématique des données non-stationnaires, la méthode de reconnaissance de formes doit être dynamique, c'est-à-dire qu'elle doit être capable de gérer l'évolution des classes. Cette évolution peut être (i) de nouvelles classes, le déplacement d'une classe, (iii) la fusion de deux classes, (iv) la scission d'une classe en deux classes ou la suppression d'une classe dite « obsolète ».

3.2.2 Apparition de nouveautés dans un espace de classification

Lorsque de nouvelles données sont acquises dans une même région ou différentes régions non occupées par les classes de la partition existante et si ces données apportent suffisamment d'informations, elles révèlent l'apparition de nouvelles classes. L'apparition de chaque nouvelle classe se caractérise par un regroupement de données similaires dans une même région de l'espace.

Dans de telles situations, l'objectif consiste à la détection et à la création des classes constituées par ces nouvelles données. Ceci se traduit par la double tâche de déterminer, d'une part, les données similaires pour former chaque classe et d'autre part, de calculer les modèles des nouvelles classes créées. La première tâche est généralement accomplie par l'utilisation d'un critère de similarité, tandis que la seconde nécessite la mise en œuvre d'une procédure d'initialisation respectant la distribution des données. Au cours de ce mécanisme de création, la partition dynamique est actualisée par l'insertion des nouvelles classes.

Nous recensons trois cas de figure pour caractériser l'apparition de nouveautés :

Cas 1 : Apparition de classes dans un espace de classification « vide »

Au départ de l'apprentissage, les premières données arrivent dans l'espace de classification. Or, cet espace étant initialement « vide », aucune classe n'existe encore. Si les premières données acquises sont similaires (géographiquement proches), elles constituent une classe, sinon elles révèlent l'apparition de plusieurs classes disjointes compte tenu de la définition 4.2. La figure 3.1.a présente une situation de création d'une classe dans un espace de classification initialement « vide ».

Cas 2 : Apparition de nouvelles classes après un changement brusque

On retrouve cette situation quand il y a déjà une ou plusieurs classes créées dans l'espace de classification. L'arrivée soudaine de nouvelles données dans une région isolée tend à l'apparition, si elles sont similaires et suffisantes, d'une nouvelle classe (figure 3.1.b).

Cas 3 : Apparition de nouvelles classes après une évolution rapide de données

Dans certaines situations, l'apparition d'une classe peut être précédée par une évolution plus ou moins rapide de données. Ces données de transition sont considérées comme des points caractérisant une dérive rapide entre deux classes. Dans le cadre de la classification, ces données de transition ne présentent pas une structure cohérente et stable pour contribuer à la formation d'une classe. Après la phase transitoire, lorsque que suffisamment de données se stabilisent dans une région isolée de la partition existante, ces données seront, quant à elles, révélatrices de l'apparition d'une nouvelle classe (figure 3-1.c). Dans certaines applications, l'analyse de cette phase transitoire décrite par la dérive de données entre deux classes présente un intérêt considérable notamment pour établir *un pronostique* (Amadou, 2002).

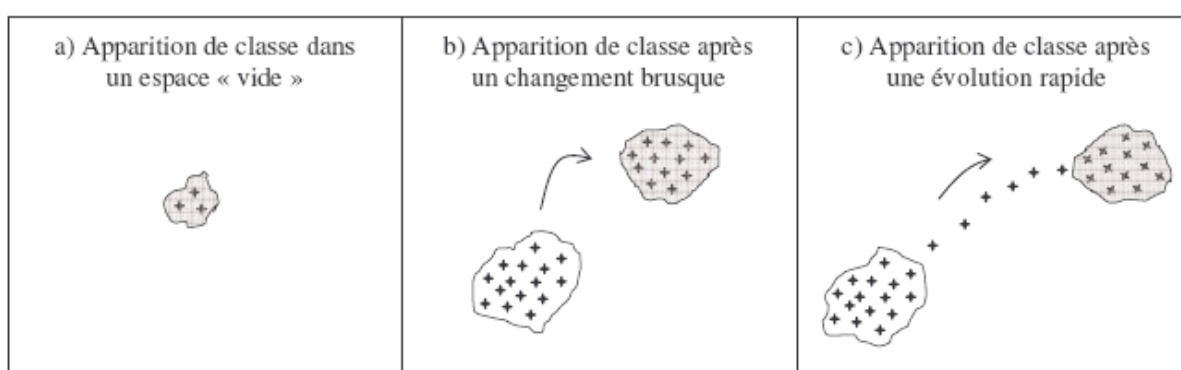


Figure 3-1: (a) Apparition de classe dans un espace de classification « vide », (b) Apparition d'une nouvelle classe après un saut brut et (c) Apparition d'une nouvelle classe après une dérive rapide (Amadou, 2002)

Le phénomène d'apparition de nouvelles classes présente plusieurs risques à cause de l'absence de superviseur. En effet, la méthode de classification peut affecter des données de classes différentes à une seule classe, ou encore attribuer des données d'une même classe à des classes différentes. Par ailleurs, la distribution des données n'étant pas connue a priori, l'initialisation d'un modèle de classe avec un nombre insuffisant de données reste difficile et peut conduire à des aberrations. D'autre part, à cause de l'influence du bruit et des perturbations, des erreurs de classification sont fréquentes. Ces erreurs peuvent conduire à la création de classes parasites affectant la qualité de modélisation. Néanmoins, ces classes parasites ne doivent pas être confondues avec des classes représentatives dont les données se présentent d'une façon périodique (par intermittence) dans une même région de l'espace. (Amadou, 2010).

3.2.3 Evolution de la partition dynamique

Contrairement au contexte de classification de données stationnaires (classes statiques), la problématique de la classification dynamique pose le problème d'évolution de la partition au cours de l'incorporation de nouvelles données. Cette problématique distingue notre étude de la plupart des travaux menés jusqu'à présent, qu'il s'agisse de classification non-supervisée ou incrémentale. Rappelons que l'intérêt de prendre en compte les évolutions des classes concerne diverses applications réelles : diagnostic industriel (évolution des modes de fonctionnement), médicale (expansion de cancers), surveillance vidéo (mouvement des cibles), ... etc. En effet, dans ce type d'applications, au fur et à mesure que les nouvelles données deviennent disponibles, leurs informations participent à la construction et à l'adaptation des classes. L'incorporation de ces données entraîne des évolutions plus ou moins complexes de la partition dynamique (figure 3-2.c).

Il est donc nécessaire dans ces situations de prendre en compte ces évolutions à travers l'adaptation des classes concernées. L'accomplissement de cette tâche passe par la réponse à deux questions : De quelle classe une nouvelle donnée est-elle suffisamment proche pour contribuer à la redéfinition de son modèle ? De quelle manière l'information portée par cette nouvelle donnée contribue-t-elle à adapter le modèle de la classe ? La réponse à la première question est généralement donnée par le critère de similarité qui détermine l'appartenance des données aux classes. Quant à la deuxième question, la réponse passe à travers la mise en œuvre de règle d'adaptation récursive.

Les situations d'évolution de la partition dynamique peuvent être divisées en deux cas.

Cas 1 : Modifications locales de classes

Les modifications locales de classes sont des mécanismes dynamiques dus à l'incorporation de nouvelles données non-stationnaires. Elles se traduisent par des variations plus ou moins importantes observables localement au sein de la structure (géométrie et contour) de classes dans leur région de définition. Les grossissements, rétrécissements et rotations de classes sont les modifications les plus courantes (figures 3-2.a et 3-2.b). La plupart des méthodes de modélisation adaptative développées avec des règles d'apprentissage incrémental, se limitent à la prise en compte des modifications locales. Au travers de ces règles, l'adaptation de modèles de classes se fait donc localement.

Cas 2 : Evolutions avec glissements de classes

Nous parlons d'évolutions avec glissements de classe quand les nouvelles données d'une même classe s'éloignent progressivement de la région occupée par les anciennes données. Contrairement aux modifications locales, les évolutions avec glissements de classes s'accompagnent de changements de la région de définition des classes. Ces évolutions se font à travers de simples déplacements de classes ou la combinaison des déplacements avec des variations de la structure de ces classes. Dans ces situations, la classification des données non-stationnaires nécessite un suivi des classes évolutives. Ce suivi des classes évolutives passe nécessairement par l'intégration des informations récentes (représentatives) apportées par les nouvelles données dans la classe et l'élagage des informations obsolètes (non-représentatives) contenues dans les données les plus anciennes de cette même classe.

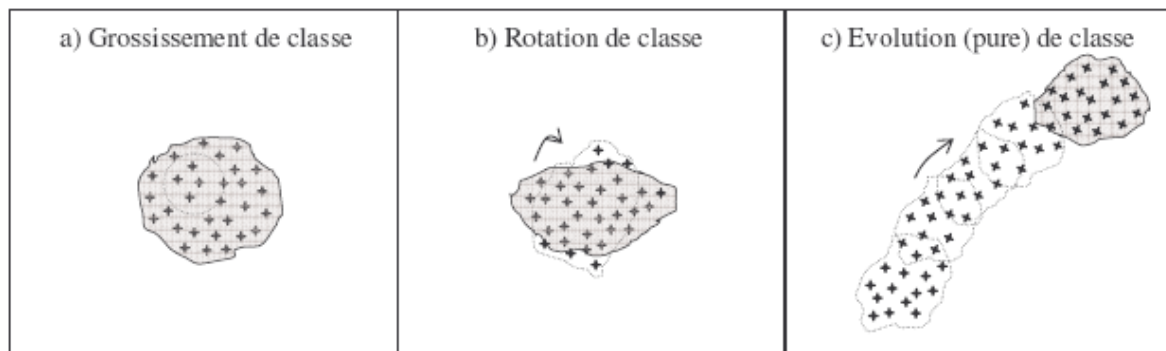


Figure 3-2: (a) Développement de classe, (b) Modification quelconque de classe et (c) Evolution de classe suite à une dérive de données non-stationnaires (Amadou, 2010)

La mise à jour récursive des modèles de classes est rendue possible grâce à des règles de mise à jour récursives permettant à la fois **l'ajout et le retrait d'informations**. Le suivi des classes évolutives constitue l'un des objectifs essentiels de la classification en environnement dynamique. La figure 3-2.c illustre une classe évolutive caractérisée par une séquence d'états transitoires.

La mise en œuvre d'une méthode de mise à jour récursive est une épreuve quelque fois très difficile selon le modèle de classification utilisé. Dans les situations d'adaptation de classes, l'influence du bruit peut causer des imperfections sur la modélisation des classes. De même, lors de la modélisation dynamique, le phénomène de dérive de données (figure 3-2.c) ne doit pas être confondu avec une évolution de classe (figure 3-2.c). Par ailleurs, la dynamique des classes évolutives peut entraîner des interactions entre les classes (coalitions,

séparations,...) pouvant affecter la qualité de la modélisation si celles-ci ne sont pas correctement traitées.

3.2.4 Fusion d'informations mutuelles

Dans un environnement non-stationnaire, suite à certains phénomènes dynamiques, des interactions peuvent se produire entre deux ou plusieurs classes. Ainsi, une ou plusieurs classes initialement disjointes peuvent se rejoindre et partager des données. Par exemple, en reconnaissance de caractères, à cause de données manquantes, deux formes initialement disjointes peuvent fusionner pour constituer un caractère après l'acquisition de données complémentaires. Lorsque deux ou plusieurs classes différentes partagent des données, ces données sont porteuses d'informations mutuelles et sont dites **ambiguës** (Dubuisson, 1990).

Afin de lever l'ambiguïté en cas d'informations mutuelles, la méthode de classification dynamique doit fusionner les classes concernées pour ne former qu'une seule classe (figure 3-3.a). La détection d'informations mutuelles peut être effectuée avec un critère de ressemblance. Par contre, la détermination du modèle résultant de la fusion, sans effectuer un réapprentissage sur toutes les données, nécessite la mise en œuvre de règles de fusion.

La mise en œuvre d'une règle de fusion est une tâche souvent compliquée qui peut exposer la partition dynamique aux problèmes d'optima locaux comme le sous-apprentissage de modèle. Un autre risque est lié au bruit dans les données qui peut se comporter comme une fausse information mutuelle et déclencher une fusion de classes non souhaitée (Amadou & Lecoeuche, 2005a ; Amadou & Lecoeuche, 2005b ; Lecoeuche & Lurette, 2003).

Beaucoup d'intérêts sont portés sur la manipulation de données non-stationnaires. En classification comme dans divers autres domaines, ces données restent très difficiles à traiter à cause de l'évolution du modèle de connaissances. L'expression de classification dynamique est ainsi introduite pour qualifier les algorithmes développés pour la classification automatique de ces données (Lurette, 2003 ; Lecoeuche, 2006).

Développé à partir des méthodes à noyau, un algorithme SAKM (Self-Adaptive Kernel Machine) est récemment introduit pour la classification dynamique (Amadou & Lecoeuche, 2005a). C'est une contribution originale pour les SVM & méthodes à noyau dans les problèmes d'apprentissage en ligne d'un système de surveillance de processus industriels et d'un système de suivi en ligne de processus non-stationnaires.

3.2.5 Séparation d'informations exclusives

Le phénomène inverse de la fusion peut se produire en environnement non-stationnaire lorsqu'une classe se scinde en deux classes au cours de l'incorporation des nouvelles données (figure 3-3.b). Par exemple, les données caractérisant le mode de fonctionnement d'un processus, peuvent être initialement mélangées avec des données parasites causées par des perturbations extérieures. Lors de l'évolution de ce mode de fonctionnement, la classe de départ peut se scinder en deux suite à une séparation de la classe normale de la classe parasite.

Lorsque de tels phénomènes se produisent, la méthode de classification dynamique doit décider de la scission en utilisant un critère de détection de scission de classe et elle doit comporter des règles de scission de classe permettant la détermination des modèles des classes résultantes de la scission.

La difficulté principale concerne la mise en œuvre du critère de détection de scission. En effet, celui-ci doit répondre à la question : les données contenues dans chaque classe sont-elles similaires ou non ? Un critère répondant de façon fiable et certaine à cette question reste évidemment très difficile à mettre en œuvre. Or, si une classe est constituée par des données non similaires, il s'ensuit une mauvaise modélisation de la partition (Amadou, 2010).

3.2.6 Suppression de connaissances obsolètes ou parasites

La classification dynamique s'effectue à travers la modélisation continue de processus sur de longues durées. Il se peut donc que les connaissances apprises dans un passé plus ou moins lointain ne soient plus valables à l'instant courant. Les classes représentant les connaissances obsolètes doivent donc être éliminées de l'espace de classification. De même, d'autres classes parasites créées à cause du bruit ou des perturbations extérieures, sont à supprimer (figure 3-3.c).

Un classifieur dynamique doit donc comporter une **procédure d'élimination** des classes non-représentatives (classes parasites ou obsolètes). La détection de classes parasites est généralement basée sur un *critère de cardinalité* (Lurette, 2003).

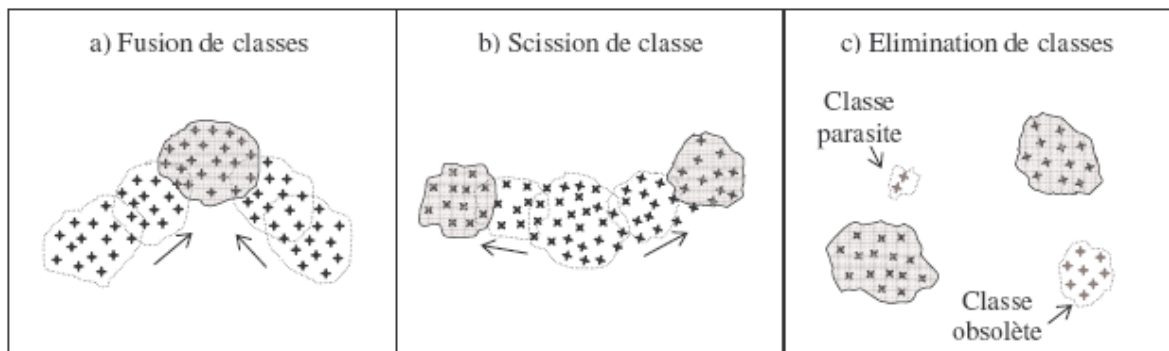


Figure 3-3: (a) Fusion de classes partageant des informations mutuelles (b) Scission de classes après séparation des informations partagées et (c) Elimination de classes obsolètes et parasites

Zhang et al. (2006) et Wong et al. (2006) ont utilisé la méthode de classification dynamique pour la détection des nouveautés rencontrées dans les systèmes mécaniques. La méthode est utilisée en deux volets, le premier consiste en la détection des observations différentes et le deuxième et d'analyser si ces nouveautés sont dues aux fluctuations causées par un bruit ou bien par un défaut. Dans le cas où la nouveauté peut apporter des informations utiles, le classifieur doit ajuster ses règles de décision et mettre à jour sa base d'apprentissage.

3.3 Système du diagnostic vibratoire

Le système du diagnostic suivi dans notre étude constitue trois étapes essentielles : (i) extraction des caractéristiques (indicateurs), (ii) classification et détection de défaut et enfin (iii) suivi des caractéristiques géométriques de la deuxième classe qui correspond au roulement défectueux.

3.3.1 Extraction des caractéristiques

La première phase de la conception d'un système de diagnostic vibratoire dynamique par RdF, consiste à extraire des caractéristiques physiques à partir des signaux vibratoires issus de différents capteurs placés sur les machines à surveiller. Le but de cette phase de diagnostic est de donner une image sur les efforts dynamiques engendrés par les pièces en mouvement ou par les sources extérieurs et donc définir l'état de santé actuel des composants (sains, défectueux ou cassés) , de suivre l'évolution de ces caractéristiques dans le temps et d'identifier l'origine de ces évolutions (défaut ou changements des paramètres intrinsèques de la machines : vitesse, charges ou autres). L'analyse de ce type de signaux permet d'extraire des indications correspondantes aux vibrations provoquées par les pièces de la machine.

Pour extraire ces caractéristiques, on peut cependant avoir recours à des techniques de traitement du signal couramment utilisées dans l'analyse vibratoire pour extraire les indicateurs de défaut (cités dans le chapitre I).

3.3.2 Détection de défaut par la méthode de classification dynamique

Pour traiter les données d'un système évolutif, il faut choisir une méthode de classification dynamique capable de traiter ce type de données en ligne, de détecter les évolutions des classes, les évolutions à l'intérieur de chaque classe et s'auto-adapter en conséquence. Dans cette section nous présentons les méthodes de classification dynamiques développées durant cette thèse : le principe, la démarche ainsi que les motivations du développement de chaque méthode développée.

Après l'arrivée de chaque nouveau signal et après l'extraction des caractéristiques, les nouvelles observations, les m dernières observations ainsi que des informations sur les résultats précédents de la classification sont récupérées d'une mémoire temporaire. La première opération effectuée consiste à vérifier si les nouvelles observations sont des nouveautés (ou outliers) pour le système (c'est-à-dire qu'elles sont 'différentes' des observations précédentes) ou sont des observations 'ordinaires' (c'est-à-dire qu'elles ressemblent aux observations précédentes). Si les observations sont considérées comme nouveautés et que ces nouveautés sont confirmées (c'est-à-dire qu'elles ne sont pas liées à la présence de bruit mais à un changement de mode de fonctionnement), une nouvelle classe est créée et le processus de classification est arrêté. Si la nouveauté n'est pas confirmée, les observations seront stockées dans un buffer pour vérifier si les observations futures vont confirmer ce changement ou non. Si les observations ne sont pas considérées comme des nouveautés, elles seront attribuées à la classe la plus proche puis une vérification est établie pour voir si l'affectation des nouvelles observations a changé la disposition des classes (par exemple une augmentation de la similarité des classes qui nécessite la fusion des classes similaires) ou si des classes devaient être oubliées.

Après l'affectation des observations à la classe la plus appropriée et l'adaptation de la partition, une décision sur le mode de fonctionnement identifié sera établie.

3.3.3 Adaptation de la méthode RT-OPTICS

Nous utilisons l'algorithme OPTICS pour les raisons suivantes : Premièrement, le graphe des distances d'accessibilité fournit par l'algorithme OPTICS est plutôt insensible aux choix des paramètres d'entrée (\mathcal{E} et $MinPts$). En général, les valeurs de ces paramètres doivent

être suffisamment grandes pour conduire à un bon résultat. Deuxièmement, la visualisation graphique de l'ordre produit par cette méthode est indépendante de la dimension de l'ensemble des données. L'algorithme OPTICS permet de donner un ordre pour tous les objets dans la base de données où nous pouvons extraire les clusters de différentes formes et densités. Par conséquent, la méthode OPTICS consiste à déterminer les zones de hautes et celles de basses densités. La première étape de cette méthode suppose que toute combinaison de zones hautes et basses forme (pics) un cluster et la deuxième étape essaie de filtrer les clusters. Les limites de cette méthode résident dans le fait qu'elle n'élimine que certains clusters et dans le cas des graphes des distances d'accessibilité avec plusieurs pics, (voir Figure 3-4), il reste toujours plusieurs groupes qui ne sont pas réellement des clusters. Pour cette raison cette méthode n'est pas efficace pour la détection automatisée des clusters à partir de cet ordre des clusters.

3.3.3.1 Présentation de la méthode RT-OPTICS

Real Time OPTICS (RT-OPTICS) est une méthode non supervisée (benmahdi et al. 2017) développée pour suivre l'évolution de l'état de santé de roulement (signaux simulés et signaux issus d'un banc expérimental).

Le choix de la méthode de suivi de l'évolution des données par méthode de classification dynamique a été établi après des essais de plusieurs méthodes de classification ; KNN, Kmeans et DBSCAN. L'étude a montré que la première méthode nécessite une connaissance à priori de la cinématique de la machine et une construction d'une base d'apprentissage. La deuxième méthode est meilleure que la première mais présente également un inconvénient qui est la fixation du nombre de classe et aussi le choix des centres initiaux cela ne permet donc pas une classification dynamique des données. La troisième méthode est justifiée par Kerroumi et al. (2013). Cette méthode est une méthode de classification basée sur la densité, elle cherche automatiquement le nombre de clusters appropriés. L'inconvénient majeur de cette méthode réside dans le choix du paramètre d'entrée qui est le rayon de voisinage ϵ , Si sa valeur est importante l'ensemble des clusters contenus dans la base de données seront regroupés dans une seule classe (ou cluster). Si la valeur d'épsilon est faible la majorité des points vont être considérés comme du bruit. De cette étude des trois méthodes, nous avons conclu que la meilleure méthode de séparation dans le cas des signaux vibratoires est celle basée sur la densité (Kerroumi et al., 2013). Pour remédier au problème de la méthode DBSCAN, nous avons proposé une nouvelle méthode de clustering basée sur la densité, OPTICS (Ordering Points To Identify the Clustering Structure) (Ankerst et al., 1999). Cette méthode n'a jamais été utilisée auparavant dans le diagnostic des roulements.

La méthode de classification OPTICS est une méthode de type non supervisée qui effectue la classification en se basant sur la densité des points. Cette méthode est bien connue pour sa capacité de visualiser les régions plus denses par rapport à ceux de faible densité même pour les grandes bases de données spatiales contrairement aux méthodes de partitionnement. Cette méthode n'est pas influencée par le choix des paramètres initiaux, le rayon de voisinage ϵ et le nombre minimum de points *MinPts*.

3.3.3.2 Nouvelle méthode pour le calcul des clusters

Notre méthode consiste à calculer les zones basses et les zones hautes et cela en créant une matrice colonnes [DebutCluster, FinCluster] d'où DebutCluster représente l'ordre du début du cluster ou bien les pics bas et le FinCluster représente l'ordre de la zone haute. Le seuil est défini quand cluster1 est égale à 1 c'est-à-dire tous les points bordures sont pris en considération et aucun point n'est considéré comme du bruit. Toute combinaison de zones qui commence par une zone basse et se termine par une zone haute et dont la hauteur est supérieure ou égale au seuil est considérée comme un cluster. Les points qui se situent entre une zone basse et haute qui vérifient la condition du seuil appartiennent sûrement au cluster, ainsi que les points des zones basses. Nous allons sélectionner parmi ces points ceux appartenant au cluster.

Dans la Figure 3-4, un cluster commence à partir du point n°2 et se termine au point n°16. Le point 2 fait partie du cluster, car sa valeur d'accessibilité élevée indique qu'il est loin du point 1. Il doit être proche du point 3, parce que le point 3 a une valeur d'accessibilité faible indiquant qu'il est proche du point 2. Un argument similaire s'applique pour le point 16. Pour automatiser ce processus, nous avons besoin de définir un critère d'arrêt mais aussi un critère pour distinguer entre une création d'un nouveau cluster qui est dû à l'effet de charge, de l'effet de la vitesse de rotation de moteur ou bien à une dégradation de roulement. Un saut brusque du cluster vers une des coordonnées ou vers les deux coordonnées en sens négatif (ici, nous avons étudié la classification avec une base de données en 2D, la valeur RMS et le kurtosis) signifie qu'il y a un changement dans la cinématique de la machine (figure 3-5).

Le kurtosis et la valeur efficace sont parmi les indicateurs temporels les plus sensibles à la détection des chocs dans les premiers stades de défaillance. Pour cette raison, on peut constater que le saut ou bien l'apparition du deuxième cluster évolue dans le sens positif des deux coordonnées en même temps.

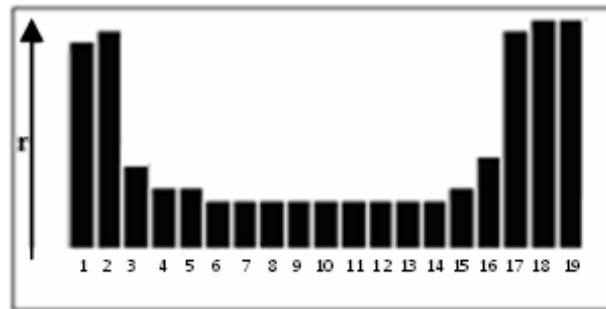


Figure 3-4: Extraction d'un cluster (Ankerst et al., 1999).

DebutCluster : contient le début de chaque zone basse et la fin de chaque région haute.

FinCluster : contient les valeurs d'accessibilité du début de chaque zone basse et du premier point après la fin de chaque zone haute qui forme le cluster.

n : le nombre d'objet dans la base de données.

Etape 1

indice =0, DebutCluster= \emptyset , FinCluster= \emptyset ; OrdreCluster= \emptyset ;

r (p): la distance d'accessibilité du point p

OrdreCluster est le vecteur de l'ordre des distances accessibilité.

Tant que indice < n faire.

 Si une région basse commence à indice

 Si indice $\neq 0$

 son ordre sera ajouté à OrdreCluster

 Debut =r (indice) ; on parcourt les points suivants dans l'ordre des objets.

 Fin=r(indice+1) ;

 Ajouter la valeur de Début et de Fin à DebutCluster et FinCluster et continuer à droite.

 - Sinon incrémenter indice.

Etape 2

- on calcule le centre de gravité de l'ensemble de données classifié en temps t.
- On calcul le centre de gravité C_a des dix nouveaux points de la nouvelle acquisition, résultant des dix tranches du signal.
- On calcule le centre de gravité de l'ensemble de données classifiés en temps $t=t+1$.
- On calcul le centre de gravité C_c des dix derniers points classés dans le fichier OrderCluster.
- On calcul l'écart entre le centre C_a et le centre C_c
- Si l'écart entre ces deux centres est non nul.
- Ou bien si la différence entre les coordonnées du centre C_c et celles du centre C_a est négative :
- On continue le processus de classification sinon on arrête et commence le calcul des caractéristiques géométriques de la deuxième classe à savoir le saut du centre, le contour, la densité, la surface et la distance maximale entre l'ensemble des

Figure 3-5: Algorithme principal de RT-OPTICS (Benmahdi et al. 2017)

L'algorithme est résumé sur la figure 3-6.

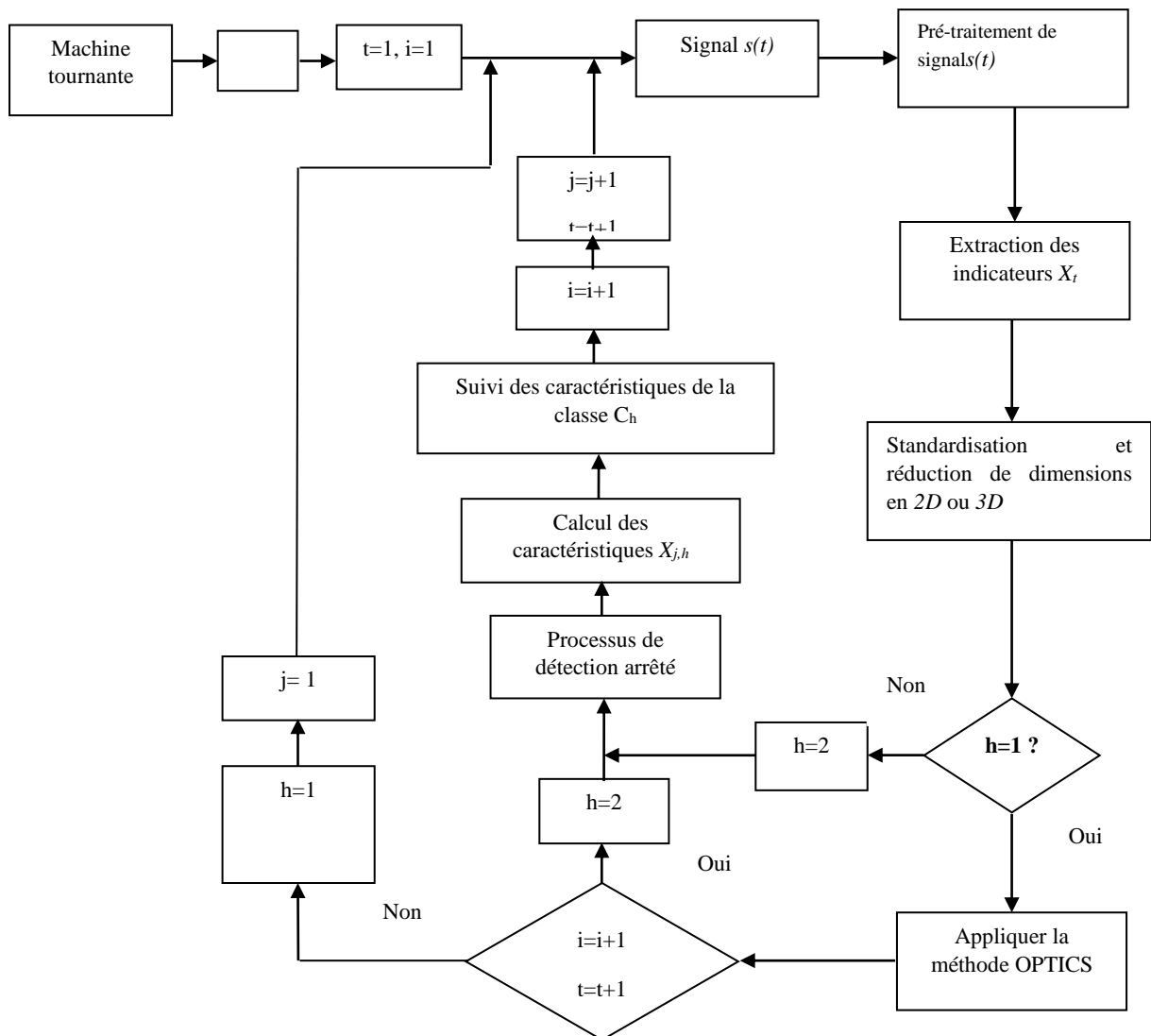


Figure 3-6: Organigramme de l'algorithme principal

3.3.3.3 Détection du défaut

Les présents travaux se concentrent sur le déploiement d'une méthode de classification dynamique. Ainsi, les indicateurs Kurtosis et RMS seront choisis pour leur performance individuelle pour une détection de défauts.

Nous voulons que l'algorithme fonctionne en temps raisonnable même pour de grandes bases de données. L'algorithme OPTICS trace pour chaque vecteur de la base de données la distance-accessibilité en fonction de son ordre en forme d'endogramme, et cela s'exécute dans une durée plus ou moins longue. Dans notre algorithme cette étape a été éliminée et remplacée par des courbes de Kurtosis en fonction de la valeur efficace pour qu'elles soient

lisible pour une personne non expertisée dans le domaine. En RT-OPTICS une stratégie simple a été employée comme critère d'arrêt : les indicateurs commencent à être dans un ordre croissant dès l'apparition de premier défaut en particulier le Kurtosis. Cette propriété a été exploitée pour détecter la naissance de défaut c.-à-d. le premier cluster.

La méthode OPTICS est appliquée à la matrice d'indicateurs obtenue pour détecter la présence de défaut dans le roulement. La notion du temps t est introduite pour traiter les données en temps réel. La méthode OPTICS s'auto exécute à chaque nouvelle prise de mesure jusqu'à la détection de défaut. La détection de défaut se traduit par l'apparition d'une nouvelle classe (C2) différente de la classe représentant le roulement jugé sain (C1). Donc deux classes C1 et C2 sont détectées. Après la détection, le suivi sera effectué à partir des paramètres caractéristiques de la matrice d'observation (indicateurs de défaut) représentant la classe C2. Nous avons choisi comme paramètres de suivi la densité Den_h , (Eq. 3.1), la distance maximale entre tous les éléments de la classe, le saut du centre de gravité de la matrice d'indicateurs en termes de distance euclidienne (Eq. 3.3), le contour c_h (Barber et al. 1996) et la surface.

La figure 3-6 résume l'algorithme principal d'adaptation de la méthode OPTICS pour la détection et le suivi de roulement. Cet algorithme comporte deux boucles imbriquées. La boucle principale consiste en l'acquisition et le prétraitement des signaux ainsi que l'extraction des indicateurs X_t (normalisation et réduction des dimensions de la base de données par la sélection des indicateurs). Le processus de classification commence avec la deuxième boucle. Il s'arrête lorsque le défaut est détecté ($h=2$, *deuxième classe détectée*). Les paramètres définissant la classe de roulement défectueux (C2) seront ensuite calculés pour chaque enregistrement pour suivre le degré de dégradation du roulement.

3.3.4 Suivi des caractéristiques géométrique

Une fois le défaut sera détecté, l'étape de la classification sera mise en arrêt et on continuera notre diagnostic en calculant les paramètres géométriques de la classe caractérisant le roulement défectueux. Les différents paramètres sont calculés comme il suit :

3.3.4.1 Enveloppe convexe (contour)

L'enveloppe convexe c_h de la base de données est le plus petit contour qui contient tous les points, Barber et al. 1996). L'algorithme de contour est présenté sur la figure 3-7.

```
Créer un simplexe de  $d + 1$  points
pour chaque facette F
pour chaque point non affecté p
si p est supérieur à F
    assigner p à l'ensemble extérieur de F
pour chaque facette F avec un ensemble extérieur non vide
    sélectionner le point le plus éloigné p du jeu extérieur de F
    initialiser l'ensemble visible V à F
    pour tous les voisins non visités N de facettes en V
        si p est supérieur à N
            ajouter N à V
    la limite de V est l'ensemble des arêtes d'horizon H
    pour chaque crête R en H
        créer une nouvelle facette à partir de R et p
        lier la nouvelle facette à ses voisins
    pour chaque nouvelle facette F '
        pour chaque point non affecté q dans un ensemble extérieur d'une facette dans V
            si q est supérieur à F '
                assigner q au jeu extérieur de F
    supprimer les facettes en V
```

Figure 3-7: Algorithme de l'enveloppe convexe (contour)

Son principe de consiste en premier lieu, de trouver le point le plus à gauche P, et le point le plus à droite Q (en cas d'égalité on choisit de manière arbitraire). Les points P et Q appartiennent à l'enveloppe convexe (Figure 3-8). On peut alors résoudre le problème, en trouvant l'enveloppe convexe des points au dessus de la droite (PQ) et l'enveloppe des points en-dessous de la droite, puis en contenant les listes de points (en ne répétant pas P et Q). Les sous-problèmes ont alors une forme plus restreinte que instance d'origine: on dispose de deux points sur une droite, telle que tous les points sont du même côté de la droite, disons à gauche de (PQ), et tous les points qui appartiennent à la droite sont dans le segment [PQ]. On trouve alors le point H le plus éloigné de la droite. L'enveloppe convexe de l'enveloppe de point est la concaténation de deux ensembles de points: l'enveloppe convexe des points à gauche de (PH) et l'enveloppe des points à gauche de (HQ). On peut calculer récursivement ces ensembles (ils sont dans la même configuration que précédemment).

L'enveloppe convexe est résumé en matlab par la fonction «convhulln».

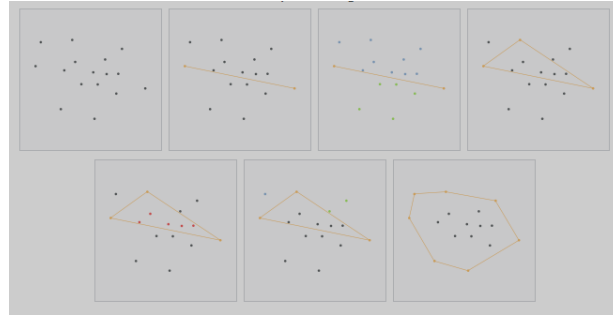


Figure 3-8: les différentes étapes de l'algorithme de l'enveloppe convexe

3.3.4.2 Densité

La densité d'un ensemble j de points est définie comme étant le rapport entre le nombre de points distribués dans un espace métrique N_h . La densité de la classe C_2 est définie par :

$$Den_j = \frac{N_h}{|A_h|} = \quad (3.1)$$

A_h est la surface de la classe h . On a considéré la surface d'un cercle en 2D (2 indicateurs) et la surface d'une sphère en 3D (3 indicateurs) de rayon d_{max} (distance maximale entre tous les éléments de la classe par rapport au centre de gravité). Le rayon est défini par la relation suivante :

$$d_{max} = \max(dist(x_j, C_h)) \quad (3.2)$$

3.3.4.3 Saut du centre

Le saut du centre est la distance intra-classe entre le centre de la classe actuelle et celui de l'ancienne classe. La présente classe est composée de $n=10$ points qui correspondent aux nombres de tranches de signaux. En $t=t-1$ la classe comporte $n.m$ points.

Le saut du centre de gravité de la classe C_2 entre le temps t et le temps $t+1$ est exprimé par la distance euclidienne entre $Cent_{2,t}$ et $Cent_{2,t+1}$ (3.4) :

$$Dist(C_{2,t}, C_{2,t+1}) = \sqrt{d^2(Cent_{2,t}; Cent_{2,t+1})} \quad (3.3)$$

$Cent_h$ est le centre du $h^{ème}$ cluster qui est définie par:

$$Cent_h = \frac{1}{N_h} \sum_{j=1}^{N_h} x_{h,j} \quad (3.4)$$

avec N_h est le nombre d'objets du cluster C_h , x_{hj} est le $j^{ème}$ objet du $h^{ème}$ cluster.

3.3.4.4 Distance maximale d_{max}

Nous avons calculé aussi un autre paramètre qui est la distance entre le point le plus proche à l'origine et le point le plus éloigné. Ce paramètre peut nous donner une idée sur l'élargissement de la deuxième classe, la classe de roulement défectueux.

3.3.4.5 La surface

Seulement deux indicateurs ont été pris en considération dans la phase de clustering par classification, la valeur efficace et le kurtosis. De ce fait, la surface calculée est la surface d'un cercle d'un rayon de valeur égale à la distance entre tous les points constituant cette classe. Vu la forme aléatoire de la classe, la surface calculée n'est pas la surface réelle occupée par ces indicateurs.

3.3.5 Limitations de la classification par RT-OPTICS

La méthode de classification RT-OPTICS est une méthode de classification par densité. Cette méthode est du type non supervisé, elle ne nécessite pas donc une base d'apprentissage. RT-OPTICS a été conçue pour pouvoir classifier les observations reçues en deux classes reçues au fur et à mesure. Ainsi, des modifications majeures ont été effectuées sur la méthode de la base OPTICS. La première modification est la conservation des outiliers. Une étude comparative de calcul des distances entre les points en utilisant la distance euclidienne et la distance mahalanobis a été faite. D'après le test, on a conclu que la distance euclidienne est la plus performante pour cette méthode de classification. Et finalement l'intégration des outils de suivi des évolutions de la partition obtenue dans le temps.

Le *RT-OPTICS* permet de classifier dynamiquement des données en détectant les évolutions des classes et de mettre à jour la partition au fur et à mesure, tout en améliorant la qualité de la classification.

Pour montrer l'efficacité de la méthode OPTICS, on va tracer le résultat de classification par le tracé de l'accessibilité des points. Cela montre que la distinction entre les

régions les plus denses ne nécessite pas la connaissance ou le choix optimal des paramètres initiaux

3.4 DDBSCAN

Kerroumi et al. (2013) ont proposé une méthode de classification dynamique basée sur le principe DBSCAN pour suivre l'état de dégradation de roulement. Ils ont utilisé l'indicateur scalaire RMS calculé pour des signaux réels prélevés par deux capteurs piézoélectriques placés axialement et radialement sur un banc expérimental. Une étude comparative entre cette méthode et la méthode K-means a montré que le DBSCAN est le plus adapté aux signaux vibratoires. Vu la difficulté de son exploitation dans le domaine industriel, ils ont intégré la notion du **temps** dans l'algorithme pour remédier son aspect statique. L'exécution de l'algorithme proposé a montré clairement la progression des défauts où chaque classe formée correspond à un nouvel état de dégradation.

L'inconvénient majeur de la méthode DBSCAN réside dans le choix du rayon ϵ ; si sa valeur est trop petite la majorité des individus vont être considérés comme des bruits et si elle est importante les différents clusters peuvent être regroupés ensemble.

3.5 Conclusion

Dans ce chapitre, nous avons présenté deux méthodes de classification non supervisée basées sur la notion de densité, OPTICS et DBSCAN. Ces deux méthodes ne nécessitent pas un nombre important de paramètres à l'entrée des algorithmes. Ces méthodes requièrent les concepts suivants : densité, noyau, point limite, accessibilité et connexité qui sont liés à la notion de plus proches voisins d'un point. Une classe croit dans n'importe quelle direction où se trouve la densité. Et donc les algorithmes correspondant sont aptes à découvrir des classes de formes arbitraires et fournissent une protection naturelle contre les outliers. Enfin ils peuvent être utilisés pour des données de grande taille sans poser de problème particulier. Notre contribution est d'utiliser ces deux méthodes pour le pronostic des signaux de roulement.

Chapitre 4 :

Contribution à l'étude de diagnostic de défauts par la
méthode de classification dynamique

4 Validation numérique et expérimentale de la méthode RT-OPTICS

4.1 Introduction

Pour permettre la validation de la méthode adaptée, deux types de validation ont été mis en œuvre : une validation numérique en créant une base de données à partir d'un modèle de défaut de roulement et une validation expérimentale de l'approche proposée.

Le diagnostic sera effectué en deux étapes : (i) classification et le suivi. Le but de l'étape de classification consiste à détecter la naissance du défaut sur la bague extérieure de roulement à galet ensuite effectuer le suivi des caractéristiques de la classe représentant le défaut plus facilement pour chaque reprise de mesures. (ii) Une fois le défaut est détecté, le processus de classification sera arrêté et le suivi débute par le calcul des caractéristiques géométriques de la classe caractérisant le défaut.

Une classification statique en utilisant la méthode classique OPTICS sera réalisée pour la visualisation de la distribution des différentes zones de densité dans la deuxième classe. Une étude comparative entre la méthode RT-OPTICS et la méthode DDBSCAN (Kerroumi et al., 2013) sera également étudiée dans cette partie.

4.2 Validation numérique de la méthode RT-OPTICS

4.2.1 Création du modèle mathématique

Le comportement dynamique des roulements et des éléments roulants des machines tournantes sous défauts localisés a été un sujet de recherche intensive, menant à un certain nombre de modèles mathématiques qui reflètent leurs mouvements (Yiakopoulos et al., 2011). Pour la génération d'un signal temporel comportant un défaut, le modèle mathématique obtenu réalisant un produit de convolution entre un signal d'une force excitatrice et la réponse impulsionnelle de la structure est présentée.

Des modèles du comportement dynamique des roulements sous défauts localisés (défaut de bague extérieure, défaut de bague intérieure et défaut de bille) ont été développés par (Antoni & Randall, 2002 ; MacFadden & Smith, 1984b ; Randall, 1987). Les éléments de base d'un tel modèle sont comme suit : Les impacts répétitifs produits par un défaut de roulement localisé peuvent être décrits en un train de delta de Dirac de fonctions $d(t)$ de la forme suivante :

$$d(t) = \begin{cases} q(t)d_0 \sum_{k=0}^N \delta(t - kT_d) \text{ cas de défaut sur la bague intérieure} \\ d_0 \sum_{k=0}^N \delta(t - kT_d) \text{ cas de défaut sur la bague extérieure} \end{cases} \quad (4.1)$$

Où d_0 est l'amplitude de la force d'impulsion caractérisant la gravité du défaut. La période de répétition des impulsions T_d est l'inverse de la fréquence caractéristique de passage des billes sur la bague extérieure (BPFO : $f_{bpfo} = 1/T_d$) ou bien sur la bague intérieure du roulement (BPFI : $f_{bpfi} = 1/T_d$), selon le type de défaut. Ces deux fréquences sont proportionnelles à la vitesse de rotation de l'axe f_{shaft} , et leur valeur dépend des caractéristiques géométriques de roulement. La fonction $q(t)$ est la répartition de la charge autour des éléments roulants de roulement sous la charge radiale qui est typiquement approximée par l'équation bien connue de Stribeck (Yiakopoulos et al., 2011):

$$q(t) = \begin{cases} q_0 [1 - (1/2\varepsilon)(1 - \cos \theta)]^n \text{ pour } |\theta| < \theta_{max} \\ 0 \text{ ailleurs} \end{cases} \quad (4.2)$$

Avec q_0 est l'intensité de charge maximum, ε est le facteur de la répartition des charges, θ_{max} est l'ampleur angulaire de la zone de charge et $n = 3/2$ pour les roulements à billes. Les termes q_0 , ε et θ_{max} sont tous des fonctions du jeu diamétral du roulement et de la charge appliquée. Pour un roulement avec un jeu positif, $\varepsilon < 0.5$ et le $\theta_{max} < \pi/2$.

Cependant, des variations significatives de la réponse dynamique d'un élément roulant défectueux peuvent exister lorsque certains paramètres de l'équation ci-dessus sont supposés présenter plutôt un comportement stochastique qu'un comportement déterministe. L'amplitude des impacts dépend principalement de la répartition de la charge autour de la circonférence du roulement, ainsi que d'autres paramètres, tels que la variation de la rigidité dynamique de l'assemblage, l'ondulation des éléments et des bagues de roulement et l'existence de billes hors-taille dans le complément de balle. Par conséquent, la série d'impacts devrait être plutôt considérée comme modulée de façon aléatoire en amplitude.

Parallèlement, les roulements à rouleaux présentent un mouvement de glissement qui introduit des effets de non-linéarité dans le système. En conséquence, les impulsions ne se produisent jamais exactement à la même position d'un cycle à l'autre.

De là, un modèle plus réaliste pour le train d'impact des éléments roulant d'un roulement défectueux devrait être de la forme suivante (Yiakopoulos et al., 2011):

$$D(t) = \begin{cases} q(t) \sum_{k=0}^N A_k \delta(t - kT_d - \tau_k) \text{ cas de défaut sur la bague intérieure} \\ d_0 \sum_{k=0}^N A_k \delta(t - kT_d - \tau_k) \text{ cas de défaut sur la bague extérieure} \end{cases} \quad (4.3)$$

D'où A_k sont des variables aléatoires de l'amplitude de la $k^{\text{ième}}$ force d'impulsion avec une fonction de densité de probabilité généralement supposée être normale (gaussienne), τ_k est une variable aléatoire pour le décalage entre deux impacts dû à la présence d'un glissement, qui est généralement supposé être de moyenne nulle et de fonction de densité de probabilité normale (gaussienne).

Lorsqu'un train d'impulsions est appliqué à l'élément roulant de roulement, il provoque des résonances structurelles. La réponse structurelle impulsive d'un tel système à chaque impulsion peut être exprimée par la fonction suivante:

$$x(t) = \sum_{i=1}^M B_i e^{-2\pi\xi_i f_{ni} t} \quad (4.4)$$

$$f_{oi} = f_{ni} \cdot \sqrt{1 - \xi_i^2} \quad (4.5)$$

où $i = 1, \dots, M$ sont les modes excités, et pour chaque mode i , f_{ni} est la fréquence de résonance et ξ_i est le facteur d'amortissement. Par conséquent, la réponse dynamique $x(t)$ résultant d'un défaut induit dans un roulement peut être exprimée par :

$$d(t) = \begin{cases} \sum_{k=0}^N q(t) A_k \delta(t - kT_d - \tau_k) \otimes \sum_{i=1}^M B_i e^{-2\pi\xi_i f_{ni} t} \cos(2\pi f_{oi} t) + br(t) & \text{cas de défaut sur la bague intérieure} \\ \sum_{k=0}^N A_k \delta(t - kT_d - \tau_k) \otimes \sum_{i=1}^M B_i e^{-2\pi\xi_i f_{ni} t} \cos(2\pi f_{oi} t) + br(t) & \text{cas de défaut sur la bague extérieure} \end{cases} \quad (4.6)$$

Où le symbole \otimes dénote la convolution et $br(t)$ est un bruit de fond additif, qui tient compte les effets des autres sources de vibration du système et de l'environnement externe. C'est le bruit imprévisible de mesure qui est présent dans n'importe quel système de mesure pratique.

Une application du modèle de l'équation (eq.4.6) pour générer la réponse de vibration attendue d'un roulement défectueux avec un défaut sur la bague externe et intérieure. Le roulement tourne à 25 Hz, le signal est conditionné avec une fréquence d'échantillonnage de $f_e = 20$ KHz. Comme il peut être observé, un phénomène de modulation caractérise les formes d'onde du domaine temporel, tandis que la réponse du domaine fréquentiel est caractérisée par des bandes latérales autour de la fréquence de résonance structurelle, espacées de distances

égales à la fréquence *BPFO* ou *BPFI* correspondante. De plus, dans le cas du défaut de la bague intérieure, il existe des bandes latérales supplémentaires autour des fréquences de défaut de la bague intérieure, espacées à des distances égales à la vitesse de rotation de l'arbre....

4.2.2 Rapport signal sur bruit RSB

Le bruit étant présent dans toutes les mesures, il est important de savoir simuler des signaux bruités afin de vérifier la pertinence des méthodes d'analyse. Un bruit blanc gaussien, qui représente le modèle le plus fréquent de bruit est généré sous MATLAB par la fonction *randn*. Cette fonction permet de générer une matrice ou un vecteur aléatoire dont les éléments correspondent à un bruit blanc, au sens où ils ne sont pas corrélés entre eux (il n'y a pas de lien entre deux échantillons). La variance de bruit est simplement fixée par un facteur multiplicateur. Ce facteur est déterminé empiriquement pour un rapport signal sur bruit *RSB* inférieur ou égal à -20 pour une amplitude du signal de l'ordre de 10^{-17} .

Le rapport signal sur bruit (*RSB*) est une quantité mesurée en *dB* qui donne un indice de la puissance du signal par rapport à celle du bruit. Il conditionne toute observation ; s'il est faible on ne voit rien. Il est mesuré par la formule :

$$RSB = 10 * \log_{10} \left(\frac{\text{Energie du Signal}}{\text{Energie de Bruit}} \right) \quad (4.7)$$

La puissance moyenne d'un signal centré est mesurée sous MATLAB par le carré de la fonction *std*. La puissance moyenne est aussi appelée variance pour les signaux aléatoires de moyenne nulle, telle que le bruit.

Un *RSB* positif signifie que le signal est plus énergétique que le bruit. Pour un *RSB*=0, il y a autant de bruit que de signal et pour les valeurs négatives du *RSB*, on dit que le signal est dégradé. Le bruit est généré par l'équation :

$$n(t) = m(\text{randn}(1, N)) \quad (4.8)$$

L'organigramme suivant résume la procédure de création de la base de données ainsi que le bruit généré.

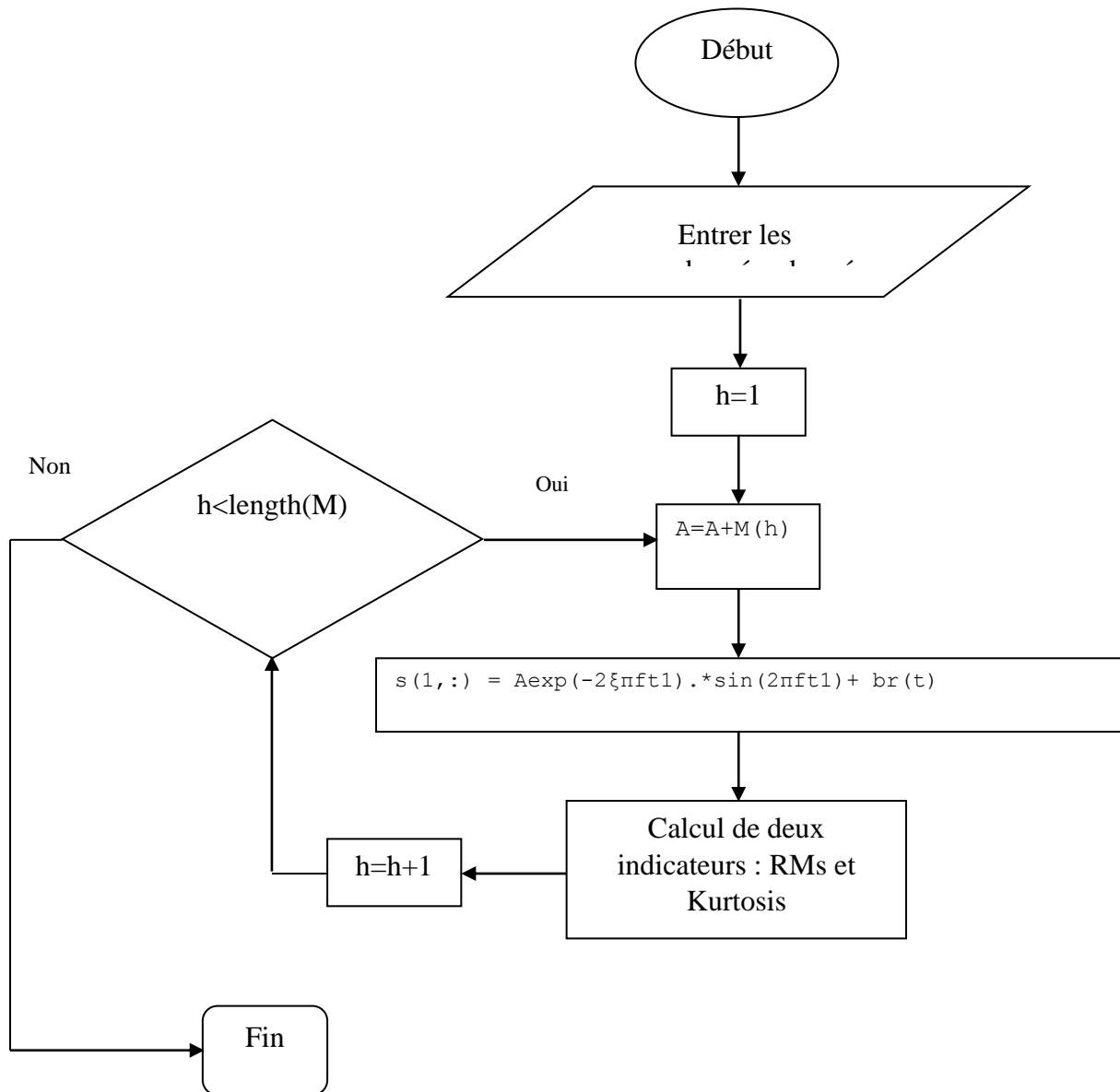


Figure 4-1: Création de la base de données

Le bruit est gaussien. Pour éviter toutes coïncidences entre ce dernier et le signal choc (signal avec une amplitude A très faible de l'ordre de 10^{-7}), on a gardé le même signal $randn(1,N)$ pour toutes les tranches et on a fait varier le facteur m d'un petit pourcentage pour chacune d'elles. Et cela pour obtenir variance du bruit variable. Ce bruit sera utilisé pour tous les signaux.

4.2.3 Validation numériques

La méthode RT-OPTICS est appliquée à la détection et le suivi de défaut de roulement simulé. Cette simulation permettra de voir l'influence des paramètres de simulation sur la méthode OPTICS. Un modèle mathématique représentant la signature vibratoire d'un

roulement avec un défaut sur la bague extérieure (*BPFO*) est utilisé. Ce modèle décrit l'impact produit par un défaut de bague extérieure à chaque passage d'un élément roulant en fonction du temps t (Chieumentin, 2007). Il est décrit par les équations suivantes :

$$x(t) = \begin{cases} A \exp(-\xi 2\pi f t) \sin(2\pi f t) \otimes \sum_{k=1}^N \Pi_{T_{BPFO}} \left(t - \frac{k}{f_{BPFO}} \right) + br(t) & \text{si } t \geq 0 \\ 0 & \text{si } t < 0 \end{cases} \quad (4.9)$$

$$f_{BPFO} = \frac{1}{2} f_r n_b \left(1 - \frac{d_{bille}}{D_m} \cos \alpha \right) \quad (4.10)$$

A est l'amplitude du signal, ξ l'amortissement, f la fréquence de résonance structurelle, k est le nombre d'harmoniques (impulsions) ; T_{BPFO} est la période des impulsions ; f_{BPFO} est la fréquence caractéristique de défaut bague extérieure. Un défaut sur la bague extérieure d'un roulement à billes de type SKF6206 avec une vitesse de rotation f_r de 25 Hz est simulé. Les caractéristiques sont dans le tableau 4-1. Les signaux sont échantillonnés à la fréquence d'échantillonnage de 20kHz pour un nombre de points de 8192 points (N). Le bruit blanc $br(t)$ a pour valeur de facteur de 0.1.

D	diamètre extérieur, $D= 62 \text{ mm}$
d	diamètre intérieur, $d= 30 \text{ mm}$
D_m	diamètre primitif, $D_m= (D +d)/2$
n_b	nombre de billes, $n_b =9$
d_{bille}	diamètre des billes, $d_{bille} = 9.525 \text{ mm}$
α	angle, $\alpha=0^\circ$

Tableau 4-1: Dimensions de roulement SKF6206.

Dans ce modèle, pour un roulement donné, trois paramètres peuvent influencer la signature vibratoire : l'amplitude A , la vitesse de rotation f_r et l'amortissement ξ . A partir du modèle mathématique et pour chaque condition, cinq signaux sont calculés dont la variabilité dépend du bruit $br(t)$. Cela permettra d'obtenir un nuage de points représentant les indicateurs qui résultent de la variation de ces paramètres. La méthode OPTICS proposée sera appliquée

pour la détection de la deuxième classe (C2) ; en pratique la première classe représente le roulement sain et la deuxième classe représente la naissance de défaut, et le suivi de la classe (C2). La figure 4-2 montre les signaux des deux types de défauts créés, bague intérieure et bague extérieure, pour une amplitude $A=2$. Dans la suite l'amplitude sera considérée comme la taille de défaut.

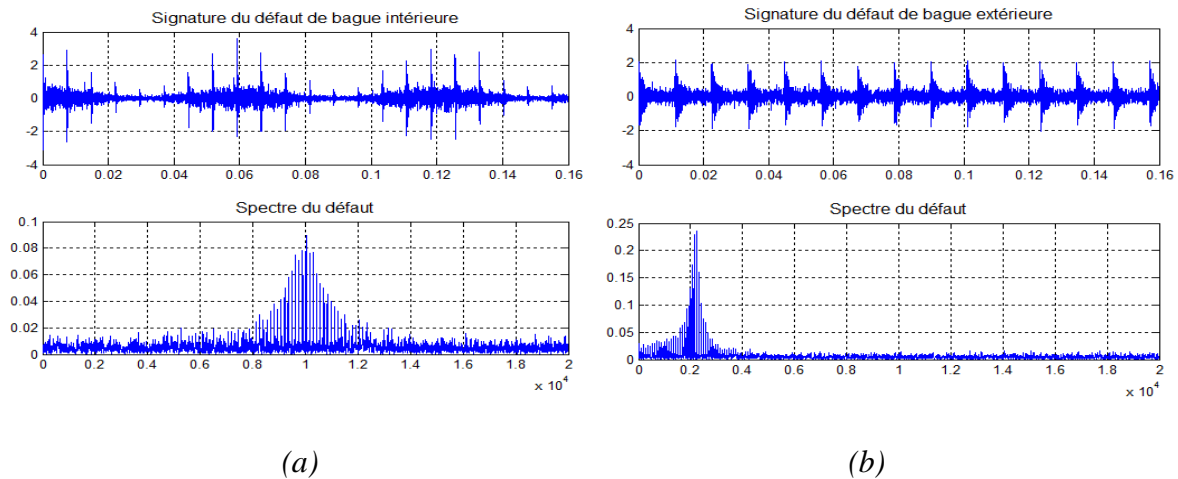


Figure 4-2: Signaux simulés (a) bague intérieure ; (b) bague extérieure, pour une amplitude $A=2$

4.2.3.1 Calcul d'indicateurs scalaires

Pour vérifier la sensibilité de l'approche temporelle dans la détection des défauts de roulement, plusieurs indicateurs scalaires ont été calculés dans différentes configurations prises dans la simulation.

La figure 4-3 montre l'évolution de la valeur efficace RMS en fonction du kurtosis. Les deux indicateurs sont sensibles à l'augmentation de l'amplitude. En théorie, pour les défauts localisés, le défaut est détecté pour une valeur de kurtosis supérieure à 3 (Honarvar & Martin, 1997). Visuellement, l'apparition des chocs est claire pour une valeur du kurtosis de 3.26. A ce stade, l'ensemble des points représentant chaque variation d'amplitude commence à être distingué, ainsi on peut admettre que la détection est obtenue pour une amplitude de $5 \cdot 10^{-2}$ (signal n°12). Le kurtosis tend vers une stabilisation alors que la valeur efficace augmente graduellement avec l'augmentation de l'amplitude.

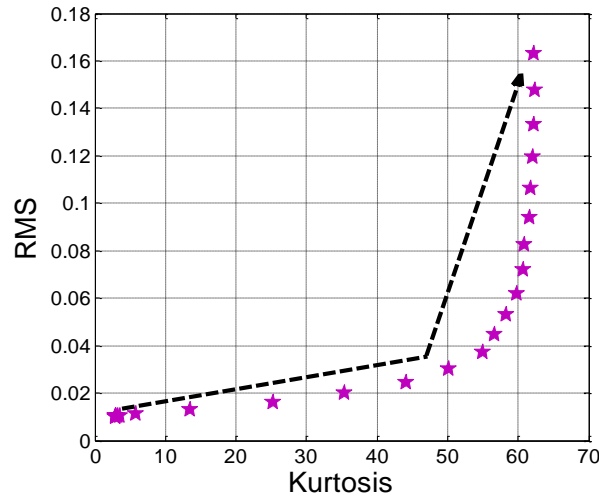


Figure 4-3: Évolution de la valeur efficace en fonction du kurtosis lors d'une augmentation de l'amplitude pour un défaut de bague extérieure

4.2.3.2 Influence de la variation de l'amplitude

Dans notre étude, on a choisi de simuler le défaut de bague extérieure de roulement. Pour se faire, on fait varier l'amplitude du signal (A). Pour chaque amplitude on découpe le signal en cinq tranches et on calcule deux indicateurs statistiques (kurtosis et RMS).

Pour étudier l'influence de l'amplitude, l'amplitude A prend successivement 30 valeurs : 0, 10^{-6} , $2 \cdot 10^{-6}$, 10^{-5} , $2 \cdot 10^{-5}$, 10^{-4} , $2 \cdot 10^{-4}$ et de 10^{-3} jusqu'à 0.23 avec un pas de 0.01). Les signaux sont numérotés en fonction de l'amplitude A dans l'ordre croissant. Le signal n°1 correspond à l'amplitude $A = 0$. La vitesse de rotation est égale à $fr = 1000 \text{ tr/mn}$ et le taux d'amortissement $\xi = 0.05$.

La figure 4-4 (4-4.a, 4-4.b et 4-4.c) montre les signaux obtenus pour les amplitudes $A=0$, $A = 4 \cdot 10^{-2}$ et $A = 5 \cdot 10^{-2}$ et qui ont respectivement pour rapport signal/bruit de -1.60 dB , -0.23 dB et 2.45 dB . Les signatures vibratoires montrent que les chocs peuvent être distingués clairement à partir du signal qui correspond à une amplitude de $5 \cdot 10^{-2}$ (Signal n°12, figure 4-4.c). La présence des chocs périodiques traduit le défaut de la bague extérieure.

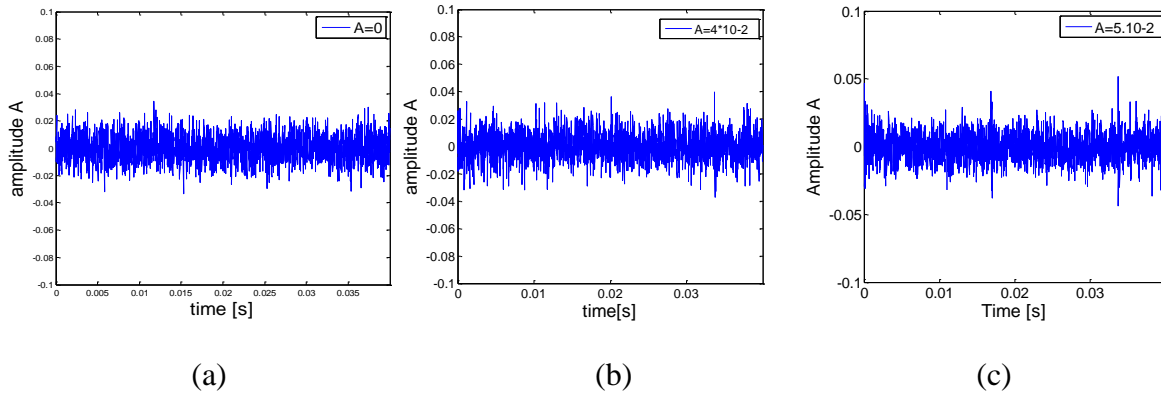


Figure 4-4: Signaux temporels pour (a) $A = 0$ (RSB = $-1.60dB$), (b) $A = 4 \cdot 10^{-2}$ (RSB = $-0.23dB$), (c) $A = 5 \cdot 10^{-2}$ (RSB = $2.45dB$)

4.2.3.3 Variation de la vitesse

L'influence de la vitesse de rotation (augmentation de la vitesse) sur la détection d'une deuxième classe en cas de présence de défaut a été aussi étudiée. Cette influence est évaluée pour une amplitude donnée $A = 0.1$ et un taux d'amortissement de 0.05 où les chocs sont visibles. La figure 4.5 illustre la variation de la valeur efficace en fonction du kurtosis en utilisant la méthode OPTICS. Cette figure montre clairement que, la méthode OPTICS ne détecte pas de deuxième classe et par conséquent la variation de la vitesse de rotation n'a pas d'influence sur la création d'une nouvelle classe mais sur le déplacement de la classe C1.

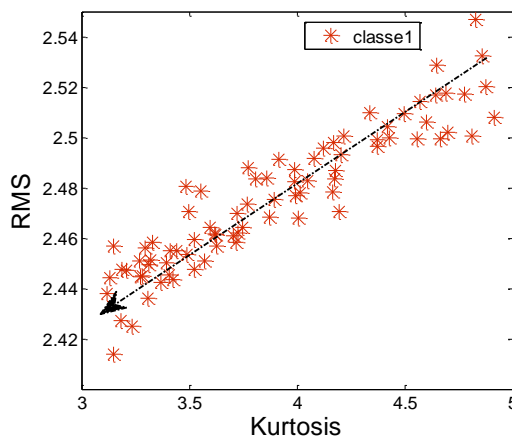


Figure 4-5: Évolution de RMS en fonction du kurtosis pour une amplitude $A=0.1$

On remarque que la valeur de *RMS* diminue avec l'augmentation de la vitesse de rotation, par contre le kurtosis prend une direction très différente de celle de *RMS*.

L'évolution de la valeur efficace en fonction du kurtosis suit une trajectoire différente de celle obtenue dans le cas de la variation d'amplitude A (figures 4-3 et 4-5).

4.2.3.4 Variation du taux d'amortissement

La figure 4-6 montre l'évolution de la valeur efficace en fonction du kurtosis pour un taux d'amortissement qui varie de 0.05 à 0.075, pour une amplitude donnée $A=0.1$ et une vitesse de rotation fr de 1000 tr/mn. De même cette évolution suit une trajectoire différente des deux précédentes (variation d'amplitude et variation de vitesse). De plus, la variation du taux d'amortissement, comme la variation de la vitesse de rotation n'a pas d'influence sur la création d'une nouvelle classe mais sur le déplacement de la classe C1.

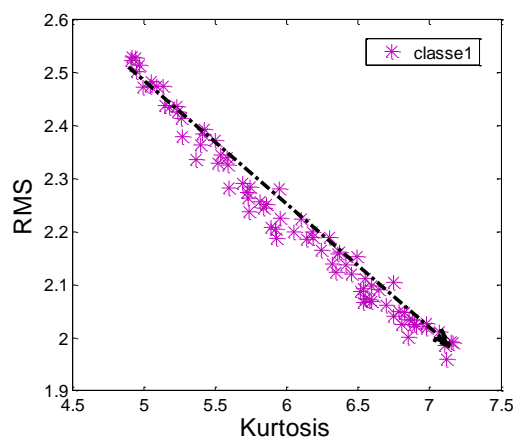


Figure 4-6: Évolution de la valeur efficace en fonction du kurtosis pour une amplitude $A=0.1$

La simulation montre que (i) la méthode proposée permet de déceler un état défectueux avec la création d'une nouvelle classe, et (ii) elle permet de suivre la dynamique d'une classe caractéristique des changements de charge et de vitesse, (iii) les trajectoires suivies par les nuages de points dues à la variation d'amplitude, de charges et de la vitesse de rotation sont différentes. Cela permet de fiabiliser la détection de défaut et le suivi de l'évolution du défaut.

4.2.4 Détection du défaut

La sélection d'indicateurs n'est pas traitée dans ce papier. Les présents travaux se concentrent sur le déploiement d'une méthode de classification dynamique. Ainsi, les indicateurs Kurtosis et RMS seront choisis pour leur performance individuelle pour une détection de défauts.

La méthode OPTICS est appliquée à la matrice d'indicateurs obtenue pour détecter la présence de défaut dans le roulement. La notion du temps t est introduite pour traiter les

données en temps réel. La méthode OPTICS s'auto exécute à chaque nouvelle prise de mesure jusqu'à la détection de défaut. La détection de défaut se traduit par l'apparition d'une nouvelle classe (C2) différente de la classe représentant le roulement jugé sain (C1). Donc deux classes C1 et C2 sont détectées.

Un signal à un instant t correspond à une simulation du modèle bruité adopté pour une valeur évolutive de la valeur de l'amplitude A , la méthode OPTICS est mise en œuvre pour la détection des deux classes (roulement sain et roulement défectueux). La figure 4-7 montre les valeurs du Kurtosis en fonction des valeurs RMS issues de la méthode OPTICS pour un roulement sain (signal n°1 et signal n°10). Le roulement sain correspond à une amplitude variant entre 0 et 2.10^{-2} (signaux n° 1 jusqu'au n° 10 avec un $RSB = -3.28dB$). La méthode OPTICS détecte le choc (une deuxième classe C2) à partir du 11^{ème} signal avec un rapport signal sur bruit de $-0.23 dB$ alors que la méthode classique (suivi des indicateurs seuls dans le temps) détecte ce même choc pour un rapport signal/bruit $RSB = 2.45 dB$ ce qui correspond au 12^{ème} signal. La figure 4-7.c représente le résultat obtenu pour la détection de défaut (signal n°11) par la méthode OPTICS.

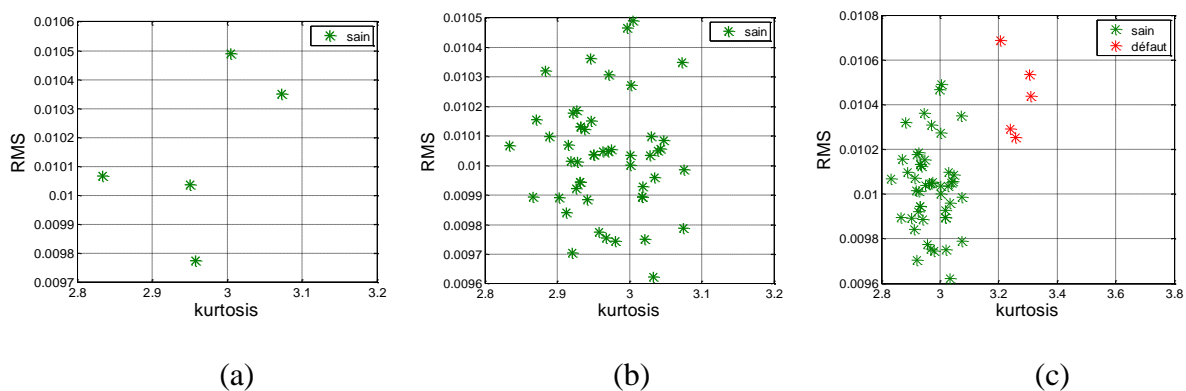


Figure 4-7: Dynamique de la détection de défaut par la méthode OPTICS. (a) roulement sain ($A = 0$). (b) roulement sain ($A = 3.10^{-2}$, $RSB = -3.28dB$). (c) roulement défectueux ($A = 4.10^{-2}$ et $RSB = -0.23 dB$).

Le tableau 4-2 montre les résultats obtenus par les deux méthodes pour la détection de défaut, en fonction du RSB . La méthode adaptée permet d'utiliser les informations complémentaires de ces deux indicateurs pour détecter le défaut de façon précoce comparé à l'utilisation d'un indicateur seul.

signal	$RSB = -3.28$	$RSB = -0.230$	$RSB = 2,45$	$RSB = 5.04$
kurtosis	2.94	3.26	3.60	5.86
RMS	0.0101	0.0104	0.0105	0.0112
Méthode classique	Non	Non	Oui	Oui
RT-OPTICS	Non	Oui	Oui	Oui

Tableau 4-2 : Comparaison entre la méthode classique et la méthode RT-OPTICS

4.2.4.1 Suivi de la classe C2

Après la détection de la classe C2, le suivi de l'évolution des signaux sera effectué à partir des paramètres caractéristiques de la matrice d'observation (indicateurs de défaut) représentant la classe C2 (contour du nuage des points, le saut du centre de gravité et densité de la classe). Le saut du centre de gravité (figure 4-8.a) comme pour le contour (figure 4-8.b) présente deux zones. Dans la première zone ces deux paramètres évoluent légèrement et leurs allures sont presque linéaires. Cela signifie que le bruit est toujours important. La deuxième zone a une allure exponentielle. Le bruit devient négligeable par rapport aux chocs. Dans ce cas, seule la valeur RMS évolue mais le kurtosis tend vers une saturation. La densité de la classe C2 a été calculée. La figure 4-8.c montre son évolution en fonction de l'amplitude A . on remarque un saut très important entre la première et la deuxième valeur, puis la densité diminue lentement qui traduit la saturation.

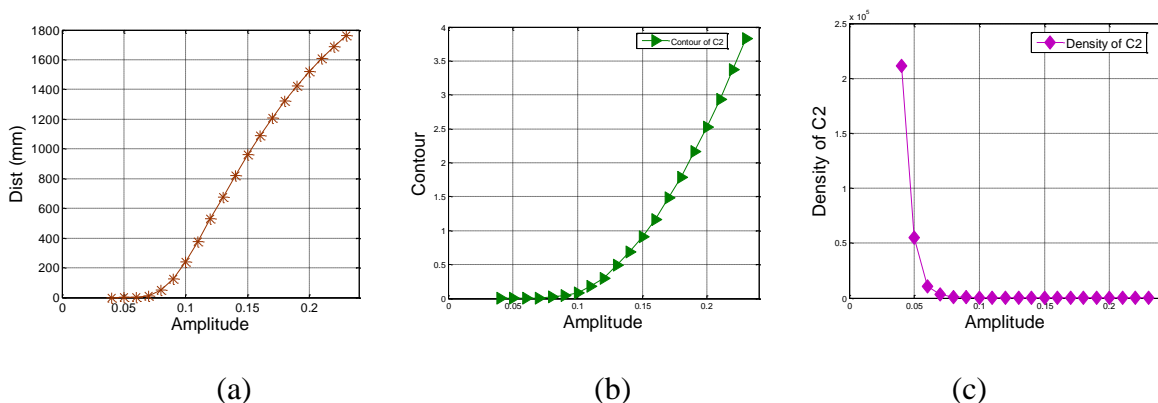


Figure 4-8: (a) Evolution de centre de la classe C1, (b) Evolution de contour de la classe C2, (c) Evolution de la densité de la classe C2

Ces trois paramètres ont donné une signification physique sur l'évolution de l'amplitude A des signaux simulés. De ce fait, ils peuvent être considérés comme indicateurs de défauts. Nous allons les utiliser par la suite dans le cas des signaux réels des défauts de roulement.

4.3 Validation expérimentale et étude comparative

4.3.1 Description du banc expérimental

L'étude est portée sur des signaux de défauts de roulements collectés sur le banc SURVIB au sein de laboratoire (GRESPI/MAN) à l'Université de Reims Champagne-Ardenne. Le banc est un carter accouplé à un moteur de puissance maximale de 10KW par l'intermédiaire d'un arbre et deux roulements ; roulement à bille sain (roulement à bille 6206) et un roulement à galets à dégrader (N.206.E.G15). La rotation est commandée par un variateur de vitesse. L'arbre est chargé radialement par un vérin via un câble acier (figure 4-9.a). Pour chaque taille de défaut et pour différentes charges (6 différentes charges de 100 à 250 daN avec un pas de 30 daN), la vitesse de rotation de moteur varie dans la plage 1405 à 1560 avec un pas de 15 tr/min. Les différentes tailles de défauts sont consignées dans le tableau 4-4. Un accéléromètre piézoélectrique est placé radialement sur le palier à roulements en question pour la mesure des signaux vibratoires.

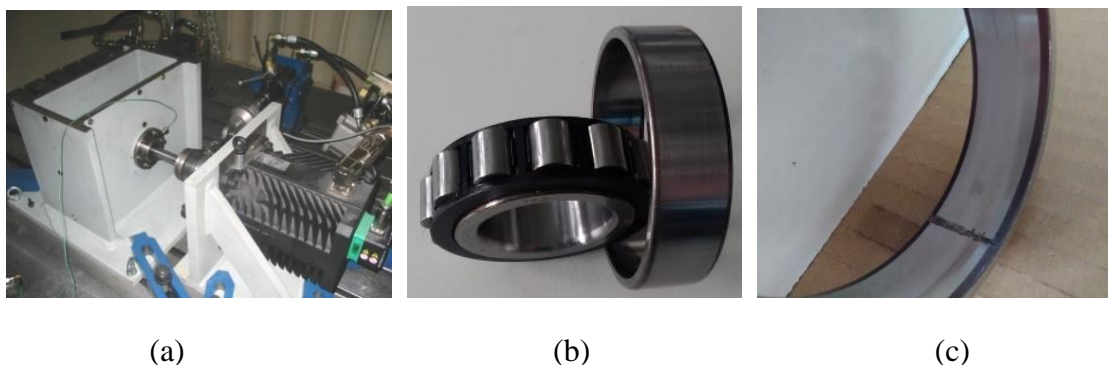


Figure 4-9: (a) Banc expérimental. (b) Roulement N.206.E.G15. (c) Exemple de défaut créé

Pour chaque essai, les signaux issus des accéléromètres sont collectés par le système d'acquisition OROS OR34. Les signaux sont collectés durant 2 secondes avec une fréquence d'échantillonnage de 51200 Hz et un nombre total de points $NT = 102400$ pts. Grâce aux caractéristiques du roulement utilisé et à la cinématique du système, les fréquences caractéristiques peuvent être calculées, tels que :

- Fréquence de défaut de la bague intérieure : $FBI = 7.76 \text{ fr}$
- Fréquence de défaut de la bague extérieure : $FBE = 5.24 \text{ fr}$

Une base de données de signaux à l'état neuf et à différents niveaux d'endommagement a été générée.

4.3.2 Création des défauts sur la bague extérieure

Dans kerroumi et al. (2013), les défauts ont été créés par le procédé d'électro-érosion. Ces défauts sont des trous débouchant et de ce fait ils ne traduisent pas le cas réel de dégradation des organes des roulements ; le défaut de roulement est en général similaire à un défaut d'écaillage (défaut superficiel). Notre étude consiste en la détection d'un défaut très naissant et pour cela nous avons utilisé un autre moyen pour créer des défauts superficiels, le stylo à gravure. Les défauts ont été créés sur la bague extérieure du roulement à galet. L'avantage de cet outil est de créer facilement des défaut de formes aléatoires sur la bague extérieure de roulement. Ces défauts sont pratiquement similaires aux défauts réels d'écaillage, comme le montre la figure 4.9-c. Par contre la taille de défaut créée est très difficile à mesurer avec les outils classiques, surtout quand il s'agit d'un défaut naissant (de l'ordre de dizaine de nanomètres). Pour remédier à ce problème, nous avons proposé une méthode basée sur la mesure d'empreinte de défaut en question en utilisant une patte très lisse "plastiform". Pour la prise des empreintes nous avons réalisé un support pour maintenir la bague défectueuse en position d'équilibre ensuite à l'aide d'un pistolet, on a injecté la pâte visceuse sur le défaut. La pâte se solidifié très rapidement et elle est facile à être décoller. Finalement, on a utilisé un profilomètre '*Taylor-Hobson subtronic 3P*', comme le montre la figure 4-10.b, pour mesurer la taille du profil obtenu comme étant une rugosité.



(a)



(b)

Figure 4-10: (a) défaut créé, (b) mesure de la taille de défaut

La taille de défaut est caractérisée par trois grandeurs à l'aide de cet outil qui sont la rugosité arithmétique Ra , la rugosité totale Rt et la largeur du défaut l . Leurs valeurs sont consignés dans le tableau 4-3. Durant tous les essais, nous avons utilisé un seul roulement. Nous avons augmenté la taille de défaut après chaque essai. On remarque que la taille du premier défaut, $Ra = 0.33 \mu m$, est trop petite à être détectée soit par les méthodes classiques ou bien par les méthodes de classification. D'ailleurs, on a trouvé une difficulté pendant la mesure de sa taille.

Défauts	Sn	Df1	Df2	Df3	Df4	Df5	Df6	Df7	Df8
$Rt (\mu m)$	-	2.42	3.00	8.25	10.50	10.55	18.17	18.67	21.42
$Ra (\mu m)$	-	0.33	0.41	0.73	1.32	1.52	1.82	2.36	2.97
$Largeur (mm)$	-	0.16	0.21	0.45	0.74	1.09	1.78	2.03	2.32

Tableau 4-3: Dimensions des défauts créés sur la bague extérieure de roulement

Chaque passage du défaut sur une bille se traduira par une impulsion. Ainsi, la période de ces chocs sera celle de la vitesse de rotation. Or l'arbre tourne à la vitesse, par exemple, de 1430 *rpm*, soit une fréquence 23.83 Hz. Il en résulte donc un train de chocs de période 0.0419 s. Les harmoniques de la fréquence 23.83 Hz devront apparaître. Le roulement comporte 13 rouleaux, est montée sur l'arbre tournant à la fréquence de 23.83 Hz. Ce dernier sera alors soumis à 13 chocs par tour.

4.3.3 Calcul des indicateurs

Pour vérifier la sensibilité de l'approche temporelle et fréquentielle dans la détection des défauts de roulements et pour étudier l'apport de ces caractéristiques et ainsi choisir ceux qui nous permettront de bien discerner les signaux issus d'un roulement sain de ceux d'un roulement défectueux même quand la machine fonctionne sous conditions non stationnaires, plusieurs indicateurs scalaires ont été calculés pour différentes tailles de défauts et dans des caractéristiques variables de système, parlant de la vitesse de rotation de moteur et la charge sous laquelle est soumis l'axe de rotation. Les signaux ont été collectés durant une période de 2 seconds et avec une fréquence d'échantillonnage de 51200 Hz. Les signaux sont découpés en dix tranches pour avoir une moyenne sur toute la longueur du signal en question, et ainsi avoir un nombre suffisant de points qui remplissent la condition d'un cluster.

Dans la section suivante on va étudier chaque type d'indicateurs utilisés et le tester pour voir son comportement pour des signaux réels. Il faut noter que les indicateurs extraits sont normalisés afin d'éviter l'effet d'échelle et de bien cerner les indicateurs ayant de faible valeurs. Les indicateurs scalaires calculés sont : la valeur efficace, le Kurtosis, la valeur crête à crête, le skewiness, le facteur F . les indicateurs fréquentiels cités dans le chapitres 1.

4.3.3.1 Indicateurs scalaires

Dans le domaine temporel (Figure 4-11), les signaux ont été traités pour extraire deux indicateurs permettant une première appréciation de l'état vibratoire d'une installation: la valeur efficace et le kurtosis. Ces indicateurs sont d'ailleurs habituellement prescrits par les normes internationales. Malheureusement, pour un diagnostic approfondi identifiant de façon précise les organes défailants et surtout les défauts naissants, ils sont limités.

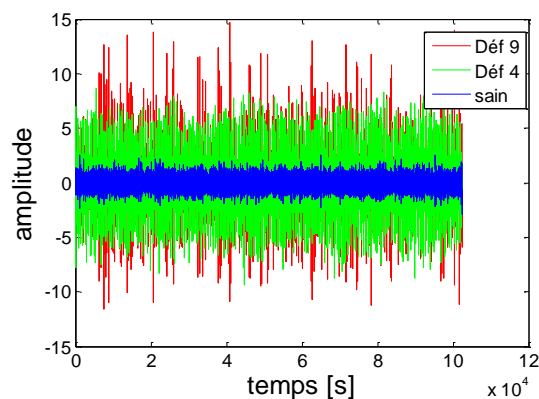


Figure 4-11: Signal temporel, défauts (sain, défaut 4 et défaut9)

Les signaux collectés sont découpés en dix tranches avant d'être traités. Ainsi pour chaque état de roulement, la charge et la vitesse de rotation varient (6 charges \times 10 vitesses pour chaque taille de défaut), ce qui représente une base de données de 600 points pour chaque état.

La figure 4-12 illustre l'évolution de la valeur efficace en fonction du kurtosis pour toutes les tailles de défauts. On distingue deux nuages de points sur la figure, un nuage est obtenu pour des faibles valeurs de RMS et de Kurtosis ce qui peut représenter le roulement en bon état et un autre nuage correspondrait au roulement avec défaut. Le kurtosis et a valeur RMS augmentent graduellement avec l'augmentation de la taille de défaut, Les valeurs de kurtosis sont comprises entre 3.30 et 4.20 dans le cas d'un roulement sain et pour toutes les variations des paramètres (vitesse et charge). Ces valeurs dépassent le seuil théorique ($Kurtosis = 3$) pour un roulement sain. Pour la première taille de défaut (0.16mm), les valeurs

du kurtosis varie entre 3.20 et 3.90. On remarque des valeurs communes entre les deux états de roulement (sain et défectueux). Le saut entre ces deux nuages (taille de défaut de 0.16 mm) sera constaté par l'application de la méthode proposée (RT-OPTICS).

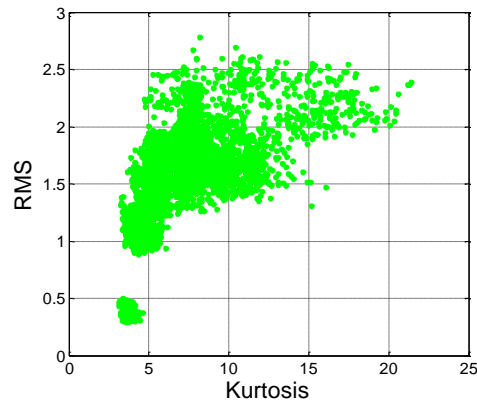


Figure 4-12: Evolution de la valeur RMS en fonction du kurtosis pour toutes les tailles de défauts et dans le cas de tous les paramètres (9 défauts \times 10 vitesses \times 6 charges \times 10 tranches = 5400points)

Les valeurs des indicateurs scalaires, calculés pour une vitesse et une charge constantes pour toutes les tailles de défauts sont illustrées sur la figure 4-13. Presque tous les indicateurs augmentent avec l'augmentation de la taille de défaut, sauf le skewness.

Le kurtosis augmente graduellement avec l'augmentation de la taille de défaut, par contre on remarque une diminution de la valeur RMS à partir de défaut 6 et cela est dû peut être aux défauts créés.

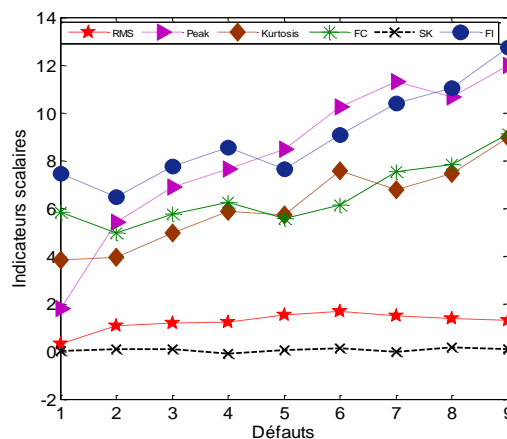


Figure 4-13: Evolution des indicateurs en fonction de la taille de défauts, vitesse et charge constantes

4.3.3.1.1 Effet de la vitesse de rotation sur les indicateurs scalaires

L'influence de la variation de la vitesse de rotation du moteur sur l'évolution des indicateurs et ainsi sur la détection de défaut est réalisée en maintenant la charge constante. La figure 4-14 montre l'évolution des indicateurs scalaires en fonction des défauts pour toutes les vitesses. On remarque que tous les indicateurs ne suivent pas la même allure et n'évoluent pas d'une façon progressive et cela est dû peut être à l'effet d'hystérésis ou bien au pas de variation de la vitesse qui est de 15 tr/min et cela peut influencer la densité des classes et par conséquent le résultat de la classification. En pratique, on verra que la fréquence de défaut va être différente à celle qui correspond à la vitesse théorique.

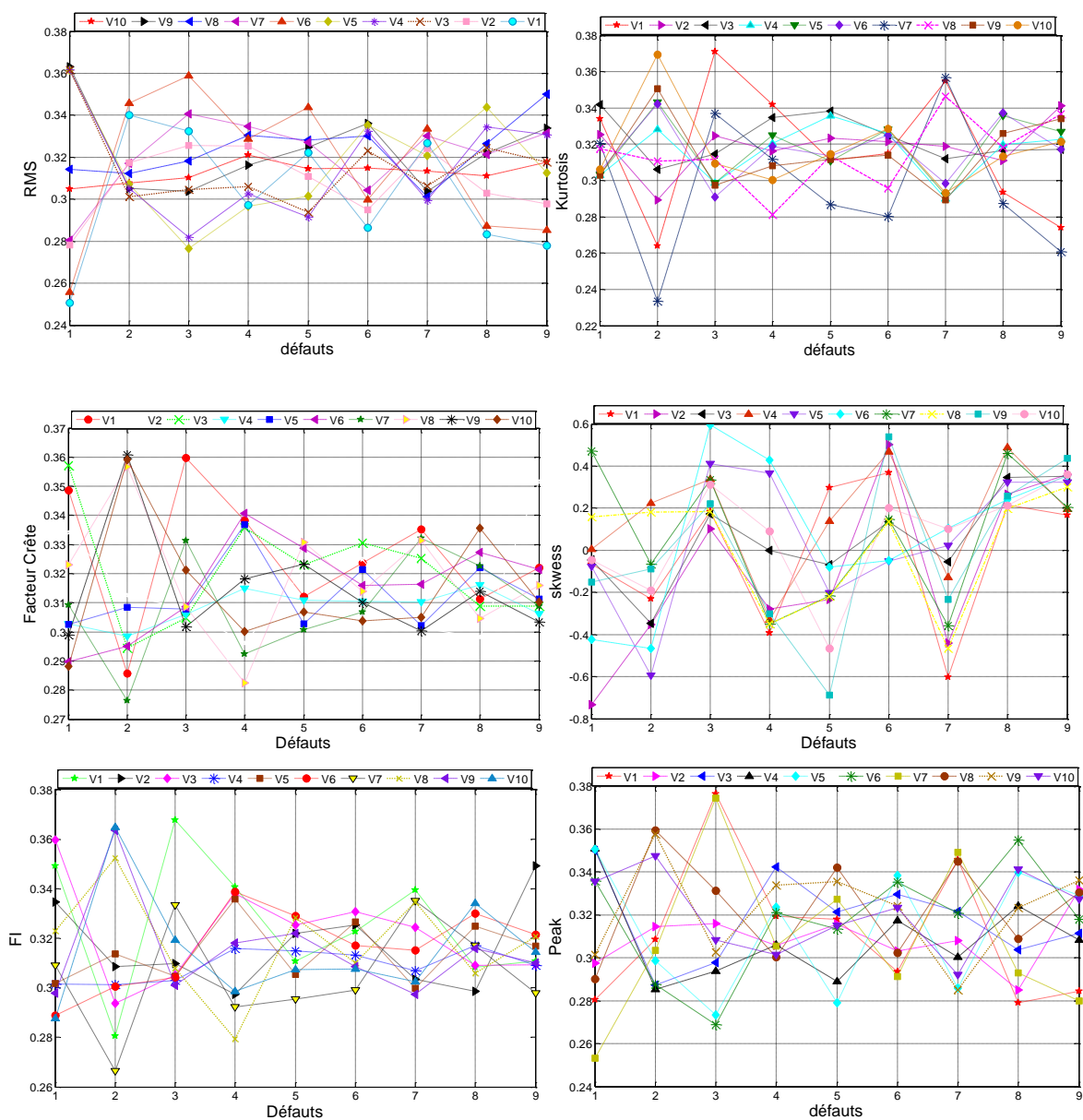


Figure 4-14: influence de la vitesse de rotation sur l'évolution des indicateurs scalaires

La figure 4-15 montre l'évolution de la valeur efficace en fonction du kurtosis pour une seule taille de défaut de 0.21 mm, pour une charge constante de 100 daN et pour une vitesse variable. Cette évolution prend une direction très différente de celle qui correspond à l'évolution du défaut (figure 4-6 de simulation). On conclut que la vitesse de rotation influence seulement sur le déplacement de la classe.

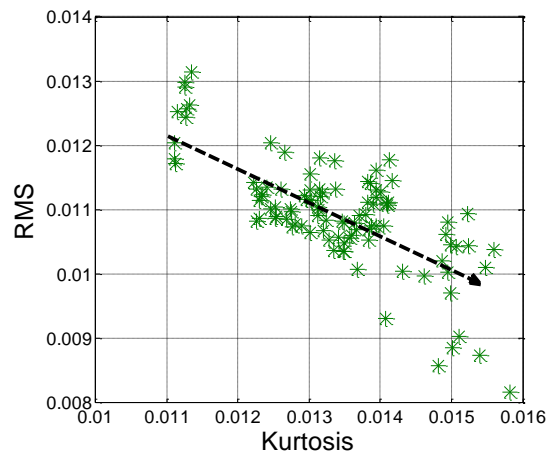


Figure 4-15: Évolution de la valeur efficace en fonction du kurtosis pour une taille de défaut de 0.21 mm, charge= 100 daN et une vitesse variable

4.3.3.1.2 Effet de la charge sur les indicateurs scalaires

L'influence de la charge à laquelle est soumis le roulement sur l'évolution des indicateurs scalaires est aussi étudiée (augmentation de la charge). Cette influence est évaluée pour chaque taille de défaut pour une vitesse constante et une charge variable (6 charges). Les résultats sont résumés sur la figure 4-16.

On peut remarquer que pour presque la totalité des indicateurs, l'allure de leur évolution reste presque la même pour toutes les charges. Par contre on peut remarquer une très légère augmentation de leurs valeurs à cause du pas d'augmentation de la charge qui est pratiquement insuffisant pour avoir une idée claire sur l'influence de cette caractéristique cinématique sur l'évolution de ces indicateurs. D'un autre côté, la valeur de la charge ne dépasse pas la limite supportée par le roulement qui est fixée par le constructeur.

L'influence de la charge sur l'évolution de la valeur RMS en fonction du kurtosis pour la même taille de défaut et pour une vitesse constante de 1405 tr/mn est présentée sur la figure 4-17. Il est très clair que son allure est différente de celle de la vitesse et celle de

l'influence du défaut. On constate que ces paramètres, vitesse et charge, influent sur le déplacement de la classe et non pas sur le niveau de dégradation de roulement.

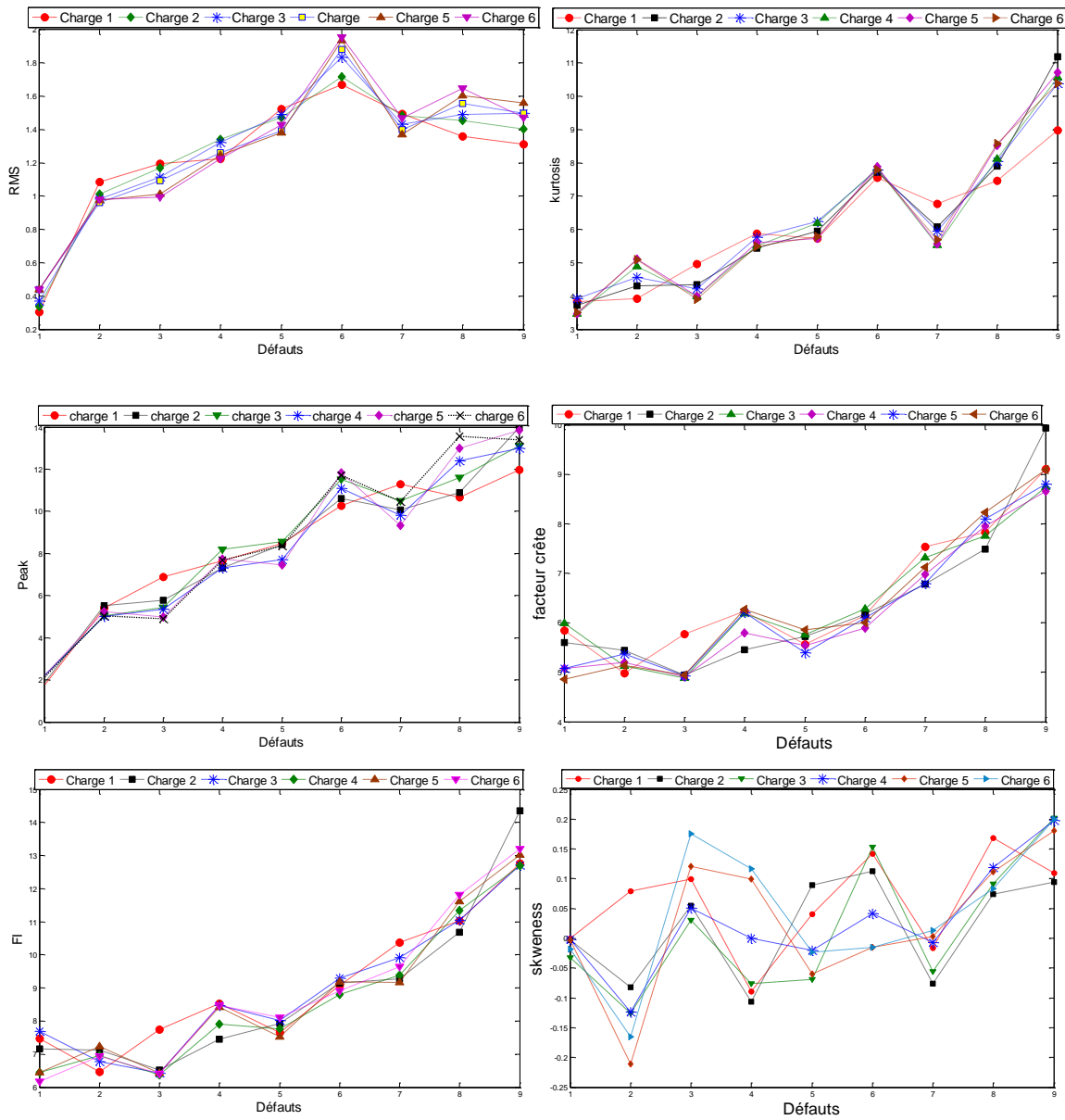


Figure 4-16: Influence de la charge sur l'évolution des indicateurs scalaires (chaque point est une moyenne de dix valeurs)

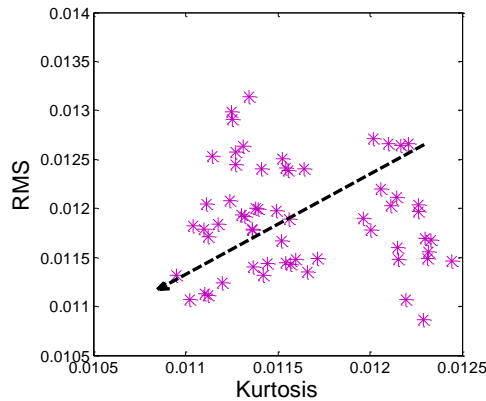


Figure 4-17: Évolution de la valeur efficace en fonction du kurtosis pour taille de défaut de 0.21 mm, vitesse= 1405 tr/mn et une charge variable

4.3.3.2 Calcul des indicateurs fréquentiels

Pour vérifier la sensibilité de l'approche fréquentielle dans la détection des défauts de roulement, plusieurs indicateurs ont été calculés à partir des spectres de puissance des signaux vibratoires pour toutes les tailles de défauts de même vitesse et charge. Dans la figure 4-18, on peut voir clairement l'évolution des indicateurs fréquentiels en cas de la présence du défaut, ce qui nous donne une idée générale sur leur capacité de distinguer entre un roulement sain et un roulement défectueux.

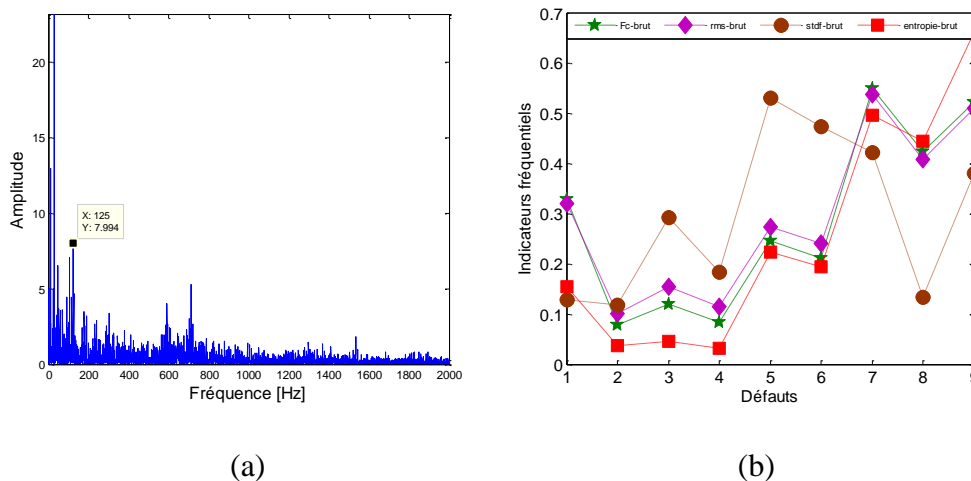


Figure 4-18:(a) spectre de puissance du signal vibratoire pour une vitesse de 1430 rmp et (b) Les indicateurs fréquentiels calculés à partir de la puissance de l'enveloppe du spectre

Bien que les indicateurs fréquentiels montrés dans la figure 4-18 montrent un écart remarquable entre un roulement sain et un roulement défectueux mais de valeur négative. Cet écart peut mener à un mauvais diagnostic et donc peut confondre l'état sain à l'état défaut avancé.

Plus loin, les indicateurs issus du spectre d'enveloppe des signaux pour une même vitesse de rotations du moteur et pour la même charge ont été calculés. Cette section présente les résultats d'analyse d'enveloppe du signal du roulement étudié dans les deux cas : roulement sain et roulement défectueux. Durant les tests, la vitesse de rotation est maintenue faible (entre 1405 et 1560 rpm) en raison du bruit excessif produit et du risque imminent de dommages du banc d'essai (l'axe de rotation a été changé grâce à son endommagement).

La fréquence du défaut de la bague extérieure pour une vitesse de rotation nominale par exemple de 1420 tr/min est de: $5.24fr$ qui est égale à 124 Hz. La fréquence caractéristique bague extérieure dépend de la vitesse réelle durant l'essai (qui peut être différente de la vitesse nominale) et également de la BPFO réelle (qui peut être aussi un peu différente de celle donnée par le constructeur du roulement). On verra que la fréquence réelle du défaut vaut 126.5.

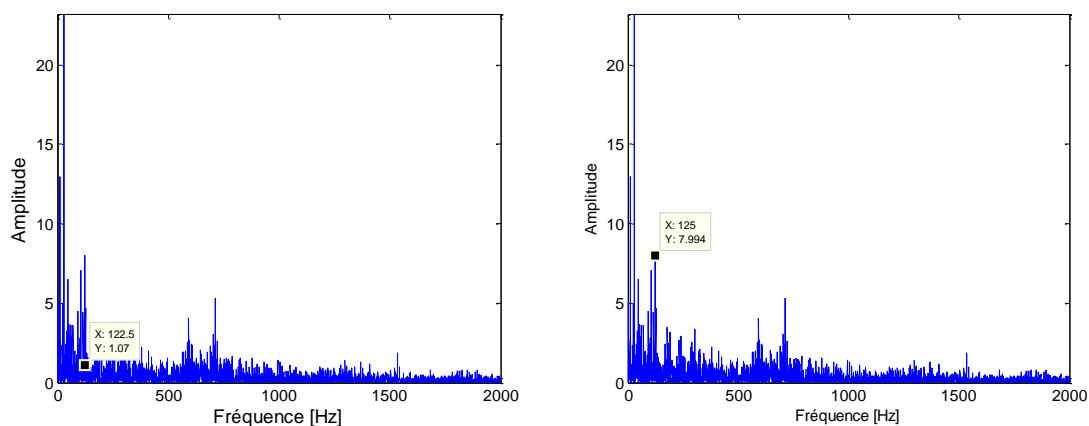


Figure 4-19: (a) Spectre de roulement sans défaut pour une vitesse de 1405 rpm, (b) Spectre de roulement avec défaut ($Ra= 0.33 \mu m$) pour une vitesse de 1435 rpm

La figure 4-19 illustre respectivement les spectres de l'enveloppe dans les cas des vitesses de 1405 et 1535 rpm. Le spectre du signal montre des pics proches autour des deux vitesses 1405 et 1435 Hz. Cependant il est difficile de confirmer le type de défaut associé à cette région.

Comme première conclusion sur le signal du spectre, nous pouvons confirmer que le roulement est endommagé. Toutefois, il est difficile d'identifier le défaut en se basant uniquement sur l'analyse d'enveloppe.

Les indicateurs extraits à partir de l'application de l'analyse de l'enveloppe sur les signaux sont présentés dans la figure 4-20. L'application de cette méthode n'a pas amélioré la détection précoce de défaut et même dans le cas de deuxième défaut les valeurs de ces indicateurs restent plus basses par rapport à l'état sain et cela est dû peut être à la taille de défaut qui est de l'ordre de centième de nanomètre. Notant que la taille du premier défaut est mesurée difficilement, l'empreinte prise est très peu remarquable et cela est dû à la pate utilisé qui n'a pas bien émergé dans le défaut créé. Ces indicateurs ne donnent aucune information sur la détection de défaut et cela peut nous amener à un diagnostic très tardif.

On peut remarquer que les valeurs de ces indicateurs dans les premiers stades de défaillance de traduisent pas l'état de roulement sauf la valeur std_f et le SPRO.

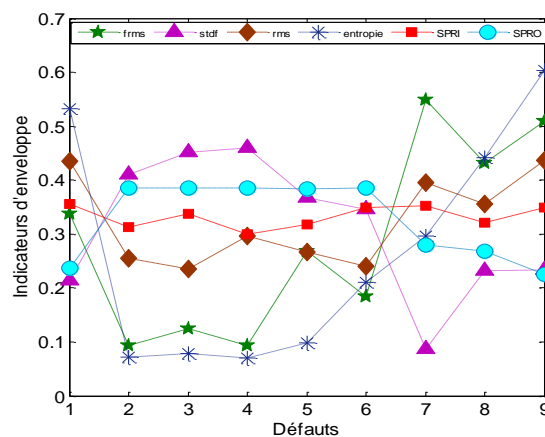


Figure 4-20: Les indicateurs calculés à partir du spectre d'enveloppe

4.3.3.2.1 Effet de la vitesse de rotation sur les indicateurs fréquentiels

L'effet de la variation de la vitesse de rotation sur les indicateurs extraits du spectre de puissance brut est aussi étudié dans cette partie. La figure 4-21 montre que presque tous les indicateurs suivent une même allure dans le cas de la variation de la vitesse de rotation quelque soit la taille de défaut. Donc cela ne peut pas influencer la détection de défaut.

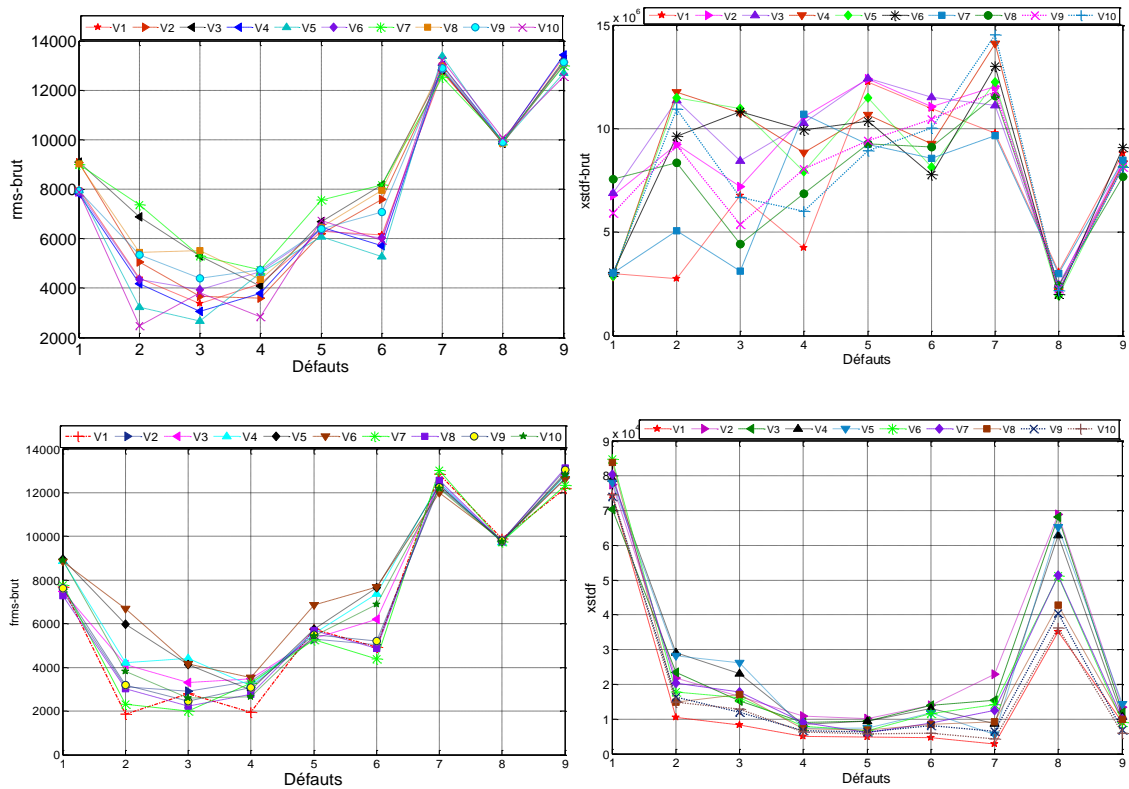


Figure 4-21: Evolution des indicateurs fréquentiels en cas de présence de défaut, et en cas de différentes vitesses

4.3.3.2.2 Effet de la charge sur les indicateurs fréquentiels

L'effet de la variation de la charge sur la détection de défaut et ainsi sur l'évolution des indicateurs a montré les résultats présentés sur la figure 4-22. On peut voir que tous les indicateurs augmentent et ne changent pas d'allure avec l'augmentation de la charge appliquée.

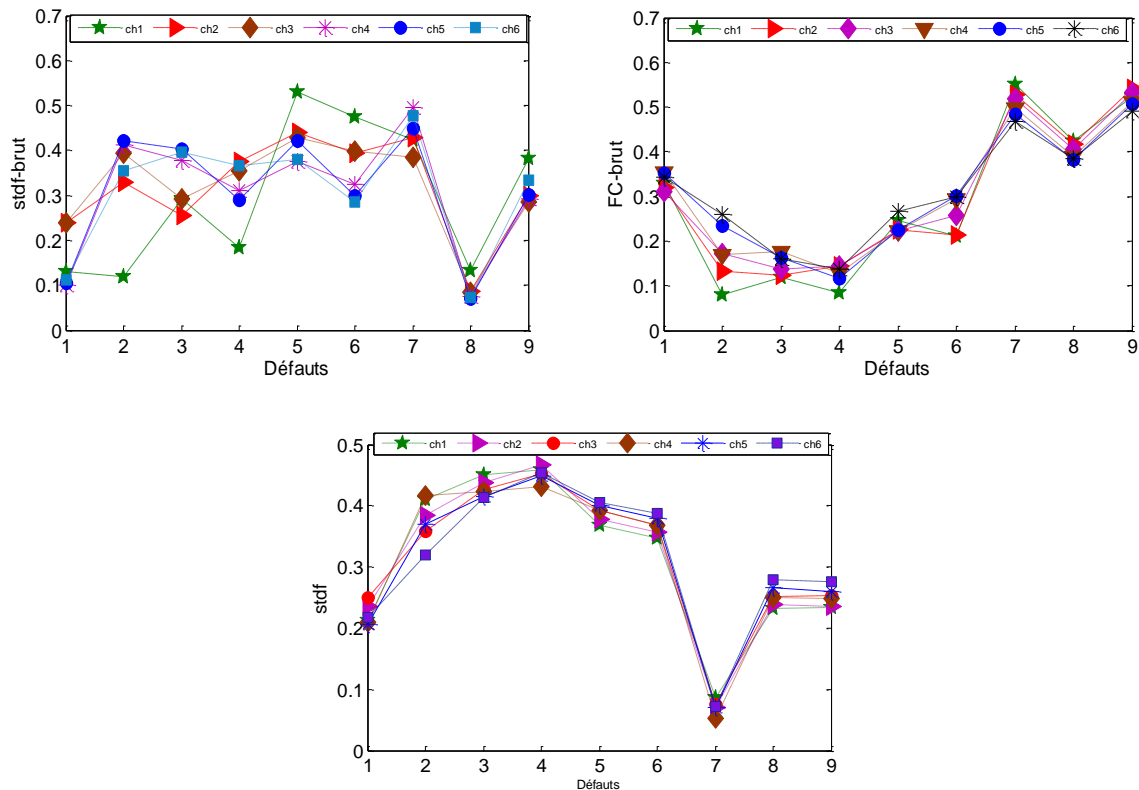


Figure 4-22: Evolution des indicateurs fréquentiels en cas de présence de défaut et en cas de différentes charges

4.3.3.3 Analyse par les ondelettes

La plupart de des méthodes de traitement de signal donnent des résultats acceptables quand il s'agit de déceler l'existence d'anomalies courantes dans l'installation. Pour la localisation de défaut tant dans l'espace que dans le temps nous avons l'une des techniques d'analyse du signal dans le domaine temps-fréquence, la décomposition en ondelettes.

En premier temps, nous avons calculé les coefficients réels et positifs de la transformée en ondelettes continue du signal en question ensuite nous avons calculé le spectre d'enveloppe de ces coefficients. Deux indicateurs peuvent être extraits de ce spectre d'enveloppe : le W_{rms} ou la valeur efficace des fréquences et le P_{CWT} qui représente la valeur moyenne des amplitudes du spectre d'enveloppe. La procédure de l'extraction de ces deux indicateurs est présentée dans la figure 4-23.

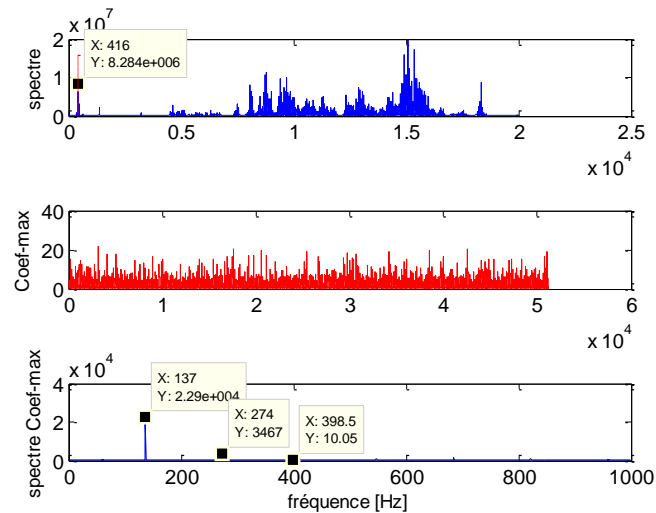


Figure 4-23: Procédure de l'extraction de ces deux indicateurs W_{rms} et P_{CWT}

L'évolution de ces indicateurs est présentée dans la figure (4-24), les deux indicateurs ont bien montré un écart important entre les deux états de santé de roulement sain et défectueux. L'indicateur W_{rms} diminue progressivement et cela n'explique pas la présence du défaut mais il peut aider à la détection de défaut par les méthodes de classification. L'indicateur P_{CWT} montre une meilleure séparation entre état sain et état défectueux, et donc une meilleure capacité à la détection.

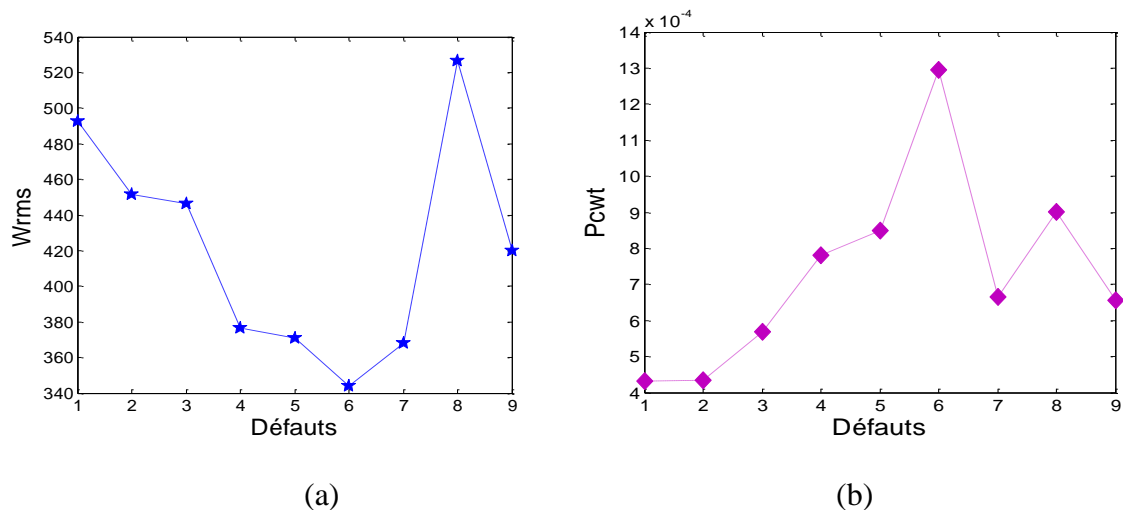


Figure 4-24: Extractions des indicateurs : (a) W_{rms} et (b) P_{cwt} A partir de la transformée en ondelettes continues

4.3.3.3.1 Effet de la vitesse

Lorsque la vitesse de rotation augmente la température de fonctionnement s'élève à cause des frottements. Si cette température dépasse une certaine limite, le lubrifiant n'assure plus son pouvoir lubrifiant, ce qui peut conduire au grippage du roulement. Pour cette raison, nous avons étudié l'influence de la vitesse de rotation sur l'évolution des indicateurs. La vitesse de rotation maximale à laquelle peut tourner un roulement en continu dépend principalement du type de roulement, de sa taille, du type de cage, de la lubrification, du refroidissement éventuel. La figure 4-25 montre l'influence de la vitesse de rotation du moteur sur les indicateurs extraits de l'analyse par ondelette.

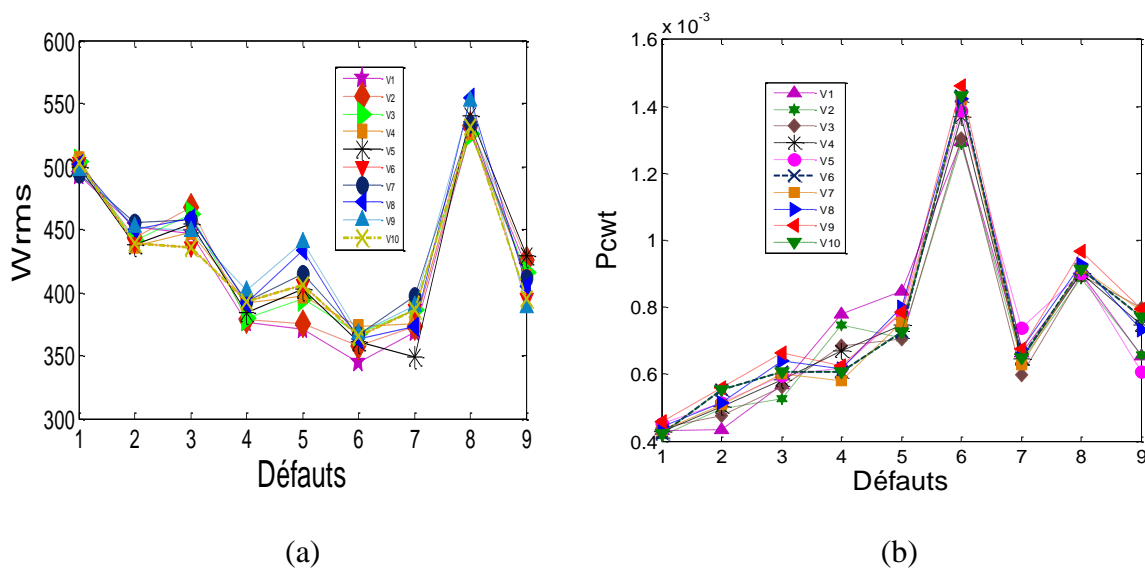


Figure 4-25: Influence de la vitesse sur les indicateurs : (a) W_{rms} et (b) P_{cwt}

4.3.3.3.2 Effet de la charge

L'effet de la charge sur la sensibilité de la détection temporelle et fréquentielle est étudiée dans cette phase. Le roulement, étant non chargé au début, est soumis par la suite à six charges avec une différence régulière de 100 daN à 250 daN avec un pas de 30 daN. Celles-ci sont assurées par un frein monté sur l'arbre mené. On constate que l'augmentation de la charge de 100 à 250 daN pour l'arbre menant provoque une légère augmentation de l'indicateurs W_{rms} (figure 4-26). Donc la détection du défaut est beaucoup plus évidente dans le cas où le roulement est chargé. De même que l'indicateur P_{cwt} . En conclusion une augmentation du chargement de roulement ne permet pas seulement la détection mais influe aussi sur les amplitudes du signal d'où la variation importante de son énergie. On verra dans la suite que

malgré cette augmentation, la méthode de classification adaptée, RT-OPTICS, ne détectera un état défectueux de roulement sauf quand le défaut est présent.

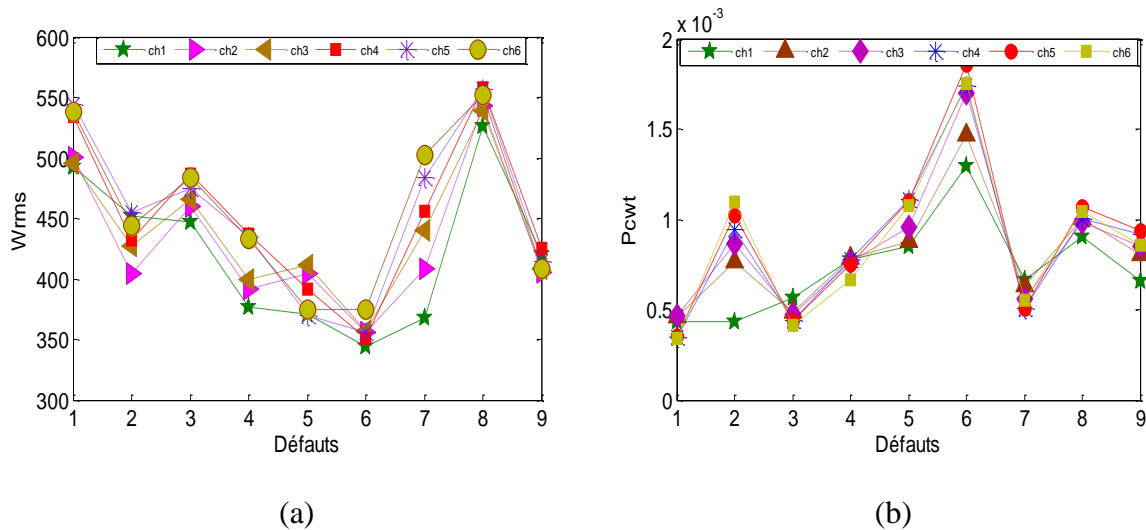


Figure 4-26: Influence de la charge sur les indicateurs : (a) W_{rms} et (b) W_{Pcwt}

4.3.4 Validation de la méthode proposée

La validation de la méthode proposée a été réalisée dans quatre cas : cas de la vitesse et de la charge constantes, cas de la vitesse constante et la charge variable, cas de la vitesse variable et la charge constante et enfin dans le cas de la charge et de la vitesse variables.

4.3.4.1 Cas d'une vitesse et d'une charge constante

En premier lieu, la méthode a été testée sur des signaux avec une charge et une vitesse constante pour toutes les tailles de défauts. La méthode a été utilisée dans deux volets : le premier consiste en la détection juste de défaut et le deuxième pour visualiser les différentes régions de densité dans la deuxième classe, classe caractérisant le défaut, par le tracer de l'ordre des distance-accessibilité.

Seulement deux indicateurs ont été pris en compte dans l'étape de diagnostic par classification. La figure 4-27 montre l'évolution de la valeur efficace en fonction du kurtosis dans le cas des neuf défauts. On Remarque une augmentation très marquée de la valeur RMS, un saut de 0.26 à 1.4 et cela est dû peut être au bruit blanc. Par contre la valeur du kurtosis varie très légèrement de 3.8 à 3.9. La valeur de la RMS et du kurtosis dans le cas du défaut 7 diminuent et cela est dû peut être au défaut crée, plus profond dans les cotés de la bague extérieure qu'au centre, et donc le choc résultant du contact des éléments roulants avec

le défaut est minime. Ou bien cette diminution est due au mauvais emplacement de roulement ; défaut placé du côté de la charge.

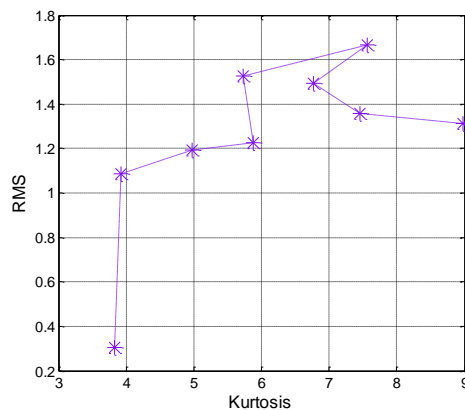


Figure 4-27: Evolution de la valeur efficace RMS en fonction du kurtosis des 9 défauts

4.3.4.1.1 Détection du défaut avec la méthode adaptée

Les figures 4-28 représentent la dynamique de la détection de défaut par l'application de la méthode RT-OPTICS, en représentant la valeur efficace en fonction de kurtosis, on a tracé aussi l'ordre de la distance-accessibilité en fonction de l'ordre de points en utilisant la méthode OPTICS, et cela dans le but de visualiser la distribution des régions de différentes densité et de comparer cette distribution avec nos résultats.

D'après la figure 4-28 (a et b), on peut remarquer clairement que notre méthode RT-OPTICS détecte la deuxième classe qui correspond au roulement défectueux dès le deuxième signal qui correspond au premier défaut créé. Donc, pour les signaux stationnaires la méthode adaptée paraît très efficace.

La figure 4-28 (c) montre l'ordre de la distance-accessibilité de toute la base de données. On Remarque un pic très marquée à la 11^{ème} points. Cela signifie par définition la présence de deux classes. La première classe contient les points d'ordres situés à gauche de ce pic et le reste des points représente la deuxième classe. Dans la deuxième classe on remarque différentes zones qui traduisent les différentes taille de défaut.

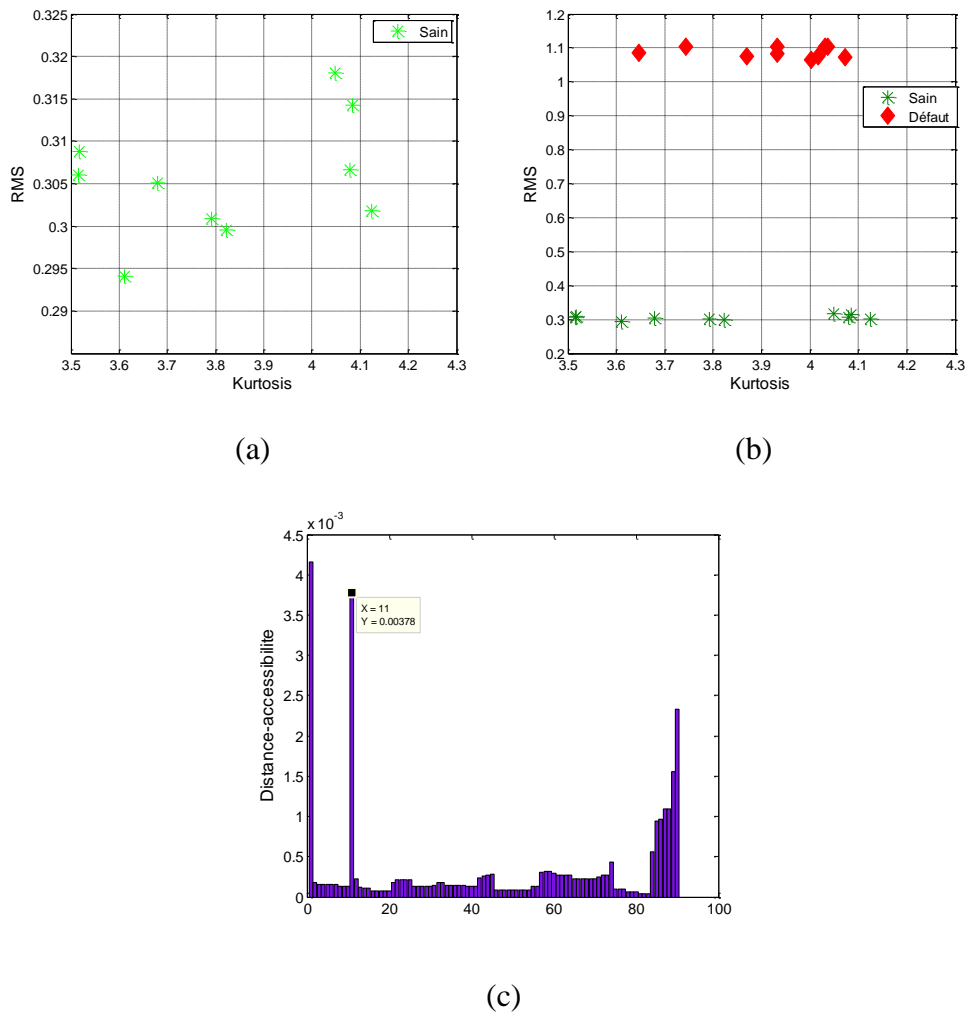


Figure 4-28: (a), (b) dynamique de la détection de défaut et (c) ordre de la distance-accessibilité

On peut conclure que la méthode OPTICS est très efficace pour la visualisation des différentes zones de densité sans être influencé par les paramètres d'entrée (rayon de voisinage et le nombre minimum de points).

4.3.4.1.2 Suivi de la classe de roulement défectueux

La figure 4-29 montre l'évolution des différentes caractéristiques géométriques. On remarque un saut très clair entre le cas d'un roulement sain et d'un roulement défectueux. Ces caractéristiques se développent avec l'augmentation de la taille de défaut et cela peut être utilisé comme indicateurs pour le suivi de l'état du degré de sévérité du défaut.

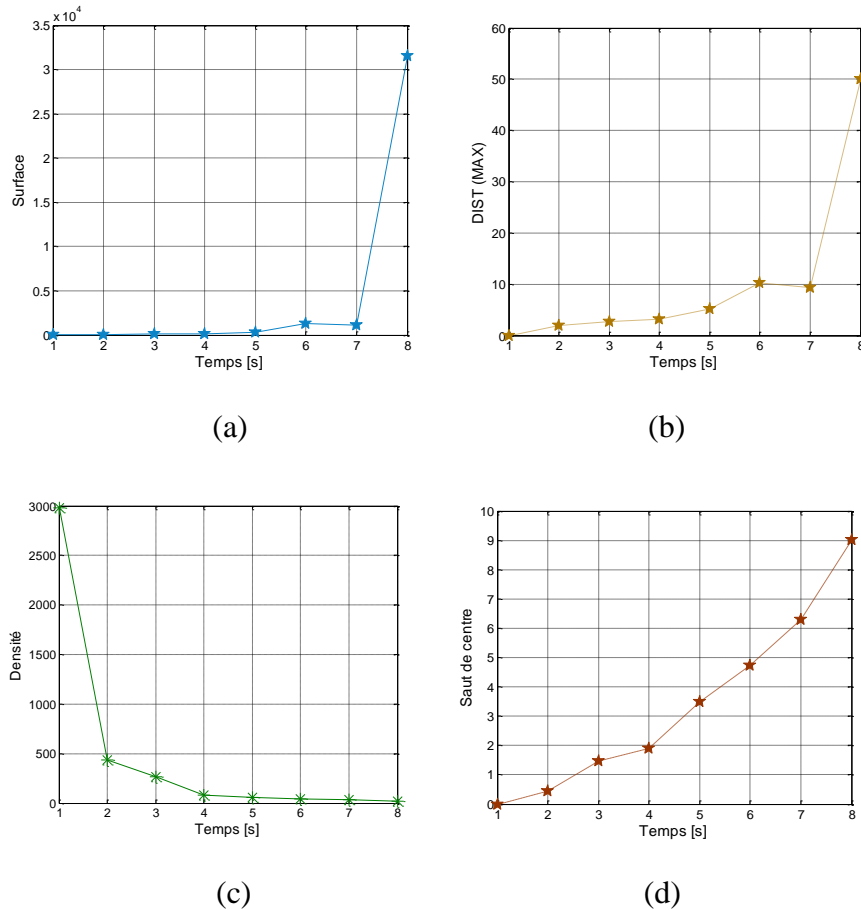


Figure 4-29: Evolution des caractéristiques géométriques: (a) surface, (b) Distance maximale, (c) densité, (d) saut de centre.

4.3.4.2 Cas d'une vitesse variable

J'effet de la vitesse de rotation sur la détection de défaut est étudié. Nous avons tracé l'évolution de la valeur efficace en fonction du kurtosis dans le cas d'un défaut évolutif et une vitesse variable pour chaque taille de défaut (figure 4-30). On remarque qu'il est facile de détecter la classe de roulement sain, mais on ne peut plus distinguer entre les différents niveaux de dégradation de roulement.

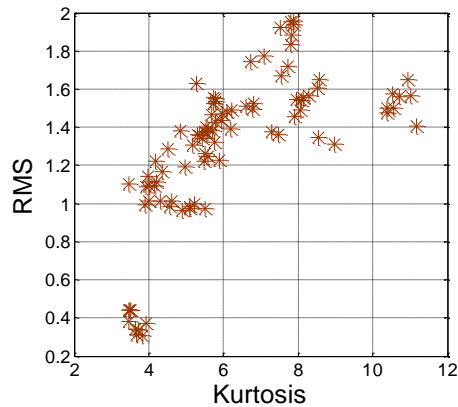


Figure 4-30: Evolution de la valeur efficace RMS en fonction du kurtosis pour différentes vitesses et une charge constante

4.3.4.2.1 Détection du défaut avec la méthode adaptée

La dynamique de la détection de défaut est montrée sur la figure 4-31.a.b et c le défaut est détecté dès le premier signal qui correspond au roulement défectueux, 11^{ème} signal.

La méthode OPTICS n'est pas influencée par la variation des paramètres cinématiques de la machine et elle détecte les moindres augmentations des indicateurs, en particulier le kurtosis parce qu'il est considéré comme l'indicateur le plus sensible aux chocs dans les premiers stades de dégradation de roulement.

La figure 4-31.d montre l'ordre des distances d'accessibilité. On peut remarquer que deux zones de densité distinctes. La première zone de (0- 101) qui correspond au roulement sain et la deuxième zone de (101 à la fin de la base de données) qui correspond au roulement défectueux. De cette distribution, nous pouvons dire que la vitesse influe le suivi de niveau de dégradation de roulement.

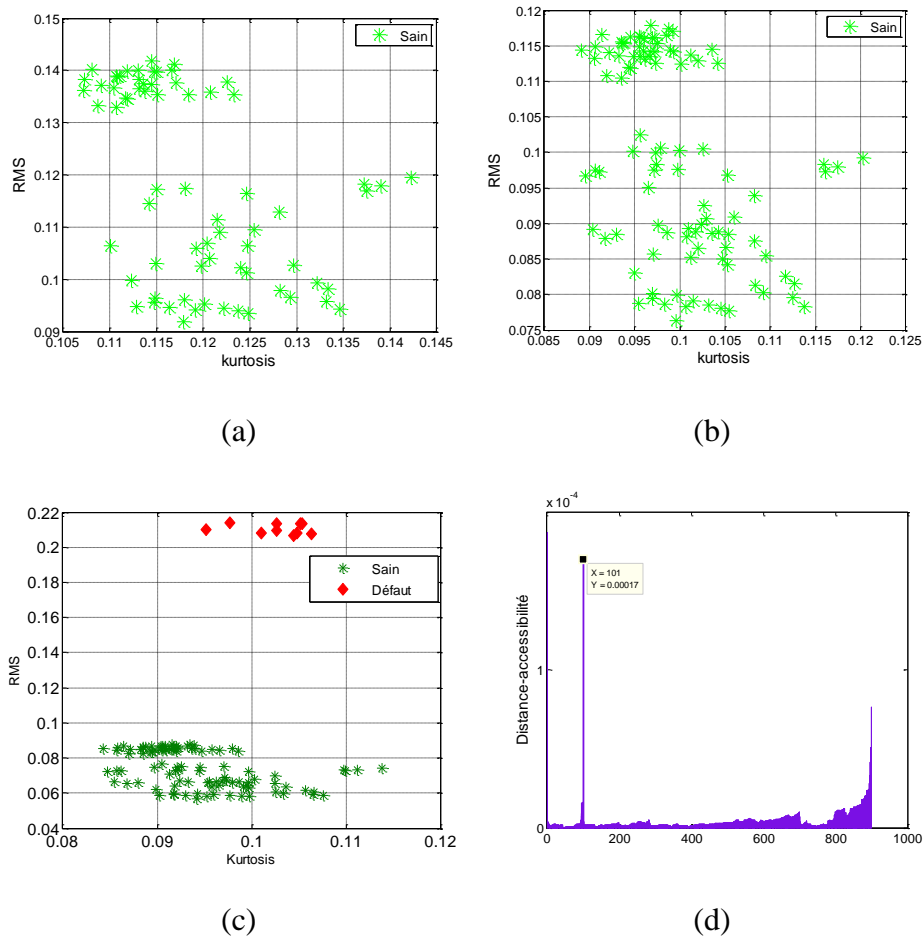
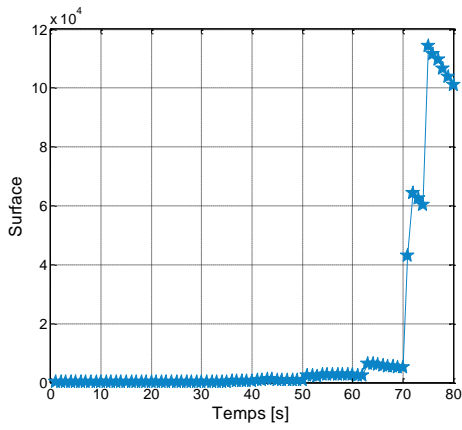


Figure 4-31: (a), (b) et (c); dynamique de la détection de défaut.

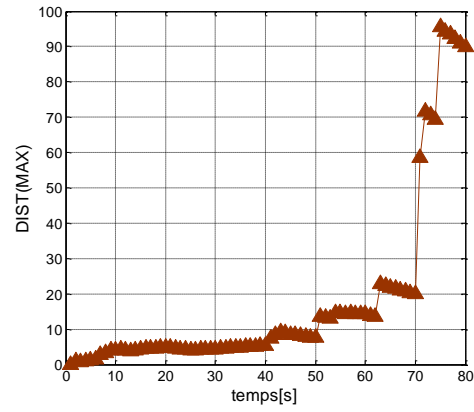
(d) ordre de la distance-accessibilité

4.3.4.2.2 Suivi de la classe de roulement défectueux

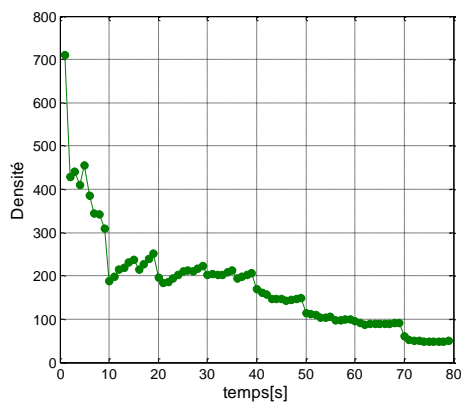
Une fois la détection du défaut est réalisée, le processus de la classification est arrêté et l'étape du suivi des caractéristiques géométriques se débute. La figure 4-32 montre l'évolution de ces dernières dans le temps. Pour le contour, la densité et le saut du centre de classe leur évolution ne montrent pas vraiment des paliers qu'on peut les différencier facilement les uns des autres. Les deux autres caractéristiques, la distance maximale et la surface gardent leur efficacité de suivi.



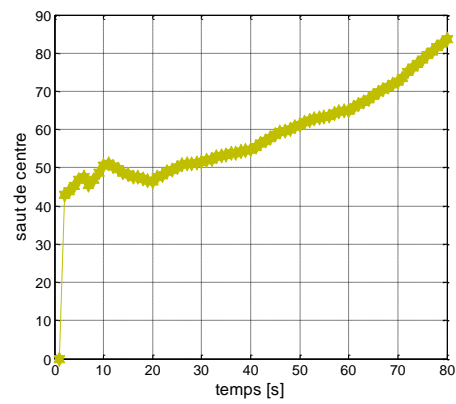
(a)



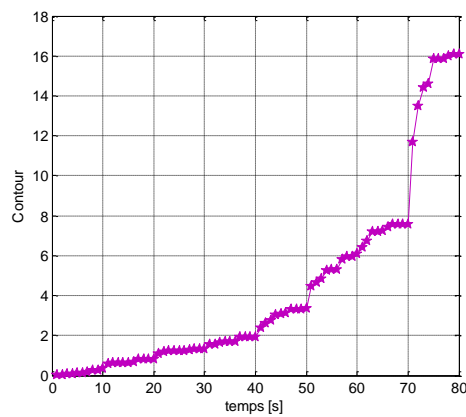
(b)



(c)



(d)



(e)

Figure 4-32: Evolution des caractéristiques géométriques: (a) surface, (b) Distance maximale, (c) densité, (d) saut de centre et (e) le contour.

4.3.4.3 Cas d'une charge variable (vitesse constante)

Dans cette partie, nous avons testé la méthode adaptée sur des données où la vitesse de rotation de moteur est constante (1405 rpm) et une charge variable et bien sûr un défaut évolutif. La figure 4-33 présente l'évolution de la valeur efficace en fonction du kurtosis. On Remarque bien que l'allure présente différentes zones présentant les différents stades de dégradation de roulement.

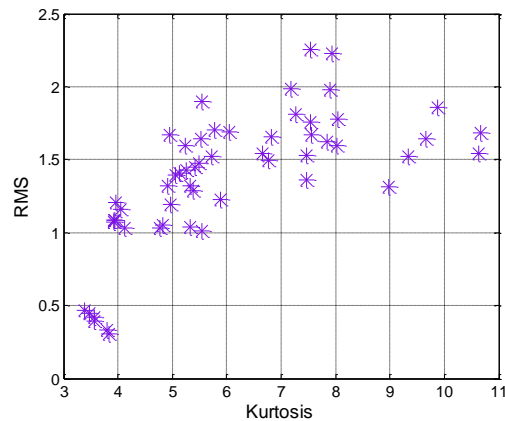


Figure 4-33: Evolution de la valeur efficace RMS en fonction du kurtosis pour différentes charges et une vitesse constante

4.3.4.3.1 Détection du défaut avec la méthode adaptée

La détection de la présence de défaut sur le roulement est présentée sur la figure 4-34.a.b et c. L'évolution de la charge n'a pas d'influence sur la détection de défaut. On Remarque que le défaut est détecté au temps bien précis. On peut conclure que la charge n'influe pas sur la détection de défaut et de cout on peut distinguer le taux de dégradation de roulement.

La figure 4-34.d montre le résultat de classification, on remarque que la courbe présente différentes zones. La première zone est très distincte, elle présente le roulement à l'état sain (de 0 à 60). Les distances accessibilité 61 à 540 présentent le roulement défectueux. A l'intérieur de cette zone on peut distinguer entre différentes autres zones de densité, chacune présente un niveau de dégradation de roulement. En conclusion, la méthode OPTICS permet de visualiser les différentes zones de densité et ainsi la détection de la naissance de défaut.

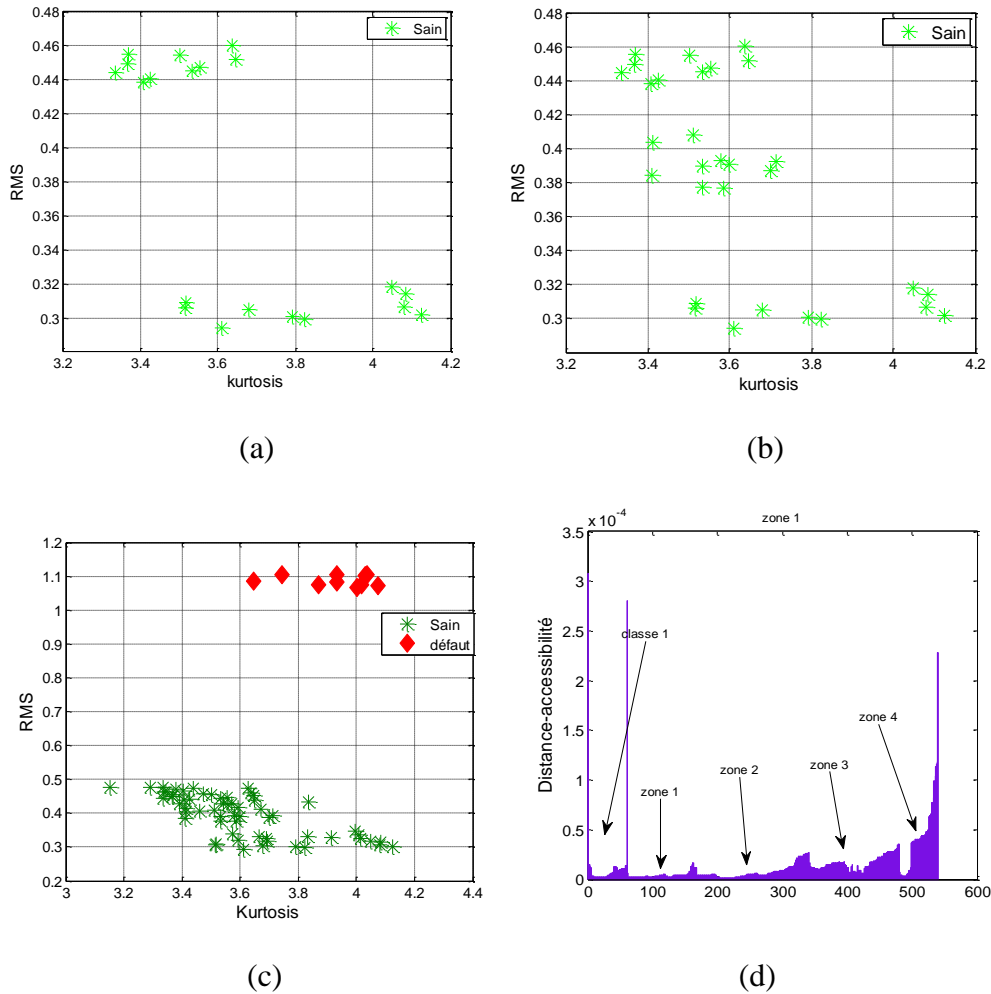


Figure 4-34: (a), (b) et (c); dynamique de la détection de défaut.

(d) ordre de la distance-accessibilité

4.3.4.3.2 Suivi de la classe de roulement défectueux

Le suivi des caractéristiques de la classe présentant le roulement avec défaut est montré sur la figure 4-35. L'allure de contour, de la densité et de saut du centre présente différents paliers traduisant les différents niveaux de dégradation de roulement. Ces caractéristiques montrent leur efficacité comme indicateurs de suivi de développement de défaut.

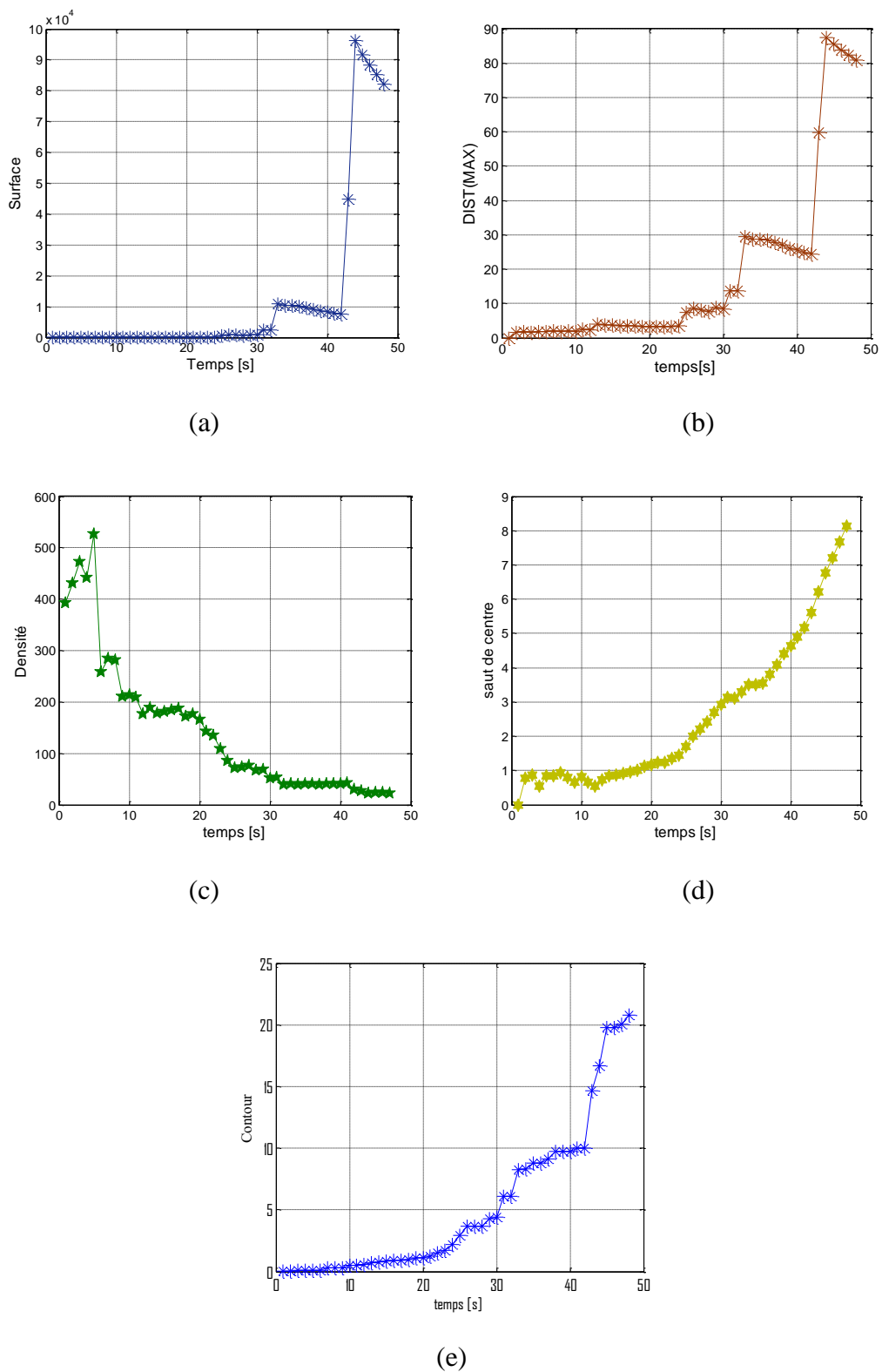


Figure 4-35: Evolution des caractéristiques géométriques: (a) surface, (b) Distance maximale, (c) densité, (d) saut de centre et (e) le contour

4.3.4.4 Cas d'une vitesse et d'une charge variables

4.3.4.4.1 Détection du défaut avec la méthode adaptée

Nous avons mis en œuvre la méthode de classification OPTICS pour la détection de défaut. Les figures (36.a, 36.b et 36.c) montrent la dynamique de la détection de défaut par la méthode RT-OPTICS. Le roulement est jugé toujours comme sain par la méthode adaptée (36.a, 36.b). Dans la figure 4-36.c, le défaut est détecté et la méthode considère deux classes caractérisant l'état de roulement (roulement sain et roulement défectueux).

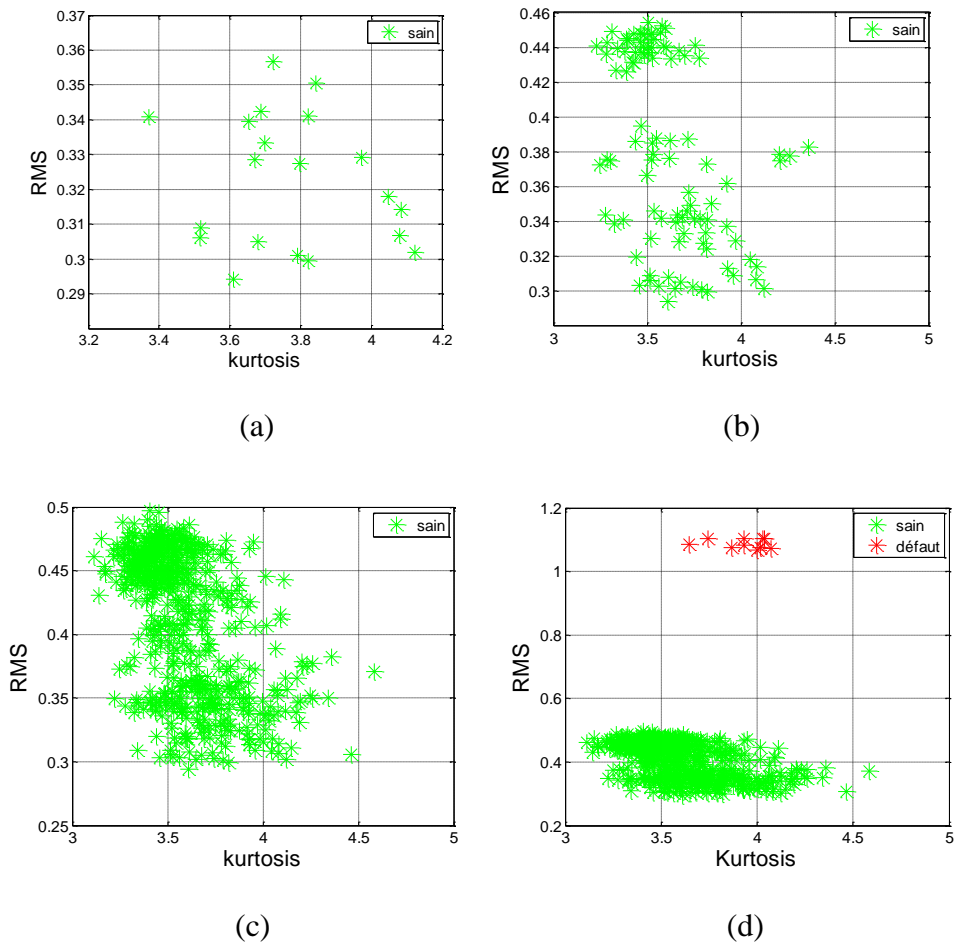


Figure 4-36: Dynamique de la méthode adaptée. (a) Signal 2 : roulement sain ; vitesse = 1405 tr/mn, charge = 130 daN. (b) Signal 8 : roulement sain ; vitesse = 1405 tr/mn, charge = 130 daN. (c) Signal 60 : roulement sain ; vitesse = 1560 tr/mn, charge = 250 daN. (d) Signal 61 : Roulement défectueux, deux classes détectées

L'application de la méthode RT-OPTICS a donné un taux de succès de 100% de l'apparition de défaut, soixante premiers signaux donc 600 points, malgré la variation des caractéristiques de la cinématique de la machine et la perturbation de l'évolution des

indicateurs, en particulier le kurtosis qui est pratiquement le seul indicateur le plus sensible aux chocs. Les résultats montrent clairement que la méthode RT-OPTICS est capable de détecter un petit changement de l'état de santé de roulement. On remarque de plus que le résultat de classification n'est pas influencé par la variabilité des caractéristiques cinématiques et dynamiques de la machine (charge et vitesse). Ceci indique la robustesse de la méthode face à ces changements.

4.3.4.4.2 Suivi de la classe de roulement défectueux

Une fois la classification effectuée et le défaut détecté, un suivi des caractéristiques de la deuxième classe sont calculées.

La figure 37-a illustre l'évolution du saut du centre de la classe représentant le roulement défectueux. L'allure présente différentes zones qui correspondent aux différentes tailles de défauts créés. On remarque une évolution aléatoire du centre de classe suite à l'évolution aléatoire des surfaces des défauts créés.

Le contour de la classe de roulement défectueux est montré sur la figure 4-37.b. On remarque plusieurs zones de différentes pentes qui évoluent linéairement. Ces zones traduisent la dispersion des indicateurs pour les différentes tailles de défauts. A partir de là, ces paramètres peuvent nous donner une idée sur le nouvel état de roulement en comparaison avec les différentes zones.

L'évolution de la densité de la classe de roulement avec défaut est présentée sur la figure 4-37.c. Cette densité traduit la concentration des points en une surface. Dans notre cas, nous avons considéré la surface d'un cercle de diamètre qui présente la distance maximale de tous les points par rapport au centre de la classe. Noter que cette surface n'est plus la surface réelle occupée par les points constituant cette classe. L'évolution de la densité suit une allure opposée à celle du contour. De même que le contour, l'évolution de la densité montre différentes zones qui peuvent être expliqués par le niveau de dégradation de l'élément étudié.

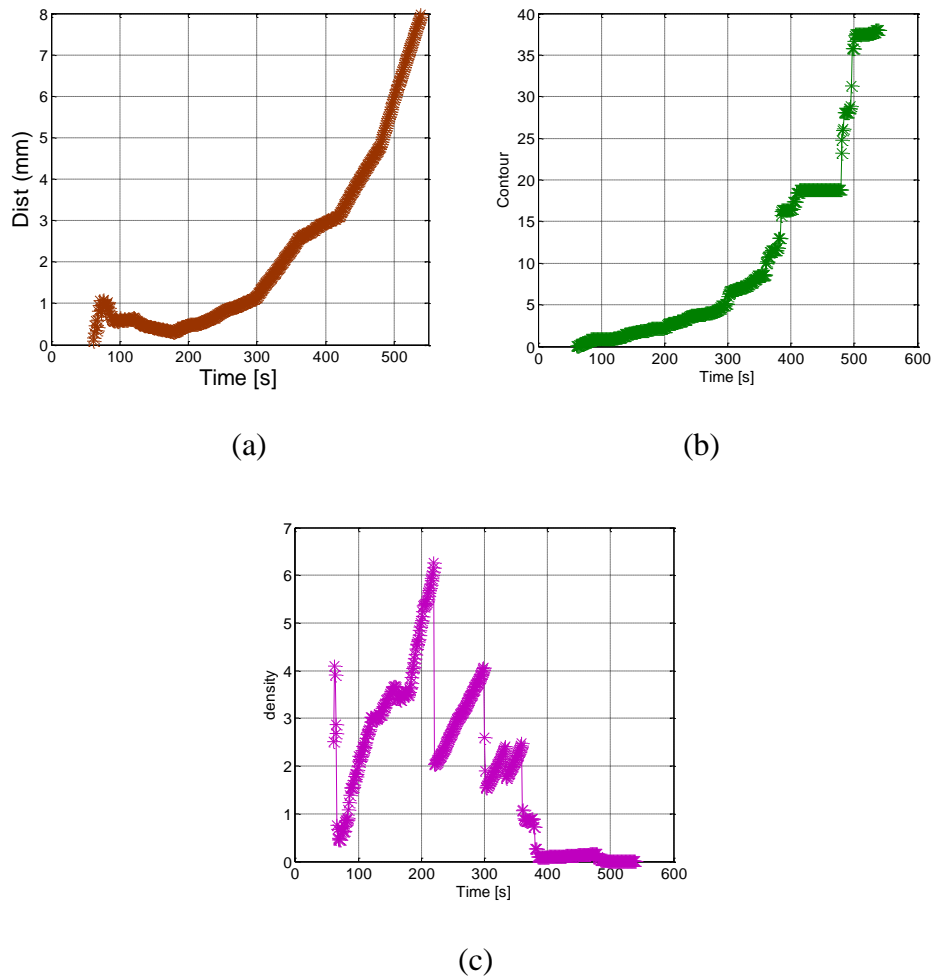


Figure 4-37: Evolution des caractéristiques géométriques: (a) surface, (b) Distance maximale, (c) densité.

4.3.5 Comparaison entre la méthode RT-OPTICS et DDBSCAN

Les résultats du clustering obtenus avec la méthode OPTICS ont été comparés avec ceux obtenus par la méthode de même concept, DBSCAN (Kerroumi et al., 2013). Et cela pour tester la convergence de la méthode adaptée vers une détection très précoce de défaut.

Nous confrontons l'utilisation de la méthode proposée avec la méthode décrite par Kerroumi et al. (2013). L'algorithme permet de suivre l'état de dégradation d'une bague d'un roulement soumis à des fortes charges. Le nombre minimum de points *MinPts* est pris égal au nombre de subdivision d'un seul signal qui égal à 10 alors que le rayon de voisinage est choisi d'une façon empirique et suite aux résultats de Kerroumi et al. (2013). L'algorithme comporte deux principales phases : détection et le suivi. Quand le nouveau signal est acquis, les indicateurs sont extraits, et le processus de classification s'auto-exécute. Une fois la détection

est assurée, la classification est arrêtée et la classification est arrêtée et le suivi des caractéristiques du deuxième groupe est démarré.

Les résultats de classification sont présentés sur la figure 4-38. La méthode DDBSCAN détecte le défaut de façon plus tardive en comparaison avec la méthode RT-OPTICS. Ainsi, le défaut détecté par cette méthode correspond au 62ème signal qui correspond au premier défaut pour la deuxième vitesse de rotation 1415 rpm et une charge de 100 daN (première charge). Les individus du 61ème signal sont considérés comme bruit (deuxième cluster non détecté) alors que la méthode RT-OPTICS détecte le défaut en ce moment.

Cela est dû aux paramètres d'entrée *MinPts* et le rayon de voisinage qui ne remplissent pas la propriété de densité des points dans le cas de soixante deuxième signal. Dans ce cas, la méthode DBSCAN considère ces points comme étant du bruit. Un nouveau cluster sera détecté peut être pour les prochaines mesures.

Le tableau 4-4 résume les résultats obtenus par les deux méthodes en terme de détection. La méthode RT-OPTICS est basée essentiellement, comme son nom l'indique, sur l'ordre des points. Donc, pour de petites augmentations des valeurs des indicateurs, principalement le kurtosis, le second cluster sera détecté.

Signal	Signal 60	Signal 61	Signal 62
DDBSCAN	Non	Non	Oui
RT-OPTICS	Non	Oui	Oui

Tableau 4-4: Comparaison de la détection de défaut entre DBSCAN et RT-OPTICS. *Oui: détection, Non: non détecté.*

La méthode RT-OPTICS a pu détecter le défaut à un taux de 100% par contre le taux de succès de détection précoce de défaut par la méthode DDBSCAN est de 98.33%.

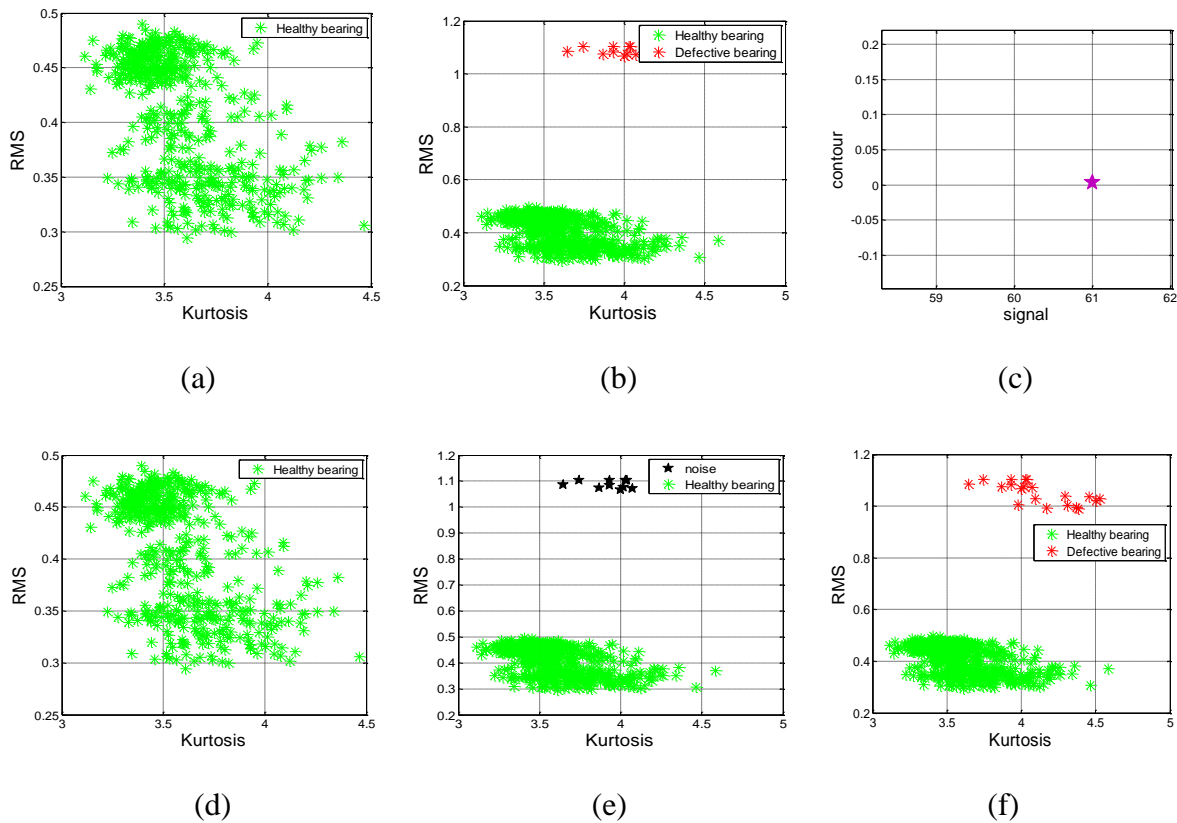


Figure 4-38: Comparaison de la dynamique de détection

La méthode RT-OPTICS: (a) Signal 50 : roulement sain; vitesse=1405 rpm, charge= 130 daN. (b) Signal 61 : roulement défectueux, deux clusters détectés. (c) Signal 62 : suivi du contour

DBSCAN Dynamique: (d) Signal 50 : roulement sain; Vitesse =1405 rpm, charge= 130 daN.(e) Signal 61 : roulement sain; vitesse=1560 rpm, charge= 250 daN) (f) Signal 62 : roulement défectueux, deux clusters détectés

4.4 Conclusion

Une approche dynamique de classification non supervisée de la méthode OPTICS (ordering point to identify clustering structure), a été proposée dans le cadre de la maintenance prédictive pour le diagnostic des défauts sur la bague extérieure de roulement. L’aspect dynamique de cette approche, en intégrant la notion du temps, cible deux volets : la détection et le suivi bidimensionnel de la dégradation de la bague extérieure en utilisant deux indicateurs scalaires, kurtosis et valeur RMS.

Notre approche a été validée sur deux types de signaux : signaux simulés et signaux issus d'un banc expérimental avec des variations de charge et de vitesse de rotation f_r . Dans les deux cas, l'approche proposée a pu différencier entre la présence de défaut et la variation de ces deux derniers paramètres. Nous avons constaté que dans les cas des signaux simulés, l'évolution de ces indicateurs pour une amplitude donnée et pour une charge ou un taux d'amortissement variable prennent des directions très distinctes entre elles et différentes de celle de la variation de l'amplitude. Compte tenu des avantages de cette approche, la détection précoce de défaut sans avoir recours à des méthodes de sélection ou de prétraitement et son efficacité d'automatisation sans influence des caractéristiques cinématiques de la machine, fait qu'elle peut être utile pour la maintenance industrielle. Nous souhaiterons utiliser cette approche sur des signaux vibratoires plus complexes et de renforcer le deuxième volet, le suivi, par une recherche plus approfondie des paramètres du suivi.

Conclusion générale & perspectives

Conclusion générale & perspectives

Dans cette thèse, notre motivation est portée d'abord sur la nécessité d'utilisation des algorithmes dynamiques de clustering dans le diagnostic des défauts mécaniques. De plus, la plupart des versions de ces algorithmes se basent sur les versions statiques qui génèrent des résultats de clustering dépendant du choix du nombre de cluster, aussi de la taille des bases de données étudiées et d'un système apprenant, qui n'est pas évident pour le diagnostic automatisés des machines tournantes.

Nous avons opté pour l'utilisation de l'algorithme de clustering OPTICS car il fournit un ordre pour les objets de la base de données, qui est insensible aux paramètres d'entrées et permet d'identifier la structure de clustering tel que soit la forme arbitraire des clusters et même dans le cas des clusters chevauchés.

L'objectif principal de ce projet est de proposer un algorithme dynamique sans fixation ni du nombre de classe ni d'un système d'apprentissage. Notre approche, Real Time OPTICS, RT-OPTICS, est basée sur la distribution des régions de densité et de niveau d'évolution des indicateurs calculés. La méthode OPTICS permet la visualisation de la distribution des différentes régions de densité.

L'étude est portée sur des signaux de défauts de roulements collectés sur le banc SURVIB au sein de laboratoire (GRESPI/MAN) à l'Université de Reims Champagne-Ardenne. Le banc est un carter accouplé à un moteur de puissance maximale de 10 kW par l'intermédiaire d'un arbre et deux roulements ; roulement à bille sain (roulement à bille 6206) et un roulement à galets à dégrader (N.206.E.G15). La rotation est commandée par un variateur de vitesse. L'arbre est chargé radialement par un vérin via un câble en acier. Les défauts ont été créés par un stylo à gravure sur la bague extérieure du roulement à galet. Les défauts varient légèrement et leurs tailles sont mesurées à l'aide d'une pâte "plastiform". Le profil obtenu est caractérisé comme étant une rugosité avec un profilomètre "Taylor-Hobson subtronic 3P".

Trois cas de signaux ont été étudiés : i) signaux stationnaires, ii) effet de la vitesse de rotation du moteur et iii) effet de la charge. En premier lieu, le diagnostic est réalisé sur des signaux stationnaires, vitesse et charge constante. En deuxième lieu, on a varié la vitesse de rotation en troisième lieu la vitesse est constante et la charge variable et en fin les deux paramètres cinématiques, vitesse et charge, variables. Le but de cette variété est déployée de voir l'effet de chaque paramètre sur l'évolution des indicateurs et sur la détection de défaut de

roulement en utilisant les méthodes de classification non supervisées. Notre approche est basée sur trois étapes principales suivantes :

La première étape consiste en l'utilisation des méthodes de traitement de signal en calculant des indicateurs avec différentes approches et de les suivre dans le temps, parlant des indicateurs temporels, fréquentiels et les indicateurs issus de l'analyse par ondelette. Cette étude montre qu'il n'y a pas un indicateur fiable pour la maintenance automatisée de roulement.

D'après la littérature, le kurtosis et la valeur efficace sans les indicateurs les plus sensibles aux premiers chocs dus de défauts d'écaillage sur les éléments roulant de roulement. Pour cette raison ces deux indicateurs ont été choisis pour construire la matrice des données pour la méthode de clustering choisis.

La deuxième étape consiste à proposer une nouvelle version de la méthode de clustering OPTICS pour la détermination des clusters à partir des distances d'accessibilités et leurs ordres et aussi nous avons introduit la notion de direction de l'évolution des nuages de points. Avec ces informations, nous pouvons détecter le deuxième cluster qui représente le roulement défectueux quelque soit le types des signaux.

Une fois le défaut est détecté, la dernière étape est mise en route, le suivi des caractéristiques géométriques du deuxième cluster dans le temps. Dans cette étude, nous avons calculé la densité de la classe, le saut de centre, le contour, la distance maximale entre les différents points et la surface de la classe. Cependant pour visualiser les étapes de développement de la gravité de défaut et pour tester la validité de ces caractéristiques la méthode OPTICS dans sa version originale a été appliquée sur les données qui correspondent au roulement défectueux.

Une étude comparative entre la méthode proposée et la méthode de classification DDBSCAN, a été réalisée afin de tester la validité de notre approche en temps réel de détection de défaut.

Nous avons réalisé une solution pour l'exécution de la méthode OPTICS indépendamment de la version statique. Dans l'évaluation expérimentale, nous avons montré que notre approche de clustering dynamique donne de bons résultats de diagnostic et en temps d'exécution réduit. L'approche proposée a pu différencier entre la présence de défaut et la variation de ces deux paramètres cinématiques de la machine. Compte tenu des avantages de cette approche, la détection précoce de défaut sans avoir recours à des méthodes de sélection

ou de prétraitement et son efficacité d'automatisation sans influence des caractéristiques cinématiques de la machine, fait qu'elle peut être utile pour la maintenance industrielle. En revanche, elle a pu détecter la présence de défaut en temps avancé en comparaison avec la méthode de clustering DBSCAN. Cette dernière est influencée par le choix des paramètres initiaux, le rayon de voisinage et le nombre de points.

L'application de la méthode OPTICS sur les données de la deuxième classe a montré la présence de différentes zones de densité différentes. Cela justifie notre choix des caractéristiques géométrique cité auparavant. L'évolution de ces caractéristiques dans le temps montre la présence de différents paliers qui traduisent la présence de différentes zones. Ces zones représentent les différents stades de défaillance de roulement. De ce fait, ils peuvent être considérés comme de bons indicateurs de suivi de la sévérité de défauts dans le temps.

Perspectives

Comme perspectives de notre travail, nous proposons tout d'abord d'exécuter cette approche sur plusieurs machines tournantes pour mieux voir l'efficacité de la détection des défauts par la méthode proposée et celui des caractéristiques proposées pour le suivi du niveau de dégradation des organes contrôlés.

Naturellement dans les machines tournantes il peut y avoir plusieurs défauts qui apparaissent sur un composant mécanique (roulement, engrenage, courroie, etc.). C'est pourquoi nous proposons d'appliquer cette approche sur une base de données d'un système contenant plusieurs types de défauts, d'utiliser ainsi plusieurs indicateurs et de mettre en œuvre également des méthodes de sélection des indicateurs les plus pertinents.

Références bibliographiques

- Abbasion S., Rafsanjani A., Farshidianfar A., & Irani N., (2007), Rolling element bearings multi-fault classification based on the wavelet denoising and support vector machine. *Mechanical Systems and Signal Processing*, 21(7): 2933-2945, 26.
- AFNOR. Norme AFNOR X 60-000. Paris Afnor, 2002. 8.
- AFNOR. Norme AFNOR X 60-010. Paris Afnor. 1990.
- Agrawal R., Gehrke J., Gunopulos D., & Raghavan P., (1998). Automatic subspace clustering of high dimensional data for data mining applications. *Proceeding of the ACM SIGMOD Conference*, pp 94-105, Seattle, WA.
- Amadou-Boubacar H. (2010). Supervision adaptative et pronostic de défaillance pour la maintenance prévisionnelle de systèmes évolutifs complexes. PhD thesis, Université des Sciences et Technologies de Lille, 2010.
- Amadou-Boubacar, H. (2002). Outils pour la Détection de Changement de Comportement d'un Système. DEA Automatique & Informatique Industrielle, Université des Sciences et Technologies de Lille (USTL). France.
- Amadou-Boubacar, H., Lecoeuche, S. (2005a). A new kernel-based algorithm for online clustering. Springer-Verlag, Duch, W. and al. (Eds): *International Conference on Neural Networks (ICANN05)*, pp: 583-588. Warsaw, Poland.
- Amadou-Boubacar, H., Lecoeuche, S. (2005b). System Drifts Monitoring using Neural Supervision System. *IASTED proceedings of Artificial Intelligence and Applications (AIA05, ed.)*, pp: 803-808. Innsbruck.
- Ankerst, M., Breunig, M. M., Kriegel, H., & Sander, J. (1999). OPTICS: Ordering Points To Identify the Clustering Structure. *SIGMOD '99 Proceedings of the 1999 ACM SIGMOD international conference on Management of data*, Vol. 8(2), pp. 49-60.
- Antoni, J., & Randall, R. B. (2002). Differential diagnosis of gear and bearing faults. *Journal of Vibration and Acoustics*. *Journal of Vibration and Acoustics*, Vol. 124(2), pp. 165–171.
- Ashraf Abdel-Rahim Ali, (2007). Center-Hung Induced Draft Fan Axial Vibration. *Vibrations*, Volume 23, Number 2, pp 14-16.

- Austria. Amarnath M., Sugumarnan V., & Kumar H. (2013). Exploiting sound signals for fault diagnosis of bearing using decision tree. *Measurement*, xi, 49, 50, 51, 59.
- Aye S.A & Heyns P.S., (2011). Effect of speed and torque on statistical parameters in tapered bearing fault detection. *World Academy of Science, Engineering and Technology*, 5, 20.
- Badaoui M., (1999). Contribution au Diagnostic Vibratoire des Réducteurs Complexes à Engrenages par l'Analyse Cepstrale. PhD thesis, Université Jean Monnet, 1999. 23
- Barber, C. B., Dobkin D.P., & Huhdanpaa H.T.(1996). The Quickhull Algorithm for Convex Hulls. *ACM Transactions on Mathematical Software*, Vol. 22, No. 4, p. 469-483.
- Baydar N. & Ball A. (2001). A comparative study of acoustic and vibration signals in detection of gear failures using wigner-ville distribution. *Mechanical Systems and Signal Processing*, Vol. 15(6), pp. 1091-1107.
- Benchaabane C., Djebala A., Ouelaa N. & Guenfoud S., (2008). Diagnostic vibratoire des défauts d'engrenages base sur les indicateurs scalaires. *Groupe omportement Dynamique des Structures et Maintenance Industrielle*, Guelma, Algerie, P 10.
- Benmahdi D. , Rasolofondraibe L. , Chimentin X. , Murer S., & Felkaoui A., (2017). RT-OPTICS : real-time classification based on OPTICS method to monitor bearings faults, *Journal of Intelligent Manufacturing*, pp. 1-14.
- Bezdek J.C., & Pal N.R. (1998), Some new indexes of cluster validity, *IEEE Trans. On systems, MAN, and cybernetics* VOL. 28, NO. 3.
- Blansché A., (2016). Classification non supervisée avec pondération d'attributs par des méthodes évolutionnaires, PhD thesis, Université Louis Pasteur, Strasbourg, 2006.
- Bonnardot F. (2004). Comparaison entre les analyses angulaire et temporelle des signaux vibratoires de machines tournantes. Etude du concept de cyclostationnarité floue. PhD thesis, Institut National Polytechnique de Grenoble, 36, 37.
- Bonnardot F. (2005). Comparaison entre les analyses angulaire et temporelle des signaux vibratoires de machines tournantes. Etude du concept de cyclostationnarité floue, thèse de doctorat. LAPSI, Université Jean Monnet, Saint-Etienne, janvier 2005.

- Bonnardot, F. (2004). Comparaison entre les analyses angulaire et temporelle des signaux vibratoires de machines tournantes. Etude du concept de cyclostationnarité floue. PhD thesis, Institut National Polytechnique de Grenoble, 2004.36, 37.
- Bouillaut L. & Sidahmed M. (2001). Cyclostationary approach and bilinear approach: comparison, applications to early diagnosis for helicopter gearbox and classification method based on hocs. *Mechanical Systems and Signal Processing*, 15(5): pp923-943.
- Boulenger A., & Pachaud C. (1995). Surveillance des machines par analyse des vibrations : du dépistage au risque, AFNOR, ISBN : 2-12-309012-3, 215.
- Breneur C. (2002). Eléments de maintenance préventive de machines tournantes dans le cas de défauts combinés d'engrenages et de roulements. Thèse de doctorat INSA , 2002.
- Campagna-Varenne A., (1985). Étude des roulements par analyse spectrale, *Vibrations Actualités* 3.
- Capdessus C, Sidahmed M, & Lacoume J.L. (1995). Apport de la théorie des processus cyclostationnaires à l'analyse et au diagnostic des engrenages, 2ème Conférence Internationale des Méthode de Surveillance, Senlis-France, 10-12 Octobre 1995, Vol. 1, pp. 391-401.
- Capdessus C. (1992). Aide au diagnostic des machines tournantes par traitement du signal. Thèse de doctorat ,Institut National Polytechnique de Grenoble, 1992.
- Capdessus S. C., Sid Ahmed M. & Lacoume J.L. (2000). Cyclostationary processes: application In gear faults early diagnosis'' *Mechanical Systems and Signal Processing*. 14(3), pp 371-385.
- Chavent M. & Lechevallier Y. (2002). Dynamical clustering of interval data : Optimization of an adequacy criterion based on hausdor_ distance. In *Classification, clustering, and data analysis*, pages 53_60. Springer, 57.
- Chen Y., Miao D., & Wang R. (2010). A rough set approach to feature selection based on ant colony optimization. *Pattern Recognition Letters*, Vol. 31(3), pp. 226-233.
- Chimentin X., (2007). Localisation et quantification des sources vibratoires dans le cadre d'une maintenance préventive conditionnelle en vue de fiabiliser le diagnostic et le suivi de l'endommagement des composants mécaniques tournants : application aux

roulements à billes. PhD thesis, Université de Reims Champagne-Ardenne, France, 2007.

Davies D. et Bouldin D.(1979). A cluster separation measure. IEEE transactions on Pattern and machine Intelligence, 1(2): 224-227.

Decoulon F. , Théorie et traitement de signal. Edition Dunod 1984.

Dellomo M. R. (1999). Helicopter Gearbox Fault Detection: A Neural Network Based Approach. Journal of Vibration and Acoustics, Vol. 121(3), pp. 265-272.

Didier G., (2004). Modélisation et diagnostic de la machine asynchrone en présence de défaillance, phd thesis, Université Henri Poincaré Nancy 1. 2004.

Dong S., & Luo T., (2013). Bearing degradation process prediction based on the PCA and optimized ls-svm model. Measurement, 46(9) :3143-3152.

Dron J.P., Bolars.F, & Rasolofobdraibe L., (2004), Improvement of the sensitivity of the scalar indicators (crest factor, kurtosis) using a de-noising method by spectral subtraction: Application to the detection of defects in ball bearings, Journal of Sound and Vibration 270 (1-2), pp. 61-73.

Dubuisson, B. (1990). Diagnostic et reconnaissance des formes. Hermès. Paris, 34, rue Eugène Flachat, 75017.

Duda R., Hart P., & Stork D., (2001). Pattern Classification, Willey Interscience Second edition,

Dunn J.C. (1974). Well separated clusters and optimal fuzzy partitions. Journal of cybernetics, 4: 95-104.

Dybala J. & Zimroz R., (2012). Application of empirical mode decomposition for impulsive signal extraction to detect bearing damage : industrial case study. Springer, 36.

Dyer D., & Stewart R.M., (1978), Detection of rolling element bearing damage by statistical vibration analysis, Transactions of the ASME, Journal of Mechanical Design, 100 (2), pp 229-235.

Elghazel H., (2007). Classification et Préviation des Données Hétérogènes :Application aux Trajectoires et Séjours Hospitaliers. PhD thesis, Université Claude Bernard Lyon 1, 2007. 69.

- Ester M., Kriegel H.P., Sander J., & Xu X., (1996). A density-based algorithm for discovering clusters in large spatial databases with noise. In *The 2nd International Conference on Knowledge Discovery and Data Mining*, Portland, . xi, 63, 65.
- Fix E., & Hodges, J.L. (1951), *Discriminatory analysis. Non-parametric discrimination*. USAF School of Aviation Medicine, Randolph Field, TX, 4 Technical Report 4.
- Fouchal S., (2011). *Partitionnement d'objets complexes : méthodes et algorithmes*. Ph.D Thesis, Université Paris 8, Vincennes Saint Denis, France.
- Ghaemi M., Etefagh M. & Yazdanian A., (2014). Bearing fault diagnosis using hybrid genetic algorithm k-means clustering. In *Innovations in intelligent systems and applications*, Alberobello, 2014.
- Ghaemi, M., & Feizi-derakhshi, M. (2014). Expert Systems with Applications Forest Optimization Algorithm. *Expert Systems With Applications*, Vol. 41(15), pp. 6676-6687.
- Ghahramani Z. (2003). *Advanced Lectures on Machine Learning*. Springer-Verlag, 36, 38.
- Ghribi M., Cuxac P., Lamirel J.C., & Lelu A., (2010). Mesures de qualité de clustering de documents : prise en compte de la distribution des mots clés. In *Conférence Internationale Francophone sur l'Extraction et la Gestion des Connaissances - EGC*, Hammamet, Tunisia, 2010. 68.
- Goumas S.K., Zervakis M .E. & Stavrakakis G.S., (2001). Intelligent on-line quality control of washing machines using discrete wavelet analysis features and likelihood classification. *Engineering Applications of Artificial Intelligence*, 2001.
- Grasso M., Pennacchi P., & Colosimo B. M. (2014). Empirical mode decomposition of pressure signal for health condition monitoring in waterjet cutting. *The International Journal of Advanced Manufacturing Technology*, Vol 72, pp347-364.
- Guo, Q., Yu, H. & Xu, A. (2005). Modified Morlet wavelet neural networks for fault detection, *International Conference on Control and Automation (ICCA2005)*, pp.1209-1214.
- Halikidi M., Batistakis Y. & Vazirgiannis M. (2001). On Clustering Validation Techniques, *Journal of Intelligent Information Systems*, 107–145.

- Hastie T., Tibshirani R., & Friedman J., (2009). *The Elements of Statistical Learning :Data Mining, Inference, and Prediction*. Springer, 43, 47.
- Honarvar F., & Martin H.R, (1997), New statistical moments for diagnostics of rolling element bearings, *Journal of Manufacturing Science and Engineering, Transactions of the ASME* 119 (3), pp.425-432.
- Huang N. E., Shen Z., & Long S. R (1998). The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis in: *Proceedings of the Royal Society of London Series*, Vol 454, pp. 903-995.
- Huang N. E., Shen Z., Long S.R., Wu M.C., Shih H. H., Zheng Q., Yen N.C., Tung C.C., & Liu H.H. (1998). The empirical mode decomposition and hilbert spectrum for nonlinear and non-stationary time series analysis. *Proceedings of the Royal Society of London. Series A : Mathematical, Physical and Engineering Sciences*, 29, 34.
- Hubert L., et Arabie P. (1985) . Comparing partitions. *Journal of classification*, 2 : 193-218 .
- Jack, L. B., & Nandi, A. K. (2002). Fault detection using support vector machines and artificial neural networks, augmented by genetic algorithms. *Mechanical Systems and Signal Processing*, Vol. 16(2-3), pp. 373–390.
- Jain A.K., Duin R.P.W., & Mao J. (2000). Statistical pattern recognition : A review. *IEEE*, 38, 39, 42.
- Janecek A.G.K., & Gansterer W. N., (2008). DemelandM.A., & EckerG. F., (2008). On the relationship between feature selection and classification accuracy . In *New challenges for feature selection in data mining and knowledge discovery*, Belgium, 2008. 42,
- Jardine A.K.S., Lin D., & Banjevic D., (2006). A review on machinery diagnostics and prognostics implementing condition-based maintenance. *Mechanical Systems and Signal Processing*, 2006. 12, 24, 25, 33.
- Jiang L., Cao Y., Yin H., & Deng K. (2013). An improved kernel k-mean cluster method and its application in fault diagnosis of roller bearing. *Engineering*, Vol.5(1), pp. 44-49.
- Johnson, S. C. (1967). Hierarchical clustering schemes. *Psychometrika*, 32(3), 241-254.
- Kambhatla N. & Leen T.,(1997). Dimension reduction by local principal component analysis . *Neural Computation*, 9(7) :1493-1516.

- Karabadji N., Seridi H., Khelf I., Azizi N., & Boulkroune R. (2014). Improved decision tree construction based on attribute selection and data sampling for fault diagnosis in rotating machines. *Engineering Applications of Artificial Intelligence*. 49, 50, 59.
- Kerroumi S., (2016). Extraction des paramètres et classification dynamique dans le cadre de la détection et de suivi de défaut de roulements. PhD thesis, Université de Reims Champagne-Ardenne, France.
- Kerroumi, S., Chimentin, X., & Rasolofondraibe, L. (2013). Dynamic classification method of fault indicators for bearings monitoring, *Mechanics & Industry*, Vol 14(2), pp. 115-120.
- Kim Y.H., Tan A.C.C., Mathew J., & Yang B.S. (2006). Condition monitoring of low speed bearings: A comparative study of the ultrasound technique versus vibration measurements. In *Engineering Asset Management*, pages 182-191. Springer London, 20.
- Konar, P., & Chattopadhyay, P. (2011). Bearing fault detection of induction motor using wavelet and Support Vector Machines (SVMs). *Applied Soft Computing Journal*, Vol. 11(6), pp. 4203-4211.
- Kumar H. S., Srinivasa P.P., Sriram N.S., & Vijay G.S. (2013). ANN based evaluation of performance of wavelet transform for condition monitoring of rolling element bearing. *Procedia Engineering*, 53.
- Kurek, J., & Osowski, S. (2010). Support vector machine for fault diagnosis of the broken rotor bars of squirrel-cage induction motor, *Neural Computing and Applications*, Vol. 19(4), pp. 557–564.
- Larose D.T. (2012). *Exploration de données*. Vuibert, 2012. 8.
- Lebart L., Morineau A., & Fenelon J.P, *Traitement des données statistiques*. Dunod, 1979. 68.
- Lecoeuche S., (2006). *Apprentissage Récuratif: Quelques méthodes d'identification et de classification pour le suivi de systèmes évolutifs*. Rapport de HDR, Université de Sciences et Technologies de Lille, France.
- Lecoeuche S., Lurette, C. (2003). Auto-adaptive and Dynamical clustering Neural Network. Springer, ICANN'03 proceedings, pp: 350-358.

- Lee J.H. & Lee Y.T. (2003) Robust technique for estimating the bearings of cyclostationary signals. *Signal Processing*, 36.
- Li C. Q. & Pickering C. J. D. (1992). Robustness and sensitivity of nondimensional amplitude parameters for diagnosis of fatigue spalling. In *Condition Monitoring and Diagnostic Technology*, 20, 140.
- Li H., Deng X., & Dai H. (2007) Structural damage detection using the combination method of emd and wavelet analysis. *Mechanical Systems and Signal Processing*, 36.
- Li H., Lian X., Guo C., & Zhao P. (2013a). Investigation on early fault classification for rolling element bearing based on the optimal frequency band determination. *Journal of Intelligent Manufacturing*, doi:10.1007/s10845-013-0772-8.
- Li H., Zhang Y. & Zheng H. (2009). Gear fault detection and diagnosis under speed-up condition based on order cepstrum and radial basis function neural network. *Journal of Mechanical Science and Technology*, Vol 23,pp 2780-2789.
- Li J. & Ma C., (1997). Wavelet decomposition of vibrations for detection of bearing localized defects. *NDT & E International*,1997. 26.
- Li R., Sopon P., & He D. (2012). Fault features extraction for bearing prognostics. *Journal of Intelligent Manufacturing*, Vol 23, pp313-321.
- Li W.,Gong W, Liang Y. & Chen W. (2005). *Pattern recognition and image analysis* ». Springer, 2005.
- Li Z., Yan X., Tian Z., Yuan C., Peng Z., & Li L. (2013b). Blind vibration component separation and nonlinear feature extraction applied to the nonstationary vibration signals for the gearbox multi-fault diagnosis. *Measurement*, Vol 46, pp 259-271.
- Lurette, C. (2003). Développement d'une technique neuronale auto-adaptative pour la classification dynamique des données évolutives: Application à la supervision d'une presse hydraulique. Thèse de doctorat Université des Sciences et Technologies de Lille, France.
- MacQueen J. (1965). Some methods for classification and analysis of multivariate observations. Dans *proceeding of the 5th Berkeley Symposium on Mathematical Statistics and Probability*. Pages 281-297, Berkeley, CA. University of California Press.

- Mahgoune H., Bekka R.E., & Felkaoui A. (2011). Etude Comparative Entre la Transformée de Hilbert- Huang et la Transformée en Ondelettes Continue Dans la Détection des Défauts de Roulements, Journées d'Etudes Nationales de Mécanique, JENM'2011 Ouargla. 07-08 Mars, 2011.
- McCormick, A. C., & Nandi, A. K., (1997). Classification of the rotating machine condition using artificial neural networks. Proceedings of Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science, Vol. 211(6), pp. 439-450.
- McFadden P. & Smith J., (1984b). Vibration monitoring of rolling element bearings by the high-frequency resonance technique: a review. Mechanical Systems and Signal Processing, 27
- McFadden P.D. & Smith J.D., (1985). The vibration produced by multiple point defects in a rolling element bearing. Journal of Sound and Vibration, Vol 98(2):263-273.
- Misiti T. Les ondelettes et leurs applications. Paris. Hermès Science Publications. : Lavoisier., 2003.
- Monchy F., la fonction maintenance », MASSON, ISBN 2-225-855518-8, Paris 1996.
- Moosavian A., Ahmadi H., & Tabatabaeefar A., (2012). fault diagnosis of main engine journal bearing based on vibration analysis using fisher linear discriminant, k-nearest neighbor and support vector machine. Journal of Vibroengineering, 14(2), 34.
- Morel J. (2005). Surveillance vibratoire et maintenance prédictive. Techniques de l'ingénieur Comportement en service des systèmes et composants mécaniques, TIB180DUO, xi, 17, 19.
- Moret M. (1990), Roulements et butées à billes et à rouleaux, Techniques de l'Ingénieur, ref. B5 370.
- Mori K., Kasashima N., Yoshioka T., & Ueno Y., (1996), Prediction of spalling on a ball bearing by applying the discrete wavelet transform to vibration signals. Wear, Vol. 195, pp. 162-168.
- Mouchaweh M. S. & Lughofer E., (2010) . Learning in Non-Stationary Environments. Springer. 73
- Nakache J.P. & Confais J., 2004. Approche pragmatique de la classification : arbres hiérarchique, partitionnement. Etidion TECHNIP. pp 150-165.

- Nikolaou N. & Antoniadis I., (2002). Demodulation of vibration signals generated by defects in rolling element bearings using complex shifted morlet wavelets. *Mechanical Systems and Signal Processing*, 2002. 34.
- Pachaud C., Crest factor and kurtosis contributions to identify defects inducing periodical impulsive forces. *Mechanical Systems and Signal Processing* 11, 1997, pp. 903-916.
- Pan M.C., & Tsao.W.C (2013) . Using appropriate IMFs for envelope analysis in multiple fault diagnosis of ball bearings, *International Journal of Mechanical Sciences*. Vol 69,pp 114-124.
- Parey A., El Badaoui M., Guillet F., & Tandon N.(2006). Dynamic modeling of spur gear pair and application of empirical mode decomposition-based statistical analysis for early detection of localized tooth defect. *Journal of Sound and Vibration*, 36.
- Peng Z.K. & Chu F.L., (2004), Application of the wavelet transform in machine condition monitoring and fault diagnostics : a review with bibliography. *Mechanical Systems and Signal Processing*, 2004. 27.
- Peng Z.K., Chua F.L., & Peter W. T., (2003): Detection of the rubbing-caused impacts for rotor–stator fault diagnosis using reassigned scalogram. *Mechanical Systems and Signal Processing* 19. pp 391-409.
- Qian Y., Xu L., Li X., Lin X., Kraslawski L., & Lubres A. (2008). An expert system development and implementation for real-time fault diagnosis of a lubricating oil refining process. *Expert Systems with Applications*, Vol 35(3), pp1251-1266.
- Raad A. (2003), Contributions aux statistiques cycliques d'ordre supérieur : applications au diagnostic des défauts d'engrenage, Thèse de doctorat, Université Technologie de Compiègne,2003.
- Rafiee, J., Arvani, F., Harifi, A., & Sadeghi, M. H. (2007). Intelligent condition monitoring of a gearbox using artificial neural network. *Mechanical Systems and Signal Processing*, Vol 21(4), pp 1746–1754.
- Randall R. B. (2011). *Vibration-based condition monitoring : industrial, aerospace and automotive applications*. John Wiley & Sons, Ltd.2011.
- Randall R. B., (2001). The relationship between spectral correlation and envelope analysis in the diagnostics of bearing faults and other cyclostationary machine signals. *Mechanical Systems and Signal Processing*, 29, 36.

- Randall R. B., Antoni J., & Chobsaard S. (2001). The relationship between spectral correlation and envelope analysis in the diagnosis of bearing faults and other cyclostationary machine signals. *Mechanical Systems and Signal Processing*, Vol 15, pp 945-962.
- Randall R.B., (2011). *Vibration-based Condition Monitoring : Industrial, Aerospace and Automotive Applications*. Wiley, 29, 36.
- Rousseuw P. J. (1987). Silhouettes : A graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics*, 69.
- Rubini R. & Meneghetti U., (2001). An application of the envelope and wavelet transform analyses for the diagnosis of incipient faults in ball bearings, *Mechanical Systems and Signal Processing*. Vol 15(2), 287/302.
- Safizadeh M.S., (1999). *Diagnostic des machines dans le plan temps fréquence ; thèse doctorat 1999, école polytechnique de Montréal, CANADA*.
- Saimurugana M., Ramachandran K.I., Sugumaran V., & Sakthivel N.R., (2011). Multi component fault diagnosis of rotational mechanical system based on decision tree and support vector machine. *Expert Systems with Applications*, 51, 59.
- Samanta B, & Al-Balushi K.R. (2003). Artificial neural network based fault diagnosis of rolling element bearing using time-domain features. *Mechanical Systems and Signal Processing*, Vol. 17(2), pp. 317-328.
- Samanta B., & Nataraj C. (2009). Use of particle swarm optimization for machinery fault detection. *Engineering Applications of Artificial Intelligence*, Vol 22 , pp 308-316.
- Samanta B., Al-Balushi K. R., & Al-Araimi S. A. (2003). Artificial neural networks and support vector machines with genetic algorithm for bearing fault detection. *Engineering Applications of Artificial Intelligence*, Vol. 16(7-8), pp. 657-665.
- Samanta B., Al-Balushi K.R, & Al-Araimi S.A. (2001). Use of genetic algorithm and artificial neural network for gear condition diagnostics. *Proceedings of COMADEM*, University of Manchester, UK, pp. 449-456.
- Samanta, B. (2003). Artificial neural networks and support vector machines with genetic algorithm for bearing fault detection. *Engineering Applications of Artificial Intelligence*, Vol. 16(7-8), pp.657-665.

- Samanta, B. (2004). Gear fault detection using artificial neural networks and support vector machines with genetic algorithms, *Mechanical Systems and Signal Processing*, Vol.18(3), pp. 625-644.
- Samuel P. D. & Pines D. J. (2005). A review of vibration-based techniques for helicopter transmission diagnostics. *Journal of Sound and Vibration*, 282 :475- 508, 20.
- Shen C., Wang D., Kong F., & Tse P. W. (2013). Fault diagnosis of rotating machinery based on the statistical parameters of wavelet packet paving and a generic support vector regressive classifier. *Measurement*, Vol 46, pp 1551-1564.
- Staszewski W.J., Worden K. & Tomlinson G.R. Time-frequency analysis in greabox fault detection using wigner-ville distribution and pattern recognition . *MSSP, Mechanical system and signal processing*, 2002.
- Stepanic P., Latinovic I. V., & Djurovic Z. (2009). A new approach to detection of defects in rolling element bearings based on statistical pattern recognition. *International Journal of Advanced Manufacturing Technology*, Vol 45, pp 91-100.
- Su Y.T. & Lin S.J. (1992). On initial fault detection of a tapered roller bearing:Frequency domain analysis. *Journal of Sound and Vibration*, Vol 155(1):75-84.
- Székely G. J. & Rizzo M. L., (2005). Hierarchical clustering via Joint Between-Within Distances: Extending Ward's Minimum Variance Method. , *Journal of Classification*, vol. 22, no 2, septembre 2005, p. 151-183.
- Tan P.N., Steinbach M., & Kumar V., (2005). *Introduction to Data Mining*. Addison Wesley. 36, 62.
- Thomas M. (2011), *Fiabilité, maintenance prédictive et vibration des machines*, ISBN 978-2-7605-3357-8, pp 633.
- Tyagi C.S. (2008). A Comparative Study of SVM Classifiers and Artificial Neural Networks Application for Rolling Element Bearing Fault Diagnosis using Wavelet Transform Preprocessing. *World Academy of Science, Engineering and Technology*, Vol 19, pp 309-317.
- Vapnik V.N. (1998). *Statistical Learning Theory*. wiley Interscience publication, NewYork, 1998.

- Wang F., Jiang H., Shao H., Duan W. & Wu S., (2017), 'An adaptive deep convolutional neural network for rolling bearing fault diagnosis, *Measurement Science and Technology*, Volume 28.
- Wang Y., Kang S., Yang G., & Song L., (2012). Classification of fault location and the degree of performance degradation of a rolling bearing based on an improved hyper sphere structured multi-class support vector machine. *Mechanical Systems and Signal Processing*, 32.
- Wang, C. & Gao, R.X. Wavelet transform with spectral post-processing for enhanced feature extraction. *IEEE Transactions on Instrumentation and Measurement*. 2003, 52, pp. 1296-1301.
- Wong M.L.D., Jack L.B., & Nandi A.K., (2006). Modified self-organising map for automated novelty detection applied to vibration signal monitoring. *Mechanical Systems and Signal Processing*, 20, 88.
- Wu F. & Qu, L. (2008). Diagnosis of subharmonic faults of large rotating machinery. *Mechanical Systems and Signal Processing*.
- Xie X. et Beni G. (1991). A validity measure for fuzzy clustering. *IEEE transactions on Pattern Analysis and machine Intelligence*, 13(8): 841-847.
- Yang Z. L., Wang B., Dong X. H., & Liu H. (2012). Expert System of fault Diagnosis for gear box in wind turbine. *Systems Engineering Procedia*, Vol. 4, pp. 189-195.
- Yang Y., Dejie Yu., & Junsheng C.(2006) A roller bearing fault diagnosis method based on emd energy entropy and ann. *Journal of Sound and Vibration*, 36, 109.
- Yang Y., Yu D., & Cheng J. (2007). A fault diagnosis approach for roller bearing based on IMF envelope spectrum and SVM, *Measurement*, Vol 40, pp943-950.
- Yiakopoulos, C. T., Gryllias, K. C., & Antoniadis, I. A. (2011). Expert Systems with Applications Rolling element bearing fault detection in industrial environments based on a K -means clustering approach. *Expert Systems With Applications*, Vol. 38(3), pp. 2888-2911.
- Yu, Y., Dejie, Yu & Junsheng C., (2006). A roller bearing fault diagnosis method based on EMD energy entropy and ann. *Journal of Sound and Vibration*, 53.

- Zhang J., Yan Q., Zhang Y., & Huang Z., (2006). Novel fault class detection based on novelty detection methods. In International Conference on intelligent computing ICIC, Intelligent computing in signal processing and pattern recognition, Kunming, China. 88.
- Zhang X. & Zhou J., (2013b). Multi-fault diagnosis for rolling element bearings based on ensemble empirical mode decomposition and optimized support vector machines. *Mechanical Systems and Signal Processing*. 48.
- Zhang Y., Zuo H., & Bai F. (2013). Classification of fault location and performance degradation of a roller bearing. *Measurement*, Vol. 46(3), pp. 1178-1189.
- Zhao X, Zuo M.J., Liu Z. & Hoseini M.R. (2013). Diagnosis of artificially created surface damage levels of planet gear teeth using ordinal ranking, *Measurement* Vol 46, pp 132-144.
- Zhu X., Shen L., & Yum T.S.P. (2009). Hausdorff clustering and minimum energy routing for wireless sensor networks. *Transactions on Vehicular Technology IEEE*, 58(2): 990-997, 57.
- Ziani R., Zegadi R., & Felkaoui A. (2012). Bearing Fault Diagnosis Using Neural Network and Genetic Algorithms with the Trace Criterion. In Condition monitoring of machinery in non-stationary operations : proceedings of the Second International Conference "Condition Monitoring of Machinery in Non-Stationary Operations. CMMNO'2012 Springer pp 89-96.

Résumé

Contribution à l'étude du diagnostic des défauts mécaniques par classification non supervisée

Cette thèse a pour objectif de combiner l'analyse vibratoire avec la méthode de reconnaissance des formes basée sur la densité, RT-OPTICS (Real Time Ordering Points to Identify the Clustering Structure), afin d'améliorer la détection de défaut de roulement et de suivre l'évolution de la dégradation dans le temps en utilisant les caractéristiques géométriques de cluster caractérisant le roulement défectueux. Cette approche a été validée sur des signaux non stationnaires, numériques et expérimentale. Une étude comparative est réalisée entre l'approche proposée et la méthode DBSCAN ainsi que le suivi d'indicateurs scalaires et fréquentiels dans le temps a été également réalisé.

Mots clés : Analyse vibratoire, diagnostic, indicateurs, suivi de cluster, roulement, classification non supervisée, DBSCAN, OPTICS.

الملخص:

المساهمة في دراسة تشخيص العيوب الميكانيكية حسب التصنيف غير الخاضع للرقابة.

تهدف هذه الرسالة إلى الجمع بين التحليل الاهتزازي وطريقة التعرف على الأشكال المعتمدة على الكثافة ، RT-OPTICS (الوقت الفعلي لترتيب النقاط لتحديد بنية المجموعات) ، من أجل تحسين اكتشاف عيوب المدحرجات و متابعة تطور التدهور مع مرور الوقت باستخدام الخصائص الهندسية للمجموعة التي تميز المدحرجة. تم التحقق من صحة هذا النهج على إهتزازات غير ثابتة رقمية وتجريبية. كما أجريت دراسة مقارنة بين النهج المقترح و الطريقة DBSCAN و كذا متابعة لمؤشرات العددية والتردد مع مرور الوقت.

كلمات مفتاحية : تحليل الاهتزاز ، التشخيص ، المؤشرات ، المراقبة ، المدحرجات ، التصنيف غير الخاضع للرقابة ، DBSCAN ، OPTICS.

ABSTRACT:

Contribution to the study of the diagnosis of mechanical faults by unsupervised classification

This thesis aims to combine vibration analysis with the method of pattern recognition based on the density, RT-OPTICS (Real Time Ordering Points to Identify the Clustering Structure) to improve fault detection and monitoring the evolution of the degradation over time using the geometric characteristics characterizing the defective bearing cluster. This approach has been validated on non-stationary signals, digital and experimental. A comparative study is carried out between the proposed approach and the DBSCAN method as well as the monitoring of scalar and frequency indicators over time was also carried out.

Keywords: Vibratory analysis, diagnosis, indicators, monitoring cluster, bearing, unsupervised classification, OPTICS.