

UNIVERSITÉ DE REIMS CHAMPAGNE-ARDENNE

ÉCOLE DOCTORALE SCIENCES TECHNOLOGIE SANTE (547)

THÈSE

Pour obtenir le grade de

DOCTEUR DE L'UNIVERSITÉ DE REIMS CHAMPAGNE-ARDENNE

Discipline : AUTOMATIQUE, SIGNAL, PRODUCTIQUE, ROBOTIQUE

Présentée et soutenue publiquement par

Sanaa KERROUMI

Le 21 octobre 2016

Extraction des paramètres et classification dynamique dans le cadre de la détection et du suivi de défaut de roulements

Thèse dirigée par **LANTO RASOLOFONDRAIBE**

JURY

M. François GUILLET,	, Professeur,	à l'Université de Saint Etienne Jean Monnet ,	, Président
M. Moamar Sayed MOUCHAWEH,	, Professeur,	EMD DOUAI ,	, Rapporteur
M. Valéry BOURNY,	, Maître de Conférences HDR,	à l'Université Amiens Picardie Jules Verne ,	, Examineur
M. Xavier CHIEMENTIN,	, Maître de Conférences HDR,	à l'Université Reims Champagne-Ardenne ,	, Examineur
M. Jean Paul DRON,	, Professeur,	à l'Université Reims Champagne-Ardenne ,	, Examineur
Mme. Danielle NUZILLARD,	, Professeur,	à l'Université Reims Champagne-Ardenne ,	, Examineur
M. Lanto RASOLOFONDRAIBE,	, Maître de Conférences HDR,	à l'Université Reims Champagne-Ardenne ,	, Examineur
M. Roger SERRA,	, Maître de Conférences HDR,	INSA Centre Val de Loire ,	, Examineur



Abstract

Vibration analysis remains the most popular and most effective tool for monitoring the internal state of an operating machine. Through vibration analysis, fault indicators are extracted then monitored to detect/ locate the presence of a defect. However, counting solely on the evolution of these fault indicators to diagnose a machine can cause false alarms and question the reliability of the diagnosis. In this thesis, we combined vibration analysis tools with pattern recognition method to firstly improve fault detection reliability of bearings, secondly to assess the severity of degradation by closely monitor the defect growth and finally to estimate their remaining useful life. For these reasons, we have designed a pattern recognition process capable of; identifying defect even in machines running under non stationary conditions, processing evolving data and can handle an online learning. Three dynamic classification methods have been developed : Dynamic DBSCAN was developed to capitalize on the time evolution of the data and their classes, Evolving Scalable DBSCAN that was created to include the online processing and finally Dynamic Fuzzy Scalable DBSCAN a dynamic fuzzy and semi-supervised version of ESDBSCAN. With these techniques we were able to enhance the reliability of fault detection by identifying the origin of the fault indicators evolution. The application of the designed process on real data helped us prove the legitimacy of the proposed techniques in identifying the different states of bearings over time (healthy or normal, defective) and the origin of the observations' evolution with a low error rate, a reliable diagnosis and a low memory occupation.

Résumé

Parmi les techniques utilisées en maintenance, l'analyse vibratoire reste l'outil le plus efficace pour surveiller l'état interne des machines tournantes en fonctionnement. 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. 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 développé trois méthodes de classification dynamique : Dynamic DBSCAN qui a été développée pour suivre les évolutions des classes, Evolving scalable DBSCAN qui représente une version en ligne et évolutive de DBSCAN et finalement Dynamic Fuzzy Scalable DBSCAN qui est une version dynamique et floue ESDSCAN adaptée pour un apprentissage en ligne. Ces méthodes distinguent les variations des observations liées au changement du mode de fonctionnement de la machine (variation de charges) et les variations liées au défaut. Ainsi, Elles permettent de détecter, de façon précoce, l'apparition de défaut. L'application sur des données réelles a permis d'identifier les différents états du roulement au cours du temps (sain ou normal, défectueux) et l'évolution des observations liée à la variation de charges avec un taux d'erreur faible et d'établir un diagnostic fiable.

To Him, my eternal companion
To my sun, my moon , my angel and my better me
(dad, mom, Malak and Saloua)
To my very own cheerleaders and my die-hard fans
(my girls, my family)
To you whom I haven't met yet.

Remerciements

La travail de synthèse et de précision qui nécessite la rédaction d'une thèse a été pour moi un vrai challenge, mais à présent que je me trouve face à mes remerciements je me sens toute aussi impuissante. Par qui commencer mes remerciements ? comment exprimer ma reconnaissance et ma gratitude envers des personnes qui sans leurs soutiens ce travail n'aurait pas vu le jour ?

A mes directeurs de thèse,

Je commence par toi **Lanto Rasolofondraibe**, Merci de m'avoir choisi pour être ta doctorante, merci d'avoir cru en moi dans le temps que moi-même j'avais de la peine à croire en moi, merci d'avoir su me diriger quand il fallait et me laisser voler de mes propres ailes quand je pouvais. Merci de m'avoir consacré énormément du temps et d'efforts pour que je puisse accomplir cette thèse dans les meilleures conditions possibles. Merci d'avoir aidé à traduire mes idées parfois incohérentes en des mots et expressions compréhensibles. Merci pour tes encouragements, tes conseils, ta disponibilité, ta bienveillance. J'ai énormément appris à tes cotés. Merci pour tout.

Xavier Chimentin, quand je t'ai rencontré la première fois, la première chose que je me suis dite après "il est plus jeune que je l'ai pensé" a été "Ca va être amusant de travailler avec lui " et je suis contente que je n'avais pas tort. Merci Xavier d'avoir su me rendre toute de suite à l'aise, de m'avoir faire sentir comme une collègue, et de m'avoir bien accueillie. Aussi merci de m'avoir doucement poussé à découvrir un domaine que je connaissais peu et que maintenant me fascine (La reconnaissance de formes et l'intelligence artificielle). Merci de m'avoir montré que le monde de la recherche pouvait être un univers passionnant.

Aux membres du jury,

Je tiens à vous remercier chaleureusement d'avoir accepté de juger mes travaux et d'avoir dégagé du temps pour vous y consacrer. J'ai été touchée que vous ayez répondu si rapidement, et avec enthousiasme, à ma demande. Merci à Professeur **Moamar Sayed-Mouchaweh**, d'avoir bien voulu être rapporteur de mes travaux. Je vous suis reconnaissante d'avoir porté votre regard d'expert sur mon manuscrit et de vos remarques on ne peut plus pertinentes pour l'améliorer. Votre contribution est pour moi essentielle dans la mesure où la reconnaissance de formes a été pour moi la pierre angulaire dans la conception de mon projet et dans l'orientation de mes travaux. Professeur **François Guillet**, vous avez aussi accepté d'être rapporteur dans mon jury. Je vous suis vraiment grée pour votre investissement dans l'évaluation de mon travail et pour votre implication très positive. Au cours de la période assez anxiogène de préparation de cet examen, la conversation téléphonique ainsi que le rapport sur mon manuscrit que vous avez transmis à l'école doctorale m'ont fait beaucoup de bien. Merci beaucoup au Professeur **Jean Paul Dron** d'avoir accepté de participer à mon jury et au Professeur **Danielle Nuzillard** de m'avoir accueillie au sein de l'équipe SIC. Votre implication à mes côtés a toujours été sans faille. J'ai eu la chance de vous côtoyer au cours de ma thèse, et vous m'avez toujours reçue avec patience et gentillesse. Cela a été pour moi un vrai privilège d'interagir avec vous et j'en garderai un excellent souvenir.

Les remerciements suivants reviennent aux messieurs **Roger Serra** et Monsieur **Valery Bourny** d'avoir accepté d'examiner mes travaux de recherche et de se déplacer pour assister à la présentation de ce travail.

Aux membres et Ex-membres de laboratoire Crestic et Grespi,

Merci de m'avoir accueilli parmi vous, de m'avoir montré d'intérêt à mes avancés et mes recherches. Un gros merci à **Ida Lenclume**, qui est pour moi beaucoup plus qu'une simple collègue mais une amie précieuse, Merci pour ton écoute, pour ton aide et soutien, Merci d'avoir toujours veiller sur moi. Merci à Mme **Chantal Vermeil** d'avoir rendu notre travail quotidien

plus agréable avec votre gentillesse (merci pour les cadeaux de Noël déposés sur mon bureau très tôt le matin). Merci à tous les doctorants de Crestic et Grespi, merci de m'avoir fait sentir que je ne suis pas la seule à stresser et de m'avoir d'oublier momentanément le travail dans des soirées, repas, sorties ou autres. Merci d'avoir partagé mon sens d'humeur peu commun. Merci pour toutes les personnes formidables que j'ai rencontré durant ma thèse.

A mes amis et ma famille,

Toute ma gratitude revient à la **famille Alami** qui m'a complètement adopté comme leur fille, de m'avoir accueillie chez eux pendant la partie la plus difficile de ma thèse. Merci de m'avoir montré tant de tendresse et de générosité. Je ne pourrai jamais vous remercier assez. Merci à toutes les filles merveilleuses, excentriques et complètement folles que j'avais la chance de rencontré (*Aya, Abla, Katie, Nadaa, Nehla, Lara, Magalie, Khatou, Fatimetou, Rajae, Dina, Amouna, Houda, fatima zahera, . . .*) Merci pour les anniversaires surprises, pour votre écoute et soutien. Merci pour tout.

Enfin, les mots les plus simples étant les plus forts, j'adresse tout mon affection à ma famille, en particulier à **mes parents** qui m'ont soutenu tout au long de ma thèse moralement et physiquement et d'avoir toujours été fiers de moi, merci à ma sœur jumelle **Saloua** (my better half) qui a toujours été à mon écoute, et merci à ma petite sœur (**Malak**) merci pour ton amour et tes encouragements. Merci à mes grands-parents, mes oncles et tantes et mes petits et grands cousins. Merci d'avoir fait de moi ce que je suis aujourd'hui. Je vous adore.

Je tourne une page et j'en ouvre une autre vivement la suite

Table des matières

Table des figures	xi
-------------------	----

Liste des tableaux	xvii
--------------------	------

1	Traitement du signal pour la maintenance	5
1.1	Stratégies de maintenance	6
1.1.1	Définition	6
1.1.2	Rôle de la maintenance	7
1.1.3	Types de maintenance	7
	A. Maintenance corrective	7
	B. Maintenance préventive	8
1.1.4	Choix du type de maintenance	10
1.2	Diagnostic des machines tournantes par analyse vibratoire	10
1.2.1	Présentation générale de l'analyse vibratoire	11
1.2.2	Les différents niveaux d'analyse vibratoire	12
1.3	Les principaux défauts de roulements	13
1.3.1	Définition de défaut	13
1.3.2	Défauts des roulements	14
1.4	Les techniques de traitement du signal appliquées à l'analyse vibratoire . .	17
1.4.1	Analyse temporelle	18
1.4.2	Analyse spectrale	21
1.4.3	Analyse d'enveloppe	22
1.4.4	Analyse cepstrale	23
1.4.5	Analyse temps-fréquence	24
1.4.6	Analyse temps-échelle	26

TABLE DES MATIÈRES

1.4.7	Décomposition modale empirique (EMD)	29
1.4.8	Cyclo-stationnarité	31
1.4.9	Approches statistiques	32
1.4.10	Approches issues de l'intelligence artificielle	33
1.5	Limitations des méthodes classiques de diagnostic et suivi	34
2	Méthodes de reconnaissance de formes dans le cadre de diagnostic des machines tournantes	35
2.1	Présentation générale	36
2.2	Approches de reconnaissance de formes	37
2.2.1	La comparaison à des modèles (Template matching)	37
2.2.2	L'approche statistique (statistical pattern recognition)	37
2.2.3	L'approche syntaxique ou structurelle (Syntactic or structural matching)	38
2.2.4	Les Réseaux de neurones (Neural networks)	38
2.3	Méthode de reconnaissance de formes statistique	39
2.3.1	Acquisition des données	39
2.3.2	Prétraitement	39
2.3.3	Extraction de caractéristiques	40
2.3.4	Sélection de caractéristiques pertinentes	42
2.3.5	Classification	44
A.	Méthodes de classification du type apprentissage supervisé	45
B.	Méthodes de classification du type apprentissage non supervisé	53
2.3.6	Post-traitement	66
A.	Mesures de performance de la classification supervisée	66
B.	Mesures de performance des méthodes de classification non supervisées	68
2.3.7	Règles de décision	70
A.	Approche analytique	70
B.	Approche statistique	71
2.4	Limitations de méthodes de reconnaissance des formes employées dans le cadre de diagnostic	72

3	Processus de développement des systèmes de reconnaissance de formes dynamiques pour le diagnostic des roulements	75
3.1	La reconnaissance de formes pour le suivi	77
3.1.1	Le contexte général	77
3.1.2	Evolution des observations caractéristiques d'un système de production	77
A.	Saut d'observation	78
B.	Dérive d'observation	78
3.1.3	Evolutions d'une classe	79
3.1.4	Les différentes contraintes à gérer pour concevoir une méthode de RdF adaptée au suivi des composants mécaniques	80
A.	Les contraintes liées à la de gestion de flux	80
B.	Les contraintes liées au suivi d'un composant	81
C.	Les contraintes liées à la conception de la méthode RdF	81
3.1.5	Les caractéristiques des méthodes de RdF pour les systèmes évolutifs	82
A.	L'adaptabilité du classifieur	82
B.	La non-exhaustivité des données	83
3.2	La classification dynamique	85
3.2.1	Définitions et principes	85
3.2.2	Aspect de la classification dynamique	87
A.	Traitement de nouveautés	87
B.	Dynamique des classes	88
3.2.3	Conclusion sur la classification dynamique	93
3.3	Conclusion du chapitre	93
4	Contribution à la classification dynamique pour le suivi de roulements	95
4.1	Présentation de l'architecture de la méthode de reconnaissance de formes conçue	95
4.2	Extraction des indicateurs de défauts	96
4.3	L'extraction des caractéristiques non physiques par les méthodes RD	97
4.3.1	L'extraction par préservation de structure locale de données	98
4.3.2	Online Kernel Principal components analysis	100
4.3.3	Sélection des caractéristiques les plus informatives	101

TABLE DES MATIÈRES

4.3.4	Bilan comparatif sur les méthodes d'extraction des caractéristiques non physiques	103
4.4	Classification	104
4.4.1	Architecture générale des méthodes de classification proposées . . .	104
4.4.2	La détection des nouveautés	106
4.4.3	La méthode utilisée pour détecter les outliers en ligne	107
4.4.4	La création d'une nouvelle classe	107
4.5	Méthodes de classification dynamiques proposées	107
4.6	Dynamic DBSCAN <i>DDBSCAN</i>	108
4.6.1	Présentation générale	108
4.6.2	<i>DDBSCAN</i>	108
A.	Phase 1 : Classification	109
B.	Phase 2 : Détection et Adaptation	111
C.	Phase 3 : Validation	111
4.6.3	Bilan et limitations de la classification par Dynamic DBSCAN <i>DDBSCAN</i>	111
4.7	Evolving Scalable DBSCAN	112
4.7.1	Présentation générale	112
4.7.2	Scalable DBSCAN	112
4.7.3	Evolving Scalable DBSCAN <i>ESDBSCAN</i>	115
A.	Classification	115
B.	Adaptation et validation	121
4.7.4	Bilan et limitations de la classification par <i>Evolving scalable DBSCAN</i>	121
4.8	Dynamic Fuzzy Scalable DBSCAN <i>DFSDBSCAN</i>	121
4.8.1	Présentation générale	121
4.8.2	Fuzzy Neighborhood DBSCAN	123
4.8.3	Scalable fuzzy Neighborhood DBSCAN ou <i>SFN – DBSCAN</i> . . .	126
4.8.4	Dynamic fuzzy scalable DBSCAN <i>DFSDBSCAN</i>	129
A.	La classification ponctuelle : Désignation des points représentatifs	130
B.	Détection et adaptation	133
C.	La validation	134

4.8.5	Bilan et limitations de la classification par <i>DFSDBSCAN</i>	134
4.9	Bilan sur les classifieurs proposés	135
4.10	Exploitation de la classification dynamique dans le suivi	135
4.11	Conclusion du chapitre	136
5	Validation expérimentale et mise en œuvre sur un banc de fatigue	137
5.1	Mise en œuvre de la méthode de classification dynamique sur un banc expérimental	139
5.1.1	Description du modèle mécanique et du banc expérimental	139
5.1.2	Extraction des caractéristiques physiques	140
A.	Indicateurs issus de l'analyse temporelle	140
B.	Indicateurs de défaut issus de l'analyse spectrale	142
C.	Conclusion sur les indicateurs de défaut	144
5.1.3	Sélection et extraction des caractéristiques non physiques	147
A.	Sélection des caractéristiques informatives	147
B.	Extraction des caractéristiques non physiques par les mé- thodes de réduction de dimension	149
5.1.4	Classification dynamique	155
A.	Dynamic <i>DBSCAN</i>	155
B.	Evolving scalable <i>DBSCAN</i>	165
C.	Dynamic Fuzzy Scalable <i>DBSCAN</i>	173
D.	Conclusion	181
5.1.5	Bilan sur la classification dynamique	182
5.1.6	Suivi de dégradation du roulement par les paramètres caractéris- tiques des classes	183
5.2	Mise en œuvre de la méthode de classification dynamique sur un banc de fatigue	188
5.2.1	Description du modèle mécanique et du banc expérimental	188
5.2.2	Extraction des indicateurs de défaut	189
5.2.3	Classification dynamique	192
A.	<i>KPCA – DDBSCAN</i>	192
B.	<i>NPE – ESDBSCAN</i>	195
C.	<i>LPP – DFSDBSCAN</i>	198

TABLE DES MATIÈRES

5.2.4	Bilan sur les résultats obtenus par les classifieurs proposés	201
5.2.5	Suivi de dégradation du roulement par les paramètres caractéristiques des classes	201
5.3	Conclusion du chapitre	205
6	Conclusion générale et perspectives	209
A	Annexes - Décomposition empirique modale pour le diagnostic de roulements	213
A.1	Décomposition modale empirique	213
A.2	La décomposition en valeurs singulières SVD des IMFs	214
A.3	Indicateurs issus de la décomposition empirique modales	215
	Références	219

Table des figures

1.1	Les stratégies de maintenance	8
1.2	Thomson 2001 (TM01)	15
1.3	Bonnet 2008(BY08)	15
1.4	Ecaillage	16
1.5	Géométrie d'un roulement (Mor05)	17
1.6	Signal temporel d'émission acoustique issu d'un système d'engrenage . . .	23
1.7	la récurrence d'un défaut à 10Hz (roue primaire) est traduite par un peigne de Dirac de fondamental à 100 ms	23
1.8	Coordonnées tridimensionnelles montrant le temps, la fréquence et l'am- plitude d'un signal (Oul14)	24
1.9	les étapes de l'application de la transformée d'ondelettes	28
2.1	Diagramme d'un système de reconnaissance des formes statistique	40
2.2	Les différentes approches utilisées dans la RdF statistique	45
2.3	Les étapes formant le processus de classification supervisée	46
2.4	Schéma du principe de l'arbre de décision	50
2.5	L'arbre de décision présenté dans l'article (ASK13)	51
2.6	L'architecture d'un réseau de neurones	52
2.7	Un exemple de dendrogramme (CHA)	60
2.8	méthode des centres mobiles	60
2.9	des clusters avec des densités différentes (EK SX96)	63
2.10	Création d'un cluster avec DBSCAN	65
3.1	Saut d'observation : a) saut vers une région non définie b) saut vers une classe définie	78

TABLE DES FIGURES

3.2	Dérive d'observation : a) dérive vers une classe définie b) dérive vers une région non définie	79
3.3	Exemples des évolutions des classes : a) rotation de la classe C . b) déplacement de la classe C . c) fusion de deux classes C_1 et C_2 en C_3 . d) scission de la classe C_1 en deux classes C_2 et C_3	80
3.4	Création d'une nouvelle classe	89
3.5	Types d'évolutions d'une classe dynamique	90
3.6	Fusion des classes	91
3.7	Scission de la classe C_1 en classes C_2 et C_3 après la classification des observations nouvellement arrivées	92
3.8	La suppression des classes non représentatives	93
4.1	L'organigramme général de la méthode RdF proposée	96
4.2	La sélection dans un environnement non supervisé (<i>Wrapper</i>)	103
4.3	l'architecture générale de la méthode de classification dynamique pour le suivi	105
4.4	Deux types d'outliers rencontrés dans le suivi	106
4.5	L'organigramme général de la classification dynamique	109
4.6	Exemples de <i>CovRad</i> et <i>CovCnt</i> pour deux points p_1 et p_2	114
4.7	L'organigramme de la classification par <i>ESDBSCAN</i>	116
4.8	Les points P_1 et P_2 ont la même densité selon <i>DBSCAN</i> , alors qu'ils ont des densités différentes	123
4.9	Définition ferme de densité (<i>DBSCAN</i>)	124
4.10	Définition floue par une fonction d'appartenance linéaire	124
4.11	Définition floue par une fonction d'appartenance exponentielle	124
4.12	Différence entre une définition ferme et floue de voisinage d'un point x . .	124
4.13	Outlier par SFN-DBSCAN	129
4.14	L'organigramme de la classification par <i>DFSBSCAN</i>	134
5.1	Procédure de suivi par la classification dynamique	138
5.2	Banc expérimental et les roulements utilisés durant l'essai	139
5.3	Evolution des indicateurs en présence de défaut en cas d'un roulement chargé à 15DaN	141

5.4	Comparaison des indicateurs en présence de défaut pour un roulement chargé à 30 DaN	141
5.5	Comparaison des indicateurs en présence de défaut pour un roulement chargé à 45 DaN	141
5.6	les indicateurs fréquentiels calculés à partir de la puissance de spectre . . .	142
5.7	Les indicateurs d'enveloppe calculés à partir de la puissance de spectre extrait par les ondelettes	143
5.8	L'évolution de l'indicateur X_{Wrms} pour des roulements sous différentes charges	144
5.9	L'évolution de l'indicateur X_{PCWT} pour des roulements sous différentes charges	144
5.10	Ressemblance des courbes de tendances de kurtosis	145
5.11	Ressemblance des courbes de tendances de FC	146
5.12	La sélection par SFS des indicateurs de défauts extraits des signaux vibratoires issus de roulements avec taille de défaut variable (cas charge constante 15DaN)	147
5.13	La sélection par SFS des indicateurs de défauts extraits des signaux vibratoires issus de roulements avec taille de défaut variable (cas charge variable)	148
5.14	Sélection des indicateurs par SFS non supervisée	148
5.16	L'application de la $KPCA$ sur des roulements sous différentes charges (taille de défauts de $00mm^2$ à $06mm^2$ et charge de 15DaN à 45DaN) . . .	151
5.15	L'application de la $KPCA$ sur des roulements sous la même charge (15DaN) . . .	151
5.17	Extraction des caractéristiques par LPP	152
5.18	Extraction des caractéristiques par NPE	153
5.19	Résultats de classification quand on augmente la charge d'un roulement sain	157
5.20	Classification des observations reçues pour des roulements au début de dégradation et sous charge variable	158
5.21	Classification des observations reçues des roulements sous charge variable et avec un niveau de dégradation avancé	159
5.22	Classification des observations reçues d'un roulement sain sous charge variable et un roulement au début de dégradation sous charge variable . . .	161

TABLE DES FIGURES

5.23	Classification des observations reçues des roulements sous charge variable et avec un niveau de dégradation avancé	162
5.24	Classification des observations reçues pour des roulements sous charge variable et avec un niveau de dégradation avancé	163
5.25	Classifications des observations reçues des roulements en différents états et sous charge variable	164
5.26	Classification des premières observations reçues pour un roulement sain sous différentes charges et des roulements au début de dégradation sous charge variable	166
5.27	Classification des observations reçues des roulements sous charge variable et avec un niveau de dégradation avancé	167
5.28	Classification des premières observations reçues pour un roulement sain sous différentes charges et des roulements au début de dégradation sous charge variable par <i>LPP – ESDBSCAN</i>	168
5.29	Classification des observations reçues des roulements sous charge variable et avec un niveau de dégradation avancé par <i>LPP – ESDBSCAN</i>	169
5.30	Classification des premières observations reçues pour un roulement sain sous différentes charges et des roulements au début de dégradation sous charge variable par <i>NPE – ESDBSCAN</i>	170
5.31	Classification des observations reçues des roulements sous charge variable et avec un niveau de dégradation avancé par <i>NPE – ESDBSCAN</i>	171
5.32	Classification des premières observations reçues pour un roulement sain sous différentes charges et des roulements au début de dégradation sous charge variable par <i>SFS – ESDBSCAN</i>	172
5.33	Classification des observations reçues des roulements sous charge variable et avec un niveau de dégradation avancé par <i>SFS – ESDBSCAN</i>	173
5.34	Classification des premières observations reçues pour un roulement sain sous différentes charges et des roulements au début de dégradation sous charge variable par <i>KPCA – DFSDBSCAN</i>	174
5.35	Classification des observations reçues des roulements sous charge variable et avec un niveau de dégradation avancé par <i>KPCA – DFSDBSCAN</i>	175

5.36	Classification des premières observations reçues pour un roulement sain sous différentes charges et des roulements au début de dégradation sous charge variable par <i>LPP – DFSDBSCAN</i>	176
5.37	Classification des observations reçues des roulements sous charge variable et avec un niveau de dégradation avancé par <i>LPP – DFSDBSCAN</i> . . .	177
5.38	Classification des premières observations reçues pour un roulement sain sous différentes charges et des roulements au début de dégradation sous charge variable par <i>NPE – DFSDBSCAN</i>	178
5.39	Classification des observations reçues des roulements sous charge variable et avec un niveau de dégradation avancé par <i>NPE – DFSDBSCAN</i> . .	179
5.40	Classification des premières observations reçues pour un roulement sain sous différentes charges et des roulements au début de dégradation sous charge variable par <i>SFS – DFSDBSCAN</i>	180
5.41	Classification des observations reçues des roulements sous charge variable et avec un niveau de dégradation avancé par <i>SFS – DFSDBSCAN</i> . . .	181
5.42	Illustration des évolutions de la séprabilité des classes avec les trois méthodes de classification dynamique	184
5.43	Illustration des évolutions de la densité des classes avec les trois méthodes de classification dynamique	185
5.44	Illustration des évolutions de la surface des classes avec les trois méthodes de classification dynamique	186
5.45	Le banc expérimental et la butée à billes utilisée durant l'essai	188
5.46	L'évolution des indicateurs de défaut durant la période d'essai, moment de détection d'un défaut de 2,47mm est marqué en noir	190
5.47	L'évolution des indicateurs de défauts issus de l'analyse d'enveloppe . . .	191
5.48	Classification des observations reçues avant la détection et durant les premières phases de dégradation	193
5.49	La classification des observations par <i>DDBSCAN</i> après la détection . . .	194
5.50	La classification des observations reçues avant la détection et durant les premiers stades de dégradation	196
5.51	La classification des observations par <i>ESDBSCAN</i> après la détection de défaut	197

TABLE DES FIGURES

5.52	La classification des observations avant la détection de défaut et durant ses premières phases de dégradation	199
5.53	La classification des observations par <i>DFSDBSCAN</i> après la détection de défaut	200
5.54	Illustration des évolutions de la séprabilité des classes avec les trois méthodes de classification dynamique	203
5.55	Illustration des évolutions de la surface de la classe 'Défectueuse' trouvée avec les trois méthodes de classification dynamique pendant la période d'essai	204
5.56	Illustration des évolutions de la densité de la classe 'Défectueuse' trouvée avec les trois méthodes de classification dynamique pendant la période d'essai	205
A.1	les Imfs issues pour un signal vibratoire d'un roulement avec un défaut de 2 mm^2	216
A.2	Illustration des évolutions des indicateurs extraits directement du signal brut et à partir des trois premiers <i>IMFs</i>	217
A.3	Illustration des évolutions des indicateurs extraits directement de premier <i>IMF</i> pour des roulements sous différentes charges	218

Liste des tableaux

4.1	les indicateurs de défauts retenus	97
4.2	Tableau Comparatif des caractéristiques trouvées par les différentes méthodes d'extraction	103
4.3	Tableau Comparatif des classifieurs proposés	135
5.1	Tableau comparatif des résultats obtenus avec classifieurs proposés	183
5.2	Tableau comparatif des résultats obtenus avec classifieurs proposés	202

LISTE DES TABLEAUX

Liste des algorithmes

1	L'algorithme général des méthodes hiérarchiques	59
2	L'algorithme général de classification par partition	61
3	L'algorithme de la classification par <i>DBSCAN</i>	65
4	L'algorithme de <i>KPCA</i> version en-ligne	101
5	Un point représentatif	116
6	CreateRepList() : Création des nouveaux points représentatifs	117
7	UpdateExRepList() : Mise à jour des points représentatifs existants	118
8	Mise à jour de la liste <i>RepList</i> existante par des nouveaux points repré- sentatifs	119
9	UpdateRepList() :Mise à jour des points représentatifs en ligne	120
10	L'algorithme de la classification par FN-DBSCAN	126
11	L'algorithme de la classification par SFN-DBSCAN	127
12	UpdateExRepList() : Mise à jour des points représentatifs existants	130
13	Mise à jour de la liste <i>RepList</i> existante par des nouveaux points repré- sentatifs	131
14	Classification globale	132
15	Sifting process (SP)	213
16	L'algorithme de EMD	214

Introduction générale

De nos jours, les machines tournantes sont utilisées dans des domaines aussi variés que le transport (trains, véhicules motorisés, etc.), la production électrique (alternateurs), l'industrie de production, ou encore l'électroménager. Une machine tournante est le siège de forces dynamiques d'origine mécanique engendrées par les pièces en mouvement. Par exemple un déséquilibre d'un arbre provoque un balourd (excitation sinusoïdale) et un écaillage sur une piste de roulement à billes provoque un choc (excitation impulsionnelle) au passage de chaque bille sur le défaut. Ces forces provoquent des vibrations qui se propagent dans l'ensemble de la structure de la machine. Ces vibrations sont utilisées comme outil pour révéler l'état interne de la machine. Au fur et à mesure que son état se détériore (déséquilibre d'un arbre - défaut de roulement ou de boîte de vitesses...), le niveau de vibration augmente. Les propriétés statistiques ainsi que les différents paramètres du signal vibratoire changent. En mesurant et en surveillant ces paramètres, on obtient un ou des indicateur(s) de l'état des composants sensibles de la machine que l'on peut appeler "indicateurs de défauts".

L'analyse vibratoire repose sur le suivi d'indicateurs de défauts extraits des signaux vibratoires. Les indicateurs peuvent être temporels (valeur efficace, facteur crête, facteur d'impulsion, Kurtosis, Kurtogramme, etc.) ou fréquentiels (amplitude des fréquences caractéristiques de défaut, Kurtosis fréquentiel, centre de fréquence, etc.). 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. Pour des machines complexes les techniques récentes de traitement du signal permettent de mieux diagnostiquer les machines tournantes. Parmi ces méthodes on trouve la décomposition empirique modale (DEM), la décomposition en ondelettes, la cyclo stationnarité, les méthodes statistiques et les méthodes issues de l'intelligence artificielle. Ces

méthodes permettent d'isoler la signature vibratoire d'un composant et de mettre en évidence l'apparition d'un défaut en exploitant les propriétés temporelles et fréquentielles de la signature vibratoire de chaque composant. Néanmoins la présence du bruit dans les signaux vibratoires nécessite souvent des prétraitements comme le débruitage du signal avant la mise en œuvre de ces méthodes.

L'efficacité de toutes ces méthodes dépendent essentiellement des conditions de fonctionnement de la machine, de la position du capteur et aussi le bruit environnant. Par ailleurs les indicateurs de défauts d'un composant (roulement, moteur, boîte de vitesse, courroie, etc.) dépendent également de plusieurs facteurs comme la vitesse de rotation, la charge ou le montage/démontage de la machine tournante. Ces facteurs rendent le diagnostic plus difficile car l'évolution des indicateurs de défauts ne signifie pas nécessairement l'apparition ou l'évolution d'un défaut mécanique vers un état critique.

Pour remédier à cette problématique, des travaux de recherche se sont orientés vers des outils alternatifs et complémentaires qui peuvent être employés pour fiabiliser et automatiser le processus du diagnostic. Parmi ces méthodes on peut citer les techniques de reconnaissance des formes. La reconnaissance des formes (RdF) est un ensemble de techniques et méthodes qui a pour objectif de détecter et caractériser un objet qui peut être un caractère, un comportement ou un changement puis l'associer à une action. Dans le diagnostic et le suivi des défauts dans les machines tournantes, les méthodes de reconnaissance des formes peuvent être employées pour identifier la présence de défaut et caractériser sa sévérité en analysant des paramètres (indicateurs de défauts) extraits du signal vibratoire, en étudiant leur changement dans le temps par rapport à une base de référence ou par rapport aux informations extraites des enregistrements précédentes afin de décider leur appartenance à une classe (état de santé). Ces méthodes apportent une complémentarité aux méthodes existantes car elles permettent : *(i)* d'automatiser la procédure de diagnostic et de suivi, *(ii)* d'extraire des informations supplémentaires cachées dans la masse de données (comme par exemple la tendance de la variation des indicateurs dans le temps, la variation de la vitesse de rotation, la variation des charges, etc.), *(iii)* de faciliter l'accès aux informations, les méthodes de RdF représentent les données d'une manière synthétique et facile à comprendre même pour des personnes non averties, *(iv)* d'être configurées pour pouvoir être intégrées dans un système de décision en temps réel. Néanmoins les méthodes de RdF utilisées dans le domaine de diagnostic et du suivi ont été déployées en vue de l'automatisation du diagnostic à un instant t donné.

L'introduction du paramètre temps dans l'analyse permettrait d'étudier les variations des indicateurs dans le temps et d'ajouter une nouvelle dimension à l'étude des roulements, à l'instar de la météorologie.

C'est pourquoi mes travaux de recherche s'orientent vers la méthode de reconnaissance des formes dans un contexte dynamique pour suivre les différents changements d'état de fonctionnement du système. Ces changements peuvent être dus à l'apparition d'un défaut, à un changement de vitesse de rotation ou à une variation de charge appliquée. Cette méthode devrait être capable d'identifier l'origine du changement observé sur les signaux vibratoires, elle doit permettre de différencier l'évolution d'un défaut à l'évolution de la vitesse de rotation ou à l'évolution de la charge. La méthode doit être intégrée en ligne dans le processus de diagnostic, c'est à dire, elle doit être capable d'analyser chaque signal arrivé et de décider l'état interne du composant mécanique et plus particulièrement le roulement, sans dépendre des connaissances préalables du système. En d'autre terme elle doit être autonome (stand-alone) et ses résultats doivent être facile à interpréter.

Pour atteindre ces objectifs, le chapitre 1 réalise une étude bibliographique sur le domaine de la maintenance conditionnelle. Après une description des différentes stratégies de maintenance et des principaux défauts de roulements, le chapitre 1 présente les méthodes de diagnostic, les différentes techniques du traitement du signal employées, les approches statistiques et les approches issues de l'intelligence artificielle. Cette étude bibliographique montre les avantages et les limites des méthodes classiques face aux différentes grandeurs comme la variation de charge ou la vitesse de rotation d'où l'intérêt d'employer d'autres outils comme les méthodes de reconnaissance des formes.

Le chapitre 2 consiste à étudier la structure générale des méthodes de reconnaissance des formes, à détailler ses différentes composantes : 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 le cœur de notre méthode. Un état de l'art présente les différents type de classifications, les approches de partitionnement, le métrique de similarité, les différentes techniques d'apprentissage. Par ailleurs, une revue bibliographique sur les méthodes de classification employées dans

le diagnostic des machines tournantes est réalisée. Enfin nous introduirons, dans ce chapitre, la limitation des méthodes statiques souvent utilisées dans le diagnostic et l'apport de la classification dynamique par rapport la classification statique.

Le chapitre 3 a pour objectif d'expliquer le processus de conception d'une méthode de reconnaissance des formes dynamique développée pour le diagnostic des roulements, d'énumérer les contraintes à respecter, d'étudier les choix de méthodes qu'on peut employer pour chaque composante de la méthode de RdF.

Le Chapitre 4 consiste à introduire l'architecture de la méthode de reconnaissance de formes développée pour le suivi de roulements et de présenter les méthodes de classification dynamiques développées durant cette thèse pour répondre aux exigences du suivi.

La validation de la méthodologie déployée dans le chapitre 4 passe par une application sur des signaux réels issus de roulements avec différentes états de roulements et sous différentes conditions de fonctionnement (évolution de l'endommagement, changement de vitesse, changement de charge) afin de maîtriser les paramètres influents. Le chapitre 5 est consacré à étudier le comportement vibratoire du roulement avec une charge qui évolue dans le temps, un défaut qui apparait et les deux qui varient en même temps pour pouvoir finalement trouver des caractéristiques qui peuvent nous aider à identifier automatiquement chaque situation.

Ce chapitre est dédié également à la validation expérimentale de la méthode proposée. Nous présentons d'une part le banc expérimental, la structure des données enregistrées, l'application pas à pas de la méthode et son paramétrage et d'autre part les résultats obtenus en fonction des différents paramètres de réglage (vitesses de rotation, charges appliquées). Une étude comparative entre les résultats obtenus par les différentes méthodes de classification a permis de différencier l'évolution des différentes classes en fonction de ces paramètres.

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

Chapitre 1

Traitement du signal pour la maintenance

Introduction

De nombreuses défaillances peuvent apparaître sur les machines tournantes provoquant des arrêts intempestifs. Le premier pas vers la réduction des arrêts non planifiés réside dans la mise en œuvre d'une bonne stratégie de maintenance. Cela nécessite l'utilisation d'une méthode de diagnostic efficace afin d'établir un système de surveillance capable d'assurer la disponibilité, la fiabilité et la sûreté de fonctionnement des machines. Ce chapitre présente d'abord les différentes stratégies de maintenance avec leurs avantages et leurs inconvénients. Nous présentons ensuite les techniques de diagnostic des machines tournantes par analyse vibratoire en passant par les principaux défauts de roulements susceptibles d'apparaître. En effet l'analyse vibratoire apparaît comme l'un des outils les plus utilisés et les plus efficaces pour détecter, diagnostiquer et suivre les défauts des composants mécaniques dans une machine tournante. Finalement nous présentons les différentes techniques de traitement du signal les plus utilisées en analyse vibratoire.

1.1 Stratégies de maintenance

1.1.1 Définition

Les machines tournantes, comme tout équipement industriel, tendent à se détériorer dans le temps. Cette détérioration peut être provoquée par de multiples causes dues au fonctionnement (usure, déformations, etc.) ou aux autres agents corrosifs (agents chimiques, atmosphériques, etc.). La défaillance d'un ou de plusieurs composants mécaniques d'une machine peut avoir des conséquences graves en termes de coût de production (arrêt de fonctionnement, diminution des capacités de production), de coût de réparation et peut mettre en péril la sécurité des personnes. Elle peut également entraîner d'autres dommages sévères sur d'autres composants de la machine ou de l'installation. Pour remédier à ce problème, des stratégies de maintenance sont employées. La notion de maintenance a été définie par l'AFNOR en 1994 (norme NFX 60-010), comme « l'ensemble des actions permettant de maintenir ou de rétablir un bien dans un état spécifié ou en mesure d'assurer un service déterminé », puis légèrement modifiée en 2011 (NF EN 13306 X 60-319) par « Ensemble de toutes les actions techniques, administratives et de management durant le cycle de vie d'un bien, destinées à le maintenir ou à le rétablir dans un état dans

lequel il peut accomplir la fonction requise. ». Ces deux définitions exposent la relation qui existe entre la notion de maintenance définie par la norme et le domaine industriel.

1.1.2 Rôle de la maintenance

Le rôle de la maintenance ne se résume pas seulement à la simple action de réparation des équipements dégradés mais aussi à la prévention de la dégradation, en garantissant que les machines fonctionnent efficacement, en toute sécurité et de façon fiable (AJ12, Gan13). Les objectifs de la maintenance peuvent être exprimés en termes quantitatifs ou qualitatifs :

- Atteindre une productivité maximale en :
 - assurant un fonctionnement continu et satisfaisant de la machine tout au long de sa durée de vie prévue - ou même plus.
 - parvenant à une utilisation maximale de la machine avec un minimum de temps d'arrêt et de réparations.
 - améliorant continuellement le processus de production.
- Optimiser les performances de la machine.
- Assurer la sécurité de fonctionnement.

1.1.3 Types de maintenance

Il existe deux stratégies de maintenance : corrective ou préventive. Chacune d'elles correspond à un concept particulier et répondant à un besoin particulier (voir figure 1.1).

A. Maintenance corrective

La maintenance curative ou corrective est certainement la plus ancienne et la plus traditionnelle des techniques de maintenance. Son principe repose sur l'exécution d'une action de maintenance après la détection d'une défaillance ou après l'arrêt définitif de la machine. Elle peut également entraîner des dommages sévères sur d'autres composants de la machine ou de l'installation. La maintenance corrective bien qu'elle est la plus ancienne n'est pas pour autant celle qui est la moins coûteuse, d'abord parce que, pour une même intervention elle peut forcer à engager des moyens importants justifiés par la criticité de

1. TRAITEMENT DU SIGNAL POUR LA MAINTENANCE

la défaillance, d'autre part parce que l'interruption non programmée du service ou de la production peut avoir des conséquences préjudiciables pour l'entreprise (norme FD X 60-000).

La maintenance corrective est toujours utilisée pour des équipements qui ne présentent pas un intérêt majeur dans le système et qui peuvent être facilement remplacés. Par exemple dans les industries où la perte d'une machine pour un court laps de temps n'est pas critique pour la production, et où le dysfonctionnement ne peut pas générer des conséquences catastrophiques (Lar12).

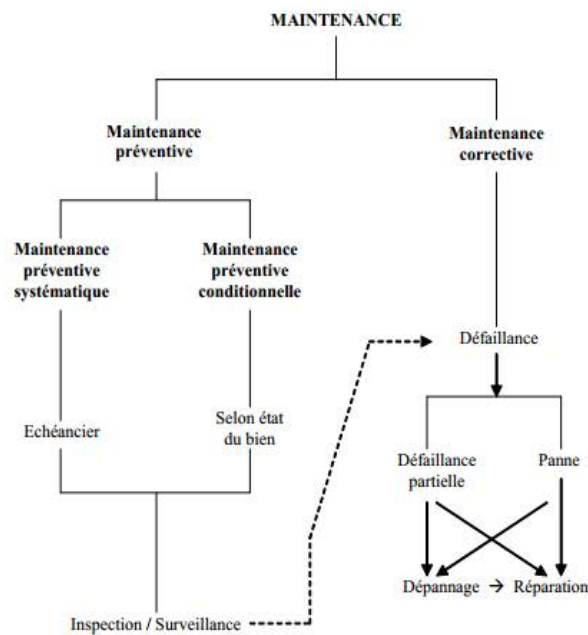


FIGURE 1.1 – Les stratégies de maintenance

B. Maintenance préventive

La maintenance préventive est une maintenance effectuée dans l'intention de réduire la probabilité de défaillance d'une machine ou d'un composant. Elle correspond à une attitude proactive : «on agit avant la défaillance ». La maintenance préventive consiste à établir un échéancier, basé sur le temps de bon fonctionnement des composants (Mon03, AFN02). L'avantage de cette méthode est que la plupart des arrêts pour la maintenance sont prévus à l'avance ce qui réduit la possibilité d'avoir une défaillance catastrophique.

On distingue, la maintenance préventive systématique, et la maintenance préventive conditionnelle.

Maintenance préventive systématique

La maintenance préventive systématique peut être utile dans le cas où l'on peut prédire avec une précision raisonnable la durée de vie d'un composant. Cependant, la durée de vie définie par le constructeur peut ne pas correspondre à la durée de vie réelle des composants ce qui peut se traduire par une usure prématurée.

Maintenance préventive conditionnelle

La démarche de la maintenance conditionnelle consiste à extraire des paramètres choisis, caractéristiques des informations relatives de l'état de fonctionnement du (ou des) composant(s) constitutif(s) d'une machine de production et à suivre dans le temps l'évolution de ces indicateurs pour prévoir le remplacement du (ou des) composant(s). Le suivi de ces indicateurs permet d'une part d'augmenter la durée de fonctionnement du composant et d'autre part d'estimer la durée de vie résiduelle afin de programmer son remplacement (HON14), (JWM08). Ce type de maintenance a des avantages évidents par rapport aux deux autres types comme l'augmentation de la productivité, la diminution des coûts de fonctionnement (limitation du nombre d'interventions sur les machines de production et du nombre de stock de composants de rechange). Ainsi la remise en état du composant est réalisée uniquement lorsque celui-ci présente des signes de dysfonctionnement (dégradation, symptômes, panne à la sollicitation) pouvant mettre en cause ses performances (HON14).

La maintenance conditionnelle diffère de la maintenance préventive systématique non seulement par le fait qu'elle est basée sur l'état actuel de la machine et non sur un échéancier défini préalablement mais aussi par l'intégration de la notion de pronostic. La maintenance conditionnelle permet donc de prédire la panne par l'utilisation des techniques de surveillance de l'état de dégradation des composants. Ces techniques permettent de déterminer l'état d'un (ou plusieurs) composant(s) à un instant t et de prévoir la durée de vie utile restante (JWM08).

1.1.4 Choix du type de maintenance

Les paramètres dont il faut tenir compte pour adopter telle ou telle politique de maintenance sont principalement d'ordre économique et humain. La maintenance corrective peut être une alternative raisonnable quand une panne est rare et l'occurrence d'une défaillance ne peut pas affecter la sécurité. Quand les pannes sont plus fréquentes ou la conséquence d'une défaillance est plus coûteuse, la maintenance préventive est la plus adaptée. En revanche, la politique de maintenance qui présente le meilleur compromis (coût/efficacité) reste la maintenance conditionnelle car elle permet d'éviter les tâches de maintenance inutiles en prenant des mesures d'intervention seulement quand il y a des preuves de comportements anormaux d'un composant. La maintenance conditionnelle, si elle est bien établie et efficacement mise en œuvre, permet de réduire considérablement les coûts de maintenance en réduisant le nombre d'opérations inutiles de maintenance préventive programmées.

La maintenance conditionnelle a deux aspects importants : le diagnostic et le pronostic. Le diagnostic a pour objectif de détecter, localiser, et identifier un défaut quand il se produit. La détection de défaut indique le dysfonctionnement du système suivi par la localisation et l'identification de la nature du défaut détecté. Le pronostic permet la prédiction de panne avant qu'elle se produise. La prédiction des défauts permet de déterminer si une panne est imminente et d'estimer combien de temps et quelle est la probabilité qu'une panne aura lieu.

Parmi les techniques utilisées en maintenance conditionnelle, l'analyse vibratoire reste l'outil le plus utilisé pour détecter et suivre la dégradation des composants mécaniques d'une machine tournante et plus particulièrement les roulements.

1.2 Diagnostic des machines tournantes par analyse vibratoire

A partir des signaux vibratoires régulièrement recueillis sur une machine tournante, on peut détecter des défauts, identifier leurs origines, surveiller en continu l'état de la machine et prédire la durée de vie restante des composants mécaniques d'une machine.

1.2.1 Présentation générale de l'analyse vibratoire

Toutes les machines en fonctionnement produisent des vibrations, images des efforts dynamiques engendrés par les pièces en mouvement. Ainsi, une machine neuve en excellent état de fonctionnement produit très peu de vibrations. La détérioration du fonctionnement conduit le plus souvent à un accroissement du niveau des vibrations. En observant l'évolution de ce niveau, il est par conséquent possible d'obtenir des informations très utiles sur l'état de la machine. La modification du niveau vibratoire d'une machine constitue souvent la première manifestation physique d'une anomalie, cause potentielle de dégradation, voire de pannes. Ces caractéristiques font de la surveillance par analyse des vibrations, un outil indispensable pour une maintenance moderne, puisqu'elle permet, par un dépistage ou un diagnostic approprié des défauts, d'éviter la casse et de n'intervenir sur une machine qu'au bon moment et pendant des arrêts programmés de production.

L'analyse vibratoire est de loin la méthode la plus répandue pour la surveillance des machines car elle offre plusieurs avantages par rapport aux autres méthodes. Elle réagit immédiatement au changement de l'état de fonctionnement de la machine, et elle peut donc être utilisée pour la surveillance permanente ou intermittente. De plus, de nombreuses techniques puissantes de traitement du signal peuvent être appliquées aux signaux de vibration pour extraire des paramètres significatifs de l'état de fonctionnement des composants mécaniques même à un niveau faible (Mor91).

Les vibrations véhiculent beaucoup d'informations sur l'état de fonctionnement de la machine, toutes les caractéristiques des vibrations (amplitude, répartition des amplitudes, fréquence, ...) peuvent être exploitées et analysées pour révéler des détails précis sur l'état interne de la machine. L'analyse vibratoire est aujourd'hui très fortement répandue dans l'industrie en y trouvant sa place au sein des stratégies de maintenance conditionnelle car elle offre une panoplie des méthodes qui permettent de détecter la présence des défauts (diagnostic) et d'analyser leur évolution temporelle (la surveillance). La maintenance par l'analyse vibratoire se compose de trois étapes clés :

- L'étude cinématique de la machine : calcul des fréquences caractéristiques de défauts des composants à surveiller.
- L'acquisition et le traitement des données vibratoires : extraction des paramètres significatifs de l'état de dégradation des composants ;

- La décision : Diagnostic et pronostic de pannes.

1.2.2 Les différents niveaux d'analyse vibratoire

On note trois niveaux d'analyse ; la surveillance, le diagnostic et le suivi et pronostic.

La surveillance est considérée comme le premier niveau d'analyse vibratoire. Elle consiste à surveiller l'état de fonctionnement général des machines à travers des indicateurs, le plus souvent globaux. Ces indicateurs sont extraits à intervalles de temps réguliers puis comparés aux valeurs précédentes par le biais de la courbe de tendance. Un seuil peut être donné par le fabricant ou choisi en étudiant l'historique de la machine. Si un indicateur dépasse le seuil admissible, soit des actions correctives sont effectuées soit on passe au niveau supérieur d'analyse : le diagnostic.

Le diagnostic consiste à mettre en évidence la défaillance des composants et d'identifier la source de cette défaillance par des techniques de traitement du signal. De nombreuses techniques existent pour extraire des informations significatives de l'état de dégradation des composants (Ceb12)(Ibr09)(Bod04)(Kho09)(Des10). Ces informations seront utilisées pour le suivi dans le temps de l'évolution de la sévérité de la défaillance (suivi et pronostic).

L'étape de suivi et de pronostic constitue le niveau le plus élevé dans l'analyse vibratoire et son objectif est la qualification, la quantification et la prédiction en termes de fiabilité(Chi07). Dans ce niveau on s'intéresse soit à l'estimation de la probabilité qu'une défaillance survienne à un instant donné, soit à la prédiction du temps résiduel avant défaillance communément appelé RUL (Remaining Useful Life) (JLB06). Cela permet de planifier un arrêt programmé de la machine pour le remplacement du (ou des) composant(s).

Il existe plusieurs techniques pour estimer la durée de vie résiduelle d'un composant. Le taux de déviation des indicateurs de défauts peut être utilisé pour estimer une durée résiduelle approximative (Kou10). Dans la littérature on rencontre trois types de pronostic de pannes ; le premier est basé sur un modèle physique (MMZ08), le second est guidé par les données caractéristiques de l'état de dégradation du composant et le troisième type est basé sur les données de retour d'expérience (CMGT02).

1.3 Les principaux défauts de roulements

On ne peut pas surveiller correctement une machine que l'on ne connaît pas, d'où la nécessité de comprendre d'abord la nature de la machine, de connaître la cinématique de l'installation à surveiller et d'analyser les origines de changement de comportements vibratoires de la machine (Aug14). Généralement, les forces à l'origine des vibrations des machines tournantes en fonctionnement dépendent :

- de l'état mécanique de la machines (dissymétrie des masses par exemple)
- des paramètres de fonctionnement (température, vitesse, charge, etc.).

L'expérience acquise sur les machines tournantes a conduit à un répertoire de dysfonctionnements des composants mécaniques. On peut citer le déséquilibre massique des rotors, la défaut de lignage d'arbre, le balourd, etc. Les causes liées à chaque type de défaut peuvent être multiples et leurs manifestations peuvent être différentes. Dans cette section nous citons que les défauts concernant les roulements.

1.3.1 Définition de défaut

Tous les matériaux et par conséquent toutes les structures présentent des anomalies à l'échelle nano/microstructurale : la difficulté est de décider quand une structure est considérée comme « endommagée ». En raison de changement de composition et des procédés utilisés au niveau de la fabrication, les composants mécaniques peuvent être légèrement différents au niveau micro-structurel et peuvent contenir des inclusions aux nombre ou forme différentes, des creux et d'autres défauts. Ce type d' « anomalie » ne peut pas être considéré comme un endommagement du composant mécanique. Dans cette thèse on s'intéresse aux endommagements qui apparaissent d'une manière naturelle dans une machine en fonctionnement. Ce type d'endommagement commence par des microfissures qui évoluent puis finissent par des modifications des propriétés du matériau (Gan13). Dans (WDB04), des définitions logiques de défaut, endommagement et anomalie sont établis :

Défaut

Un défaut est défini comme changement de l'état de la structure, produisant une réduction inacceptable de qualité. Ainsi elle ne peut plus fonctionner d'une manière

1. TRAITEMENT DU SIGNAL POUR LA MAINTENANCE

satisfaisante. La qualité d'une structure ou d'un système peut être définie comme l'aptitude ou la capacité du système à répondre à des besoins des utilisateurs ou des clients selon les critères définis par ces derniers.

Endommagement

Une structure endommagée ne fonctionne plus en son état idéal mais peut toujours fonctionner d'une manière satisfaisante.

Anomalie

Une anomalie est inhérente au matériel et statistiquement tous les matériaux contiennent une certaine quantité inconnue d'anomalies. La présence d'anomalie signifie que la structure peut fonctionner à son état de conception même si les matériaux que lui sont constitués contiennent des anomalies.

Des relations hiérarchiques peuvent être développées à partir des définitions ci-dessus : les anomalies mènent à des endommagements et les endommagements mènent aux défauts. Ainsi, la détection de défaut signifiera réellement la détection des endommagements qui mèneraient à un défaut.

1.3.2 Défauts des roulements

Les roulements sont parmi les composants les plus sollicités des machines et représentent une source de panne fréquente. De multiples études statistiques sur des machines ont été effectuées depuis les années 80 jusqu'à présent. Une étude statistique récente faite par Bonnet (BY08)(voir figures 1.2, 1.3) sur les machines asynchrones de grande puissance, exploitées dans l'industrie, montre que 69% de pannes se situent sur les roulements, un pourcentage qui a connu une grande augmentation (41% en 2001 par Thomson) (TM01). Cette répartition confirme bien que les défauts des machines proviennent principalement de roulements.

La dégradation du roulement peut être le résultat de 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 (Chi07) :

1.3 Les principaux défauts de roulements

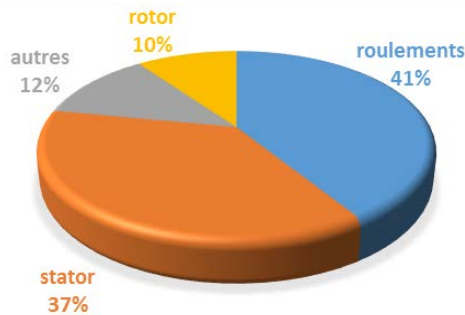


FIGURE 1.2 – Thomson 2001 (TM01)

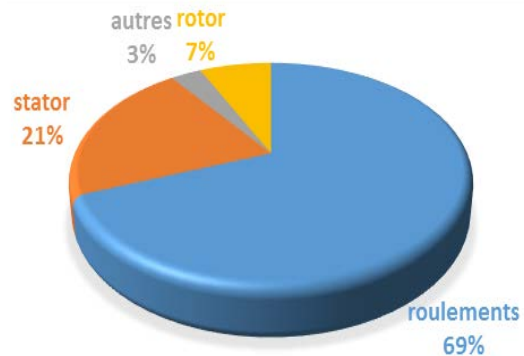


FIGURE 1.3 – Bonnet 2008(BY08)

- 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.

Tous ces défauts ont un point commun : ils se traduisent tôt ou tard par une perte de fragments de métal. Ce défaut précurseur de la destruction, et ce quel que soit les conditions d'utilisation et de fonctionnement, est l'écaillage (voir 1.4). Ce défaut survient



FIGURE 1.4 – Ecaillage

sous l'effet de la fatigue due aux contraintes de cisaillement alternées qui sévissent en sous-couche.

Les avaries les plus fréquentes sur les roulements sont les défauts d'écaillage. L'écaillage est un processus continu qui s'accélère plus ou moins après l'apparition des premières fissures. Lors de la mise en rotation, un train d'impulsion est généré par ces défauts, à une fréquence bien définie que l'on appelle « fréquence caractéristique » de défaut du roulement. Ce signal périodique est l'origine de nombreuses méthodes de détection de défaut de roulement (TC99). Les fréquences caractéristiques sont déterminées à partir de la géométrie du roulement (voir figure 1.5) et de la cinématique de la machine étudiée. Elles sont données par les équations suivantes :

- Défaut bague intérieure, Ball Pass Frequency Inner race (BPFI)

$$f_I = \frac{N}{2} \left[1 - \frac{D}{d} \cos(\theta) \right] f_r \quad (1.1)$$

- Défaut bague extérieure, Ball Pass Frequency Outer race (BPFO)

$$f_E = \frac{N}{2} \left[1 + \frac{D}{d} \cos(\theta) \right] f_r \quad (1.2)$$

- Défaut de bille, Ball Spin Frequency (BSF)

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

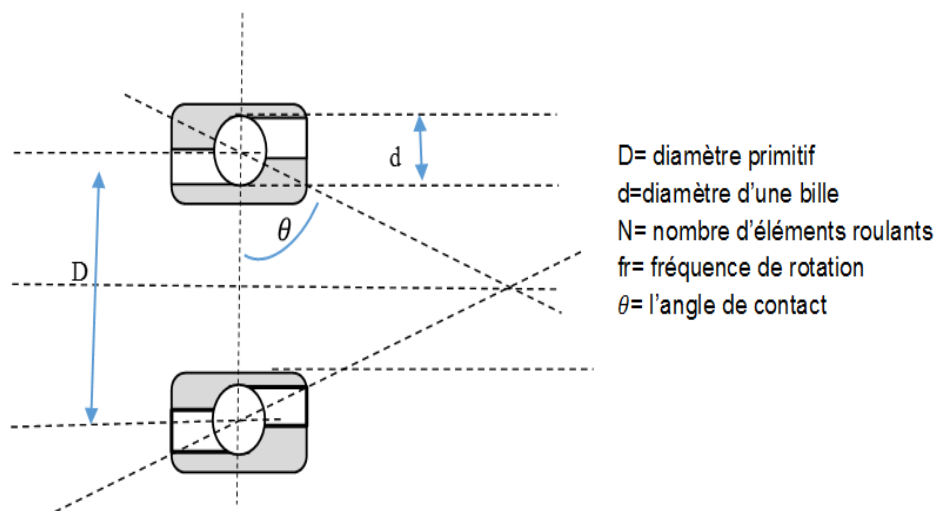


FIGURE 1.5 – Géométrie d'un roulement (Mor05)

- Défaut de cage

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

où f_r est la vitesse de rotation, N est le nombre d'éléments roulants, θ est l'angle de contact, d et D sont respectivement le diamètre de l'élément roulant et le diamètre primitif.

Dans cette thèse, on s'est focalisé sur l'étude de diagnostic et de suivi des roulements. Les méthodes citées dans les sections à venir sont axées sur les défauts de roulements.

1.4 Les techniques de traitement du signal appliquées à l'analyse vibratoire

La détection de défauts nécessite d'une part une prise de mesure du signal vibratoire et d'autre part une exploitation du signal recueilli. La prise de mesure est une étape essentielle dans la procédure d'analyse vibratoire car elle conditionne l'efficacité des méthodes de traitement mises en œuvre. Le choix du ou des point(s) de mesure, le mode de fixation des capteurs, les paramètres de l'appareil de mesure (gamme de fréquence, résolution spectrale) sont des paramètres à prendre en considération (Est04).

1.4.1 Analyse temporelle

L'analyse temporelle est basée sur l'analyse des "indicateurs de défauts" associés à un signal vibratoire enregistré. Elle consiste à étudier le comportement vibratoire de la machine à partir de ces indicateurs. Un indicateur temporel est une grandeur qui caractérise la puissance, l'amplitude ou la répartition des amplitudes du signal vibratoire. L'évolution de ces indicateurs est significative de l'apparition d'un défaut et donc de son aggravation. Ces indicateurs évaluent l'état de fonctionnement global des équipements mais ne localisent pas le défaut. De nombreux indicateurs existent dans la littérature et certains sont le résultat de la combinaison de plusieurs d'entre eux.

La valeur efficace

La valeur efficace ou la valeur *RMS* (Root Mean Square), noté x_{RMS} est l'indicateur scalaire le plus couramment utilisé, il permet de mesurer l'énergie moyenne du signal, il est utilisé pour détecter des dissipations d'énergie anormalement élevées accompagnant la naissance d'un défaut. La *RMS* est la racine carrée de la moyenne quadratique du signal vibratoire temporel discrétisé $x(n)$ de longueur N et de moyenne empirique $\bar{x}$.

$$x_{RMS} = \sqrt{\frac{\sum_{i=1}^N (x(i)^2)}{N}} \quad (1.5)$$

Malheureusement la valeur efficace n'évolue pas de manière significative au cours de la première phase de dégradation, il ne commence à croître que pendant la deuxième phase de dégradation (Fra93). Ceci rend la détection précoce impossible et représente un inconvénient majeur dans la maintenance prédictive. De plus l'augmentation du niveau de bruit peut entraîner une mauvaise interprétation de la valeur *RMS*.

Facteur crête

Contrairement à la valeur efficace, les indicateurs spécifiques comme le facteur crête est mieux adapté pour représenter un signal induit par des forces impulsionnelles telles que les écaillages de roulements. Le facteur crête est défini comme étant le rapport entre la valeur crête, noté x_{peak} , et la valeur efficace.

$$x_{Fc} = \frac{x_{peak}}{x_{RMS}} \quad (1.6)$$

1.4 Les techniques de traitement du signal appliquées à l'analyse vibratoire

Le facteur crête a l'avantage de détecter le défaut dès son apparition et donne une information très précoce de la prédiction. Ceci provient du fait que pour un roulement sans défaut, le rapport reste sensiblement constant et augmente lorsqu'un début d'écaillage apparaît. Ceci est dû à la présence des chocs dans le signal vibratoire.

Kurtosis

Le kurtosis est un indicateur qui permet de caractériser le degré d'aplatissement d'une distribution ce qui permet la détection précoce d'un défaut de roulement. Dans le cas d'un roulement sans écaillage, la distribution des amplitudes du signal recueilli est gaussienne ce qui entraîne une valeur de kurtosis proche de 3. Lorsqu'un défaut apparaît, sa valeur devient supérieure à 3.

La détection de défaut de roulement par le kurtosis peut être réalisée dans différentes bande de fréquences liées aux résonnances de la structure.

$$x_{Ku} = \frac{1/N \sum_{i=1}^N [x(i) - \bar{x}]^4}{\sigma^4} \quad (1.7)$$

Avec σ est l'écart type.

Il est à noter que dans le cas d'une forte détérioration du roulement, l'allure de la distribution de l'amplitude redevient gaussienne avec x_{Ku} voisin de 3 et la valeur *RMS* augmente sensiblement. Comme pour le facteur crête, il y a lieu de tenir compte simultanément de l'évolution des deux indicateurs : Kurtosis et valeur *RMS* (Mor05).

Le facteur K

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

$$x_K = x_{peak} \times x_{RMS} = x_{peak} \times \sqrt{\frac{1}{N} \sum_{i=1}^N x(i)^2} \quad (1.8)$$

L'interprétation du facteur crête 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.

Skewness

Le coefficient de dissymétrie (skewness) correspond à une mesure de l'asymétrie de la distribution d'une variable aléatoire réelle.

$$x_{SK} = \frac{\sum_{i=1}^N [x(i) - \bar{x}]^3}{(N-1)\sigma^3} \quad (1.9)$$

Les indicateurs présentés ici ne forment pas une liste exhaustive de tous les indicateurs utilisés dans l'industrie ou dans la littérature scientifique. Cependant ils présentent les indicateurs les plus utilisés. L'ensemble de ces indicateurs est très facile à mettre en œuvre. Ces indicateurs sont issus soit de l'étude de la distribution de la densité de probabilité comme le kurtosis ou le skewness soit de l'analyse statistique du signal tels que la valeur efficace (x_{RMS}), la valeur maximale (x_{peak})... pour ne citer que les plus utilisés. Le plus souvent, ils sont traités simultanément et calculés sur plusieurs bandes de fréquences pour fiabiliser la détection. Il existe également des indicateurs hybrides qui sont une combinaison des indicateurs classiques (par exemple le facteur crête (x_{Fc})) (1.4.1).

Ces indicateurs sont utilisés pour indiquer une modification du comportement vibratoire de la machine mais ne permettent pas la localisation de l'élément qui modifie ce comportement. La fiabilité de ces indicateurs pour diagnostiquer un défaut dépend de plusieurs facteurs comme la complexité de la machine, la charge (DP05), la vitesse de rotation (AH11, KTMY06). Plusieurs études ont été effectuées pour trouver le ou les indicateurs adaptés pour détecter fiablement le défaut même en condition non stationnaires. Comme, Li et Pickering ont montré que le kurtosis et le facteur crête sont les indicateurs les plus sensibles à la présence de défaut d'écaillage localisé et les plus insensible aux changements de condition de fonctionnement (LP92). Par ailleurs, au fur et à mesure que la sévérité du défaut augmente, le signal vibratoire devient de plus en plus aléatoire, ce qui influence les valeurs de certains indicateurs de défaut comme le facteur de crête et le kurtosis. En effet ces valeurs diminuent pour atteindre des niveaux associés à un roulement sain. Néanmoins, les performances de certains indicateurs peuvent être améliorées en normalisant ces indicateurs par des valeurs de références (Par exemples les valeurs des indicateurs pour un roulement sain et sous condition normal(FQG00)). Sun *et al* ont proposé la normalisation de la fonction de densité de probabilité (FDP) du signal vibratoire (SCZX04) pour rendre les indicateurs insensibles aux variation de

1.4 Les techniques de traitement du signal appliquées à l'analyse vibratoire

vitesse et de charges. En effet, la FDP normalisée ne varie pas avec les changements de la vitesse ou de la charge mais celle ci est sensible lorsque l'état du composant se détériore.

1.4.2 Analyse spectrale

L'analyse en fréquence est devenue l'outil fondamental pour le traitement des signaux vibratoires. Elle s'appuie sur la transformée de Fourier du signal vibratoire. Cet outil permet de connaître le contenu spectral du signal et de localiser les fréquences caractéristiques de défauts des composants mécaniques. La transformée de Fourier est définie par l'équation suivante :

$$X(f) = \int_{-\infty}^{+\infty} x(t).e^{-j2\pi ft} dt \quad (1.10)$$

$X(f)$ est la transformée de Fourier, t est la variable temps et f est la variable de fréquence.

L'analyse spectrale des vibrations émises par la machine peut être un outil révélateur de présence des défauts d'origines mécaniques ou électriques. L'analyse spectrale permet de localiser le défaut sur un composant tournant (roulement engrenage, moteur, transmission par courroie, etc.). En effet, chaque élément tournant est caractérisé par une ou plusieurs fréquences caractéristiques de défaut. Ces fréquences fondamentales dépendent de la géométrie du composant et de sa vitesse de rotation (voir chapitre 1 section 1.4.2). Cependant les fréquences de défauts calculées ne correspondent pas souvent aux fréquences qui apparaissent dans le spectre de vibration. Ce qui peut être attribuable à une augmentation de charges au dessus du normal provoquant le roulement à fonctionner à un angle de contact différent ou à une variation de la vitesse de rotation.

On peut également extraire d'autres caractéristiques statistiques à partir du spectre de puissance. Parmi les caractéristiques statistiques qu'on peut calculer dans le domaine fréquentiel on trouve le centre de fréquence X_{fc} qui est un indicateur permettant d'indiquer la concentration des fréquences dans le spectre, la RMS fréquentielle X_{rmsf} (YZYX08), la valeur efficace du spectre de puissance du signal $X_{f_{rms}}$ et l'écart-type des fréquences X_{stdf} (YHB13) :

$$X_{fc} = \frac{\sum_{j=1}^K f_j(S(f_j))}{\sum_{j=1}^K S(f_j)} \quad (1.11)$$

$$X_{rmsf} = \sqrt{\frac{\sum_{j=1}^K f_j^2(S(f_j))}{\sum_{j=1}^K S(f_j)}} \quad (1.12)$$

$$X_{frms} = \frac{\sum_{j=1}^K (S(f_j))}{K} \quad (1.13)$$

$$X_{stdf} = \sqrt{\frac{\sum_{j=1}^K (f_j - X_{fc})^2(S(f_j))}{\sum_{j=1}^K S(f_j)}} \quad (1.14)$$

Avec $S(j)$ est le spectre de puissance du signal $x(n)$, pour $j = 1, 2, \dots, K$. K est le nombre de lignes spectrales et f_j est la fréquence du $j^{\text{ème}}$ ligne.

L'inconvénient de l'analyse spectrale classique est son caractère global. L'aspect temporel du signal disparaît. En effet, la transformée de Fourier suppose que les signaux sont stationnaires, elle ne fournit donc pas d'informations sur l'évolution du spectre du signal en fonction du temps (RU01). De plus les signaux issus des machines tournantes sont typiquement non-stationnaires.

1.4.3 Analyse d'enveloppe

L'analyse d'enveloppe est particulièrement appréciée et largement utilisée pour la localisation des défauts de roulements. Cette méthode utilise la modulation de l'amplitude de la fréquence de résonance du roulement par la fréquence caractéristique de défaut et plus particulièrement par le défaut de la bague extérieure. En effet le choc généré par l'élément roulant sur le défaut excite la structure sur toutes les fréquences et donc la fréquence de résonance de la structure. A chaque fois que ce phénomène se produit, il génère une vibration à la fréquence de résonance. 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 (R.B11).

La technique de détection d'enveloppe (HFRT, High frequency resonance technique) est basée sur la transformée de Hilbert et se décompose en plusieurs étapes (EVS⁺98).

1.4 Les techniques de traitement du signal appliquées à l'analyse vibratoire

D'abord, on réalise un filtrage passe-bande du signal $x(t)$ autour d'une fréquence particulière (en générale la fréquence de résonance), ensuite, on calcule le transformée de Hilbert du signal filtré pour isoler l'enveloppe du signal modulé en amplitude. La transformée de Fourier de l'enveloppe permet ainsi de retrouver la fréquence caractéristique de défaut et les harmoniques de cette fréquence. 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 (Ran01).

1.4.4 Analyse cepstrale

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 :

$$C(\tau) = TF^{-1} \left(\ln(|S(f)|^2) \right) \quad (1.15)$$

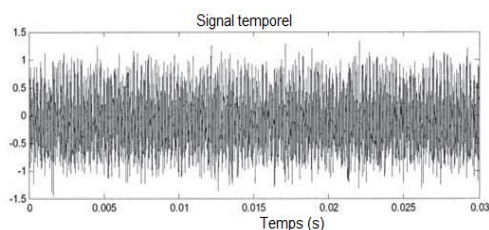


FIGURE 1.6 – Signal temporel d'émission acoustique issu d'un système d'engrenage

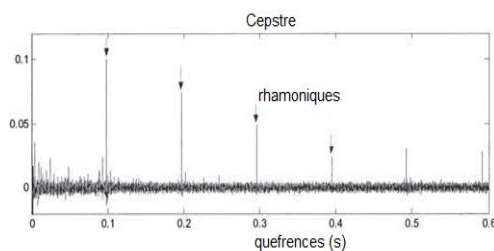


FIGURE 1.7 – la récurrence d'un défaut à 10Hz (roue primaire) est traduite par un peigne de Dirac de fondamental à 100 ms

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 "rharmique" et leur position sur l'axe des abscisses s'appelle le "quefrency". 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 (Bad99).

1.4.5 Analyse temps-fréquence

La présence des chocs dus aux défauts de roulements donne au signal vibratoire un caractère non stationnaire, ce qui interdit en principe l'utilisation de la transformée de Fourier qui suppose une stationnarité d'ordre 2 (JLB06),(Den06). Cela nécessite donc des outils permettant une analyse des caractéristiques spectrales dépendantes du temps. L'analyse temps-fréquence permet de décrire l'évolution temporelle des caractéristiques fréquentielles du signal. Ainsi, cette analyse permet d'avoir une description plus satisfaisante que l'analyse spectrale classique grâce aux deux degrés de libertés offerts par les deux dimensions du plan temps-fréquence.

Les méthodes temps-fréquence permettent (Oul14) :

- 1- de fournir une représentation du signal en trois dimensions (amplitude-temps-fréquence)(voir figure 1.8).
- 2- de détecter et de suivre le développement des défauts qui génèrent une faible puissance vibratoire.
- 3- de superviser des machines dans lesquelles le processus de fonctionnement normal produit une amplitude élevée des chocs périodiques.

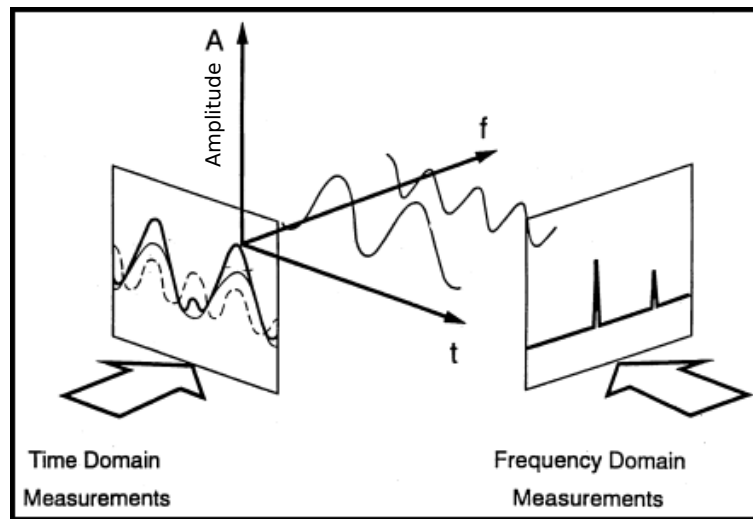


FIGURE 1.8 – Coordonnées tridimensionnelles montrant le temps, la fréquence et l'amplitude d'un signal (Oul14)

On peut noter deux variantes de l'analyse temps-fréquence : la transformée de Fourier à

1.4 Les techniques de traitement du signal appliquées à l'analyse vibratoire

court terme et la transformée de Wigner-Ville.

La transformée de Fourier fenêtrée ou à court terme TFCT (SLM05) évite l'inconvénient du caractère global de la transformée de Fourier, une idée naturelle consiste à « tronquer » le signal en fractions supposées localement stationnaires. On localise ainsi l'analyse en sélectionnant une portion autour d'une position temporelle, puis en calculant la transformée de Fourier de ce segment. On peut ensuite recommencer pour d'autres positions, ce processus est appelé la transformée de Fourier fenêtrée ou la transformée de Fourier à court terme TFCT. La transformée de Fourier fenêtrée est la plus ancienne des méthodes temps-fréquence, elle consiste à réaliser une transformée de Fourier sur une fenêtre du signal, $x(t)$, qui glissera temporellement.

$$S_x(\tau, f) = \int x(t) h(t - \tau) e^{-j2\pi ft} dt \quad (1.16)$$

Le signal est découpé au moyen d'une fenêtre $h(t)$ où l'indice τ représente le positionnement temporel de cette fenêtre et donc le positionnement de ce spectre. La série de spectre ainsi reconstituée représente une forme de transformée temps-fréquence du signal appelé spectrogramme (JLB06). On peut cependant noter un inconvénient majeur de la méthode de TFCT sur le fait que le signal est supposé stationnaire durant la durée de la fenêtre. Ainsi la longueur de fenêtre est choisie pour respecter cette hypothèse. La taille de la fenêtre influence également la résolution temporelle et donc fréquentielle. Lorsque la taille de la fenêtre est petite la résolution en temps est grande mais la résolution en fréquence est médiocre, et vice versa, selon le principe d'incertitude de Heisenberg. Comme la fenêtre est de longueur fixe cela représente un handicap important lorsqu'on veut traiter des signaux dont les variations peuvent avoir des ordres de grandeur très variables. Il serait donc intéressant d'adapter les fenêtres d'observation successives aux variations de structure du signal de façon que les hypothèses de stationnarité locale soient satisfaites.

La transformation de Wigner-Ville (TWV) est un outil d'analyse temps fréquence du signal (Den06),(JLB06),(BB01). Elle fournit un moyen efficace d'analyse des phénomènes physiques non stationnaires permettant de décrire l'évolution temporelle du spectre de ces phénomènes. La TWV possède des avantages significatifs sur d'autres méthodes voisines d'analyse temps-fréquence car elle ne fait pas appel aux méthodes issues du cas stationnaire comme la TFCT (LI12).

1. TRAITEMENT DU SIGNAL POUR LA MAINTENANCE

La TWV associe à un signal temporel $x(t)$ d'énergie finie la fonction $W_x(t, \gamma)$ des deux variables temps t et fréquence ν définie par l'équation :

$$W_x(t, \nu) = \int_{-\infty}^{+\infty} x(t + \frac{\tau}{2})x^*(t - \frac{\tau}{2})e^{-j2\pi\nu\tau}d\tau \quad (1.17)$$

Pour limiter l'effet des termes d'interférences, il existe d'autres variantes de cette distribution comme par exemple la distribution de Wigner Ville lissée (FR90).

En résumé la transformée de Fourier et la transformée de Fourier à fenêtre présentent un inconvénient majeur. La transformée de Fourier est une transformation globale tandis que la TFCT ou la transformée de Wigner-ville est locale mais toutes les deux sont de résolution temporelle fixe. La transformée en ondelettes (ou transformation temps-échelle) permet de pallier cet inconvénient car il est nécessaire de disposer d'un outil qui adapte sa résolution à la taille de l'objet ou du détail analysé.

1.4.6 Analyse temps-échelle

La transformée en ondelettes 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 (SAFN07), soit pour réaliser un diagnostic en analyse vibratoire (LM97).

Une ondelette désigne une fonction qui oscille sur un intervalle de longueur finie (un temps donné si la variable est de type spatial). L'ondelette, notée $\psi(t)$, est une fonction continue, a des moments nuls, et est nulle au-delà d'un segment de $\mathbb{R}$. Plus précisément, la condition d'admissibilité pour ψ est définie par cette équation :

$$\int_{\mathbb{R}_+} \frac{|\psi(s)|^2}{|s|} ds < +\infty \quad (1.18)$$

La transformation en ondelettes continu (TOC) consiste à calculer un " index de ressemblance " entre le signal et l'ondelette $\psi_{a,b}(t)$ obtenue par dilatation d'un facteur a positif et de décalage d'une position b de l'ondelette de référence (ou ondelette mère) $\psi(t)$.

$$C_{a,b} = \int_{-\infty}^{+\infty} x(t)\psi_{a,b}(t)dt \quad (1.19)$$

$$\psi_{a,b}(t) = \frac{1}{\sqrt{a}}\psi\left(\frac{t-b}{a}\right) \quad (1.20)$$

1.4 Les techniques de traitement du signal appliquées à l'analyse vibratoire

Le facteur d'échelle (ou dilatation) a est lié à la notion de fréquence tandis que le décalage b est lié à la notion de position temporelle.

L'application de cette transformée, dans le domaine de la détection et du diagnostic des roulements, a été développée depuis environ 20 ans avec un engouement particulier (PJ84), (PC04). Mori, en 1996, fut l'un des pionniers avec l'utilisation de l'ondelette de Morlet pour le diagnostic des roulements. Ensuite, de nombreuses études ont amélioré son utilisation et ont étendu le nombre d'ondelettes mères. 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 en 2004. Pour des signaux physiques présentant des variations très rapides, comme dans le cas d'un défaut de bague extérieure d'un roulement, l'analyse par ondelettes est adaptée car l'ondelette va détecter ces variations et les analyser. Ainsi, 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 des objectifs à atteindre (débruitage ou détection) pour avoir des résultats satisfaisants.

En général, la transformée en ondelettes est utilisée pour débruiter le signal vibratoire par une transformation discrète DWT. Le principe du débruitage par ondelettes est assez facile du fait que les ondelettes fournissent une méthode simple et basée sur des algorithmes rapides. Le débruitage par ondelettes est réalisé principalement en trois étapes que l'on peut résumer comme suit :

- décomposition,
- sélection ou seuillage des coefficients,
- reconstruction du signal débruité.

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 (RU01, NA02a, NA02b, Chi07).

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

1. TRAITEMENT DU SIGNAL POUR LA MAINTENANCE

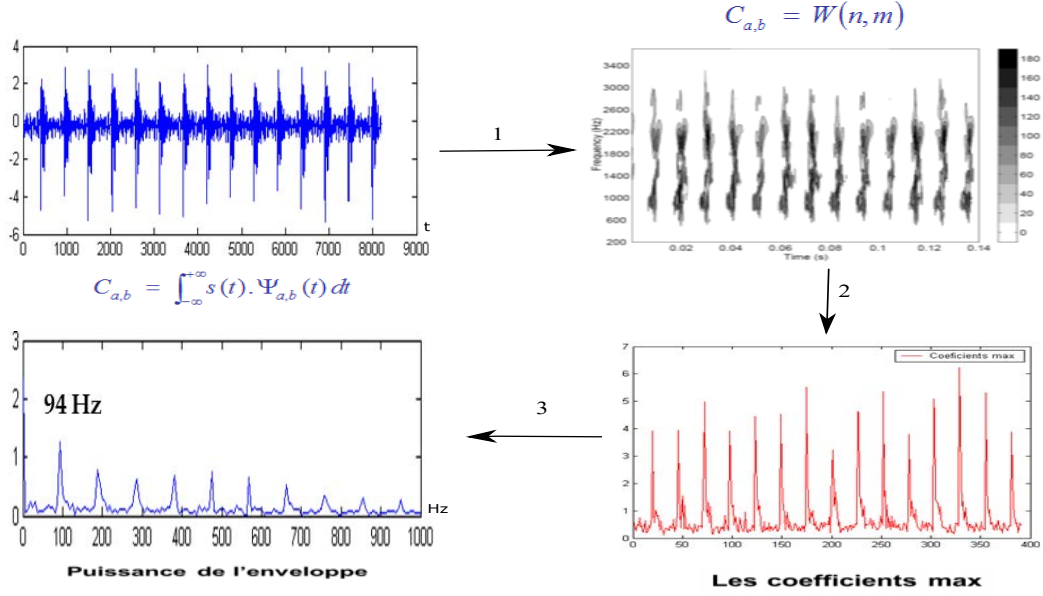


FIGURE 1.9 – les étapes de l'application de la transformée d'ondelettes

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 des coefficients. Le spectre du signal enveloppe est obtenu à partir de la transformée de Fourier de $x_{max}(t)$. La figure 1.9 représente les différentes étapes pour obtenir le spectre d'enveloppe du signal par la transformation en ondelettes.

Deux indicateurs peuvent être extraits de ce spectre d'enveloppe : le X_{Wrms} ou la valeur efficace des fréquences et le X_{PCWT} qui représente la valeur moyenne des amplitudes du spectre d'enveloppe.

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

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

1.4 Les techniques de traitement du signal appliquées à l'analyse vibratoire

$W_S(j)$ 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.

Les indicateurs SPRI et SPRO permettent d'identifier la présence des pics aux alentours des fréquences de défauts de la bague intérieure et extérieure d'un roulement.

$$SPRI = \frac{k \sum_{j=1}^H p_I(h)}{\sum_{k=1}^K S(f_k)} \quad (1.23)$$

$$SPRO = \frac{k \sum_{j=1}^H p_O(h)}{\sum_{k=1}^K S(f_k)} \quad (1.24)$$

$p_O(h)$ et $p_I(h)$ représentent les amplitudes de la h^{me} harmonique des fréquences caractéristiques de la bague extérieure BPFO et de la bague intérieure BPFI (voir les équations en 1.3.2) et $h = 1, 2, \dots, H$. H est le nombre d'harmoniques.

1.4.7 Décomposition modale empirique (EMD)

La décomposition modale empirique (ou EMD pour Empirical Mode Decomposition) est une méthode de décomposition adaptative, non paramétrique et locale de signaux non stationnaires. Cette méthode a été mise au point en 1998, par N.E. Huang (EZH⁺98) ingénieur à la NASA, pour l'étude des données océanographique. Elle a pour objectif de décomposer tout signal en une somme de composantes oscillantes extraites directement de celui-ci de manière adaptative. Ces composantes (ou IMF pour "Intrinsic Mode Functions") s'interprètent comme des formes d'ondes non stationnaires (i.e., modulées en amplitude et en fréquence) pouvant être éventuellement associées à des oscillations non linéaires.

Tout signal peut être considéré comme la superposition d'une composante lente $m(t)$ (basse fréquence) appelée approximation (ou moyenne locale) et une composante rapide $d(t)$ (hautes fréquences) appelée détail. Ces composantes sont des IMF (Fonctions modales Intrinsèques) interprétées comme étant des ondes non stationnaires.

On peut illustrer le principe de fonctionnement de l'EMD à l'échelle 1 en quatre étapes de base (FG04) :

1. TRAITEMENT DU SIGNAL POUR LA MAINTENANCE

1. On détermine les extremas de $s(t)$ (les maxima et minima locaux du signal) qui sont interpolés pour trouver une enveloppe supérieure $E_{max}(t)$ et une enveloppe inférieure $E_{min}(t)$,
2. On détermine la moyenne locale $m(t)$ de ces enveloppes $m(t) = 0.5 \times (E_{max}(t) + E_{min}(t))$,
3. On soustrait $m(t)$ du signal $s(t)$ pour obtenir $h_i(t)$ ($h_i(t) = s(t) - m(t)$).
4. Si $h_i(t)$ rentre dans les conditions d'un IMF ($h_i(t)$) sera la première composante oscillante ($c_1(t) = h_i(t)$), sinon on itère les $h_i(t)$ jusqu'à on retrouve une qui respecte les conditions,
5. Séparer $c_1(t)$ de $s(t)$ et retrouver le résidu $r_1(t) = s(t) - c_1(t)$.
6. Itérer toutes les opérations sur le résidu $r_1(k)$ qui sera considéré comme un nouveau signal ($s(t) = r_1(t)$) jusqu'à ce que "enveloppe moyenne = 0".

Il existe différentes méthodes d'interpolation pour les enveloppes. La plus utilisée est l'interpolation spline cubique, et nous considérerons dans la suite la définition des enveloppes avec cette méthode. Une IMF (pour Intrinsic Mode Function) est une fonction oscillante de moyenne nulle, c'est-à-dire une fonction :

- dont tous les maxima sont positifs, et tous les minima négatifs.
- dont la moyenne locale, au sens de la définition précédente, est nulle en tout point.

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 (YYJ06), (DZ12), (PBG06), (LDD07).

La décomposition modale empirique (*EMD*) est un outil de traitement du signal récemment introduit dans le diagnostic des roulements pour son caractère auto-adaptatif et pour sa facilité d'utilisation. Cette décomposition peut être un outil puissant dans le diagnostic des roulements puisqu'il peut aider à débruiter le signal pour ne garder que les composantes informatives dans le diagnostic et le suivi des roulements.

Plusieurs variantes de l'EMD ont été proposées dans la littérature comme l'*EEMD* (Ensemble Empirical Mode Decomposition) ou le *CEEMDAN* (Complete Ensemble

1.4 Les techniques de traitement du signal appliquées à l'analyse vibratoire

Empirical Mode Decomposition with Adaptive Noise). Des indicateurs temporels ou fréquentiels peuvent être calculés à partir des *IMFs* comme (YYJ06) qui extrait l'entropie de l'énergie des *IMFs* pour surveiller l'état d'un roulement. L'*EMD* et ses variantes ont été également combinées avec des méthodes statistiques comme la *SVD* ou l'*ACP* pour caractériser le défaut d'un roulement (MWM11), (YHB13), (YYJ06), (ZLJY15).

1.4.8 Cyclo-stationnarité

Un signal est dit cyclostationnaire si ses propriétés statistiques d'ordre 1 (moyenne, variance) ou d'ordre 2 (fonction de corrélation) sont périodiques (GS94), (SG94). Les propriétés de cyclo-stationnarité couvrent une typologie statistique du signal permettant de générer une fonction nouvelle appelée corrélation spectrale définie comme la TF à deux dimensions suivant le temps et le cycle de la fonction de corrélation (Gho08). Développée dans le domaine des télécommunications, cette théorie connaît beaucoup de développements et d'applications dans le domaine du diagnostic par analyse acoustique et vibratoire des machines tournantes (LL03), (R.B11), (Ran01), (Bon04). L'analyse de la cyclo-stationnarité permet de révéler la présence de la modulation dans le signal vibratoire ce qui permet de trouver :

- la fréquence modulante (fréquence cyclique) qui permet de trouver la signature de défaut.
- la fréquence de la porteuse qui dépend de la structure ou en d'autres termes la fonction de transfert entre la source des vibrations et le capteur. Cela facilite la détection des défauts de structure.

L'approche de la cyclo-stationnarité est très bien adaptée au diagnostic des machines tournantes pour plusieurs raisons. Tout d'abord, l'apparition d'un défaut dans un composant va produire un dégagement d'énergie vibratoire répétitif, ce qui va générer un signal avec un comportement cyclostationnaire. La dimension supplémentaire offerte par le phénomène cyclostationnaire (temps / axe angulaire, l'axe de fréquence cyclique) a permis également une meilleure caractérisation des signaux vibratoires comme la re-échantillonnage angulaire permettant l'exploitation de la vitesse instantanée dans le cadre du diagnostic (Bon04), (Ibr09).

1.4.9 Approches statistiques

Des travaux de recherche se sont orientés vers des approches statistiques pour la détection et le diagnostic des défauts des composants mécaniques des machines tournantes (C09), (Ceb12), (BM12). Ces méthodes sont basées sur des algorithmes de fouille de données pour déterminer l'état de santé de la machine. Ces algorithmes utilisent des indicateurs de défaut pour détecter le défaut, l'identifier puis de quantifier sa sévérité. Cependant la fiabilité de diagnostic par ces procédures dépendent des indicateurs utilisés et leurs sensibilités aux bruits, aux variations de vitesse ou aux variations de charges.

Après le choix de(s) l'indicateur(s) de défauts. La problématique rencontrée par les méthodes issues des approches statistiques réside dans le choix de la bonne méthode de classification. Cette méthode de classification sera responsable du regroupement des signaux dans différentes classes ou catégories, ou chaque classe représentera, par exemple, un type de défaut. Ce regroupement sera fait sur la base de la similitude des caractéristiques ou des indicateurs (WKYS12), (YHB13), (ZZ13a). 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. L'analyse par la classification statistique est souvent utilisée pour diagnostiquer une machine ou un composant en associant la signature vibratoire enregistrée à une des classes définies. Cette association ou attribution à la bonne classe se fait en mesurant la similarité ou la différence entre les caractéristiques de cet enregistrement et les caractéristiques 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 (MEA14), la distance de Mahalanobis (SWT02), la distance bayésienne (GZS01). Parmi les méthodes de classification couramment utilisées dans le diagnostic des roulements on peut citer K-means pour classifier des signaux enregistrés pour des roulements avec différents défauts (YGA11).

Ces méthodes seront détaillées dans le chapitre suivant.

1.4.10 Approches issues de l'intelligence artificielle

Dans la littérature, les techniques de l'Intelligence Artificielle (Artificial intelligence AI) sont de plus en plus proposées pour diagnostiquer les machines tournantes et elles semblent plus efficaces que les approches conventionnelles, en théorie (JLB06). En revanche, il n'est pas facile d'appliquer ces techniques dans la pratique en raison de l'absence de procédures efficaces pour obtenir des données qui peuvent être utilisées pour la phase d'apprentissage et des connaissances spécifiques, qui sont nécessaires pour former les modèles. Jusqu'ici, la plupart des applications dans la littérature utilisent simplement des données expérimentales pour la formation de modèle. Dans la littérature également, deux techniques d'IA sont souvent employés pour le diagnostic des machines tournantes (Khe14) : Les réseaux de neurones artificiels (Artificial Neural Network ou ANN) (RHS96)(FL02)(BAB03) et les systèmes experts (Expert system : ES) (LSN96)(EP08). Un ANN est un modèle de calcul qui imite la structure du cerveau humain. Il se compose d'éléments de traitement simples connectés entre eux pour former une structure de couche complexe qui permet de générer un modèle approximatif d'une fonction non-linéaire complexe avec entrées et sorties multiples. Un élément de traitement comprend un nœud et un poids. Le ANN modélise la fonction inconnue en ajustant ses poids avec les observations d'entrée et de sortie. Ce processus est généralement appelé entraînement d'un ANN. Il existe différents modèles de réseaux neuronaux. Feed-forward neuronal network(FFNN) est la structure la plus largement utilisée de réseau de neurones dans le diagnostic des machines (JLB06).

Contrairement aux réseaux de neurones, qui apprennent des connaissances par entraînement sur des données observées avec entrées et sorties connues, ES utilisent des connaissances d'experts dans un programme d'ordinateur avec un moteur d'inférence automatisé pour le raisonnement pour la résolution de problèmes. Trois principales méthodes de raisonnement pour ES sont utilisées dans le domaine du diagnostic de machines : le raisonnement par règles (CBBL89), le raisonnement par cas (WJCM03) et le raisonnement basé sur un modèle (SCR03).

1.5 Limitations des méthodes classiques de diagnostic et suivi

Les techniques de diagnostic traditionnelles basées sur l'analyse vibratoire extraient des caractéristiques statistiques du signal brut en sa forme temporelle et spectrale ; l'évolution temporelle de ces caractéristiques est utilisée pour détecter la présence du défaut qui est souvent lié à un changement brusque ou graduel de ces indicateurs. Cependant ce changement peut être dû à d'autres facteurs que la présence de défaut comme les facteurs non linéaires qui peuvent affecter la machine tournante et ajoutent à la complexité du système, par exemple le changement de la vitesse, les charges ou la friction pour ne citer que les cas les plus fréquents (AS10).

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 ((ST13), (AHA12)) 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 de leurs mises en oeuvre dans le diagnostic des machines tournantes.

Chapitre 2

Méthodes de reconnaissance de formes dans le cadre de diagnostic des machines tournantes

2.1 Présentation générale

La 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 (TSK05). La reconnaissance a commencé à prendre sa forme actuelle et devenir une discipline spécifique dans les années 60 (TD60). Les techniques de reconnaissance des formes ont permis l'avancement de la technologie qu'on connaît actuellement. Grâce à ces techniques plusieurs applications en temps réel ont connu le jour, en particulier dans le domaine des applications auditives et visuelles (LWLC05).

Les domaines d'application des méthodes de reconnaissance de formes sont très variés. On peut citer la reconnaissance de parole (YGD07), la reconnaissance d'image (W85), l'identification des défauts (DRP⁺15). Ces méthodes peuvent être employées pour : assurer la sécurité, automatiser des tâches, diagnostiquer des processus industriels ou des personnes malades, ou tout simplement pour humaniser les ordinateurs en leur donnant la capacité de reconnaître et d'exprimer des émotions, de répondre intelligemment l'émotion humaine et d'utiliser les mécanismes de l'émotion qui contribuent à la prise de décision rationnelle.

Le terme « forme » dans la reconnaissance des formes est défini par Watanabe (S.W85) « l'opposé du chaos, une entité, vaguement définie, à qui on peut donner un nom ». Une forme peut être, par exemple, une empreinte digitale, un mot manuscrit, un visage humain, ou un signal . . . Du point de vue statique, une forme représente l'observation du système en fonction de plusieurs paramètres à un instant donné. L'espace de représentation contient l'ensemble des formes connues observées à différents instants.

La reconnaissance des formes se présente sous deux aspects : l'identification et la classification. L'identification consiste à déterminer que l'observation, Y , est une manifestation de l'individu, « I » préalablement connu (Gha03). Tandis que la classification consiste à déterminer que l'observation, Y , est une manifestation d'un membre de la classe, « K ». Dans les deux cas, les individus ou les classes sont caractérisés par un vecteur de propriétés. La reconnaissance des formes est basée sur la comparaison des propriétés (par exemple la distance, le trajectoire, des caractéristiques statistiques...). L'efficacité de celle-ci réside essentiellement dans la détermination des propriétés discriminantes.

2.2 Approches de reconnaissance de formes

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.

2.2.1 La comparaison à des modèles (Template matching)

C'est l'une des plus simples et des plus anciennes approches de la reconnaissance des formes (RdF). Elle est utilisée pour déterminer la similarité entre deux entités (point, courbe, forme) de même nature (S.W85). Le modèle de référence (souvent une forme à 2 dimensions) avec lequel on compare les autres entités est établi à partir de l'ensemble d'apprentissage. La méthode de comparaison doit prendre en considération toutes les postures possibles de l'objet (translation, rotation, changement d'échelle). La similarité mesurée est souvent une corrélation qui peut être améliorée en utilisant les données disponibles dans l'ensemble d'apprentissage. Cette technique est exigeante en termes de calculs, mais avec l'évolution technologique son utilisation est devenu possible. Cependant la rigidité de cette approche bien qu'efficace dans certains domaines d'application a un certain nombre d'inconvénients. Par exemple, dans le processus de traitement d'image, cette méthode échouerait si les entités sont déformées, s'il y a un changement de point de vue ou s'il y a une large variation des intra-classes entre les formes (Bru09).

2.2.2 L'approche statistique (statistical pattern recognition)

Dans l'approche statistique, chaque forme est représentée par d caractéristiques ou mesures et elle est considérée comme un point dans l'espace de d dimensions (S.W85). L'objectif est de choisir des caractéristiques qui permettent aux vecteurs de formes appartenant à différentes classes d'occuper des régions compactes et disjointes dans un espace de d dimensions. Plus les classes sont bien séparées (la distance entre classe dans l'espace de décision est grande), meilleur est le choix des caractéristiques (HDSIW06).

Étant donné un ensemble de formes d'apprentissage de chaque classe, l'objectif de cette approche est d'établir des limites de décision dans l'espace des caractéristiques qui permettent de séparer des formes appartenant à des classes différentes. Dans l'approche

2. MÉTHODES DE RECONNAISSANCE DE FORMES DANS LE CADRE DE DIAGNOSTIC DES MACHINES TOURNANTES

statistique théorique de la décision, les limites de décision sont déterminées par les distributions de probabilité des formes appartenant à chaque classe. Ces distributions de probabilité doivent être spécifiées ou apprises.

2.2.3 L'approche syntaxique ou structurelle (Syntactic or structural matching)

Dans le cas où les formes sont complexes, il est plus approprié d'adopter une perspective hiérarchique dans laquelle une forme est vue comme étant composée des sous-formes qui sont elles mêmes composées de sous-formes plus simples. Ces sous formes élémentaires sont appelées « primitives ». A partir des relations entre ces primitives on peut représenter des formes plus complexes (S.W85).

En reconnaissance des formes syntaxiques, une analogie formelle est établie entre la structure des modèles et la syntaxe d'une langue. Les motifs ou formes sont considérés comme des phrases appartenant à une langue, les primitives sont considérées comme des alphabets de la langue, et les phrases sont générées selon une grammaire. Ainsi, une grande collection de modèles complexes peut être décrite par un petit nombre de primitives et des règles grammaticales (Fuk90). La grammaire pour chaque classe de modèle doit être déduite à partir des échantillons disponibles. La RdF structurelle est immédiate car, en plus de la classification, cette approche fournit également une description de la façon dont le motif donné est construit à partir des primitives. Ce modèle a été utilisé dans des situations où les motifs ont une structure définie qui peut être établie en termes d'un ensemble de règles, telles que des formes représentant des ondes, les images de texture, et l'analyse de la forme des contours (JDJ00).

2.2.4 Les Réseaux de neurones (Neural networks)

Un réseau de neurones est un système de traitement d'informations. Les réseaux de neurones 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 (Gha03). Les unités de traitement travaillent en coopération les uns avec les autres et elles peuvent réaliser des calculs complexes en utilisant un traitement massif parallèle. Les réseaux de neurones s'inspirent de certaines fonctionnalités du cerveau biologique et les systèmes neuronaux. Les avantages des réseaux de

neurones sont leurs apprentissages adaptifs, leurs auto-organisations et leurs capacités de tolérance aux pannes (JDJ00).

Bien qu'on puisse trouver la majorité de ces méthodes employées dans le diagnostic des machines tournantes, les méthodes de reconnaissance de formes du type "statistique" sont celle qui offrent le meilleur compromis facilité et utilité.

La reconnaissance en général est une suite d'étapes commençant par la collecte des données et finissant par l'établissement d'une décision.

2.3 Méthode de reconnaissance de formes statistique

Dans Cette thèse, On s'intéresse au suivi de roulement en analysant ses signaux vibratoires. La RdF de type statistique, où les entrées sont sous formes de vecteurs de mesures, est plus adaptée à nos objectifs. La méthode de RdF est réalisée en plusieurs étapes (voir figure 2.1). Chaque étape est essentielle dans le processus de reconnaissances des formes pour une application donnée.

2.3.1 Acquisition des données

Dans une application, les données acquises peuvent être de type et de nature différentes et dépendent de ce que l'on recherche. Suivant la nature du signal, un capteur est nécessaire pour acquérir le signal sous forme numérique ou analogique. Dans cette étape, la pertinence des données acquises vis-à-vis de l'application n'est pas étudiée. Par contre, il est nécessaire de prendre en compte l'incertitude liée aux différents capteurs (SJ01). L'emplacement du capteur, le bruit, la sensibilité et l'étendue du capteur sont des facteurs qui peuvent influencer la qualité des données acquises et font que ces données nécessitent un prétraitement pour qu'elles soient exploitables.

2.3.2 Prétraitement

Le prétraitement consiste à faire ressortir un motif intéressant de l'arrière-plan. En règle générale, cette étape consiste à filtrer et à transformer les données enregistrées pour obtenir des données plus adaptées qui vont faciliter la recherche de caractéristiques informatives. Généralement on utilise des méthodes de dé-bruitage (ARFI07), de lissage (ARM00), de

2. MÉTHODES DE RECONNAISSANCE DE FORMES DANS LE CADRE DE DIAGNOSTIC DES MACHINES TOURNANTES

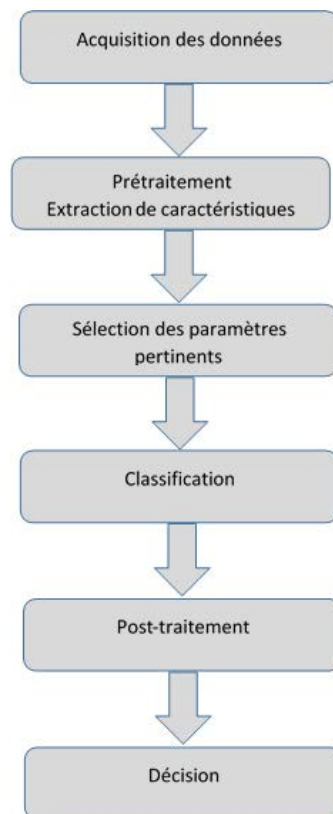


FIGURE 2.1 – Diagramme d'un système de reconnaissance des formes statistique

filtrage (LYT02), de normalisation (SC95), de sous-échantillonnage (OKK98), ou d'autres méthodes de prétraitement. L'essentiel, c'est de pouvoir "nettoyer" les données enregistrées et de les préparer à l'étape suivante qui consiste à définir les caractéristiques représentatives ou descriptives des données.

2.3.3 Extraction de caractéristiques

L'objectif de cette étape est d'extraire des caractéristiques informatives qui permettront de mieux discriminer les modes de fonctionnement et de bien séparer les classes. Ces caractéristiques doivent être robustes, insensibles au bruit et limiter toute variabilité indésirable (par exemple des caractéristiques qui sont sensibles au défaut mais insensible au changement de conditions de fonctionnement).

Il existe deux types de caractéristiques utilisées en reconnaissance des formes : les ca-

caractéristiques qui portent une signification physique claire tels que les indicateurs de défaut dans le diagnostic des machines tournantes, et les caractéristiques qui n'ont pas de sens physique direct mais qui peuvent être complémentaires au premier type de caractéristique. Les caractéristiques physiques permettent d'interpréter le résultat final tandis que les caractéristiques non physiques, elles, permettent de faciliter la classification en améliorant la séparabilité des classes. Les caractéristiques non physiques permettent également de voir les données sous une autre perspective. Certes, ce type de caractéristique peut augmenter la complexité des calculs. Cependant elles peuvent améliorer la qualité de classification.

Dans l'extraction des caractéristiques physiques, on peut avoir recours à plusieurs types de méthodes dont le choix dépend d'abord du domaine d'application. En diagnostic des machines, par analyse vibratoire, on peut utiliser des méthodes de traitement du signal tels que la transformée de Fourier, les méthodes Temps-Fréquence (Wigner-Ville (FG04, SWT02), Gabor), les ondelettes (BKS⁺10)), les filtres Auto-régressif (AR), à moyenne mobile (Moving Average (JATH12) ou MA) ou auto-régressif à moyenne mobile (ARMA) (WSJ⁺12) etc. Ce type de caractéristiques peuvent être calculés à partir de différentes méthodes de transformation du signal ou directement à partir du signal temporel (débruité ou non). Parmi les paramètres les plus utilisés on peut citer le nombre de pics présents dans un signal, l'écart type des données, la moyenne quadratique (RMS), la valeur maximale ou minimale (SKTR15), le coefficient d'aplatissement (Kurtosis) (HAWF05), etc.

L'extraction des caractéristiques non physiques ou physiquement non interprétables repose sur deux principes distincts : le premier consiste à créer des sous-ensembles de caractéristiques qui permettent d'enlever des éléments redondants ou non pertinents de l'ensemble de données car elles peuvent conduire à une réduction de la précision de la classification et à une augmentation inutile des coûts de calcul. Le second repose sur l'extraction des nouveaux paramètres en employant des techniques de réduction de dimension afin que la taille de l'espace d'attributs (espace de caractéristiques) puisse être réduite sans perdre beaucoup d'informations de l'espace de l'attribut d'origine. Les méthodes de réduction de dimension (RD) font référence à des algorithmes et des techniques qui créent de nouveaux attributs en combinant des attributs d'origine afin de réduire la dimension d'un ensemble

2. MÉTHODES DE RECONNAISSANCE DE FORMES DANS LE CADRE DE DIAGNOSTIC DES MACHINES TOURNANTES

de données (JDJ00). Parmi les techniques les plus utilisées dans la RD pour des caractéristiques linéaires on trouve l'analyse discriminante linéaire (LDA) (SDL10), l'analyse en composantes principales (ACP) (NT97), l'analyse en composantes indépendantes (ICA) (AE00), etc. Ces méthodes produisent de nouveaux attributs comme des combinaisons linéaires des variables initiales. En revanche, d'autres méthodes sont utilisées pour des caractéristiques non linéaires comme l'analyse en composantes principale à noyau (kernel Principal component analysis KPCA) (LWLC05), le réseau auto-associatif non linéaire, l'analyse multidimensionnelle (MDS) (WB13), la carte auto adaptative (Self organized Map SOM) (BKS⁺10), les projections à localité préservés (Locality Preserving Projections, LPP) (AHP14). La méthode *KPCA* a déjà été utilisée dans le cadre du diagnostic des roulements, notamment par Zhang *et al* dans (ZZB13) pour souligner la supériorité de KPCA par rapport à la méthode *ACP*.

2.3.4 Sélection de caractéristiques pertinentes

La sélection repose sur le choix des caractéristiques les plus discriminantes de l'ensemble des caractéristiques disponibles (JGDE08), 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.

L'une des difficultés principales liées à la représentation des données est la dimension des données. Le problème de la dimension des données concerne le nombre et la qualité des variables descriptives caractérisant chacun des individus. Ce problème peut se résumer par la phrase de Liu et al dans (LMSZ10), « Less is more » qui signifie que si l'on désire extraire de l'information utile et compréhensible à partir de nos données, il convient en premier lieu de retirer les parties non pertinentes et d'éliminer toute redondance possible.

Les méthodes de sélection choisissent le sous-ensemble de paramètres les plus informatifs (JGDE08). 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.

Le problème de sélection peut être formulé comme suit : Etant donnée un ensemble de paramètres ou caractéristiques $X = \{x_i | i = 1 \dots N\}$, il faut trouver un sous-ensemble

Y_M avec $M < N$ qui maximise une fonction objectif $J(Y)$

$$Y_M = \{x_{i1}, x_{i2}, \dots, x_{iM}\} = \arg_{M, i_M} \max J\{x_i | i = 1 \dots N\} \quad (2.1)$$

Les méthodes de sélection sont classées généralement en deux groupes : Les méthodes « filtre » et les méthodes « wrapper ».

1. La première approche (**méthode "filtre"**), inspirée de l'apprentissage non supervisé, utilise des mesures statistiques calculées à partir des caractéristiques afin de filtrer les caractéristiques peu informatives. Cette approche est indépendante de la méthode de classification. Elles présentent des avantages au niveau de leur efficacité en termes de calcul et au niveau de leur robustesse face au sur-apprentissage. Par contre elle ne tient pas compte des interactions entre les caractéristiques et tendent à sélectionner des caractéristiques comportant des informations redondantes plutôt que des informations complémentaires. Parmi ces méthodes filtres on trouve : la partition Fisher (Fisher score)(BBS14) ou de corrélation de Pearson (HTF09), le gain de l'information, etc. L'objectif de ces méthodes est de fournir une mesure ou un score qui évaluent les capacités des caractéristiques choisies pour séparer l'ensemble de données en classes.
2. La seconde approche (**méthodes enveloppantes** ou « wrapper ») est une méthode dite d'apprentissage supervisée. Cette approche est plus coûteuse en temps de calcul, mais en contrepartie, elle est souvent plus précise. Les wrappers sont des méthodes de rétroaction qui intègrent l'algorithme d'apprentissage dans le processus de sélection des caractéristiques. Elles s'appuient sur la performance d'un classificateur spécifique pour évaluer la qualité d'un ensemble de caractéristiques. Les méthodes enveloppantes cherchent l'espace des sous-ensembles de caractéristiques et calculent la précision estimée d'un algorithme d'apprentissage unique pour chaque caractéristique qui peut être ajoutée ou retirée du sous-ensemble (Tal05). Diverses stratégies peuvent être utilisées pour chercher l'espace des sous-ensembles de caractéristiques. Habituellement, une recherche exhaustive est trop coûteuse en temps de calcul et donc les techniques de recherche heuristique non-exhaustive, comme les algorithmes génétiques, les algorithmes « greedy stepwise » (HNZ12), les algorithmes « best first strategy » (XYC88) ou « random search » (BSM06) sont les plus utilisés.

En règle générale, plus les paramètres du système permettent de discriminer les formes, meilleurs sont les résultats de la classification. L'ensemble des paramètres trouvés par ces

2. MÉTHODES DE RECONNAISSANCE DE FORMES DANS LE CADRE DE DIAGNOSTIC DES MACHINES TOURNANTES

méthodes représente les attributs qui permettent de caractériser chaque forme. Lorsque les données issues de l'observation du fonctionnement d'un système sont représentées par des paramètres statistiques, 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 paramètres sont bien déterminés, les classes sont bien discriminées et elles sont situées dans différentes régions de l'espace de représentation.

Chaque classe est associée à un mode de fonctionnement (normal ou défaillant). Ces formes, avec leurs assignements à une classe, sont représentées par n caractéristiques ou attributs, ce qui permet de les voir en tant que vecteurs de n dimensions, c'est à dire des points dans l'espace de représentation.

2.3.5 Classification

La reconnaissance des formes repose sur la classification des formes (vecteurs de caractéristiques de l'état de fonctionnement du système) dans des classes, où chaque classe caractérise un état de fonctionnement du système. Chaque classe a ses propres caractéristiques et elle peut être définie par un modèle qui est représenté par une fonction d'appartenance. Ces modèles qui résument toutes les règles (règles de décision) définissent pourquoi une forme est attribuée à une certaine classe et la valeur d'appartenance de cette forme à sa classe attribuée. Dans le domaine du diagnostic et du suivi des machines tournante, la forme n'est autre que l'observation simplifiée de l'état de la machine. La décision d'attribution d'une forme à une certaine classe est réalisée par un classifieur en respectant des règles de décision prédéfinies préalablement et ensuite peut être mise à jour au cours de la classification.

La méthode de classification se résume en la caractérisation du modèle et de l'étiquette de chaque classe associée à une forme. Cela nécessite des techniques qui permettent de regrouper, par apprentissage, les formes similaires. En diagnostic des machines tournantes, la reconnaissance des formes devient alors un problème de classification ; le mode de fonctionnement actuel du système est déterminé en connaissant la classe de l'observation actuelle de l'état de fonctionnement du système (Ond06).

Il existe plusieurs types de procédure d'apprentissage (voir figure 2.2). En diagnostic des machines, on trouve essentiellement deux types d'apprentissage dans la littérature : l'apprentissage supervisé et l'apprentissage non supervisé.

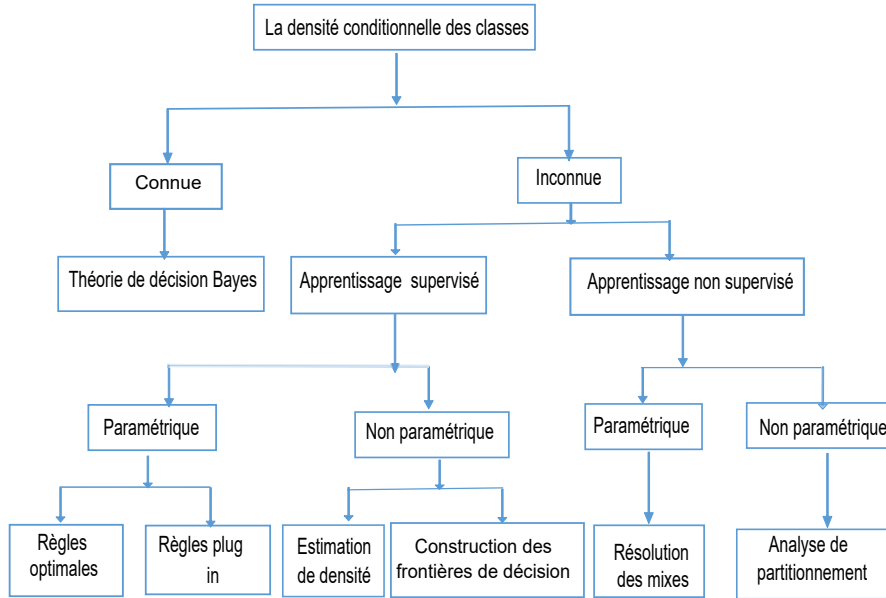


FIGURE 2.2 – Les différentes approches utilisées dans la RdF statistique

A. Méthodes de classification du type apprentissage supervisé

L'apprentissage supervisé suppose qu'on dispose déjà d'exemples dont la classe est connue et étiquetée. La classification supervisée (appelée aussi classification inductive) a pour objectif « d'apprendre » par l'exemple. Elle cherche à expliquer et à prédire l'appartenance des objets à des classes connues *a priori*. C'est donc un ensemble de techniques qui vise à prédire l'appartenance d'un individu à une classe en s'aidant uniquement des valeurs qu'elle prend (G.G03).

La classification supervisée peut se résumer par la recherche d'une fonction d'appartenance généralisée à partir des exemples donnés. Prenons l'exemple d'un objet d , décrit par n caractéristiques, l'objectif est de chercher à prédire l'appartenance de d à une classe C_k parmi n_C classes, en se servant d'un ensemble A des exemples d'objets classés dans lesquels pour chaque objet d_i on connaît *a priori* sa classe C_{d_i} (2.2)

2. MÉTHODES DE RECONNAISSANCE DE FORMES DANS LE CADRE DE DIAGNOSTIC DES MACHINES TOURNANTES

$$A = (d_1, C_{d_1}), (d_2, C_{d_2}), \dots, (d_n, C_{d_n}), d_i \in \mathbb{R}^n \text{ et } C_{d_i} \in C \quad (2.2)$$

A partir de l'ensemble A , la fonction de décision va associer à toute nouvelle observation d une classe C_k . Cette classe est choisie de sorte que le mauvais classement est minimisé, c'est à dire que le nombre de fois où la classe prédite est différente de la classe connue *a priori* ($\Gamma(d) \neq C_d$).

L'apprentissage supervisé se compose de trois étapes (voir figure 2.3) : l'apprentissage, la validation et le test. Dans chaque étape on utilise un ensemble de données propre à cette étape.

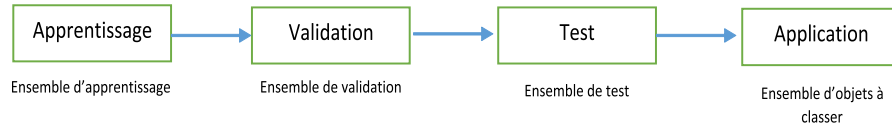


FIGURE 2.3 – Les étapes formant le processus de classification supervisée

L'apprentissage (Learning/training)

L'apprentissage consiste à déterminer un ou plusieurs modèles candidats qui peut expliquer la relation qui relie les objets à leurs classes, et de choisir également les paramètres optimaux pour un modèle donné. L'ensemble d'apprentissage se compose toujours des objets dont on connaît *a priori* leurs classes. L'évaluation de certains jeu de paramètres du modèle à l'aide de l'ensemble d'apprentissage devrait donner une estimation sans biais de la fonction de coût. La fonction de coût représente une mesure bruitée du nombre d'erreurs d'apprentissage. Pourtant le fait de choisir des paramètres qui optimisent l'estimation de la fonction de coût à partir de jeu d'apprentissage biaise l'estimation qu'ils fournissent. Les paramètres du modèle sont choisis en fonction de leurs performances sur l'ensemble de l'apprentissage. Par conséquent, le rendement apparent de ces paramètres, telle qu'évaluée sur le jeu d'apprentissage, est toujours bon.

La validation

L'ensemble de validation sert à choisir le meilleur modèle et d'évaluer la justesse du modèle choisi. L'ensemble de validation se compose de nouveaux objets pour

lesquels on connaît à l'avance leur classe d'appartenance et qu'ils n'étaient pas utilisés pendant l'apprentissage. Ces objets vont être utilisés pour vérifier que les résultats donnés par le modèle et leur vraie classe d'appartenance correspondent. En d'autres termes la validation consiste à estimer le taux d'erreur de classification pour décider si le modèle est juste ou si on a besoin de l'ajuster.

Le test

L'ensemble de test sert à indiquer que le choix final du modèle est bon. Il donne une estimation non biaisée de la performance réelle qu'on va obtenir lors de l'exécution du test. Dans cette phase, il est nécessaire d'utiliser un troisième ensemble de nouveaux objets classés qui est différent de l'ensemble de validation et de l'ensemble d'apprentissage. On ne peut pas utiliser l'ensemble d'apprentissage pour tester le modèle parce que les paramètres sont biaisés envers les objets d'apprentissage. Par ailleurs on ne peut pas utiliser les données de la validation parce que le modèle lui-même est biaisé en faveur de l'ensemble de validation, d'où, la nécessité d'un troisième ensemble de nouveaux objets. Dans le cas où on ne dispose pas de plusieurs modèles candidats, l'étape de validation devient l'étape test.

Parmi les méthodes d'apprentissage supervisé les plus utilisées, on peut citer les méthodes SVM (Séparateurs à Vastes Marges ou Support Vector Machine ou Machines à vecteurs de supports), l'arbre de décision, les réseaux de neurones.

A.1. Machine à vecteurs de support (SVM)

Les machines à vecteurs de support (*SVM*) sont des méthodes d'apprentissage relativement nouvelles (introduite par Vapnik en 1995 (V.V95)), conçues principalement pour la classification binaire (en deux classes). L'idée de base est de trouver un hyperplan qui sépare parfaitement les données de d dimensions dans deux classes. En règle générale, les données ne sont pas linéairement séparables, SVM a introduit la notion de "fonction noyau" qui permet de projeter les données dans un espace de dimension supérieure où les données sont linéairement séparables (HTF09).

Dans le diagnostic des machines tournantes, plusieurs variantes de *SVM* ont été utilisées, (ZZB13, WSJ⁺12) par exemple (WSJ⁺12) Wang *et al* ont utilisé la variante « Hyper-sphere-structured multi-class SVM » afin de déterminer le degré de dégradation d'un

2. MÉTHODES DE RECONNAISSANCE DE FORMES DANS LE CADRE DE DIAGNOSTIC DES MACHINES TOURNANTES

roulement en classifiant des caractéristiques calculées à partir des signaux vibratoires traités par plusieurs méthodes de traitement du signal en plusieurs groupes de données. Cette variante de SVM utilise un séparateur ou un hyperplan qui a une forme sphérique (voir figure 1.3). Zhang a utilisé une autre variante de *SVM* appelé « Particle Swarm Optimization for *SVM* (*SVM – PSO*) ou Optimisation par essaims particulaires pour le *SVM* », qui représente une version de SVM avec une fonction noyau de type « Radial Basis Function (*RBF*) » (ZZB13) qui nécessite l'initialisation de plusieurs paramètres dus à l'emploi de la méthode d'optimisation par essaims particulaires qui permet de trouver les valeurs optimales pour ces paramètres. Dans cet article la *SVM – PSO* a été employée après l'extraction et la sélection des indicateurs de défauts. Des signaux vibratoires pour un roulement en différents états de fonctionnement ont été choisis pour servir d'ensemble d'apprentissage de la méthode.

Dans l'article (ZZ13b), l'optimisation multi-classes de *SVM* par la distance interclasse (Multi-class *SVM* optimized by inter-cluster distance) est une autre variante de *SVM* qui a été employée pour le diagnostic des roulements. La particularité de cette variante est qu'elle permet de trouver les paramètres du noyau choisis en fonction de la séparabilité des classes dans l'espace des caractéristiques. En effet cette optimisation de *SVM* permet de définir le degré de séparabilité des classes ou la distance qui sépare les classes si on choisit de façon aléatoire une valeur pour initialiser les paramètres de la fonction noyau. Donc avec cette méthode on peut avoir l'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.

A.2. Arbre de décision

Un arbre de décision est la représentation graphique d'une procédure de classification. Elle peut aussi servir de modèle pour un problème de décision séquentielle incertain. Un arbre de décision décrit graphiquement la prise de décisions, les événements qui peuvent survenir et les résultats associés à des combinaisons de décisions/événements. L'arbre de décision permet d'assigner une probabilité à chaque événement dont sa valeur est déterminée pour chaque résultat. Généralement le principal objectif de ce type d'analyse est

de déterminer les meilleures décisions à prendre et le coût associé à chaque décision, et finalement de réaliser une classification parfaite avec un nombre minimal de décisions.

Un arbre de décision est un arbre au sens informatique du terme. Les ADs se constituent d'une racine, plusieurs nœuds, des branches et des feuilles. Le chemin reliant une racine à une feuille, appelé par branche, représente une séquence de possibilités qui conduit à une décision (une classe). L'attribut choisi comme nœud racine est l'attribut qui permet la meilleure séparation des échantillons à n classes (ASK13). Les nœuds internes sont appelés nœuds de décision. Chaque nœud de décision est étiqueté par un test qui peut être appliqué à toute description d'un individu de la population. En général, chaque test examine la valeur d'un unique attribut de l'espace des descriptions. Les réponses possibles au test correspondent aux étiquettes des arcs issus de ce nœud. Les feuilles sont étiquetées par une classe appelée classe par défaut (voir figure 2.4).

La construction d'un arbre de décision (AD) comprend généralement deux étapes : l'expansion et l'élagage. Dans la phase d'expansion, les données de formation (échantillons) sont réparties en deux ou en plusieurs sous-ensembles descendants à plusieurs reprises, selon certaines règles de fractionnement, jusqu'à ce que toutes les instances de chaque sous-ensemble enveloppent la même classe (pur) où un critère d'arrêt a été atteint. En général, cette phase d'expansion génère un grand arbre de décision qui comprend des exemples d'apprentissage et considère de nombreuses incertitudes dans les données (particularité, le bruit et la variation résiduelle). La procédure d'élagage basée sur des heuristiques a pour but de prévenir le problème de sur-apprentissage en supprimant toutes les sections de l'arbre de décision qui peut être basée sur des données bruitées et / ou erronées. Cela réduit la complexité et la taille de l' AD . La phase d'élagage peut sous-élaguer ou sur-élaguer l'expansion de l' AD (KSI⁺14).

L'arbre de décision a été utilisé dans le cadre du diagnostic des roulements en utilisant des paramètres issus des signaux vibratoires et aussi des signaux acoustiques. Plusieurs variantes d'arbre de décision ont été employées comme méthodes de classification des données. 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).

2. MÉTHODES DE RECONNAISSANCE DE FORMES DANS LE CADRE DE DIAGNOSTIC DES MACHINES TOURNANTES

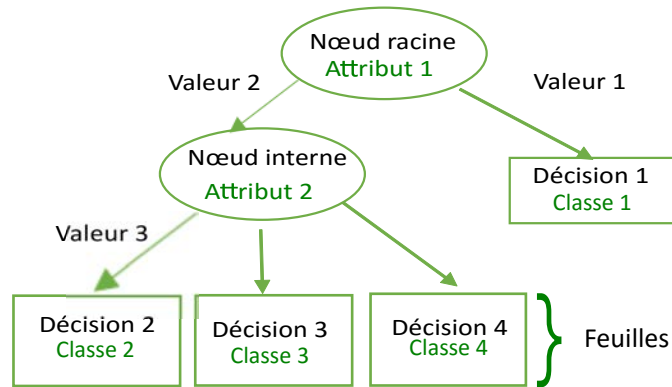


FIGURE 2.4 – Schéma du principe de l'arbre de décision

Dans l'article (KSI⁺14) les auteurs présentent une comparaison des performances de plusieurs versions de l'AD utilisées dans le domaine industriel et proposent de combiner l'arbre de décision avec des méthodes de sélection des caractéristiques et une méthode de réduction de dimensions afin de construire une version améliorée de l'arbre de décision non élagué (Improved Unpruned Decision Tree *IUDT*). Ceci convertit le problème de la construction de l'arbre de décision en une exploration combinatoire de l'espace de recherche graphique. Ces méthodes utilisent les techniques de réduction de dimensions combinées avec des techniques de sélection de données (Wrapper ou enveloppante pour choisir les attributs les plus pertinents) pour surmonter les deux challenges de l'AD que sont la complexité et le sur/sous apprentissage. Cette comparaison a 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 moyenne entre les fréquences des quatre amplitudes maximales (par les ondelettes)) est le meilleur choix de l'AD pour le diagnostic des roulements. 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.

Amarnath *et al* dans l'article (ASK13) 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). Le choix des indicateurs et leur ordre d'apparition ont été basés sur une analyse faite par les auteurs (voir la figure 2.5).

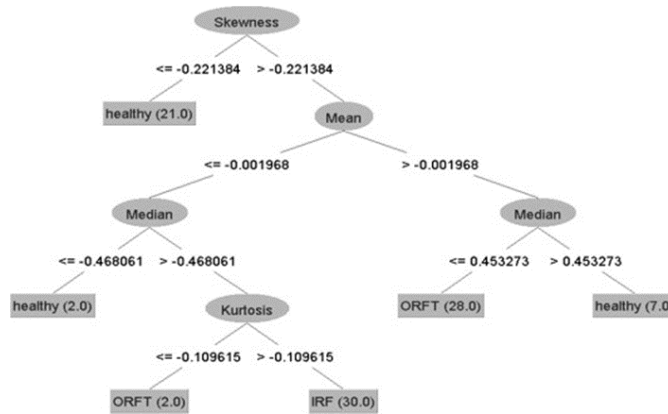


FIGURE 2.5 – L’arbre de décision présenté dans l’article (ASK13)

Dans l’article (SRSS11) les auteurs ont présenté une autre façon pour employer l’arbre de décision. Les auteurs ont utilisé l’*AD* pour sélectionner les caractéristiques des vibrations les plus pertinentes. La méthode de classification choisie fut de type *SVM*. L’*AD* a été employé pour sélectionner les indicateurs les plus pertinents c’est à dire les indicateurs qui contribuent dans la classification afin d’améliorer la précision de la méthode de classification *SVM* et de réduire sa complexité.

A.3. Réseaux de neurones

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. L’*ANN* est un système adaptatif qui modifie sa structure basée sur l’information qui circule à travers le réseau (voir figure 2.6).

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

2. MÉTHODES DE RECONNAISSANCE DE FORMES DANS LE CADRE DE DIAGNOSTIC DES MACHINES TOURNANTES

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-adaptative). 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.

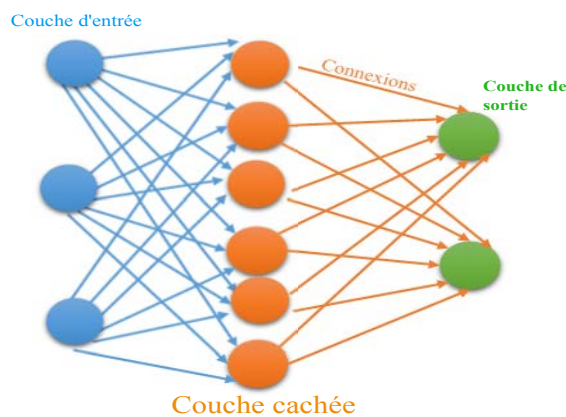


FIGURE 2.6 – L'architecture d'un réseau de neurones

Les réseaux de neurones sont peu utilisés dans le domaine du diagnostic des roulements. Dans l'article (AFS⁺15) un réseau de neurones utilisé se compose de quatre couches, une couche d'entrée avec 18 nœuds ou chaque nœud correspond à un indicateur de défaut extrait du signal vibratoire, deux couches cachées, la première couche contient 20 nœuds et la seconde 18 nœuds et finalement une couche de sortie avec 4 nœuds. Chaque couche a un poids supérieur en provenance de la couche précédente. Le poids est adapté par la fonction de transfert sigmoïde tangente hyperbolique. L'algorithme back-propagation (BP) a été employé pour former le réseau de neurones. Ce réseau a été utilisé pour classer les données en sept classes finales : roulement sain, bague intérieure dégradée, bague extérieure dégradée, bague intérieure défaillante, bague extérieure défaillante, élément

roulant dégradé, élément roulant défaillant. Le taux de reconnaissance cité dans l'article est de 93%.

Kumar *et al* (KSSV13) ont choisi d'utiliser le réseau de neurones de type perceptron multicouches qui a pour objectif de classer les 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 débruitées. Ce réseau de neurones a été combiné avec la méthode des ondelettes.

Dans (YYC06), 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 classer les données en trois classes : sain, défaut bague intérieure, défaut bague extérieure.

B. Méthodes de classification du type apprentissage non supervisé

L'apprentissage non supervisé permet d'aborder les problèmes avec peu ou pas de connaissance *a priori* de ce que les résultats devraient ressembler. C'est la recherche d'une typologie, ou segmentation, c'est-à-dire d'une partition, ou répartition des individus en classes ou catégories. L'apprentissage non supervisé passe par l'optimisation d'un critère visant à regrouper les individus dans des classes, chacune le plus homogène possible et, entre elles, les plus distinctes possible. Les classes déterminées par l'apprentissage non supervisé ont deux propriétés : d'une part elles ne sont pas prédéfinies par procédure, d'autre part les classes regroupent des objets ayant des caractéristiques similaires et séparent les objets ayant des caractéristiques différentes. L'apprentissage non supervisé est connu aussi sous le nom de classification descriptive ou classification automatique, il se caractérise par le fait qu'il n'y a pas d'objectif cible, c'est à dire une classe prédéfinie, on ne sait pas à l'avance la classe à laquelle chaque objet appartient, même le nombre de classes n'est pas toujours fixé à l'avance.

B.1. Complexité de la classification automatique

Un calcul élémentaire combinatoire montre que le nombre de partitions possibles d'un ensemble de n éléments croît plus qu'exponentiellement avec n ; le nombre de partitions

2. MÉTHODES DE RECONNAISSANCE DE FORMES DANS LE CADRE DE DIAGNOSTIC DES MACHINES TOURNANTES

de n éléments en k classes est le nombre de Stirling. Le nombre total de partition est celui de Bell B_n (voir equation 2.3)

$$B_n = \frac{1}{e} \sum_{k=1}^{\infty} \frac{k^n}{k!} \quad (2.3)$$

Pour $n = 20$, le nombre de partitionnements possibles est de 1013. Il est donc difficile de chercher à augmenter l'homogénéité de classes trouvées en s'assurant que les classes sont les plus distinctes possible et cela sur toutes les partitionnements possibles. Les méthodes existantes se limitent à l'exécution itérative convergeant vers une « bonne » partition qui correspond en général à un optimum local.

B.2. Types de classification automatique (ou clustering)

Déterministe et 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 et 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. A l'inverse, une méthode non-incrémentale va exécuter l'ensemble de données fournies en entrée. Si, par la suite, une nouvelle donnée devait être fournie en entrée de l'algorithme, celui-ci devrait être exécuté à nouveau.

Hard et Fuzzy : Une méthode **Hard ou stricte** va associer à chaque objet une classe unique alors qu'une méthode Fuzzy va associer à chaque objet un degré d'appartenance à chaque classe. Une méthode **Fuzzy clustering ou classification floue** peut être convertie en une méthode **Hard clustering ou classification stricte** en assignant chaque donnée au cluster dont la mesure d'appartenance est la plus forte.

B.3. Structure de la classification automatique

La classification automatique se compose des étapes suivantes :

- i Représentation des formes (par l'extraction et la sélection de caractéristiques pertinentes).
- ii Définition d'une mesure de proximité des formes appropriées pour le domaine d'application (distance ou similarité).
- iii Partitionnement ou clustering.
- iv Abstraction des données (facultative).
- v Evaluation de la qualité de classification (facultative).

La première étape est déjà discutée dans les parties 2.3.3 extraction et 2.3.4 sélection, donc les étapes suivantes sont explicitées ci-après.

B.4. Définition d'une mesure de proximité des formes (distance ou similarité)

Pour regrouper les objets qui se ressemblent, il faut choisir un « critère de ressemblance ». Pour cela on examine l'ensemble des informations dont on dispose concernant les objets (comme les caractéristiques extraites et sélectionnées auparavant). Les caractéristiques de chaque objet i seront considérées comme les coordonnées d'un point $M_i = (x_i, y_i, z_i, \dots)$ de l'espace.

Métrie

La métrie est une description de la distance ou la dissimilarité entre deux objets, elle permet de quantifier à quel point ces deux objets se ressemblent ou l'inverse. Cette distance (métrie) est une application de $E \times E \rightarrow R$ tel que , $\forall i, j, k \in E$:

1. $d(i, j) = d(j, i)$ symétrie
2. $d(i, j) \geq 0$
3. $d(i, j) = 0 \Leftrightarrow i = j$
4. $d(i, j) \leq d(i, k) + d(k, j)$ inégalité du triangle

Si toutes ces propriétés sont respectées, on se trouve en présence d'une distance métrie.

2. MÉTHODES DE RECONNAISSANCE DE FORMES DANS LE CADRE DE DIAGNOSTIC DES MACHINES TOURNANTES

Types de distances

La distance entre deux points x_i et y_i peut se calculer de plusieurs façons. La distance la plus connue est la **distance euclidienne** :

$$d(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \text{ pour des vecteurs de dimensions } n \quad (2.4)$$

Cette distance n'est qu'un cas particulier pour $p = 2$ de la **distance de Minkowsky** :

$$d(x, y) = \left(\sum_{i=1}^n |x_i - y_i|^p \right)^{\frac{1}{p}} \quad (2.5)$$

Pour $p = 1$, on obtient la **distance de Manhattan** (aussi appelée distance « city-block » ou métrique absolue) :

$$d(x, y) = \sum_{i=1}^n |x_i - y_i| \quad (2.6)$$

Pour $p = \infty$, la **distance de Chebychev**, aussi appelée distance « Queen-wise » ou encore métrique maximum :

$$d(x, y) = \max_{i=1}^n |x_i - y_i| \quad (2.7)$$

Le calcul du **pourcentage de différence** (percent disagreement) est une autre mesure de distance entre deux vecteurs de n éléments. Cette distance ne tient pas compte de l'importance de la différence entre les éléments homologues. Cette distance est utile dans les cas où les valeurs des éléments représentent des variables nominales non ordonnées. Cette distance est considérée comme une certaine généralisation de la **distance de Hamming**, pour des vecteurs binaires uniquement (2.8). Celle-ci est utilisée pour calculer le nombre de bits différents dans deux vecteurs.

$$d(x, y) = \frac{|x - y|}{n} \quad (2.8)$$

La **distance Mahalonobis** désigne la distance entre deux individus définie par l'équation (2.9). Utiliser la distance de Mahalonobis revient à utiliser la distance euclidienne ordinaire sur des coordonnées normalisées de l'ACP en lieu et place des données et, dans ce cas, mesurer la **corrélation** des données.

$$d(x, y) = \sqrt{(x - y)^T Cov^{-1}(x - y)} \quad (2.9)$$

$Cov()$ est la matrice de covariance.

Pour partitionner les données on peut soit mesurer leur dissimilarité (une distance) ou mesurer leur similarité. Alors que la distance mesure le degré de « différence » entre deux vecteurs, un **indice de similarité** mesure le degré de « ressemblance » entre deux vecteurs. Dans le cas d'une distance, on cherche habituellement les éléments les plus proches, c'est-à-dire qu'on cherche la distance minimale alors que dans le cas d'une similarité, on cherche les éléments les plus similaires, c'est-à-dire l'indice de similarité maximal. Les différents coefficients de corrélation peuvent être associés à des mesures de similarité.

La distance Hausdorff est une autre mesure qu'on peut retrouver dans la littérature. Cette distance peut être employée pour déterminer la dissimilarité entre un point (ou un sous ensemble) et un (autre) sous-ensemble de données. La distance Hausdorff a été implementée plusieurs fois dans des algorithmes de classification et notamment dans des méthodes de classification non supervisées (ZSY09), (CY02)). L'intégration de la distance Hausdorff permet d'améliorer la qualité de la classification (BBC⁺07).

Dans un espace métrique $(\mathfrak{S}, \delta)$ avec la métrique δ , la distance entre un point $a \in \mathfrak{S}$ et un sous-ensemble $B \subseteq \mathfrak{S}$ est donnée par

$$\tilde{d}(a; B) = \inf_{b \in B} (\delta(a, b)) \quad (2.10)$$

(Tous les sous ensembles sont considérés comme non nuls et compacts). Pour un sous ensemble quelconque $(A \subseteq \mathfrak{S})$, la fonction qui définit la distance est :

$$\tilde{d}(A; B) = \sup_{a \in A} \tilde{d}(a; B) = \sup_{a \in A} (\inf_{b \in B} (\delta(a, b))) \quad (2.11)$$

Ceci mesure la plus grande distance des distances $\tilde{d}(a; B)$. La distance de Hausdorff entre deux sous-ensembles $(A, B \subseteq \mathfrak{S})$ est :

$$d_H(A, B) = \max(\tilde{d}(A; B), \tilde{d}(B; A)) = \max(\sup_{a \in A} (\inf_{b \in B} (\delta(a, b))), \sup_{b \in B} (\inf_{a \in A} (\delta(a, b)))) \quad (2.12)$$

La distance de Hausdorff entre A et B est la plus petite distance positive r , de telle sorte que tous les points de A est à distance r de n'importe quel point de B et chaque point de B est à une distance r de n'importe quel point de A .

2. MÉTHODES DE RECONNAISSANCE DE FORMES DANS LE CADRE DE DIAGNOSTIC DES MACHINES TOURNANTES

La distance de Hausdorff peut également être utilisée pour tester les performances d'une méthode de classification ou pour définir une mesure à partir de laquelle deux objets sont définis comme dissimilaires.

Le choix de la métrique de distance ou similarité reste toujours un choix laissé à l'initiative de l'analyste. Dans le diagnostic des machines tournantes, la métrique la plus utilisée est la distance euclidienne (MEA14) suivi de la distance Mahalanobis (WWTW13).

B.5. Partitionnement ou clustering

Il existe trois types de méthode de partitionnement : la méthode hiérarchique, la méthode de classification par partition et la méthode par densité.

La méthode hiérarchique a pour objectif de construire une suite de partitions emboîtées appelée hiérarchie. La représentation graphique de ces hiérarchies se fait par un arbre hiérarchique ou dendrogramme. Il existe deux types de classification hiérarchique : la classification hiérarchique ascendante (ou agglomérative) et la classification descendante (divisives). La classification hiérarchique ascendante part du particulier pour arriver au général tandis que la classification descendante part du général pour arriver au particulier.

L'objectif de la classification hiérarchique ascendante (CHA) est de classer des individus en groupes ayant un comportement similaire sur un ensemble de variables. Son principe est le suivant : initialement chaque individu forme une classe, ensuite on commence par agréger les deux individus les plus proches. Puis, on continue en agrégeant les éléments (individus ou groupes d'individus) les plus semblables. Ces agrégations sont effectuées deux à deux. L'algorithme continue jusqu'à ce que l'ensemble des individus se retrouve dans une unique classe. 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.

Algorithm 1 L'algorithme général des méthodes hiérarchiques

- 1: **Procédure** : (*Classification hiérarchique ascendante*)
 - 2: *Entrées* : les individus.
 - 3: *Sorties* : La classe C.
 - 4: **Initialisation** Chaque *individu* forme une *classe*
 - 5: **Calculer** la matrice des distances des classes deux à deux
 - 6: **Tant que**(Il y a plus qu'une classe)
 - 7: **Regrouper** les deux classes les plus proches au sens de la distance choisie
 - 8: **Mettre à jour** le tableau de distances en remplaçant les deux classes regroupées par la nouvelle et en calculant sa « distance » avec chacune des autres classes
 - 9: **Fin**
-

Le résultat de la CHA est représenté sous forme d'un graphe qui résume le processus d'agrégation qu'on appelle dendrogramme. Le dendrogramme est une représentation graphique, sous forme d'arbre binaire, il représente des agrégations successives jusqu'à la réunion en une seule classe de tous les individus. La hauteur d'une branche est proportionnelle à l'indice de dissemblance ou distance entre les deux objets regroupés, il représente la perte d'inertie d'interclasse (entre classes) qui est un critère à optimiser dans la classification en général. Le dendrogramme est utilisé pour déterminer le choix optimal du nombre de classes.

Pour déterminer le nombre optimal de classes on procède en coupant les branches avant qu'elles ne soient trop longues c'est à dire qu'il faut couper avant que la perte d'inertie interclasse soit trop élevée. Le nombre de classes obtenues dépend de l'emplacement de la coupe.

Les méthodes hiérarchiques sont déjà utilisées dans le diagnostic des machines tournantes soit d'une manière directe (c'est à dire pour la classification) (KSI⁺14) (ASK13) ou indirecte soit pour préparer les données à la classification durant l'apprentissage pour des méthodes supervisées soit pour trouver le nombre optimal de classes (SRSS11)et dans ce cas ces méthodes sont combinées à des méthodes de classification automatique qui nécessitent la connaissance *a priori* du nombre de classes.

Les méthodes de classification par partition

2. MÉTHODES DE RECONNAISSANCE DE FORMES DANS LE CADRE DE DIAGNOSTIC DES MACHINES TOURNANTES

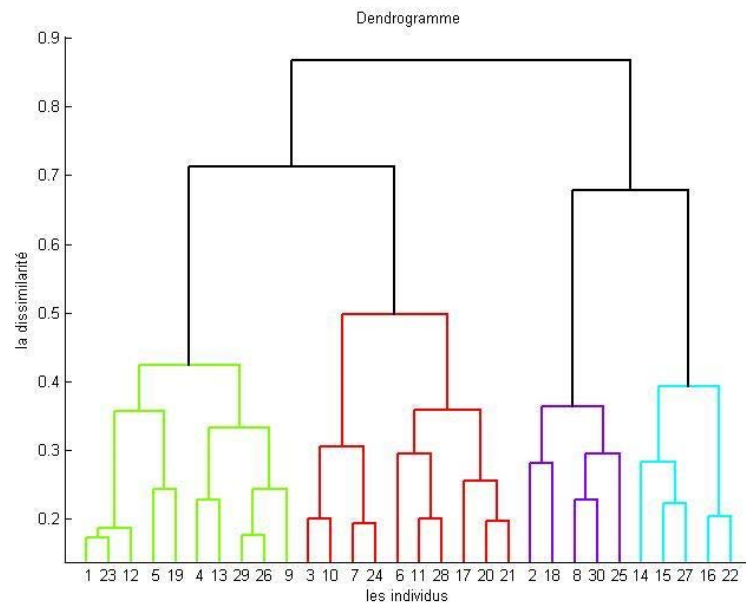


FIGURE 2.7 – Un exemple de dendrogramme (CHA)

La classification par partition (CPP) construit directement une partition contrairement aux méthodes de classification hiérarchiques qui construisent les classes progressivement.

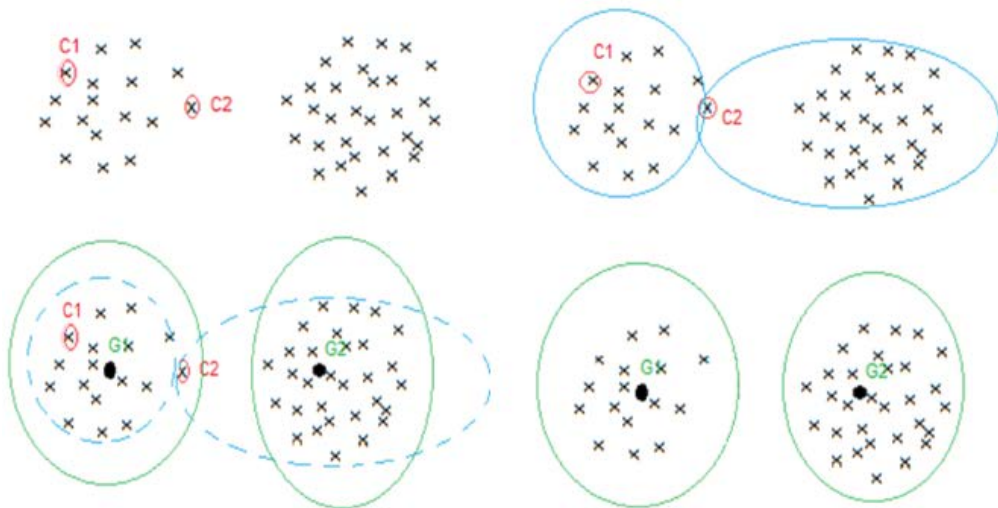


FIGURE 2.8 – méthode des centres mobiles

La CPP partitionne les données initialement en k classes où chaque classe contient au moins un individu. Chaque individu appartient à au moins une classe (dans le cas de

partitionnement floue) ou seulement à une classe (partitionnement dure). Puis la CPP cherche à améliorer la qualité de cette partitionnement en réattribuant les individus d'une classe à une autre afin d'optimiser une fonction objective qui dit qu'une classe doit être homogène et que les classes doivent être différentes. l'algorithme général de classification est décrit dans 2. La fonction objective la plus utilisée est définie par l'équation (2.13) :

$$F = \sum_{i=1}^k \sum_{x \in C_i} d^2(x, m_i) \quad (2.13)$$

Avec k le nombre de classe, x l'élément à classer et m_i le centre de la classe C_i . Cette fonction permet de calculer l'homogénéité de chaque classe.

Algorithm 2 L'algorithme général de classification par partition

Procédure : *Classification par partition*

2: *Entrée* : les individus, k le nombre de classes à trouver

Sorties : Ensemble de k classes

4: **Initialisation** On choisit **aléatoirement** k individus comme centres initiaux des classes.

Attribuer à chaque classe les objets qui sont proche d'elle

6: **Tant que**(il y a des individus qui change de classe)

Calculer les centres d'inertie de chaque classe

8: **Mettre à jour** les classes en redistribuant les objets dans la classe qui leur est la plus proche en tenant compte des nouveaux centres calculés de à l'étape précédente

Fin

L'algorithme le plus connu des méthodes de partitionnement est celui de K-moyennes ($K - means$). Cet algorithme réalise une partition stricte (« dure »), c'est-à-dire que chaque objet n'est assigné qu'à une seule classe. L'algorithme est itératif, son principe de partitionnement consiste à classer un ensemble d'éléments $M = m_1, m_2, \dots, m_n$ dans un nombre K de classes (« clusters »), K est spécifié par l'utilisateur. La partition finale de K-means est faite de telle façon que les éléments à l'intérieur d'une classe sont les plus semblables possibles et les plus distincts des éléments appartenant à d'autres groupes.

L'algorithme de $K - means$ est réalisé en deux étapes : d'abord, il faut définir les K centres ou prototypes de chaque classe (généralement le choix initial des centres est arbitraire). Chaque point ou objet est ensuite affecté au centre le plus proche. Chaque

2. MÉTHODES DE RECONNAISSANCE DE FORMES DANS LE CADRE DE DIAGNOSTIC DES MACHINES TOURNANTES

groupement de points associé à un centre devient une classe. Le centre de gravité de la classe récemment formée est mis à jour à chaque itération. Les affectations et les mises à jour sont répétées jusqu'à ce que les modifications des points n'entraînent pas de changements de classes ou jusqu'à ce que les centres de gravité n'évoluent plus.

Les avantages des méthodes de partitionnement et plus particulièrement *K - means* sont multiples ; la complexité de cette méthode est linéaire, c'est-à-dire que son temps d'exécution est proportionnel au nombre n d'individus, ce qui la rend applicable à des grands volumes de données. Cette méthode est facile à mettre en œuvre et les différentes classes obtenues sont remises en cause à chaque itération. Par contre l'algorithme de *K - means* est instable ; sa partition finale dépend essentiellement du choix arbitraire initial des centres. Il est incapable de trouver des classes d'une forme autre que la sphère, le nombre K de classes doit être connu a priori. Si ce nombre ne correspond pas à la configuration véritable du nuage des individus, la méthode de classification ne donne plus de résultats satisfaisants (Tuf12, TSK05).

La méthode *K - means* est déjà utilisée dans la cadre de la surveillance des machines tournantes. Yiakopoulos et al. (YGA11) ont employé une version modifiée de *K - means*, combinée avec des indicateurs fréquentiels et utilisée avec différentes métriques de distance (cosine, corrélation, euclidienne et cityblock) pour classifier les caractéristiques fréquentiels des signaux vibratoires en 3 classes : sain, défaut bague intérieure, défaut bague extérieure.

Les auteurs dans (MEA14) proposent de combiner *K - means* avec les algorithmes génétiques (AG) afin de surmonter les inconvénients de la classification par *K - means* et ceci en appliquant d'abord l'AG pour prétraiter les données avant de les classifier. Cette combinaison a permis d'augmenter la précision de la classification qui a passé de (64% à 97%) à (91% à 100%). Les distances interclasses ont été également améliorées avec cette méthode. Enfin, cela a permis de classifier les données en quatre classes : sain, défaut bague intérieure, défaut bague extérieure, défaut élément roulant.

Les auteurs dans (JCYD13) ont proposé une autre manière pour surmonter les limitations de *K - means* en employant une fonction noyau pour projeter les données dans un

nouvel espace avant d'appliquer l'algorithme $K - means$. Cette variante de $K - means$ a permis de diminuer le nombre d'itérations avant convergence et d'augmenter légèrement la précision de la méthode classique qui est passé de 91% à 95%. Les auteurs ont utilisé la variante « Improved kernel $K - means$ » pour classifier les données en quatre classes : sain, défaut bague intérieure, défaut bague extérieure, défaut d'élément roulant.

Les méthodes de classification par densité

Les méthodes de la classification basées sur la densité sont des méthodes hiérarchiques dans lesquelles les classes sont considérées comme des régions de haute densité qui sont séparées par des régions de faible densité. La densité est représentée par le nombre d'individus de l'ensemble des données. C'est pourquoi ces méthodes sont capables de chercher des classes de formes arbitraires. Elles ne travaillent que dans un espace métrique.

Le principe de ces méthodes est de caractériser une classe comme étant une zone où le nombre de données initiales est plus important qu'ailleurs.

Il y a deux approches dans ce type de méthodes :

- Approche basée sur la connexité de densité (DBSCAN, OPTICS, DBCLASD)
- Approche basée sur la fonction de densité (DENCLUE)

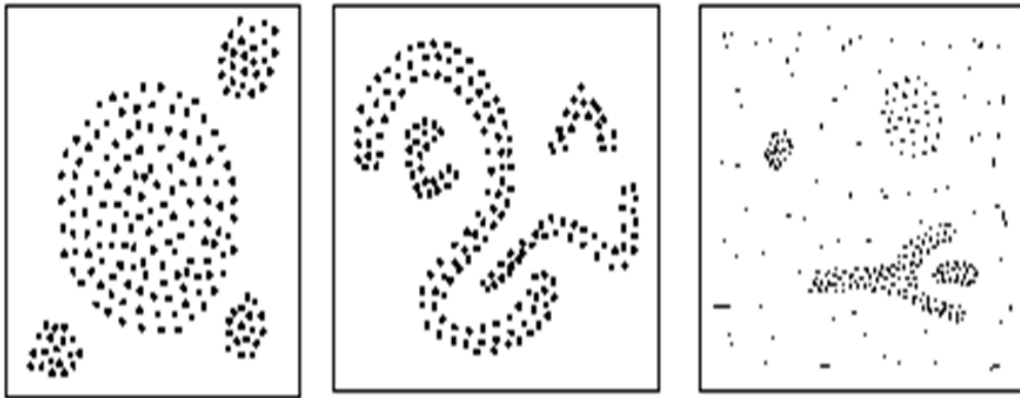


FIGURE 2.9 – des clusters avec des densités différentes (EKSX96)

L'algorithme DBSCAN (Density-Based Spatial Clustering of Applications with Noise) est un algorithme classique de la famille des méthodes de classification par densité, il

2. MÉTHODES DE RECONNAISSANCE DE FORMES DANS LE CADRE DE DIAGNOSTIC DES MACHINES TOURNANTES

présente l'intérêt de trouver lui-même l'évolution du nombre de classes. Il permet également de gérer tout type de données et de tenir compte des données aberrantes qui ne sont pas affectées aux classes identifiées (CN04, Tuf12). Il a aussi peu de paramètres à régler et n'est pas très sensible au bruit (LZW07, CL11, SZ13). L'utilisation de *DBSCAN* nécessite la définition de deux paramètres *Eps* et *MinPts*. *Eps* définit le rayon de voisinage ou la distance maximale entre deux points d'une même classe. *Minpts* définit le seuil de densité qui correspond au nombre minimal d'objets dans le voisinage d'un point.

Dans la classification par *DBSCAN* il faut définir quelques notions :

- Le voisinage *Eps* d'un point p dans une base D est noté par $Eps(p)$ est définie par $Eps(p) = \{q \in D \mid distance(p, q) \leq Eps\}$
$$N_p(q) = \begin{cases} 1 & \text{si } dist(p, q) \leq Eps \\ 0 & \text{sinon} \end{cases}$$
Avec $N_p(q)$ est le degré d'appartenance de point q à l'ensemble des points au voisinage de p .
- L'accessibilité par densité est la notion de base de DBSCAN. Un point p est accessible par densité de q si deux conditions sont satisfaites : (i) $p \in Eps(q)$ et (ii) $|Eps(q)| \geq Minpts$.

L'algorithme *DBSCAN* a pour but d'identifier les partitions et le bruit dans une base de données spatiale. Idéalement, il faudrait connaître les paramètres appropriés *Eps* et *MinPts* de chaque cluster et un point de chacun des clusters respectifs. Tous les points des données de densité accessibles peuvent être retrouvés à partir de ces paramètres. Il n'est pas facile d'obtenir ces informations à l'avance pour chaque partition de la base de données (voir algorithme 3).

Il existe une heuristique simple et efficace pour déterminer les paramètres *Eps* et *MinPts* des clusters les plus minces (BMN89). Ces paramètres de densité sont des bons candidats pour les paramètres globaux spécifiant des densités les plus basses qui ne sont pas considérées comme du bruit.

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

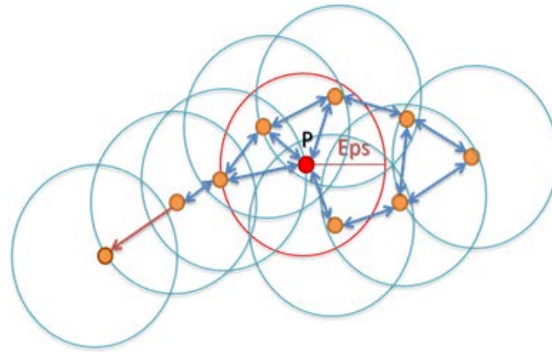


FIGURE 2.10 – Création d'un cluster avec DBSCAN

de p et *DBSCAN* visitera le prochain point de la base de données. Grâce à l'utilisation des valeurs globales *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 ensembles 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 (EKSX96).

Algorithm 3 L'algorithme de la classification par *DBSCAN*

Procédure *DBSCAN*

Entrées : les individus

3: *Sorties* : Ensembles de classes

Initialisation Choisir arbitrairement un point p

Tant que(il reste des points non traités)

6: **Chercher** tous les points densité accessible de p en utilisant *Eps* et *MinPts*

Si(p est un point de type *centre*) un cluster est formé

Sinon

9: **si**(p est point de type *frontière* et aucun point est accessible par densité de p)
le point p est considéré comme un bruit et *DBSCAN* passe au point suivant dans la base de données.

Fin

Une revue de la littérature a montré que *DBSCAN* n'a jamais été utilisé auparavant dans le diagnostic des roulements.

2. MÉTHODES DE RECONNAISSANCE DE FORMES DANS LE CADRE DE DIAGNOSTIC DES MACHINES TOURNANTES

2.3.6 Post-traitement

Cette étape consiste à évaluer les performances de la classification en utilisant des différents taux d'erreur, le taux de précision, le temps de calcul, la complexité de l'algorithme, etc. Tous ces outils vont permettre de déterminer s'il faut changer les paramètres de la méthode de classification employés ou changer la méthode de classification.

Il existe plusieurs méthodes pour évaluer une méthode de classification, ces méthodes diffèrent d'une famille de méthode de classification à une autre, il existe des méthodes développées spécialement pour les méthodes de classification supervisées et spécialement pour des méthodes de classification non supervisées.

A. Mesures de performance de la classification supervisée

Plusieurs outils sont utilisés pour mesurer les performances d'une méthode de classification. Le choix d'un outil dépend de plusieurs paramètres comme le type d'apprentissage, le coût de la mauvaise classification ou autres.

A.1. Taux de bonne classification sans coûts

La première mesure la plus intuitive est le taux de bonne classification simplifié (tbc_s) ou sans coût. Il s'agit d'un indicateur qui permet d'évaluer les performances d'un système de classification. Cette valeur, simple à calculer, correspond au nombre d'éléments correctement identifiés par le système. La définition du taux de bonne classification sans la prise en compte du rejet est définie par l'équation (2.14) :

$$tbc_s = \frac{\text{Nombre d'éléments correctement classifiés}}{\text{Nombre total d'éléments classifiés}} \quad (2.14)$$

On obtient alors le taux d'erreur te_s par

$$te_s = 1 - tbc_s \quad (2.15)$$

A.2. Taux de rejet

Le taux de rejet, t_r , définit le taux d'éléments mal classifiés. Il mesure le nombre d'éléments sur lesquels le système de classification n'a pas pris de décision.

$$t_r = 1 - tbc_s - te_s \quad (2.16)$$

A.3. Taux de bonne classification avec coûts : *tbc*

Généralement la conséquence d'une mauvaise classification d'un objet dépend de la classe dans laquelle l'objet a été mal classifié. Prenons l'exemple de deux classes, une qui nous indique qu'un roulement est défectueux et l'autre qui nous indique que le roulement est sain. La classification des paramètres extraits d'un roulement sain dans la classe « défectueux » va générer une alarme qui va entraîner le déplacement d'un technicien de maintenance et l'arrêt de la machine entraînant un coût financier. Par contre la classification des paramètres correspondants à un roulement défectueux à la classe « sain » va ralentir le processus de diagnostic mais il ne va pas entraîner une perte aussi grave que le premier cas. Dans ce cas, la classification est associée à un coût.

Le taux de bonne classification avec coût permet d'évaluer les performances d'une méthode de classification avec la prise en compte de la répartition des classes et également des coûts de bonne et mauvaise classification.

Pour un système à C classes, nous définissons les indices $i, j \in 1, \dots, C$; ω_i représentant l'étiquette de la classe et $N(\omega_i)$ le nombre d'éléments de la classe i présents dans la base. Afin de résumer au mieux les résultats de la classification, nous utilisons une matrice dite de confusion qui met en relation les décisions prises par le classifieur et les étiquettes des exemples. Associée à cette matrice, nous définissons la matrice des coûts qui fait correspondre à chaque élément de la matrice de confusion un coût. Nous introduisons d'une part, la notation $\varepsilon_{i,j}$ qui correspond au nombre d'éléments étiquetés de la classe i et identifiés comme des éléments de la classe j . Le terme de score est également utilisé pour définir cette valeur. D'autre part, nous définissons $Cost_{i,j}$ comme le coût associé à $\varepsilon_{i,j}$ [rapport stage erreur de classification].

$$tbc = \sum_{i=1}^C \left[\frac{\varepsilon_{i,i} Cost_{i,i} + \sum_{j=1, j \neq i}^C \varepsilon_{i,j} Cost_{i,j}}{N_{\omega_i}} \right] * \sum_{i=1}^C N_{\omega_i} * \frac{1}{C} \quad (2.17)$$

Dans le cas où le coût est inconnu ou il est difficile à déterminer, on peut employer d'autres méthodes d'évaluation (AH99)

2. MÉTHODES DE RECONNAISSANCE DE FORMES DANS LE CADRE DE DIAGNOSTIC DES MACHINES TOURNANTES

B. Mesures de performance des méthodes de classification non supervisées

Il existe plusieurs outils qui permettent d'évaluer les performances d'une méthode de classification non supervisée. Les plus utilisés sont les outils qui sont basés sur la distance. Les indices inertiels (LMF79) sont les plus connus et les plus utilisés pour évaluer la qualité d'une classification (GCLL10).

B.1. Inertie intra-classes

L'inertie intra-classes permet de mesurer le degré d'homogénéité entre les objets appartenant à la même classe.

$$Intra = \frac{1}{n} \sum_{C=1}^P \sum_{i=1}^{n_C} d^2(X_i, C_c) \quad (2.18)$$

Avec n_c le nombre d'éléments de la classe C et C_c son centre de gravité, n est le nombre total des éléments de la partition.

B.2. Inertie inter-classes

L'inertie inter-classes mesure le degré d'hétérogénéité entre les classes. Elle calcule les distances entre les centres de gravités des classes de la partition.

$$Inter(C_i) = \frac{1}{n} \sum_{i=1}^P n_C d^2(C_c, c_G) \quad (2.19)$$

Avec C_c le centre de la classe C , c_G est le centre du tout l'ensemble de points et P le nombre de classes dans la partition. Plus les données à l'intérieur des classes sont homogènes, plus leurs distances par rapport au point représentant la classe sont faibles. Par conséquent, une valeur faible de l'inertie intra-classes décrit une homogénéité des données à l'intérieur des classes. Plus les classes sont hétérogènes entre elles, plus les distances entre les points représentant les profils des classes sont élevées. Une valeur élevée de l'inertie interclasses traduit une hétérogénéité entre les classes.

B.3. Indice de Dunn

L'indice de Dunn permet de trouver la distance minimale qui sépare deux classes dans la partition tout en tenant compte de la distribution des éléments à l'intérieur des classes. Plus cette distance est grande meilleure est la partition.

$$Dunn = \frac{\min_{1 \leq i < j \leq n} \Delta_{i,j}}{\max_{1 \leq k \leq n} \Delta'(k)} \quad (2.20)$$

Avec $\Delta_{i,j}$ la distance entre les classes i et j et $\Delta'(k)$ est la distance intra-classe de la classe k

$$\Delta'(c) = \frac{1}{n_c} \sum_{i=1}^{n_c} (X_i - C_c) \quad (2.21)$$

$$\Delta(i, j) = d(C_i, C_j) \quad (2.22)$$

Avec $d()$ est la distance choisie dans la métrique de la méthode de classification et C_i , C_j sont les centres des classe i et j .

B.4. Indice de Davies-Bouldin

L'indice de Davies-Bouldin traite chaque classe individuellement et cherche à mesurer à quel point elle est similaire à la classe qui lui est la plus proche. L'indice DB est formulé de la façon suivante (Elg07)

$$DB = \frac{1}{P} \sum_{i=1}^k \max_{1 \leq j \leq k, j \neq i} \frac{\Delta'(i) + \Delta'(j)}{\Delta(i, j)} \quad (2.23)$$

B.5. Indice de silhouette

L'indice de silhouette est défini par Rousseeuw (Rou87) pour tout individu X_i de l'ensemble X par la formule suivante

$$\forall X_i \in X, S(X_i) = \frac{b(X_i) - a(X_i)}{\max(a(X_i), b(X_i))} \quad (2.24)$$

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 $C(X_i)$

$$a(X_i) = \frac{1}{n_{c_i}} \sum_{X_j \in C(X_i), X_j \neq X_i} d(X_j, X_i) \quad (2.25)$$

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 .

$$b(X_i) = \min_{C \in P, C \neq C(X_i)} d(X_i, C) \quad (2.26)$$

Où

2. MÉTHODES DE RECONNAISSANCE DE FORMES DANS LE CADRE DE DIAGNOSTIC DES MACHINES TOURNANTES

$$d(X_i, C) = \frac{1}{n_c} \sum_{X_j \in C} d(X_i, X_j) \quad (2.27)$$

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 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.

2.3.7 Règles de décision

La règle de décision utilisée pour le classement des nouvelles observations ainsi que la fonction de décision (ou fonction discriminante), notée $\hat{u}$, définit une partition de l'espace de représentation en autant de régions que de modes de fonctionnement et permet d'associer toute nouvelle observation X_i à l'un des K modes de fonctionnement :

$$\mathbf{R}^d \rightarrow \{1, \dots, k\} \quad (2.28)$$

$$X_i \rightarrow \hat{u}(X_i) \quad (2.29)$$

$$\text{Avec } \hat{u}(X_i) = K \text{ Si } X_i \in C^k \quad (2.30)$$

Lors de la définition de cette règle de décision, on parle couramment de frontières de décision entre les modes de fonctionnement. L'obtention de ces frontières peut se faire suivant deux approches : analytique ou statistique.

A. Approche analytique

On détermine directement la fonction discriminante en estimant les paramètres d'une fonction mathématique de manière à séparer au mieux les modes de fonctionnement, c'est à dire à minimiser la probabilité de mauvais classement. Le choix du modèle pour la fonction mathématique dépend de la complexité de la frontière de décision à mettre en place. En général, on commence par l'utilisation des fonctions linéaires, puis si les

performances de classement ne sont pas satisfaisantes, on fait appel à des fonctions quadratiques, ou encore à des modèles neuronaux.

B. Approche statistique

On considère que chaque observation X_i de $\mathbf{R}^D$ suit, dans chaque mode de fonctionnement C_k , une loi de probabilité notée $f(X_i/C^k)$. Les concepts issus de la théorie statistique de la décision et de l'analyse discriminante peuvent alors être utilisés pour établir les frontières de décision. Il existe deux types de méthodes de modélisation : (i) **les modèles paramétriques** pour lequel la fonction de probabilité est supposée connue pour chaque mode de fonctionnement (classes). La règle de décision optimale est alors définie à partir de la théorie Bayésienne, (ii) **les modèles non paramétriques** pour lesquels les lois de probabilités sont inconnues. Dans ce cas la fonction de probabilité est estimée à l'aide de méthodes comme les noyaux de Parzen, où les probabilités sont directement estimés a posteriori à l'aide de la méthode des K plus proches voisins.

Il existe un troisième type de modèle qui est un mélange de deux premiers appelé un **modèle de mélange**. La méthode la plus utilisée pour estimer les paramètres de ce modèle est Expectation-Maximization (EM).

A l'issue de cette étape (règle de décision) on définit des régions de l'espace où chaque région correspond à un mode de fonctionnement et la nature des frontières qui séparent ces régions. En définissant les règles de décisions on peut faire face à des cas où les formes à classer se trouvent trop près des frontières de décision. Ces formes peuvent engendrer une ambiguïté au système de diagnostic par RdF. Dans le cas où les formes se trouvent dans une région très éloignée de tous les modes de fonctionnement, l'affectation de ses formes au mode le plus proche inclue généralement une erreur de classification. Dans ce cas le système doit introduire la notion de rejet pour des raisons d'ambiguïté ou de distance.

2.4 Limitations de méthodes de reconnaissance des formes employées dans le cadre de diagnostic

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. En analysant cette littérature, on a observé que les méthodes de RdF employées dans ce domaine partagent deux points en commun : Premièrement elles étaient toutes appliquées sur une base de signaux enregistrés auparavant. Deuxièmement les signaux formant la base de données ont été extraits soit du même roulement avec différents états de dégradation (le même défaut mais avec des tailles différentes) ou différents roulements avec différents défauts (défaut bague extérieure, défaut bague intérieure, défaut de bille).

L'étude bibliographique nous a permis aussi de conclure que les méthodes de classification dans le diagnostic des machines ont été spécialement utilisées pour leur qualité de « séparation automatique » des données différents 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 connaissance d'une base 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.

Avec la RdF classique, la modélisation des modes de fonctionnement ne permet pas de quantifier les évolutions des observations, or un système de diagnostic doit pouvoir les détecter. 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 (ME10), les classes peuvent être créées si nécessaire, ou supprimées 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.

Un autre point essentiel est que les classes d'un système en évolution sont forcément dynamiques ; leurs caractéristiques changent au fil du temps, de façon lente et progressive, ou de façon abrupte. Le changement dans le comportement des classes est directement lié à l'état du système de fonctionnement. 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.

2. MÉTHODES DE RECONNAISSANCE DE FORMES DANS LE CADRE DE DIAGNOSTIC DES MACHINES TOURNANTES

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

Processus de développement des systèmes de reconnaissance de formes dynamiques pour le diagnostic des roulements

3. PROCESSUS DE DÉVELOPPEMENT DES SYSTÈMES DE RECONNAISSANCE DE FORMES DYNAMIQUES POUR LE DIAGNOSTIC DES ROULEMENTS

Introduction

Le chapitre précédent a montré qu'il existe plusieurs types de méthodes de RdF utilisées pour le diagnostic, mais aucune de ces méthodes n'a été utilisée pour le suivi. Cependant, le suivi en continu du fonctionnement des systèmes industriels, tels que les machines tournantes permet d'améliorer leur productivité et de faire décroître leur coût de maintenance en réduisant leur temps d'arrêt. Quand un défaut est détecté, le système de suivi exécute une procédure pour étudier la gravité de ce défaut et programme un arrêt pour changer le composant défaillant. Plusieurs méthodes peuvent être utilisées pour réaliser le suivi d'un système de production. Le choix de la méthode dépend de plusieurs facteurs comme :

- La dynamique du système (discrète, continue ou hybride),
- Les contraintes d'environnement (en ligne pour une réaction souhaitée en temps réel, ou hors ligne si la réaction peut être prise après un certain temps),
- La représentation des informations (quantitative et/ou qualitative),
- La complexité du système (grande ou simple),
- L'information disponible sur le système (structurelle, analytique, connaissances heuristiques, etc.),
- ...

Le suivi d'un composant mécanique implique certaines contraintes au processus de suivi. En effet ce processus doit apprendre à gérer *(i)* le fait qu'au début, il ne dispose pas d'une base d'apprentissage pour modéliser le système, *(ii)* les observations qui (les caractéristiques extraites du signal vibratoire) arrivent au fur et à mesure dans le temps, *(iii)* le (ou les) composant(s) qui va changer de mode de fonctionnement pour passer d'un mode normal à un mode défectueux ou un autre mode normal, ce qui implique que les observations vont connaître un changement dans le temps.

Le processus de suivi doit être capable de gérer toutes ces contraintes et de déterminer l'état du système. Les méthodes classiques de reconnaissance des formes doivent être adaptées pour effectuer le suivi d'un ou plusieurs composants mécaniques d'une machine de production.

Ce chapitre présente en détails les contraintes rencontrés lors de la conception d'une méthode RdF pour le suivi de roulements, les solutions trouvées dans la littérature pour des cas similaires, le choix de la classification dynamique pour suivre les évolutions des roulements et les étapes à suivre pour concevoir une méthode RdF qui respecte toutes les exigences citées. La première partie présente les différentes formes d'évolutions des classes que l'on peut rencontrer et les techniques de détection de ces évolutions. La deuxième partie est consacrée à la présentation des méthodes de classification dynamique et la troisième partie traite le processus d'adaptation des méthodes de RdF classique pour remplir les exigences dans le cadre du suivi.

3.1 La reconnaissance de formes pour le suivi

3.1.1 Le contexte général

Une machine tournante est soumise à des évolutions qui induisent des changements de son comportement vibratoire. Ces changements peuvent être dus soit à un changement de ses paramètres (variation de charges, de vitesse, ou autres), soit à des dégradations, soit à des défauts affectant ses caractéristiques intrinsèques et son comportement. Les variations des caractéristiques intrinsèques des systèmes industriels sont généralement imprédictibles et non contrôlées au cours du temps, ce qui signifie qu'il est impossible de connaître leur évolution d'une manière déterministe et exacte.

3.1.2 Evolution des observations caractéristiques d'un système de production

L'évolution des caractéristiques d'un système de production traduisant le passage du système d'un mode de fonctionnement à un autre. Ce passage peut être lent ou rapide. L'évolution rapide ou dérive rapide est caractérisée par un glissement des observations (Une observation est un ensemble de caractéristiques d'un système) dans l'espace de représentation, l'évolution lente ou dérive lente est caractérisée par une évolution lente du mode de fonctionnement. Généralement il y a trois modes de fonctionnement : mode normale, mode dégradé, et mode panne. Le passage d'un mode à un autre peut se faire d'une manière brusque ou progressive. Ce passage peut être observé notamment lors de la présentation de nouvelles observations au module de classification. On distingue généralement trois types d'évolutions. La première traduit un saut des observations en dehors

3. PROCESSUS DE DÉVELOPPEMENT DES SYSTÈMES DE RECONNAISSANCE DE FORMES DYNAMIQUES POUR LE DIAGNOSTIC DES ROUEMENTS

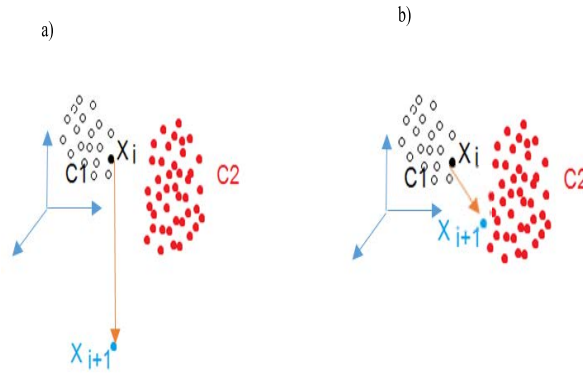


FIGURE 3.1 – Saut d’observation : a) saut vers une région non définie b) saut vers une classe définie

de la classe actuelle, la deuxième traduit un éloignement progressif des observations en dehors de la classe actuelle et la troisième considère diverses évolutions possibles d’une classe initiale (déplacement, rotation, fusion scission ...).

A. Saut d’observation

Soit X_t un vecteur d’observation pris à l’instant t . Il peut arriver qu’à l’instant $t + 1$ l’observation X_{t+1} présente des caractéristiques très différentes de X_t . elle va donc être positionnée dans une région de l’espace de représentation éloignée des observations précédentes. Cette observation peut soit rejoindre une classe déjà identifiée si elle respecte sa fonction d’appartenance ou être affectée à la "classe rejet" d’appartenance si elle est située loin de toutes classes (Fig. 3.1).

L’apparition d’une observation dans une région non définie de l’espace de représentation peut avoir deux significations : soit il s’agit d’une observation aberrante, soit la classe n’a pas été caractérisée initialement. La méthode de RdF doit être en mesure de décider s’il faut créer une nouvelle classe ou considérer l’observation comme un bruit si une classe est créée. Une mise à jour de classifieur est donc nécessaire.

B. Dérive d’observation

Un autre phénomène peut être observé lors de la surveillance d’un système évolutif quand les observations commencent à s’éloigner progressivement d’une classe connue vers une

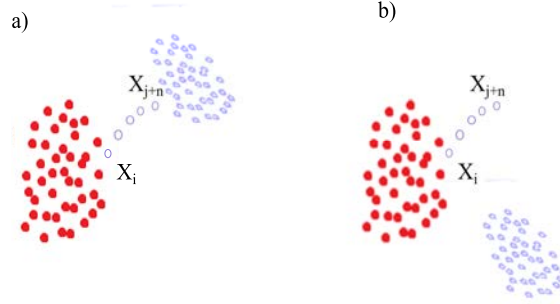


FIGURE 3.2 — Dérive d’observation : a) dérive vers une classe définie b) dérive vers une région non définie

autre classe connue ou vers une région de l’espace de représentation non identifié. C’est ce qu’on appelle une dérive d’observation (Fig. 3.2).

Dans ces situations, le problème consiste à détecter de manière précoce l’évolution des observations ainsi que l’occurrence d’une stabilisation éventuelle dans une région identifiée ou non de l’espace de représentation. La difficulté qu’une méthode RdF doit y faire face dans cette situation c’est de distinguer entre un état de stabilisation et un état transitoire. En effet le classifieur ne doit pas établir des mises à jour que s’il s’agit d’un état de stabilisation dans une nouvelle région de l’espace de représentation.

3.1.3 Evolutions d’une classe

Les observations qui caractérisent le même mode de fonctionnement sont attribués à la même classe. Il peut arriver que certains phénomènes physiques altèrent les paramètres du modèle (fonction d’appartenance) de la classe. Par exemple les observations arrivent au fil de temps et se positionnent entre les frontières de deux classes en rendant la région de représentation comme une seule classe au lieu de deux. Dans ce cas, la méthode RdF doit mettre à jour son classifieur pour fusionner les deux classes en une seule. Il se peut également qu’une classe perd son importance (classe transitoire) et dans ce cas le classifieur doit supprimer cette classe, etc.

Les classes comme les observations peuvent subir des évolutions (fusion, scission, déplacement, rotation, suppression, etc.) dans le temps que la méthode de RdF doit prendre en considération.

3. PROCESSUS DE DÉVELOPPEMENT DES SYSTÈMES DE RECONNAISSANCE DE FORMES DYNAMIQUES POUR LE DIAGNOSTIC DES ROUEMENTS

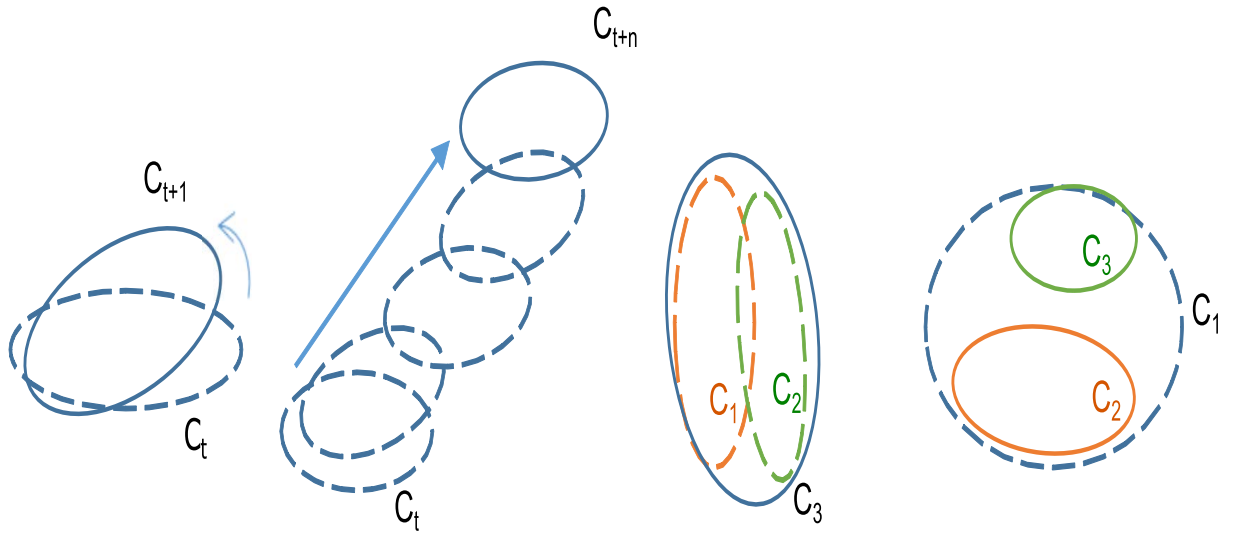


FIGURE 3.3 – Exemples des évolutions des classes : a) rotation de la classe C . b) déplacement de la classe C . c) fusion de deux classes C_1 et C_2 en C_3 . d) scission de la classe C_1 en deux classes C_2 et C_3

L'efficacité d'une méthode de RdF est fonction de sa capacité à détecter les différentes évolutions des classes et à mettre à jour les paramètres de son classifieur.

3.1.4 Les différentes contraintes à gérer pour concevoir une méthode de RdF adaptée au suivi des composants mécaniques

Une méthode RdF développée pour le suivi doit faire face à trois types de contraintes ; des contraintes liées à la gestion de flux, des contraintes liées au suivi et des contraintes liées à la conception de la méthode.

A. Les contraintes liées à la de gestion de flux

Le fait de choisir que le système de suivi décide l'état de la machine en se basant sur des données recueillies instantanément nous impose la contrainte de gestion de flux (streaming). Plusieurs paramètres liés à ce problème doivent être pris en considération :

- Le temps : le classifieur ne dispose qu'une durée limitée et constante pour classer les données et se mettre à jour ;
- La nature des données : la lecture des observations enregistrées à un instant t n'est disponible qu'une seule fois et dans leur ordre d'arrivée ;

- La mémoire : la taille mémoire allouée à ce problème est généralement fixée *à priori* ;
- La précision : la méthode doit générer un modèle proche de celui qui aurait été généré s'il n'y avait pas eu de contrainte de flux ;
- L'accessibilité : cela concerne la possibilité d'interroger le modèle et de suivre les changements de concept à tout moment.

B. Les contraintes liées au suivi d'un composant

La méthode de RdF développée pour le suivi d'un composant mécanique doit également respecter les contraintes suivantes :

- La méthode doit être capable de définir l'état du composant à tout moment ;
- La méthode de RdF doit être capable de suivre dans le temps l'état du système et de détecter tout changement de comportement du système en particulier des dérives lentes qui sont des dérives difficilement détectables ;
- La méthode doit être capable d'identifier et de caractériser chaque mode de fonctionnement du système, elle doit être capable de préciser à l'utilisateur s'il s'agit d'un changement dû à un défaut ou de changement de paramètres intrinsèques (variation de vitesse, charges ou autres) ;
- La méthode doit être capable d'exploiter les évolutions des classes pour décider sur l'état interne de la machine en intégrant des indicateurs de classes (comme la vitesse de dérive, la densité de la classe ou autres) et de fournir des informations sur l'état futur du composant ;

C. Les contraintes liées à la conception de la méthode RdF

La méthode développée doit faire face à deux autres types d'exigences :

1. Des exigences liées à la qualité de l'algorithme d'apprentissage :
 - la rapidité d'apprentissage,
 - la complexité limitée,

3. PROCESSUS DE DÉVELOPPEMENT DES SYSTÈMES DE RECONNAISSANCE DE FORMES DYNAMIQUES POUR LE DIAGNOSTIC DES ROULEMENTS

- la rapidité et facilité de mise à jour,
 - la consommation en mémoire,
2. Des exigences liées à la pertinence du classifieur :
- la précision de la méthode de classification,
 - la simplicité (nombre de paramètres),
 - la rapidité de classification,
 - la compréhensibilité (la facilité d'interprétation des résultats),
 - la généralisation,
 - la sensibilité au bruit.

Le respect de ces contraintes nous garantit que la méthode développée sera capable de détecter et de suivre dans le temps le composant mécanique, en particulier le roulement.

3.1.5 Les caractéristiques des méthodes de RdF pour les systèmes évolutifs

Le développement d'une méthode de RdF capable d'effectuer le suivi d'un composant signifie que cette méthode permettra l'apprentissage en continu, qu'elle sera capable de suivre dans le temps les évolutions du système. Les connaissances déjà apprises lors de la phase d'apprentissage évoluent également dans le temps. Tout ceci implique que le classifieur doit être capable de s'adapter en permanence pour prendre en compte des évolutions qui traduisent la non-exhaustivité de la base d'apprentissage. Le classifieur doit donc faire face à deux challenges : premièrement l'adaptabilité permanente et deuxièmement la non-exhaustivité des connaissances.

A. L'adaptabilité du classifieur

La méthode incrémentale apporte une solution pour remédier au problème de l'apprentissage adaptatif. Les méthodes de classification incrémentale font référence à des algorithmes permettant d'entraîner un classifieur de manière incrémentale, c'est à dire qu'au fur et à mesure que de nouvelles informations (données) sont présentées à l'algorithme, celui-ci apprend et ce sans avoir besoin de réapprendre le modèle à partir de zéro ni même à avoir à stocker l'ensemble des données (SL11).

Il existe quelques méthodes de classification incrémentales qui étaient développées et appliquées sur des données reçues en flux (pas forcément pour la surveillance). Parmi ces méthodes, on peut citer : la méthode *IKNN* (incremental k-nearest neighbor) (FMA⁺10) qui est une version incrémentale de K plus proches voisins développée pour surveiller les données collectées à partir des capteurs des signaux physiologiques (*EEG*, *ECG*, *EMG*, etc.). La méthode *IFPM – BS* (incremental frequent pattern matching based on bit-sequence) (DYHR11) adopte la séquence de bits avec le mode incrémental pour compresser la base de données afin d'économiser l'espace de mémoire et le temps de calcul, en évitant la mise à jour de classifieur jusqu'à ce qu'un certain nombre de nouvelles opérations ont été insérées. On trouve également une version incrémentale des méthodes séparateurs à vaste marge *ISVM* (Incremental support vector machine) (LGK06), cette variante de *SVM* a été spécialement développée pour l'apprentissage en ligne. Il y a bien d'autres méthodes incrémentales qui ont été développées, cependant les limitations de ces méthodes résident dans le fait qu'elles ne permettent pas de prendre en considération la notion de dérive des formes et des classes (une information qui est très importante dans la surveillance) car elles ne permettent pas d'accorder plus d'importance ou de poids aux nouvelles données ou d'«oublier» les données qui ne sont plus informatives (Har10). Ceci limite l'efficacité des méthodes incrémentales dans le suivi des systèmes évolutifs.

B. La non-exhaustivité des données

Généralement, la décision de l'appartenance d'une observation par rapport à une classe dans la méthode de RdF peut se faire d'une façon binaire (c'est à dire que l'observation appartient à la classe ou non). Que cette observation soit au centre de la classe ou à la frontière, les méthodes de décision classiques jugent que les deux cas sont similaires. Or dans le cadre du suivi, une observation qui se situe à la frontière de la classe peut signifier que le composant commence à dériver vers un nouveau mode de fonctionnement ou il se peut que l'observation enregistrée appartient à un état transitoire entre deux modes de fonctionnement. La méthode de RdF pour le suivi doit donc être capable de différencier entre les deux cas.

Prenons l'exemple du suivi d'un composant mécanique : initialement, le composant est dans un bon état, puis au fur et à mesure son état se dégrade jusqu'à arriver à un état d'usure « inacceptable » où il faut le remplacer pour garantir le bon fonctionnement de

3. PROCESSUS DE DÉVELOPPEMENT DES SYSTÈMES DE RECONNAISSANCE DE FORMES DYNAMIQUES POUR LE DIAGNOSTIC DES ROULEMENTS

la machine. La transition entre les deux états (bon et dégradé) peut être lente ou rapide. La vitesse avec laquelle le composant se dégrade peut aider à prévoir une date à partir de laquelle le composant doit impérativement être changé. Pour caractériser cette transition on peut avoir recours à la théorie des ensembles flous, l'intégration de cette théorie à la décision va permettre à la méthode de Rdf de gérer les types d'incertitude intervenant dans la conception d'un système de suivi. Ce qui a donné naissance à la RdF floue utilisée dans le diagnostic des systèmes évolutifs.

Parmi ces méthodes on trouve Fuzzy pattern matching *FPM* (Mou02) développée pour des systèmes évolutifs. Cette méthode ne nécessite pas de connaissance *a priori*, elle utilise la notion de rejet pour créer des nouvelles classes. Lurette (Lur03) a proposé un algorithme de classification dynamique de données évolutives, *AUDyC* (AUto-adaptive and Dynamical Clustering) est un réseau de neurones évolutif qui permet de caractériser chaque mode de fonctionnement par une classe définie par des prototypes gaussiens qui sont mis à jour au fur et à mesure que les données arrivent. L'avantage de cette méthode est qu'elle ne nécessite pas de connaissances *a priori* du mode de fonctionnement du système. Une autre méthode basée sur le *k*-plus proches voisins floue FKNN appelée *Semi supervised fuzzy K nearest neighbor* a été proposée par Hartert (Har10) pour le diagnostic des systèmes évolutifs. Cette méthode utilise une version semi supervisée de *FKNN* dans laquelle l'auteur a intégré la notion de dynamique de classes. Ondel (Ond06) a proposé une autre méthode basée aussi sur le *K*-plus proches voisins. Cette fois, la méthode *KNN* a été combinée avec un maillage spécifique de l'espace de représentation, l'auteur a établi un maillage dans les zones des classes apprises autour desquelles, il a défini des zones de restriction ce qui lui a permis de créer ou d'éliminer des classes de la base de connaissances.

N.B : *En cherchant dans la littérature des solutions proposées pour des cas similaires, on rencontre souvent trois modèles de classifieurs (en ligne/ online, flux/stream et incrémental). Ces classifieurs qui partagent plusieurs notions communes, se départagent dans leur traitement de données. Bien que les classifieur en ligne et stream sont les deux conçus pour traiter des données de type flux, et ils doivent décider de l'appartenance de chaque point i arrivé avant l'arrivée de son suivant ($i + 1$) , la décision de l'appartenance de ce point par le classifieur en ligne est basée sur les $i - 1$ points qu'ils le précèdent. Les classifieurs en ligne ont donc accès aux anciens points. Alors que les classifieurs de type stream peuvent n'avoir accès qu'à(aux) point(s) actuel(s) (one pass). Les classifieurs*

stream par contre ne sont pas obligés de décider l'appartenance d'un point unique ils peuvent ne décider qu'après l'arrivée d'un ensemble de points (SAN⁺03).

3.2 La classification dynamique

3.2.1 Définitions et principes

Les données non-stationnaires sont des données issues d'un processus dont le comportement évolue dans le temps. La non-stationnarité entraîne une évolution des caractéristiques apprises de ces données. L'expression de classification dynamique est introduite pour qualifier les algorithmes développés pour la classification automatique de ces données. Lorsque les données d'une classe évoluent dans le temps, cette classe est dite évolutive (Har10).

Une partition dynamique représente l'ensemble des regroupements homogènes des données non-stationnaires. Elle est associée à un modèle de connaissances (c'est à dire classes) susceptible d'évoluer dans le temps en fonction de l'incorporation de nouvelles données. Selon le type du système étudié, les formes obtenues peuvent être statiques ou dynamiques (lorsque leurs caractéristiques évoluent avec le temps).

Selon Hartert (Har10), on peut faire face à quatre situations qui sont les combinaisons des deux types de formes et des deux types de classes. Chaque combinaison nécessite un traitement spécifique que nous présentons.

Reconnaissance statique des formes statiques :

Les données statiques sont des données qui conservent les mêmes caractéristiques statistiques au cours du temps. Les formes doivent être classées dans des classes statiques ne se déplaçant pas dans l'espace de représentation. Le classifieur utilise comme ensemble d'apprentissage ces formes et classes statiques et il reste inchangé dans le temps.

Reconnaissance statique des formes dynamiques :

3. PROCESSUS DE DÉVELOPPEMENT DES SYSTÈMES DE RECONNAISSANCE DE FORMES DYNAMIQUES POUR LE DIAGNOSTIC DES ROULEMENTS

Les formes dynamiques sont représentées par une séquence temporelle d'observations, c'est à dire un signal ou un trajectoire. La classification statique des formes dynamiques ne se base généralement pas sur la similarité de leurs indications ponctuelles (leurs valeurs à un instant donné) mais sur la similarité de leurs structures temporelles (leurs valeurs durant une période donnée). L'ordre d'arrivée des observations et leur évolution historique sont donc des informations importantes.

Reconnaissance dynamique des formes statiques :

Dans ce cas les classes ne sont plus statiques, mais dynamiques ; elles ont des caractéristiques qui évoluent au cours du temps. Le temps est donc un aspect important à prendre en compte pour empêcher la dégradation des performances du classifieur. Cette évolution des caractéristiques des classes peut se traduire par un déplacement, une scission, une fusion, une rotation, etc. Ces modifications sont dues à des changements progressifs ou brutaux des modes de fonctionnement d'un système.

Reconnaissance dynamique des formes dynamiques :

Ce dernier cas de classification correspond au cas des formes dynamiques dont les caractéristiques évoluent avec le temps. La structure d'une forme dynamique va donc changer progressivement ou brutalement. Ce changement de caractéristiques va mener ces formes dynamiques à changer de classe au cours du temps. Dans ce cas, la méthode de RdF doit détecter un changement dans la structure des formes. Ce changement peut correspondre à une détérioration progressive ou brutale du fonctionnement d'un système. La méthode de classification dynamique doit donc, dans ce cas, reconnaître que ce changement ne fait pas partie de la structure normale d'un signal mais qu'il y a une évolution du système.

La classification dynamique des données dynamiques est la plus adaptée au suivi de roulements car la signature vibratoire d'un roulement évolue dans le temps entraînant ainsi l'évolution des classes qui les représentent.

3.2.2 Aspect de la classification dynamique

Une méthode de classification est qualifiée de dynamique si elle remplit d'autres critères que les critères classiques liés à la classification ordinaire (statique), comme la détection des nouveautés, la gestion des problèmes liés aux phénomènes dynamiques dans un environnement non stationnaire. Ces phénomènes se manifestent par l'apparition ou la disparition des nouvelles classes, l'évolution de classes, la fusion, la scission ou autres. Les méthodes de classification doivent également faire face aux conséquences de l'apprentissage en ligne, c'est à dire que la méthode doit classer une observation dès qu'elle se présente en utilisant seulement les connaissances extraites des données déjà acquises et sans attendre que toutes les informations contenues dans les données futures soient disponibles. Elles doivent faire face aux problèmes liés à la mise à jour du modèle de classification qui se fait au fur et à mesure que les données se présentent. Cela requiert la mise en place de règles d'adaptation tenant compte des changements engendrés au cours de l'incorporation des nouvelles données dans les classes (Bou06).

A. Traitement de nouveautés

La machine tournante peut subir des changements de comportement vibratoire à cause de plusieurs facteurs comme la variation de la vitesse de rotation, la variation des charges, la présence d'un ou des défauts, etc. Ces changements se manifestent aussi dans l'espace de décision par l'apparition des « nouveautés » qui sont des événements non reconnus par le classifieur. Ces événements vont remettre en question l'adéquation des règles de décision et par conséquent les mises à jour nécessaires pour traiter des cas similaires dans le futur.

La détection des nouveaux événements est une capacité très importante pour toute méthode de RdF traitant des signaux. Compte tenu du fait que nous ne pouvons pas prévoir tous les cas possibles qu'on peut rencontrer durant la réception des nouvelles observations, il devient important que le classifieur sache différencier les informations qui sont connues et de celles qui sont inconnues. La détection des nouveautés est une tâche extrêmement difficile. C'est pour cette raison qu'il existe plusieurs modèles créés pour traiter des nouveautés et qui peuvent être plus ou moins performant selon l'application. Néanmoins il n'existe pas de modèle standard applicable dans tous les cas puisque le succès de la mé-

3. PROCESSUS DE DÉVELOPPEMENT DES SYSTÈMES DE RECONNAISSANCE DE FORMES DYNAMIQUES POUR LE DIAGNOSTIC DES ROULEMENTS

thode de détection ne dépend pas seulement du modèle mais aussi des caractéristiques statistiques des données traitées (MS03a).

Le classifieur peut agir comme un détecteur de nouveauté quand il rencontre des informations non connues. Cette méthode a été utilisée notamment dans le cas de la détection de défauts des composants mécaniques par Zhang *et al* (ZYZH06) et Wong *et al* (WJN06). Le traitement de nouveautés commence par la détection et puis le jugement si cette nouveauté est causée par des fluctuations dû au bruit ou à 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.

Pour détecter des nouveautés, plusieurs méthodes peuvent être utilisées comme, entre autres, des estimateurs de densité de probabilités complémentée par des seuils, la distance (méthodes de clustering), le calcul de caractéristiques spéciales comme l'indice de nouveauté, le taux de rejet, le taux d'erreur, les réseaux de neurones exploratoires (MS03a, MS03b), etc.

Une fois la détection confirmée, on passe à l'étape suivante qui se résume par l'évaluation de l'utilité de l'évènement détecté. Durant cette étape deux facteurs interviennent : l'avis de l'expert et les connaissances *a priori*. L'avis de l'expert permet d'établir des seuils à partir desquels un évènement est jugé intéressant ou aberrant du point de vue application (suivi et diagnostic). Si l'avis de l'expert est indisponible, on peut se baser sur des caractéristiques statistiques calculées à partir de la base d'apprentissage.

Pour surmonter le problème de détection de nouveautés on suggère l'utilisation des méthodes de classification floue et utiliser les pourcentages d'appartenance fournies par la méthode combinée avec un seuil afin de bien s'assurer qu'il ne s'agit pas de bruit.

B. Dynamique des classes

Les changements du comportement vibratoire d'une machine tournante peuvent se manifester par une évolution des classes comme l'apparition d'un nouveau nuage de point dans une zone non couvert par une classe, cette apparition peut entraîner la création d'une nouvelle classe, la fusion de deux classes trop proches, la scission d'une classe en deux ou la suppression d'une classe ancienne.

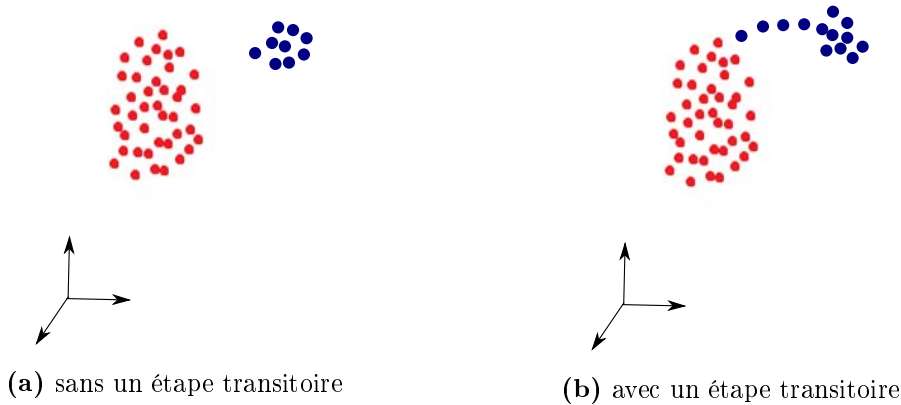


FIGURE 3.4 – Création d'une nouvelle classe

B.1. La création d'une nouvelle classe

Dans le cas où des nouvelles observations se regroupent dans une même région ou dans des régions différentes non attribuées à des classes connues et que ces observations apportent des informations utiles, ce regroupement peut s'expliquer par le passage du système d'un mode de fonctionnement à un autre. L'objectif des méthodes dynamiques est de détecter ce passage puis de créer des nouvelles classes qui définissent ce nouveau mode de fonctionnement. Pour atteindre cet objectif, il faut d'abord trouver des observations regroupées dans la même région dans l'espace de classification en utilisant un critère de similarité, s'assurer que ses observations ne sont pas le résultat d'un bruit en utilisant un seuil, puis de calculer le modèle qui correspond à la nouvelle classe créée en mettant en œuvre une procédure d'initialisation qui respecte la distribution des observations regroupées. Après la confirmation de la création de la nouvelle classe, la partition dynamique est actualisée pour inclure la nouvelle classe créée. Dans certaines situations, l'apparition d'une classe peut être précédée par une dérive rapide, l'analyse et l'exploitation de cette dérive est très intéressante pour des applications qui visent le pronostic.

B.2. Différentes formes d'évolution des classes

Les informations portées par les nouvelles observations enregistrées peuvent modifier la forme ou la position de la classe définie par la méthode de classification dynamique. Il est donc nécessaire que le classifieur prend en compte ces modifications ou évolutions et les intègre dans la construction des classes de la partition. L'incorporation des nouvelles observations peuvent se traduire par des modifications locales (grossissement, rétrécis-

3. PROCESSUS DE DÉVELOPPEMENT DES SYSTÈMES DE RECONNAISSANCE DE FORMES DYNAMIQUES POUR LE DIAGNOSTIC DES ROUEMENTS

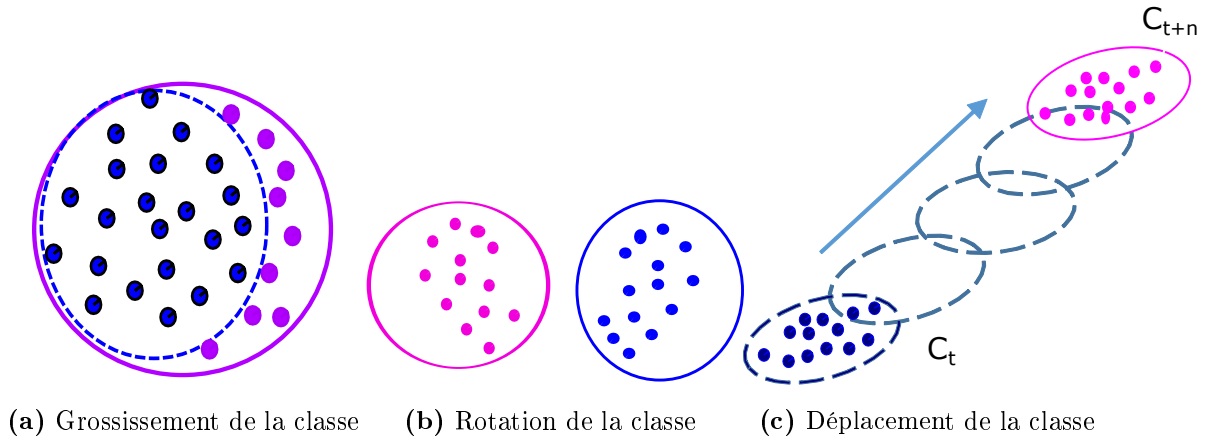


FIGURE 3.5 – Types d'évolutions d'une classe dynamique

sement ou rotation des classes) ou par des modifications avec glissement (déplacements des classes).

La méthode de classification dynamique doit être capable de détecter ces évolutions soit par l'utilisation des méthodes de modélisation adaptatives, soit par des méthodes récursives grâce à la mise à jour des règles de classification récursive afin de permettre l'ajout et la suppression des informations. L'évolution des classes peut créer des interactions entre les classes qui peuvent aboutir à la fusion, à la scission ou à la suppression des classes.

La fusion des classes

La fusion des classes se produit quand deux ou plusieurs classes initialement séparées se rejoignent et partagent les mêmes données. Ces données partagées sont généralement porteuses d'informations et aussi d'ambiguïté. Pour enlever cette ambiguïté, la méthode de classification dynamique doit fusionner les classes concernées en une seule.

La procédure de fusion de classes commence par la détection des observations partagées entre plusieurs classes en utilisant un critère de similarité puis s'assurer que ces données ne proviennent pas du bruit. Ensuite la méthode détermine le nouveau modèle résultant de la fusion.

Il existe plusieurs mécanisme de fusion parmi lesquels on peut choisir ou créer une nouvelle classe qui va regrouper des classes qui se chevauchent deux à deux ou plusieurs à la fois.

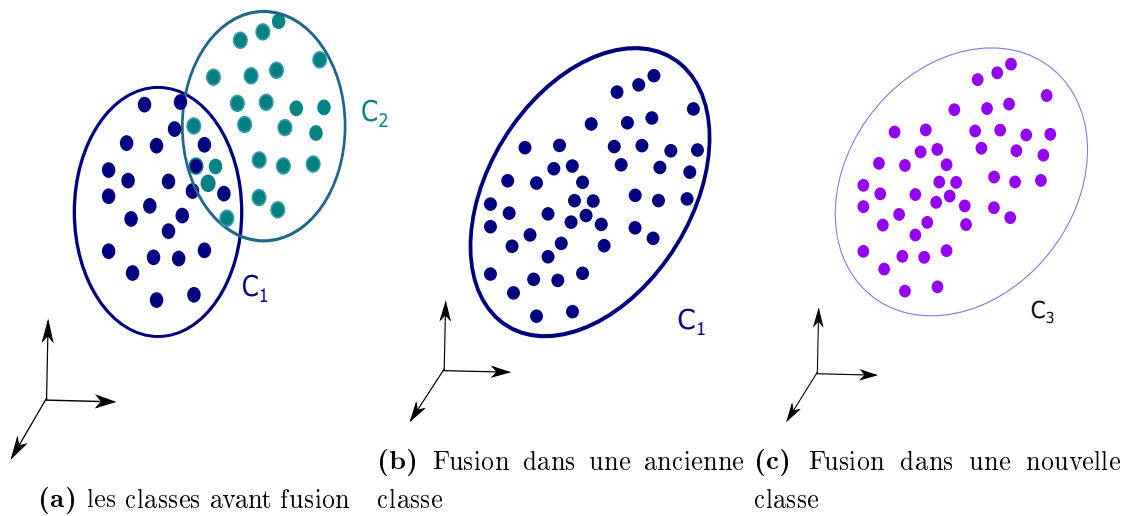


FIGURE 3.6 – Fusion des classes

On peut également attribuer les observations d'une des classes qui se chevauchent à la classe qui possède le plus d'informations. Dans les deux cas le déclenchement de processus de fusion est toujours attribué au dépassement d'un ou de plusieurs seuils définis par l'expert. Généralement un seuil est calculé à partir du nombre des observations ambiguës (ou mutuelles entre les classes concernées).

Dans le cas d'un suivi d'un roulement, la fusion de classe peut se produire durant la période dans laquelle le roulement est en bon état de fonctionnement.

La scission des classes

Le phénomène de scission se produit quand une classe se scinde en deux ou plusieurs petites classes au cours de l'incorporation des nouvelles observations. Ceci peut se produire lorsque les observations correspondant au mode de fonctionnement sont mélangées avec des observations parasites causées par des perturbations extérieures. La classe correspondante au mode de fonctionnement va évoluer et se séparer des observations qui représentent les perturbations. La méthode de classification dynamique doit être capable de détecter ce phénomène en utilisant un critère révélateur de la scission. La méthode doit également incorporer des règles à suivre dans le cas de scission pour mettre à jour les modèles des classes résultantes. La difficulté principale réside dans la définition d'un détecteur fiable qui va permettre à la méthode de déterminer si les données regroupées dans une classe sont vraiment similaires ou non.

3. PROCESSUS DE DÉVELOPPEMENT DES SYSTÈMES DE RECONNAISSANCE DE FORMES DYNAMIQUES POUR LE DIAGNOSTIC DES ROULEMENTS

Dans le cas d'un suivi d'un roulement, la scission des classes peut se produire lors du passage d'un état de dégradation vers un autre état plus critique.



FIGURE 3.7 — Scission de la classe C_1 en classes C_2 et C_3 après la classification des observations nouvellement arrivées

La suppression des classes obsolètes ou parasites

Durant l'apprentissage en ligne, il se peut que des connaissances apprises dans le passé ne soient plus valables à l'instant courant. Il se peut également que des connaissances ne soient pas, ou plus représentatives de l'état de fonctionnement de la machine à surveiller. Les méthodes de classification doivent donc se débarrasser des classes précédemment détectées et formées durant la classification et qui peuvent représenter des perturbations. Ces classes considérées comme des classes parasites vont uniquement alourdir le calcul. Il vaut mieux donc les supprimer.

La question à laquelle le classifieur dynamique doit répondre est qu'à partir de quel moment une classe devient non représentative (obsolète ou parasite)? La détection des classes parasites se fait généralement à l'aide d'un critère de cardinalité en adoptant par exemple un seuil minimal de points dans une classe pour qu'elle soit considérée comme représentative. Cependant, si on utilise la cardinalité toute seule on risque de supprimer toutes les classes nouvellement créées même celles représentatives du passage de la machine à un nouveau mode de fonctionnement. La solution est donc d'associer un autre critère à la cardinalité. Ce critère va permettre à la méthode de savoir si la classe suspectée est nouvelle ou parasite, et ceci en surveillant si cette classe continue à recevoir des nouvelles observations d'une façon intermittente, il s'agit bien d'une nouvelle classe sinon la classe est jugée parasite. L'obsolescence de connaissances (classes) est déterminée

en calculant le nombre d'observations classées dans d'autres classes depuis que la classe suspectée a reçu ses dernières observations. Il faut noter que la notion d'obsolescence reste une notion subjective qui s'appuie sur la vieillesse des informations dans les données. La décision de l'obsolescence d'une connaissance n'est possible qu'en ayant des connaissances supplémentaires sur l'application.

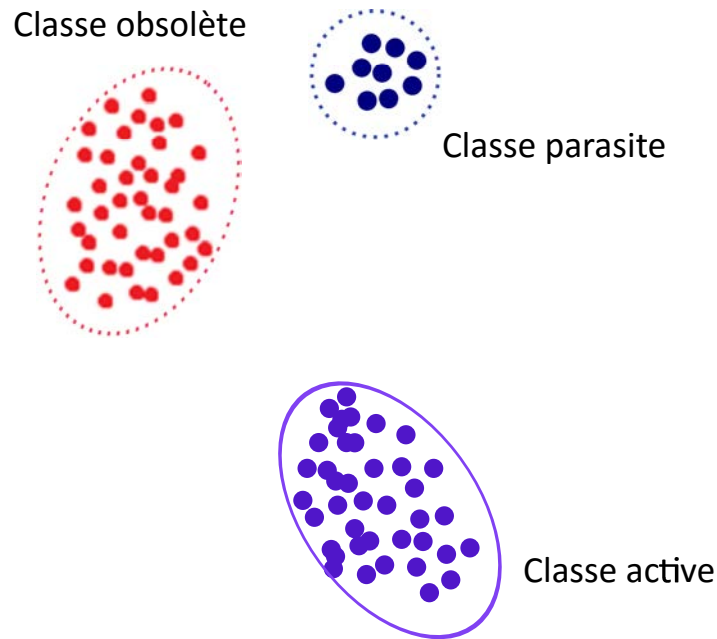


FIGURE 3.8 – La suppression des classes non représentatives

3.2.3 Conclusion sur la classification dynamique

Les méthodes de classification dynamique représentent une solution optimale pour gérer l'aspect évolutif des données. L'aspect dynamique de ces méthodes se révèle par leurs capacités à détecter les évolutions des classes (déplacement, rotation, grossissement, etc.), les changements de partition (création des nouvelles classes, suppression d'une classe, etc.), et d'adapter en permanence le classifieur pour qu'il classifie au mieux les nouvelles données arrivées tout en gardant un historique des changements que le système a subit.

3.3 Conclusion du chapitre

Afin de proposer une solution adaptée au suivi, il faut tout d'abord définir le contexte du problème étudié, ses contraintes et d'analyser les solutions possibles, ceci a été l'objectif

3. PROCESSUS DE DÉVELOPPEMENT DES SYSTÈMES DE RECONNAISSANCE DE FORMES DYNAMIQUES POUR LE DIAGNOSTIC DES ROULEMENTS

de la première partie de ce chapitre. Dans cette partie on a pu voir que pour répondre aux différentes contraintes, la méthode de RdF doit s'adapter à la nature évolutive et non stationnaire des données, elle doit être capable également de détecter les évolutions des données et/ou des classes, d'intégrer toutes les informations disponibles même incomplètes dans son processus de décision. Une des possibilités pour gérer une partie de ces contraintes est d'employer une méthode de classification dynamique.

Dans la deuxième partie de ce chapitre on a introduit la classification dynamiques avec quelques exemples et définitions. Nous avons également expliqué les différentes évolutions qu'on peut rencontrer et comment peut-on les gérer.

Chapitre 4

Contribution à la classification dynamique pour le suivi de roulements

Dans ce chapitre, nous présentons d'abord l'architecture de la méthode de reconnaissance de formes pour le suivi de roulement, ensuite les méthodes de classifications dynamiques développées durant cette thèse.

4.1 Présentation de l'architecture de la méthode de reconnaissance de formes conçue

La figure 4.1 représente l'organigramme de la méthode Rdf proposée. L'étape extraction dans cette méthode est divisée en deux parties, premièrement on extrait plusieurs indicateurs de défauts du nouveau signal enregistré. A partir de ces indicateurs, une matrice est construite. Pour réduire la dimension de cette matrice, améliorer la séparation des classes et enlever toutes redondances d'informations, on appliquera une méthode de réduction de dimension des caractéristiques (*KPCA*, *LPP*, *NPE* ou *SFS*). Ces nouvelles caractéristiques sont ensuite fournies à la méthode de classification dynamique comme des coordonnées de la(les) nouvelle(s) observation(s) à classer dans l'espace de représentation. La méthode de classification va alors affecter cette (ces) observation(s) à la classe (mode de fonctionnement) appropriée, mettre à jour ses règles de classification (en cas de besoin), et analyser l'évolution des classes après affectation de la nouvelle observation.

4. CONTRIBUTION À LA CLASSIFICATION DYNAMIQUE POUR LE SUIVI DE ROULEMENTS

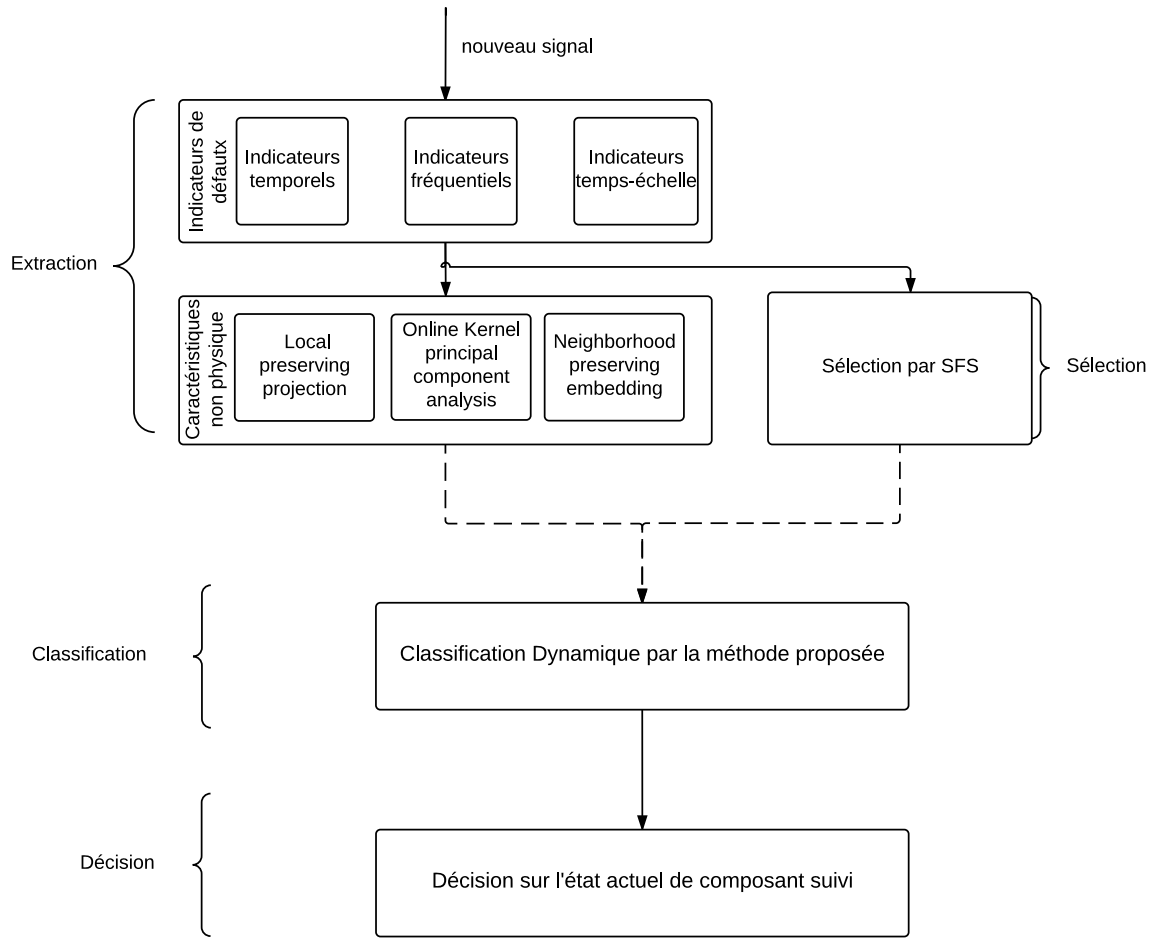


FIGURE 4.1 – L'organigramme général de la méthode RdF proposée

4.2 Extraction des indicateurs de défauts

Le choix des indicateurs de défauts retenus a été établi selon trois critères :

- La sensibilité au défaut : les indicateurs de défauts doivent aider à détecter la présence de défaut le plus précocement possible.
- L'insensibilité aux changements des conditions de fonctionnement : l'évolution observée en cas de défauts doit être différente de celle observée en cas de changement des conditions de fonctionnement (augmentation de charge ou augmentation de vitesse de rotation).
- La séparabilité entre mode sain et défectueux : certains indicateurs affichent les

4.3 L'extraction des caractéristiques non physiques par les méthodes RD

mêmes valeurs pour un composant sain que celle d'un composant en fin dégradation. On s'intéresse alors à des indicateurs qui peuvent nous aider à toujours séparer mode sain d'un mode défectueux même quand le niveau de dégradation de composant est élevé.

Il n'existe pas, dans la littérature, d'indicateurs qui respectent toutes ces contraintes quelle que soit la cinématique de la machine. Nous avons retenu les indicateurs qui pour nous sont qui respectent au mieux ces contraintes. La liste des indicateurs retenus est présenté dans le tableau 3.1.

Indicateurs temporels	RMS , $Peak$, $Kurtosis$, Fc , $Skewness$, FI et <i>écart type</i> .
Indicateurs fréquentiels	Spectre : $frms_b$, FC_b , $rmsf_b$, $xstdf_b$. Enveloppe : W_{RMS} , $PCWT$, $SPRI$ et $SPRO$

TABLE 4.1 – les indicateurs de défauts retenus

4.3 L'extraction des caractéristiques non physiques par les méthodes RD

L'extraction dans la reconnaissance de formes ne se résume pas seulement à l'extraction des caractéristiques qui portent des explications physiques mais on peut avoir recours à d'autres caractéristiques qui ne sont toujours pas directement interprétables mais qui peuvent apporter des meilleurs résultats en termes de séparation entre les différentes classes (Mu.01).

Plusieurs méthodes issues de traitement de données peuvent être employées pour **créer** des nouvelles caractéristiques résumant toute l'information véhiculée par les anciennes caractéristiques (indicateurs de défauts) tout en permettant une meilleure séparation des classes. Les plus connues sont les techniques de réduction de dimensions à citer : La décomposition en valeurs singulières SVD, l'analyse en composantes principales, l'analyse factorielle discriminative, l'analyse discriminante linéaire (linear discriminant analysis *LDA*), local preserving projection *LPP*...

Ces méthodes peuvent être trier par leur type :

- Apprentissage : Supervisé (*LDA*) non supervisée (*ACP*, *SVD*) ou configurable (*LPP*, *NPE*)

4. CONTRIBUTION À LA CLASSIFICATION DYNAMIQUE POUR LE SUIVI DE ROULEMENTS

- Linéarité : Linéaire (ACP, LPP) non Linéaire (KPCA)
- Projection : "Manifold" (LPP, NPE) ou non (ACP)

Nous avons utilisé trois méthodes de réduction de dimension de la matrice d'indicateurs de défauts : *KPCA*, *LPP* ou *NPE*. Dans cette section nous présentons la méthode LPP, NPE ainsi que la version en-ligne de la méthode *KPCA* que nous avons utilisée.

4.3.1 L'extraction par préservation de structure locale de données

Local preserving projection *LPP* est une méthode d'extraction des caractéristiques introduite par Niyogi en 2004 (Niy04) comme une méthode de réduction de dimensions linéaire alternative à l'*ACP* et qui permet de préserver la topologie locale des données même après projections dans le nouvel espace de dimension réduite.

La *LPP* est une méthode qui peut être supervisée si on a accès à des données étiquetées ou non supervisée dans le cas contraire. Elle peut être appliquée directement sur les données ou sur les données projetées par une fonction noyau. Son principe repose sur la construction d'un graphe incorporant toutes les informations de voisinage de l'ensemble de données. En utilisant la notion "Laplacian" du graphe, la matrice de transformation des points de données à un sous ensemble est calculée. Cette transformation linéaire préserve d'une manière optimale les informations du voisinage local au sens d'un critère choisi. La carte de la représentation générée par l'algorithme peut être considérée comme une approximation linéaire discrète de la carte continue initiale décrivant la topologie des données avant réduction (MP03).

L'algorithme de *LPP* est le suivant :

1. **Construction du graphe d'adjacence** : Supposons $\mathbf{G}$ un graphe de m nœuds. Une arête connecte les nœuds i et j si x_i et x_j sont "proches". Il existe deux façons pour construire ce graphe :
 - (a) ϵ -neighborhood ou ϵ -voisinage (le paramètre $\epsilon \in \mathbf{R}$). Les nœuds i et j sont connectés par une arête si $\|x_i - x_j\|^2 < \epsilon$ où la norme utilisée est la norme euclidienne en $\mathbf{R}^n$.
 - (b) K plus proches voisins (*KNN*) (le paramètre $k \in \mathbf{N}$). Les nœuds i et j sont connectés par une arête si i est parmi les k plus proches voisins de j ou si l'inverse est vrai.

4.3 L'extraction des caractéristiques non physiques par les méthodes RD

2. **Le choix des poids :** Il y a deux façons pour choisir les poids des arêtes. W est une matrice $m \times m$ symétrique creuse avec $W_{i,j}$ est le poid de l'arête connectant i et j ou 0 sinon.

(a) *Heat kernel* (le paramètre $t \in \mathbf{R}$). Si les nœuds i et j sont connectés :

$$W_{i,j} = e^{-\frac{\|x_i - x_j\|^2}{t}} \quad (4.1)$$

(b) *Simple-minded* (pas de paramètres). $W_{i,j} = 1$ Si les nœuds i et j sont connectés par une arête.

3. **Le calcul des projections :** On calcule les vecteurs propres et les valeurs propres pour le problème généralisé de vecteurs propres

$$XLX^T a = \lambda XD X^T a \quad (4.2)$$

Avec D est la matrice diagnoale avec $D_{ii} = \sum_j W_{ji}$. $L = D - W$ est la matrice laplacienne, et le i -ième colonne de X est x_i . Supposons que les colonnes des vecteurs $a_0, a_1, \dots, a_{l-1}$ sont les solutions de l'équation 4.2, ils sont ordonnés selon leur valeurs propres, $\lambda_0 < \dots < \lambda_{l-1}$:

$$x_i \rightarrow y_i = A^T x_i, A = (a_0, a_1, \dots, a_{l-1}) \quad (4.3)$$

Avec y_i est un vecteur de l dimensions et A est une matrice $n \times l$

Les projections de *LPP* sont obtenues en trouvant les approximations linéaires optimales pour les fonctions propres de l'équation de Laplace Beltrami opérées sur l'espace topologique des données.

Il y a d'autres techniques inspirées de *LPP*. On cite *NPE* (Neighborhood preserving embedding) qui suit la même démarche que *LPP*. La seule différence qu'on peut observer réside dans la construction de la matrice des poids où il faut résoudre le problème suivant :

$$\min \sum_i \|x_i\| \sum_j W_{i,j} x_j \quad (4.4)$$

Avec la contrainte $\sum_j W_{ij} = 1, j = 1, 2, \dots, m$.

L'utilisation de *LPP* et *NPE* est la même que dans l'*ACP*, il suffit de choisir le bon nombre d'axes ou la bonne dimension de représentation et de projeter les données dans

4. CONTRIBUTION À LA CLASSIFICATION DYNAMIQUE POUR LE SUIVI DE ROULEMENTS

l'espace de représentation. Le LPP a été utilisé dans le diagnostic par Yu *et al.* (Yu11) comme une méthode d'extraction de caractéristiques d'une matrice d'indicateurs afin de fiabiliser le diagnostic.

En effet, l'utilisation des versions en-ligne ("online") est plus adaptée au suivi car cette version permet d'exploiter les informations cachées dans l'historique des observations pour créer de nouvelles caractéristiques représentatives de l'état de la machine.

Bien qu'il existe des versions en-ligne de *LPP* (JCSQ11) cependant nous ne l'avons pas intégré dans la procédure suite à une manque de documentation.

4.3.2 Online Kernel Principal components analysis

La version standard de *KPCA* est largement utilisée dans le mode de traitement par lot, non pas dans le mode de traitement en ligne. La version en-ligne, dans laquelle on peut exploiter les résultats trouvés dans les itérations précédentes, afficher un taux d'erreur de reconstruction équivalent ou proche de celui calculé pour une application hors ligne tout en n'ayant accès qu'à vecteur de données récemment enregistré, est plus adaptée au suivi.

Dans la version standard de *PCA* ou de ses variantes (*KPCA* et autres), l'entrée du problème est un ensemble de vecteurs $X = \{x_1, \dots, x_n\}$ dans $\mathbb{R}^{d \times n}$, la dimension cible est $k < d$. La sortie est un ensemble de vecteurs $Y = \{y_1, \dots, y_n\}$ dans $\mathbb{R}^{k \times n}$ qui minimise l'erreur de reconstruction.

Dans la version en ligne, l'algorithme reçoit les vecteurs d'entrée x_t successivement, la sortie y_t doit toujours être prête avant de recevoir le prochain vecteurs d'entrée x_{t+1} .

La méthode *KPCA* proposée par (P.H07) permet de traiter les données en ligne. Cette méthode permet de s'affranchir des problèmes rencontrés dans les adaptations précédentes de *KPCA* pour l'apprentissage en ligne (la dépendance du nombre de modèle avec le nombre de données d'où la condition de savoir à l'avance la taille de données). La méthode proposée par Honeine (et décrit dans l'algorithme 4) permet non seulement de traiter les données en ligne mais aussi de choisir le noyau de projection adapté à chaque vecteur de données reçues.

4.3 L'extraction des caractéristiques non physiques par les méthodes RD

Algorithm 4 L'algorithme de *KPCA* version en-ligne

Initialisation

$m = 1, \omega_1 = 1, k_1 = k(x_1, x_1), \beta_1 = 1$

à chaque instant $t \geq 2$, lors de l'acquisition de x_t

1. Calculer $k(x_t)$

$k(x_t) = [k(x_{\omega_1}, x_t) \dots k(x_{\omega_m}, x_t)]^T$

2. Représentation de sous-espace $k(x_t, \cdot)$

$\beta_t = K_m^{-1} k(x_t)$

3. Calculer la distance (carré) au sous-espace

$\epsilon_t^2 = k(x_t, x_t) - k(x_t)^T \beta_t$

if Le critère sur la distance est satisfait $\epsilon_t^2 \geq 2$ **then**

Incrémenter l'ordre du modèle

$m = m + 1, \omega_m = t, \alpha_t = [\alpha_t^T 0]^T$

Mettre à jour l'inverse de matrice Gram des noyaux

$K_m^{-1} = \begin{bmatrix} K_{m-1}^{-1} & 0 \\ 0 & 0 \end{bmatrix} + \frac{1}{\epsilon_t^2} + \begin{bmatrix} -\beta_t \\ 0 \end{bmatrix} \begin{bmatrix} -\beta_t^T & 1 \end{bmatrix}$

et la carte de noyau empirique $k(x_t)$

$k(x_t) = [k(x_t)^T k(x_t, x_t)]^T$

Mettre à jour la représentation de sous-espace de $k(x_t, \cdot)$

$\beta_t = [0_{m-1}^T 1]^T$

5. La sortie $\psi_t(x_t)$

$y_t = \alpha_t^T k(x_t)$

6. Mettre à jour les coefficients

$\alpha_{t+1} = \alpha_t + \nu_t y_t (\beta_t - y_t \alpha_t)$

Les coordonnées principales de tout x

$\psi(x) = \alpha^T k(x)$

4.3.3 Sélection des caractéristiques les plus informatives

Une méthode de sélection des paramètres repose sur la détermination de plusieurs éléments :

- La qualité d'un paramètre : définir un paramètre ou une caractéristique pertinente.
- La technique de recherche utilisée pour retrouver les paramètres les plus pertinents.
- Le choix du nombre de paramètres ou caractéristiques à retenir.

4. CONTRIBUTION À LA CLASSIFICATION DYNAMIQUE POUR LE SUIVI DE ROULEMENTS

Pour la sélection des caractéristiques les plus informatives, nous avons utilisé une méthode assez connue, la *SBS* (sequential backward selection). Le critère de qualité généralement employé avec *SBS* vise à maximiser la dispersion intra-classe et minimiser la dispersion inter-classe en se servant des données fournies dans la base d'apprentissage. Dans le cadre du suivi, on ne peut pas employer des méthodes supervisées, l'objectif de sélection sera alors de trouver le plus petit sous-ensemble des caractéristiques qui révèle mieux les groupements intéressants et naturels (clusters) à partir de données selon le critère choisi. Contrairement à l'apprentissage supervisé, dans lequel on dispose des étiquettes de classe pour guider la recherche de caractéristiques les plus informatives, dans l'apprentissage non supervisé, nous devons définir ce que "intéressant" et "naturel" signifient. Ceux-ci sont généralement représentés sous la forme d'un critère choix.

Dans la sélection des paramètres pour un apprentissage non supervisé, le critère de sélection repose généralement soit sur la similarité ou la ressemblance de données(MCP02), soit sur la qualité de séparation des classes formées à partir d'un sous ensemble des caractéristiques choisies.

Le schéma décrit dans la figure 4.2 résume le processus de sélection pour une méthode de classification non supervisée. On commence par fournir à la méthode de sélection l'ensemble de caractéristiques dont on dispose, la méthode commence par effectuer une recherche d'un sous ensemble candidat de caractéristiques, ce sous ensemble est introduit à la méthode de classification non supervisée. Ensuite la méthode de classification classe ces caractéristiques en clusters ou classes. La qualité de la méthode de séparation de ces classes est évaluée par un critère d'évaluation de classification. Puis on crée un autre sous-ensemble de caractéristique et on répète les étapes de classification et évaluation jusqu'à la satisfaction d'un critère d'arrêt qui peut être la fin de recherche.

Dans cette thèse, nous avons adopté le même processus de sélection décrit dans 4.2, en choisissant la méthode *SFS* comme méthode de recherche, *K – means* comme méthode de classification non supervisée et le critère *Silhouette* comme critère d'évaluation. le critère d'arrêt choisi est la fin de la recherche.

4.3 L'extraction des caractéristiques non physiques par les méthodes RD

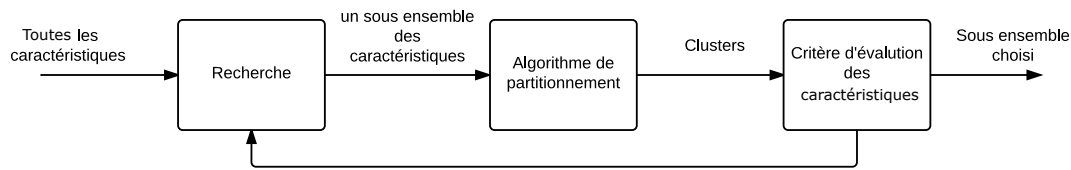


FIGURE 4.2 – La sélection dans un environnement non supervisé (*Wrapper*)

4.3.4 Bilan comparatif sur les méthodes d'extraction des caractéristiques non physiques

L'extraction des caractéristiques non physiques avait pour objectif de retrouver/ créer des nouvelles caractéristiques qui peuvent améliorer la séparation entre état sain et état défectueux. Quatre méthodes d'extraction ont été retenues *SFS*, *LPP*, *NPE* et *Online – KPCA*. Le tableau 4.2 représente un bilan récapitulatif des méthodes de réductions retenues.

TABLE 4.2 – Tableau Comparatif des caractéristiques trouvées par les différentes méthodes d'extraction

Méthode employé	En ligne	Manifold Preserving	Complexité	Signification physique
SFS	⊖	⊖	⊕	⊕
LPP	⊖	⊕	○	○
NPE	⊖	⊕⊕	○	○
Online-KPCA	⊕	⊖	⊕	⊖

Comme on peut voir sur le tableau 4.2, chaque méthode a ses avantages et ses inconvénients : Avec *SFS* on peut conserver le sens physique des caractéristiques sélectionnées ce qui peut être très utile pour interpréter les résultats obtenus et/ou les phénomènes observées après la classification. Cependant, le processus de sélection par *SFS* est gourmand en consommation de mémoire et temps du calcul. Contrairement à la réduction par 'Online- KPCA' qui permet une réduction optimale correspondante à la nature en ligne de données. Néanmoins, l'interprétation physique des résultats obtenus est très difficile. Les méthodes de réduction par conservation de la structure locale de données (*LPP* et *NPE*) permettent une meilleure réduction tout en gardant la structure d'origine des données ce qui est très utiles pour observer les changements temporels de données et qui est très apprécié par les méthodes de classification automatique. Cependant, l'interprétation physique par ces méthodes est difficilement accessible.

4.4 Classification

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 principes, la démarche ainsi que les motivations du développement de chaque méthode développée.

4.4.1 Architecture générale des méthodes de classification proposées

L'architecture de la méthode de classification développée pour le suivi est présentée dans la figure 4.3, elle suppose que les données ont été déjà traités, extraites et sélectionnées et qu'elles sont prêtes à être classifiées.

Après l'arrivée de chaque nouveau signal et après l'extraction des caractéristiques et la sélection de celles qui sont les plus informatives, 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 la partition est mise à jour. 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.

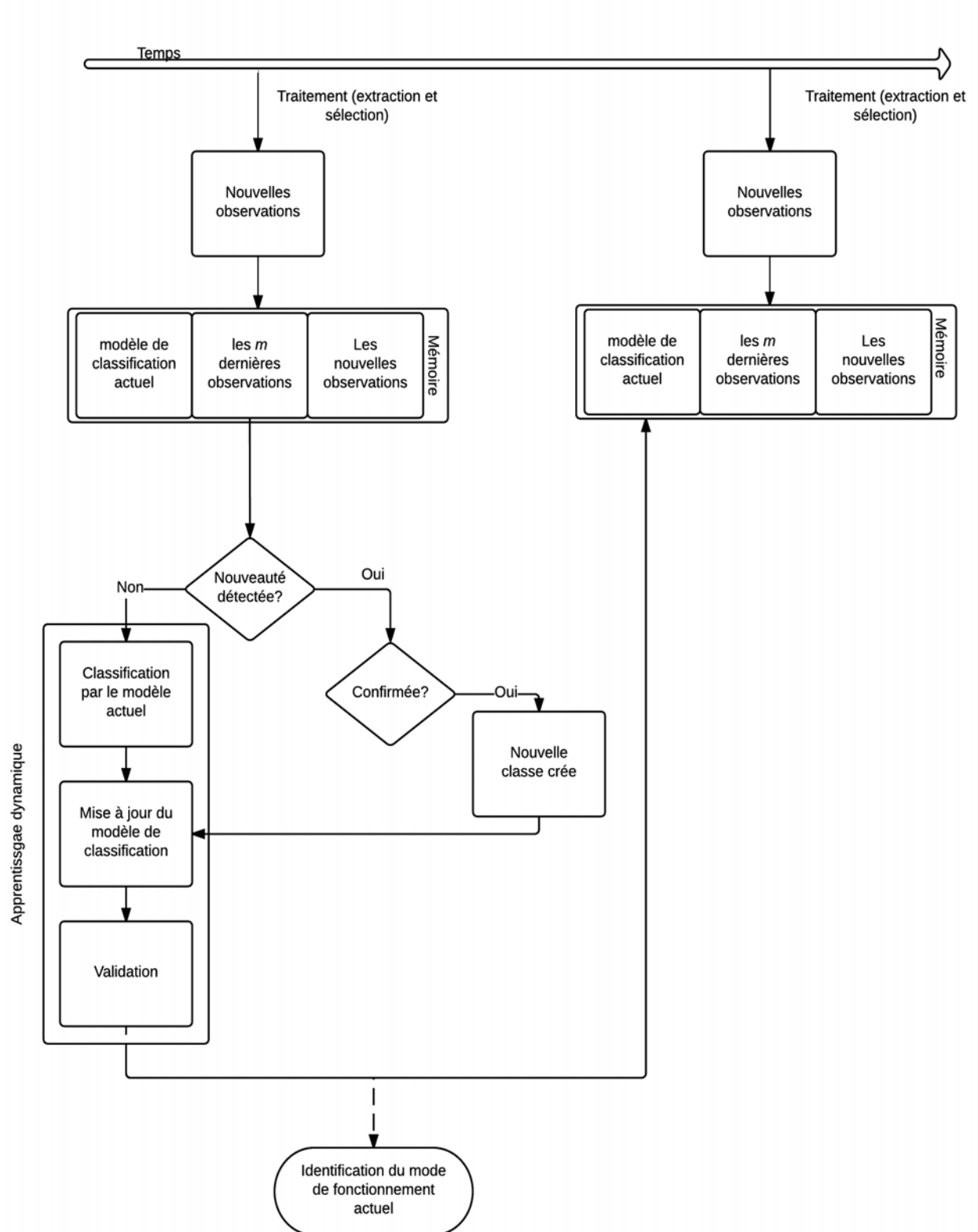


FIGURE 4.3 – l’architecture générale de la méthode de classification dynamique pour le suivi

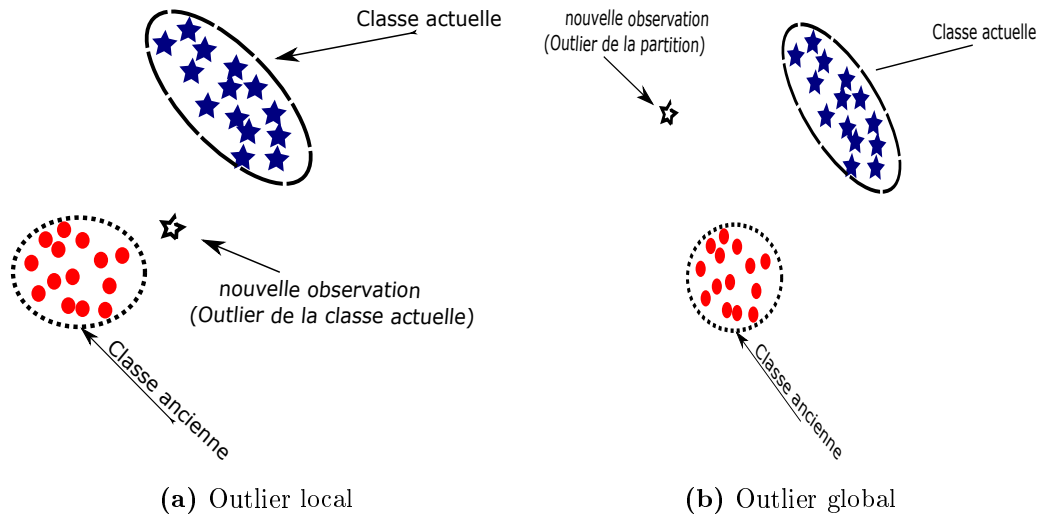


FIGURE 4.4 – Deux types d’outliers rencontrés dans le suivi

4.4.2 La détection des nouveautés

Une observation est considérée comme nouveauté ou *outlier* si elle est différente de toutes les observations retenues (par exemple s’elle est située dans une zone loin des classes connues) ou si elle est différente des observations récentes (loin de la classe actuelle). Une nouveauté détectée peut être un point aberrant résultant du bruit, ou le résultat de passage de système suivi d’un mode à un autre. Une vérification est donc nécessaire pour s’assurer que la nouveauté détectée est porteuse d’information ou un bruit.

Il n’existe pas de méthode standard de détection des nouveautés qui peut être appliquée dans tous les cas et pour tout type d’application. Il y a différents outils possibles pour la détection et chaque technique a ses avantages et ses inconvénients.

Dans le cas du suivi, on peut rencontrer deux types de outliers : locaux et globaux. On définit un outlier global par toute observation qui se situe loin de toute zone connue dans l’espace de représentation, et local par toute observation différente des observations les plus récentes et donc de la classe actuelle. Un outlier global est donc un outlier local (puisque’il est loin de la classe actuelle) qui est aussi loin de toutes les autres classes définies par le système.

Les outliers locaux et globaux peuvent être utilisés dans le cadre du diagnostic et du suivi. Nous nous intéressons aux outliers locaux pour deux raisons : (i) pour alléger la complexité du processus, (ii) la détection des outliers globaux peut être faite à partir des

outliers locaux.

4.4.3 La méthode utilisée pour détecter les outliers en ligne

Pour détecter les *outliers* locaux, on applique la méthode de détection de nouveautés "The Moving Window Filtering Algorithm *MFWA*" légèrement modifiée pour intégrer la notion de traitement en ligne, le principe de *MFWA* est d'évaluer la validité d'une nouvelle observation sur la base de sa distance des voisinage des plus proches observations déjà validées. Pour alléger cette méthode et l'adapter au traitement en ligne nous avons appliqué l'algorithme MFWA sur les observations récemment validées. Cette méthode va marquer toute observation qui est trop loin de observations, récemment identifiées comme normales, par "outlier". Ces outliers détectés seront des outliers locaux à la classe la plus récente.

Toute observation marquée comme *outlier* sera ensuite analysée puis traitée par la méthode de classification.

4.4.4 La création d'une nouvelle classe

Si les observations marquées comme *outliers* et stockées dans le buffer sont *(i)* bien séparées de la classe la plus récente, *(ii)* assez proches l'une de l'autre pour former un cluster ou une classe *(iii)* assez nombreuses pour différencier entre points aberrants et un véritable changement du système (si le nombre d'observations qui respectent les deux premières règles est supérieur à un minimum définie), alors une nouvelle classe est créée et tous les outliers qui respectent les règles seront attribués à cette nouvelle classe et supprimés du buffer. Cette classe sera considérée comme la classe actuelle (nouvelle classe). Si les trois règles ne sont pas respectées, les outliers resteront dans le buffer.

Dans le cas où les nouvelles observations ne sont pas considérées comme des nouveautés, on entame le processus d'apprentissage dynamique.

4.5 Méthodes de classification dynamiques proposées

Comme il a été expliqué auparavant, l'apprentissage dynamique est un processus à plusieurs étapes (apprentissage (si besoin), classification, adaptation et validation). Nous

4. CONTRIBUTION À LA CLASSIFICATION DYNAMIQUE POUR LE SUIVI DE ROULEMENTS

allons détailler chaque étape de ce processus pour trois méthodes de classification dynamiques différentes développées durant cette thèse (*Dynamic DBSCAN*, *Evolving scalable DBSCAN* et *DFSDBSCAN Dynamic fuzzy scalable DBSCAN*).

4.6 Dynamic DBSCAN *DDBSCAN*

4.6.1 Présentation générale

Dynamic DBSCAN est une méthode non supervisée développée pour suivre l'évolutions de l'état de roulement.

Le choix de *DBSCAN* comme une base de développement de la méthode de classification dynamique pour le suivi a été établi après une étude préliminaire et comparative entre les performances de plusieurs méthodes de classification non supervisées statiques appliquées sur des signaux vibratoires. Cette étude nous a aidé à conclure que la meilleure méthode de séparation dans le cas des signaux vibratoires est celle basée sur la densité (KXL13). La méthode de classification *DBSCAN* 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 à détecter les classes même ceux d'une forme irrégulière, à détecter les points aberrants, à distinguer le bruit, à utiliser des méthodes d'accès spatiales même pour les grandes bases de données spatiales contrairement aux méthodes de partitionnement.

La méthode de classification *DBSCAN* a été mis à niveau pour s'adapter à la nature dynamique de données.

4.6.2 *DDBSCAN*

La méthode *DDBSCAN* créée dans cette thèse, comme chaque méthode de classification dynamique ou évolutive, suit une démarche à plusieurs étapes (1/classification ,2/adaptation et 3/validation).

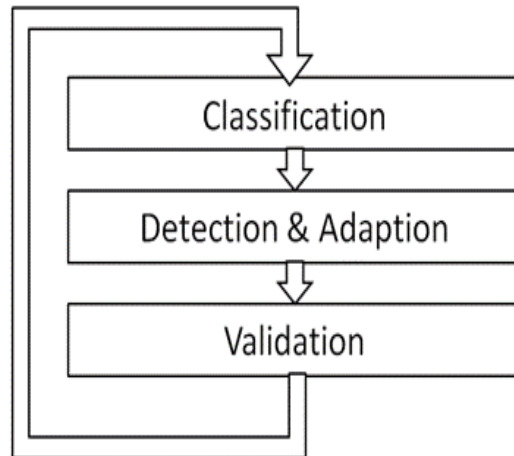


FIGURE 4.5 – L’organigramme général de la classification dynamique

A. Phase 1 : Classification

Il n’y a pas de phase d’apprentissage car ce classifieur est de type non supervisé. la méthode commence directement par la classification *DBSCAN* classique. Nous avons ajouté deux fonctionnalités :

- Des données étiquetées : Le *DBSCAN* classique ne traite que les données non étiquetées, mais dans cette version nous avons ajouté la possibilité de recevoir des données étiquetées dans le cas où l’utilisateur veut ajouter une information de contrôle comme par exemple présence de défaut ou variation de charge. Si l’algorithme ne reçoit pas des données étiquetées, il classera les données d’une façon classique.
- Le traitement des outliers : Le *DBSCAN* classique étiquette les points par trois étiquettes (core point, border point et outlier). Un point est étiqueté "core" si on peut trouver un nombre de points dans son voisinage *Eps* supérieur à *Minpts*. Un point est étiqueté "border" s’il est dans le voisinage d’un point core et qu’il n’y a pas assez de points dans son voisinage. Quand on ne trouve pas assez de points au voisinage d’un certain point et qu’il n’est pas au voisinage d’un point core, ce point est considéré comme un outlier. Cependant dans la classification en ligne, les statuts de tous les points peuvent changer dans le temps ; un "outlier" peut changer pour

4. CONTRIBUTION À LA CLASSIFICATION DYNAMIQUE POUR LE SUIVI DE ROULEMENTS

devenir un "border" ou un "core" d'une classe existante ou ce point peut appartenir à une classe nouvellement créée. Le changement de statut d'un point accompagné d'autres phénomènes peut indiquer une dérive de fonctionnement du système à surveiller. Bien que tout changement de statut des points peut être intéressant, le changement de statut d'un point outlier est plus informative pour le suivi. Pour ces raisons on s'est focalisé sur les outliers que l'algorithme originale de DBSCAN néglige pour qu'ils ne soient pas être mises à coté mais analysés pour pouvoir en tirer des informations utiles (SXL14). Dans la version modifiée de *DBSCAN* avant de marquer un point outlier nous l'avons mis dans un buffer. Si le point continue à être marqué outlier après un certain temps et qu'il ne définit pas une dérive, le point sera finalement marqué outlier et ignoré dans le traitement.

- Changement de métrique : Le *DBSCAN* classique utilise la distance euclidienne. Dans cette version nous avons remplacé la distance euclidienne par la distance Mahalanobis.

Nous avons également collecté certaines informations sur les classes pour chaque ensemble de données arrivées à savoir :

- La distance Hausdorff : on calcule la distance entre la classe actuelle et les autres classes :

$$h_{ij} = h(C_i, C_j) \quad (4.5)$$

- La densité de la classe : on calcule la compacité de la classe

$$S_C = \prod_{k=1}^{N_p} \max(x_{Ck}) - \min(x_{Ck})$$

Avec N_p est le nombre d'objets de la classe C

$$Dc = S_C / N_p \quad (4.6)$$

- Le rayon de la classe : on calcule la distance entre les deux points les plus éloignés de la classe selon chaque caractéristique

$$rad(C, a) = \max(dist(x_a, y_a)) \quad (4.7)$$

Avec x et y sont deux points de la classe C et a la caractéristique ou l'attribut.

B. Phase 2 : Détection et Adaptation

La classe qui reçoit un nouveau point x est celui qui peut évoluer. Mais avant de classer le point x dans une classe C , deux conditions doivent être vérifiées ; d'abord dans un rayon de voisinage du point x on trouve un *Minpts* points, le point x est temporairement affecté à la classe la plus proche. Si le point x a été marqué comme temporairement "outlier" et que on a pu trouver un *Minpts* de nouveaux points dans son entourage, une nouvelle classe est créée avec tous les points dans le voisinage de x , et tous les "paramètres" de classe (distance Hausdorff, la densité, et trajectoire) sont mis à jour pour l'ancienne classe à laquelle les points au voisinage de x ont été temporairement affectés. Ces paramètres sont initialisés pour la nouvelle classe créée. Si le "outlier" se trouve au voisinage d'un point "Core" le point sera alors assigné à la classe de point core.

C. Phase 3 : Validation

Dans cette phase, deux types d'évolution des classes sont traités ; la fusion et la suppression. Deux classes doivent fusionner si la surface de recouvrement entre les deux classes est supérieure à un seuil (THF) défini par l'utilisateur. La valeur de THF est exprimée en pourcentage (THF = 10 signifie que la surface de recouvrement est égal à 10% de la surface des deux classes). Dans le cas de la suppression de classe, un autre seuil THD est spécifié. Ce seuil représente le nombre de points depuis que la classe C a reçu son dernier point (si THD = 30 signifie que 30 points ont été reçus depuis qu'il a reçu son dernier point).

4.6.3 Bilan et limitations de la classification par Dynamic DBSCAN *DDBCAN*

La méthode de classification *DynamicDBSCAN* est une méthode de classification par densité. Cette méthode est du type non supervisé, elle ne nécessite pas donc une base d'apprentissage. *DynamicDBSCAN* a été conçue pour pouvoir classer les observations reçues à fur et à mesure. Ainsi, des modifications majeures ont été effectuées sur la méthode de la base *DBSCAN*. La première modification est la conservation des outliers dans une mémoire tampon, ensuite le changement de la métrique pour passer de la distance euclidienne à la distance mahalanobis et finalement l'intégration des outils

4. CONTRIBUTION À LA CLASSIFICATION DYNAMIQUE POUR LE SUIVI DE ROULEMENTS

de détection et le suivi des évolutions des classes et de la partition obtenue dans le temps.

Le *DDBCAN* permet de classer 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 classification. Cependant la taille de mémoire nécessaire pour enregistrer les données, l'accès multiple aux anciennes données et l'absence d'une règle floue pour améliorer la classification sont des challenges que la méthode *DDBCAN* a échoué à relever et que nous allons essayé de corriger dans les deux méthodes proposées *Evolving Scalable DBSCAN* et *Dynamic fuzzy scalable DBSCAN*.

4.7 Evolving Scalable DBSCAN

4.7.1 Présentation générale

Le premier challenge qu'on a essayé de relever avec cette nouvelle méthode est de conserver la nature de la classification par densité non supervisée tout en minimisant l'accès aux observations anciennes. Le but a été donc de développer une méthode premièrement adapté au traitement en ligne et auquel on peut intégrer les mécanismes de l'apprentissage dynamique. Seules quelques méthodes de classification non supervisées adaptées au traitement en ligne ont été trouvées dont *Clustream* (AJJP03), *BIRCH* (VR02) sont les plus connues. Ces deux méthodes qui partagent le principe de la classification sur deux étapes ont permis de rendre la classification non supervisée capable de traiter les données en ligne. Une méthode de classification par densité dont le principe se rapproche de ces deux méthodes est "Scalable Density Based Distributed Clustering *SDBDC*" proposée par Januzaj *et al* dans l'article (JKP04). Nous avons donc décidé d'adapter *SDBDC* au traitement en ligne et dynamique.

4.7.2 Scalable DBSCAN

La méthode *Density Based Distributed Clustering (DBDC)* (JKP04) décrit une stratégie visant à rendre *DBSCAN* distribuée. Les auteurs cherchent à regrouper la totalité de l'ensembles de données distribuées en minimisant les coûts de transition et sans nécessiter un accès complet aux données qui vont être analysées. Cet algorithme peut également être utilisé dans le cas du traitement d'un large ensemble de données divisé en petits

sous-ensembles ou dans le cas d'un traitement en ligne incrémentale.

Le principe de *DBDC* consiste à calculer la densité autour de chaque objet situé localement reflétant son aptitude à servir en tant que représentant du site local (ou d'un sous ensemble). Après avoir trier les objets en fonction de leur densité, on garde les représentants locaux pour représenter le site local. Ces objets représentant seront ensuite associés à leurs classes respectives avec une version améliorée de l'algorithme *DBSCAN*. En d'autres termes, le *DBDC* cherche (i) à déterminer un modèle local représentant de chaque sous ensemble de données, (ii) à déterminer un model global de l'ensemble des données et (iii) à mettre à jour les modèles locaux.

Afin de déterminer les représentants locaux adéquats de chaque cluster, il faut tout d'abord explorer l'ensemble de voisinage de tous les points locaux et chercher à calculer la densité de chaque points en respectant les deux conditions suivantes : Pour chaque point x de l'ensemble X à classifier on cherche l'ensemble des points à une distance inférieure à Eps

$$N(x; Eps) = \{y \in X | d(x, y) \leq Eps\} \quad (4.8)$$

Un niveau de représentativité est ensuite, affecté à chaque point. Le niveau de représentativité est calculé comme suit :

$$StatRepQ(x; Eps) = \sum_{y \in N(x; Eps)} (Eps - d(y, x)) \quad (4.9)$$

plus il y a d'objets autour de x dans un rayon de Eps , plus ces objets sont proches de x plus $StatRepQ(x; Eps)$ est élevée. Les points intéressants en vue de la classification sont ceux qui ont une grande valeur de $StatRepQ$. Il existe une autre mesure de représentativité par rapport aux autres points déjà choisis comme représentants.

$$DynRepQ(x, Eps, Rep_i) = \sum_{\substack{y \in N(x; Eps) \\ \forall r \in Rep_i; y \notin N(r; Eps)}} (Eps - d(y, x)) \quad (4.10)$$

La valeur de $DynRepQ(x, Eps, Rep_i)$ dépend du nombre et des distances des éléments trouvés à un rayon de Eps de x et qui ne sont pas encore inclus dans l'ensemble de voisinage d'un ancien point représentant local.

4. CONTRIBUTION À LA CLASSIFICATION DYNAMIQUE POUR LE SUIVI DE ROULEMENTS

Pour chaque élément représentant local, l'algorithme calcule aussi leur rayon de couverture $CovRad$ afin d'indiquer les éléments ayant une grande distance par rapport aux autres et aussi le nombre $CovCnt$ des objets couverts par le représentant local.

$$CovRad(r_{i_n}, Eps, Rep_{i_n}) = \max\{Eps - d(x, r_{i_n}) | \forall x \in X_i, \forall r \in Rep_{i_n} : x \in N(r_{i_n}; Eps) \wedge x \notin N(r; Eps)\} \quad (4.11)$$

$$CovCnt(r_{i_n}, Eps, Rep_{i_n}) = |\{x | \forall x \in X_i \forall x \in Rep_{i_n} : x \in N(r_{i_n}; Eps) \wedge x \notin N(r; Eps)\}| \quad (4.12)$$

Avec $X_i \subseteq X$ est l'ensemble d'objets situés sur le site i et $Rep_i = r_{i_1}, \dots, r_{i_n}$ est la séquence des n représentants locaux où $r_{i_1}, \dots, r_{i_n} \subseteq X_i$.

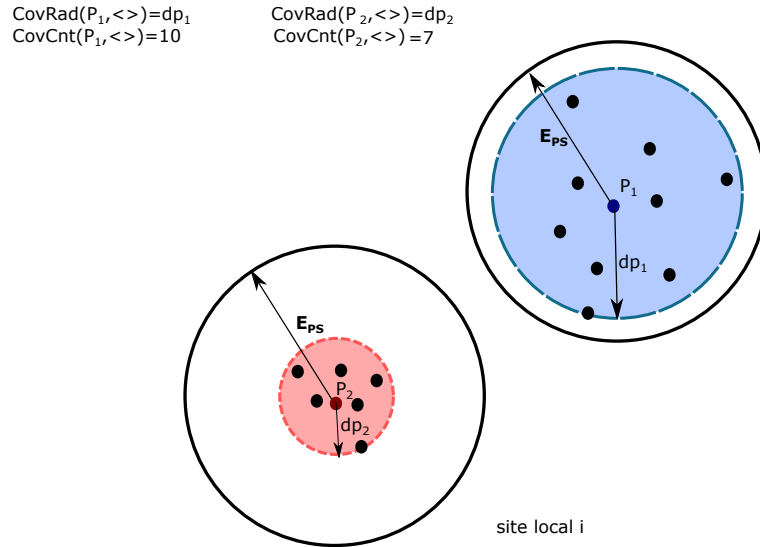


FIGURE 4.6 – Exemples de $CovRad$ et $CovCnt$ pour deux points p_1 et p_2

On peut voir des exemples des valeurs attribuées à $CovCnt$ et $CovRad$ pour deux exemples (voir figure 5.18a). Pour déterminer le modèle global de l'ensemble de données, une version améliorée de *DBSCAN* et adaptée à la classification des points locaux représentants est appliquée. Pour chaque point représentant une recherche de l'ensemble

des points situés autour de ce point à un rayon Eps est effectuée. Ainsi, une valeur $\varepsilon(r_i)$ spécifique à chaque représentant r_i est utilisée pour définir son voisinage.

$$\varepsilon(r_i) = Eps + CovRad(r_i) \quad (4.13)$$

L'algorithme *DBSCAN* original effectue une recherche pour chaque point de l'ensemble de données afin de retrouver son ensemble de voisinage sur un rayon Eps . Cette recherche ne pourra plus être effectuée dans le cas de classification distribuée pour laquelle seule une fraction de données est accessible. Pour que *DBSCAN* puisse classer les représentants locaux des différents sites sans nécessiter l'accès aux points non représentants, il faut qu'il augmente le rayon de recherche Eps par le rayon de couverture $CovRad(r_i)$ de représentant actuel r_i . Cette approche garantit que la recherche des objets situés autour d'un représentant dans un rayon élargi aboutira aux mêmes résultats qui auraient été trouvés par une recherche classique sur l'ensemble des données. Il faut noter aussi que le *Minpts* appliqué pour retrouver les points core est réduit à la valeur '2'.

4.7.3 Evolving Scalable DBSCAN *ESDBSCAN*

L'organigramme 4.7 trace les grandes lignes de la méthode de classification proposée dans cette section. La classification commence par effectuer une analyse des nouvelles observations reçues afin de déterminer si ces nouvelles observations relèvent de la nouveauté pour le système ou non. Si ces observations sont différentes des celles qui les précèdent on crée des nouveaux points représentatifs. Ces nouveaux points représentatifs seront ensuite soit affectée à des classes déjà connues au système soit la classification sera relancée de nouveau.

A. Classification

La classification par *ESDBSCAN* se fait sur deux étapes. La première étape consiste à affecter les nouvelles observations aux points représentatifs déjà trouvés ou à de nouveaux points représentatifs. La deuxième étape consiste à affecter les points représentatifs à leurs classes.

4. CONTRIBUTION À LA CLASSIFICATION DYNAMIQUE POUR LE SUIVI DE ROUEMENTS

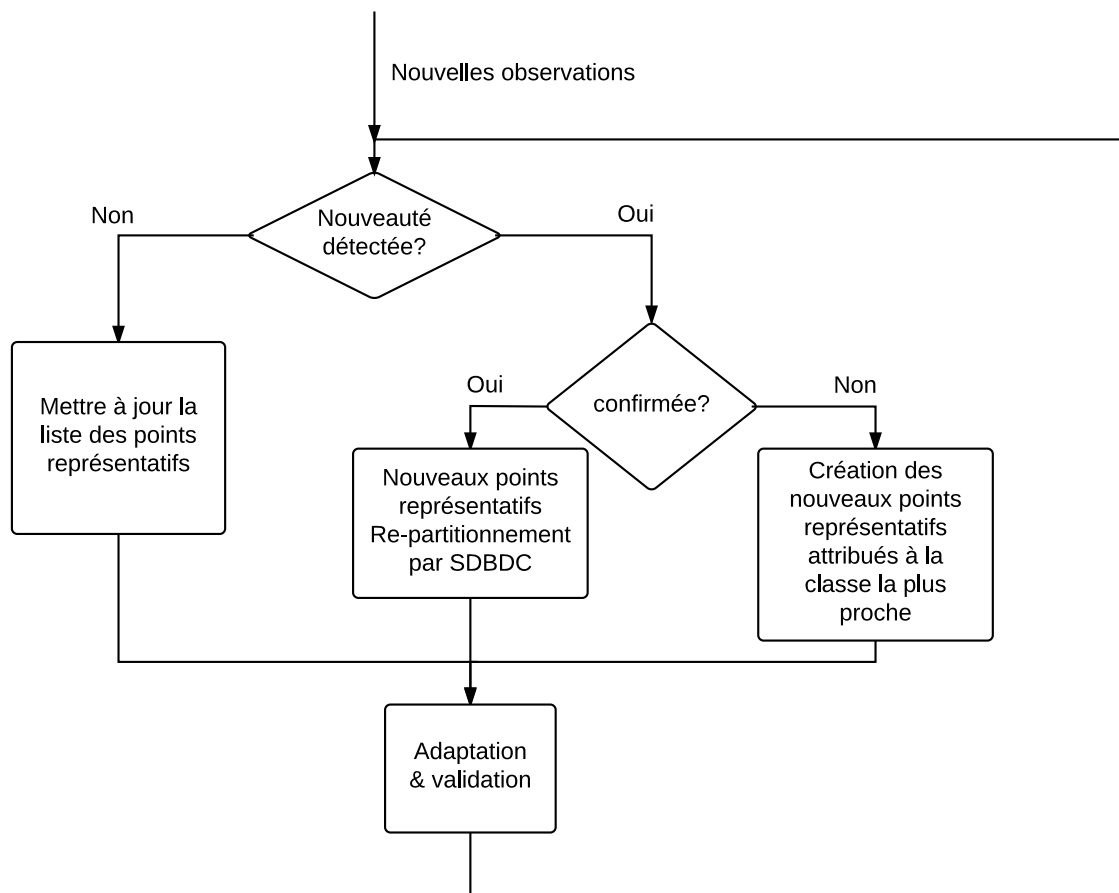


FIGURE 4.7 – L’organigramme de la classification par *ESDBSCAN*

A.1. Classification locale : recherche et mise à jour des points représentants

Dans cette étape on s’intéresse à l’identification des nouveaux points représentatifs et à la mise à jour des points représentants déjà identifiés auparavant.

Un point représentatif représente un objet pour la méthode de classification.

Algorithm 5 Un point représentatif

Propriétés

Point ;

ClusterId ;

CovCnt ;

CovRad ;

Dans l'identification des points représentants, on priorise d'abord la mise à jour des points représentatifs existants en ajoutant les nouveaux points de données qui se trouve dans un voisinage de Eps si ce n'est pas possible on crée un nouveau point représentant en calculant à chaque fois la qualité de représentation dynamique $DynRepQ(x, Eps, Rep_i)$.

L'algorithme 6 résume les différentes étapes à suivre pour retrouver des nouveaux points représentatifs.

Algorithm 6 CreateRepList() : Création des nouveaux points représentatifs

```

 $Rep_i = \{\}$ 
for  $x \in X_i$  do
    Calculer  $StatRepQ(x, Eps)$ 
Trier la liste des points représentatifs candidats SortRepList :=  $\langle (x_1, StatRepQ(x_1, Eps)), \dots, (x_{|X_i|}, StatRepQ(x_{|X_i|}, Eps)) \rangle | i \leq j \Rightarrow StatRepQ(x_i, Eps) \geq StatRepQ(x_j, Eps) \rangle$ 
while Critère d'arrêt est non vérifié do
     $Rep_i := Rep_i + SortRepList(1)$ 
    for chaque  $x \in X_i - Rep_i$  do
        Calculer  $DynRepQ(x, Eps, Rep_i)$ 
        Trier la liste des points représentatifs candidats SortRepList :=  $\langle (x_1, DynRepQ(x_1, Eps, Rep_i)), \dots, (x_{|X_i - Rep_i|}, DynRepQ(x_{|X_i - Rep_i|}, Eps, Rep_i)) \rangle | i \leq j \Rightarrow DynRepQ(x_i, Eps, Rep_i) \geq DynRepQ(x_j, Eps, Rep_i) \rangle$ 
     $Rep_i := Rep_i + SortRepList(1)$ 

```

La création de la liste commence par (1) effectuer une recherche des points candidats, ensuite (2) on ordonne les points représentatifs candidats dans un ordre descendant en se basant sur la qualité de représentativité statique $StatRepQ$ de chaque point. (3) On retire le point qui a la plus grande valeur $StatRepQ$ de la liste des candidats et on l'ajoute dans la liste des points représentatifs locaux retenus. (4) On crée une autre liste des points non retenus, en se basant cette fois sur leurs qualités dynamique $DynRepQ$ calculé pour tout objet local qui n'a pas été utilisé comme un point représentatif. (5) Cette liste est à son tour ordonnée dans l'ordre descendant en utilisant cette fois les valeurs de $DynRepQ$. (7) Si la condition d'arrêt n'est pas satisfaite, on revient à l'étape (3) et on suit les étapes suivantes jusqu'à que la condition d'arrêt est vérifiée. Cette condition peut être

4. CONTRIBUTION À LA CLASSIFICATION DYNAMIQUE POUR LE SUIVI DE ROULEMENTS

choisie en fonction d'un nombre de points représentatifs fixé ou une mesure de qualité de classification. Une fois la liste de points représentatifs est finalisée, on procède au calcul de poids et de cardinalité de chaque élément de la liste. Cet algorithme est utilisé pour créer la liste des points représentatifs dans le cas où les nouvelles données ne peuvent pas être attribuées à des anciens points représentatifs. Sinon on procède à une mise à jour des points représentatifs déjà désignés auparavant.

La mise à jour des représentatifs déjà connus après la réception d'un nouveau bloc de points est expliqué dans l'algorithme 7.

Algorithm 7 UpdateExRepList() : Mise à jour des points représentatifs existants

UpdateExRepList() :

for Chaque nouveau point P_i de P **do**

Trouver les points dans son entourage Eps $S = RangeQuery(P_i, Eps)$;

if Tester si dans l'entourage de P_i il existe un point représentatif de $RepList$ **then**

Assigner le point P_i **au point représentatif le plus proche** $Rep_j \in S$

Mettre à jour les caractéristiques du point représentatif le plus proche

$Rep_j.CovCnt = Rep_j.CovCnt + 1$,

$Rep_j.CovRad = \max(Rep_j.CovRad, dist(Rep_j.rep, P_i))$;

MarkedAsAssigned(P_i) ;

La première étape de la mise à jour des points représentatifs existants consiste à mettre à jour leurs caractéristiques $CovRad$ et leurs cardinalités $CovCnt$ si un ou plusieurs nouveaux points sont dans un rayon Eps de ces points retenus (voir algorithme 7). Sinon on crée des nouveaux points représentatifs à ajouter à la liste actuelle en suivant l'algorithme 8.

Algorithm 8 Mise à jour de la liste *RepList* existante par des nouveaux points représentatifs

```

for Chaque  $P_i$  non assigné à aucun point représentatif de RepList do
    Calculer  $DynRepQ(P_i, Eps, RepList)$ 
    Trier Liste des points représentatifs candidats  $SortRepList =$ 
 $Sort(DynRepQ, 'descend')$   $Rep_i = SortRepList(1)$ 
    Calculer la cardinalité  $CovCnt$  de  $Rep_i$   $Rep_i.CovCnt =$ 
 $CntObjets(RangeQuery(Rep_i, Eps))$ 
    Calculer la couverture  $CovRad$  de  $Rep_i$   $Rep_i.CovRad =$ 
 $max(distance(RangeQuery(Rep_i, Eps)))$ 
    Ajouter  $Rep_i$  à RepList

```

Il est à noter que dans certains cas, le nombre de points représentatifs retenus d'une classe dépasse un seuil défini. Dans ce cas on peut créer des points *SuperReprésentatif*, c'est à dire des points qui représentent des points représentatifs et qui peuvent résumer toute l'information continue dans la classe. Dans ce cas on établit la création d'une nouvelle liste de points représentatifs à partir d'une liste constituée des points représentatifs.

Dans la liste des points représentatifs trouvés, on peut trouver des points représentatifs très denses et d'autres moins denses voir même vides (c'est à dire qu'il ne représente qu'eux même ou un nombre de points moins que *MinPts*). Ces points moins dense ou presque vide peuvent graduellement s'agrandir dans le temps et devenir des points denses comme ils peuvent rester tout le temps vides. Pour ne pas encombrer la mémoire et l'algorithme avec des points qui représentent peut-être des outliers au système, nous avons pensé à ajouter un paramètre en fonction de temps (un âge) qui associe aux points représentatifs moins denses que la limite acceptable une limite d'âge. Ce paramètre qui peut nous permettre de retrouver facilement ceux qui sont restés vides trop longtemps. Pour gérer ce genre de situation nous avons développé l'algorithme (9) qui n'ajoute un point représentatif à la liste des points retenus que s'ils sont assez denses (c'est-à-dire qu'ils sont des points core), sinon ces points représentatifs seront mis dans un *Buffer* (ces points seront dorénavant désignés par *o-Rep* pour *outlier Représentatif*) afin de voir si ces points vont devenir assez denses dans le futur après un certain temps (limite d'âge) ou après un certain nombre d'occurrence des autres cas similaires (limite sur nombre). Ces points seront considérés comme des "outliers" et supprimés du buffer. Le point représentatif de type outlier a une caractéristique différente de celle d'un point

4. CONTRIBUTION À LA CLASSIFICATION DYNAMIQUE POUR LE SUIVI DE ROULEMENTS

Algorithm 9 *UpdateRepList()* : Mise à jour des points représentatifs en ligne

```

for Pour chaque point(s) reçu(s) do
    UpdateExRepList() Ajouter les nouveaux points reçus aux points repré-
sentatifs existants de RepList
    if Les nouveaux points n'ont pas été ajouté aux membres de RepList then
        UpdateExRepList() Ajouter les nouveaux points aux membres de o -
RepList
        CreateRepList() pour créer des nouveaux points représentatifs à partir
des points non ajoutés à RepList ou à o - RepList
        Les o - Rep mis à jour de o - RepList sont considérés comme des
        nouveaux points représentatifs et supprimés de o - RepList
        Trier les points nouveaux représentatifs créés par leur densité
        for Chaque points représentatifs Repj avec une cardinalite CardRepj < MinPts
do
            Ajouter Repj à la liste o - RepList
            Supprimer tout o - Rep de o - RepList qui a dépassé le statut de
limitation
            for Chaque Repj crée avec CardRepj > MinPts do
                Ajouter Repj à RepList

```

représentatif normal, puisqu'il faut lui attribuer un age (un facteur d'oubli). En plus de (*point*, *CovCnt*, *CovRad* et *ClusterId*), il est également caractérisé par (*CT*, et *LV*). Le *CT* ou *creation time* indique l'instant de sa création et *LV Last visit* indique combien de fois le buffer a été visité sans que le point soit mis à jour.

A.2. Deuxième étape (Classification globale) : Attribution des classes Dans cette étape on peut envisager deux situations. Première situation : le(s) nouveau(x) point(s) représentatif(s) est proche(s) d'une classe connue (c'est à dire le nouveau point représentatif *Rep_i* est loin de la classe *j* d'une distance moins de *Eps + Rep_i.CovRad*). Dans ce cas le(s) nouveau(x) point(s) représentatif(s) sera affecté à cette classe et les caractéristiques de la classe seront mis à jour. Deuxième situation : Le(s) nouveau(x) est loin de toutes les classes de la partition. Dans ce cas une re-classification sera nécessaire. L'appartenance de toutes les points représentatifs déjà trouvées sera remis à zéro et une relancement de la classification globale présenté dans l'algorithme *DBDC* est effectué.

B. Adaptation et validation

L'étape d'adaptation et validation sont semblables à ceux de *DDBSCAN*.

4.7.4 Bilan et limitations de la classification par *Evolving scalable DBSCAN*

Durant la conception de la méthode *Evolving Scalable DBSCAN*, notre objectif a été de limiter l'accès aux anciennes observations tout en gardant la nature dynamique de la méthode de classification et d'améliorer les résultats de la classification obtenus auparavant. Nous avons alors adopté le principe des points représentatifs inspiré de la méthode *SDBC* pour minimiser le nombre de points conservés dans le mémoire ce qui réduira par conséquent le temps du calcul. Avec *Evolving Scalable DBSCAN*, l'étape de classification est désormais sur deux étapes : une classification locale où les points représentatifs sont désignés et une classification globale où les points représentatifs sont affectés à des classes. Les étapes d'adaptation et de validation n'ont pas changé.

Cependant la méthode de classification *ESDBSCAN*, les classes ainsi que les voisinages des points cores sont régis par des fonctions d'appartenances strictes. Alors que des fonctions floues sont plus adaptées pour suivre l'évolution d'un système.

4.8 Dynamic Fuzzy Scalable DBSCAN *DFSDBSCAN*

4.8.1 Présentation générale

Le deuxième challenge qu'on a voulu relevé a été d'intégrer la notion des ensembles floues aux règles de décision de *DBSCAN*.

L'intégration des ensembles flous dans le processus de classification par *DBSCAN* a fait que les résultats de d'appartenance d'un élément x à une classe n'est plus sous forme d'un booléen (vrai,faux) mais sous forme d'un degré d'appartenance calculé à partir d'une fonction d'appartenance μ_A .

$$A = (x, \mu_A(x)) | x \in X \quad (4.14)$$

La fonction d'appartenance μ_A quantifie le degré d'appartenance d'un élément x à un ensemble fondamental X . Si la valeur de la fonction est égale à 0 cela signifie que l'élément n'est pas inclus dans l'ensemble donné et si cette valeur est égale à 1 l'élément est

4. CONTRIBUTION À LA CLASSIFICATION DYNAMIQUE POUR LE SUIVI DE ROULEMENTS

pleinement inclus. Les valeurs strictement comprises entre 0 et 1 caractérisent les éléments flous.

Généralement, l'introduction des ensembles flous dans la classification non supervisée se fait par la définition des fonctions d'appartenance en fonction de la distance, de sorte que les degrés d'appartenance expriment la proximité des entités de données des centres des classes. Cependant cette définition ne peut être valable que pour des classes d'une forme sphérique ou ellipsoïdale et pour lesquelles on peut leur attribuer des centres. La définition des centres pour des classes avec des formes irrégulières est très compliquée. Pour cela nous nous sommes intéressés à la classification par **densité** floue.

Différentes stratégies ont été proposées dans la littérature pour introduire les ensembles flous à la classification par densité (SZ13, UN12, GD14, EG14, KM05, NU09, PHK10). Parmi les méthodes proposées, nous avons cherché une méthode qui peut permettre une classification floue en conservant au mieux la nature de *DBSCAN* sans augmenter la complexité du calcul et qui peut faciliter la transition au mode dynamique et au traitement en ligne par la suite.

La méthode "Scalable fuzzy Neighborhood DBSCAN" proposée par Parker *et al* dans leur article (PHK10) est à notre avis l'alternative la plus intéressante. En effet, cette méthode permet d'intégrer les notions des ensembles flous dans la classification. Elle permet également d'alléger le temps de calcul et la mémoire en utilisant la notion de points représentatifs et non toute la bases de données pour effectuer la classification. De plus c'est une méthode de classification distribuée ce qui va faciliter l'adaptation de cette méthode à l'apprentissage en ligne et aux caractéristiques de la classification dynamique. La méthode *SFN-DBSCAN* proposée par Parker *et al*, est en effet la combinaison de deux méthodes : la version floue de *DBSCAN* appelée "Fuzzy neighborhood DBSCAN *FN-DBSCAN*" proposée par (NU09) et la méthode "Scalable Density Dased Distributed Clustering *SDBDC*".

Pour pouvoir expliquer l'algorithme *SFN – DBSCAN* qui servira de base pour la méthode que nous avons développée, Il faut d'abord expliquer les notions utilisées dans les deux méthodes sur lesquelles la méthode s'est inspirée.

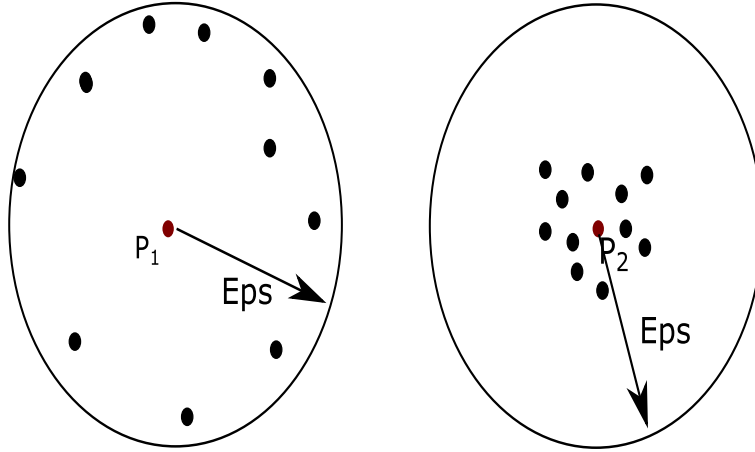


FIGURE 4.8 – Les points P_1 et P_2 ont la même densité selon *DBSCAN*, alors qu'ils ont des densités différentes

4.8.2 Fuzzy Neighborhood DBSCAN

La classification par *FN-DBSCAN*(NU09) est similaire à celle de *DBSCAN*, tout en surmontant un de ses défauts. La *DBSCAN* utilise une définition stricte de densité : un ensemble de n points à une distance Eps d'un point p_1 aura la même densité selon *DBSCAN* qu'un autre ensemble de n points étroitement regroupés autour d'un second point p_2 (voir figure 4.8).

Pour pouvoir donner différentes densités aux deux points p_1 et p_2 dans la figure 4.8, Nasibov *et al* ont proposé de remplacer la définition stricte de densité par une autre floue en utilisant une fonction d'appartenance inspirée de la méthode "fuzzy joint point *FJP*" (NG05) et sa variante "Noise robust fuzzy joint point *NRFJP*" (NU07). La densité à un point p est devenue la somme des degrés d'appartenance des points à distance Eps du point p . Nasibov a proposé différentes fonctions d'appartenance (linéaire, exponentielle et trapézoïdale) pour décrire la façon avec laquelle $N_x(y)$ varie pour les différents points au voisinage Eps (voir figure 4.12).

4. CONTRIBUTION À LA CLASSIFICATION DYNAMIQUE POUR LE SUIVI DE ROULEMENTS

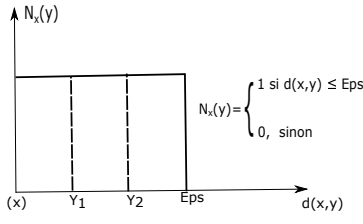


FIGURE 4.9 – Définition ferme de densité (DBSCAN)

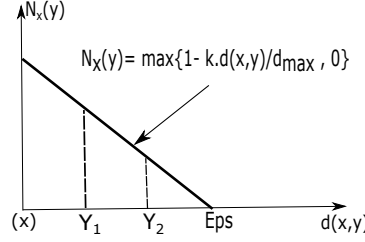


FIGURE 4.10 – Définition floue par une fonction d'appartenance linéaire

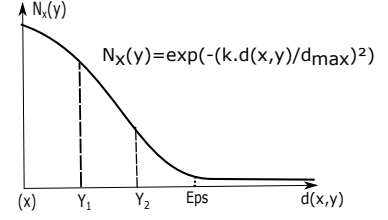


FIGURE 4.11 – Définition floue par une fonction d'appartenance exponentielle

FIGURE 4.12 – Différence entre une définition ferme et floue de voisinage d'un point x

N.B : Il faut noter que les données utilisées dans cette méthode doivent être normalisées avant l'application de FN-DBSCAN, les valeurs normalisées des données et la valeur de Eps sont donc toujours entre $[0,1]$.

L'intégration des fonctions d'appartenance floues a impliqué des changements des définitions des points cores et l'ensemble de voisinage.

- L'ensemble de voisinage avant l'intégration des ensembles floues a été définie comme suit : Etant donnée un ensemble de voisinage de points $x \in X$ avec le paramètre Eps (Eps définie la limite de voisinage)

$$N(x; Eps) = \{y \in X | d(x, y) \leq Eps\} \quad (4.15)$$

L'ensemble de voisinage déterminé avec l'équation 4.15 sera constitué de tous les points trouvés dans un rayon de Eps de point x . Cet ensemble pourra être écrit sous la forme suivante :

$$N_x(y) = \begin{cases} 1 & \text{si } d(x, y) \leq Eps \\ 0 & \text{sinon} \end{cases}$$

Avec $N_x(y)$ est le degré d'appartenance d'un point y à l'ensemble de voisinage de point x .

Après le changement des fonctions d'appartenance pour intégrer les ensembles floues, l'ensemble floue de voisinage est devenu :

$$FN(x; \varepsilon_1) = \{\langle y, N_x(y) \rangle | y \in X, N_x(y) \geq \varepsilon_1\} \quad (4.16)$$

$N_x : X \rightarrow [0, 1]$ est une fonction d'appartenance choisie pour déterminer la relation de voisinage entre les points (par exemple une fonction linéaire présentée dans la figure 4.10 ou une fonction exponentielle figure 4.11).

Il faut noter que le paramètre ε_1 utilisé dans l'équation 4.16 détermine le seuil minimal de degré d'appartenance alors que Eps détermine le seuil maximal de la distance entre deux points.

- La définition d'un point core : Dans *DBSCAN*, un point $x \in X$ avec les paramètres Eps et $MinPts$ est définie comme un point core, s'il satisfait la condition suivante :

$$|N(x; Eps)| \geq MinPts \quad (4.17)$$

$|N(x, Eps|$ définit la cardinalité de l'ensemble $N(x; Eps)$. Cette définition sera modifiée, et la définition d'un point core floue x avec les deux paramètres ε_1 et ε_2 sera :

$$Card\ FN(x; \varepsilon_1, \varepsilon_2) \equiv \sum_{y \in N(x; \varepsilon_1)} N_x(y) \geq \varepsilon_2 \quad (4.18)$$

Pour tout point $x \in X$,

$$N(x; \varepsilon_1) = \{y \in |N_x(y) \geq \varepsilon_1\} \quad (4.19)$$

Le paramètre ε_2 est lié au paramètre $MinPts$ de *DBSCAN* par la relation suivante :

$$\varepsilon_2 = \frac{MinPts}{\omega_{max}} \quad (4.20)$$

Avec $\omega_{max} = \max_{i=1, \dots, n} \omega_i$ et ω_i est la cardinalité de l'ensemble des points trouvés à un rayon Eps du point x_i .

$$\omega_i = |N(x_i; Eps)| \quad (4.21)$$

Les définitions 4.16 et 4.18 sont utilisées dans l'algorithme *FN-DBSCAN* à la place des définitions 4.15 et 4.17 de *DBSCAN*. La définition 4.18 diffère de la définition 4.17 dans la caractérisation de l'ensemble de voisinage en se basant sur un niveau minimal d'appartenance et non sur une distance et elle utilise un concept floue de cardinalité pour

4. CONTRIBUTION À LA CLASSIFICATION DYNAMIQUE POUR LE SUIVI DE ROULEMENTS

déterminer un point core floue à la place d'une cardinalité classique.

L'algorithme de la méthode FN-DBSCAN est comme suit :

Algorithm 10 L'algorithme de la classification par FN-DBSCAN

- Etape 1 Attribuer des valeurs aux paramètres $\varepsilon_1, \varepsilon_2, f$ la fonction d'appartenance floue
 - Etape 2 Marquer tous les points de la base de données comme "non classifiés". fixer $t=1$.
 - Etape 3 Trouver un point core floue non classifié avec les paramètres ε_1 et ε_2 .
 - Etape 4 Marquer p comme un point "à classifier". Créer un nouveau cluster C_t et affectuer p à ce cluster.
 - Etape 5 Créer un ensemble S vide. Trouver tous les points non classifiés dans $N(p; \varepsilon_1)$ et affecter tous ces points à l'ensemble S .
 - Etape 6 Obtenir un point p de l'ensemble S , marquer q "à classifier", attribuer q au cluster C_t , et supprimer q de l'ensemble S .
 - Etape 7 Vérifier si q est un core floue avec les deux paramètres ε_1 et ε_2 dans l'affirmative, ajouter tous les points non classifiés de $N(q; \varepsilon_1)$ à l'ensemble S .
 - Etape 8 Répéter les étapes 6 et 7 jusqu'à ce que l'ensemble S se vide.
 - Etape 9 Trouver un nouveau point core floue p avec les paramètres ε_1 et ε_2 et répéter les étapes de 4 à 7.
 - Etape 10 Marquer tous les points qui n'appartiennent à aucun cluster comme 'bruit'.
-

Il faut noter que *FN-DBSCAN* devient similaire à *DBSCAN* quand la fonction d'appartenance choisi dépend de la distance est non du degré d'appartenance.

4.8.3 Scalable fuzzy Neighborhood DBSCAN ou *SFN – DBSCAN*

La méthode *SFN-DBSCAN* ou *Scalable fuzzy Neighborhood DBSCAN* combine les deux méthodes *DBDC* et *FN-DBSCAN* dans le sens où elle applique le principe de *DBDC* sur une base de données divisée en sous ensembles où chaque sous ensemble est représenté par sa densité et son poids. Les points retenus dans chaque sous ensemble sont combinés et une version pondérée de *FN-DBSCAN* est appliquée pour retrouver les clusters des points. Finalement chaque point est affecté au cluster adéquat.

L'algorithme de la méthode *SFN-DBSCAN* est décrit comme suit :

Algorithm 11 L'algorithme de la classification par SFN-DBSCAN

```

1. Attribuer des valeurs à  $Eps$ ,  $MinCard$ , et  $f$ 
2. Diviser La base de données en sous-ensembles de la même taille où chaque sous-ensemble contient un point sélectionné aléatoirement
3. Calculer  $MaxRetainedPts$  le nombre maximal de points à retenir dans chaque sous-ensemble
for Chaque sous-ensemble do
    Calculer la densité locale de chaque point en calculant la somme des valeurs d'appartenance floue de tous les points trouvés dans un rayon de  $Eps$ .
    Trier les points par leur densité
    Ajouter les points les plus denses à la listes des points retenus ou représentant  $RetainedPtList$ 
    while  $Size(RetainedPtList) < Size(MaxRetainedPts)$  do
        Ajouter les points les plus denses et qui sont en dehors du rayon  $Eps$  de tous les points actuellement présents dans  $RetainedPtList$  à cette liste
        Attribuer la densité de chaque point dans la liste  $RetainedPtList$  comme son poids
5. Combiner les points de  $RetainedPtList$  de chaque sous-ensemble en une seule base de données.
6. Réduire  $MinCard$  en respectant le nombre des sous-ensembles.
7. Exécuter la version pondérée de FN-DBSCAN sur la base de données combinées.
8. Trouver les points core de la nouvelle base de données pondérée.
9. Distribuer les points cores et calculer le degré d'appartenance de la base de données entière vis à vis de ces points cores de la liste pondérée.

```

Il faut noter que dans la 3ème étape, le nombre des points retenus ou le nombre de représentants par sous-ensemble est calculé en fonction du nombre de sous-ensembles. Dans la 5ème étape tous les points sont recombinaés et le nombre de points retenus est au maximum égal au nombre de sous ensembles. Dans la 6ème étape, la réduction de la valeur de $MinCard$ fait baisser la valeur de ε_2 de *DBDC*. Cela implique une réduction de la densité requise pour former un cluster ce qui est nécessaire quand il s'agit d'un nombre réduit des exemples. Dans cette implémentation, cette réduction de densité est effectuée en divisant $MinCard$ par le logarithme du nombre des sous-ensembles. L'utilisateur peut aussi choisir de réduire $MinCard$ d'une autre façon si il le désire. Dans la 7ème étape, une version pondérée de *FN-DBSCAN* est exécutée. A chaque point x on attribue un poids ω et la cardinalité d'un point est calculée comme suit :

4. CONTRIBUTION À LA CLASSIFICATION DYNAMIQUE POUR LE SUIVI DE ROULEMENTS

$$Card(x) = \sum_{i=1}^n N_x(y_i)\omega_i \quad (4.22)$$

$N_x(y_i)$ représente la fonction d'appartenance basée sur la proximité de x à y_i et ω_i est le poids de y_i .

Bien que la méthode *SFN-DBSCAN* permet d'élargir le champ d'utilisation de la classification par densité en introduisant la logique floue dans son processus et en permettant une classification distribuée, elle présente plusieurs inconvénients. Par exemple la scission des clusters normalement fusionnés ; ce phénomène a été observé pendant une série de tests sur plusieurs bases de données. Des clusters trouvés par *FN-DBSCAN* ont été divisés en deux ou en plusieurs par *SFN-DBSCAN* alors qu'ils ne devraient constituer qu'une seule classe. Cette scission peut être attribuée à l'affectation arbitraire des points aux sous-ensembles. Si une classe de forme irrégulière est reliée par un seul point, l'omission de ce point dans la liste de points retenus résultera en la scission de cette classe en deux. Le deuxième problème rencontré par *SFN-DBSCAN* concernent les points considérés comme du bruit ; la caractérisation du bruit dans *SFN-DBSCAN* n'est pas pareille que celle dans *DBSCAN*. Dans *DBSCAN* un outlier est un point qui n'a pas assez de points dans son rayon, alors que dans *SFN-DBSCAN* les points qui sont marqués comme du bruit sont des points représentatifs ou des points retenus c'est dire qu'ils ont déjà un nombre de points dans leurs entourages mais qui sont loin des autres points représentants de la classe (voir figure 5.18b). Ces points qui peuvent être considérés par *DBSCAN* comme des points de bordure ou "border point" sont considérés comme des outliers dans *SFN-DBSCAN*.

Dans l'exemple montré dans la figure 5.18b, le point P_3 est loin de tous les autres points. De plus, $Eps + CovRad(P_3)$ sera qualifié comme un outlier par *SFN-DBSCAN* alors qu'il se peut que le nombre des points dans $CovRad()$ est supérieur à la limite définie par $Minpts$. Dans ce cas, ce point ainsi que les points qu'il représente sera attribué à une nouvelle classe. Pour résoudre ce problème et introduire la notion d'apprentissage dynamique à cette méthode nous avons proposée la méthode "*Dynamic Fuzzy DBSCAN DFDBSCAN*".

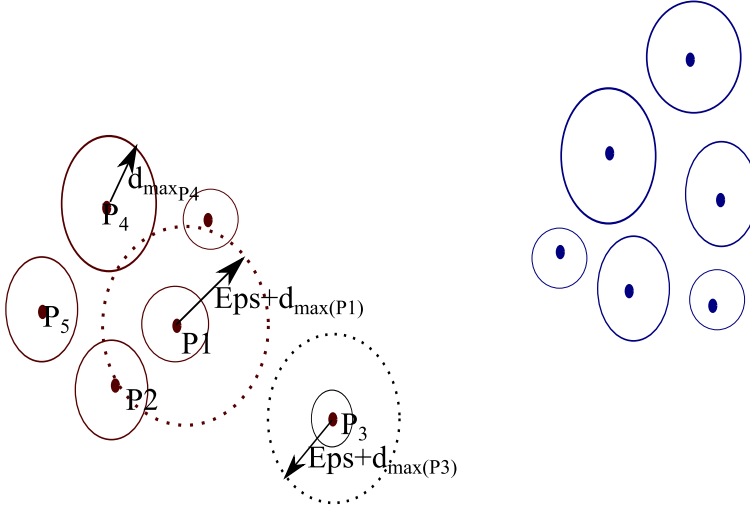


FIGURE 4.13 – Outlier par SFN-DBSCAN

4.8.4 Dynamic fuzzy scalable DBSCAN *DFSDBSCAN*

Pour rendre la méthode *SFN-DBSCAN* dynamique et pour l'adapter à l'apprentissage en ligne, il faut d'abord s'assurer que cette méthode peut détecter les évolutions des classes et d'intégrer les informations de ces évolutions dans son processus de classification. Il faut s'assurer également que cette méthode peut créer des nouvelles classes en cas de besoin, de fusionner des classes qui sont trop proches et de diviser une classe en plusieurs. Il faut ensuite gérer les limitations de *SFN-DBSCAN* en termes d'identification des points bordures comme des outliers. Finalement, il faut donc adapter *SFN-DBSCAN* à l'apprentissage en ligne par l'ajout des fonctionnalités de traitement ponctuel de chaque enregistrement reçu tout en gardant un accès aux résultats globaux.

Dans notre première adaptation de *SFN-DBSCAN*, nous avons essayé d'établir un processus de classification à deux étapes. Dans la première phase, le processus commence par retrouver les points représentatifs et à l'arrivée de chaque enregistrement, il va chercher à voir si on peut intégrer ou non ce point dans un point représentatif déjà existant. Si ce n'est pas possible il va créer un nouveau point représentatif. Dans la deuxième phase, il va chercher à attribuer le point représentatif à sa classe en calculant son degré d'appartenance. Une fois le point représentatif classifié, on va voir si cette attribution a changé la structure de la partition (chevauchement de deux classes par exemple) et ainsi valider

4. CONTRIBUTION À LA CLASSIFICATION DYNAMIQUE POUR LE SUIVI DE ROULEMENTS

la partition trouvée. Parallèlement, on va étudier l'évolution des points à l'intérieur de chaque classe en utilisant la fenêtre glissante *MWFA* introduite dans la section 4.4.3. La nouvelle version de *SFN-DBSCAN* que nous proposons d'appeler *DFSDBSCAN* pour *Dynamic fuzzy scalable DBSCAN* suit alors trois étapes : classification, adaptation et validation. L'étape classification se divise en deux parties ; une classification ponctuelle afin de désigner les points représentants après réception des nouvelles données et une classification globale afin d'attribuer les données à leurs classes appropriées.

A. La classification ponctuelle : Désignation des points représentatifs

La désignation des points représentatifs est semblable à celle présentée dans la méthode *ESDBSCAN* avec quelques changements au niveau de la configuration d'un point représentatif où nous avons intégré la notion du poids et de la cardinalité. Cette intégration va altérer quelques étapes dans les algorithmes de *UpdateRepList* et *UpdateExRepList* 7 et 8 qui se feront remplacé par les algorithmes 12 et 13. Les nouveaux algorithmes sont comme suit. La mise à jour des représentatifs déjà connus après la réception d'un nouveau bloc de points est expliquée dans l'algorithme 7.

Algorithm 12 *UpdateExRepList()* : Mise à jour des points représentatifs existants

UpdateExRepList() :

for Chaque nouveau point P_i de P **do**

Trouver les points dans son entourage Eps $S = RangeQuery(P_i, Eps)$;

if Tester si dans l'entourage de P_i il existe un point représentatif de *RepList* **then**

Assigner le point P_i **au point représentatif le plus proche** $Rep_j \in S$

$w_{Rep_j} = w_{Rep_j} + 1$,

$Card_{Rep_j} = Card_{Rep_j} + N_{P_i} w_{P_i}$,

MarkedAsAssigned(P_i) ;

La première étape de la mise à jour des points représentatifs existants consiste à mettre à jour leur poids (w_{Rep_i}) et leurs cardinalités (Card) si un ou plusieurs nouveaux points sont dans un rayon Eps de ces points retenus (voir algorithme 12). Sinon on crée des nouveaux points représentatifs à ajouter à la liste actuelle en suivant l'algorithme 13.

Algorithm 13 Mise à jour de la liste *RepList* existante par des nouveaux points représentatifs

```

for Chaque  $P_i$  non assigné à aucun point représentatif de RepList do
    Calculer  $DynRepQ(P_i, Eps, RepList)$ 
    Trier Liste des points représentatifs candidats  $SortRepList =$ 
 $Sort(DynRepQ, 'descend')$   $Rep_i = SortRepList(1)$ 
    Calculer le poids  $w_{Rep_i}$  de  $Rep_i$   $w_{Rep_i} = CntObjects(RangeQuery(Rep_i, Eps))$ 
    Calculer la cardinalité de  $Rep_i$   $Card(Rep_i) = \sum_{i=1}^n N_{Rep_i}(y_i)w_{y_i}$ 
    Ajouter  $Rep_i$  à RepList

```

A.1. La classification globale : attribution des points représentants à leurs classes appropriées

L'attribution des points représentatifs à leurs classes appropriées dépend de plusieurs facteurs :

- Si les nouveaux points enregistrés ne nécessitent pas la création de nouveaux points représentatifs mais la mise à jour des points représentatifs existants, les nouveaux points auront le même ClusterId que leurs points représentants.
- Si les nouveaux points enregistrés nécessitent la création des nouveaux points représentatifs :
 - Si ces points n'ont pas été marqués par l'algorithme de détection de nouveautés, les nouveaux points représentatifs seront affectés à la classe la plus proche.
 - Si ces points ont été marqués comme nouveautés et qu'il n'y a aucun point représentatif assez proche d'eux (c'est à dire qu'ils sont loin de la classe actuelle et aussi de toutes les classes déjà trouvées). Si les conditions de création d'une nouvelle classe sont satisfaites, une nouvelle classe sera créée. Sinon ces points seront mis dans le buffer pour être analysés après (une classe transitoire est soupçonnée). Ces points seront revisités durant les futurs enregistrements, s'ils restent dans le buffer plus que la limite définie, ils seront considérés comme outliers et seront supprimés de la mémoire tampon (buffer).

4. CONTRIBUTION À LA CLASSIFICATION DYNAMIQUE POUR LE SUIVI DE ROULEMENTS

- Si ces points ont été détecté comme nouveautés mais qu'ils sont proche d'un ou plusieurs point(s) représentatif(s) existant, ces points vont être assigner à la classe la plus proche(une dérive peut être soupçonnée).

Algorithm 14 Classification globale

Obtenir les nouveaux points de données

if *UpdateRepList()* a créé des nouveaux points représentatifs **then**

if Une nouveauté est détectée **then**

Attribuer les points représentatifs à leurs classes en exécutant la version pondérée de *FDBSCAN* et en gardant l'appartenance des points représentatifs déjà classifié

else

Attribuer les points représentatifs à la classe actuelle

Mettre à jour les classes de la partition

Pour garder une certaine trace de l'évolution des classes trouvées par *DFSDBSCAN*, on calcule certain paramètres après chaque insertion d'un nouveau point à une classe de partition. Les paramètres permettant de collecter des informations sur l'évolution des classes dans le temps sont :

- La distance Hausdorff : on calcule la distance entre la classe actuelle et les autres classes de la partition :

$$h_{ij} = h(C_i, C_j); \forall j \leq N | i \neq j \quad (4.23)$$

Avec N le nombre de classes et C_i la classe actuelle.

- La densité de la classe : on calcule la compacité de chaque classe :

$$D_{C_i} = \sum_{j=1}^{n_{c_i}} CovCnt(r_{ij}) \quad (4.24)$$

Avec r_{ij} le j 'ième point retenu de la classe C_i et n_{c_i} le nombre de points retenus dans la classe.

- Le couverture de la classe : C'est à dire que l'on va calculer la zone couverte par la classe en se servant des informations *CovRad* des points retenus

$$Cov_{C_i} = \sum_{j=1}^{n_{c_i}} CovRad(r_{ij}) \quad (4.25)$$

- On calcule aussi le centre de la chaque classe.

$$C_{C_i} = \frac{C_{C_{i-1}} \times n_{c_i}}{n_{c_i} + 1} + \frac{r_{inc_i}}{n_{c_i} + 1} \quad (4.26)$$

Avec r_{inc_i} le nouveau point retenu de la classe C_i

Après la classification vient l'étape détection/adaptation et puis validation.

B. Détection et adaptation

L'inconvénient majeur de *Scalable fuzzy DBSCAN* est qu'il peut classifier des points représentatifs naturellement classés dans une seule classe alors qu'ils sont classés dans des classes différentes, ce qui impose l'introduction d'un outil permettant la fusion des classes similaires. En plus, durant la classification des données évolutives, il se peut aussi que deux classes initialement disjointes se chevauchent en créant une région partagée par les deux classes ce qui enfreint la condition de continuité de la région décrite par chaque classe. La solution est donc de créer une règle de fusion des classes trop proches pour en créer une nouvelle. Généralement, la règle de fusion est souvent établie sur la cardinalité de la région partagée par les classes qui se chevauchent en définissant un seuil de similarité entre classes. Une fois ce seuil est dépassé les classes qui se chevauchent doivent être fusionnées. Cependant, la définition du seuil est une tâche difficile et dépend de l'application. Pour *DFDBSCAN*, la règle de fusion que nous avons choisi d'utiliser permet de mesurer la similarité en basant sur les valeurs d'appartenance des points représentatifs des classes de la partition. Cette mesure a été proposée par Frigui *et al* dans (HK96) et elle est définie par :

$$\delta_{iz} = 1 - \frac{\sum_{x \in C_i \text{ et } x \in C_z} |\pi_i(x) - \pi_z(x)|}{\sum_{x \in C_i} \pi_i(x) + \sum_{x \in C_z} \pi_z(x)} \quad (4.27)$$

$\pi_i(x)$ et $\pi_z(x)$ représentent les valeurs d'appartenance de x à C_i et à C_z . De cette expression, on peut conclure que plus δ_{iz} se rapproche de 1, plus les deux classes sont similaires. La valeur du suil à partir duquel on juge que deux classes sont trop similaires et que les deux classes doivent être fussionnées (th_f) est définie dans la partie application.

4. CONTRIBUTION À LA CLASSIFICATION DYNAMIQUE POUR LE SUIVI DE ROULEMENTS

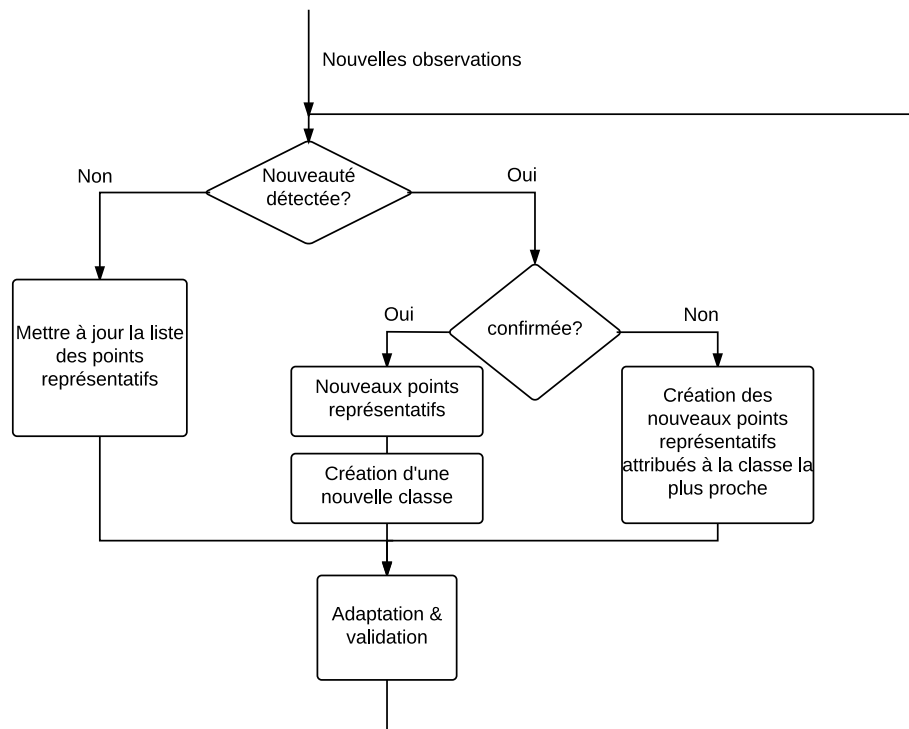


FIGURE 4.14 – L’organigramme de la classification par *DFSBSCAN*

C. La validation

L’étape validation a pour objectif de s’assurer que la partition ne contient que les classes utiles en éliminant le bruit, les classes parasites et les classes obsolètes. On peut créer une nouvelle règle pour gérer le cas des classes transitoires ou le cas des classes obsolètes. En associant aux classes un facteur d’oubli, on peut déterminer l’âge de la classe et déterminer la date de la dernière action effectuée sur cette classe ce qui nous permettra de vérifier si la classe est toujours active ou si elle est devenue obsolète ou s’il s’agit d’une classe transitoire définissant un phénomène éphémère.

4.8.5 Bilan et limitations de la classification par *DFSDBSCAN*

Comme on peut voir sur l’organigramme affiché dans la figure 4.14, la structure globale de la méthode *DFSBSCAN* ressemble à celui de *ESDBSCAN* (voir figure 4.7) avec quelques modifications dans la construction de voisinage, le choix des points cores et de la nature non-supervisée qui est devenue semi-supervisée.

Cependant, comme toutes méthodes de ce type, la qualité de la classification dépend de l'initiation des différents paramètres.

4.9 Bilan sur les classifieurs proposés

Trois méthodes de classification ont été proposé dans cette thèse. Chaque méthode de classification a été développée pour répondre à une contrainte spécifique concernant la nature en ligne et non stationnaire des données.

Dans le tableau 4.3, les méthodes proposées ainsi que la méthode de base *DBSCAN* se comparent pour souligner les caractéristiques de chacune.

TABLE 4.3 – Tableau Comparatif des classifieurs proposés

Méthode employé	Complexité	Dynamique	En ligne	Mémoire	Définition de la densité
DBSCCAN	$\oplus\oplus\oplus$	$\ominus$	$\ominus$	$\ominus$	$\ominus$
DDBSCAN	$\oplus\oplus$	$\oplus$	$\circ$	$\ominus$	$\ominus$
ESDBSCAN	$\oplus$	$\oplus$	$\oplus$	$\oplus$	$\oplus$
DFSDBSCAN	$\circ$	$\oplus$	$\oplus$	$\oplus$	$\oplus\oplus$

4.10 Exploitation de la classification dynamique dans le suivi

Durant la classification, nous avons défini des paramètres de classes afin de pouvoir tracer l'évolution de ces paramètres dans le temps et voir si on peut associer ces évolutions à des phénomènes physiques. On peut rencontrer deux types d'évolution dans la classification dynamique : une évolution interne à la classe et une évolution de la partition. Les deux types d'évolution sont intéressantes du point de vue suivi et les deux peuvent être porteurs d'information.

L'évolution interne à la classe peut se manifester par un glissement du centre de la classe, par un changement de densité, etc.. Les évolutions internes de la classe peuvent être utilisées pour détecter un changement de condition de fonctionnement (augmentation de charge par exemple), une évolution de la taille de défaut d'un composant, etc.. Détecter ce type d'évolution peut nous aider à quantifier la vitesse avec laquelle le composant suivi se dégrade.

4. CONTRIBUTION À LA CLASSIFICATION DYNAMIQUE POUR LE SUIVI DE ROULEMENTS

L'évolution de la partition peut toujours être liée à l'état interne du composant suivi, qu'elle soit sous forme de création d'une nouvelle classe, de fusion de deux classes ou d'autres formes d'évolution. Dans le cas du suivi de roulement, nous avons pu constater que la naissance d'un défaut se manifeste toujours par un saut d'observation conduisant à la création d'une nouvelle classe. Par contre, la fin de vie de roulement se manifeste par une dérive d'observations dirigée vers la classe qui représente un roulement sain. Ces deux types d'évolutions peuvent nous aider à établir un diagnostic précoce et un pronostic fiable.

4.11 Conclusion du chapitre

Dans ce chapitre, nous avons présenté la méthode de RdF conçue pour le diagnostic et suivi de roulements. On a commencé par présenter l'architecture générale de la méthode proposée, nous avons aussi détaillé chaque technique utilisée dans l'étape extraction et sélection. Dans la partie classification nous avons présenté les méthodes développées pour le suivi de roulement. La première méthode proposée *DDBSCAN* a été inspirée de l'algorithme *DBSCAN* pour lequel nous l'avons ajouté deux couches 'détection & adaptation' et 'validation'. Ces deux étapes supplémentaires ont permis de rendre l'algorithme dynamique. Cependant, la méthode *DDBSCAN* nécessite l'accès à la totalité de données. Nous avons alors cherché à résoudre ce problème dans la deuxième méthode proposée *ESDBSCAN* qui est une version de *DDBSCAN* adaptée au traitement en ligne et capable de classer correctement en ne gardant en mémoire que quelques exemples de données déjà classées par la méthodes. Une troisième méthode de classification *DFSDBSCAN* qui intègre la logique floue dans sa définition de voisinage de points.

Ces méthodes de classification dynamiques détectent les évolutions des classes et s'auto-adaptent en conséquence. Ces évolutions qui peuvent nous donner accès à des informations cachées et que nous nous pouvons pas y accéder avec les techniques statiques. Le fait que ces techniques s'auto-adaptent va améliorer la séparation des différentes classes (modes de fonctionnement).

Chapitre 5

Validation expérimentale et mise en œuvre sur un banc de fatigue

Introduction

La validation des méthodes proposées dans le chapitre précédent est réalisée à travers deux systèmes tournants. Nous avons d'abord testé ces méthodes sur un banc de roulements avec des défauts artificiellement initiés sous différentes charges. Nous avons ensuite réalisé un suivi de roulements sur un banc de fatigue sous une charge constante et vitesse constante. La procédure de suivi par classification dynamique est présentée par l'organigramme suivant.

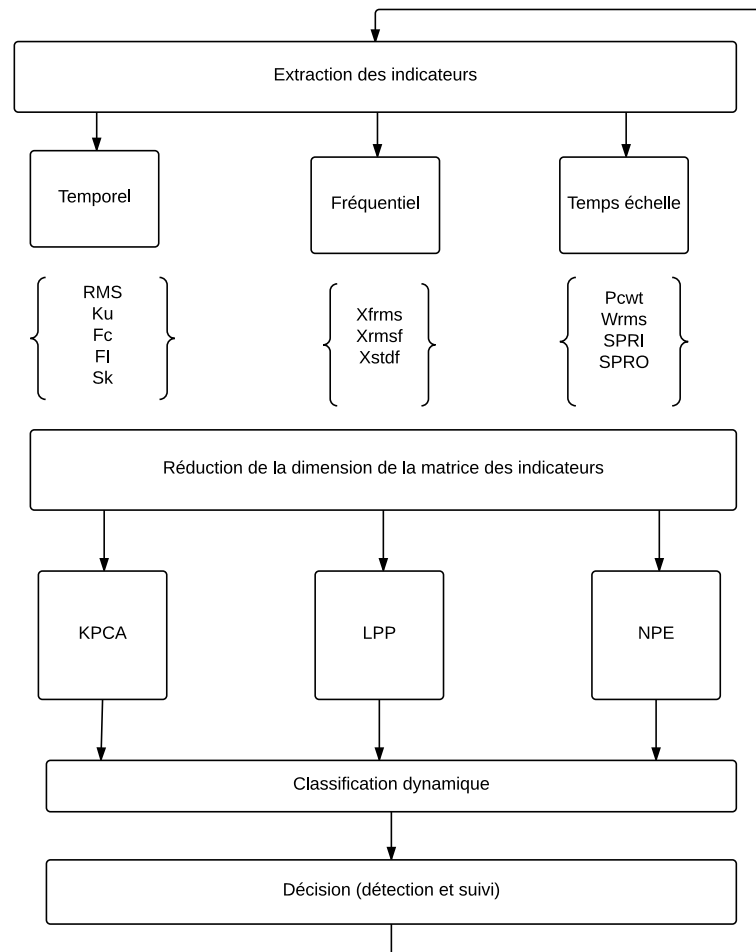


FIGURE 5.1 – Procédure de suivi par la classification dynamique

5.1 Mise en œuvre de la méthode de classification dynamique sur un banc expérimental

5.1.1 Description du modèle mécanique et du banc expérimental

La validation expérimentale est réalisée sur un banc "SURVIB" de l'université de Reims Champagne Ardenne. Ce banc est constitué d'un bâti massif en béton sur lequel est montée une plaque rainurée en acier permet de fixer le module de motorisation. Viennent ensuite s'ajouter des modules démontables d'étude de composants mécaniques (un module roulement, un module engrenage et un module fatigue) et enfin, un module moteur frein vient créer le couple. L'ensemble est boulonné sur une structure en béton pour l'isoler des basses fréquences engendrées par l'environnement extérieur 5.2. Le module utilisé est un carter constitué d'un arbre et deux roulements, un chargement sur l'arbre est réalisé. Par électroérosion, des défauts de différentes tailles sont générés sur la bague extérieure (2mm^2 jusqu'à 16mm^2).



(a) Banc expérimental.



(b) Les roulements

FIGURE 5.2 – Banc expérimental et les roulements utilisés durant l'essai

Un capteur piézoélectrique est placé radialement sur le palier du roulement (sain et défectueux), considéré comme le meilleur point de mesure. Grâce aux caractéristiques du roulement (roulement 6206) et à la cinématique du système ($1000\text{tr}/\text{mn}$), les fréquences caractéristiques peuvent être calculées : 59.5Hz pour fréquence de passage des billes sur la

5. VALIDATION EXPÉRIMENTALE ET MISE EN ŒUVRE SUR UN BANC DE FATIGUE

piste extérieure BPFO et 90.5Hz pour BPFI fréquence de passage des billes sur la piste intérieure. Nous avons utilisé le système d'acquisition "Siglab". Nous avons enregistré 8192 points du signal vibratoire dans une gamme de fréquence de 0-20 kHz pour avoir un nombre de cycle significatif. Chaque signal enregistré a ensuite été découpé en quatre signaux avec un taux de recouvrement de 25%.

5.1.2 Extraction des caractéristiques physiques

Pour étudier l'apport de ces caractéristiques et choisir ceux qui nous permettront de bien discerner les signaux extraits d'un roulement sain de ceux d'un roulement défectueux même quand la machine fonctionne sous conditions non stationnaires, dans la section suivante on va étudier chaque type d'indicateurs utilisés et le tester pour voir son comportement dans les différents cas. Il faut noter que la qualité des indicateurs extraits est définie par la différence entre les grandeurs de ces indicateurs pour un roulement sain et leurs grandeurs pour un roulement défectueux ce qui fiabilisera le diagnostic.

A. Indicateurs issus de l'analyse temporelle

Nous avons évalué les performances de certains indicateurs temporels (à savoir : RMS, Kurtosis, Skewness, Fc et FI) afin de mesurer leur sensibilité en fonction de la taille du défaut et de la variation de la charge appliquée sur le roulement. Nous avons utilisé 13 roulements dont 12 sont défectueux avec des tailles de défaut différentes. La figure 5.3 montre l'évolution des valeurs des indicateurs en fonction de la taille du défaut de roulement. Sur cette figure on peut voir que les valeurs de certains indicateurs diminuent quand le défaut atteint une certaine taille (Par exemple Kurtosis). Cette diminution de valeur peut être expliquée par le fait que la surface de défaut est devenue assez grande pour couvrir deux éléments roulants.

Les figures 5.4 et 5.5 montrent l'évolution des valeurs des indicateurs temporels en fonction de la taille du défaut de roulement pour différentes valeurs de la charge appliquée sur le roulement. Ainsi, on peut tirer les conclusions suivantes sur chaque indicateur utilisé :

- **Valeur efficace x_{RMS} et la valeur maximale x_{Peak} :**
 - Dans tous les cas de figures, l'évolution de x_{RMS} et celle de la valeur crête (x_{Peak}) sont semblables .

5.1 Mise en œuvre de la méthode de classification dynamique sur un banc expérimental

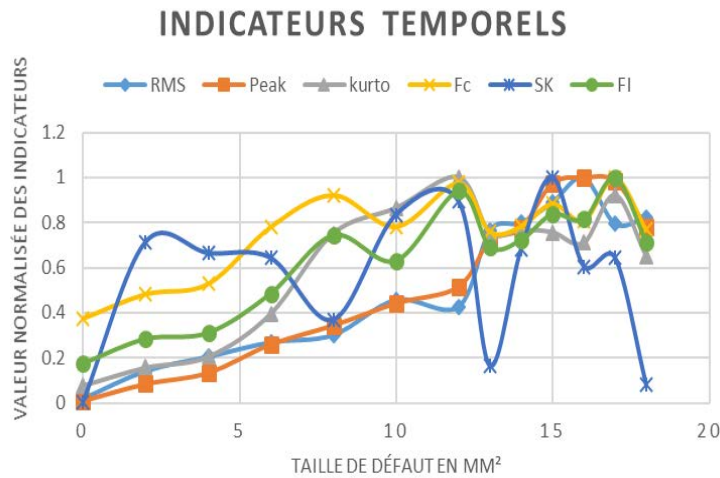


FIGURE 5.3 – Evolution des indicateurs en présence de défaut en cas d'un roulement chargé à 15DaN

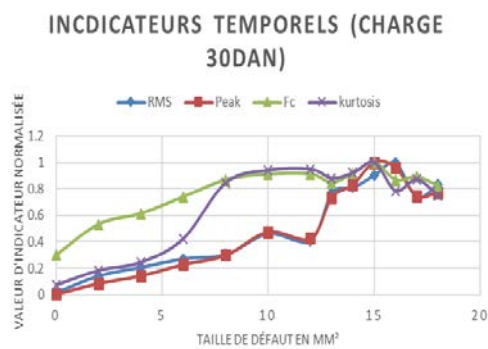


FIGURE 5.4 – Comparaison des indicateurs en présence de défaut pour un roulement chargé à 30 DaN

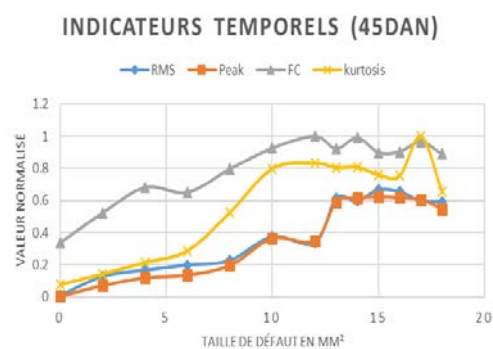


FIGURE 5.5 – Comparaison des indicateurs en présence de défaut pour un roulement chargé à 45 DaN

- La valeur de x_{RMS} et x_{Peak} augmente en présence de défaut et continue à augmenter avec la détérioration de défaut, puis commence à diminuer après que le défaut atteigne la taille $16 mm^2$

- **kurtosis :**

- Dans tous les cas de figures, la valeur du *kurtosis* augmente au fur et à mesure que la taille de défaut augmente jusqu'à atteindre son maximum puis diminue

5. VALIDATION EXPÉRIMENTALE ET MISE EN ŒUVRE SUR UN BANC DE FATIGUE

pour une surface de défaut spécifique.

- La taille de défaut qui entraîne la diminution de la valeur du *kurtosis* change en changeant la charge.

- **Facteur crête x_{Fc}**

- Même en l'absence de défaut, la valeur de x_{Fc} augmente en fonction de la charge.
- La valeur de x_{Fc} augmente également en présence de défaut et continue à augmenter avec l'avancement de dégradation.

En contraste avec ce que nous avons trouvé dans la littérature (LP92), où les auteurs ont affirmé que le *kurtosis* et Fc sont les indicateurs les plus sensibles à la présence de défaut et les plus insensible aux changements de conditions de fonctionnement, le *kurtosis*, comme les autres indicateurs temporels classiques, n'est pas insensible aux variations de la charge.

B. Indicateurs de défaut issus de l'analyse spectrale

Pour voir la réaction des indicateurs de défaut issus de l'analyse spectrale (x_{frms} , x_{rmsf} et x_{stdf}) dans la présence des défauts de roulements, ces indicateurs ont été calculés à partir des spectres de puissance des signaux vibratoires. Dans la figure 5.6, on peut

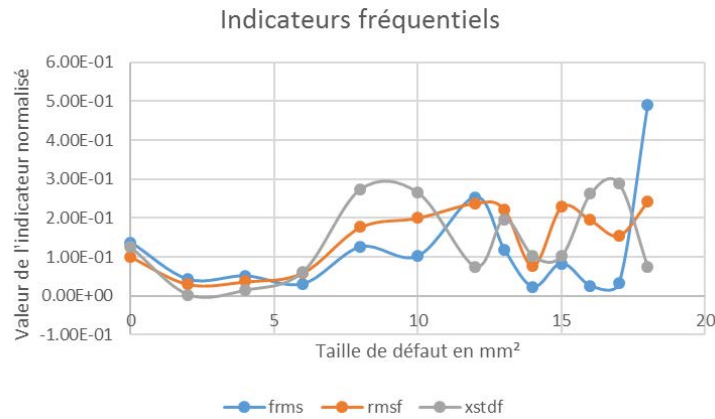


FIGURE 5.6 – les indicateurs fréquentiels calculés à partir de la puissance de spectre

voir l'évolution des indicateurs fréquentiels en cas de défaut, ce qui nous donne une

5.1 Mise en œuvre de la méthode de classification dynamique sur un banc expérimental

idée générale sur leur capacité à distinguer entre un roulement sain et un roulement défectueux.

Bien que les indicateurs fréquentiels affichés dans la figure 5.6 montrent un écart remarquable entre un roulement sain et un roulement avec un petit défaut (2 mm^2) cet écart se réduit avec le développement de défaut ce qui peut mener le système de diagnostic automatique à confondre l'état sain de l'état défaut avancé. L'application de l'analyse de

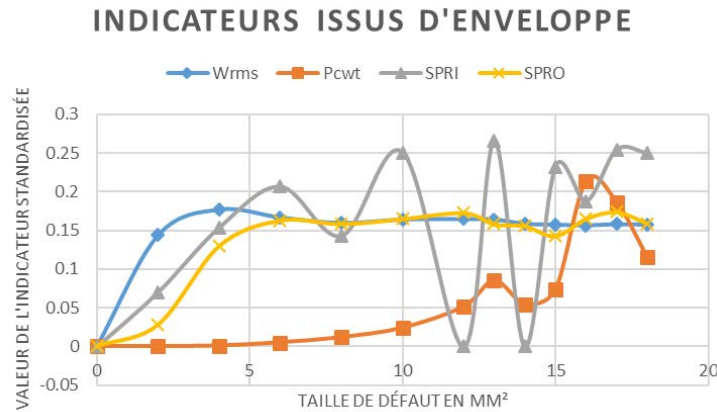


FIGURE 5.7 – Les indicateurs d'enveloppe calculés à partir de la puissance de spectre extrait par les ondelettes

l'enveloppe sur les signaux a nettement amélioré la détection comme on peut le voir sur la figure 5.7 que certains indicateurs ont affiché une meilleure séparation entre état sain et état défectueux.

On peut voir l'évolution des indicateurs issus de l'analyse d'enveloppe extrait par les ondelettes dans sur la figure (5.7), les deux indicateurs X_{PCWT} et $X_{W_{RMS}}$ ont bien montré qu'ils réagissent à la présence de défaut par une augmentation quand le roulement passe d'un état sain à un état défectueux. Mais si on doit choisir entre les deux indicateurs, l'indicateur X_{PCWT} montre une meilleure séparation entre l'état sain et l'état défectueux, et donc une meilleure capacité de détection. On peut aussi voir que l'indicateur $SPRI$ a aussi évolué avec l'évolution de défaut, alors que le défaut étudié est celui d'une bague extérieure, l'interprétation de ce phénomène peut être dû au chevauchement des harmoniques correspondantes au défaut bague extérieure de celles correspondantes au défaut de bague intérieure comme par exemple la 3ème harmonique de défaut bague extérieure ($\pm 178.5 \text{ Hz}$) et la deuxième harmonique de la bague intérieure ($\pm 181 \text{ Hz}$).

5. VALIDATION EXPÉRIMENTALE ET MISE EN ŒUVRE SUR UN BANC DE FATIGUE

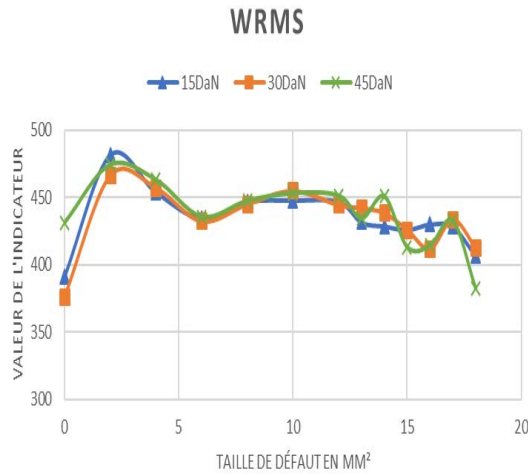


FIGURE 5.8 – L'évolution de l'indicateur X_{Wrms} pour des roulements sous différentes charges

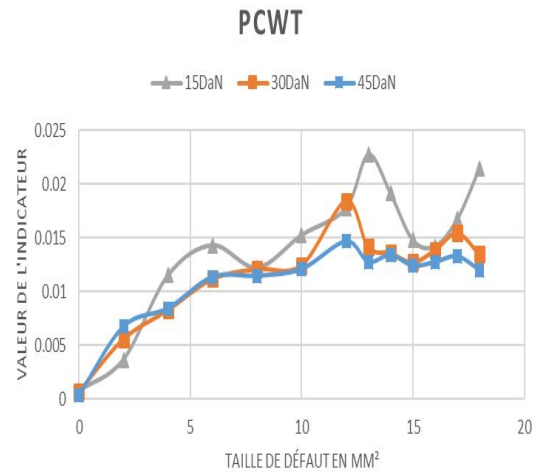


FIGURE 5.9 – L'évolution de l'indicateur X_{PCWT} pour des roulements sous différentes charges

Pour voir si le changement de charge affectent l'évolution de ces indicateurs, on a calculé ces deux indicateurs mais pour des roulements sous différentes charges. les résultats sont affichés dans les figures 5.8, 5.9. On peut déduire de l'analyse de ces deux figures (5.8,5.9) que même en changeant la charge appliquée sur le roulement, l'évolution de l'indicateur X_{Wrms} ne change pas pour un roulement défectueux. Cependant en changeant la charge appliquée sur un roulement sain la valeur de l'indicateur change avec la charge. Bien que l'évolution de l'indicateur X_{PCWT} évolue avec le changement de la charge et avec l'apparition de défaut mais l'évolution due au défaut est différente de celle constatée en cas de variation de charge.

C. Conclusion sur les indicateurs de défaut

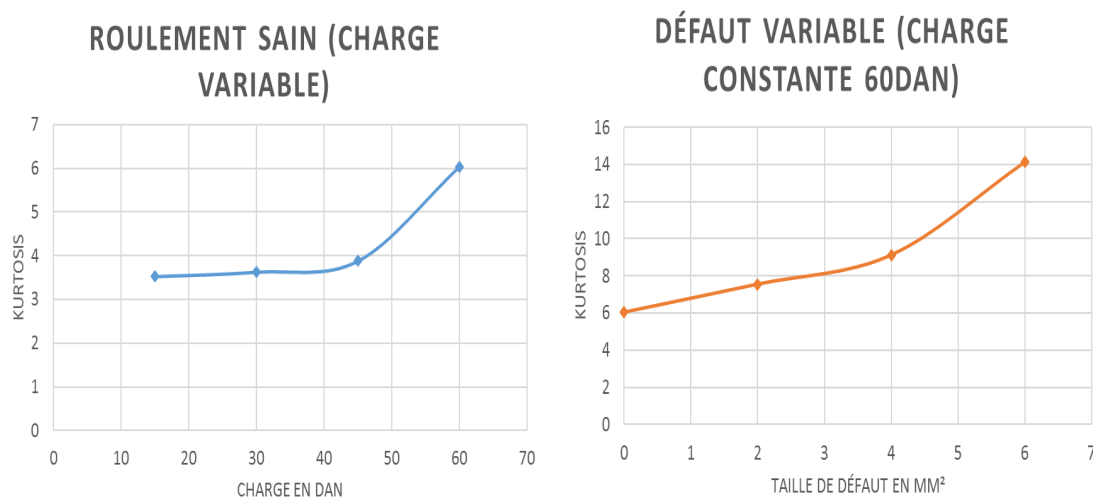
Cette étude préliminaire établie avait pour objectif de tester la sensibilité des indicateurs classiques, des indicateurs récemment introduits dans la littérature et d'autres proposés dans cette thèse (W_{RMS} et P_{cwt}).

Cette étude nous a permis de tirer quelques conclusion vis à vis des performances des

5.1 Mise en œuvre de la méthode de classification dynamique sur un banc expérimental

indicateurs en cas de défaut et en cas de variation de charge et sélectionner ceux qui sont le plus adapté aux contraintes d'un système de reconnaissance des formes. Les conclusions qu'on a pu tirer sont comme suit :

- En étudiant les indicateurs temporels, fréquentiels et temps-échelle, on a pu tirer que bien que ces indicateurs réagissent à l'apparition de défaut (qu'il soit accompagné d'une variation de charge ou avec une charge constante) et à la variation de charge en absence de défaut. Leurs évolutions en cas de défaut et en cas de variation de charge peuvent être semblables (voir figures 5.10a 5.10b et figures 5.11a, 5.11b).



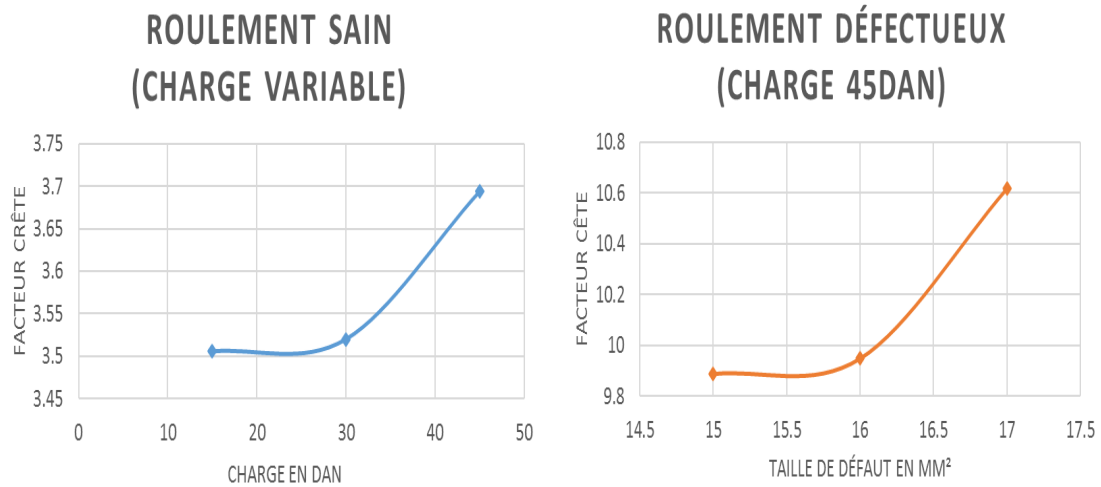
(a) L'évolution du Kurtosis pour un roulement sain mais sous charge variable (b) L'évolution du Kurtosis pour un roulement défectueux mais sous charge constante

FIGURE 5.10 – Ressemblance des courbes de tendances de kurtosis

- Les indicateurs issus de spectre de puissance sont moins performant que ceux issus d'enveloppe.
- Les indicateurs issus de la décomposition modale empirique offrent une meilleure visibilité de défaut même en cas de variation de charge.

En conclusion, dans cette étude on n'a trouvé qu' aucun indicateur parmi cette liste est sensible au défaut et insensible à la variation de charge (ou vice versa), et que certains indicateurs ont un meilleur rendement que d'autres. Cependant, ce qu'on on a pu retirer

5. VALIDATION EXPÉRIMENTALE ET MISE EN ŒUVRE SUR UN BANC DE FATIGUE



(a) L'évolution du facetur crête pour un roule- (b) L'évolution du facetur crête pour un roule-
ment sain mais sous charge variable ment défectueux mais sous charge constante

FIGURE 5.11 – Ressemblance des courbes de tendances de FC

de cette étude est que même s'ils sont sensibles au deux, la réaction des indicateurs à la présence de défaut et à la variation de charge n'est pas pareille et peuvent être séparables. En effet l'efficacité de ces indicateurs dépend de plusieurs facteurs notamment le type de défaut, la cinématique de la machines ou/et autres. L'objectif de cette partie d'étude a été d'établir une liste d'indicateurs qui peuvent nous permettre de diagnostiquer la machine (peu importe ses conditions de fonctionnement ou sa cinématique) en toute fiabilité. Cet objectif ne peut être accompli qu'on combinant ces indicateurs avec des techniques complémentaires. Ces méthodes complémentaires auront pour objectif de surmonter ce problème, en combinant ces indicateurs et puis sélectionner après automatiquement ceux qu'ils sont plus informatifs en point de vue séparation (Classification), ou créer des nouvelles caractéristiques qui peuvent ne pas être physiquement interprétables mais qui offrent une meilleure séparation entre les changements vibratoires causés par le défaut de ceux causés par la variation des conditions de fonctionnement (la charge).

5.1.3 Sélection et extraction des caractéristiques non physiques

Après une première extraction des indicateurs de défaut, on peut procéder soit à une sélection des indicateurs les plus informatifs à partir de lot des indicateurs extraits du signal vibratoire ou à créer de nouvelles caractéristiques à partir des combinaisons linéaires ou non des indicateurs extraits auparavant.

A. Sélection des caractéristiques informatives

La méthode qu'on a choisi d'appliquer est la sélection par *SFS* Sequential Forward Selection ou la sélection ascendante séquentielle. Cette méthode a été appliquée sur une matrice constitué des indicateurs extraits par des méthodes citée dans la section précédente. On a d'abord appliqué une version supervisée de SFS, on a choisi quelques exemples pour constituer une base d'apprentissage.

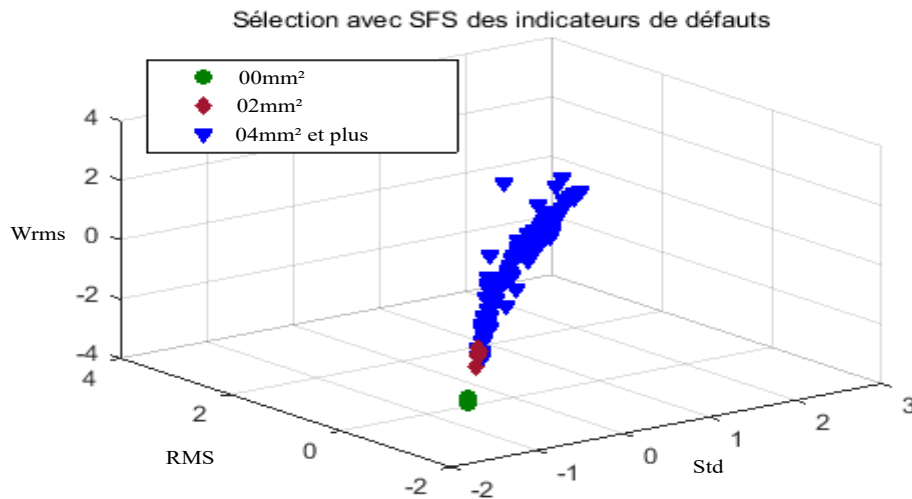


FIGURE 5.12 – La sélection par SFS des indicateurs de défauts extraits des signaux vibratoires issus de roulements avec taille de défaut variable (cas charge constante 15DaN)

La figure 5.12 montre qu'en choisissant les bons indicateurs l'état "sain" est bien séparée de l'état "défectueux" et que l'évolution de l'état "défectueux" suit un trajet plus ou moins linéaire. Les indicateurs retenus par cette méthode dans le cas charge constante sont : *RMS*, *standard deviation* et *Wrms* et dans le cas de charge variable sont : *F_C*, *xfc* et *Pcwt* (voir figure 5.13). Les figures (5.12,5.13) montrent clairement que l'application de SFS a permis de choisir les indicateurs qui permettent une meilleure séparation entre

5. VALIDATION EXPÉRIMENTALE ET MISE EN ŒUVRE SUR UN BANC DE FATIGUE

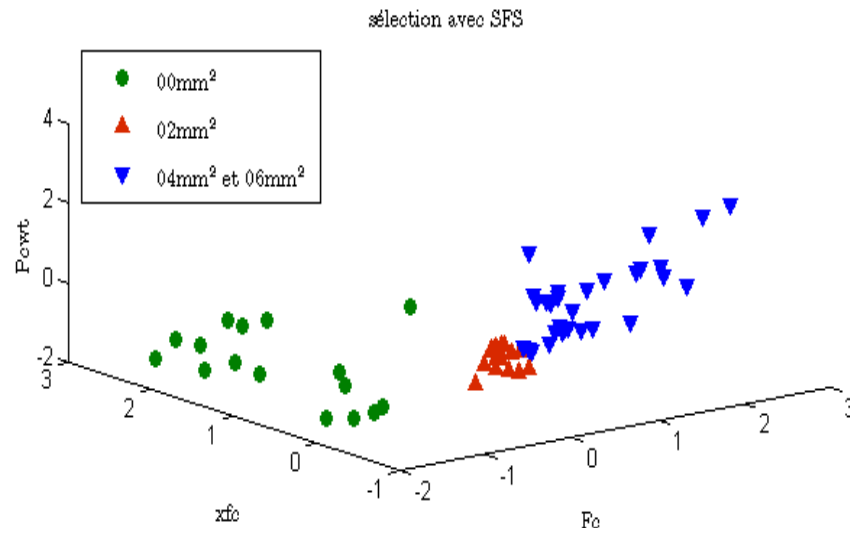


FIGURE 5.13 – La sélection par SFS des indicateurs de défauts extraits des signaux vibratoires issus de roulements avec taille de défaut variable (cas charge variable)

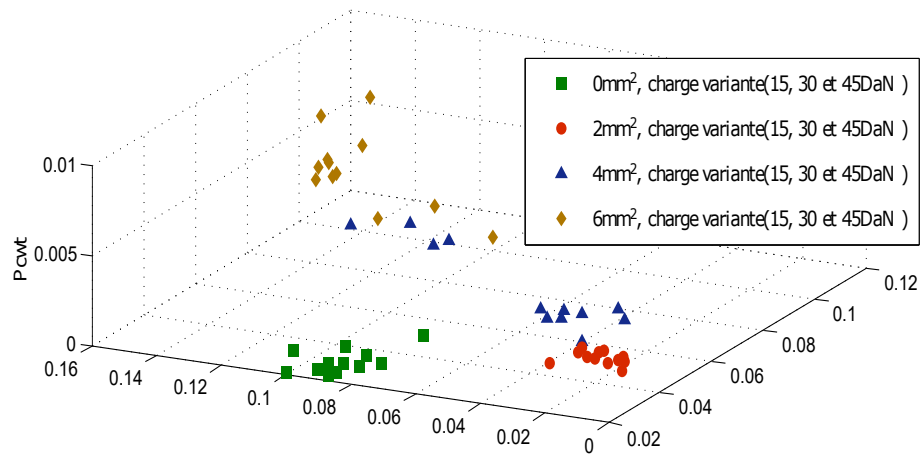


FIGURE 5.14 – Sélection des indicateurs par SFS non supervisée

cas "sain" et "défectueux" son avoir besoin de chercher les indicateurs manuellement ou créer de nouvelles caractéristiques.

Néanmoins, dans un cas réel, on n'a pas forcément accès à des exemples étiquetés qui peuvent nous aider à apprendre à notre méthode de sélection de trouver le critère de sélection qui permet d'améliorer la classification. Pour ceci on a choisi de travailler avec SFS combinée avec une méthode de classification non supervisée (K-means) et un critère de sélection (Silhouette).

5.1 Mise en œuvre de la méthode de classification dynamique sur un banc expérimental

Les résultats de la sélection ont été moins bon que dans le cas supervisé, la distance séparant les observations pour un roulement sain sous différentes charges et d'un roulement en début de dégradation est plus petite qu'avec la sélection supervisée. On remarque que pour la même taille de défaut en changeant la charge on peut avoir un nuage de points séparé. Comme on peut voir sur la figure 5.14, le cas de roulement avec défaut de $4mm^2$, on peut remarque que les points présentant ce roulement sont séparée en deux nuages (premier nuage proche de $2mm^2$ (charge 15DaN et 30DaN) et l'autre est plutôt plus proche de nuage correspondant à $06mm^2$ ($04mm^2$ charge 45DaN). Cependant on retrouve tout de même séparation entre ces nuages de points. Les indicateurs qui étaient sélectionnés sont *Fc*, *Pcwt* et *Stdf*.

La sélection par SFS permet de choisir le sous ensemble d'indicateurs qui offrent la meilleure séparation en comparant le score de chaque indicateur par rapport à un critère. La définition de ce critère exige une compréhension complète et profonde de données et de problème, ce qui rend l'emploi de ce type de données difficile.

Une autre alternative sera donc d'employer des techniques de création ou d'extraction de nouvelles caractéristiques non physiques. Ces techniques qui peuvent révéler des informations cachées et qui pourraient nous être utiles.

B. Extraction des caractéristiques non physiques par les méthodes de réduction de dimension

Pour pouvoir analyser l'apport de des méthodes de réduction de dimensions dans le suivi, on a tout d'abord appliqué ces méthodes sur une matrice constituée des indicateurs de défaut extraits à partir des signaux vibratoires enregistrés pour des roulements avec des défauts de différentes tailles. Ces roulements sont premièrement sous les mêmes conditions de fonctionnement, et ensuite sous des différentes charges.

La constitution de la matrice des indicateurs est comme suit :

$$M_{ind} = \left[\begin{array}{c} \left[\begin{array}{c} Temp \end{array} \right] \quad \left[\begin{array}{c} Freq \end{array} \right] \quad \left[\begin{array}{c} WT \end{array} \right] \end{array} \right] \quad (5.1)$$

$$Temp = \begin{pmatrix} RMS & Kurtosis & Fc & Skewness & FI & St \\ x_t(1,1) & & \cdots & & \cdots & x_t(1,6) \\ \vdots & & & \ddots & & \vdots \\ x_t(4,1) & & \cdots & & \cdots & x_t(4,6) \end{pmatrix} \quad (5.2)$$

$$Freq = \left[\begin{array}{c} \left[\begin{array}{c} puissance \end{array} \right] \left[\begin{array}{c} enveloppe \end{array} \right] \end{array} \right] \quad (5.3)$$

$$Freq_p = \begin{pmatrix} frms_b & FC_b & rmsf_b & xstdf_b \\ x_p(1,1) & & \cdots & x_p(1,4) \\ \vdots & & \ddots & \vdots \\ x_p(4,1) & & \cdots & x_p(4,4) \end{pmatrix} \quad (5.4)$$

$$WT_e = \begin{pmatrix} WRMS & Pcwt & SPRI & SPRO \\ x_e(1,1) & & \cdots & x_e(1,4) \\ \vdots & & \ddots & \vdots \\ x_e(4,1) & & \cdots & x_e(4,4) \end{pmatrix} \quad (5.5)$$

B.1. Application de l'analyse par composantes principales à noyau *KPCA*

L'application de *KPCA* avec un noyau gaussien sur des différents roulements sous la même charge (15DaN) a amélioré la séparation entre les différents états de santé (voir figure 5.15), sur la figure on peut voir que les points correspondants à un roulement sain sont séparables de ceux d'un roulement défectueux.

5.1 Mise en œuvre de la méthode de classification dynamique sur un banc expérimental

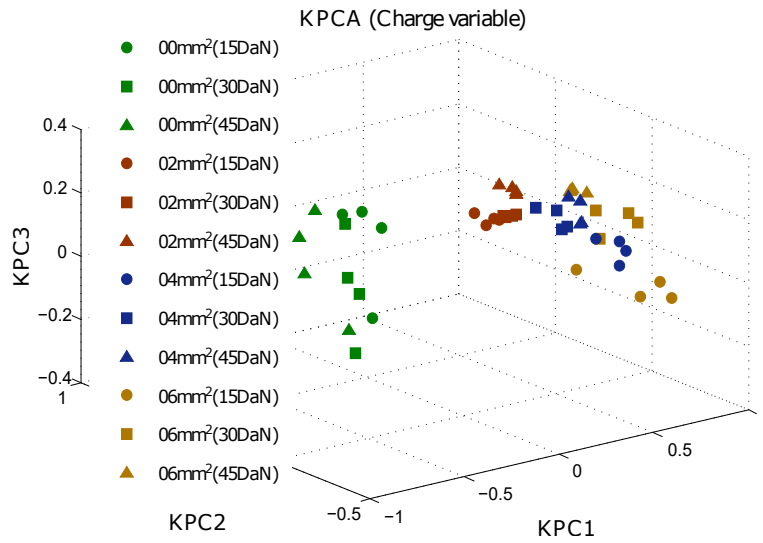


FIGURE 5.16 – L'application de la *KPCA* sur des roulements sous différentes charges (taille de défauts de 00mm^2 à 06mm^2 et charge de 15DaN à 45DaN)

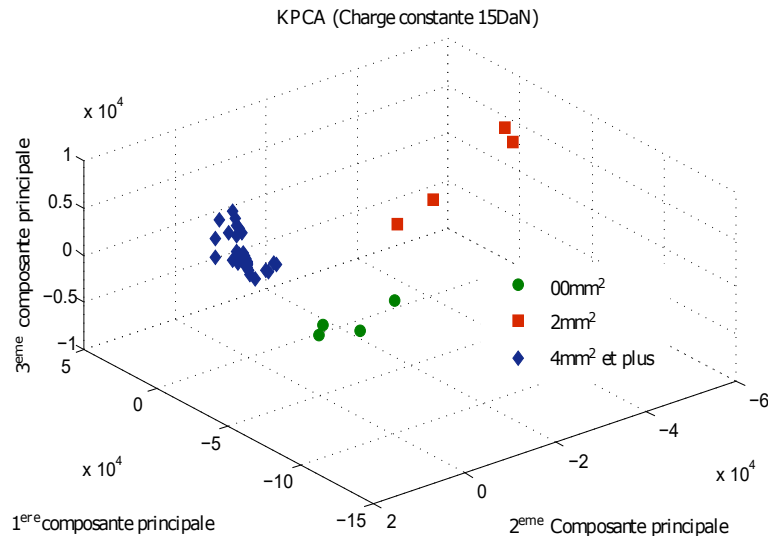


FIGURE 5.15 – L'application de la *KPCA* sur des roulements sous la même charge (15DaN)

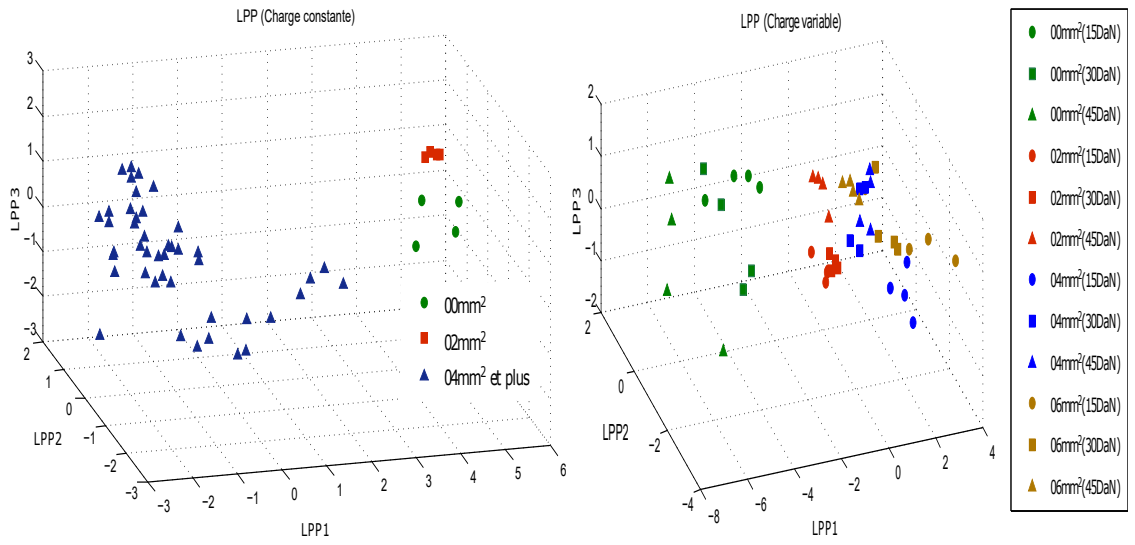
Quant à la séparation de état défectueux et sain quand les roulements sont sous différentes charges, on peut voir sur la figure 5.16 qu'avec *KPCA* gaussien on arrive bien à séparer sain et défectueux. Les points "sain" (vert) sont tous placés d'une coté et points "défectueux" dans l'autre.

5. VALIDATION EXPÉRIMENTALE ET MISE EN ŒUVRE SUR UN BANC DE FATIGUE

B.2. Application de projection par préservation de la topologie de données (*NPE* et *LPP*)

Local preserving projection

On commence par appliquer *LPP* sur les signaux correspondants à des différents roulements sous la même charge (voir figure 5.17a), ensuite on fait varier la charge (voir figure 5.17b). Sur les deux figure, on peut voir qu'avec *LPP* on peut toujours différencier entre état sain et état défectueux même en variation de charge.



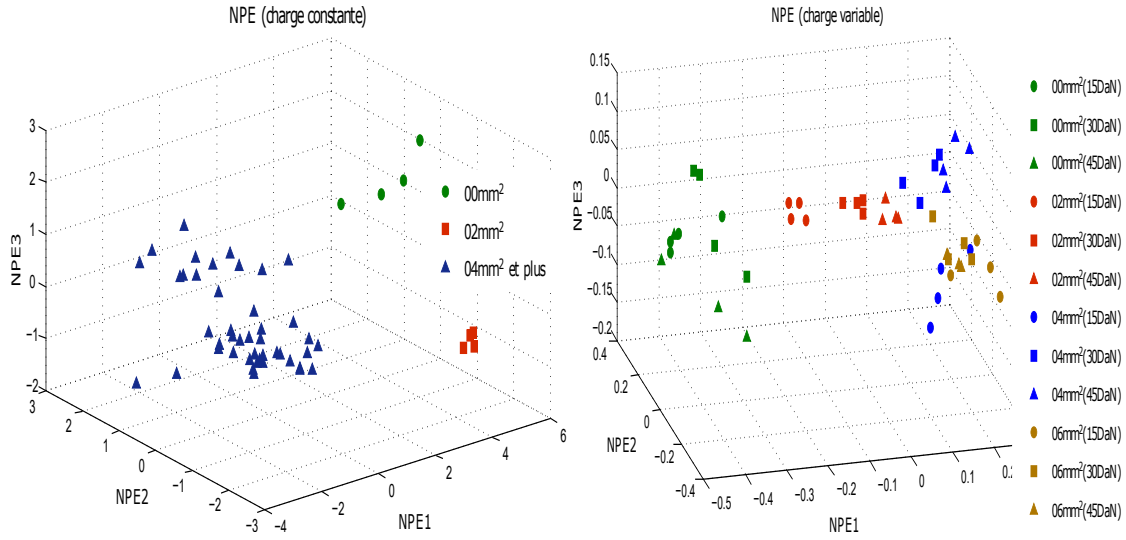
(a) L'application de *LPP* sur des données extraits de roulements sains et défectueux sous la même charge (b) L'application de *LPP* sur des données extraits de roulements sains et défectueux sous différentes charges

FIGURE 5.17 – Extraction des caractéristiques par *LPP*

Neighbor preserving embedding

L'application de *NPE* directement sur les indicateurs extraits à partir des signaux vibratoires de roulements en différents états a donné les résultats affichés dans la figure (5.18a), on peut voir les points correspondants à l'état sain sont mieux séparés de ceux correspondants à l'état défectueux qu'avec cas de *LPP*. Même en variant la charge le *NPE* peut aussi séparer sain de défectueux mieux que *LPP* (voir 5.18b).

5.1 Mise en œuvre de la méthode de classification dynamique sur un banc expérimental



(a) L'application de NPE sur des données ex- (b) L'application de NPE sur des données ex-
traits sur des roulements sain et défectueux sous traits sur des roulements sain et défectueux sous
la même charge différentes charges

FIGURE 5.18 — Extraction des caractéristiques par NPE

B.3. Conclusion sur l'extraction des caractéristiques non physiques Dans cette section on a pu voir que les méthodes d'extraction améliorent la séparation entre état sain et défectueux même en cas de variation de condition de fonctionnement (notamment la charge). Le degré de séparation entre les deux cas dépend bien évidemment du choix de la méthode et de la bonne configuration de ses paramètres.

Bien que les techniques d'extraction de type supervisées peuvent offrir une meilleure séparabilité entre état « sain » et « défectueux » de roulement (dans les deux cas, charge constante ou charge variable), mais les techniques non supervisées donnent des résultats convenables et spécialement les méthodes de préservation de topologie de données (LPP et NPE).

Comme on a pu constater dans cette section, les méthodes de sélection et d'extraction ont permis d'améliorer l'écart entre état sain et état défectueux des roulements étudiants même sous charge variable. Cette séparation va venir en aide à la méthode de classification qui pourra identifier les données efficacement.

Les étapes précédentes (extraction des indicateurs et sélection ou extraction des caractéristiques physiques) sont considérés comme des étapes préparatoires à l'étape de la

5. VALIDATION EXPÉRIMENTALE ET MISE EN ŒUVRE SUR UN BANC DE FATIGUE

classification. Dans cet étape les données seront attribuées à leurs classes appropriées et où chaque classe représente un état de fonctionnement du roulement suivi.

5.1.4 Classification dynamique

Afin de valider les méthodes de classification proposées dans le chapitre précédent, et de choisir le meilleur couple méthode d'extraction ou sélection et méthode de classification. Nous avons appliqué chaque méthode de classification (*DDBSCAN*, *ESDBSCAN*, *FDSDBSCAN*) sur les données extraites par les différentes techniques mentionnées dans la section précédente (*SFS*, *KPCA*, *LPP*, *NPE*). Le

A. Dynamic *DBSCAN*

L'application de la méthode *DDBSCAN* sur les signaux vibratoires, nous permet de valider nos choix d'indicateurs et de pouvoir associer l'état de roulement à une classe. Pour tester cette méthode nous avons utilisé des signaux vibratoires issus de roulements sous différentes charges et où chaque roulement a un niveau de dégradation différent. Le déroulement de test est comme suit :

- de $t=1$ jusqu'à $t=3$: on introduit à la méthode Rdf des signaux vibratoires issus d'un roulement sain avec une charge qui augmente ($t=1$ charge=15DaN, à $t=2$ la charge passe à 30DaN et à $t=3$ la charge atteint la valeur 45DaN),
- de $t=4$ à $t=6$: on introduit à la méthodes Rdf des signaux vibratoires issus d'un roulement avec un petit défaut de $2mm^2$ la charge change de la même manière que le roulement sain (à $t=4$ charge=15DaN, à $t=5$ la charge passe à 30DaN et à $t=6$ la charge atteint la valeur de 45 DaN),
- de $t=7$ à $t=9$: on introduit à la méthodes Rdf des signaux vibratoires issus d'un roulement avec un défaut de $4mm^2$ la charge change de la même manière qu'avant (à $t=7$ charge=15DaN, à $t=8$ la charge passe à 30DaN et à $t=9$ la charge atteint la valeur de 45 DaN),
- ⋮
- de $t=37$ à $t=39$: on on introduit à la méthodes Rdf des signaux vibratoires issus d'un roulement avec un défaut de $18mm^2$ la charge change à chaque $t + 1$ (à $t=37$ charge=15DaN, à $t=38$ la charge passe à 30DaN et à $t=39$ la charge atteint la valeur de 45 DaN),

5. VALIDATION EXPÉRIMENTALE ET MISE EN ŒUVRE SUR UN BANC DE FATIGUE

A.1. *KPCA – DDBSCAN*

Les résultats de la classification par *DDBSCAN* montrent que la variation de charge pour un roulement sain n'induit pas l'apparition d'une nouvelle classe. Par contre la présence de défaut sur la bague extérieure (défaut de $2mm^2$) fait apparaître une nouvelle classe qui représente l'état du roulement ('état défectueux').

Les figures 5.19a, 5.19b et 5.19c montrent l'évolution de la classe dite 'saine' en fonction de la variation de charge (15 daN, 30 daN et 45 daN) entre l'instant t_1 et t_3 . Il n'y a pas d'apparition de nouvelle classe, seul le centre de la classe "saine" se déplace.

5.1 Mise en œuvre de la méthode de classification dynamique sur un banc expérimental

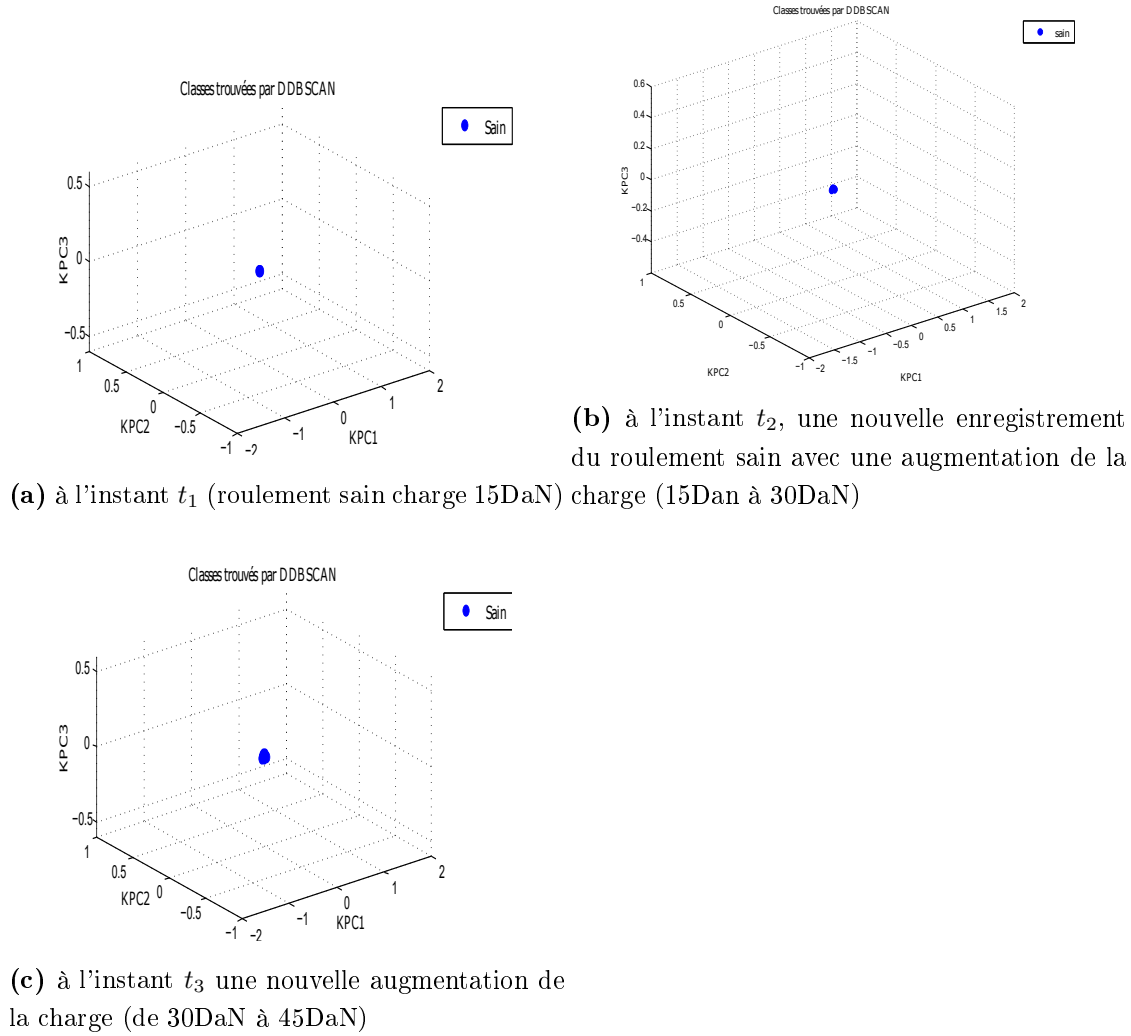


FIGURE 5.19 – Résultats de classification quand on augmente la charge d'un roulement sain

A t_4 , un saut d'observation vers un espace de décision non reconnu par la méthode de classification apparaît provoquant ainsi la création d'une nouvelle classe. Ce saut d'observation se coïncide avec l'apparition d'un défaut de $2mm^2$. La variation de la charge, de 15 daN à 45 daN, pour ce défaut de taille $2mm^2$ n'induit pas l'apparition d'une nouvelle classe, les observations restent dans la même classe.

5. VALIDATION EXPÉRIMENTALE ET MISE EN ŒUVRE SUR UN BANC DE FATIGUE

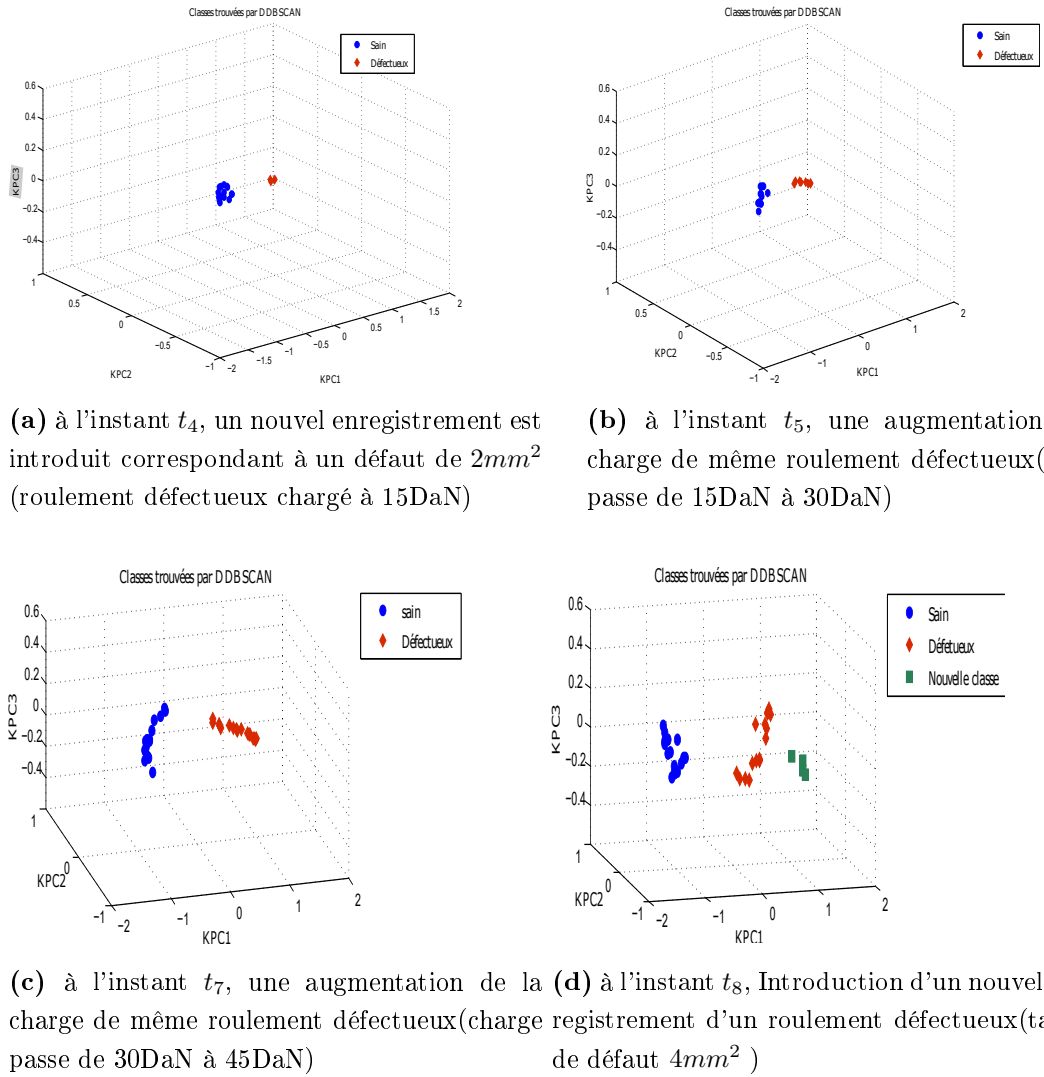
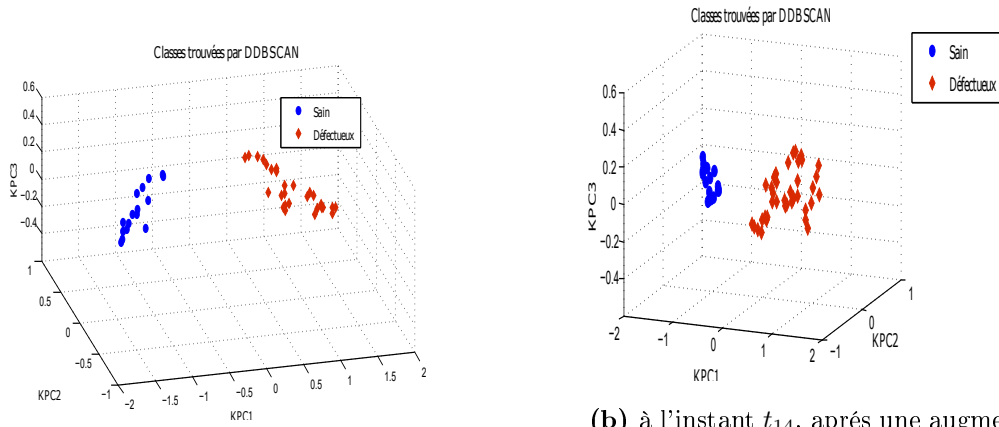


FIGURE 5.20 – Classification des observations reçues pour des roulements au début de dégradation et sous charge variable

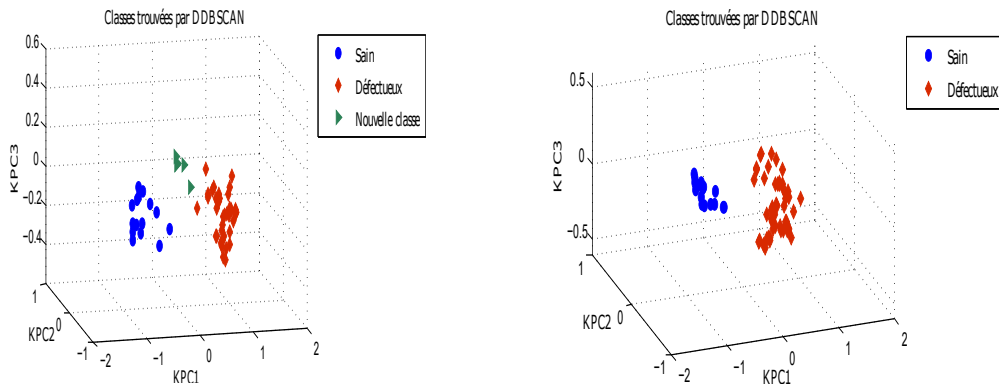
De t_5 à t_7 les observations reçues correspondent à l'augmentation de la charge. La méthode de classification a continué à les attribuer à la même classe ("défectueuse") provoquant un déplacement du centre de gravité de la classe. A t_8 , l'introduction d'un nouvel enregistrement correspondant à un roulement avec un défaut de $4mm^2$. Cette évolution de la taille de défaut s'est manifesté par un nouveau saut d'observation provoquant la création d'une nouvelle classe.

5.1 Mise en œuvre de la méthode de classification dynamique sur un banc expérimental



(a) à l'instant t_{11} , une augmentation de la charge de même roulement défectueux (charge passe de 15DaN à 30DaN puis à 45DaN)

(b) à l'instant t_{14} , après une augmentation de la taille de défaut à t_{12} (taille de défaut atteint les $6mm^2$), une agnmentation de la charge est appliquée (la charge passe de 15 DaN à 30DaN puis 45DaN)



(c) à l'instant t_{15} , la taille de défaut passe à $8mm^2$ (la charge 15DaN) résultant en un nouveau saut d'observation qui se traduit par la création d'une nouvelle classe

(d) à l'instant t_{17} , augmentation de la charge (la charge passe de 15DaN à 30 DaN puis à 45DaN), la classe qui a été créée se fusionne avec la classe défectueux

FIGURE 5.21 – Classification des observations reçues des roulements sous charge variable et avec un niveau de dégradation avancé

A chaque augmentation de la taille de défaut, un nouveau saut d'observation est observée. Ce saut entraîne la création d'une nouvelle classe qui sera par la suite fusionnée avec la classe 'défectueuse'. On peut dire que l'évolution de la taille de défaut se manifeste avec *DDBSKAN* par un saut d'observation alors que l'augmentation de la charge se traduit

5. VALIDATION EXPÉRIMENTALE ET MISE EN ŒUVRE SUR UN BANC DE FATIGUE

par une dérive lente.

Durant ce test, le *DDBSCAN* a réalisé une classification sans erreur en attribuant correctement les observations correspondant à un roulement sain à la classe 'Sain ' et les observations correspondant à un roulement défectueux à la classe 'Défectueuse'. Cependant, quand le défaut atteint un niveau critique les observations se rapprochent de la classe 'saine', et on remarque que la méthode de classification a attribué quelques observations d'un roulement avec défaut développé à la classe 'saine' en faisant grimper le taux de mauvaise classification à 0.25%.

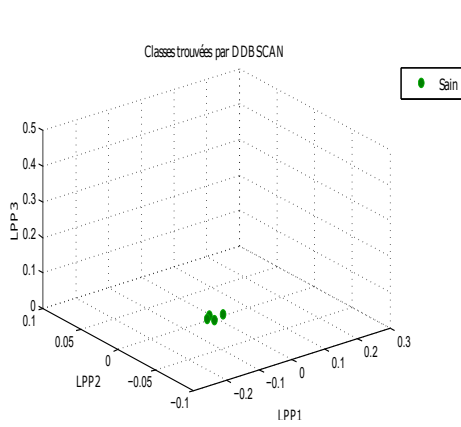
Comme la méthode classification *DDBSCAN* nécessite l'initialisation de plusieurs paramètres dont les principaux sont *Eps* et *Minpts*, les résultats de la classification en dépendent fortement. Les valeurs choisies durant ce test ont été réglés comme suit (*Minpts* = 4 qui correspond aux nombres d'observations extraites par enregistrement) et que nous jugeons suffisant pour valider la nécessité de la création d'une nouvelle classe. Quant à l'initialisation du paramètre *Eps* nous avons choisi de lui attribuer la valeur *Eps* = 0.005.

Comme le paramètre *Eps* permet de définir la limite de voisinage de points et qu'en choisissant une petite valeur on arrive à retrouver des petits regroupement de points, alors qu'en choisissant une valeur grande on retrouve que les grands regroupements. Généralement, Le choix de la valeur de *Eps* peut se faire à l'aide de graphe de K plus proche voisins (k-ppv) dans le cas de classification statique ce qui n'est plus possible pour le cas de traitement en ligne. Le réglage de *Eps* pour traitement en-ligne exige alors une certaine connaissance de données.

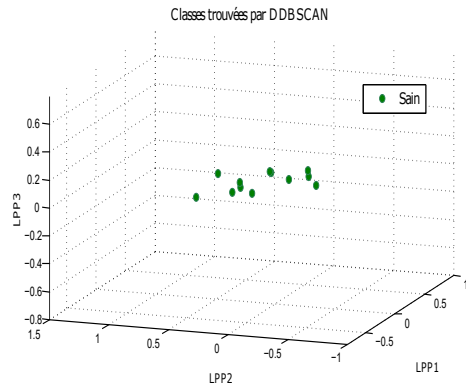
A.2. *LPP – DDBSCAN*

On retrouve les mêmes phénomènes (création et fusion) qu'on a observé pendant la classification des observations extraites par *KPCA*. On retrouve toujours une dérive lente en augmentant la charge, le saut d'observation en augmentant la taille de défaut, et la fusion de chaque nouvelle classe créée avec la classe 'Défectueuse' au cours de la réception des observations suivantes.

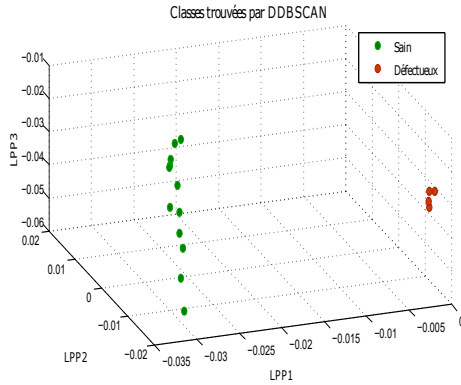
5.1 Mise en œuvre de la méthode de classification dynamique sur un banc expérimental



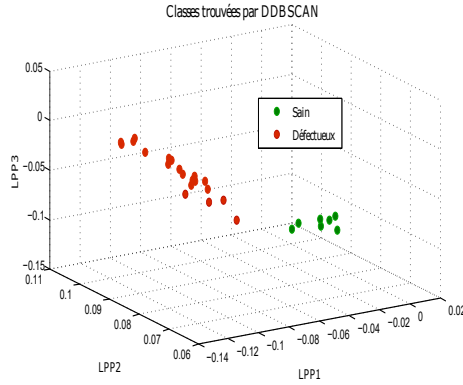
(a) à l'instant t_1 , roulement sain, charge 15DaN)



(b) à l'instant t_3 , après une augmentation successive de la charge appliquée sur le roulement sain (la charge passe de 15 DaN à 30DaN puis 45DaN)



(c) à l'instant t_4 , un nouvel enregistrement est introduit correspondant à un défaut de $2mm^2$ (roulement défectueux charge 15DaN)



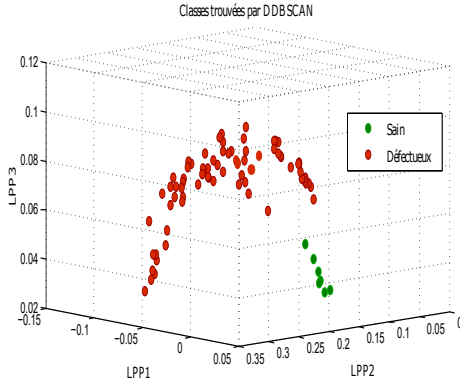
(d) à l'instant t_9 , après le passage de la taille de défaut de $2mm^2$ à $4mm^2$ suivi par une augmentation successive de la charge appliquée sur le roulement 15DaN à 30DaN puis à 45DaN

FIGURE 5.22 – Classification des observations reçues d'un roulement sain sous charge variable et un roulement au début de dégradation sous charge variable

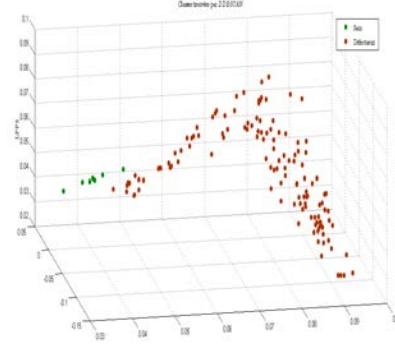
On remarque une nette amélioration de taux d'erreur qui baisse à moins de 0.2% accompagné d'une augmentation de la distance séparant le centre de la classe 'Saine' de celui de classe correspondant au début de défaut par 23%. On remarque aussi une baisse du nombre outliers non utilisés. Ces outliers qui ont été détectés par la méthode de classification *DDBSCAN*, ils étaient ré-introduit comme des nouvelles observations durant la

5. VALIDATION EXPÉRIMENTALE ET MISE EN ŒUVRE SUR UN BANC DE FATIGUE

classification.



(a) à l'instant t_{21} , la taille de défaut atteint la taille de $12mm^2$ (la charge appliquée a passé de 15DaN à 30DaN puis à 45DaN)



(b) à l'instant t_{39} , la taille de défaut a atteint la taille de $18mm^2$ suivi par une augmentation successive de la charge appliquée sur le roulement 15DaN à 30DaN puis à 45DaN

FIGURE 5.23 – Classification des observations reçues des roulements sous charge variable et avec un niveau de dégradation avancé

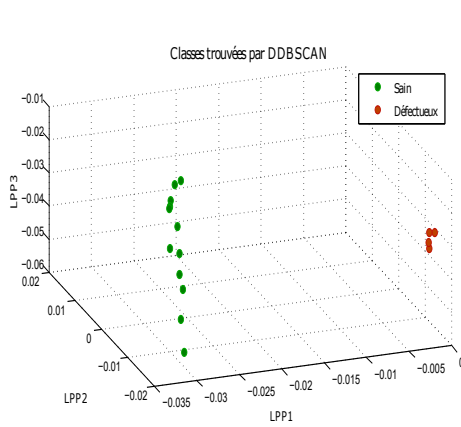
Les paramètres de *DDBSCAN* ont été choisis comme suit : $MinPts = 4$ et $Eps = 0.02$.

A.3. NPE – DDBSCAN

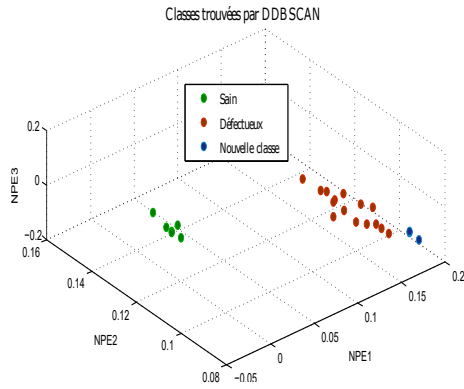
On retrouve les mêmes phénomènes observés dans le cas d'évolution de la taille de défaut et dans le cas de la variation de charge, on remarque toujours le saut d'observation en présence de défaut et dans le cas d'évolution de la taille, une dérive lente en variation de charge. Le taux de mauvaise classification augmente un peu pour atteindre la valeur 0.35%. La distance entre la classe 'saine' et le début de défaut est supérieure à celle calculée pour le *LPP* par 10%.

Dans les figures 5.22a à 5.24d, on peut voir les résultats de classification dans des instants choisis.

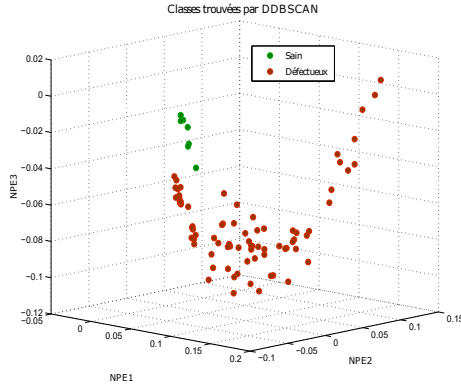
5.1 Mise en œuvre de la méthode de classification dynamique sur un banc expérimental



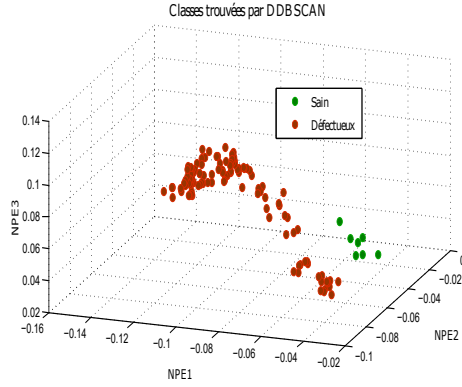
(a) à l'instant t_4 , début de défaut $2mm^2$, charge 15DaN



(b) à l'instant t_{10} , un nouvel enregistrement est introduit correspondant au passage de défaut à la taille $6mm^2$ (roulement défectueux charge 15DaN)



(c) à l'instant t_9 , après le passage de la taille de défaut à $12mm^2$ charge appliquée sur le roulement est de 15DaN



(d) à l'instant t_{39} , la taille de défaut a atteint la taille de $18mm^2$ suivi par une augmentation successive de la charge appliquée sur le roulement 15DaN à 30DaN puis à 45DaN

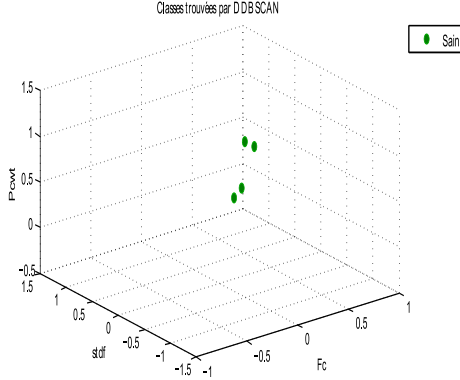
FIGURE 5.24 – Classification des observations reçues pour des roulements sous charge variable et avec un niveau de dégradation avancé

A.4. SFS – DDBSCAN

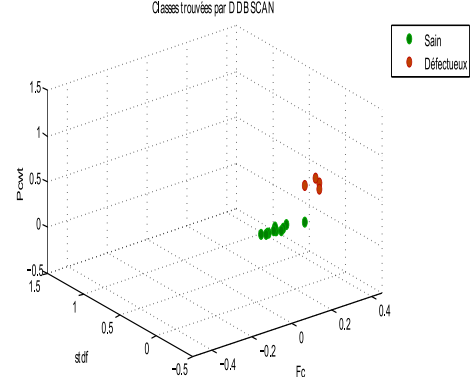
Cette fois on a appliqué la méthode de classification *DDBSCAN* sur les indicateurs de défauts choisis par *SFS*. Dans les figures (fig5.25a à fig5.25d), on retrouve l'apparition des même phénomènes observés pour les cas précédents. Le taux d'erreur reste moins de

5. VALIDATION EXPÉRIMENTALE ET MISE EN ŒUVRE SUR UN BANC DE FATIGUE

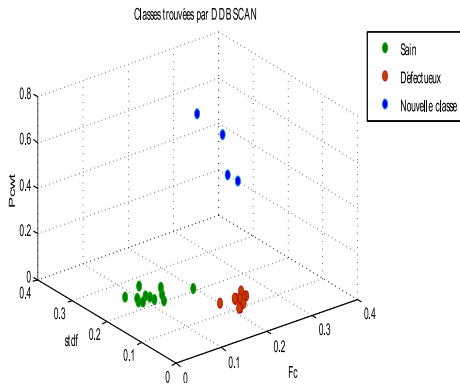
0.3% et on perd en distance entre les deux centres de gravité (centre de classe sain, et début de défaut) par 15% par celle de *NPE*.



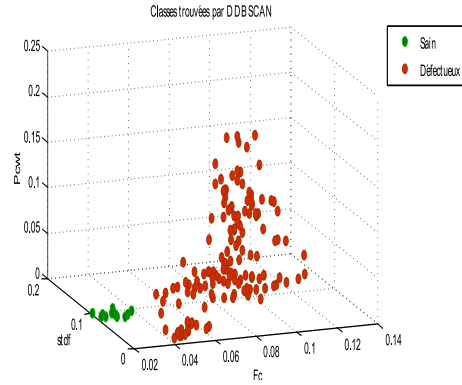
(a) à l'instant t_1 , roulement sain, charge 15DaN)



(b) à l'instant t_4 , début de défaut $2mm^2$, charge 15DaN



(c) à l'instant t_7 , après le passage de la taille de défaut à $04mm^2$ charge appliquée sur le roulement est de 15DaN



(d) à l'instant t_{39} , la taille de défaut a atteint la taille de $18mm^2$ suivi par une augmentation successive de la charge appliquée sur le roulement 15DaN à 30DaN puis à 45DaN

FIGURE 5.25 – Classifications des observations reçues des roulements en différents états et sous charge variable

A.5. Conclusion sur la classification par *DDBSCAN*

Bien que, l'application de *DDBSCAN* sur les caractéristiques créées par les trois méthodes de réduction de dimension ou celles sélectionnés par *SFS* nous a permis de bien classifié les observations reçues dans leurs classes appropriées. Le taux de mauvaise clas-

sification est resté en dessous des 1%. On peut voir que la meilleure combinaison extraction/classification en termes de séparation et taux d'erreur sont les deux combinaisons sont *LPP – DDBSCAN* et *KPCA – DDBSCAN*.

Cependant, la classification par *DDBSCAN* nécessite l'accès à la totalité des anciennes observations. Le stockage des observations, dont la taille peut être infinie, pour pouvoir identifier l'état d'endommagement de composant suivi est une contrainte qu'on peut pas tolérer dans le cas d'un suivi. D'où la nécessité d'employer une méthode de classification adapté au traitement en ligne, qui pourrait nous garantir une classification avec un taux d'erreur acceptable et qui permet de minimiser la taille de mémoire occupée par les observations tout en nous offrant tous outils dont on avait l'accès avec *DDBSCAN*.

B. Evolving scalable *DBSCAN*

La première alternative que nous avons développé pour contourner les problèmes rencontrés avec *DDBSCAN* a été *ESDBSCAN*. Nous avons introduit des nouvelles fonctionnalités dans cette méthode pour l'adapter à la nature dynamique des données.

B.1. *KPCA – ESDBSCAN*

La nature de classification et de traitement des observations par *ESDBSCAN* est différente de celle par *DDBSCAN*. Les phénomènes observés par *DDBSCAN* sont beaucoup moins visibles, on retrouve toujours le saut d'observation quand la taille de défaut augmente.

Comme on peut voir dans les figures 5.26a, 5.26b, la variation de charge n'a pas provoqué la création d'une nouvelle classe, elle n'a pas non plus provoqué la création d'un nouveau point représentatif. Cependant, l'apparition d'un défaut (5.26c)s'est manifesté par un saut d'observation vers une zone loin des points représentatifs existants et assez loin de la classe 'saine' que la méthode de classification les a attribué un nouveau point représentatif et à une nouvelle classe.

5. VALIDATION EXPÉRIMENTALE ET MISE EN ŒUVRE SUR UN BANC DE FATIGUE

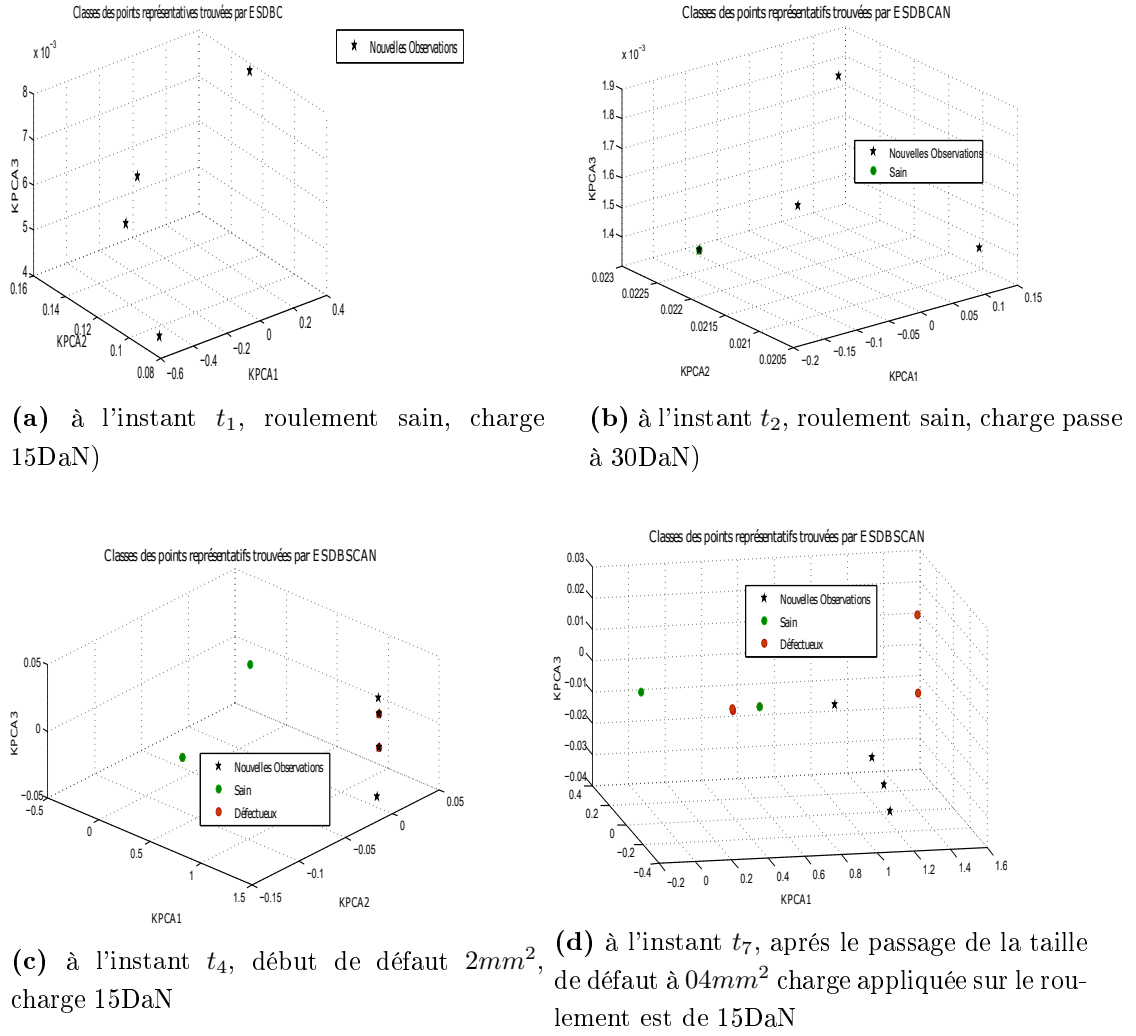


FIGURE 5.26 – Classification des premières observations reçues pour un roulement sain sous différentes charges et des roulements au début de dégradation sous charge variable

A fur et à mesure que le défaut se développe, les observations continuent à être attribuées à des nouveaux points représentatifs de la même classe dans le cas de variation de charge ou à des nouveaux points représentatifs d'une nouvelle classe quand la taille de défaut augmente comme on peut voir sur les figures 5.27a et 5.27b. Ce phénomène sera détaillé dans la section suivi.

5.1 Mise en œuvre de la méthode de classification dynamique sur un banc expérimental

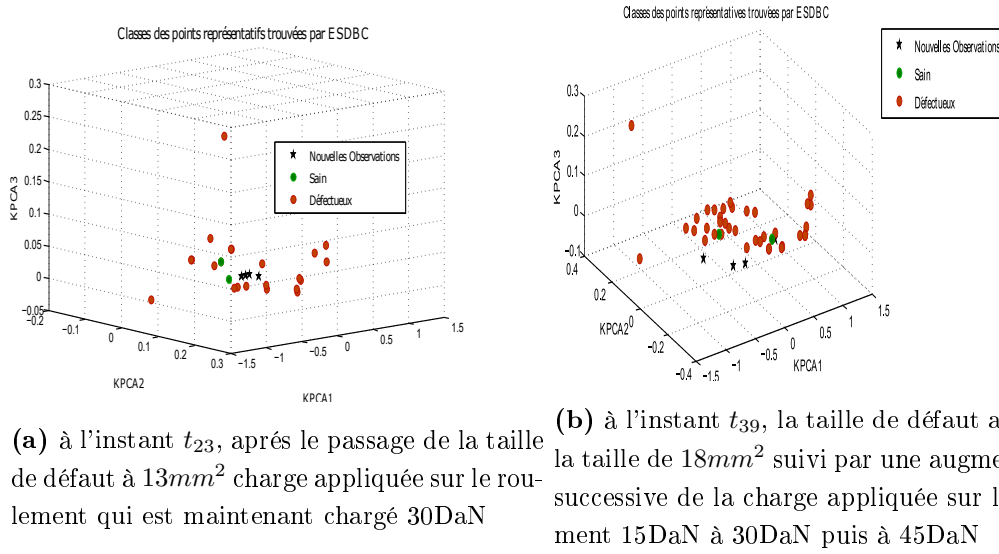


FIGURE 5.27 – Classification des observations reçues des roulements sous charge variable et avec un niveau de dégradation avancé

Ce qu'on peut aussi retenir de ces deux dernières figures ; plus le niveau de dégradation avance plus la distance séparant ces observations et le centre de la classe saine se rétrécit entraînant une augmentation du taux d'erreur.

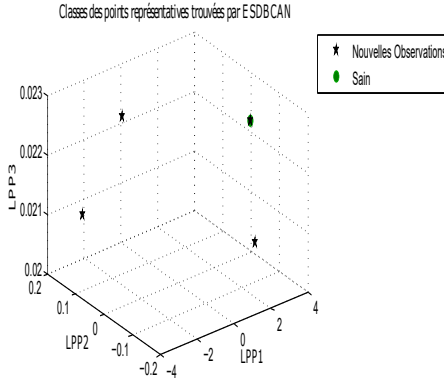
Bien que le taux de mauvaise classification augmente par rapport à celui calculé pour *DDBSCAN*, on arrive à avoir une classification correcte en gardant accès qu'à 25% des observations originales. Ce gain en mémoire peut être encore amélioré en introduisant la notion des *SuperRepPoints* ou des super points représentatifs qui représentent des points représentatifs et non les observations. Vu le nombre des observations ne nécessitant pas une réduction aussi majeure, nous n'avons pas implémenté cette fonctionnalité dans la procédure *ESDBSCAN* validée sur ces signaux.

B.2. LPP – ESDBSCAN

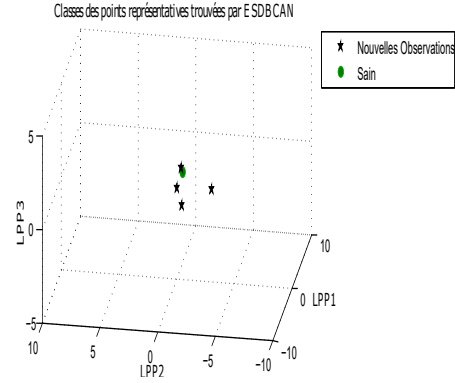
Nous avons remarqué une nette amélioration du taux de mauvaise classification qui baisse à 0.006% sachant que cette méthode n'avait accès qu'à 15% de la totalité des observations. La séparation entre classe 'saine' et classe 'défectueuse' est meilleure que celle observée avec celle de *ESDBSCAN* combinée avec *KPCA*. La distance séparant classe 'saine' et

5. VALIDATION EXPÉRIMENTALE ET MISE EN ŒUVRE SUR UN BANC DE FATIGUE

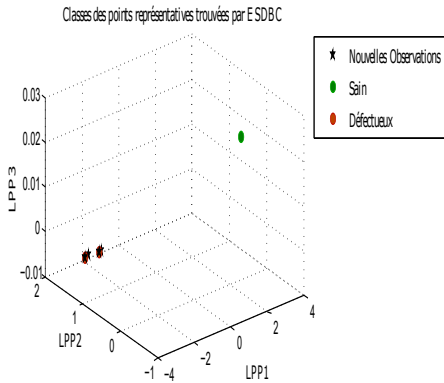
classe 'défectueuse' s'est amélioré par rapport à ce que nous avons retrouvé avec *KPCA*. Nous avons remarqué les mêmes phénomènes observés avec la méthode *DDBCAN*. En effet nous avons également retrouvé le saut d'observation provoqué par l'évolution de défaut. Ce saut entraîne la création des nouveaux points représentatifs dont la méthode de classification les attribue temporairement à une nouvelle classe. Cette classe sera ensuite fusionnée avec classe 'défectueuse'.



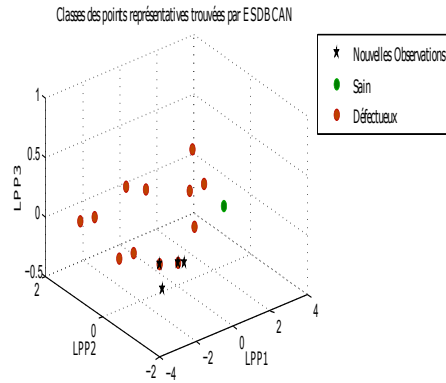
(a) à l'instant t_1 , roulement sain, charge 15DaN)



(b) à l'instant t_3 , roulement sain, charge passe à 30DaN puis à 45DaN)



(c) à l'instant t_4 , début de défaut $2mm^2$, charge 15DaN

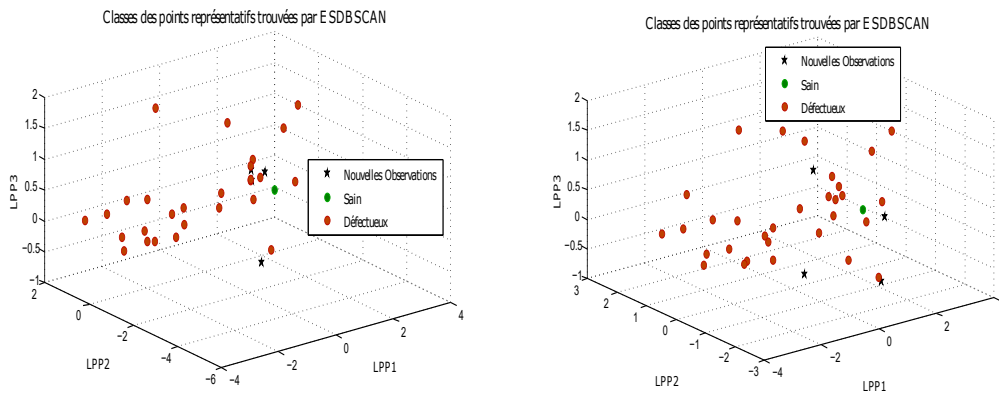


(d) à l'instant t_{10} , après le passage de la taille de défaut à $06mm^2$ charge appliquée sur le roulement est de 15DaN

FIGURE 5.28 – Classification des premières observations reçues pour un roulement sain sous différentes charges et des roulements au début de dégradation sous charge variable par *LPP – ESDBSCAN*

5.1 Mise en œuvre de la méthode de classification dynamique sur un banc expérimental

Cependant, la variation de charge, avec cette méthode de classification, ne conduit pas à la même conséquence dans la présence de défaut qu'en absence de défaut : En présence de défaut, elle provoque la création des nouveaux points représentatifs, ces points représentatifs se trouvent toujours assez proche d'une classe existante. La méthode de classification les attribue par la suite à la classe la plus proche. En absence de défaut, la variation de charge ne provoque pas la création des nouveaux points représentatifs ; les observations corresependantes à cette variation seront alors attribuées aux points représentatifs existants. Ce point qui sera assez dense qu'il représentera à lui seule la première classe trouvée 'Saine'.



(a) à l'instant t_{28} , après le passage de la taille de défaut à $15mm^2$ charge appliquée sur le roulement qui est maintenant chargé 15DaN (b) à l'instant t_{39} , la taille de défaut a atteint la taille de $18mm^2$ suivi par une augmentation successive de la charge appliquée sur le roulement 15DaN à 30DaN puis à 45DaN

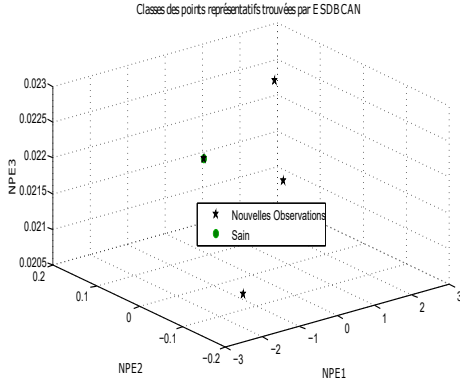
FIGURE 5.29 – Classification des observations reçues des roulements sous charge variable et avec un niveau de dégradation avancé par $LPP - ESDBSCAN$

B.3. $NPE - ESDBSCAN$

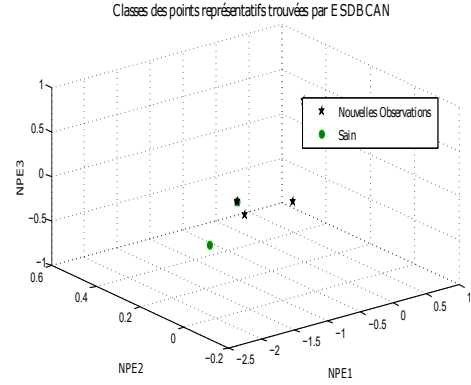
En remplaçant les caractéristiques créées par LPP par celles créées par NPE . Le taux d'erreur a augmenté par rapport à celui calculé avec LPP pour atteindre la valeur de 0.3%. La séparation entre les deux centres de gravités de deux classes s'est détériorée par 30%. Le saut d'observation associé à l'évolution de défaut est toujours détecté avec $ESDBSCAN$. Contrairement à $LPP - ESDBSCAN$, la variation de charge ne se manifeste pas par la même manière ; parfois elle provoque la création des nouveaux points

5. VALIDATION EXPÉRIMENTALE ET MISE EN ŒUVRE SUR UN BANC DE FATIGUE

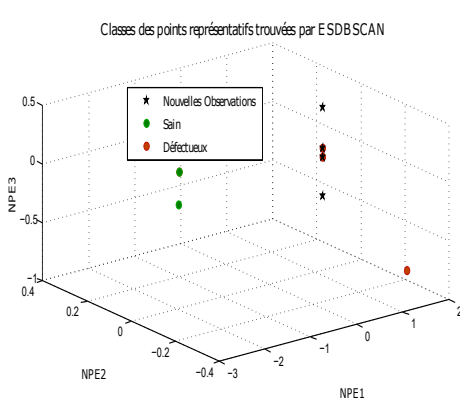
représentatifs et autres fois par la mise à jour des points représentatifs existants.



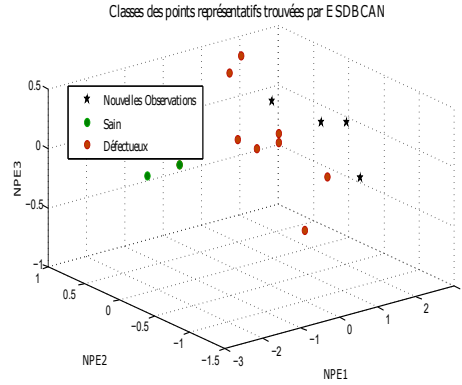
(a) à l'instant t_1 , roulement sain, charge 15DaN



(b) à l'instant t_3 , roulement sain, charge passe à 30DaN puis à 45DaN



(c) à l'instant t_4 , début de défaut $2mm^2$, charge 15DaN

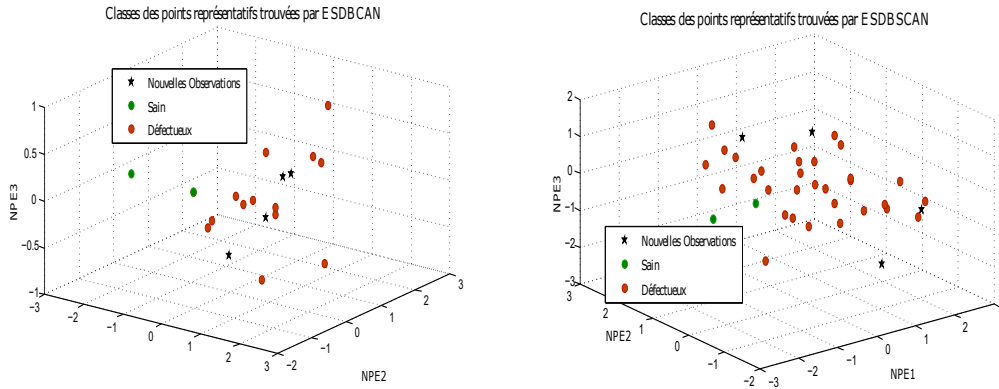


(d) à l'instant t_{10} , après le passage de la taille de défaut à $06mm^2$ charge appliquée sur le roulement est de 15DaN

FIGURE 5.30 – Classification des premières observations reçues pour un roulement sain sous différentes charges et des roulements au début de dégradation sous charge variable par $NPE - ESDBSCAN$

La combinaison de NPE avec $ESDBSCAN$ nous a permis de gagner plus de 79% de mémoire occupée par les observations utilisées dans $DDBSCAN$ affichant une amélioration en termes de taille de mémoire occupée par rapport à LPP et ceci sans implémentation des *SuperRepPoints*.

5.1 Mise en œuvre de la méthode de classification dynamique sur un banc expérimental



(a) à l'instant t_{21} , après le passage de la taille de défaut à $12mm^2$ charge appliquée sur le roulement est de 45DaN
(b) à l'instant t_{39} , la taille de défaut a atteint la taille de $18mm^2$ suivi par une augmentation successive de la charge appliquée sur le roulement 15DaN à 30DaN puis à 45DaN

FIGURE 5.31 – Classification des observations reçues des roulements sous charge variable et avec un niveau de dégradation avancé par $NPE - ESDBSCAN$

B.4. $SFS - ESDBSCAN$

Dans cette méthode, le taux de mauvaise classification a atteint 1%, ce qui représente une détérioration de la qualité de classification. Le gain de mémoire s'est amélioré car on a pu libérer plus de 84% de la mémoire utilisée par les observations employées par $DDBSCAN$.

5. VALIDATION EXPÉRIMENTALE ET MISE EN ŒUVRE SUR UN BANC DE FATIGUE

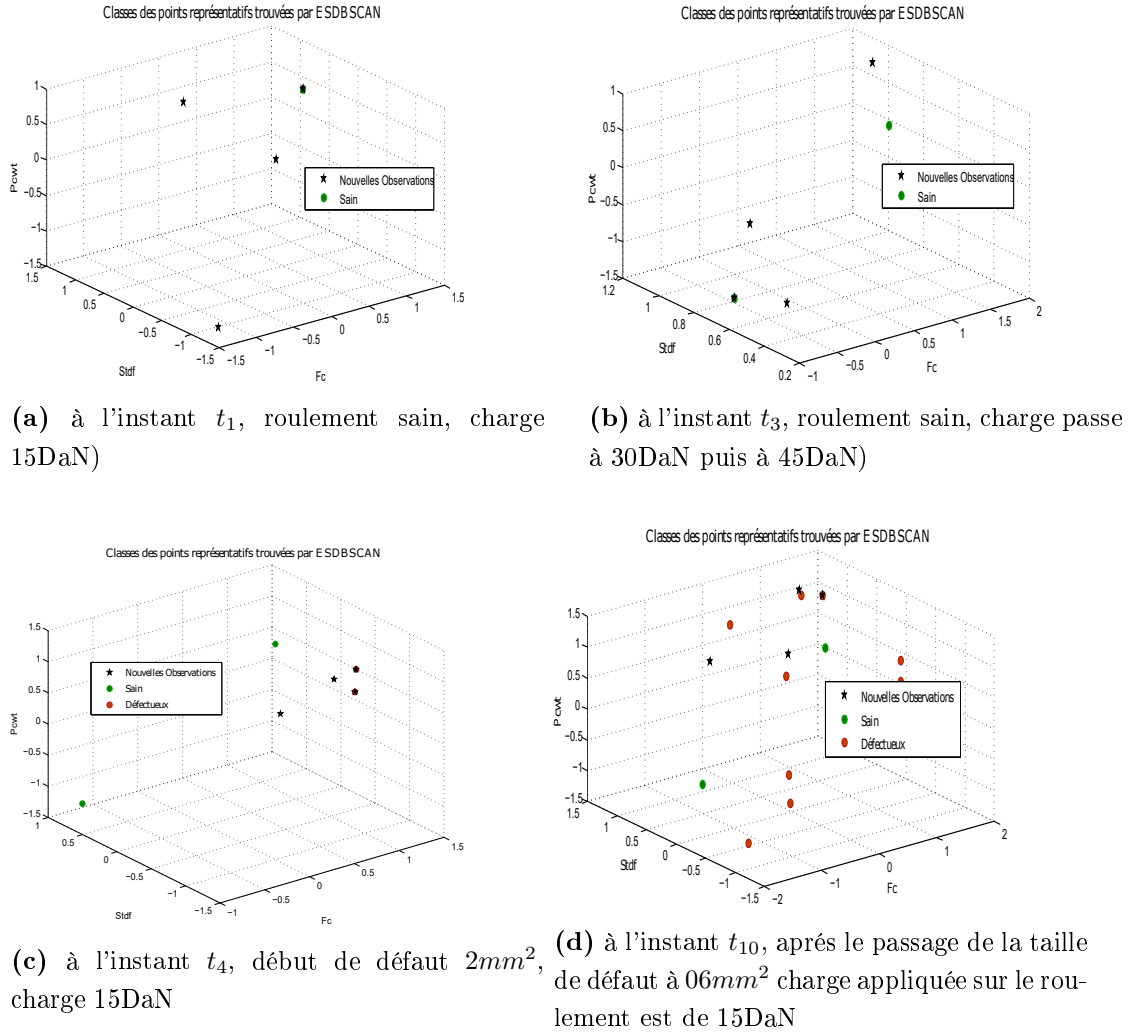


FIGURE 5.32 – Classification des premières observations reçues pour un roulement sain sous différentes charges et des roulements au début de dégradation sous charge variable par *SFS – ESDBSCAN*

La distance séparant les deux classes 'Saine' et 'Défectueuse' s'est réduite par rapport à celle trouvée avec *NPE*.

5.1 Mise en œuvre de la méthode de classification dynamique sur un banc expérimental

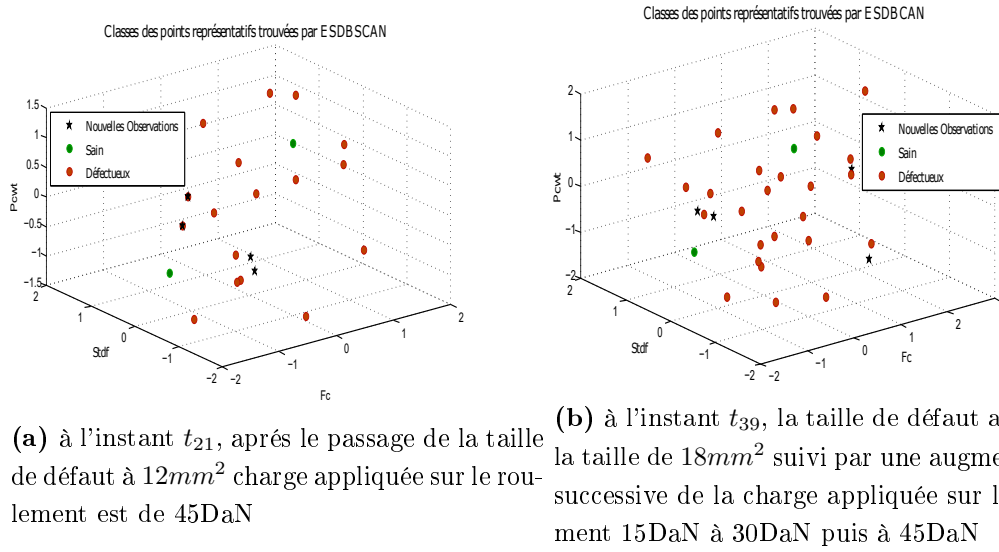


FIGURE 5.33 – Classification des observations reçues des roulements sous charge variable et avec un niveau de dégradation avancé par *SFS – ESDBSCAN*

B.5. Conclusion sur les résultats obtenus par la classification *ESDBSCAN*

L'application de *ESDBSCAN* sur les caractéristiques sélectionnées ou créées par des méthodes de réduction de dimensions a permis de libérer en moyenne 80% de la mémoire utilisées dans la méthode précédente tout en gardant un taux d'erreur de moins de 1%. La combinaison qui a permis d'obtenir un meilleur compromis taille de mémoire libérée/ taux d'erreur et distance entre classe 'Saine' et classe 'Défectueuse' est *ESDBSCAN – LPP*. Nous avons correctement classifié les observations en deux classes 'Saine' et 'Défectueuse' et gardé une bonne distance entre les deux classes, tout en respectant les contraintes du traitement en-ligne. Cette classification est de nature 'stricte' et non dynamique et elle devrait re-lancer la classification pour chaque détection d'un saut d'observation vers une zone non reconnue.

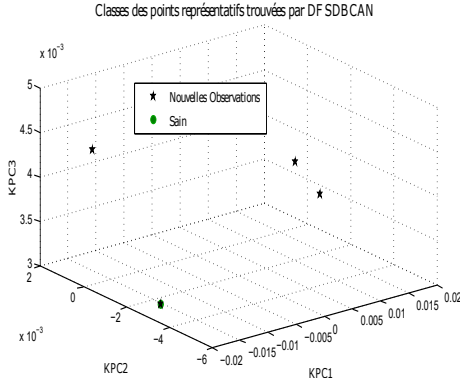
C. Dynamic Fuzzy Scalable *DBSCAN*

C.1. *KPCA – DFSDBSCAN*

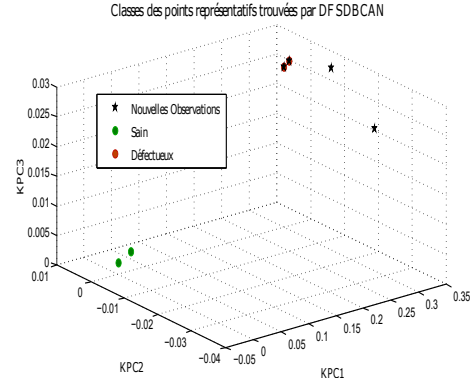
L'application de la méthode *DFSDBSCAN* sur les composantes principales obtenues par *KPCA* a permis de détecter la présence de défaut correctement. Bien qu'on ne peut

5. VALIDATION EXPÉRIMENTALE ET MISE EN ŒUVRE SUR UN BANC DE FATIGUE

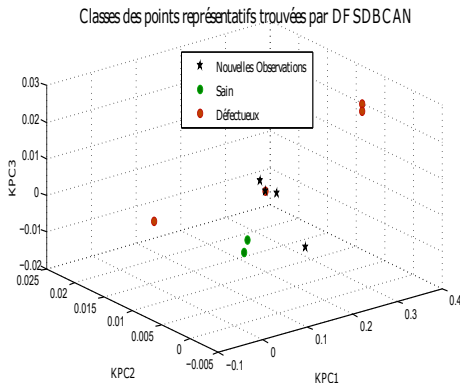
pas voir sur les figures 5.34a à 5.35b qui affichent les résultats finaux de la classification (après l'étape validation). La méthode *DFSDBSCAN* crée une nouvelle classe à chaque passage de défaut à un niveau de dégradation supérieur. Cependant, la variation de charge se manifeste par la création des nouveaux points représentatifs proche de la classe actuellement créée.



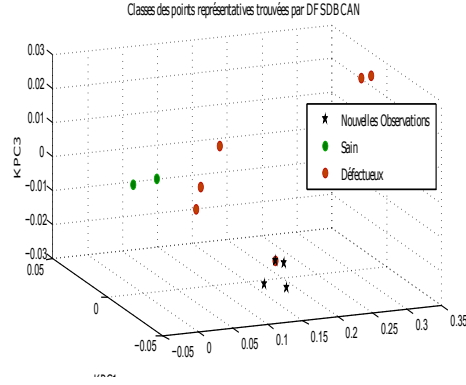
(a) à l'instant t_1 , roulement sain, charge 15DaN



(b) à l'instant t_4 , début de défaut $2mm^2$, charge 15DaN



(c) à l'instant t_{12} , taille de défaut atteint $4mm^2$, charge est à 15DaN



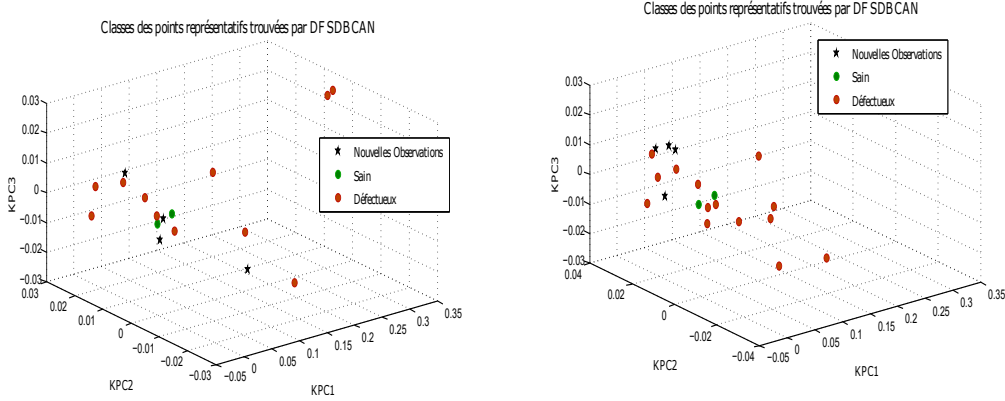
(d) à l'instant t_{10} , après le passage de la taille de défaut à $06mm^2$ charge appliquée sur le roulement est de 15DaN

FIGURE 5.34 – Classification des premières observations reçues pour un roulement sain sous différentes charges et des roulements au début de dégradation sous charge variable par *KPCA – DFSDBSCAN*

Le taux d'erreur ou le taux de la mauvaise classification augmente par rapport aux

5.1 Mise en œuvre de la méthode de classification dynamique sur un banc expérimental

autres méthodes de classification pour atteindre 0.5%. La séparation entre les deux classes 'Saine' et 'Défectueuse' est moyenne.



(a) à l'instant t_{21} , après le passage de la taille de défaut à $12mm^2$ charge appliquée sur le roulement est de 45DaN

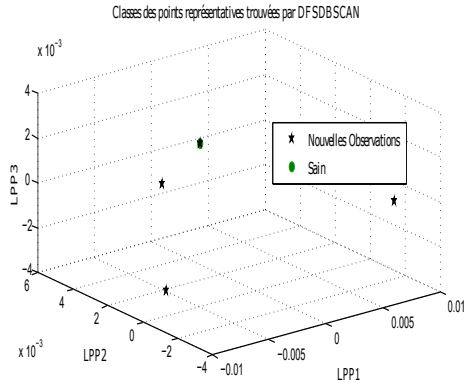
(b) à l'instant t_{39} , la taille de défaut a atteint la taille de $18mm^2$ suivi par une augmentation successive de la charge appliquée sur le roulement 15DaN à 30DaN puis à 45DaN

FIGURE 5.35 – Classification des observations reçues des roulements sous charge variable et avec un niveau de dégradation avancé par *KPCA – DFSDBSCAN*

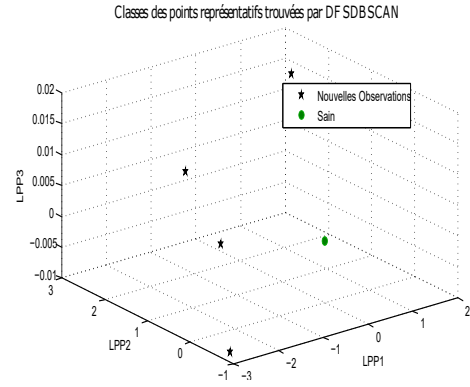
C.2. *LPP – DFSDBSCAN*

La combinaison *LPP – DFSDBSCAN* a permis d'améliorer le taux de classification qui s'est approché de la valeur 0%. Le taux d'observations gardé en mémoire est moins de 10% de la totalité des observations. Pendant la classification des observations par *LPP – DFSDBSCAN*, l'évolution de la taille de défaut provoque également un saut d'observation entraînant la création d'une nouvelle classe. La variation de charge se manifeste par l'augmentation de la densité des points représentatifs qui se trouvent au voisinage des observations correspondantes à ce changement.

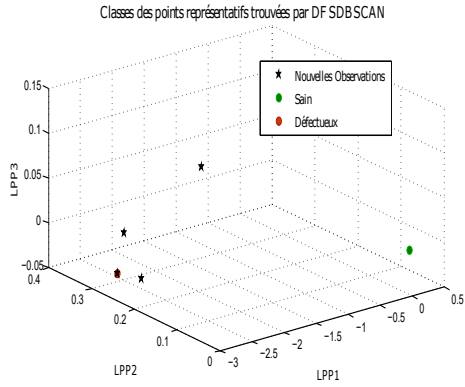
5. VALIDATION EXPÉRIMENTALE ET MISE EN ŒUVRE SUR UN BANC DE FATIGUE



(a) à l'instant t_1 , roulement sain, charge 15DaN



(b) à l'instant t_3 , roulement sain, la charge a passé de 15DaN à 30DaN puis à 45DaN



(c) à l'instant t_4 , début de défaut $2mm^2$, (d) à l'instant t_{12} , taille de défaut atteint $4mm^2$, charge est à 15DaN

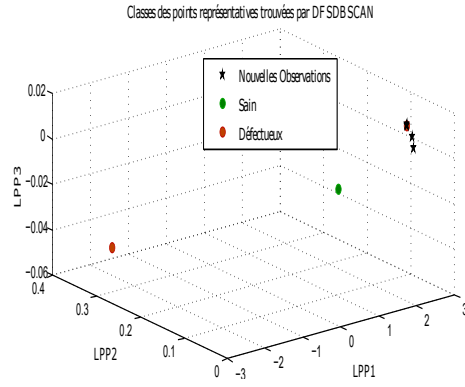
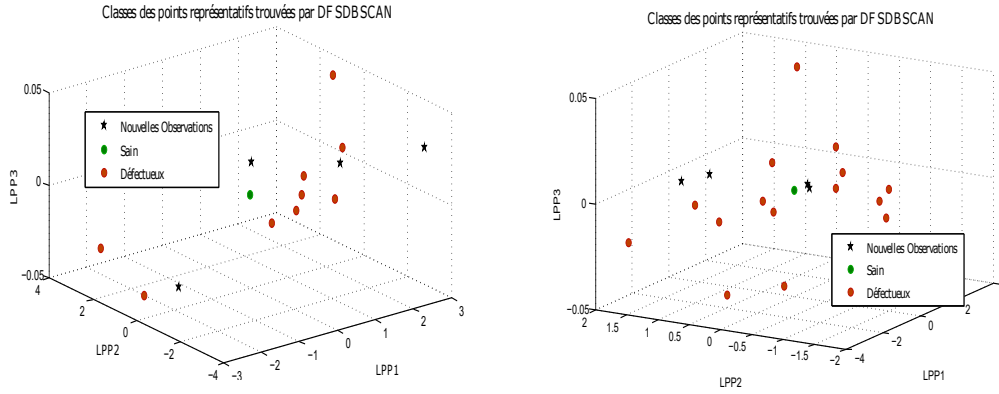


FIGURE 5.36 – Classification des premières observations reçues pour un roulement sain sous différentes charges et des roulements au début de dégradation sous charge variable par $LPP - DFSDSCAN$

Dans les figures 5.36a et 5.36b on peut voir que la variation de charge n'a pas provoqué la création d'une nouvelle classe ou d'un nouveau point représentatif, les observations correspondantes à la variation de charge du roulement sain ont toutes été attribuées au même point représentatif de la même classe 'Saine'.

Quand le défaut apparaît (5.36c), les nouvelles observations ont provoqué la création des nouveaux points représentatifs entraînant la création d'une nouvelle classe. De même, à chaque fois que la taille du défaut évolue (les figures 5.36d, 5.37a, 5.37b).

5.1 Mise en œuvre de la méthode de classification dynamique sur un banc expérimental



(a) à l'instant t_{21} , après le passage de la taille de défaut à 12mm^2 charge appliquée sur le roulement est de 45DaN
 (b) à l'instant t_{39} , la taille de défaut a atteint la taille de 18mm^2 suivi par une augmentation successive de la charge appliquée sur le roulement 15DaN à 30DaN puis à 45DaN

FIGURE 5.37 – Classification des observations reçues des roulements sous charge variable et avec un niveau de dégradation avancé par $LPP - DFSDBSCAN$

la distance séparant la classe 'Saine' et la classe 'Défectueuse' est meilleure que celle calculée pour $KPCA - DFSDBSCAN$. Cependant, cette distance se rétrécit au fur et à mesure que le niveau de dégradation évolue.

C.3. $NPE - DFSDBSCAN$

En appliquant la combinaison $NPE - DFSDBSCAN$ sur les observations extraites des signaux vibratoires, nous avons pu constater que les observations correspondantes à une variation de charge ont été affectées aux points représentatifs existants (voir figure 5.38a et 5.38b) alors que les observations correspondantes à l'apparition d'un défaut ou à son évolution ont entraîné un saut d'observation provoquant la création des nouveaux points représentatifs ainsi que la création de nouvelles classes temporaires.

Le taux de mauvaise classification fut aussi identique à celui de $LPP - DFSDBSCAN$ en utilisant que 0.08% du totale d'observations. Cependant, pour avoir ces résultats, cette méthode a considéré plusieurs observations utiles comme outliers.

5. VALIDATION EXPÉRIMENTALE ET MISE EN ŒUVRE SUR UN BANC DE FATIGUE

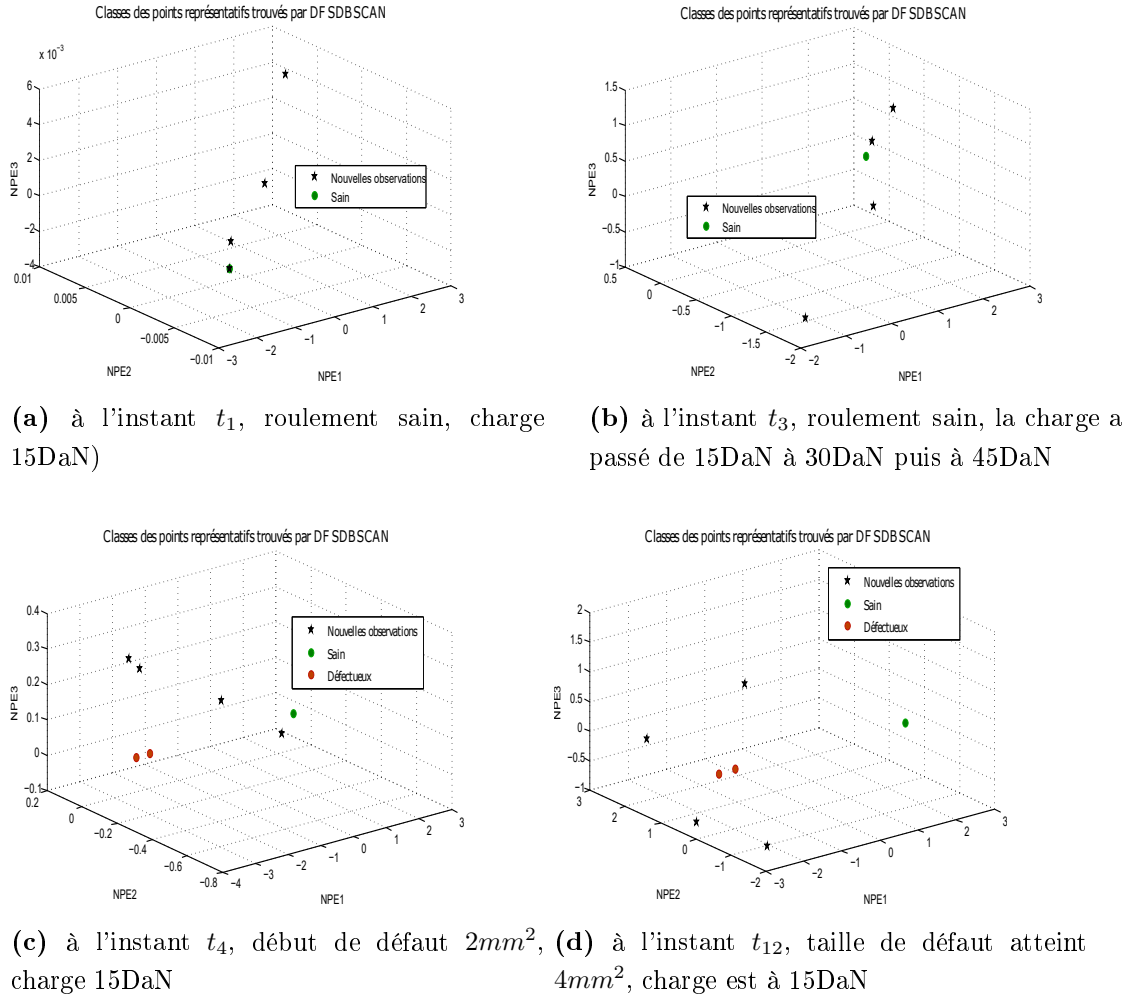


FIGURE 5.38 – Classification des premières observations reçues pour un roulement sain sous différentes charges et des roulements au début de dégradation sous charge variable par $NPE - DFSDBSCAN$

La séparation entre classe 'Saine' et 'Défectueuse' s'est améliorée par rapport à $LPP - DFSDBSCAN$ et on remarque le même phénomène observé que dans $LPP - DFSDBSCAN$; la distance séparant les deux classes se rétrécit au fur et à mesure que l'état de dégradation de roulement évolue.

5.1 Mise en œuvre de la méthode de classification dynamique sur un banc expérimental

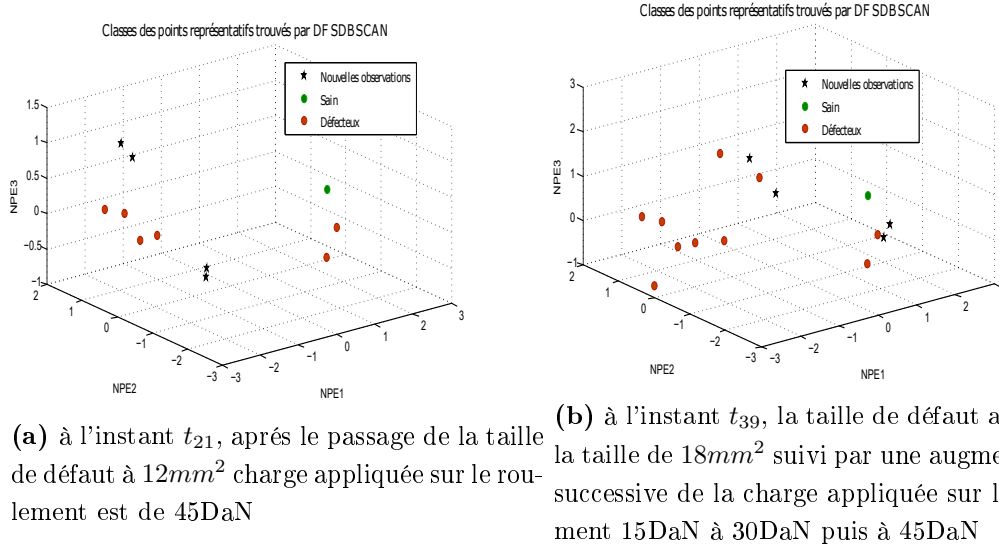


FIGURE 5.39 – Classification des observations reçues des roulements sous charge variable et avec un niveau de dégradation avancé par $NPE - DFSDBSCAN$

C.4. $SFS - DFSDBSCAN$

On observe les mêmes phénomènes que les autres méthodes associés à SFS . Les observations issues d'un roulement sain sont attribuées aux mêmes points représentatifs même si la charge évolue (voir figures 5.40a et 5.40b), alors que les observations issues d'une évolution de la taille défaut sont attribuées à des nouveaux points représentatifs (voir les figures 5.40c, 5.41a et 5.41b) qui sont affectés à leur tour à des nouvelles classes temporaires qui finissent à se fusionner avec la classe 'Défectueuse'. Le taux d'occupation des observations est de 16% de la totalité des observations.

5. VALIDATION EXPÉRIMENTALE ET MISE EN ŒUVRE SUR UN BANC DE FATIGUE

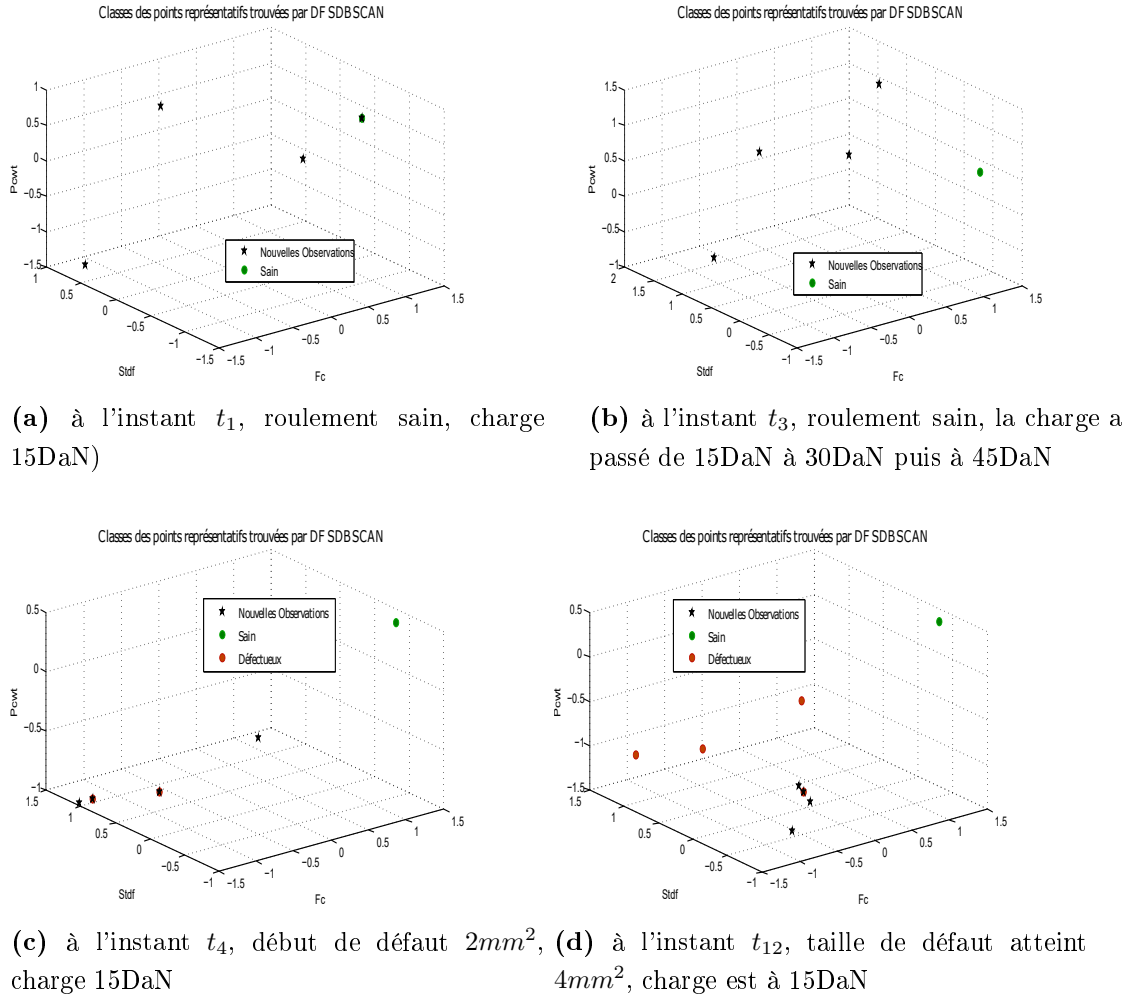
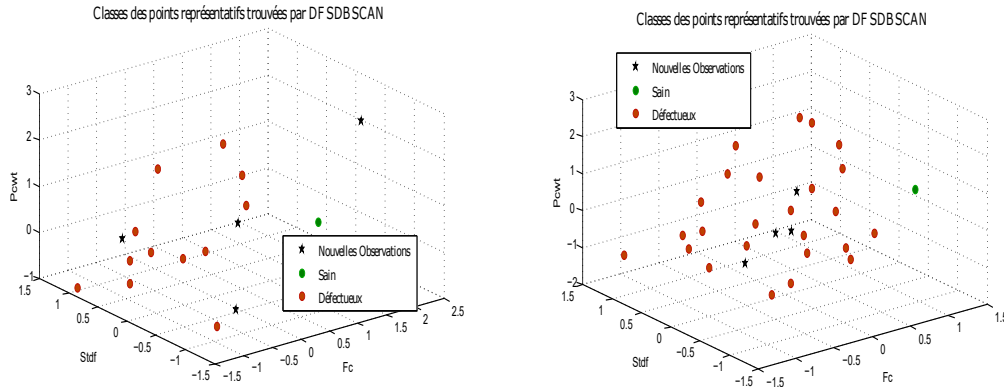


FIGURE 5.40 – Classification des premières observations reçues pour un roulement sain sous différentes charges et des roulements au début de dégradation sous charge variable par *SFS – DFSDBSCAN*

La distance séparant les deux classes 'Saine' et 'Défectueuse' a été moins importante que celle retrouvée par les autres méthodes (*LPP*, *NPE* et *KPCA*). Le taux d'erreur a été de 0.005%. Cependant plusieurs points utiles ont été jugé comme outliers et ils étaient écarté de la classification, ce qui peut être considéré comme une perte d'information.

5.1 Mise en œuvre de la méthode de classification dynamique sur un banc expérimental



(a) à l'instant t_{21} , après le passage de la taille de défaut à $12mm^2$ charge appliquée sur le roulement est de 45DaN
(b) à l'instant t_{39} , la taille de défaut a atteint la taille de $18mm^2$ suivi par une augmentation successive de la charge appliquée sur le roulement 15DaN à 30DaN puis à 45DaN

FIGURE 5.41 – Classification des observations reçues des roulements sous charge variable et avec un niveau de dégradation avancé par *SFS – DFSDBSCAN*

C.5. Conclusion sur les résultats obtenus par *DFSDBSCAN*

L'application de *DFSDBSCAN* combinée avec les techniques de réduction de dimension a permis d'améliorer la classification et la distance séparant classe 'Saine' du classe 'Défectueuse'. Nous avons pu révéler les mêmes phénomènes : Saut observation pour marquer l'évolution de la taille du défaut, le rapprochement de la classe 'Saine' et de la classe 'Défectueuse' quand le défaut atteint un niveau avancé, l'augmentation de la densité des points représentatifs et des classes quand la charge varie. En comparant les couples *KPCA – DFSDBSCAN*, *LPP – DFSDBSCAN*, *NPE – DFSDBSCAN* et *SFS – DFSDBSCAN*, les deux couples qui présentent le meilleur compromis sont *LPP – DFSDBSCAN*, *NPE – DFSDBSCAN*.

D. Conclusion

Les trois méthodes de classification ont permis d'affecter les observations issues d'un roulement défectueux à la classe 'Défectueuse' et les observations issues d'un roulement sain à la classe 'Saine'. Les trois méthodes de classification validées sur ce banc sont toutes dynamiques/évolutives. L'application de ces méthodes ont permis de voir la nature non stationnaires de données qui s'est manifestée par des sauts d'observations, des dérives

5. VALIDATION EXPÉRIMENTALE ET MISE EN ŒUVRE SUR UN BANC DE FATIGUE

lentes et par la création et la fusion des classes. Ces phénomènes observés lors de la classification dynamique représentent l'évolution des conditions de fonctionnement du roulement (variation de charge ou de vitesse et apparition de défaut mécanique). Les méthodes de classification statique ne permettent pas de détecter tous ces phénomènes observés par la méthode dynamique.

La méthode *DDBSCAN* a été la première méthode proposée durant cette thèse. Bien qu'elle ait permis de classer les données correctement, et de détecter l'évolution de ces données, la méthode *DDBSCAN* ne permet pas d'assurer un suivi continu de roulement car elle nécessite une capacité mémoire importante. En effet, cette méthode a besoin d'accéder à la totalité des données enregistrées depuis le début pour garantir un suivi fiable.

La méthode *ESDBSCAN* permet de contourner la nature en-ligne de traitement et de respecter la nature non stationnaire de données. Cette méthode a permis de classer correctement les observations reçues au fur et à mesure de leur réception et de détecter leurs évolutions.

Nous avons introduit la méthode *DFDBSCAN* pour améliorer la méthode *ESDBSCAN* qui est de nature stricte en terme de classification et qui est non supervisée. Cette méthode qui combine la nature évolutive de *ESDBSCAN* et l'apprentissage semi-supervisé nous a permis d'apporter plus d'information quant à l'évolution des données. Elle a permis également de diminuer le taux de mauvaise classification et le taux d'occupation des observations dans la mémoire.

En combinant ces méthodes de classification avec les méthodes de réduction de dimension (*KPCA*, *NPE*, *LPP* et *SFS*) et en analysant les résultats obtenus, nous pouvons conclure que les deux méthodes qui représentent le meilleur compromis taux d'erreur/occupation de mémoire ont été *LPP – DFSDSCAN* et *NPE – ESDBSCAN*.

5.1.5 Bilan sur la classification dynamique

A partir de toutes les combinaisons étudiées (méthode RD et classifieurs développés), les trois combinaisons *KPCA–DDBSCAN*, *LPP–ESDBSCAN* et *LPP–DFSDSCAN* ont présenté les meilleurs résultats. Les résultats de ces combinaisons sont affichés sur le tableau 5.1.

5.1 Mise en œuvre de la méthode de classification dynamique sur un banc expérimental

TABLE 5.1 – Tableau comparatif des résultats obtenus avec classifieurs proposés

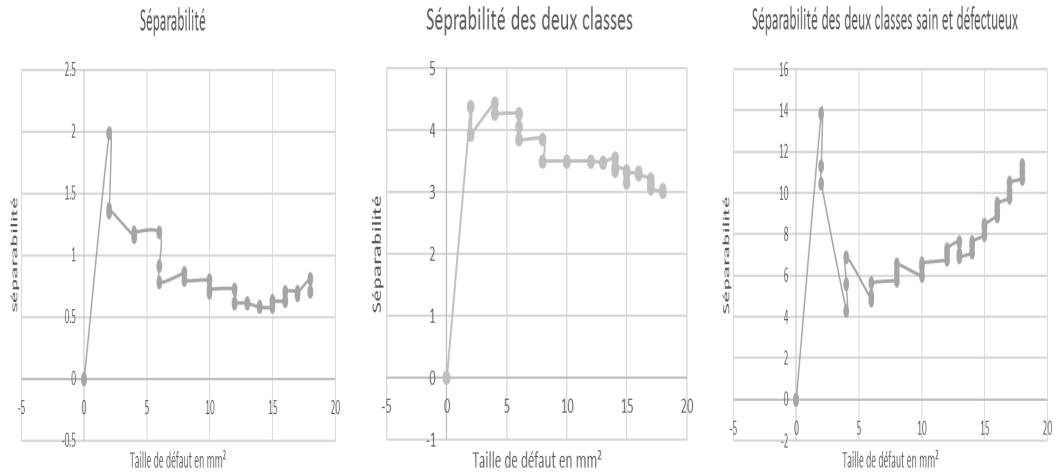
	KPCA-DDBSCAN	LPP-ESDBSCAN	LPP-DFSDBSCAN
Taux de mauvaise classification	<1%	<0.1%	<0.5%
Moment de détection	$2mm^2$	$2mm^2$	$2mm^2$
Taux de retard de détection	0%	0%	0.04%
Taux de fausses alarmes	0.13%	0.1%	0.01%
Taux de gain en mémoire	0%	85%	90%
Évolutions détectées	Saut, dérive, fusion	Saut, dérive, fusion	Saut, dérive, fusion

Comme on peut voir sur le tableau, les taux d'erreur, de fausses alarmes et de retard de détection sont très bas (presque nul). Ces valeurs dépendent de l'initiation des paramètres des classifieurs ainsi que de la qualité des signaux enregistrés.

5.1.6 Suivi de dégradation du roulement par les paramètres caractéristiques des classes

En observant les résultats de classification des observations issues de différents roulements, nous avons pu observer plusieurs phénomènes intéressants que nous pouvons associer à l'état interne du roulement et qui peuvent être utiles pour le suivi du roulement. Nous avons associé la détection de défaut à l'événement de la création d'une nouvelle classe, l'évolution de la taille de défaut à la création d'une classe transitoire et le déplacement de centre de la classe à la variation du conditionnement de fonctionnement.

5. VALIDATION EXPÉRIMENTALE ET MISE EN ŒUVRE SUR UN BANC DE FATIGUE



(a) La séparabilité par *LPP* – (b) La séparabilité par *NPE* – (c) La séparabilité par *DFSDBSCAN* *ESDBSCAN* *KPCA – DDBSCAN*

FIGURE 5.42 – Illustration des évolutions de la séparabilité des classes avec les trois méthodes de classification dynamique

Chaque classe est caractérisée par des paramètres tels que la surface de la classe, la densité, la distance séparant les classes, le nombre de points représentatifs (*ESDBSCAN* et *FNDBSCAN*), le nombre de points attribués à chaque classe, le déplacement du centre de gravité et l'écart type des observations...

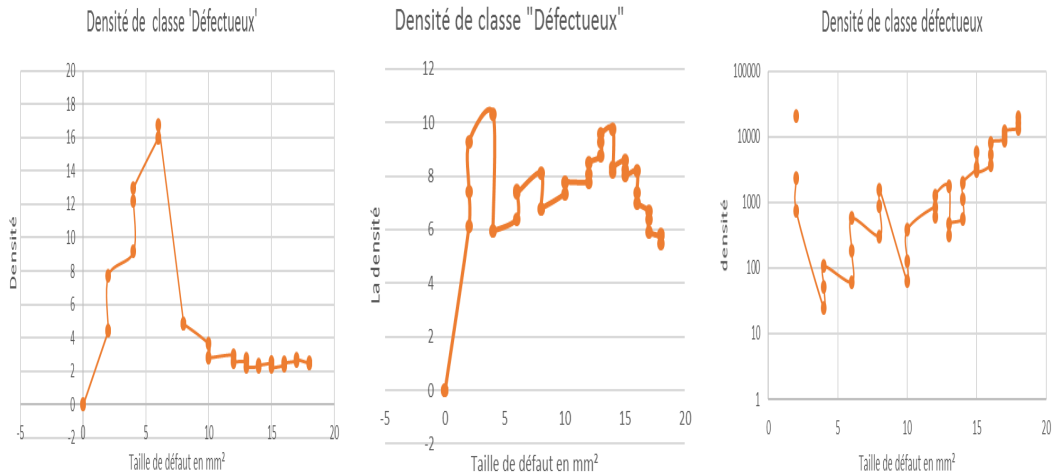
Les paramètres que nous avons retenu sont la surface de la classe, la séparabilité et la densité de la classe. Ces paramètres, en effet, nous ont permis de mieux suivre la dégradation du roulement même dans des conditions non stationnaires (variation de la charge).

La valeur de séparabilité des classes varie selon la méthode de classification employée. Cependant on peut observer la même tendance quelle que soit la méthode. En effet, la séparabilité diminue au fur et à mesure que la taille du défaut augmente (Voir figures 5.42c, 5.42b et 5.42a).

La variation de charge avait peu d'effet sur cette caractéristique. La séparabilité dans *KPCA-DDBSCAN* a été influencé par le défaut et la charge surtout dans les premières phases de dégradation où l'on peut voir que l'augmentation de la charge fait augmenter

5.1 Mise en œuvre de la méthode de classification dynamique sur un banc expérimental

la séparabilité et l'évolution de défaut la fait diminuer (voir figure 5.42c). Cette caractéristique permet d'évaluer l'état d'avancement de la dégradation. Plus sa valeur diminue dans le temps plus l'état de roulement est critique.



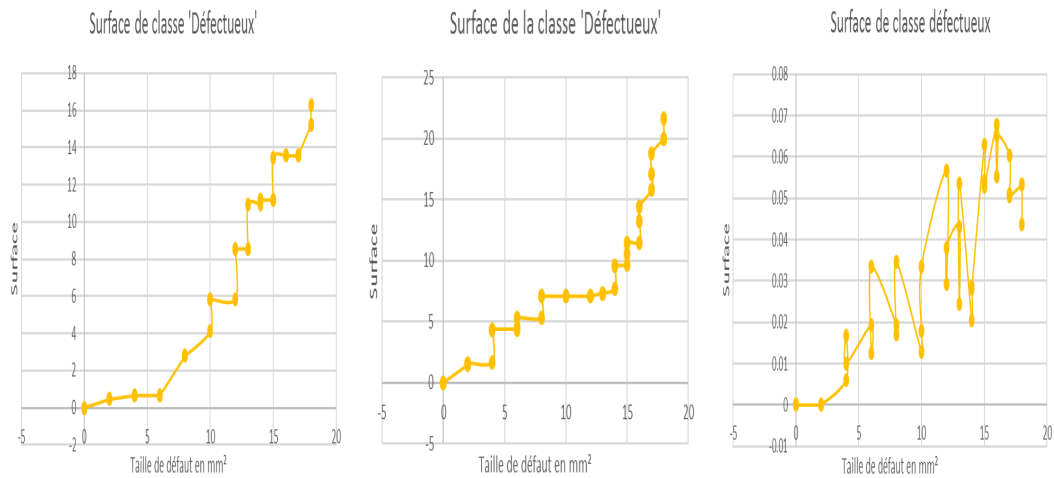
(a) La densité de la classe 'Défectueuse' par $LPP - DFDBSCAN$ (b) La densité de la classe 'Défectueuse' par $NPE - ESDBSCAN$ (c) La densité de la classe 'Défectueuse' par $KPCA - DDBSCAN$

FIGURE 5.43 – Illustration des évolutions de la densité des classes avec les trois méthodes de classification dynamique

L'évolution de la densité de la classe varie selon la méthode employée. Dans la méthode $LPP - DFDBSCAN$, la densité de la classe augmente légèrement en fonction de la taille du défaut et augmente considérablement si la charge augmente. Quand le défaut atteint la taille de $8mm^2$ la densité chute avec l'augmentation de la surface de cette classe. On retrouve les mêmes phénomènes dans la méthode $NPE - ESDBSCAN$, la densité augmente avec la charge diminue avec l'évolution de défaut 5.43b.

Par contre la densité des classes augmente avec l'augmentation de la charge et diminue avec l'augmentation de la taille de défaut dans la méthode $KPCA - DDBSCAN$. Ce paramètre peut être utile pour différencier l'évolution de défaut de celle de la charge.

5. VALIDATION EXPÉRIMENTALE ET MISE EN ŒUVRE SUR UN BANC DE FATIGUE



(a) La surface de la classe 'Défectueuse' par $LPP - DFSDBSCAN$ (b) La surface de la classe 'Défectueuse' par $NPE - ESDBSCAN$ (c) La surface de la classe 'Défectueuse' par $KPCA - DDBSCAN$

FIGURE 5.44 – Illustration des évolutions de la surface des classes avec les trois méthodes de classification dynamique

Quant au paramètre surface de la classe 'Défectueuse', elle augmente avec la taille de défaut dans les trois méthodes. La figure 5.44a montre l'évolution de la surface de la classe 'Défectueuse' de $LPP - DFSDBSCAN$. On peut voir que la surface augmente pour des défauts relativement faible et l'augmentation de la charge n'influence pas cette surface. Par contre, à partir du défaut de taille 10 mm de diamètre, la surface reste stable à la variation de la de la charge mais augmente en fonction taille du défaut.

Ceci s'explique par le fait que l'évolution de défaut provoque la création d'une nouvelle classe autre que la classe 'Défectueuse' la surface de cette classe ne s'ajoute à la surface de la classe 'Défectueuse' que par la suite, généralement cette fusion se coïncide avec la réception d'autres signaux de même roulement mais sous une charge différente.

On retrouve également ces phénomènes dans $NPE - ESDBSCAN$.

La surface de classe 'Défectueuse' trouvée par $KPCA - DDBSCAN$ évolue différemment des autres méthodes (voir figure 5.44c) ceci s'explique par le fait qu'avec 'DDBSCAN' la classification se relance à chaque réception d'un nouveau enregistrement ce qui signifie

5.1 Mise en œuvre de la méthode de classification dynamique sur un banc expérimental

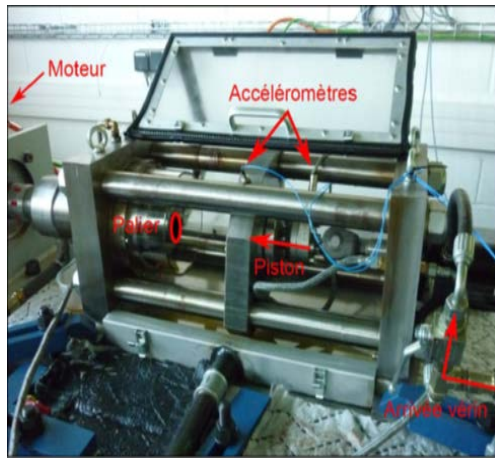
que les points constituant la classe défectueux dans la première itération peuvent être considérée comme outliers dans l'itération suivante.

La surface de la classe 'Défectueuse' peut être également utiliser pour quantifier le niveau de dégradation en analysant le taux d'augmentation de la surface dans le temps.

5.2 Mise en œuvre de la méthode de classification dynamique sur un banc de fatigue

5.2.1 Description du modèle mécanique et du banc expérimental

Nous avons mis en œuvre ces techniques sur un banc expérimental 5.45a constitué d'un moteur entraînant la ligne d'arbre en rotation à 1800 tours/min, d'un palier accueillant une des deux bagues de la butée testée, d'un piston sur lequel est posée l'autre bague de la butée, d'un vérin permettant d'exercer la précontrainte par l'intermédiaire du piston. Deux accéléromètres, l'un est disposé radialement et l'autre axialement sont placés sur le bâti accueillant la bague de la butée.



(a) Banc expérimental.



(b) Butée à billes à simple effet (SNR 51207). Mise en évidence de l'écaillage

FIGURE 5.45 – Le banc expérimental et la butée à billes utilisée durant l'essai

L'opération consiste à placer une bague de la butée sur le palier fixe et à placer l'autre bague avec les billes sur le palier mobile. Le piston est ensuite actionné pour mettre l'assemblage des bagues en contact. On règle la pression afin d'obtenir une charge de 30000N. Le système est mis en rotation par l'intermédiaire du moteur jusqu'à l'apparition d'un défaut d'écaillage sur la piste de la bague de la butée. A ce stade, on procède à une inspection visuelle de la taille du défaut d'écaillage de façon régulière 5.45b. Après l'inspection, la butée est remise en place et le système remis en marche. Cette opération est répétée jusqu'à ce que l'on considère que le défaut devienne trop important ou jusqu'à

5.2 Mise en œuvre de la méthode de classification dynamique sur un banc de fatigue

la ruine de la butée. La butée testée est une butée à billes à simple effet dont la référence SNR est 51207. Elle possède 12 billes et a une capacité de charge dynamique ISO C de 39000 N (5.45b). Nous avons enregistré 8192 points du signal vibratoire dans une gamme de fréquence de 20 kHz pour avoir un nombre de cycle significatif. La fréquence d'enregistrement a été variable : les premiers 50heures de fonctionnement, la fréquence d'enregistrement a varié entre 10h, 5h et 3h, ensuite dès 50heures de fonctionnement un signal a été enregistré chaque 3 heures jusqu'à 98heures de fonctionnement. A partir de 107h de fonctionnement jusqu'à 115h on enregistre un signal chaque heure. Le banc a été démonté après 114heures pour inspection et un défaut de 2.47 mm a été mesuré, la fréquence d'enregistrement a été accélérée après 115h pour enregistrer un signal chaque 30 minutes puis toutes les 5 minutes. Le banc a été démonté 8 fois en total, la première fois a été à t=114h où un défaut de 2.47 mm a été découvert, ensuite à t= 116h le diamètre de défaut s'est élargi pour atteindre 22.78 mm, à t=116h20min le défaut a été de 28.38 mm , à t=117h10min défaut a été de 53.05 mm , à t=118h le diamètre a atteint 83.51 mm de dimaètre et finalement à t=118h25 min où le dimètre finale de défaut a été de 104.6 mm. Après chaque démontage, l'essai a été effectué dans les mêmes conditions initiales.

5.2.2 Extraction des indicateurs de défaut

Nous avons extrait les mêmes indicateurs de défaut choisis dans la section précédente, à savoir

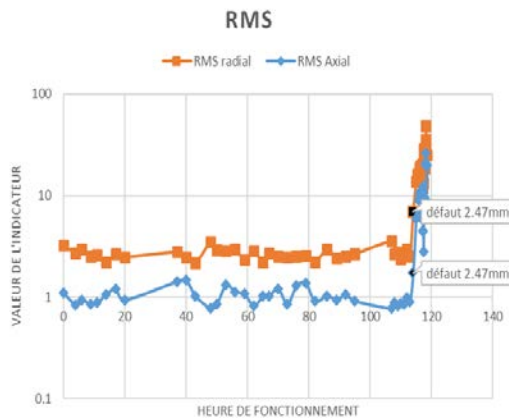
Indicateurs temporels : RMS, Kurtosis, Fc, Skewness, FI

Indicateurs fréquentiels : frms,rmsf, stdf

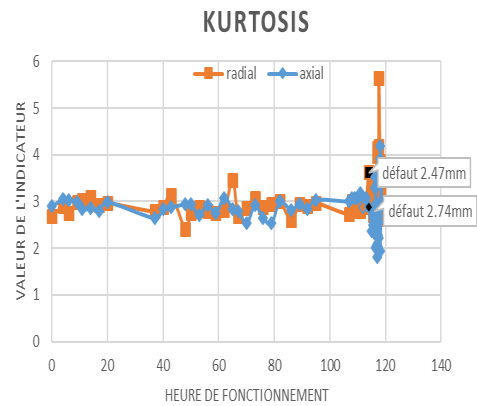
Indicateurs temps-échelle : Pcwt, Wrms, SPRI

on n'a pas gardé que l'indicateur SPRO puisque pour une butée à billes, puisque les deux bagues ont les mêmes caractéristiques à savoir la fréquence de passage des billes sur la bague.

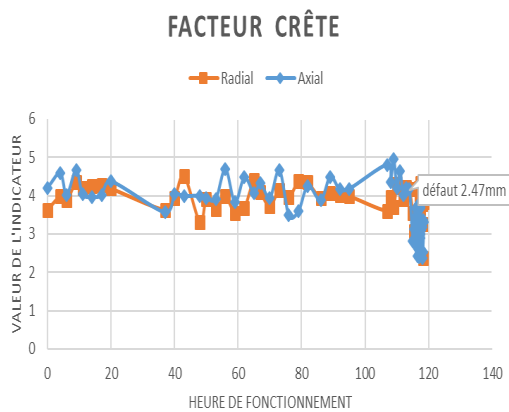
5. VALIDATION EXPÉRIMENTALE ET MISE EN ŒUVRE SUR UN BANC DE FATIGUE



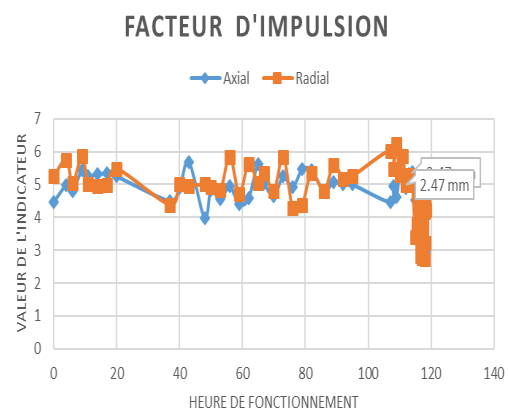
(a) L'évolution de RMS axial et radial durant la période d'essai, mise en évidence de la taille de défaut



(b) L'évolution du Kortosis axial et radial



(c) L'évolution du Facteur crête axial et radial



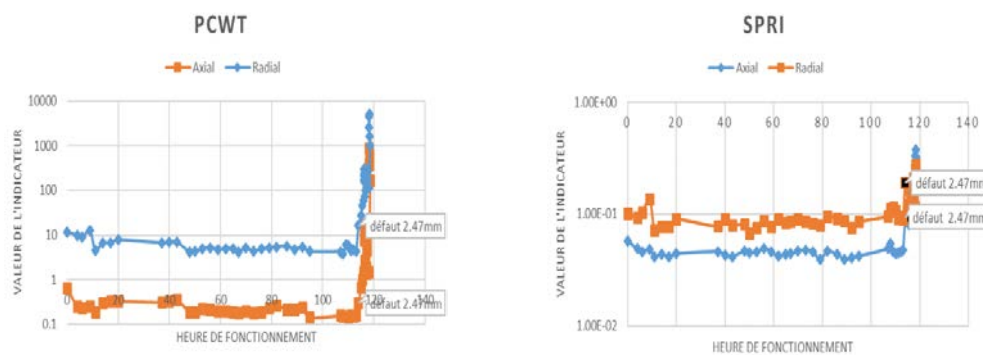
(d) l'évolution de la FI axial et radial

FIGURE 5.46 – L'évolution des indicateurs de défaut durant la période d'essai, moment de détection d'un défaut de 2,47mm est marqué en noir

L'évolution de certains indicateurs est moins significative en termes de détection par rapport à celle retrouvé précédemment. La valeur du *kurtosis* par exemple reste proche de 3 (la valeur associé à un roulement sain) même après l'apparition d'un défaut. La valeur de *Fc* et *FI* fluctuent autour d'une valeur moyenne durant toute la période de l'essai même après la détection de défaut (Fig 5.46c et 5.46d).

5.2 Mise en œuvre de la méthode de classification dynamique sur un banc de fatigue

La valeur RMS affiche une augmentation au moment de la détection d'un défaut de 2,47mm sur la bague de la butée à billes (voir figure 5.46a). Les indicateurs issus de l'analyse d'enveloppe sont meilleurs que ceux issus de l'analyse temporelle avec lesquelles on peut voir clairement une augmentation de leurs valeurs au moment de la détection. Les valeurs des X_{PCWT} radial et axial par exemple ont connu une augmentation au moment de la détection d'un défaut de 2,47mm sur la bague 5.47b, même chose pour le $SPRI$ comme on peut voir sur la figure 5.47a. Une détection basée seulement sur ces indicateurs où la majorité d'eux affichent une évolution insignifiante au moment de détection peut être imprécise ou difficile.



(a) L'évolution du PCWT axial et radial

(b) l'évolution de SPRI axial et radial

FIGURE 5.47 – L'évolution des indicateurs de défauts issus de l'analyse d'enveloppe

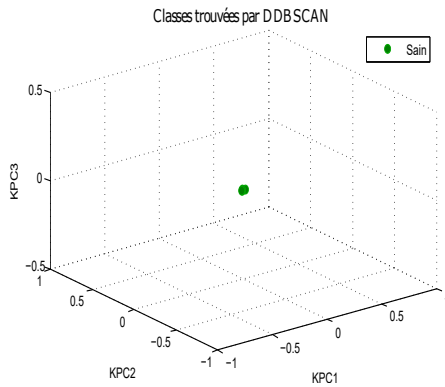
Nous avons donc utilisé tous ces indicateurs même si certains indicateurs ne sont pas très sensibles à l'apparition du défaut.

5.2.3 Classification dynamique

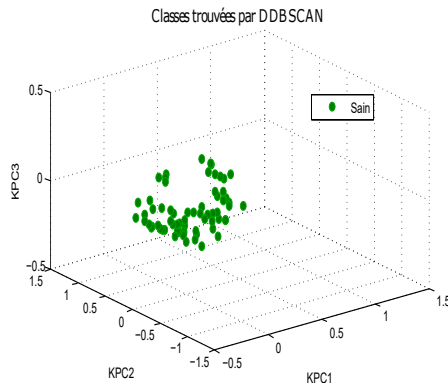
A. *KPCA – DDBSCAN*

Les paramètres de *DDBSCAN* ont été choisis comme suit ; $MinPts = 4$ qui représente le nombre de points par enregistrement et $Eps = 0.5$. Comme on peut voir sur les figures de 5.48a, 5.48b et 5.48c, la méthode *DDBSCAN* classe toutes les observations reçues de $t = 0$ h à $t = 113$ h dans la même classe 'Saine', et la surface de la classe s'agrandit . A $t = 114$ h une nouvelle classe se crée, cette classe coïncide avec la détection d'un défaut d'un diamètre de 2,47mm 5.48d.

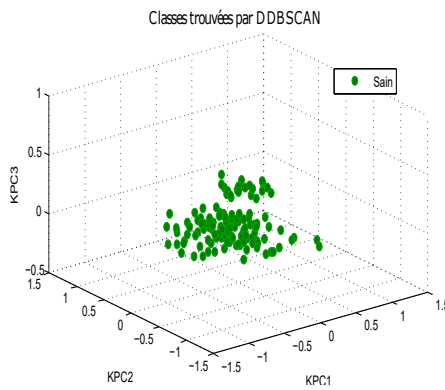
5.2 Mise en œuvre de la méthode de classification dynamique sur un banc de fatigue



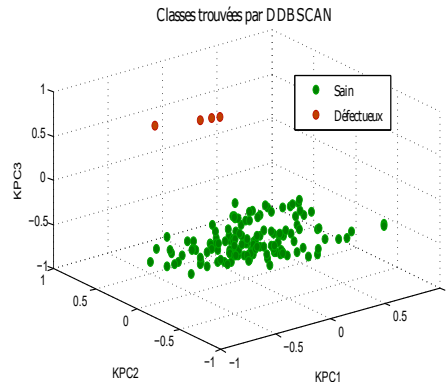
(a) La classification des nouvelles observations par *KPCA - DDBSCAN* à $t=0$ h



(b) La classification des nouvelles observations par *KPCA - DDBSCAN* à $t=62$ h



(c) La classification des nouvelles observations par *KPCA - DDBSCAN* à $t=113$ h



(d) La classification des nouvelles observations par *KPCA - DDBSCAN* à $t=114$ h. Le défaut observé est d'un diamètre plus de 2,47 mm

FIGURE 5.48 – Classification des observations reçues avant la détection et durant les premières phases de dégradation

Après montage du roulement défectueux, on observe un phénomène de création de classes transitoires provoquées par l'évolution du défaut (Fig. 5.49a) et puis une fusion de ces classes avec la classe "défectueuse". Après la fusion, une nouvelle classe "défectueuse" est ensuite créée (Fig. 5.49b) puis fusionnée avec classe 'Défectueuse' (Fig. 5.49c).

5. VALIDATION EXPÉRIMENTALE ET MISE EN ŒUVRE SUR UN BANC DE FATIGUE

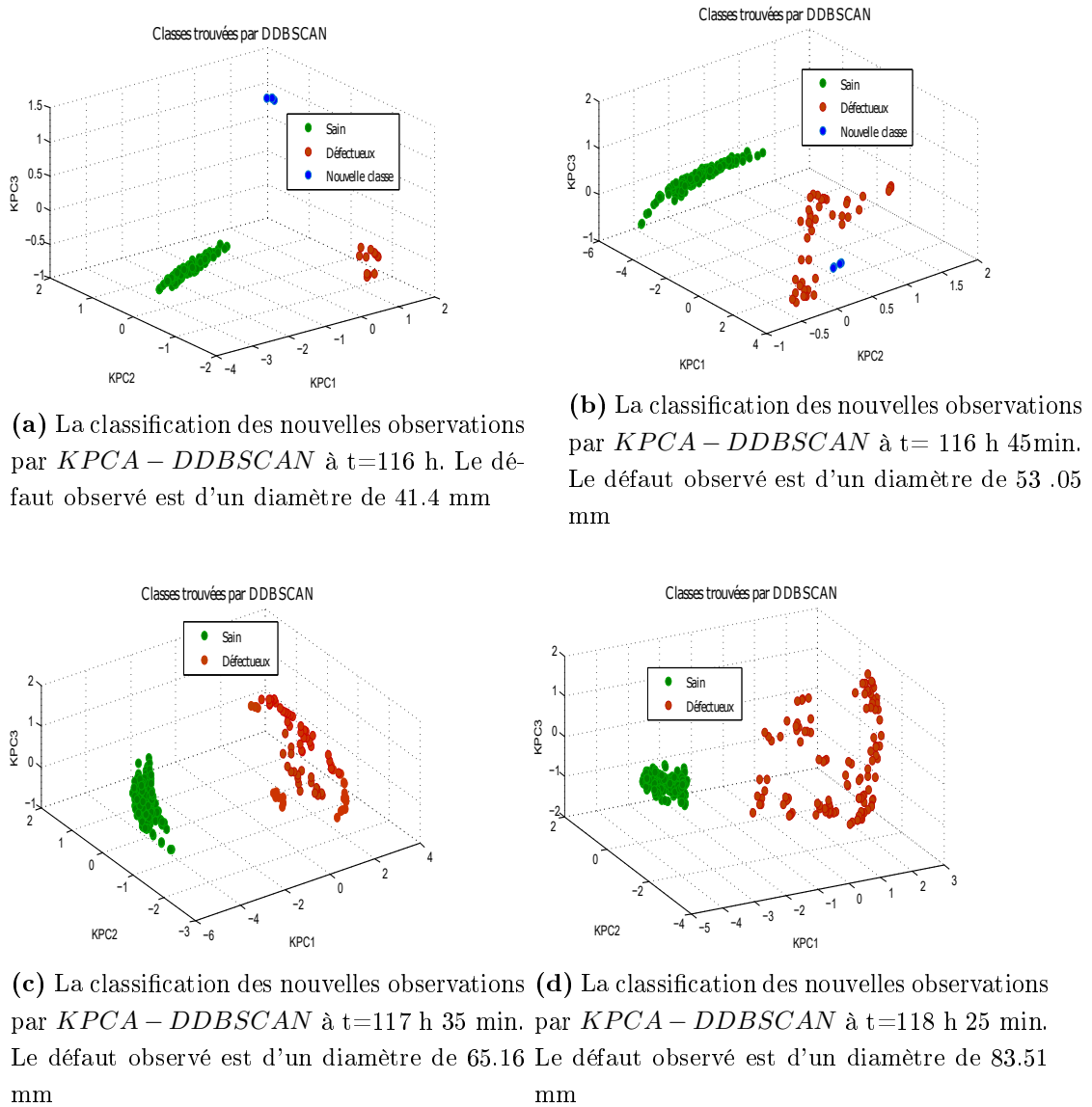


FIGURE 5.49 – La classification des observations par $DDBSCAN$ après la détection

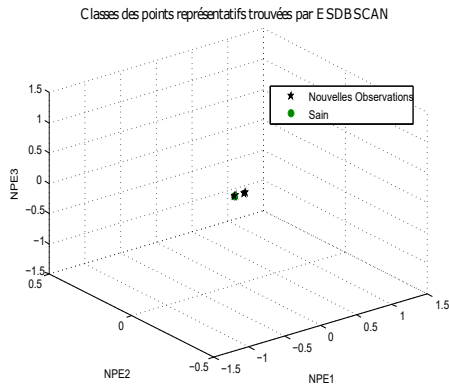
Le suivi du roulement par la méthode $KPCA - DDBSCAN$ nous a permis de valider l'hypothèse que l'évolution de défaut se manifeste par des sauts d'observations qui provoquent la création de nouvelles classes. Ces classes se fusionnent avec la classe 'Défectueuse' par les mécanismes du fusion implémentés dans la méthode de classification. Elle nous a permis également de valider la proposition d'associer l'événement de la création

d'une nouvelle classe avec la détection de défaut et de valider notre choix de paramètres puisque le taux d'erreur est resté tout au long de la période d'essai égal à 0% .

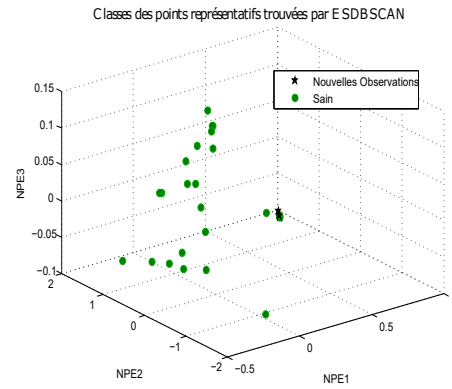
B. *NPE – ESDBSCAN*

Comme on a expliqué dans la section précédente, la méthode *DDBSCAN* nécessite l'accès à la totalité des observations reçues, une limite qu'on ne peut pas respecter dans un cas réel, où le nombre d'observations est important. L'application du couple *NPE – ESDBSCAN* sur la matrice des indicateurs standardisés a donnée les résultats visibles sur les figures de 5.50a à 5.51d. La méthode a classifié les observations reçues entre $t=0$ h à $t=112$ h dans la même classe 'Saine' (comme on peut voir dans les figures 5.50a , 5.50b). Des points représentatifs ont été créés au fur et à mesure de la réception des nouvelles observations. A $t=113$ h une nouvelle classe a été créée, une heure avant la détection d'un défaut de 2.47 mm de diamètre.

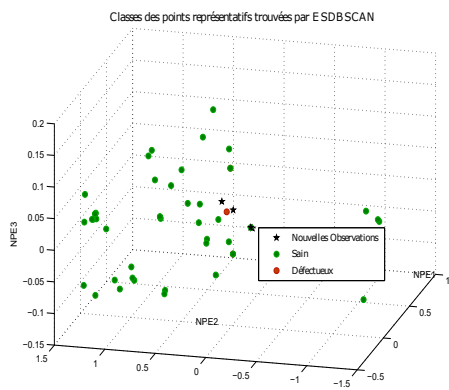
5. VALIDATION EXPÉRIMENTALE ET MISE EN ŒUVRE SUR UN BANC DE FATIGUE



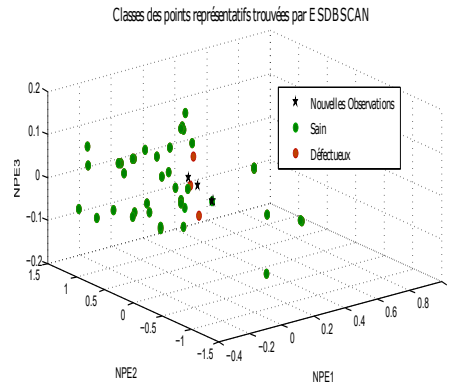
(a) La classification des nouvelles observations par $NPE - ESDBSCAN$ à $t=0$ h



(b) La classification des nouvelles observations par $NPE - ESDBSCAN$ à $t=62$ h



(c) La classification des nouvelles observations par $NPE - ESDBSCAN$ à $t=113$ h



(d) La classification des nouvelles observations par $NPE - ESDBSCAN$ à $t=114$ h. Le défaut observé est d'un diamètre de 2,47 mm

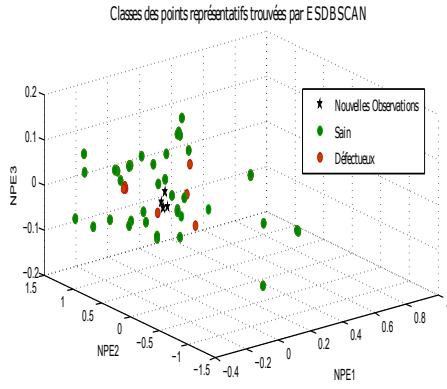
FIGURE 5.50 – La classification des observations reçues avant la détection et durant les premiers stades de dégradation

Après la détection de défaut, la classe 'Défectueuse' a continué à s'agrandir et les observations ont continué à être attribuées à des nouveaux points représentatifs ou à des points représentatifs existants de la classe. On remarque aussi la création de nouvelles classes transitoires après la détection de défaut pour marquer le passage de défaut à un niveau de dégradation plus avancé notamment à $t=116$ h, à $t=117$ h 10 min et à $t=118$ h).

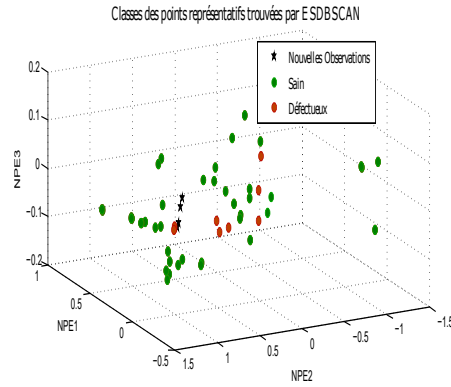
Bien que les points représentatifs appartenant de deux classes apparaissent chevauchés

5.2 Mise en œuvre de la méthode de classification dynamique sur un banc de fatigue

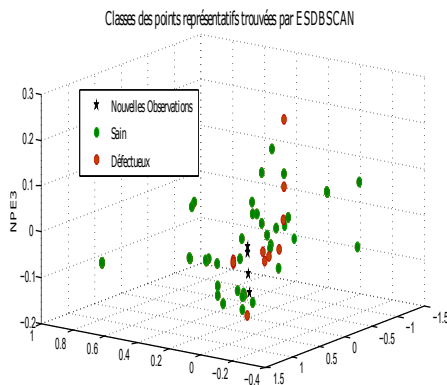
sur les figures, ils ne le sont pas réellement. La distance séparant ces points représentatifs de différentes classes est plus grande que celle séparant les points de la même classe. Il faut aussi penser que ces points représentatifs englobent à leur tour d'autres observations précédemment reçues.



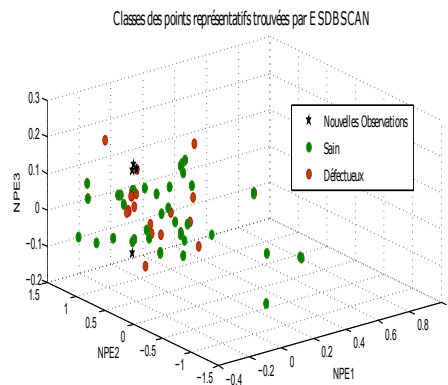
(a) La classification des nouvelles observations par $NPE - ESDBSCAN$ à $t=116$ h 45 min. Le défaut observé est d'un diamètre de 41.4 mm



(b) La classification des nouvelles observations par $NPE - ESDBSCAN$ à $t=117$ h. Le défaut observé est d'un diamètre de 53.05 mm



(c) La classification des nouvelles observations par $NPE - ESDBSCAN$ à $t=117$ h 35 min. Le défaut observé est d'un diamètre de 65.16 mm



(d) La classification des nouvelles observations par $NPE - ESDBSCAN$ à $t=118$ h 10 min. Le défaut observé est d'un diamètre de 83.51 mm

FIGURE 5.51 – La classification des observations par $ESDBSCAN$ après la détection de défaut

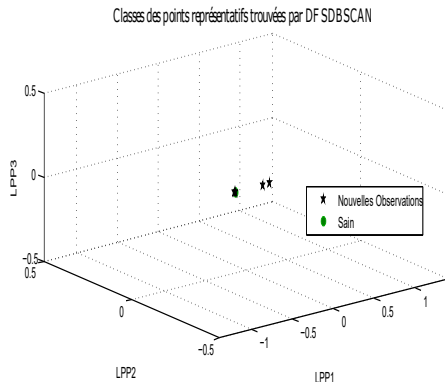
5. VALIDATION EXPÉRIMENTALE ET MISE EN ŒUVRE SUR UN BANC DE FATIGUE

Cette méthode *NPE – ESDBSCAN* détecte plus tôt le défaut que sa précédente (*KPCA-DDBSCAN*), le défaut a été détecté à $t = 113$ h alors que *KPCA – DBSCAN* a détecté le défaut à $t = 114$ h. Le taux d'erreur calculé durant la période d'essai a été de moins de 1% (quand le défaut est dans un état avancé, des observations que normalement correspondent à une classe défectueuse, la méthode les a attribué à la classe sain) et le taux d'occupation de mémoire a été moins de 25% la taille totale des observations reçues.

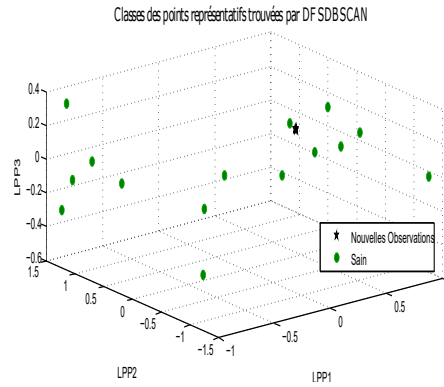
C. *LPP – DFSDBSCAN*

En appliquant la méthode *LPP – DFSDBSCAN* sur la matrice des indicateurs normalisée, on a retrouvé une meilleure visibilité de défaut. La méthode a classifié les observations reçues dans l'intervalle $t = 0$ h et $t = 113$ h à la classe 'Saine' en créant des points représentatifs quand les observations ne peuvent pas être attribuées à des points représentatifs existants (voir figures 5.52a, 5.52b). A $t = 112$ h les observations reçues ainsi que celles stockés dans le buffer ont été attribuées à des nouveaux points représentatifs assez loin de tous les points représentatifs de la classe 'Saine' provoquant ainsi la création d'une nouvelle classe 'Défectueuse' (voir 5.52c). Ces points représentatifs, que la classification stricte par (*ESDBSCAN*) les a attribués à la classe 'Saine', la classification floue par *DFSDBSCAN* les a affectées à la nouvelle classe 'Défectueuse'.

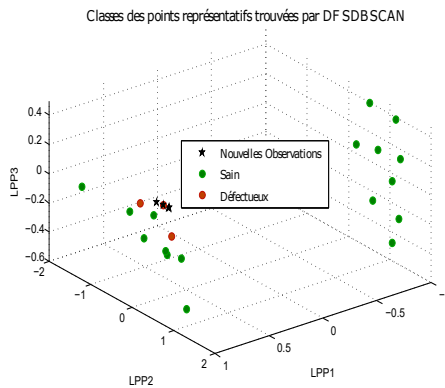
5.2 Mise en œuvre de la méthode de classification dynamique sur un banc de fatigue



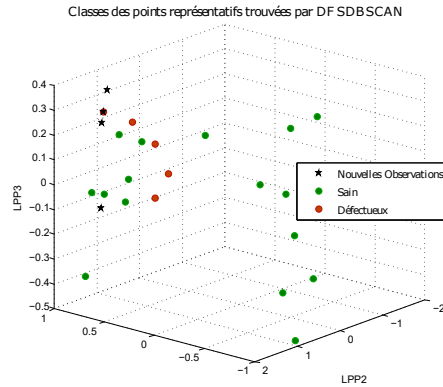
(a) La classification des nouvelles observations par $LPP - DF SDBSCAN$ à $t=0$ h



(b) La classification des nouvelles observations par $LPP - DF SDBSCAN$ à $t=62$ h



(c) La classification des nouvelles observations par $LPP - DF SDBSCAN$ à $t=112$ h



(d) La classification des nouvelles observations par $LPP - DF SDBSCAN$ à $t=114$ h. Le défaut observé est d'un diamètre de 2,47mm

FIGURE 5.52 – La classification des observations avant la détection de défaut et durant ses premières phases de dégradation

A $t=114$ h où on a observé un défaut de 2.47 mm, la classe 'Défectueuse' est déjà constituée de 5 points représentatifs regroupant plus de 12 observations. Bien qu'on ne peut pas savoir exactement la taille de défaut détecté à $t=112$ h, on peut dire que la détection a été la plus précoce par rapport à toutes les autres méthodes de classification testées sur ce banc.

N.B : Les points représentatifs de la classe 'Défectueuse' apparaissent proches de la classe 'Saine' dans les figures, alors que les distances séparant ces points de ceux de la classe

5. VALIDATION EXPÉRIMENTALE ET MISE EN ŒUVRE SUR UN BANC DE FATIGUE

'Saine' sont plus grandes en réalité. Ceci peut être dû aux choix de l'angle de vue, des axes de représentation (L'affichage a été fait sur 3 axes de LPP alors que la classification se fait sur tous les vecteurs de LPP) et le fait que ces points représentatifs englobent des observations invisibles dans les figures.

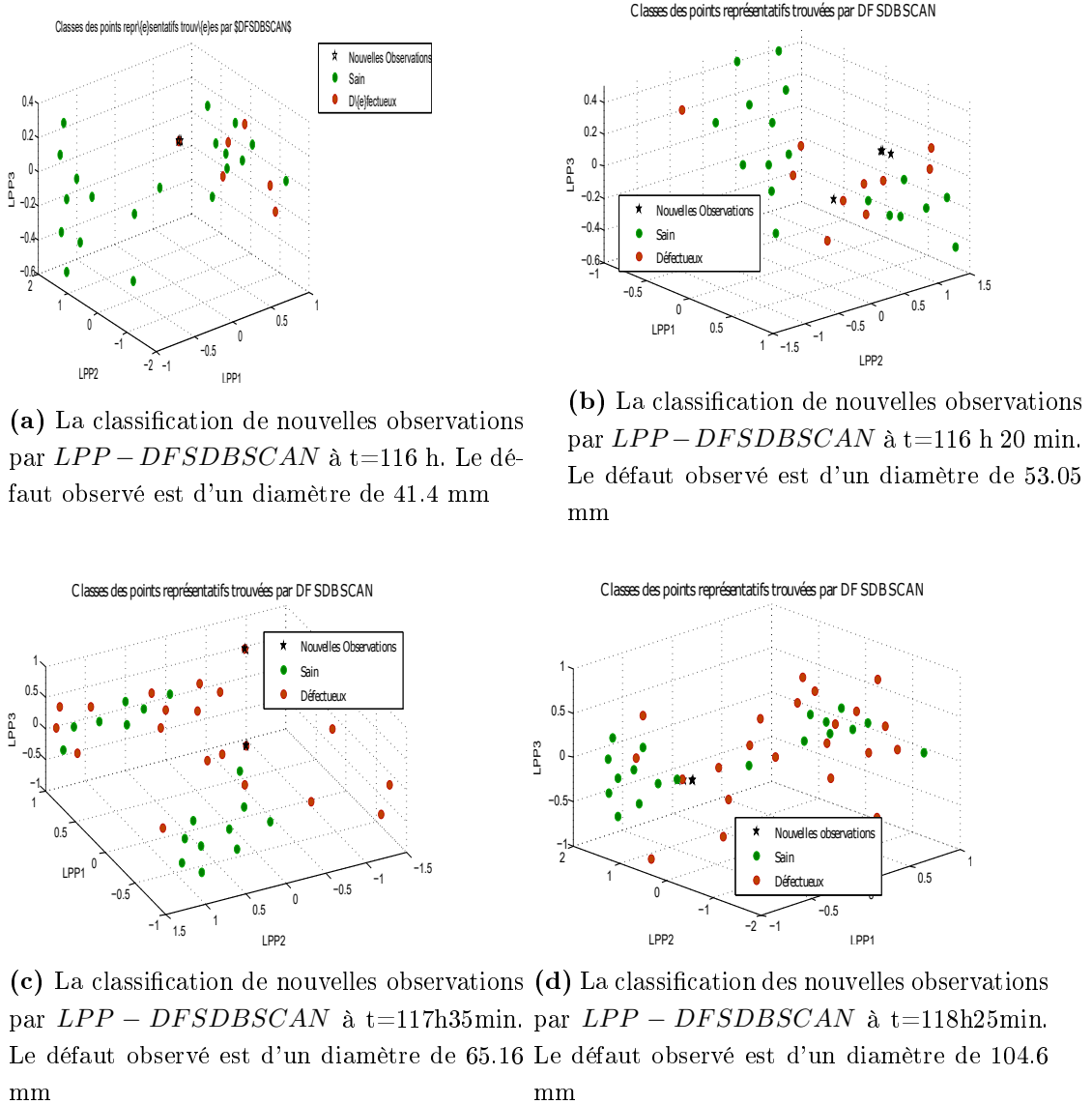


FIGURE 5.53 – La classification des observations par $DFSDBSCAN$ après la détection de défaut

Après la détection, on a observé les mêmes phénomènes de création de nouvelles classes transitoires à $t=116$ h 15min et à $t=117$ h 40 min et à $t=118$ h 15 min. Ces classes se fusionnent avec la classe 'Défectueuse' par la suite (comme on peut voir sur les figures 5.53b, 5.53c et 5.53d). La cadence de création des points représentatifs augmente au fur et à mesure que le roulement se dégrade. Le taux d'erreur a augmenté quand l'état de roulement est devenu critique pour atteindre la valeur de 2.5%. Quant au taux d'occupation de mémoire, il a atteint 15 % de la taille des observations reçues..

La détection avec cette méthode *LPP – DFSDBSCAN* a été plus précoce que ses précédentes, le défaut a été détecté à $t=112$ h contrairement à *NPE – ESDBSCAN* où le défaut n'a été détecté qu'à $t=113$ h et *KPCA – DDBSCAN* à $t=114$ h.

5.2.4 Bilan sur les résultats obtenus par les classifieurs proposés

La classification dynamique nous a permis de détecter le défaut précocement et de voir l'évolution de la taille de défaut autrement. La classification par *ESDBSCAN* et *DFSDBSCAN* ont permis d'optimiser la mémoire occupée durant l'essai tout en détectant le défaut à 1h de *NPE – ESDBSCAN* à 2h de *KPCA – DDBSCAN*. Ces méthodes ont détecté bien avant les indicateurs classiques de défaut (à 114h). En comparant les résultats obtenus après l'association des méthodes de réduction avec les classifieurs développés durant cette thèse, on peut voir que trois combinaisons présentent les meilleurs résultats en termes de taux d'erreur, taux de fausses alarmes, précocité de détection et la visibilité des évolutions des classes associées au changement de l'état du roulement surveillé. Ces combinaisons sont (*KPCA-DDBSCAN*, *LPP-ESDBSCAN*, *LPP-DFSDBSCAN*).

Comme on peut voir sur le tableau 5.2, *LPP – ESDBSCAN* et *LPP – DFSDBSCAN* ont permis de détecter le défaut plus précocement (1h à 2h) qu'avec *KPCA – DDBSCAN* et en utilisant que 25% ou 15% de données employés par *DDBSCAN*.

5.2.5 Suivi de dégradation du roulement par les paramètres caractéristiques des classes

Nous avons également suivi les caractéristiques de classes pour chaque méthode de classification. Les figures (fig 5.54, fig 5.55 et fig 5.56) montrent l'évolution de ces caractéristiques.

5. VALIDATION EXPÉRIMENTALE ET MISE EN ŒUVRE SUR UN BANC DE FATIGUE

TABLE 5.2 – Tableau comparatif des résultats obtenus avec classifieurs proposés

	KPCA-DDBSCAN	LPP-ESDBSCAN	LPP-DFSDSCAN
Taux de mauvaise classification	<0.5%	<0.3%	<2.5%
Moment de détection	114h (défaut 2,47mm)	113h (taille inconnue)	112h (taille inconnue)
Taux de retard de détection	0.25%	0.14%	$0.1\% < T < 0.14\%$
Taux de fausses alarmes	0.25%	0.16%	$0.007\% < T < 0.1$
Taux de gain en mémoire	0%	75%	85%
Évolutions détectées	Saut	Saut, dérive lente vers classe saine en fin de dégradation	Saut, dérive lente vers classe saine en fin de dégradation

Le premier paramètre de suivi est la séparabilité des classes qui représente la distance séparant les deux classes 'Saine' et 'Défectueuse'. Comme on peut voir sur les figures Fig.5.54a et Fig.5.54b, l'évolution de ce paramètres est la similaire pour les deux méthodes *LPP-DFSDSCAN* et *NPE-ESDBSCAN*. Dans la méthode *LPP-DDBSCAN* la valeur de la séparabilité augmente après la détection et chute quand la taille du défaut atteint 22.78 mm. Cette valeur se stabilise quand la taille du défaut atteint 53.05 mm alors que dans la méthode NPE-ESDBSCAN elle augmente à nouveau.

Dans la méthode *KPCA-DDBSCAN*, la valeur de la séparabilité augmente avec l'évolution du défaut jusqu'à ce que la taille de défaut atteint un niveau critique (83.51 mm) puis elle diminue.

5.2 Mise en œuvre de la méthode de classification dynamique sur un banc de fatigue

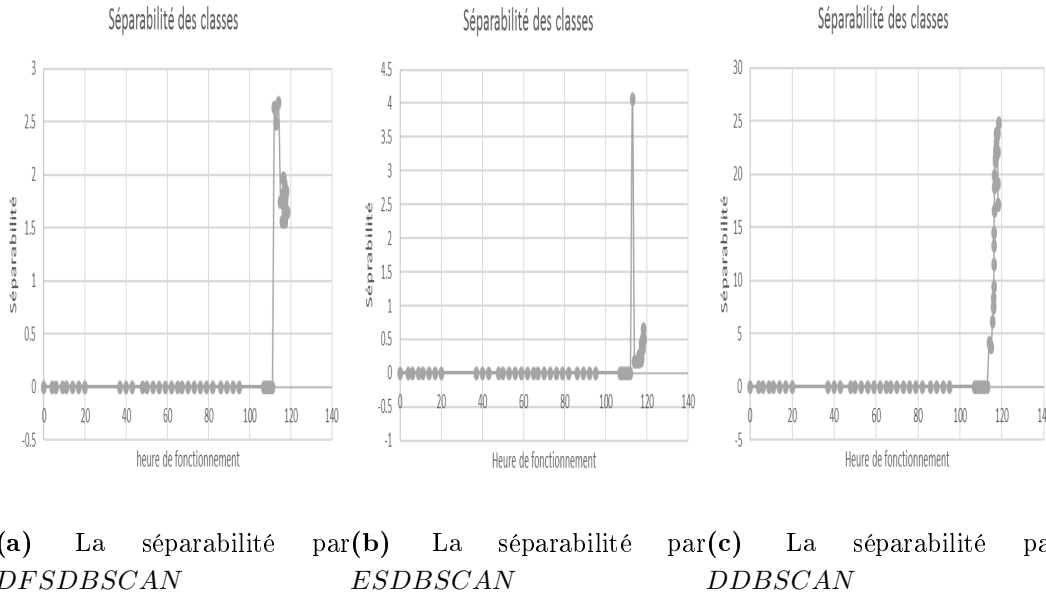
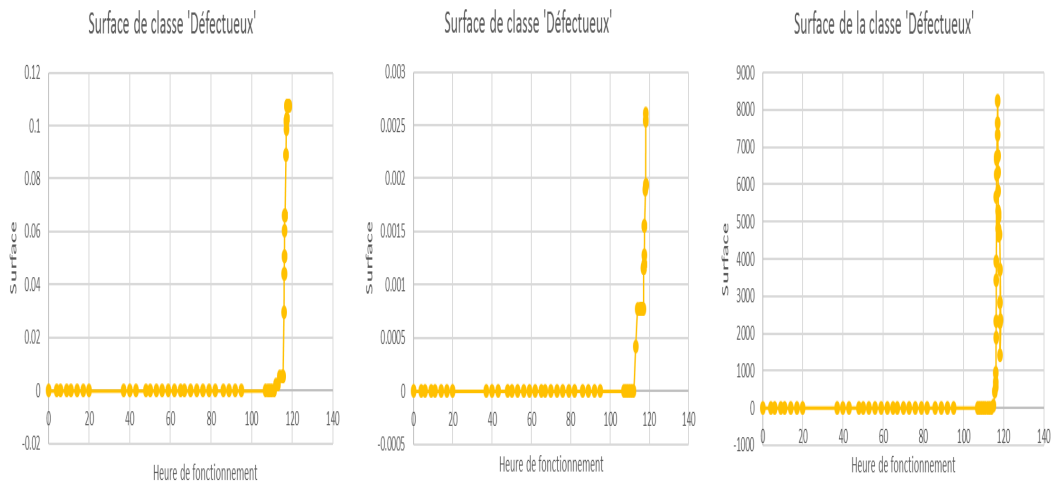


FIGURE 5.54 – Illustration des évolutions de la séparabilité des classes avec les trois méthodes de classification dynamique

La surface de la classe défectueux dans les deux méthodes (*NPE – ESDBSCAN* et *LPP – DFSDSCAN*) suit le même type d'évolution : la surface augmente avec l'évolution de la taille de défaut, elle augmente légèrement au début puis fortement quand l'état de dégradation devient critique (voir la figure 5.56).

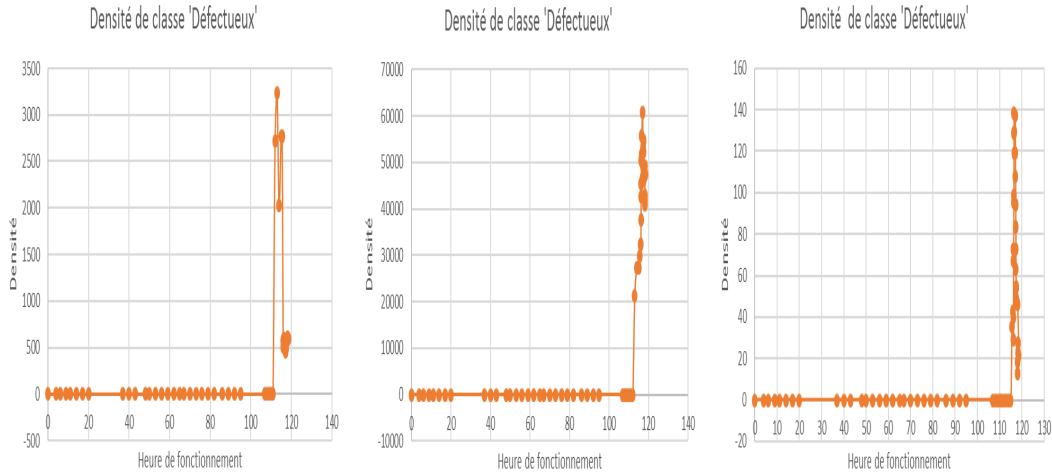
La densité de la classe 'Défectueuse' par *LPP – DFSDSCAN*, comme affiché sur la figure 5.43a augmente après la détection puis diminue quand la taille de défaut atteint une certaine taille puis se stabilise quand l'état de dégradation de roulements est critique. La densité par *NPE – ESDBSCAN* et *KPCA – DDBSCAN* ont la même tendance ; la valeur de la densité augmente avec l'évolution de défaut jusqu'à ce que le roulement atteigne un niveau de dégradation avancé puis commence à diminuer (voir figures Fig. 5.43b et Fig. 5.43c).

5. VALIDATION EXPÉRIMENTALE ET MISE EN ŒUVRE SUR UN BANC DE FATIGUE



(a) La surface de la classe 'Défectueuse' par *DFSDBSCAN* (b) La surface de la classe 'Défectueuse' par *ESDBSCAN* (c) La surface de la classe 'Défectueuse' par *DDBSCAN*

FIGURE 5.55 — Illustration des évolutions de la surface de la classe 'Défectueuse' trouvée avec les trois méthodes de classification dynamique pendant la période d'essai



(a) La densité de la classe 'Défectueuse' par *DFSDBSCAN* (b) La densité de la classe 'Défectueuse' par *ESDBSCAN* (c) La densité de la classe 'Défectueuse' par *DDBSCAN*

FIGURE 5.56 – Illustration des évolutions de la densité de la classe 'Défectueuse' trouvée avec les trois méthodes de classification dynamique pendant la période d'essai

Ces paramètres peuvent être employés pour suivre et quantifier la dégradation de roulement en analysant par exemple l'évolution de la surface de classe, le sens d'évolution de la séprabilité et de la densité. Une étude approfondie de l'évolution de ces paramètres sur plusieurs essais et des bancs expérimentaux sous différentes conditions pourra nous aider pour confirmer l'apport de ces paramètres dans le suivi de roulements.

5.3 Conclusion du chapitre

L'objectif de ce chapitre est de valider les différentes méthodes proposées sur d'un banc expérimental et un banc de fatigue. Dans la première partie de ce chapitre, nous avons mené une étude comparative des différents indicateurs de défaut issus de l'analyse vibratoire afin de sélectionner ceux qui resteront sensible au défaut même dans le cas des conditions de fonctionnement non stationnaires (à savoir une variation de charge). Les résultats de cette étude nous ont permis de conclure que la majorité des indicateurs réagissent à la présence de défaut et aux changements de condition de fonctionnement (variation de charge). Cependant l'évolution de l'indicateur dû à un défaut et celle dû à

5. VALIDATION EXPÉRIMENTALE ET MISE EN ŒUVRE SUR UN BANC DE FATIGUE

une changement de charge ne sont pas les mêmes.

Les méthodes de réduction de dimensions utilisées (*LPP* , *NPE* et *KPCA*) ont permis d'améliorer la séparation entre état sain et état défectueux même en cas de variation de condition de fonctionnement (notamment la charge). Nous avons également proposé une méthode de sélection des caractéristiques *SFS* paramétrée pour fonctionner dans un environnement non supervisé. Cette méthode a permis de sélectionner directement de la matrice des indicateurs ceux qui portent plus d'informations utiles à la détection. Dans la partie classification dynamique : La première méthode de classification *DDBSCAN* combinée avec *LPP* et *KPCA* a présenté le meilleur taux d'erreur et la meilleure séparation des classes 'Saine' et 'Défectueuse'. En analysant les résultats de cette classification on a pu révéler l'apparition des phénomènes qui nous n'étaients pas visibles avec les méthodes de classification statiques à savoir la création d'une nouvelle classe provoquée par le saut d'observation lors de l'apparition d'un défaut, la création des classes temporaires qui se fusionnent avec la classe 'Défectueuse' à chaque passage de roulement à un nouveau état de dégradation. Aussi, le centre de classe se déplace quand la charge varie. Bien que les résultats de la classification par *DDBSCAN* ont été plus que satisfaisants, cette méthode ne pourra pas être utile que si la taille des données est limitée alors que dans un traitement en ligne la taille des données est présumée infinie.

Pour remédier à ce problème, nous avons développé *ESDBSCAN*. Cette méthode a été également combinée avec les différentes méthodes de réduction de dimension et sélection pour choisir la meilleure combinaison. La combinaison qui a affiché les meilleures résultats à été *NPE – ESDBSCAN* avec le taux d'erreur le plus faible et une bonne séparation entre les deux classes même si la charge varie.

Une classification floue est plus adaptée dans suivi qu'une classification stricte et un apprentissage semi supervisé est également plus adapté qu'un apprentissage non supervisé. C'est pourquoi nous avons développé la méthode *DFSDBSCAN*. Cette méthode combinée avec *LPP* a permis d'améliorer la détection et la séparation entre classe 'Saine' et classe 'Défectueuse'. Elle a permis également de baisser le taux d'erreur ainsi que le taux d'observations utilisés dans la classification.

A partir des résultats de la classification dynamique, nous avons alors extraits des caractéristiques des classes après chaque réception des nouvelles observations puis nous avons

suivi leur évolution en fonction de défaut. Après avoir analysé les différentes caractéristiques possibles, la surface, la densité et la séparabilité ont été jugé les plus informatifs. Le suivi des évolutions des ces caractéristiques en fonction de défaut (même dans des conditions non stationnaires) peuvent être employées pour quantifier la dégradation d'un roulement ou pour différencier un changement dû à l'évolution de défaut à celle dû à une augmentation de charge.

Dans la deuxième partie de chapitre, nous avons validé les méthodes proposées sur des données issues d'un banc de fatigue. Ces méthodes, dont le choix des paramètres et des indicateurs a été basé sur les résultats de la première partie de chapitre, ont montré leur efficacité quant à la détection et au suivi de défaut. On a appliqué d'abord la méthode *KPCA-DBSCAN* sur ces observations (vu que l'essai s'est interrompu après 119h de fonctionnement la taille finale des observations a été gérable). La méthode *DBSCAN* a permis de classer correctement les observations avec un taux d'erreur de 0%. On a observé les mêmes phénomènes que dans la première partie : création d'une nouvelle classe à l'apparition de défaut et la création des classes transitoires avec l'évolution de défaut, la méthode a détecté un défaut de 2.47mm. L'application de *NPE-ESDBSCAN* sur les observations a permis de détecter le défaut plus tôt qu'avec *DBSCAN* (une heure). La tailles d'observations gardées dans la mémoire pour être utilisées dans la classification ainsi que le taux d'erreur ont été satisfaisants. L'application de *DFSDBSCAN* a permis de détecter le défaut 2 heures avant qu'il atteigne le diamètre de 2.47mm. Le taux d'erreur et le taux d'occupation de la mémoire ont été satisfaisants.

Le suivi de caractéristiques (surface, densité et séparabilité) a été informatif et plus particulièrement la surface de la classe défectueux.

Chapitre 6

Conclusion générale et perspectives

Ce projet de recherche a porté sur la détection et le suivi de défaut de roulement par la méthode de classification dynamique. Ce travail s'inscrit dans le cadre d'un projet de recherche réunissant trois laboratoires : CRESTIC (URCA), GRESPI (URCA) et le laboratoire LTI (Modélisation, Informatique et Systèmes, UPJV). Il s'agit de développer un système de diagnostic et de suivi fiable d'une structure à contact glissant et qui fait appel à des compétences pluridisciplinaires, mécanique, traitement du signal, système embarqué et transmission sans fil, automatique dans le cadre de la recherche des méthodes de classification pertinentes. Ce projet de recherche concerne l'extraction des paramètres pertinents des signaux de natures différentes (vibratoires, acoustiques, électriques) par des techniques de traitement du signal et le développement des méthodes de classification en vue d'établir un diagnostic et un suivi de dégradation des composants ou des outils-pièces. Le choix des paramètres dits " indicateurs de l'état du composant " est essentiel pour la surveillance et le suivi de chaque composant sensible. Ces paramètres doivent être définis pour chaque type de signal (vibratoire, acoustique et électrique). La classification de ces paramètres et l'intégration d'un processus d'aide à la décision basé sur l'évolution de ces paramètres dans un système embarqué fourniront un dispositif de diagnostic fiable pour tout système à contact glissant. Dans ce travail, nous nous sommes concentrés sur les méthodes de classification dynamique des données vibratoires.

Après avoir rappelé les différentes stratégies de maintenance, le premier chapitre a présenté les différentes techniques de traitement du signal pour la maintenance des machines de production par analyse vibratoire. Ce chapitre présente également les différents types

6. CONCLUSION GÉNÉRALE ET PERSPECTIVES

de défauts de roulements susceptibles d'apparaître ainsi que les méthodes utilisées pour les détecter.

Le deuxième chapitre est un état de l'art sur les méthodes de reconnaissance de formes dans le cadre de diagnostic des machines tournantes. Cette étude bibliographique nous a permis de constater que ces méthodes ont été utilisées sur une base de données provenant soit du même roulement avec différents états de dégradation (même type de défaut mais avec des tailles différentes) soit des roulements différents avec différents type de défauts (défaut bague extérieure, défaut bague intérieure, défaut de bille). Cette étude bibliographique nous a également permis de conclure que les méthodes de classification dans le cadre de diagnostic des machines ont été spécialement utilisées pour leur qualité de "séparation automatique" des données différentes ou pour trouver un "modèle" qui peut expliquer la relation reliant les données avec leurs classes. L'utilisation de ces méthodes présente toutefois des inconvénients dans le cadre du suivi car elles supposent la connaissance a priori de toutes les observations nécessaires à la caractérisation des classes. Ces méthodes de RdF classique utilisées dans le cadre du diagnostic des machines tournantes ne permettent pas de suivre les évolutions des modes de fonctionnement du système à surveiller et donc les observations ainsi que les différentes classes caractérisant leurs différents états de fonctionnement. L'introduction de la théorie de la classification dynamique a permis de mettre en œuvre des nouvelles méthodes de suivi des machines tournantes. Nous développons dans le chapitre 3 des nouvelles procédures de classification dynamique pour le suivi de roulements. Nous passons d'abord en revue les différents travaux de recherche sur les processus de développement des systèmes de reconnaissance des formes dynamiques. En effet, des méthodes de classification dynamique ont été déjà développées dans le domaine de l'automatique/robotique et dans le domaine médical (analyse des signaux EEG, ECG, etc.). Les méthodes que nous avons mises en œuvre doivent répondre aux différentes contraintes imposées par les différents modes de fonctionnement de la machine tournante. La méthode de RdF dynamique doit s'adapter à la nature évolutive et non stationnaires des données vibratoires, intégrer toutes les informations disponibles même incomplètes dans son processus de décision.

Nous avons enfin validé les procédures de classification dynamique proposées sur un banc de roulements dans lequel nous avons placé successivement des roulements sains et défectueux avec des défauts de bague extérieure de différentes tailles. Nous avons comparé les résultats obtenus avec ceux obtenus par les méthodes de détection et de

suivi classiques. Cette validation a permis de voir l'évolution des différentes classes (classe "saine" et classe "défectueuse") en fonction de la charge et de la taille du défaut. Cette étape a permis de montrer le bien-fondé de la méthode que nous avons développée. Nous avons ensuite validé la méthode sur un banc de fatigue pour surveiller un roulement sous différentes conditions de fonctionnement.

Les perspectives envisageables au terme de ce travail sont divers ; d'autres paramètres comme les paramètres acoustiques, électriques peuvent être intégrés pour améliorer le diagnostic et le suivi d'une machine tournante. Une méthode de sélection de la méthode de projection pourrait être envisagée car cela permettrait d'utiliser plusieurs types de projection des matrices de données et de choisir la meilleure projection dans la boucle de la méthode RdF dynamique. Aussi, une étude approfondie sur l'évolution des classes ("saine" et "défectueuse") en fonction de changement des conditions de fonctionnement (variation de vitesse, etc) sera bien utile pour compléter l'étude menée dans ce travail. Les nouvelles procédures proposées dans ce projet devraient être intégrées dans un système d'aide à la décision dans un système embarqué pour le diagnostic et la sûreté de fonctionnement des structures à contact glissant (machines tournantes, systèmes mécaniques articulés, outil-pièce dans les procédés d'usinage à grande vitesse, robots) dont les retombées peuvent être importantes.

6. CONCLUSION GÉNÉRALE ET PERSPECTIVES

Annexe A

Annexes - Décomposition empirique modale pour le diagnostic de roulements

A.1 Décomposition modale empirique

Algorithm 15 Sifting process (SP)

- 1: **Procédure : algorithme SP**
 - 2: *Entrée* : signal $s(t)$.
 - 3: *Sorties* : moyenne $m(t)$, IMF $d(t)$
 - 4: **Initialisation**
 - 5: $d(t) \leftarrow s(t)$
 - 6: *Tant que* (le critère d'arrêt n'est pas satisfait)
 - 7: **Calculer**
 - 8: (E_{sup} et E_{inf}) $\triangleright$ les enveloppes supérieures et inférieures de $d(t)$, interpolant les maximas et les minimas respectivement
 - 9: **Effectuer**
 - 10:
$$d(t) \leftarrow d(t) - \left(\frac{E_{sup} + E_{inf}}{2} \right) \tag{A.1}$$
 - 11: **Poser**
 - 12: $d(t) \leftarrow (s(t) - d(t))$
 - 13: **Fin**
-

Algorithm 16 L'algorithme de EMD

Procédure *algorithme EMD*

2: *Entrée* : signal $s(t)$
Sorties : IMFs $d_1(t), d_2(t), \dots, d_N(t)$, résidu $r(t)$

4: **Initialisation** $r(t) \leftarrow s(t); i(t) \leftarrow 0$
Tant que ($r(t)$ possède plus de 3 extrêmes) $i \leftarrow i + 1$

6: **Calculer**
 $m(t) \leftarrow SP(r(t))$

8: **Poser**
 $d_i(t) \leftarrow (r(t) - m(t))$

10: $r(t) \leftarrow m(t)$

Renvoyer

12: les IMF $d_i(t)$
le résidu $r(t)$

14: **Fin**

On appelle Sifting Process (SP) l'opérateur qui consiste à soustraire à un signal sa moyenne locale, plusieurs fois de suite jusqu'à obtenir une moyenne (quasi) nulle. On note $m = SP(s)$. Cette définition de l'opérateur SP bien qu'elle est imprécise, elle est algorithmiquement simple. Pour construire $m = SP(s)$, on réalise l'algorithme décrits dans l'algorithme 15 :SP.

Le problème majeur posé par cet algorithme est celui de la convergence.

L'algorithme de l'EMD est relativement simple (voir l'algorithme 16 : EMD). En pratique, la procédure précédente, dite de « sifting », est généralement itérée sur le détail local pour affiner la décomposition.

A.2 La décomposition en valeurs singulières SVD des IMFs

La décomposition en valeurs singulières (*SVD*) est une méthode de traitement de données qui est souvent combinée avec les techniques de traitement du signal pour améliorer la détection en s'affranchissant des interférences de changements de condition de fonctionnement sur le diagnostic (GCL⁺14). la *SVD* peut être utilisée de différentes façons : elle peut être utilisée directement pour débruiter les signaux vibratoire (WT03) ou combinée avec une technique de traitement du signal comme : les ondelettes (YTYW11) pour remédier au problème de changement des conditions de fonctionnement ou encore avec

la décomposition modale empirique (WSJ⁺12, ZZB13). L'extraction des caractéristiques par la décomposition en valeurs singulières est une technique qui était récemment introduite dans le diagnostic des roulements, elle est généralement utilisée pour extraire des caractéristiques à partir des *IMFs* obtenus par *EEMD* ou *EMD* (CYY06).

La décomposition en valeurs singulières (*SVD*) est un outil permettant l'extraction des composantes principales d'une matrice. Dans le cas de signaux vibratoires, ces composantes principales sont liées à des structures de données maximisant l'énergie du signal et qu'on peut les extraire par la décomposition modale empirique *EMD*.

L'extraction des indicateurs issus de l'*EMD* par *SVD* commence par l'extraction des *IMFs* de chaque signal vibratoire suivi par la décomposition de la matrice des *IMFs* obtenus en valeurs singulières (*SVD*). Par conséquent, on obtient un vecteur de valeurs singulières $\lambda = [\lambda_1; \lambda_2; \dots \lambda_n]$. De ce vecteur, on peut ensuite calculer l'entropie de Shannon qui nous permettra de quantifier la quantité d'information dans chaque *IMF* (WSJ⁺12).

L'entropie de Shannon peut se calculer comme suit :

$$H_{TN} = - \sum_{i=1}^n q_i \log q_i \quad (\text{A.2})$$

Avec

$$q_i = \frac{\lambda_i}{\lambda} \quad (\text{A.3})$$

et

$$\lambda = \sum_{i=1}^n \lambda_i. \quad (\text{A.4})$$

n correspond au nombre de *IMFs* trouvés pour chaque signal.

ces caractéristiques non physiques issues de la *SVD* permettent d'isoler des phénomènes aléatoires de la distribution d'énergie dans les différentes *IMF*. La présence de défaut se manifeste principalement par la réduction de l'entropie calculé à partir des *IMFs*.

A.3 Indicateurs issus de la décomposition empirique modales

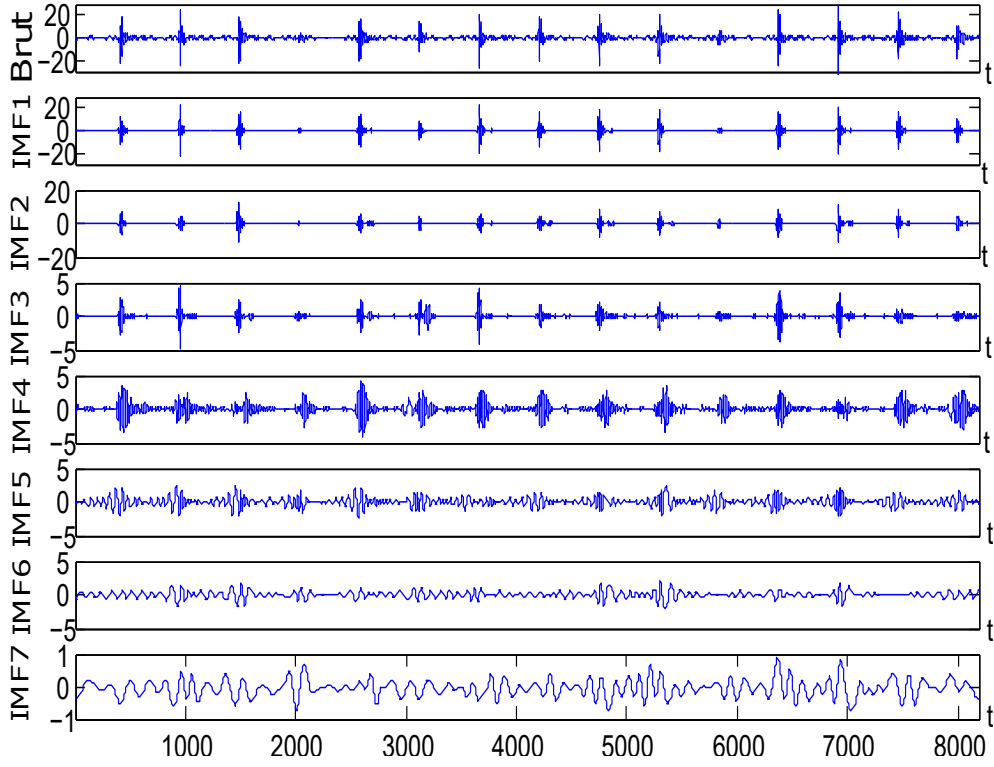


FIGURE A.1 — les Imfs issues pour un signal vibratoire d'un roulement avec un défaut de 2 mm^2

La décomposition modale empirique d'un signal vibratoire pour un signal d'un défaut se présente comme affiché sur la figure A.1.

Pour illustrer les avantages de cette méthode on l'a appliqué sur les même signaux que les sections précédentes. On a commencé par extraire les premiers *IMFs* de chaque signal. On a ensuite calculé les indicateurs temporels classiques à partir de ces *IMFs*.

L'objectif de la comparaison des changements des comportements de même indicateur extrait directement du signal brut ou à partir des trois premiers *IMFs* a été d'étudier l'hypothèse qui dit que la qualité de détection s'améliore en calculant le même indicateur à partir d'un signal décomposé par rapport à un signal brut (voir figure A.2).

Sur la figure A.2, on peut voir que l'indicateur *RMS* de *IMF1* a un peu près les même comportements que ceux extraits du signal brut. Le niveau de cet indicateur diminue quand il est extrait à partir des *IMF3* et *IMF2* (voir figures A.2a ??). En général,

A.3 Indicateurs issus de la décomposition empirique modales

l'extraction de ces indicateurs à partir du signal décomposé ne semble pas améliorer la séparabilité des états de santé, on peut même dire que l'extraction à partir du signal décomposé semble nuit à la qualité de séparabilité des différentes états de roulements (surtout pour les *IMFs* 2 et 3).

Cependant, la détection par les indicateurs Kurtosis et Facteur crête Fc semblent s'améliorer en appliquant la décomposition. Puisque pour les deux indicateurs, la valeur des indicateurs pour un roulement sain est assez différente de celle pour un défectueux et surtout s'ils sont extrait à partir de premier *IMF* (voir figures A.2b A.2c).

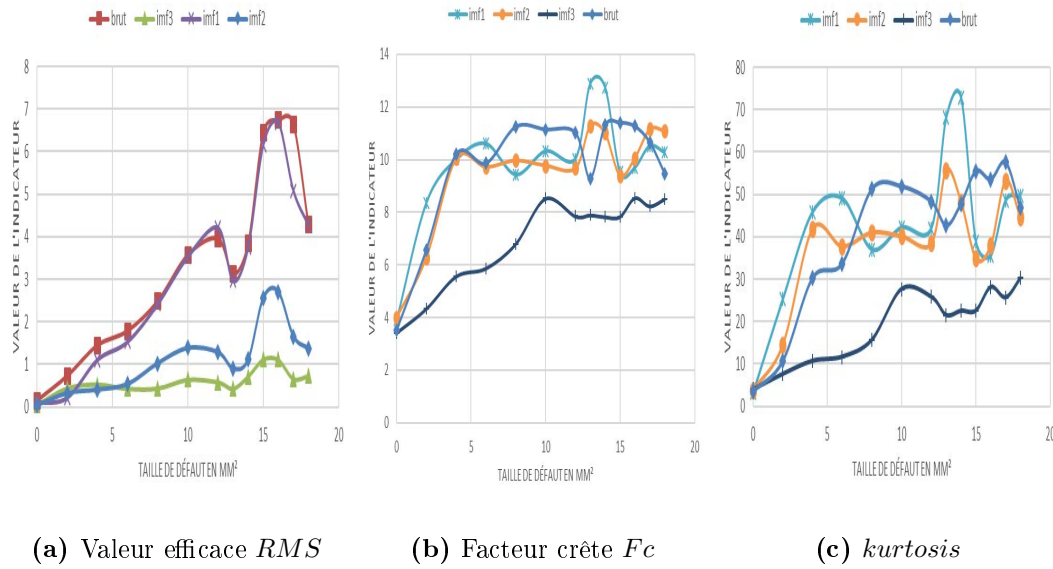


FIGURE A.2 – Illustration des évolutions des indicateurs extraits directement du signal brut et à partir des trois premiers *IMFs*

Deuxième étape consiste à étudier la qualité de détection de ces indicateurs mais cette fois en cas de variation de charge. Afin de révéler l'effet de celle-ci sur la détection de défauts, on a calculé les mêmes indicateurs extraits à partir de premier *IMFs* et pour des roulements sous différentes charges (voir A.3). Sur cette figure on peut voir que même en changeant la charge, l'indicateurs RMS manifeste le même comportement ce qui signifie que la charge n'affecte presque pas la détection de défaut. Quant aux indicateurs Fc et $kurtosis$, certes le changement de la charge affecte l'évolution de ces indicateurs mais

A. ANNEXES - DÉCOMPOSITION EMPIRIQUE MODALE POUR LE DIAGNOSTIC DE ROUEMENTS

n'affecte pas la capacité de détection de défaut de ces indicateurs.

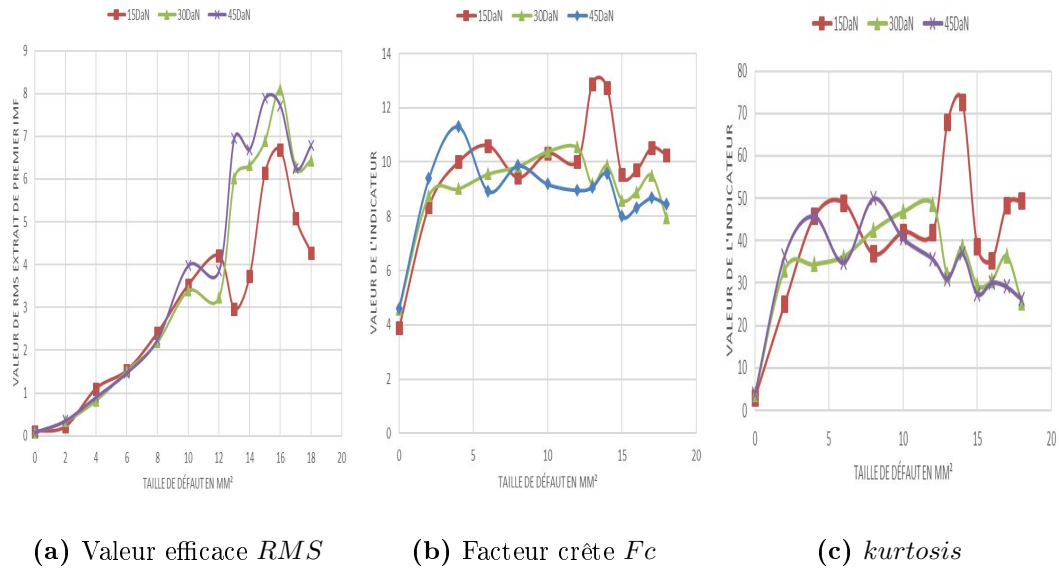


FIGURE A.3 – Illustration des évolutions des indicateurs extraits directement de premier IMF pour des roulements sous différentes charges

En conclusion l'extraction des indicateurs classiques à partir des $IMFs$ peut améliorer la détection de défaut surtout pour les deux indicateurs $Kurtosis$ et F_c et peut aussi apporter une certaine stabilité des indicateurs quand les conditions de fonctionnement ne sont pas stationnaires (l'évolution de RMS en cas de variation de charge est la même que dans le cas contraire).

Références

- [AE00] A.Hyvärinen and E.Oja. Independent component analysis : algorithms and applications. *Neural networks*, 13(4) :411–430, 2000. 42
- [AFN02] AFNOR. *Norme AFNOR X 60-000*. Paris Afnor, 2002. 8
- [AFS⁺15] J. Ben Ali, N. Fnaiech, L. Saidi, B. Chebel-Morello, and F. Fnaiech. Application of empirical mode decomposition and artificial neural network for automatic bearing fault diagnosis based on vibration signals. *Applied Acoustics*, 2015. 52
- [AH99] N.M. Adams and D.J. Hand. Comparing classifiers when the misallocation costs are uncertain. *Pattern Recognition*, 1999. 67
- [AH11] S.A Aye and P.S. Heyns. Effect of speed and torque on statistical parameters in tapered bearing fault detection. *World Academy of Science, Engineering and Technology*, 5, 2011. 20
- [AHA12] A.Moosavian, H.Ahmadi, and A.Tabatabaeefar. 814. 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), 2012. 34
- [AHP14] M.E. Ashalatha, M.S. Holi, and P.R.Mirajkar. Face recognition using local features by lpp approach. In *International Conference on Circuits, Communication, Control and Computing (I4C)*, pages 382–386, 2014. 42
- [AJ12] A.Bilosova and J.bilos. *Vibration diganostics*. Textbook Ostrava, 2012. 7

RÉFÉRENCES

- [AJJP03] C.C. Aggarwal, J.Han, J.Wang, and P.S.Yu. A framework for clustering evolving data streams. In *Proceedings of the 29th international conference on Very large data bases-Volume 29*, pages 81–92. VLDB Endowment, 2003. 112
- [ARFI07] S. Abbasion, A. Rafsanjani, A. Farshidianfar, and N. Irani. Rolling element bearings multi-fault classification based on the wavelet denoising and support vector machine. *Mechanical Systems and Signal Processing*, 2007. 39
- [ARM00] A.K.Jain, R.P.W.Duin, and M.Jianchang. Statistical pattern recognition : a review. *Transactions on Pattern Analysis and Machine Intelligence, IEEE*, 22(1) :4–37, 2000. 39
- [AS10] F. Ardjani and K. Sadouni. Optimization of svm multiclass by particle swarm (PSO-SVM). *Modern Education and Computer Science*, 2010. 34
- [ASK13] M. Amarnath, V. Sugumarnan, and H. Kumar. Exploiting sound signals for fault diagnosis of bearing using decision tree. *Measurment*, 2013. xi, 49, 50, 51, 59
- [Aug14] D. Augeix. Analyse vibratoire des machines tournantes. *Techniques de l'ingénieur Vibrations en milieu industriel, mesures, surveillance et contrôle*, TIB424DUO., 2014. 13
- [BAB03] B.Samanta and K. R. Al-Balushi. Artificial neural network based fault diagnostics of rolling element bearings using time-domain features. *Mechanical systems and signal processing*, 17(2) :317–328, 2003. 33
- [Bad99] M. Badaoui. *Contribution au Diagnostic Vibratoire des Réducteurs Complexes à Engrenages par l'Analyse Cepstrale*. PhD thesis, Université Jean Monnet, 1999. 23
- [BB01] N. Baydar and A. Ball. A comparative study of acoustic and vibration signals in detection of gear failures using wigner-ville distribution. *Mechanical Systems and Signal Processing*, 2001. 25

- [BBC⁺07] N. Basalto, R. Bellotti, F. De Carlo, P. Facchi, E. Pantaleo, and S. Pascazio. Hausdorff clustering of financial time series. *Physica A : Statistical Mechanics and its Applications*, 379(2) :635–644, 2007. 57
- [BBS14] C. Bouveyron and C. Brunet-Saumard. Discriminative variable selection for clustering with the sparse fisher-em algorithm. *Springer-Verlag*, 29, 2014. 43
- [BKS⁺10] K.M. Bhavaraju, P. K. Kankar, Satish, C. Sharma, and S. P. Harsha. A comparative study on bearings faults classification by artificial neural networks and self-organizing maps using wavelets. *International Journal of Engineering Science and Technology*, 2010. 41, 42
- [BM12] E.M. El Adel B. Mnassri, M. Ouladsine. Diagnostic de défauts par l’approche rbc ratio. In *Conférence Internationale Francophone d’Automatique*, Grenoble, 2012. 32
- [BMN89] A. E. Brouwer, A. M. Cohen, and Neumaier. Distance-regular graphs. *Springer-Verlag*, 1989. 64
- [Bod04] D. Bodin. *Modèle d’endommagement cyclique : Application à la fatigue des enrobés bitumineux*. PhD thesis, Université de Nantes, 2004. 12
- [Bon04] F. Bonnardot. *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. 31
- [Bou06] H.A. Boubacar. *classification dynamique de données non stationnaires apprentissage et suivi de classes évolutives*. PhD thesis, Université des sciences et technologies de Lille, 2006. 87
- [Bru09] R. Brunelli. *Template Matching Techniques in Computer Vision : Theory and Practice*. Wiley, 2009. 37

RÉFÉRENCES

- [BSM06] B. Waske, S. Schiefer, and M. Braun. Random feature selection for decision tree classification of multi-temporal sar data. In *IEEE International Conference on Geoscience and Remote Sensing Symposium (IGARSS) 2006*, pages 168–171, July 2006. 43
- [BY08] A. H. Bonnett and C. Yung. Increased efficiency versus increased reliability. *IEEE Industry Applications Magazine*, 14 :29–36, 2008. xi, 14, 15
- [C09] E. Côme. *Apprentissage des modèles génératifs pour le diagnostic de systèmes complexes avec labélisation douce et contraintes spatiales*. PhD thesis, Université de Technologie de Compiègne, 2009. 32
- [CBBL89] G. C. Collins, J. R. Bourne, A. J. Brodersen, and C. F. Lo. Comparison of rule-based and belief-based systems for diagnostic problems. In *Proceedings of the 2nd International Conference on Industrial and Engineering Applications of Artificial Intelligence and Expert Systems*, volume 2, pages 785–793, 1989. 33
- [Ceb12] A. Ceban. *Apprentissage des modèles génératifs pour le diagnostic de systèmes complexes avec labélisation douce et contraintes spatiales*. PhD thesis, Université Lille de Nord, 2012. 12, 32
- [Chi07] X. Chimentin. *localisation et quantification des sources dans le cadre d’une maintenance préventive conditionnelle en vue de fiabiliser le diagnostic et le suivi de l’endommagement des composants mécanique tournants : application aux roulements à billes*. PhD thesis, Université de Reims Champagne Ardenne, 2007. 12, 14, 27
- [CL11] Z. Chen and Y.F. Li. Anomaly detection based on enhanced dbscan algorithm. *Procedia Engineering*, 2011. 64
- [CMGT02] C. Byington, M. Roemer, G. Kacprzynski, and T. Galie. Prognostic enhancements to diagnostic systems for improved condition-based maintenance. In *Aerospace Conf.*, Big Sky, USA, 2002. 12
- [CN04] J. Confais and J.P. Nakache. *Arbres hiérarchiques, partitionnements*. Editions Technip, 2004. 64

- [CY02] M. Chavent and Y. Lechevallier. Dynamical clustering of interval data : Optimization of an adequacy criterion based on hausdorff distance. In *Classification, clustering, and data analysis*, pages 53–60. Springer, 2002. 57
- [CYY06] C. Junsheng, Y. Dejie, and Y. Yu. A fault diagnosis approach for roller bearings based on emd method and ar model. *Mechanical Systems and Signal Processing*, 20(2) :350–362, 2006. 215
- [Den06] L. E. Denaud. *Analyses vibratoires et acoustiques du déroulage*. PhD thesis, l’Ecole Nationale Supérieure d’Arts et Métiers, 2006. 24, 25
- [Des10] M. Desbazeille. *Diagnostic de groupes électrogènes diesel par analyse de la vitesse de rotation du vilebrequin*. PhD thesis, Université Jean Monnet-Saint-Etienne, 2010. 12
- [DP05] P. D. Samuel and D. J. Pines. A review of vibration-based techniques for helicopter transmission diagnostics. *Journal of Sound and Vibration*, 282 :475–508, 2005. 20
- [DRP⁺15] D. D. Doana, E. Ramassoa, V. Placetb, S. Zhangb, L. Boubakarb, and N. Zerhouni. An unsupervised pattern recognition approach for ae data originating from fatigue tests on polymer-composite materials. *Mechanical Systems and Signal Processing*, 64, 2015. 36
- [DYHR11] W. Dong, J. Yi, H. He, and J. Ren. An incremental algorithm for frequent pattern mining based on bit-sequence. *International Journal of Advancements in Computing Technology*, 2011. 83
- [DZ12] J. Dybala and R. Zimroz. Application of empirical mode decomposition for impulsive signal extraction to detect bearing damage : industrial case study. *Springer*, 2012. 30
- [EG14] A. Erdem and T. I. Gundem. M-fdbscan : A multicore density-based uncertain data clustering algorithm. *Turkish Journal of Electrical Engineering and Computer Science*, 22(1) :143–154, 2014. 122

RÉFÉRENCES

- [EKSX96] M. Ester, H.-P. Kriegel, J. Sander, and X. Xu. 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, 1996. xi, 63, 65
- [Elg07] H. Elghazel. *Classification et Prévission des Données Hétérogènes :Application aux Trajectoires et Séjours Hospitaliers*. PhD thesis, Université Claude Bernard Lyon 1, 2007. 69
- [EP08] Stephan Ebersbach and Zhongxiao Peng. Expert system development for vibration analysis in machine condition monitoring. *Expert systems with applications*, 34(1) :291–299, 2008. 33
- [Est04] P. Estocq. *Une approche méthodologique numérique et expérimentale d’aide à la détection et au suivi vibratoire de défauts d’écailage de roulements à billes*. PhD thesis, Université de Reims champagne Ardenne, 2004. 17
- [EZS⁺98] N. E.Huang, Z.Shen, S.R.Long, M.C. Wu, H. H.Shih, Q.Zheng, N.C.Yen, C.C. Tung, and H.H.Liu. 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*, 1998. 22, 29
- [FG04] P. Flandrin and P. Goncalvès. Empirical mode decomposition as data-driven wavelet-like expansions. *International Journal of Wavelets, Multiresolution and Information Processing*, 2004. 29, 41
- [FL02] Yimin Fan and C James Li. Diagnostic rule extraction from trained feedforward neural networks. *Mechanical Systems and Signal Processing*, 16(6) :1073–1081, 2002. 33
- [FMA⁺10] K. Förster, S. Monteleone, A.Calatroni, D. Roggen, and G. Tröster. Incremental knn classifier exploiting correct - error teacher for activity recognition. *IEEE Computer Society*, 2010. 83
- [FQG00] X. Fengfeng, S. Qiao, and K. Govindappa. Bearing Diagnostics Based on Pattern Recognition of Statistical Parameters. *Journal of Vibration and Control*, 6 :375–392, 2000. 20

- [FR90] P. Flandrin and O. Rioul. Affine smoothing of the wigner-ville distribution. *IEEE*, 1990. 26
- [Fra93] Ph. P. Framatone. Diagnostic, le facteur défaut pour la surveillance des roulements. *Maintenance & entreprise*, 1993. 18
- [Fuk90] K. Fukunaga. *Introduction to statistical pattern recognition*. Academic Press, 1990. 38
- [Gan13] E. Gandino. *Diagnostics of machines and structures : dynamic identification and damage detection*. PhD thesis, Politecnico di Torino, 2013. 7, 13
- [GCL⁺14] G.Ruocci, G. Cumunel, T. Le, P. Argoul, N. Point, and L. Dieng. Damage assessment of pre-stressed structures : A svd-based approach to deal with time-varying loading. *Mechanical Systems and Signal Processing*, 47(1) :50–65, 2014. 214
- [GCLL10] M. Ghribi, P. Cuxac, J.C. Lamirel, and A. Lelu. 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
- [GD14] G.Bordogna and D.Ienco. Fuzzy core dbscan clustering algorithm. In *Information Processing and Management of Uncertainty in Knowledge-Based Systems*, pages 100–109. Springer, 2014. 122
- [G.G03] G.Govaert. *Analyse des données (Traité IC2, Série Traitement du signal et de l'image)*. Hermes Science, 2003. 45
- [Gha03] Z. Ghahramani. *Advanced Lectures on Machine Learning*. Springer-Verlag, 2003. 36, 38
- [Gho08] M. Ghozzi. *Détection cyclostationnaire des bandes de fréquences libres*. PhD thesis, Université de Rennes I, 2008. 31
- [GS94] W. A. Gardner and C.M. Spooner. The cumulant theory of cyclostationary time-series, part i : foundations. *IEEE Trans. on Signal Processing*, 1994. 31

RÉFÉRENCES

- [GZS01] S.K. Goumas, M.E. Zervakis, and G.S. Stavrakakis. Intelligent on-line quality control of washing machines using discrete wavelet analysis features and likelihood classification. *Engineering Applications of Artificial Intelligence*, 2001. 32
- [Har10] L. Hartert. *Reconnaissance des formes dans un environnement dynamique appliquée au diagnostic et au suivi des systèmes évolutifs*. PhD thesis, Université de Reims Champagne Ardenne, 2010. 83, 84, 85
- [HAWF05] H.Sohn, D.W. Allen, K. Worden, and C. R. Farrar. Structural damage classification using extreme value statistics. *Journal of Dynamic Systems, Measurement, and Control*, 2005. 41
- [HDSIW06] Huang, De-Shuang, Irwin, and G. William. *Intelligent Computing in Signal Processing and Pattern Recognition*. Springer-Verlag Berlin Heidelberg, 2006. 37
- [HK96] H.Frigui and R. Krishnapuram. A robust algorithm for automatic extraction of an unknown number of clusters from noisy data. *Pattern Recognition Letters*, 17(12) :1223–1232, 1996. 133
- [HNZ12] R. Hunt, K. Neshatian, and M. Zhang. Simulated evolution and learning. *Springer-Verlag*, 7673, 2012. 43
- [HON14] X. HONGTAO. *Study on Intelligent Condition Diagnosis Based on Vibration Information and Support Vector Machine for Plant Machinery*. PhD thesis, Graduate School of Bioresources Mie University, 2014. 9
- [HTF09] T. Hastie, R. Tibshirani, and J. Friedman. *The Elements of Statistical Learning :Data Mining, Inference, and Prediction*. Springer, 2009. 43, 47
- [Ibr09] A. Ibrahim. *Contribution au diagnostic de machines électromécaniques : Exploitation des signaux électriques et de la vitesse instantanée*. PhD thesis, Université de Saint Etienne, 2009. 12, 31
- [JATH12] J.R.Gordon, N.M. Adams, D. K. Tasoulis, and D.J. Hand. Exponentially weighted moving average charts for detecting concept drift. *Pattern Recognition Letters*, 33, 2012. 41

- [JCSQ11] Z. Jiusun, G. Chuanhou, L. Shihua, and L. Qihui. Online process monitoring based on incremental lpp. In *Control Conference (CCC), 2011 30th Chinese*, pages 4200–4204. IEEE, 2011. 100
- [JCYD13] L. Jiang, Y. Cao, H. Yin, and K. Deng. An improved kernel k-mean cluster method and its application in fault diagnosis of roller bearing. *Engineering*, 2013. 62
- [JDJ00] A.K. Jain, R.P.W. Duin, and J.Mao. Statistical pattern recognition : A review. *IEEE*, 2000. 38, 39, 42
- [JGDE08] A.G.K. Janecek, W. N. Gansterer, M.A. Demeland, and G. F. Ecker. On the relationship between feature selection and classification accuracy. In *New challenges for feature selection in data mining and knowledge discovery*, Belgium, 2008. 42
- [JKP04] E. Januzaj, H.P. Kriegel, and M. Pfeifle. Scalable density-based distributed clustering. In *Knowledge Discovery in Databases : PKDD*, pages 231–244. Springer, 2004. 112
- [JLB06] A.K.S. Jardine, D. Lin, and D. Banjevic. A review on machinery diagnostics and prognostics implementing condition-based maintenance. *Mechanical Systems and Signal Processing*, 2006. 12, 24, 25, 33
- [JWM08] P. Jayaswal, A.K. Wadhvani, and K.B. Mulchandani. Machine fault signature analysis. *International Journal of Rotating Machinery*, 2008. 9
- [Khe14] I. Khelf. *Diagnostic des machines tournantes par les tchniques de l'intelligence artificielle*. PhD thesis, Université de Badj Mokhtar-Annaba, 2014. 33
- [Kho09] M. Khov. *Surveillance et diagnostic des machines synchrones à aimants permanents : Détection des courts-circuits par suivi paramétrique*. PhD thesis, Institut National Polytechnique de Toulouse, 2009. 12
- [KM05] H.P. Kriegel and M.Pfeifle. Density-based clustering of uncertain data. In *Proceedings of the eleventh ACM SIGKDD international conference on Knowledge discovery in data mining*, pages 672–677. ACM, 2005. 122

RÉFÉRENCES

- [Kou10] M. Koujok. *Contribution au pronostic industriel : intégration de la confiance à un modèle prédictif neuro-flou*. PhD thesis, Université Franche-Comté, 2010. 12
- [KSI⁺14] N. Karabadji, H. Seridi, I.Khelf, N. Azizi, and R. Boulkroune. Improved decision tree construction based on attribute selection and data sampling for fault diagnosis in rotating machines. *Engineering Applications of Artificial Intelligence*, 2014. 49, 50, 59
- [KSSV13] H. S. Kumar, P.P. Srinivasa, N.S. Sriram, and G.S. Vijay. ANN based evaluation of performance of wavelet transform for condition monitoring of rolling element bearing. *Procedia Engineering*, 2013. 53
- [KTMY06] Y.H. Kim, A.C.C. Tan, J. Mathew, and B.S. Yang. 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, 2006. 20
- [KXL13] S. Kerroumi, X.Chimentin, and L.Rasolofondraibe. Dynamic classification method of fault indicators for bearings’ monitoring. *Mechanics & Industry*, 14(2) :115–120, 2013. 108
- [Lar12] D.T. Larose. *Exploration de données*. Vuibert, 2012. 8
- [LDD07] H. Li, X. Deng, and H. Dai. Structural damage detection using the combination method of emd and wavelet analysis. *Mechanical Systems and Signal Processing*, 2007. 30
- [LGK06] P. Laskov, C. Gehl, and S. Krüger. Incremental support vector learning : Analysis, implementation and applications. *Journal of Machine Learning Research*, 2006. 83
- [LI12] R. LI. *Rotating Machine Fault Diagnostics Using Vibration and Acoustic Emission Sensors*. PhD thesis, University of Illinois, 2012. 25
- [LL03] J.H. Lee and Y.T. Lee. Robust technique for estimating the bearings of cyclostationary signals. *Signal Processing*, 2003. 31

- [LM97] J. Li and C. Ma. Wavelet decomposition of vibrations for detection of bearing localized defects. *NDT & E International*, 1997. 26
- [LMF79] L. Lebart, A. Morineau, and J.P. Fenelon. *Traitement des données statistiques*. Dunod, 1979. 68
- [LMSZ10] H. Liu, H. Motoda, R. Setiono, and Z. Zhao. Feature selection : An ever evolving frontier in data mining. In *The Fourth Workshop on Feature Selection in Data Mining*, India, 2010. 42
- [LP92] C. Q. Li and C. J. D. Pickering. Robustness and sensitivity of nondimensional amplitude parameters for diagnosis of fatigue spalling. In *Condition Monitoring and Diagnostic Technology*, 1992. 20, 142
- [LSN96] T.I. Liu, J.H. Singonahalli, and N.R.Iyer. Detection of roller bearing defects using expert system and fuzzy logic. *Mechanical Systems and Signal Processing*, 10(5) :595–614, 1996. 33
- [Lur03] C. Lurette. *Developpement d’une technique neuronale auto-adaptative pour la classification dynamique de données évolutives application à la supervision de la presse hydraulique*. PhD thesis, Université des sciences et technologies de Lille, 2003. 84
- [LWLC05] W. Li, W.Gong, Y. Liang, and W. Chen. *Pattern recognition and image analysis*. Springer, 2005. 36, 42
- [LYT02] L.Ma, Y.Wang, and T.Tan. Iris recognition based on multichannel gabor filtering. In *Proc. Fifth Asian Conf. Computer Vision*, volume 1, pages 279–283, 2002. 40
- [LZW07] P. Liu, D. Zhou, and N. Wu. Vdbscan :varied density based spatial clustering of applications with noise. In *International Conference on Service Systems and Service Management*, Chengdu, China, 2007. 64
- [MCP02] P. Mitra, C.Murthy, and S.K. Pal. Unsupervised feature selection using feature similarity. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 24(3) :301–312, 2002. 102

RÉFÉRENCES

- [ME10] M. S. Mouchaweh and E.Lughofer. *Learning in Non-Stationary Environments*. Springer, 2010. 73
- [MEA14] M. Ghaemi M. Ettefagh and M. Yazdanian Asr. Bearing fault diagnosis using hybrid genetic algorithm k-means clustering. In *Innovations in intelligent systems and applications*, Alberobello, 2014. 32, 58, 62
- [MMZ08] K. Medjaher, A. Mechraoui, and N. Zerhouni. Diagnostic et pronostic de défaillances par réseaux bayésiens. In *Journées Francophone sur les Réseaux Bayésiens*, Lyon, France, 2008. 12
- [Mon03] F. Monchy. *Maintenance, méthodes et organisations*. Dunod, 2003. 8
- [Mor91] J. Morel. *Vibrations des machines et diagnostic de leur état mécanique*. Eyrolles, 1991. 11
- [Mor05] J. Morel. Surveillance vibratoire et maintenance prédictive. *Techniques de l'ingénieur Comportement en service des systèmes et composants mécaniques*, TIB180DUO, 2005. xi, 17, 19
- [Mou02] M. S. Mouchaweh. *Conception d'un système de diagnostic adaptatif et prédictif basé sur la méthode Fuzzy Pattern Matching pour la surveillance en ligne des systèmes évolutifs*. PhD thesis, Université de Reims Champagne Ardenne, 2002. 84
- [MP03] M.Belkin and P.Niyogi. Laplacian eigenmaps for dimensionality reduction and data representation. *Neural computation*, 15(6) :1373–1396, 2003. 98
- [MS03a] M. Markou and S. Singh. Novelty detection : a review part 1 : statistical approaches. *Signal Processing*, 2003. 88
- [MS03b] M. Markou and S. Singh. Novelty detection : a review part 2 : neural network based approaches. *Signal Processing*, 2003. 88
- [Mu.01] Mu.Zhu. *Feature Extraction and Dimension Reduction with Applications to Classification and the Analysis of Co-occurrence Data*. PhD thesis, Stanford University, 2001. 97

- [MWM11] Q. Miao, D. Wang, and M. Pecht. Rolling element bearing fault feature extraction using emd-based independent component analysis. *IEEE, Prognostics and Health Management*, June 2011. 31
- [NA02a] N. Nikolaou and I. Antoniadis. Demodulation of vibration signals generated by defects in rolling element bearings using complex shifted morlet wavelets. *Mechanical Systems and Signal Processing*, 2002. 27
- [NA02b] N.G. Nikolaou and I.A. Antoniadis. Rolling element bearing fault diagnosis using wavelet packets. *NDT & E International*, 2002. 27
- [NG05] E.N. Nasibov and G. Ulutagay. A new approach to the problem of clustering using the fuzzy joint points method. *Automatic Control and Computer Sciences*, 39(6) :8–17, 2005. 123
- [Niy04] X. Niyogi. Locality preserving projections. In *Neural information processing systems*, volume 16, page 153. MIT, 2004. 98
- [NT97] N. Kambhatla and T. Leen. Dimension reduction by local principal component analysis. *Neural Computation*, 9(7) :1493–1516, 1997. 42
- [NU07] E.N. Nasibov and G. Ulutagay. A new unsupervised approach for fuzzy clustering. *Fuzzy Sets and Systems*, 158(19) :2118–2133, 2007. 123
- [NU09] E. N. Nasibov and G. Ulutagay. Robustness of density-based clustering methods with various neighborhood relations. *Fuzzy Sets and Systems*, 160(24) :3601–3615, 2009. 122, 123
- [OKK98] O. Yamaguchi, K. Fukui, and K. Maeda. Face recognition using temporal image sequence. In *Third IEEE International Conference on Automatic Face and Gesture Recognition*, pages 318–323. IEEE, 1998. 40
- [Ond06] O. Ondel. *Diagnostic par reconnaissance des formes : Application à une ensemble convertisseur, machine asynchrone*. PhD thesis, L’Ecole centrale de Lyon, 2006. 44, 84
- [Oul14] A. Oulmane. *Surveillance et diagnostic des défauts des machines tournantes dans le domaine temps-fréquences utilisant les réseaux de neurones et la logique floue*. PhD thesis, École Polytechnique de Montréal, 2014. xi, 24

RÉFÉRENCES

- [PBG06] A. Parey, M. El Badaoui, F. Guillet, and N. Tandon. Dynamic modelling 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*, 2006. 30
- [PC04] Z.K. Peng and F.L. Chu. Application of the wavelet transform in machine condition monitoring and fault diagnostics : a review with bibliography. *Mechanical Systems and Signal Processing*, 2004. 27
- [P.H07] P.Honeine. *Méthodes à noyau pour l'analyse et la décision en environnement non-stationnaire*. PhD thesis, Troyes, 2007. 100
- [PHK10] J.K. Parker, L.O. Hall, and A. Kandel. Scalable fuzzy neighborhood dbscan. In *International Conference on Fuzzy Systems (FUZZ)*, pages 1–8. IEEE, 2010. 122
- [PJ84] P.McFadden and J.Smith. Vibration monitoring of rolling element bearings by the highfrequency resonance technique : a review. *Mechanical Systems and Signal Processing*, 1984. 27
- [Ran01] R. B. Randall. The relationship between spectral correlation and envelope analysis in the diagnostics of bearing faults and other cyclostationary machine signals. *Mechanical Systems and Signal Processing*, 2001. 23, 31
- [R.B11] R.B.Randall. *Vibration-based Condition Monitoring : Industrial, Aerospace and Automotive Applications*. Wiley, 2011. 22, 31
- [RHS96] M.J. Roemer, C.A Hong, and S.H.Hesler. Machine health monitoring and life management using finite-element-based neural networks. *Journal of engineering for gas turbines and power*, 118(4) :830–835, 1996. 33
- [Rou87] P. J. Rousseeuw. Silhouettes : A graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics*, 1987. 69
- [RU01] Rubini R. and Meneghetti U. Application of the envelope and wavelet transform analyses for the diagnosis of incipient faults in ball bearings. *Mechanical Systems and Signal Processing*, 2001. 22, 27

- [SAFN07] S.Abbasion, A.Rafsanjani, A. Farshidianfar, and N.Irani. Rolling element bearings multi-fault classification based on the wavelet denoising and support vector machine. *Mechanical Systems and Signal Processing*, 21(7) :2933–2945, 2007. 26
- [SAN⁺03] S.Guha, A.Meyerson, N.Mishra, R.Motwani, and L.O’Callaghan. Clustering data streams : Theory and practice. *IEEE Transactions on Knowledge and Data Engineering*, 15(3) :515–528, 2003. 85
- [SC95] S.C.Pei and C.N.Lin. Image normalization for pattern recognition. *Image and Vision computing*, 13(10) :711–723, 1995. 40
- [SCR03] S.Simani, C.Fantuzzi, and R.J.Patton. *Model-based Fault Diagnosis in Dynamic Systems Using Identification Techniques*. Springer, 2003. 33
- [SCZX04] Q. Sun, P. Chen, D. Zhang, and F. Xi. Pattern recognition for automatic machinery fault diagnosis. In *ASME*, volume 126, Kunming, China, 2004. 20
- [SDL10] F. Song, D.Mei, and H. Li. Feature selection based on linear discriminant analysis. In *International Conference on Intelligent System Design and Engineering Application (ISDEA)*, volume 1, pages 746–749, 2010. 42
- [SG94] C.M. Spooner and W. A. Gardner. The cumulant theory of cyclostationary time-series, part ii : development and applications. *IEEE Trans. on Signal Processing*, 1994. 31
- [SJ01] C. Saint-Jean. *Classification paramétrique robuste partiellement supervisée en reconnaissance des formes*. PhD thesis, Université de La Rochelle, 2001. 39
- [SKTR15] T. Segretoa, S. Karama, R. Tetia, and J. Ramsing. Feature extraction and pattern recognition in acoustic emission monitoring of robot assisted polishing. In *Procedia CIRP*, 2015. 41
- [SL11] C. Salperwyck and V. Lemaire. Classification incrémentale supervisée : un panel introductif. *Revue des Nouvelles Technologies de l’Information, Apprentissage Artificiel et Fouille de Données* :121–148, 2011. 82

RÉFÉRENCES

- [SLM05] M.S. Safizadeh, A.A. Lakis, and M.Thomas. Using short-time fourier transform in machinery diagnosis. In *Proceedings of the fourth WSEAS international conference on electronic, signal processing and control*, volume 20, pages 1–7. World Scientific and Engineering Academy and Society (WSEAS) Stevens Point, Wisconsin, USA, 2005. 25
- [SRSS11] M. Saimurugana, K.I. Ramachandran, V. Sugumaran, and N.R. Sakthivel. Multi component fault diagnosis of rotational mechanical system based on decision tree and support vector machine. *Expert Systems with Applications*, 2011. 51, 59
- [ST13] S.Dong and T.Luo. Bearing degradation process prediction based on the pca and optimized ls-svm model. *Measurement*, 46(9) :3143–3152, 2013. 34
- [S.W85] S.Watanabe. *Pattern recognition : human and mechanical*. Wiley, 1985. 36, 37, 38
- [SWT02] W.J. Staszewski, K. Worden, and G.R. Tomlinson. Time-frequency analysis in greabox fault detection using wigner-ville distribution and pattern recognition. *MSSP*, 2002. 32, 41
- [SXL14] S.Kerroumi, X.Chimentin, and L.Rasolofondraibe. Online classification for spalling detection and vibratory behavior monitoring. *Mechanics & Industry*, 15(6) :517–524, 2014. 110
- [SZ13] A. Smiti and Z.Eloudi. Soft dbscan : Improving dbscan clustering method using fuzzy set theory. In *The 6th International Conference on Human System Interaction (HSI)*, pages 380–385. IEEE, 2013. 64, 122
- [Tal05] L. Talavera. *Advances in intelligent data analysis VI*. Springer, 2005. 43
- [TC99] N. Tandon and A. Choudhury. A review of vibration and acoustic measurement methods for the detection of defects in rolling element bearings. *Tribology International*, 1999. 16
- [TD60] T.Marill and D.M.Green. Statistical recognition functions and the design of pattern recognizers. *Electronic Computers, IRE Transactions on*, EC-9(4) :472–477, 1960. 36

- [TM01] W.T. Thomson and M.Fenger. Current signature analysis to detect induction motor faults. *IEEE Ind. Applications Magazine*, 2001. xi, 14, 15
- [TSK05] P.N Tan, M. Steinbach, and V. Kumar. *Introduction to Data Mining*. Addison Wesley, 2005. 36, 62
- [Tuf12] S. Tuffery. *Data mining et statistique décisionnelle*. Editions technip, 4 edition, 2012. 62, 64
- [UN12] G. Ulutagaya and E. Nasibovb. Fuzzy and crisp clustering methods based on the neighborhood concept : A comprehensive. *Journal of Intelligent & Fuzzy Systems*, 23 :1–11, 2012. 122
- [VR02] V.Ganti and R.Ramakrishnan. *Mining and monitoring evolving data*. Springer, 2002. 112
- [V.V95] V.Vapnik. *The nature of statistical learning theory*. Springer Verlag, 1995. 47
- [W85] Z. M. Wójcik. A natural approach in image processing and pattern recognition : Rotating neighbourhood technique, self-adapting threshold, segmentation and shape recognition. *Pattern Recognition*, 18, 1985. 36
- [WB13] Q. Wang and K.L. Boyer. Feature learning by multidimensional scaling and its applications in object recognition. In *26th SIBGRAPI - Conference on Graphics, Patterns and Images (SIBGRAPI)*, pages 8–15, 2013. 42
- [WDB04] K. Worden and J. M. Dulieu-Barton. An overview of intelligent fault detection in systems and structures. *Structural Health Monitoring*, 3(1), 2004. 13
- [WJCM03] Z. Wen, J.Crossman, J. Cardillo, and Y.L. Murphey. Case-base reasoning in vehicle fault diagnostics. In *Neural Networks, 2003. Proceedings of the International Joint Conference on*, volume 4, 2003. 33
- [WJN06] M.L.D. Wong, L.B. Jack, and A.K. Nandi. Modified self-organising map for automated novelty detection applied to vibration signal monitoring. *Mechanical Systems and Signal Processing*, 20, 2006. 88

RÉFÉRENCES

- [WKYS12] Y. Wang, S. Kang, G. Yang, and L. Song. 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 SignalProcessing*, 2012. 32
- [WSJ⁺12] Y. Wang, S.Kang, Y. Jiang, G. Yang, L. Song, and V.I. Mikulovich. 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*, 2012. 41, 47, 215
- [WT03] W.X.Yang and W. T.Peter. Development of an advanced noise reduction method for vibration analysis based on singular value decomposition. *Ndt & E International*, 36(6) :419–432, 2003. 214
- [WWTW13] S.D. Wu, C.W. Wu, T.Y.Wu, and C.C. Wang. Multi-scale analysis based ball bearing defect diagnostics using mahalanobis distance and support vector machine. *Entropy*, 2013. 58
- [XYC88] L. Xu, P. Yan, and T. Chang. Best first strategy for feature selection. In *9th International Conference on Pattern Recognition*, volume 2, pages 706–708, Nov 1988. 43
- [YGA11] C.T. Yiakopoulos, K.C. Gryllias, and I.A. Antoniadis. Rolling element bearing fault detection in industrial environments based on a K-means clustering approach. *Expert Systems with Applications*, 2011. 32, 62
- [YGD07] S. Yella, N. K. Gupta, and M. S. Dougherty. Comparison of pattern recognition techniques for the classification of impact acoustic emissions. *Transportation Research Part C : Emerging Technologies*, 15, 2007. 36
- [YHB13] Y.Zhang, Z. Hongfu, and B.Fang. Classification of fault location and performance degradation of a roller bearing. *Measurement*, 2013. 21, 31, 32
- [YTYW11] Y.Jiang, B. Tang, Y.Qin, and W.Liu. Feature extraction method of wind turbine based on adaptive morlet wavelet and svd. *Renewable energy*, 36(8) :2146–2153, 2011. 214

- [Yu11] J.B. Yu. Bearing performance degradation assessment using locality preserving projections. *Expert Systems with Applications*, 38(6) :7440–7450, 2011. 100
- [YYC06] Y. Yu, , YuDejie, and C.Junsheng. A roller bearing fault diagnosis method based on EMD energy entropy and ann. *Journal of Sound and Vibration*, 2006. 53
- [YYJ06] Y.Yang, YuDejie, and C. Junsheng. A roller bearing fault diagnosis method based on emd energy entropy and ann. *Journal of Sound and Vibration*, 2006. 30, 31
- [YZYX08] Y.Lei, Z.He, Y.Zi, and X.Chen. New clustering algorithm-based fault diagnosis using compensation distance evaluation technique. *Mechanical Systems and Signal Processing*, 22(2) :419–435, 2008. 21
- [ZLJY15] X. Zhang, Y. Liang, J.Zhou, and Y.Zang. A novel bearing fault diagnosis model integrated permutation entropy, ensemble empirical mode decomposition and optimized SVM. *Measurement*, 69 :164–179, 2015. 31
- [ZSY09] X. Zhu, L. Shen, and T.S.P. Yum. Hausdorff clustering and minimum energy routing for wireless sensor networks. *Transactions on Vehicular Technology IEEE*, 58(2) :990–997, 2009. 57
- [ZYZH06] J. Zhang, Q. Yan, Y. Zhang, and Z. Huang. 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, 2006. 88
- [ZZ13a] X. Zhang and J. Zhou. Multi-fault diagnosis for rolling element bearings based on ensemble empirical mode decomposition and optimized support vector machines. *Mechanical Systems and Signal Processing*, 2013. 32
- [ZZ13b] X. Zhang and J. Zhou. Multi-fault diagnosis for rolling element bearings based on ensemble empirical mode decomposition and optimized support vector machines. *Mechanical Systems and Signal Processing*, 2013. 48

RÉFÉRENCES

- [ZZB13] Y. Zhang, H. Zuo, and F. Bai. Classification of fault location and performance degradation of a roller bearing. *Measurement*, 2013. 42, 47, 48, 215

EXTRACTION DES PARAMETRES ET CLASSIFICATION DYNAMIQUE DANS LE CADRE DE LA DETECTION ET DU SUIVI DE DEFAUT DE ROULEMENTS

Parmi les techniques utilisées en maintenance, l'analyse vibratoire reste l'outil le plus efficace pour surveiller l'état interne des machines tournantes en fonctionnement. En et 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. 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 développé trois méthodes de classification dynamique : *DynamicDBSCAN* qui a été développée pour suivre les évolutions des classes, *Evolving scalable DBSCAN* qui représente une version en ligne et évolutive de *DBSCAN* et finalement *Dynamic Fuzzy Scalable DBSCAN* qui est une version dynamique et floue *ESDBSCAN* adaptée pour un apprentissage en ligne. Ces méthodes distinguent les variations des observations liées au changement du mode de fonctionnement de la machine (variation de charges) et les variations liées au défaut.

Ainsi, Elles permettent de détecter, de façon précoce, l'apparition de défaut. L'application sur des données réelles a permis d'identifier les différents états du roulement au cours du temps (sain ou normal, défectueux) et l'évolution des observations liée à la variation de charges avec un taux d'erreur faible et d'établir un diagnostic fiable.

Diagnostic et suivi, Roulements, Méthodes de reconnaissance de formes, Analyse vibratoire, Apprentissage en ligne, Classification dynamique.

EXTRACTION OF NEW FEATURES AND INTEGRATION OF DYNAMIC CLASSIFICATION TO IMPROVE BEARING FAULT MONITORING

Vibration analysis remains the most popular and most effective tool for monitoring the internal state of an operating machine.

Through vibration analysis, fault indicators are extracted then monitored to detect/ locate the presence of a defect.

However, counting solely on the evolution of these fault indicators to diagnose a machine can cause false alarms and question the reliability of the diagnosis.

In this thesis, we combined vibration analysis tools with pattern recognition method to firstly improve fault detection reliability of bearings, secondly to assess the severity of degradation by closely monitor the defect growth and finally to estimate their remaining useful life.

For these reasons, we have designed a pattern recognition process capable of; identifying defect even in machines running under non stationary conditions, processing evolving data and can handle an online learning.

Three dynamic classification methods have been developed: *Dynamic DBSCAN* was developed to capitalize on the time evolution of the data and their classes, *Evolving Scalable DBSCAN* that was created to include the online processing and finally *Dynamic Fuzzy Scalable DBSCAN* a dynamic fuzzy and semi-supervised version of *ESDBSCAN*.

With these techniques we were able to enhance the reliability of fault detection by identifying the origin of the fault indicators evolution. The application of the designed process on real data helped us prove the legitimacy of the proposed techniques in identifying the different states of bearings over time (healthy or normal, defective) and the origin of the observations' evolution with a low error rate, a reliable diagnosis and a low memory occupation.

Diagnosis and monitoring, Bearings, Pattern recognition, Vibratory analysis, Online learning, Dynamic classification.

Discipline : AUTOMATIQUE, SIGNAL, PRODUCTIQUE, ROBOTIQUE

