



الجمهورية الجزائرية الديمقراطية الشعبية
République Algérienne Démocratique et Populaire
وزارة التعليم العالي و البحث العلمي
Ministère de l'enseignement supérieur et de la recherche scientifique
جامعة 20 أوت 1955 سكيكدة
Université - 20 Août 1955 - Skikda
كلية التكنولوجيا
Faculté de Technologie
قسم : الهندسة الكهربائية
Département : Génie Electrique



THESE

Présentée en vue de l'obtention du diplôme de

Doctorat en Sciences

Option : **Machines Electriques**

THEME

Diagnostic et pronostic des défauts des Moteurs Asynchrones

Présentée par :

Tahar BOUKRA

Soutenue publiquement le : 11 Février 2016

Devant le jury composé de :

Président :	Mr. BOUKADOUM AHCENE	Pr.	Université Skikda
Rapporteur :	Mr. LEBAROUD Abdessalam	Pr.	ENP Constantine
Examineurs :	Mr. BOUZEKRI Hacene	M.C.A.	Université Skikda
	Mr. CHAÂBANE Mabrouk	Pr.	Université Batna
	Mr. AZOUI Boubaker	Pr.	Université Batna

Réf : D012116001D

RESUME

L'importance des enjeux et les prérogatives de la fonction maintenance ont évolué ces dernières décennies accompagnant le développement industriel pour leur rôle dans le maintien et l'amélioration de la disponibilité et la sûreté des systèmes, aussi bien que la qualité du produit. L'évolution vers de nouvelles pratiques de maintenance plus proactives a fait émerger des stratégies prévisionnelles dont le processus clé est le pronostic.

A travers une revue riche de la littérature sur la thématique de diagnostic et de pronostic des défauts des machines électriques, nous avons orienté notre travail, concernant la première partie de notre thèse en l'occurrence le diagnostic, vers l'amélioration du taux de classification correcte des défauts du moteurs asynchrones tels que les défauts des roulements, les défauts déséquilibre des phases statoriques et les défauts cassures de barres rotoriques.

A cette fin, nous avons proposé un classifieur issu directement du plan d'auto-ambiguïté. Ce dernier est choisi comme espace d'extraction des caractéristiques pertinentes. Pour la partie sélection, la dimension du vecteur de forme a été optimisée en minimisant un modèle probabiliste d'erreur de classification proposé pour l'approche statistique. Les résultats de cette approche comparés à des classifieurs intelligents à base de réseaux de neurones sont très satisfaisants.

Concernant la partie pronostic, Vu l'absence de caractéristiques de pronostic pouvant être identifiées de bien représenter la progression de défaut. Nous avons proposé une approche basée sur l'identification des caractéristiques avancées de pronostic susceptibles d'améliorer la prédiction de la durée de vie utile restante (RUL). L'algorithme d'optimisation par essaim particulaire (PSO) a été employé afin d'automatiser la sélection de ces caractéristiques et le *modèle semi-Markov caché* (HSMM) pour estimer la prédiction de la RUL. Bien que l'approche proposée peut être appliquée au pronostic des autres composants de machine, les roulements sont choisis comme objet de recherche car leur défaillance est la cause majeure des pannes des ces machines. Les résultats des caractéristiques identifiées présentent une légère amélioration de la tendance de progression. On devrait penser, donc, à de nouvelles mesures d'aptitude.

MOTS CLES

Diagnostic, pronostic, supervision, maintenance préventive, maintenance conditionnelle, durée de vie résiduelle

ABSTRACT

The importance of the stakes and the prerogatives of the maintenance function have evolved in recent decades with industrial development for their role in maintaining and improving the availability and security of systems, as well as product quality. The move towards new, more proactive maintenance practices have made emerged forecasting strategies in which the key process is the prognosis.

Through an extensive literature review on machine diagnosis and prognosis, we have limited the scope of our work in diagnostic part to improve the correct classification rate of induction motor faults such as bearing faults, imbalance of stator phases and broken bars rotor faults.

To this end, we have proposed a classifier directly in the auto-ambiguity plan chosen as extraction domain of the relevant characteristics. For the selection part, the size of the feature vector has been optimized by minimizing a proposed probabilistic classification error model for the statistical approach. The results of this approach compared to intelligent classifiers based on neural networks are very promising.

For the prognosis part, and in the absence of prognostic features that can be identified to properly represent the fault progression. We have proposed an approach based on the identification of new advanced prognostic features that can improve the prediction of the remaining useful life (RUL). The particle swarm optimization (PSO) algorithm is used to automate the selection of these characteristics and the *hidden semi-Markov model* (HSMM) to estimate this prediction. Although the proposed approach can be applied to prognosis of other machine components, bearings are selected as research object because their failure is the major cause of machine breakdowns. The results of the identified features present a slight improvement in the trend progression. One should think for new fitness measures.

KEY WORDS

Diagnosis, prognosis, supervision, preventive maintenance, Condition-Based Maintenance (CBM), remaining useful life prediction.

ملخص

إن أهمية الرهانات وامتيازات وظيفية الصيانة قد تطورت في العقود الأخيرة مواكبة للتطور الصناعي و ذلك لدوره في الحفاظ وتحسين توافر وأمن النظم، فضلا عن جودة المنتج. هذا التطور جاء نحو ممارسات صيانة جديدة أكثر إستباقية أبرزت استراتيجيات توقعية و التي يعتبر التنبؤ المفتاح رئيسي في العملية. من خلال مراجعة دراسات واسعة على التشخيص و التنبؤ بأعطاب الآلات، تم تحديد نطاق عملنا فيما يخص التشخيص على تحسين معدل التصنيف الصحيح لعيوب المحرك التحريضي مثل عيوب المحامل، عدم توازن أطوار الثابت و عيوب كسر قضبان الدوار.

تحقيقا لهذه الغاية، اقترحنا مصنف مباشرة في مستوى ذاتي-الغموض و الذي اختير ليكون منطقة استخراج الخصائص ذات الصلة. بالنسبة لمرحلة إختيار الخصائص، تم استمثال شعاع الخصائص عن طريق تقليل خطأ تصنيف نموذج احتمالي مقترح للمقاربة الإحصائية. نتائج هذه المقاربة مقارنة مع مصنفات ذكية تقوم على أساس الشبكات العصبية تبدو جد واعدة.

بالنسبة لجزء التنبؤ، و نظرا لعدم وجود خصائص تنبؤية يمكن الإعتماد عليها في التمثيل الصحيح لتقدم الخطأ. تم إقتراح مقارنة تستند على تحديد خصائص تنبؤية متقدمة يمكن أن تحسن من تنبؤ ما تبقى من العمر الإنتاجي (RUL). وقد تم لهذا الغرض استخدام خوارزمية تحسين سرب الجسيمات (PSO) لأتمتة إختيار هذه الخصائص و نموذج شبه ماركوف مخفي (HSMM) لحساب هذا التوقع. على الرغم من أن المقاربة المقترحة يمكن تطبيقها للتنبؤ بالخطأ على مكونات أخرى للمحرك، تم إختيار المحامل كموضوع بحث لأن أعطابهم يعتبر السبب الرئيسي لفشل هذه الآلات. نتائج الخصائص التنبؤية التي تم تحديدها أبرزت تحسن طفيف في نوعية الميل. يجب التفتيش في مقاييس جديدة للدالة الأمثلية.

كلمات مفتاحية

تشخيص، تنبؤ، صيانة إستباقية، صيانة ظرفية، باقي العمر الإنتاجي

TABLE DE MATIERE

CHAPITRE I. INTRODUCTION GENERALE

I. INTRODUCTION GENERALE

I.1 OBJECTIFS	2
I. 2 CONTRIBUTIONS DE CETTE THESE	2
I. 3 LISTE DE PUBLICATIONS.....	3
I. 4 ORGANISATION DE LA THESE	3

CHAPITRE II. CONCEPT DE MAINTENANCE ET ETAT DE L'ART DES APPROCHES DIAGNOSTIC ET PRONOSTIC

II. 1 INTRODUCTION	5
II. 2 DEFINITIONS :	5
II. 3 LES METHODES DE MAINTENANCE.....	6
II. 4 ARCHITECTURE « TYPE » DE MAINTENANCE CONDITIONNELLE ET PREVISIONNELLE CBM.....	7
II. 5 LES TECHNIQUES ACTUELLES DE TRAITEMENT DE DONNEES.....	8
II. 5. 1 CONDITIONNEMENT DES DONNEES	9
II. 5. 2 TECHNIQUES DE TRAITEMENT DU SIGNAL DANS LE DOMAINE TEMPOREL.....	11
II. 5. 3 TECHNIQUES DE TRAITEMENT DU SIGNAL DANS LE DOMAINE FREQUENTIEL..	13
II. 5. 4 TECHNIQUES DE TRAITEMENT DU SIGNAL EN DOMAINE TEMPS-FREQUENCE ..	15
II. 6 LES METHODES ACTUELLES DE DIAGNOSTIC DE DEFAUT	18
II. 7 CLASSIFICATION DES METHODES DE DIAGNOSTIC	20
II. 7. 1 APPROCHES BASEES SUR LA CONNAISSANCE.....	20
II. 7. 1. 1 ANALYSE DES MODES DE DEFAILLANCE ET DE LEURS EFFETS	20
II. 7. 1. 2 ARBRE DE DEFAILLANCE.....	21
II. 7. 1. 3 SYSTEMES EXPERTS	21
II. 7. 2 APPROCHES BASEES SUR LE TRAITEMENT DE DONNEES	22
II. 7. 2. 1 TECHNIQUES NEURONALES	23
II. 7. 2. 2 TECHNIQUES FLOUES.....	23
II. 7. 3 APPROCHES A BASE DE MODELES	24
II. 7. 3. 1 APPROCHES FDI	24
a) ESPACE DE PARITE.....	25
b) ESTIMATION D'ETAT : OBSERVATEURS.....	26
c) ESTIMATION PARAMETRIQUE	26
d) LOCALISATION ET IDENTIFICATION DES DEFAUTS	26
II. 7. 3. 2 APPROCHE DX	27
II. 8 PRONOSTIC ET ETAT DE L'ART DES TRAVAUX	29
II. 8. 1 DEFINITION :.....	29
II. 9 CLASSIFICATION DES APPROCHES DE PRONOSTIC	30
II. 9. 1 APPROCHES BASEES SUR UN MODELE PHYSIQUE.....	31

II. 9. 1. 1	ESPACE DE PARITE.....	32
II. 9. 1. 2	OBSERVATEURS.....	32
II. 9. 1. 3	ESTIMATION PARAMETRIQUE	32
II. 9. 2	APPROCHES DE PRONOSTICS GUIDEES PAR LES DONNEES	33
II. 9. 2. 1	TECHNIQUES STATISTIQUES	34
II. 9. 2. 2	APPROCHES PAR ANALYSE DES SERIES TEMPORELLES.....	34
II. 9. 2. 3	TECHNIQUES DE REGRESSION	34
II. 9. 2. 4	AUTOREGRESSIF MOYENNE MOBILE INTEGREE (ARIMA)	35
II. 9. 2. 5	APPROCHES A BASE INTELLIGENCE ARTIFICIELLE (IA).....	37
II. 9. 3	APPROCHES BASEES SUR L'EXPERIENCE	40
II. 9. 4	TENDANCE VERS L'IDENTIFICATION DE NOUVELLES CARACTERISTIQUES DE PRONOSTIC.....	41
II. 10	CONCLUSION.....	42
II. 11	REFERENCES	44
CHAPITRE III. APPROCHES STATISTIQUES ET NEURONALES POUR L'AMELIORATION DE LA CLASSIFICATION DES DEFAUTS DES MACHINES ASYNCHRONES		
III. 1	INTRODUCTION.....	52
III. 2	ESPACE TEMPS-FREQUENCE COMME DOMAINE D'EXTRACTION	52
III. 3	ALGORITHME CLASSIFICATION	54
III. 3. 1	EXTRACTION DE CARACTERISTIQUES DE DIAGNOSTIC.....	54
III. 3. 2	CONCEPTION DU NOYAU DE CLASSIFICATION	54
III. 3. 3	SELECTION DES CARATERISTIQUES.	56
III. 3. 4	MODELE DE PROBABILITE	57
III. 4	CLASSIFICATION PAR RESEAUX NEURONES ARTIFICIELS (RNA)	61
III. 4. 1	RESEAU DE NEURONES « FEEDFORWARD » (FFNN)	61
III. 4. 2	LES RESEAUX DE NEURONES A ONDELETTES « WAVELETS NEURAL NETWORKS » WNN	62
III. 5	DESCRIPTION DU PROCEDE EXPERIMENTAL.....	64
III. 5. 1	PRESENTATION DU BANC MOTEUR.....	64
III. 5. 2	ACQUISITIONS DES DONNEES	64
III. 5. 3	MODES DE FONCTIONNEMENT ETUDIES	65
III. 5. 4	ENSEMBLE D'APPRENTISSAGE	65
III. 5. 5	DECISION PAR MESURE DE DISTANCE	67
III. 5. 6	DECISION PAR RESEAUX DE NEURONES.....	69
III. 6	CONCLUSION	70
III. 7	REFERENCES	71
CHAPITRE IV. IDENTIFICATION DE NOUVELLES CARACTERISTIQUES DE PRONOSTIC POUR L'AMELIORATION DE LA PREDICTION DE LA DUREE DE VIE UTILE RESTANTE		
IV. 1	INTRODUCTION.....	74
IV. 1. 1	POSITIONNEMENT DU PROBLEME	74
IV. 1. 2	ORGANISATION DU CHAPITRE	74

IV. 2 METHODOLOGIE PROPOSEE POUR LE PRONOSTIC	74
IV. 3 ALGORITHME D'OPTIMISATION PAR ESSAIM DE PARTICULES (PSO)	75
IV. 3. 1 CONCEPTION DE FONCTION FITNESS	77
IV. 4 UTILISATION DU HMM ET HSMM POUR PREDICTION DE LA DUREE DE VIE UTILE RESTANTE.....	78
IV. 4. 1 BACKGROUND THEORIQUE.....	78
IV. 4. 2 ELEMENTS D'UN HMM.....	78
IV. 4. 3 HMM POUR LA PREDICTION DE L'ETAT FUTURE.....	80
IV. 5 HSMM POUR PREDICTION DE LA DUREE DE VIE UTILE RESTANTE.....	81
IV. 5. 1 ÉLÉMENTS D'UN HSMM	82
IV. 5. 2 ADAPTATION DU HSMM AU PRONOSTIC	83
IV. 6 MISE EN ŒUVRE EXPÉRIMENTALE.....	84
IV. 7 EXTRACTION DES CARACTERISTIQUES ET EXPLORATION	86
IV. 8 PRÉDICTION DE LA RUL.....	91
IV. 9 CONCLUSION	92
IV. 10 REFERENCES.....	93
CHAPITRE V. CONCLUSION GENERALE	
V CONCLUSIONS ET TRAVAIL FUTUR.....	95
V .1 CONCLUSIONS	95
V .2 TRAVAIL FUTUR	96
ANNEXE A.....	98
A. ALGORITHME AVANT-ARRIERE	98
B. ALGORITHME DE BAUM-WELCH	100
C. ALGORITHME DE VITERBI.....	101
ANNEXE B.....	103
A. ALGORITHME DE BAUM-WELCH POUR HMMD	103
B. ALGORITHME DE VITERBI POUR HMMD	105

LISTE DES TABLEAUX

TABLEAU III. 1: MODES DE FONCTIONNEMENT POUR CONSTITUER L'ENSEMBLE D'APPRENTISSAGE ET DE TEST [34].....	66
TABLEAU III. 2: RESULTATS DE CLASSIFICATION DE LA DISTANCE MAHALANOBIS (MAHAL) ET EUCLIDIENNE (EUCL) :.....	68
TABLEAU III. 3: RESULTATS DE CLASSIFICATION AVEC LES DIFFERENTS CLASSIFIEURS:	69
TABLEAU IV. 1: CARACTERISTIQUES DES ROULEMENTS ETUDIES	85
TABLEAU IV. 2: RESULTATS APRES L'ESSAIMAGE	91

LISTE DES FIGURES

FIGURE II. 1: METHODES DE MAINTENANCE	6
FIGURE II. 2 : LES ETAPES CLES DE LA MAINTENANCE CONDITIONNELLE CBM.....	8
FIGURE II. 3: CHRONOLOGIE DE PROGRESSION DE DEFAILLANCE D'APRES [33].....	19
FIGURE II. 4 : COMPLEMENTARITE DES ACTIVITES DE DETECTION, DIAGNOSTIC, PRONOSTIC	19
FIGURE II. 5 : CLASSIFICATION DES METHODES DE DIAGNOSTIC	20
FIGURE II. 6 : AMDEC	21
FIGURE II. 7: DIAGNOSTIC PAR RECONNAISSANCE DE FORME	22
FIGURE II. 8 : APPROCHE A BASE DE MODELES	24
FIGURE II. 9: APPROCHE FDI A BASE DE RESIDUS	25
FIGURE II. 10: APPROCHE LOGIQUE BASEE SUR LA COHERENCE.....	27
FIGURE II. 11: CLASSIFICATION DES APPROCHES DE PRONOSTIC	31
FIGURE II. 12 : ARCHITECTURE D'UN RESEAU DE NEURONE A TROIS COUCHES	37
FIGURE III. 1: CONCEPTION DU NOYAU	55
FIGURE III. 2: NORMALEMENT $I_k^{i\varphi}$ EST AFFECTE A LA CLASSE i	57
FIGURE III. 3: SELECTION D'UN NOMBRE OPTIMAL DE CARACTERISTIQUES UTILISANT LA DISTANCE MAHALANOBIS (NOYAU DEFAUT STATOR).....	60
FIGURE III. 4: FONCTIONS DE TRANSFERT.....	61
FIGURE III. 5: SYSTEME D'ACQUISITION DES SIGNAUX	65
FIGURE III. 6: BANC D'ESSAI DU MOTEUR A INDUCTION.....	66
FIGURE III. 7: DEFAUTS ROULEMENTS.....	66
FIGURE III. 8: ROTORS AVEC DES BARRES CASSEES.....	67
FIGURE III. 9: EMBLACEMENT DES CARACTERISTIQUES DANS LE PLAN DOPPLER-RETARD POUR CHACUN DES TROIS NOYAUX.....	67
FIGURE III. 10: SELECTION DU NOMBRE OPTIMAL DE CARACTERISTIQUES UTILISANT LA DISTANCE MAHALANOBIS (NOYAU DEFAUT STATOR).....	67
FIGURE III. 11: COURBE D'APPRENTISSAGE POUR LE ANN UTILISANT L'ALGORITHME LEVENBERG MARQUARDT.	68
FIGURE III. 12: COURBE D'APPRENTISSAGE POUR LE WNN UTILISANT L'ALGORITHME GRADIENT CONJUGUE.....	68
FIGURE IV. 1 : SYSTEME TYPIQUE DE SURVEILLANCE	75
FIGURE IV. 2 : ORGANIGRAMME DE LA METHODE PSO.....	77
FIGURE IV. 3 : TRANSITIONS DES ETATS DU ROULEMENT DE L'ETAT NORMAL A LA DEFAILLANCE.....	78
FIGURE IV. 4 : LES VRAISEMBLANCES DE CHAQUE ETAT DU COMPOSANT EN FONCTION DES SEQUENCES D'OBSERVATIONS.....	80
FIGURE IV. 5 : HMM AVEC DUREE (HMMD)	81
FIGURE IV. 6 : BANC D'ESSAI DE ROULEMENTS.....	84
FIGURE IV. 7 : DEFAUTS DES ROULEMENTS	85
FIGURE IV. 8 : CARACTERISTIQUES RMS DE VIBRATION DES ROULEMENTS (1, 2, 3 ET 4)	86
FIGURE IV. 9 : CARACTERISTIQUES DE DISSYMETRIE DE VIBRATION DES ROULEMENTS (1, 2, 3 ET 4).	86
FIGURE IV. 10 : CARACTERISTIQUES KURTOSIS DE VIBRATION DES ROULEMENTS (1, 2, 3 ET 4)	87
FIGURE IV. 11 : CARACTERISTIQUES QUANTILES DE VIBRATION DES ROULEMENTS (1, 2, 3 ET 4).....	87
FIGURE IV. 12 : AMPLITUDE BSF (BALL SPIN FREQUENCY) DES ROULEMENTS (1, 2, 3 ET 4).....	87
FIGURE IV. 13 : AMPLITUDE BPF (BALL PASSING FREQUENCY INNER-RACE) DES ROULEMENTS (1, 2, 3 ET 4).....	88
FIGURE IV. 14 : AMPLITUDE BPFO (BALL PASSING FREQUENCY OUTER-RACE) DES ROULEMENTS (1, 2, 3 ET 4).....	88

FIGURE IV. 15 : AMPLITUDE FTF (FOUNDATION TRAIN FREQUENCY) DES ROULEMENTS (1, 2, 3 ET 4)	88
.....	
FIGURE IV. 16 : AMPLITUDE DE PUISSANCE VHF DES ROULEMENTS (1, 2, 3 ET 4).....	89
FIGURE IV. 17 : AMPLITUDE DE PUISSANCE VHF DU ROULEMENT (3)	89
FIGURE IV. 18 : AMPLITUDE DE PUISSANCE HF DES ROULEMENTS (1, 2, 3 ET 4).....	90
FIGURE IV. 19 : AMPLITUDE DE PUISSANCE MF DES ROULEMENTS (1, 2, 3 ET 4).....	90
FIGURE IV. 20 : AMPLITUDE DE PUISSANCE LF DES ROULEMENTS (1, 2, 3 ET 4).....	90
FIGURE IV. 21 : CARACTERISTIQUES DE PRONOSTIC IDENTIFIEES POUR LES ROULEMENTS (1, 2, 3 ET 4)	91
.....	

NOMENCLATURE

Abréviations

ALE	Adaptive Line Enhancer
ALM	Accelerated Life Model
AMDE	Analyse des Modes de Défaillance et de leurs Effets
AMDEC	Analyse des Modes de Défaillance de leurs Effets et de leur criticité
ANC	Adaptive Noise Cancellation
ANN	Artificial Neural Network
ARIMA	Autoregressive Integrated Moving Average
ARMA	Autoregressive Moving Average
BD	Blind Deconvolution
BPFI	Ball Pass Frequency of Inner race
BPFO	Ball Pass Frequency of Outer race
BPNN	Back Propagation Neural Network
BSF	Ball Spin Frequency
CBM	Condition-Based Maintenance
CM	Condition Monitoring
CWT	Continuous Wavelet Transform
DCS	Dépendante de la Classe du Signal
DWNN	Dynamic Wavelet Neural Network
DWT	Discrete Wavelet Transform
Eucl	Euclidienne
FC	Frequency Centre
FDI	Fault Detection and Isolation
FDR	Fisher Discriminant Ratio
FFNN	Feed-Forward Neural Network
FFT	Fast Fourier Transform
FTF	Fundamental Train Frequency
GA	Genetic Algorithm
HF	High Frequency
HFRT	High Frequency Resonance Technique
HMM	Hidden Markov Model
HMMD	Hidden Markov Model with Duration
HSMM	Hidden Semi Markov Model
IA	Intelligence Artificielle
IHM	Interface Homme-Machine
ISO	International Standard Organization
Mahal	Mahalanobis
MCO	Maintien en Condition Opérationnelle
MF	Medium Frequency
MSE	Mean Square Error
MSF	Mean square Frequency
NF	Neuro-Fuzzy (Neuro-Flous)
OSA-CBM	Open System Architecture for Condition Based Maintenance
PDF	Probability Density Function
PHM	Prognostics and Health Management
PHM	Proportional hazards model
PIM	Proportional Intensity Models

PSO	Particle Swarm Optimization
RLE	Residual Life Estimate
RMS	Root Mean Square
RMSF	Root Mean Square Frequency
RNA	Réseau de Neurones Artificiel
RNN	Réseau Neuronal Récurrent
RTF	Représentation Temps Fréquence
RUL	Remaining Useful Life
RVF	Root Variance Frequency
SANC	Self-Adaptive Noise Cancellation
SCADA	Supervisory Control and Data Acquisition Systems
SNR	Signal to Noise Ratio
SOM	Self-Organizing Map
STFT	Short-Time Fourier Transform
SVM	Support Vector Machine
SVM-BN	Support Vector Machine - Bayesian Network
SWO	Sum of Weighted Objectives
TSA	Time Synchronous Averaging
VF	Variance Frequency
VHF	Very High Frequency
WAEP	Weight Application to Exponential Parameters
WAFT	Weight Application to Failure Times
WNN	Wavelet Neural Network

Lettres grecques

γ	résidu
η	décalage de fréquence discrète
λ	modèle de Markov
μd_i	moyenne de la distribution gaussienne de durée à l' $i^{\text{ème}}$ état.
μ_i	moyenne de la distribution gaussienne de probabilité d'émission au temps t compte tenu du fait que $S_t = i$
o	produit matriciel élément par élément
π_i	distribution d'état initial.
ρ_{ij}	coefficient de cross-corrélation entre deux entrées de plan d'ambiguïtés
σ^2_{ij}	variance de l'erreur.
σd_i	variance de la distribution gaussienne de durée à l' $i^{\text{ème}}$ état.
τ	temps de retard discret
φ	noyau (Kernel)
φ_{opt}	noyau optimal pour la classification
ϕ	angle de contact.
Σ_i	matrice de covariance de la distribution gaussienne de probabilité d'émission au temps t compte tenu du fait que $S_t = i$
ψ	ondelette mère

$A = \{a_{ij}\}$	distribution des probabilités de transition d'état.
$A[\eta, \tau]$	plan d'auto-ambiguïté de localisation $[\eta, \tau]$.
B_d	diamètre de bille (rouleau).
$b_j(k)$	probabilité d'émission : probabilité d'observer des symboles individuels v_k dans l'état i .
C_i	classe d'indice i .
$D(S_n)$	durée de l'état n .
d	mesure de distance.
d_{eucl}	distance Euclidienne.
d_{quad}	fonction discriminante quadratique.
d_M	distance de Mahalanobis.
F	transformée de Fourier.
F^{-1}	transformée de Fourier inverse.
f_r	fréquence de rotation.
$gBest$	meilleur global.
\overline{I}_i^φ	moyenne des signaux d'apprentissage de la classe C_i .
$I_k^{i\varphi}$	vecteur de forme représentant un signal à classer.
k	fréquence discrète.
M	nombre de symboles distincts d'observation par l'état.
$m_i[\eta, \tau]$	moyenne de la classe i de localisation (η, τ) .
m_{ij}	moyenne de l'erreur.
$m_{i-remain}[\eta, \tau]$	moyenne du reste des classes de localisation (η, τ) .
N	nombre d'états dans le modèle.
N_b	nombre de billes.
n	indice d'échantillon.
$O = O_1 O_2 \dots O_T$	séquence d'observations.
$P(O \lambda)$	probabilité de la séquence d'observation étant donné un modèle Markov λ .
$\tilde{P}(d, \varphi)$	modèle de probabilité d'erreur.
P_d	diamètre primitif.
$pBest$	meilleure particule.
$p_i(d)$	densité de durée de l'état .
Q	Q-fonction.
S_i	état i .
s^k	position actuelle.
s^{k+1}	nouvelle position.
\hat{t}_c	durée estimée au point-changement d'état suivant.
$V = \{v_1, v_2, \dots, v_M\}$	ensemble de symboles individuels.
$V_i^2[\eta, \tau]$	variance de la classe i de localisation (η, τ) .

$V_{i-remain}^2 [\eta, \tau]$ variance du reste des classes de localisation (η, τ) .

v^k vitesse actuelle.

v^{k+1} nouvelle vitesse.

x_{mk} $k^{ème}$ caractéristique (entrée de la fonction ambiguïté) du $m^{ème}$ signal de l'ensemble. d'apprentissage.

REMERCIEMENT

En premier lieu, je tiens à remercier mon directeur de thèse Monsieur **Abdessalam LEBAROUD** pour son encadrement, ses précieux conseils, sa disponibilité et sa confiance en mon travail. Ce qui m'a permis une indépendance autour des thèmes de recherche.

Je suis très sensible à l'honneur que m'a fait Monsieur **Ahcene BOUKADOUM** en acceptant de rapporter et de présider mon jury de thèse. J'adresse toute ma gratitude à Monsieur **Hacene BOUZEKRI** pour avoir accepté d'évaluer mes travaux en qualité d'examineur. Mes remerciements s'adressent également à Monsieur **Mabrouk CHAABANE** et Monsieur **Boubaker AZOUI** pour leur participation au jury en tant qu'examineurs.

Je remercie également, et sans exception, l'ensemble du personnel du Département Génie Electrique de l'université de Skikda. J'adresse mes remerciements également à Monsieur **Guy CLERC**, Professeur à l'Université Claude Bernard Lyon 1 de m'avoir accueilli au sein du laboratoire Ampère à l'Université Claude Bernard Lyon 1 durant mon stage de courte durée.

Je ne peux terminer sans avoir une pensée pour ma famille. Sans elle, je n'aurais jamais pu atteindre mes objectifs. Un grand merci à mes parents pour leur soutien indéterminé

I. INTRODUCTION GENERALE

En monde industriel, la productivité est l'objectif majeur pour les compagnies industrielles pour rester en compétitivité permanente dans un marché qui ne cesse d'accroître. Une productivité croissante peut être réalisée à travers une disponibilité croissante de la capacité de production. Le développement technologique a résulté en une complexité croissante des machines industrielles et des systèmes de production. Aussi, des exigences de qualité, de sécurité et de coût sont apparues. Donc, il n'est plus suffisant de se satisfaire d'une performance technique et économique, mais il devient nécessaire, voire obligatoire, de prendre en compte les contraintes environnementales : une usine laisse des déchets, pollue et contribue à l'effet de serre. A ceci s'ajoute le respect de la dignité humaine qui constitue une contrainte sociale. En conséquent, l'importance et les enjeux et prérogatives de la *fonction maintenance* ont évolué en parallèle ces dernières décennies avec le développement technologique pour leur rôle dans le maintien et l'amélioration de la disponibilité et la sûreté des systèmes, aussi bien que la qualité du produit.

Une action de maintenance consiste à remplacer les équipements du système qui sont en panne et qui ne sont plus capables d'accomplir leur fonction. La décision d'une action de maintenance est très complexe et doit reposer sur une surveillance et une analyse intelligente de l'état du système. Elle doit idéalement être réduite aux opérations qui consistent à remplacer, sans tâtonnement, les équipements réellement en panne. Plus la phase de maintenance est longue, plus elle est coûteuse. Un *diagnostic* de défaut est donc nécessaire pour déterminer le plus précisément possible les équipements qui doivent être réparés. Le diagnostic lorsqu'il est réalisé avec efficacité et s'il permet de détecter de façon précoce une dégradation, représente un des moyens pour contribuer à obtenir un meilleur gain de productivité. Il devient donc une partie intégrante de la fonction de la maintenance.

Pour la surveillance automatique des dispositifs complexes, il convient de développer des systèmes de diagnostic qui disposent d'une certaine faculté d'adaptation à des situations nouvelles (apprentissage), et permettent une reconnaissance fiable du mode de fonctionnement dans lequel le système évolue.

Pour minimiser l'occurrence des actions imprévues de maintenance qui n'ont pas été anticipés, une maintenance préventive doit être envisagée. Cette dernière peut être améliorée par un raisonnement *pronostic* dans le but de réaliser une maintenance adaptative des composants de la machine, autrement dit une *maintenance conditionnelle* CBM (Conditioned-Based Maintenance). Ainsi, sur la base de cette idée, nous obtiendrions une technologie capable de diagnostiquer les défaillances du type défaillances brusques (hard failures) ou brusques liées à des événements inopinés, et de pronostiquer les défaillances du type (soft failures) qui soient liées aux phénomènes de vieillissement ou dégradation. Par conséquent, une maintenance du type CBM pourrait être appliquée en prévoyant les opérations de maintenance d'après les nécessités du système au lieu du temps de service.

Si le diagnostic a atteint une certaine maturité, le pronostic est, quant à lui, une discipline émergente. Son but est de prédire la durée de vie résiduelle des composants défaillants de la machine et d'estimer l'impact des sollicitations influençant les équipements du système sur leurs durées de vie. Le pronostic est considérablement plus difficile à formuler que le diagnostic puisque sa précision est sujette à des processus stochastiques qui ne sont pas encore produit. Les systèmes industriels étant généralement complexes, il est souvent difficile de construire un modèle formel de pronostic. En général, beaucoup d'ingénieurs diagnostiqueurs ont une avance en connaissance d'événements et d'expérience sur la défaillance de la machine et son état de santé en surveillant continuellement et en analysant la condition de la machine dans l'industrie, mais il n'a pas de méthodologies systématiques claires pour la prédiction précise de la vie résiduelle de la machine comme support à la prise de décision.

Nous avons choisi d'orienter notre travail vers une approche guidée par les données car les technologies de mesure (capteurs) permettent aujourd'hui de récupérer en temps réel les données de fonctionnement des équipements ce qui permet d'envisager une stratégie de maintenance plus « dynamique » basée sur les données conditionnelles du composant au lieu d'une stratégie de maintenance préventive systématique. En plus de la difficulté de disposer de connaissance sur les phénomènes de dégradation complexes.

I.1 OBJECTIFS

Dans le cadre du développement d'un système PHM (Prognostics and Health Management) complet nous avons orientés les objectifs principaux de cette thèse sur l'amélioration des capacités diagnostic et pronostic des défauts des composants des machines tournantes. Parmi celles-ci, le moteur asynchrone qui occupe une place critique et privilégiée dans plusieurs industries. Ce travail enquête sur les défauts des roulements qui sont les causes majeures des arrêts des ces machines. Aussi sur les défaillances d'autres composants telles que les barres rotoriques et le déséquilibre des phases statoriques.

Ce travail vise à exploiter les données collectées à partir des tests de défauts artificiels sur des bancs d'essais réalisés au « Laboratoire Ampère Université de Lyon, France » menés sur un moteur asynchrone de puissance 5KW pour la partie diagnostic et des tests de vieillissement accéléré conduits sur des roulements au « Center for Intelligent Maintenance Systems, University of Cincinnati, USA » pour la partie pronostic.

Le premier objectif de cette thèse est le développement d'un modèle de diagnostic à taux de classification amélioré capable d'être intégré dans un environnement réel.

Le deuxième objectif est le développement d'un modèle de pronostic capable d'estimer la durée de vie résiduelle avec précision.

L'objectif final de cette thèse est l'exploitation des données expérimentales collectées et la validation des modèles diagnostic et pronostic en vue d'une application en monde réel de supervision.

I.2 CONTRIBUTIONS DE CETTE THESE

Cette thèse réclame les contributions originales suivantes :

Proposition d'un modèle d'erreur probabiliste pour le choix optimal du nombre de caractéristiques pertinentes pour le vecteur de forme.

Etude comparative des classifieurs à approches statistique et neuronale pour la partie diagnostic.

Proposition d'un modèle pronostic à base de modèle semi Markov caché.

Validation des modèles proposés en utilisant les données expérimentales.

I. 3 LISTE DE PUBLICATIONS

Plusieurs publications ont été produites comme partie de ce travail de recherche,

PUBLICATIONS PARTIE DIAGNOSTIC

Communications internationales

- **Boukra, T.;** Lebaroud, A. "Classification of Induction Machine Faults, " 7th *International Multi-Conference on Systems Signals and devices (SSD)*, p: 1 - 6, 2010
- **Boukra, T.;** Lebaroud, A.; Medoued, A. "Wavelet-Network for Classification of Induction Machine Faults Using Optimal Time-Frequency Representations," 7th *International Conference on Electrical and Electronics Engineering (ELECO)*, p: I-358 - I-362, 2011.
- **Boukra, T.;** Lebaroud, A.; Clerc, G. "Neural-Network Approaches for Classification of Induction Machine Faults Using Optimal Time-Frequency Representations," 10th *International Conference on Environment and Electrical Engineering (EEEIC)*, p. 1-4, 2011.

Publications Internationales

- **Boukra, T.;** Lebaroud, A.; Clerc, G. "Statistical and Neural-Network Approaches for the Classification of Induction Machine Faults Using the Ambiguity Plane Representation," *IEEE Transactions On Industrial Electronics*, Vol. 60, No. 9, p. 4034-4043, September 2013.

PUBLICATIONS PARTIE PRONOSTIC

Communications internationales

- **Boukra, T.;** Lebaroud, A. "Identifying New Prognostic Features for Remaining Useful Life Prediction," 16th *International Power Electronics and Motion Control Conference and Exposition (PEMC 2014)*, Antalya, Turkey, p. 1437-1443, Sept 2014.

I. 4 ORGANISATION DE LA THESE

L'organisation de cette thèse est comme suit :

Le chapitre I introduit une vue d'ensemble brève et l'étendue du domaine de recherche. Ce chapitre présente aussi l'importance, les objectifs et l'innovation de cette recherche. Il montre comment les objectifs de la recherche sont tracés à partir des problèmes identifiés. L'originalité et les principales contributions de ce travail sont aussi présentées.

Le chapitre II a pour objet de situer et limiter le cadre de notre travail. Il présente une revue détaillée de la littérature des techniques de supervision actuelles. En premier lieu, les informations de base sur les stratégies de maintenance sont passées en revue ainsi que leur évolution conceptuelle jusqu'à l'état actuel. Dans le cadre de la maintenance conditionnelle, la section II. 5 décrit les détails des étapes de cette stratégie en partant des techniques existantes du traitement du signal comme un pas fondamental vers l'étape de diagnostic ou de pronostic comme support à la décision. Un état de l'art des approches diagnostic et pronostic est également présenté à la section II. 7 et II. 9 respectivement.

Le chapitre III décrit les étapes principales du modèle proposé pour le diagnostic des défauts de la machine asynchrone. Ces étapes constituent les trois étapes clés de la maintenance conditionnelle (CBM). En premier lieu, l'étape d'extraction des caractéristiques pertinentes à partir du plan d'ambiguïté filtré par un noyau optimisé pour un maximum de séparabilité entre classes est présentée à la section III. 3. 2. Suivi de l'étape de sélection où un modèle de probabilité d'erreur dédié à l'optimisation du vecteur de forme est proposé à la section III. 3. 4. Enfin, les performances de différents classifieurs statistiques et neuronaux, comme support à la décision, sont présentés, comparés et validés par des données expérimentales recueillies à partir d'un banc d'essai sur un moteur asynchrone à la section III. 5.

Le chapitre IV est une extension de la partie aide à la décision pour un PHM complet en abordant le domaine pronostic par une approche toujours guidée par les données vu la difficulté de disposer de connaissance sur les phénomènes de dégradation complexes des composants, en utilisant le modèle semi Markov caché HSMM. Après la présentation des fondements théoriques du HSMM et son adaptation au pronostic à la section IV. 5. L'étape de validation, par des données expérimentales recueillies à partir des tests de vieillissement accélérés des roulements, est présentée aux sections IV. 6 - IV. 8.

Finalement, la thèse est achevée par une conclusion générale qui résume les contributions de notre travail de recherche reportées dans cette thèse et ébauche les directions du travail future.

II. 1 INTRODUCTION

Ce chapitre aborde les différentes formes de la maintenance ainsi que leur évolution conceptuelle jusqu'à l'état actuel. Par comparaison avec les stratégies traditionnelles de la maintenance, la maintenance conditionnelle (CBM) offre un potentiel pour la minimisation des instances des défaillances des équipements, une réduction dans les activités planifiées de maintenance et une maximisation de la disponibilité des équipements. Dans ce contexte, des détails sur les différentes étapes de cette stratégie sont présentés dans la section II. 5. En partant de l'étape d'acquisition des données et leur conditionnement vers l'étape de diagnostic et de pronostic comme support à la décision tout en passant par les techniques existantes de traitements de signal considérées comme un pas fondamental pour l'extraction et la sélection des caractéristiques de diagnostic et de pronostic pertinentes. Un état de l'art sur les approches de diagnostic et de pronostic est également présenté dans la section II. 7 et II. 9 respectivement.

II. 2 DEFINITIONS :

Pour unifier la terminologie en matière de maintenance, on présente les définitions suivantes :

Défaillance : altération ou cessation de l'aptitude d'un bien à accomplir la fonction requise.

Il existe 2 formes de défaillance :

Défaillance partielle : altération de l'aptitude d'un bien à accomplir la fonction requise.

Défaillance complète : cessation de l'aptitude d'un bien à accomplir la fonction requise.

Les concepts de défaut et de défaillance sont définis dans [52].

Défaut : un défaut représente une déviation non acceptable d'au moins d'une propriété caractéristique ou d'un paramètre du système.

Un défaut peut générer une défaillance du système. Après une défaillance, on considère que le système est en panne.

Panne : est définie par l'inaptitude d'une entité à accomplir une fonction requise. Une panne est la conséquence d'un défaut qui est donc toujours associée à une défaillance.

Maintenance : selon la norme [1] la maintenance peut être définie comme l'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.

Ainsi, la maintenance vise globalement à appréhender les phénomènes de défaillance et à agir en conséquence, afin d'assurer que le système (le bien) soit à même de remplir la fonction pour laquelle il a été conçu (Maintien en Condition Opérationnelle, MCO).

L'analyse des différentes formes de maintenance repose sur 4 concepts :

Les évènements qui sont à l'origine de l'action : référence à un échancier, la subordination à un type d'événement (auto diagnostic, information d'un capteur, mesure d'une usure, etc.), l'apparition d'une défaillance.

Les méthodes de maintenance qui leur seront respectivement associées : maintenance préventive systématique, maintenance préventive conditionnelle, maintenance corrective.

Les opérations de maintenance: inspection, contrôle, dépannage, réparation, etc.

Les activités connexes : maintenance d'amélioration, rénovation, reconstruction, modernisation, travaux neufs, sécurité, etc.

II. 3 LES METHODES DE MAINTENANCE

Le diagramme suivant synthétise selon la norme [1] les différentes méthodes de maintenance. Pour choisir la méthode de maintenance, il faut donc être informé des objectifs de la direction de l'entreprise, des directions politiques de maintenance, mais il faut connaître le fonctionnement et les caractéristiques des matériels, le comportement du matériel en exploitation, les conditions d'application de chaque méthode, les coûts de maintenance et les coûts de perte de production.

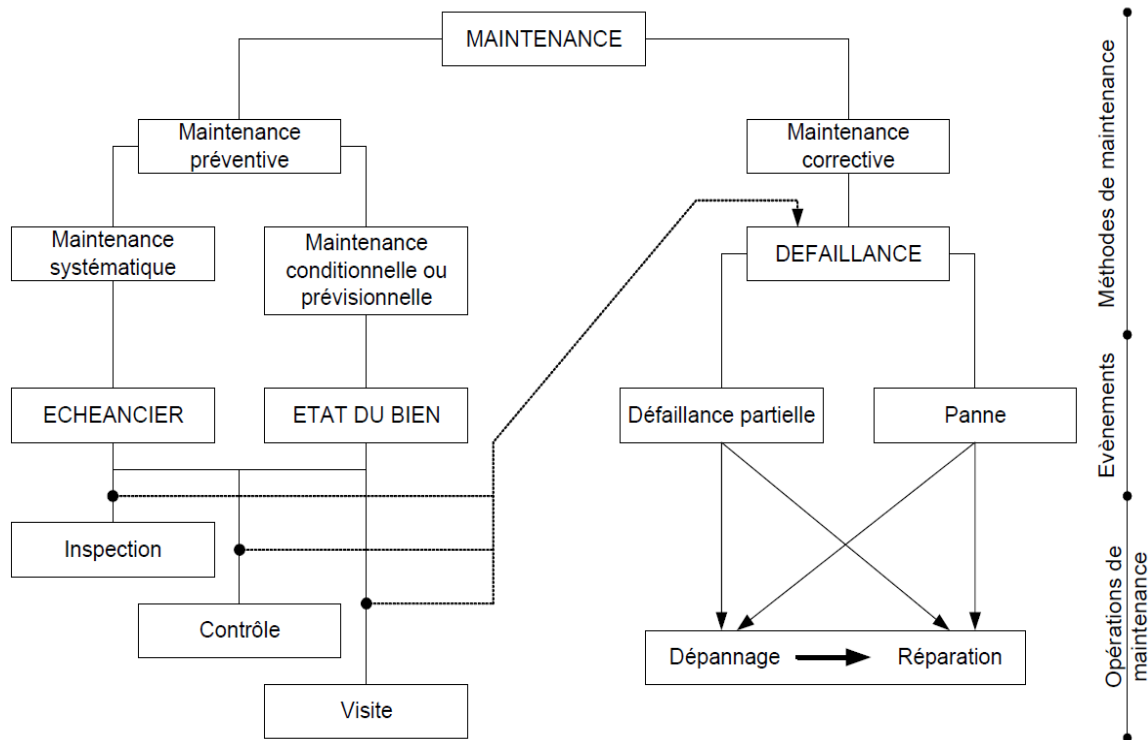


Figure II. 1: Méthodes de maintenance

Il existe deux familles principales de maintenance que l'on peut repérer sur la figure II. 1 : la *maintenance corrective* et la *maintenance préventive*.

Maintenance Corrective

La maintenance corrective a pour objet de redonner au matériel des qualités perdues nécessaires de manière à ce qu'il soit capable de remplir à nouveau ses fonctions.

Elle est généralement adoptée pour les équipements pour lesquels :

- Les conséquences de la panne ne sont pas critiques,
- la réparation est facile et ne nécessite pas beaucoup de temps,
- les coûts d'investissements sont faibles.

Deux formes de maintenance correctives peuvent être distinguées. La *maintenance curative* et la *maintenance palliative*.

Maintenance Curative

Ce type de maintenance permet de remettre définitivement en état le système après l'apparition d'une défaillance. Elle provoque donc une indisponibilité du système.

Maintenance Palliative

La maintenance palliative revêt un caractère temporaire, provisoire.

Maintenance Préventive [1]

Maintenance exécutée à des intervalles prédéterminés ou selon des critères prescrits et destinée à réduire la probabilité de défaillance ou la dégradation du fonctionnement d'un bien. Il existe trois formes particulières de maintenance préventive : la *maintenance préventive systématique*, la *maintenance préventive conditionnelle* et la *maintenance préventive prévisionnelle*.

Maintenance Préventive Systématique :

Maintenance préventive exécutée à des intervalles de temps préétablis ou selon un nombre défini d'unités d'usage mais sans contrôle préalable de l'état du bien [1]. Elle utilise, comme indicateur, des informations provenant du contexte opérationnel (temps d'utilisation ou nombre de cycles) généralement fournies par les processus de management de l'utilisation.

Maintenance-Conditionnelle

Maintenance préventive basée sur une surveillance du fonctionnement du bien et/ou des paramètres significatifs de ce fonctionnement intégrant les actions qui en découlent. La surveillance du fonctionnement et des paramètres peut être exécutée selon un calendrier, ou à la demande, ou de façon continue [1].

Elle est subordonnée à un type d'événement prédéterminé (résultat de diagnostic, donnée d'un capteur, mesure d'usure, etc...) révélateur de l'état de fonctionnement du système.

Cette stratégie s'appuie donc sur la dégradation du matériel et non sur sa défaillance. Elle permet de déclencher l'opération de maintenance lors du franchissement d'un seuil (qui représente le niveau de dégradation maximal tolérable) par un indicateur.

Maintenance Prévisionnelle :

Maintenance conditionnelle exécutée en suivant les prévisions extrapolées de l'analyse et de l'évaluation de paramètres significatifs de la dégradation du bien [1] en considérant l'ensemble des paramètres influents provenant des contextes opérationnel et environnemental futurs ainsi que des actions futurs de maintenance [2].

Cette forme de maintenance permet de réduire le nombre de défaillances imprévues, et donc l'indisponibilité du système. Elle permet de planifier les opérations de maintenance de manière à utiliser les équipements au maximum de leurs possibilités. En surveillant les équipements, il est possible de corriger des erreurs de conduite ou des anomalies qui peuvent générer des défaillances plus graves par la suite et d'améliorer la sécurité en évitant des accidents critiques. Par contre, cette forme de maintenance nécessite de mettre en place des techniques de surveillance et de mesure qui peuvent être très coûteuses [3].

II. 4 ARCHITECTURE « TYPE » DE MAINTENANCE

CONDITIONNELLE ET PREVISIONNELLE CBM

Un système intégral de maintenance conditionnelle CBM doit comprendre un certain nombre de capacités fonctionnelles:

Détection et acquisition des données, manipulation de données, surveillance de l'état de santé, évaluation/diagnostic de l'état de santé, pronostic et prise de décision. En outre, une certaine forme d'interface Homme-Machine est nécessaire pour permettre l'accès de l'utilisateur au système et fournir un moyen d'affichage des informations vitales. Ainsi, un grand nombre de

travaux sont apparus récemment pour proposer des architectures complètes de maintenance prévisionnelle. Dans ces travaux, les différences n'apparaissent qu'à un niveau de détail relativement petit, et portent essentiellement sur la forme des architectures logicielles à mettre en œuvre (locale, distribuée, par modules, etc...) [4]. Actuellement, dans le but de développer et d'encourager l'adoption des normes d'information ouvertes pour les opérations de maintenance en industrie, l'institut « Machinery Information Management Open Standards Alliance (MIMOSA)¹ » fournit une architecture standardisée pour un système de maintenance conditionnelle CBM appelée (OSA-CBM) (Open System Architecture for Condition Based Maintenance) [5]. Cette architecture a été décrite en termes de couches fonctionnelles ; de la détection et l'acquisition des données au support de décision. Les fonctions générales des couches sont spécifiées ci-dessous :

Couche 1 : Module capteurs. Ce module fournit au système des données numériques issues des capteurs ou des transducteurs.

Couche 2 : Module traitement du signal. Ce module reçoit les données à partir des capteurs ou transducteurs ou d'autres processeurs de signaux et réalise des transformations de signaux et des extractions de caractéristiques ou de descripteurs.

Couche 3 : Module de surveillance. Le module de surveillance compare les données en ligne avec certaines valeurs attendues ou connues ; il doit également être capable de générer des alertes en fonction de seuils préalablement fixés.

Couche 4 : Module de diagnostic. Ce module détermine si l'état du système, sous-système ou composant surveillé est dégradé ou non et suggère les défaillances probables.

Couche 5 : Module de pronostic. Ce module prédit l'état futur du système, sous-système ou composant surveillé. Le module s'appuie sur les données issues des modules précédents.

Couche 6 : Module d'aide à la décision. Sa fonction principale est de recommander des actions de maintenance ou d'autres alternatives pour continuer à faire fonctionner le système jusqu'à l'accomplissement de sa mission.

Couche 7 : Module de présentation. Ce module reçoit les informations de tous les modules précédents. Il peut être construit sous forme d'une Interface Homme-Machine IHM.

Les sept couches précédentes peuvent aussi être résumées en trois étapes clés comme le montre la figure II. 2.

Dans les sections suivantes, les techniques de traitement de données existantes ainsi que les méthodes de diagnostic et de pronostic sont décrites et examinées car elles sont les éléments clés d'un système de surveillance efficace.

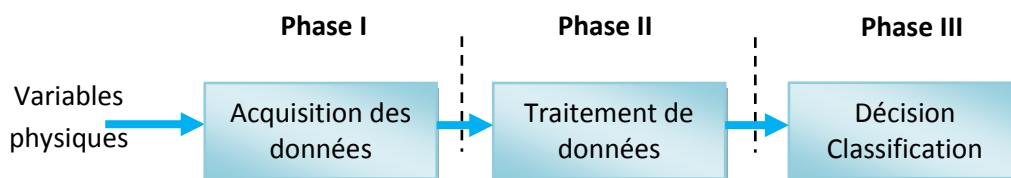


Figure II. 2 : Les étapes clés de la maintenance conditionnelle CBM

II. 5 LES TECHNIQUES ACTUELLES DE TRAITEMENT DE DONNEES

¹ www.mimosa.org

Les types de données de la maintenance conditionnelle se répartissent en deux catégories: *données d'événements* et *données de surveillance*.

Les données d'événements sont les enregistrements d'occurrences horodatés, généralement enregistrées manuellement, tels que les installations des équipements, les défaillances, les récupérations, les remplacements ou les interruptions préventives pour une éventuelle maintenance. L'acquisition des données d'événement est souvent négligée et parfois considérée comme pas importante. Ça ne doit pas être le cas, car elle contribue à l'identification du seuil de défaillance pour les données de surveillance indiquant si les données historiques se terminent par une défaillance ou une suspension. Elle contribue aussi à l'évaluation des performances des indicateurs de surveillances.

D'autre part, les données de surveillance d'état peuvent être des données vibratoires, acoustiques, analyse d'huile, de température, de pression, d'humidité ou d'environnement. Elles peuvent être enregistrées automatiquement au moyen d'un software comme LabVIEW et un matériel pour l'acquisition comme celui de National Instruments.

Avec le développement rapide de l'informatique et les technologies des capteurs, l'acquisition des données pour une mise en œuvre de la maintenance conditionnelle CBM est devenue plus puissante et moins chère. De différents types de capteurs, combinés avec les technologies de traitement du signal, ont été inventés et présentés dans des articles, mais peu qui ont trouvé leur chemin vers une application industrielle [6]. Les technologies sans fils comme le Bluetooth sont développées pour permettre une facilité et une efficacité dans la transmission des données.

Les données de surveillance collectées de l'étape d'acquisition sont versatiles et sont divisées en trois catégories [7]

Type valeur : données collectées à un temps spécifié pour une variable surveillée sont une valeur singulière. Par exemple : donnée d'analyse de l'huile, température, pression, humidité sont toutes des données de type valeur.

Type forme d'onde : données collectées à un temps spécifié pour une variable surveillée sont une série temporelle qui est souvent appelée forme d'onde. Exemple : données vibratoires ou acoustiques.

Type multidimensionnelle : données collectées à un temps spécifié pour une variable surveillée sont multidimensionnelles comme les données d'images infrarouge thermographe, image rayon X, images visuelles...

Le traitement de signal des données multidimensionnelles est plus compliqué que celui des données forme d'onde à cause de l'implication de plus d'une dimension.

L'auteur [8] a réalisé une application sur les techniques de traitement d'images pour la surveillance conditionnelle et le diagnostic des défauts.

Une attention particulière est accordée aux données de type forme d'onde car elles nécessitent plus d'une stratégie de traitement. Une variété de techniques a été développée pour leur analyse et leur interprétation.

II. 5. 1 CONDITIONNEMENT DES DONNEES

Les données brutes acquises à partir des capteurs sont contaminées de bruits et ont un rapport signal-bruit faible. Elles sont, donc, prétraitées avant d'être utilisées pour une éventuelle analyse.

La technique de base pour le conditionnement des signaux est le filtrage passe-bas, car le bruit domine le signal dans les hautes fréquences. La fréquence de coupure doit être au moins deux fois la fréquence utile du signal la plus grande. Pour filtrer les signaux acquis, la technique « Time Synchronous Averaging » (TSA) ou moyenne synchrone temporelle a été utilisée dans les défauts des réducteurs pour moyenniser les signaux bruts de vibration sur un nombre d'évolutions dans le but d'annuler les bruits aléatoires [7], [9]. Elle est formulée par :

$$\bar{s}(t) = \frac{1}{N} \sum_{n=0}^{N-1} s(t + nT), \quad 0 \leq t < T, \quad (\text{II.1})$$

où $s(t)$ est le signal, T est la période moyennée et N est le nombre des échantillons pour le moyennage.

L'auteur [10] a démontré que la rejection d'un bruit périodique d'une fréquence connue peut être optimisée par la sélection d'un nombre approprié de moyenne.

La TSA a été appliquée [11] pour les signaux de vibrations des éléments roulants dans les roulements.

L'Annulation Adaptative de Bruit « Adaptive Noise Cancellation » (ANC) est un autre algorithme efficace qui a été proposé par l'auteur [12]. Ce principe a été trouvée utile pour la séparation des signaux de vibrations des éléments roulants des roulements du reste des signaux vibratoires, aidant ainsi au diagnostic des défauts.

Les auteurs [13] ont appliqués la technique ANC pour renforcer les signatures vibratoires des éléments roulants des roulements. Cette technique nécessite l'emploi de deux capteurs : l'un pour la mesure du signal bruité et le deuxième pour la mesure du bruit de référence. Avec le bruit de référence corrélé au bruit du signal endommagé, la technique ANC annule la composante bruit du signal endommagé et laisse le signal désiré.

Les techniques « Adaptive Line Enhancer » renforceur de ligne adaptative (ALE) [14] et « Self-Adaptive Noise Cancellation » (SANC) [15] ont été proposées pour réduire à un seul le nombre des capteurs en remplaçant le bruit de référence par une version décalée du signal mesuré. Cependant, la SANC a plusieurs contrariétés dans les applications réelles, tel que l'ajustage des paramètres pendant la phase d'adaptation.

Pour palier à ce problème les auteurs [16] ont proposé l'estimation du gain de fréquence d'annulation de bruit dans le domaine fréquentiel. Cet algorithme est vérifié à travers la séparation des signaux de vibrations des roulements de ceux des réducteurs, comme les premiers sont généralement masqués par les derniers.

La technique « Blind Deconvolution » (BD) ou déconvolution aveugle a été appliquée [17] pour récupérer les signaux désirés des roulements sans une connaissance à priori du canal de réception.

Les auteurs [18] ont proposés un filtre à ondelettes pour renforcer les faibles signatures pour l'identification des défauts précoces dans les roulements.

Une fois les données « nettoyées », diverses techniques de traitement du signal ont été développées pour analyser et interpréter les données de forme d'onde dans le but d'extraire les informations utiles à des fins de diagnostic et de pronostic.

Généralement les techniques d'analyse des données forme d'ondes se répartissent en trois catégories: les techniques dans le domaine temporel, les techniques dans le domaine fréquentiel et les techniques dans le domaine temps-fréquence.

II. 5. 2 TECHNIQUES DE TRAITEMENT DU SIGNAL DANS LE DOMAINE TEMPOREL

La technique la plus directe est simplement l'inspection visuelle des portions temporelles du signal. Mais la complexité des signaux acquis rend souvent cette analyse inadéquate pour un éventuel diagnostic ou pronostic.

Les techniques dans le domaine temporel sont basées sur le comportement statistique distinct des signaux de forme d'onde. L'analyse la plus simple, dans le domaine temporel, calcule la valeur efficace (RMS) et le facteur crête (Crest Factor) d'un signal. D'autres caractéristiques couramment utilisés sont les pic, l'amplitude crête-à-crête, l'écart-type, le facteur d'asymétrie, le kurtosis et la moyenne synchrone temporelle (TSA)...etc.

Les caractéristiques décrites ci-avant sont appelées caractéristiques statistiques parce qu'elles sont fondées uniquement sur la distribution des échantillons signal. Ces caractéristiques sont appelées aussi moments ou cumulant. Dans la plupart des cas, la fonction de densité de probabilité (PDF) peut être décomposée en ses moments constitutifs. Un changement d'état entraîne un changement de la fonction de densité de probabilité du signal. Par conséquent, les moments peuvent aussi changer. Donc, le suivi de ce phénomène peut fournir des informations utiles pour le diagnostic.

Les moments des données forme d'onde peuvent être calculés en utilisant les équations suivantes :

$$m_n = E\{x^n\} = \frac{1}{N} \sum_i^N x_i^n \quad (\text{II. 2})$$

où $E\{x^n\}$ représente l'espérance mathématique de la fonction, x_i est la $i^{\text{ième}}$ donnée historique et N est le nombre de points de données.

Les quatre premiers cumulant: moyenne, écart-type, asymétrie et kurtosis, peuvent être calculés à partir des quatre premiers moments en utilisant les relations suivantes :

$$\text{Moyenne} = m_1$$

$$\text{Standard Deviation} = m_2 - m_1^2$$

$$\text{Asymétrie} = m_3 - 3m_2m_1 + 2m_1^3$$

$$\text{Kurtosis} = m_4 - 3m_2^2 - 4m_3m_1 + 12m_2m_1^2 - 6m_1^4$$

En outre, les caractéristiques non-dimensionnelles dans le domaine temporel, comme le facteur de forme et le facteur de crête sont couramment utilisées.

$$\text{Shape Factor} = x_{rms}/x_{abs}$$

$$\text{Crest Factor} = x_p/x_{rms}$$

$$\text{Clearance Factor} = \frac{x_p}{\left(\frac{1}{N} \sum_{k=1}^N \sqrt{|x(k)|}\right)^2}$$

$$(\text{Impulse Factor}) IF = \frac{x_p}{\frac{1}{N} \sum_{k=1}^N |x(k)|}$$

Où x_{rms} , x_{abs} et x_p sont la valeur quadratique, moyenne (en valeur absolue) et la valeur de crête, respectivement.

Les histogrammes qui peuvent être considérés comme une fonction discrète de densité de probabilité (PDF) sont calculés de la façon suivante. Soit d le nombre de divisions d'intervalle nécessaires, soient h_i avec $0 \leq i \leq d$ les colonnes de l'histogramme, alors

$$h_i = \sum_{i=0}^n \frac{1}{n} r_i(x_i), \quad \forall i, \quad 0 \leq i \leq d$$

$$r_i(x) \begin{cases} 1, & \text{si } \frac{i(\max(x_i) - \min(x_i))}{d} \leq x < \frac{(i+1)(\max(x_i) - \min(x_i))}{d} \\ 0, & \text{autrement} \end{cases}$$

La limite supérieure de l'histogramme (h_U) et la limite inférieure (h_L) sont définies comme,

$$h_U = \max(x_i) + \frac{\Delta}{2}$$

$$h_L = \max(x_i) - \frac{\Delta}{2}$$

$$\text{où } \Delta = \frac{\max(x_i) - \min(x_i)}{(n-1)}$$

Il est normalisé par deux paramètres: la longueur de la séquence et les divisions de colonne. Puisque le terme « somme » ci-dessus comprend un terme $1/n$, et chaque x_i doit se situer exactement dans une colonne h_i , l'effet net est que $h_i = 1$ ($i = 0, \dots, d-1$). Les divisions de la colonne sont liés à la boîte délimitée, et donc la plupart des h_i ci-dessus ne seront pas nuls. Cela est souhaitable, car elle élimine essentiellement la question de la taille d'un signal, et la basse résolution dans les signaux faibles, avec beaucoup de colonnes vides.

Dans la théorie de l'information, l'incertitude peut être mesurée par entropie. L'estimation de l'entropie est un processus en deux étapes; d'abord, un histogramme est évalué, puis l'entropie est calculée. L'estimation de l'entropie $E_s(x_i)$ et l'erreur standard $E_e(x_i)$ sont définies comme suit :

$$E_s(x_i) = - \sum P(x_i) \ln P(x_i)$$

$$E_e(x_i) = \sum P(x_i) \ln P(x_i)^2$$

Où x_i sont les signaux temporels discrets, $P(x_i)$ est la distribution de la totalité du signal.

Des approches plus intéressantes dans le domaine temporel appliquent les modèles des séries temporelles aux signaux. L'idée principale de la modélisation des séries temporelles est d'ajuster les données de forme d'onde à un modèle paramétrique et d'extraire ensuite les caractéristiques en se basant sur ce modèle [19]. Le modèle autorégressif (AR), le modèle autorégressif moyenne mobile (ARMA) et le modèle autorégressif moyenne mobile intégrée (ARIMA) sont parmi les techniques de modélisation des séries temporelles les plus favorisées.

Un modèle ARMA d'ordre (p, q) peut être exprimé par,

$$x_t = \alpha_1 x_{t-1} + \dots + \alpha_p x_{t-p} + \varepsilon_t - b_1 \varepsilon_{t-1} - \dots - b_q \varepsilon_{t-q}$$

où x est le signal forme d'onde, ε est une variable indépendante normalement distribuée avec une moyenne « 0 » et une variance constante σ^2 . α et b sont des coefficients du modèle.

Une application [20] du modèle AR a été réalisée sur les signaux vibratoires recueillis d'un moteur asynchrone en utilisant les coefficients du modèle AR comme caractéristiques pertinentes extraites.

Les auteurs [21] ont utilisé des modèles AR adaptatifs pour traiter les signaux vibratoires non-stationnaires et ont constaté qu'ils sont en mesure de fournir des informations fiables dans le domaine temps-fréquence pour la surveillance conditionnelle.

Les auteurs [22] ont comparé les performances de trois techniques de modélisation AR des séries temporelles: modèle AR, réseaux de neurones à rétro-propagation (BP) et réseaux de neurones à base de fonctions radiales et ont rapportés que les performances des réseaux de neurones à rétro-propagation dépassent les deux autres.

Une représentation d'un modèle AR du modèle d'espace d'état a été mise en œuvre pour analyser des signaux de vibration [23].

Il existe de nombreuses autres techniques d'analyse dans le domaine temporel pour traiter les données de forme d'onde pour des fins de diagnostic des machines en défaut.

Malheureusement, l'approche dans le domaine temporel seul est souvent incapable d'identifier le composant défectueux et donc insuffisante pour diagnostiquer la plupart des problèmes de la machine.

II. 5. 3 TECHNIQUES DE TRAITEMENT DU SIGNAL DANS LE DOMAINE FREQUENTIEL

Les techniques dans le domaine fréquentiel sont basées sur le fait qu'un défaut localisé génère un signal périodique avec une fréquence caractéristique unique [15]. Les méthodes d'analyse dans le domaine fréquentiel sont en mesure de surmonter les inconvénients de l'analyse dans le domaine temporel mentionnés dans le paragraphe précédent.

L'extraction de l'amplitude par la transformée de Fourier est la plus recherchée dans l'analyse spectrale et peut être utilisée comme indicateur de défaut. Généralement, si cet indicateur est sous un seuil toléré, la machine peut être considérée en état normal. Autrement, ces fréquences sont analysées et reliées à certains types de défauts pour être isoler ultérieurement. Les amplitudes peuvent aussi être vérifiées à des bandes de fréquences particulières.

L'information du spectre de phase, malgré qu'elle ne soit pas couramment utilisée comme celle de l'amplitude, elle apporte des informations importantes et complémentaires. Par exemple, [24] utilise l'angle de phase pour déterminer l'endroit du crack (fissure) d'engrenage.

Lors de l'utilisation des paramètres dans le domaine fréquentiel comme indicateurs de défauts, le diagnostic primaire est disponible par la décomposition rapide d'un signal complexe en parties plus simples. Des changements dans les paramètres du domaine fréquentiel indiquent l'apparition des défauts car des défauts différents ont des spectres différents. Les paramètres dans le domaine fréquentiel peuvent également être utilisés comme indices pour la détection précoce des défauts et des défaillances de la machine. Par conséquent, ces indices peuvent

être utilisés pour effectuer la surveillance de l'état de santé du composant, le diagnostic de défaut ainsi que le pronostic.

La technique conventionnelle dans le domaine fréquentiel est l'analyse spectrale utilisant la transformée de Fourier rapide (FFT). Pour renforcer les résultats de l'analyse du spectre, les filtres de fréquence, la démodulation, l'analyse de la structure de bande latérale et la présentation graphique sont souvent utilisés. Différents types de spectres de fréquences telles que le spectre de puissance, le Cepstre et les spectres d'ordre élevé ont été développés.

Pour un signal donné, le spectre de puissance donne un tracé de la partie de la puissance d'un signal (énergie par unité de temps) à des fréquences données. La façon la plus usuelle pour générer un spectre de puissance est l'utilisation de la transformée de Fourier discrète, mais d'autres techniques telle que la méthode d'entropie maximale peuvent également être utilisées. Les paramètres suivants, dans le domaine de fréquence, sont couramment utilisés comme indicateurs de défaut.

$$\text{fréquence Centrale (FC)} = \frac{\int_0^{+\infty} f s(f) df}{\int_0^{+\infty} s(f)}$$

$$\text{Fréquence quadratique Moyenne (MSF)} = \frac{\int_0^{+\infty} f^2 s(f) df}{\int_0^{+\infty} s(f)}$$

$$\text{Racine de Fréquence quadratique Moyenne (RMSF)} = \sqrt{MSF}$$

$$\text{Variance de fréquence (VF)} = \frac{\int_0^{+\infty} (f - FC)^2 s(f) df}{\int_0^{+\infty} s(f)}$$

$$\text{Racine de Variance de Fréquence (RVF)} = \sqrt{VF}$$

où $s(f)$ est le spectre de puissance du signal. Le FC , MSF et $RMSF$ montrent le changement de position des principales fréquences. Le VF et RVF décrivent la convergence de la puissance du spectre.

Le spectre d'ordre élevé tels que le Bispectre ou le Trispectre sont capables d'extraire plus d'informations pour les signaux non gaussiens que le spectre de puissance [19].

Les auteurs [25] ont proposé l'ordre bi-spectre et ont affirmé qu'il permet l'élimination du problème des bavures et la modulation qui se posent souvent dans le spectre conventionnel de puissance et le bi-spectre.

L'un des problèmes de mesure d'un signal de défaut est l'incapacité de détecter les signaux de petites amplitudes. Un signal de défaut dans le domaine temporel est très étroit, ce qui entraîne une propagation de la composante d'énergie sur un large intervalle de fréquences; par conséquent, les amplitudes harmoniques de la fréquence de défaut sont enterrées dans le bruit. La technique enveloppe est utilisé dans le but de renforcer les petits signaux. Ce procédé sépare d'abord les signaux de fréquences élevées de ceux des fréquences basses par un filtre passe-bande.

La technique Moyennée en domaine fréquentiel peut être répartie en deux types : la Moyennée synchrone et Moyennée spectrale. La Moyennée synchrone est très utile dans la réduction de la composante aléatoire de bruit de mesure ou la réduction de l'effet des signaux parasites tels que les composantes de bruit à proximité d'une machine.

A la différence de la Moyennée synchrone, la Moyennée spectrale ne réduit pas le bruit. Par contre, elle détermine l'amplitude moyenne à chaque fréquence. Une série de spectres individuels sont additionnés et la somme est divisée par le nombre de spectres.

Le Cepstre a la capacité de détecter les harmoniques et les formes de bandes latérales dans le spectre de puissance. L'application du Cepstre de puissance pour la détection du défaut de machine est basée sur la capacité à détecter la périodicité dans le spectre telle que la famille des harmoniques et les bandes latérales uniformément espacées tout en étant insensible à la voie de transmission du signal depuis la source interne au point de mesure extérieur. La valeur du pic principal du Cepstre s'avère un excellent paramètre de tendance. Elle représente la moyenne sur un grand nombre d'harmoniques individuels et de fluctuations suite à des variations de charge par exemple. La valeur de Cepstre largement moyenné donne une courbe de tendance lisse et monotone avec le temps.

Des exemples de détection d'impulsions d'excitation en utilisant le Cepstre ont été présentés par les auteurs [26].

La technique de résonance à haute fréquence « High Frequency Resonance Technique » (HFRT), développée en 1970 comme un outil d'identification des défauts précoces dans les roulements, tire parti du fait que la majeure partie de l'énergie du signal généré par un défaut est concentrée dans la zone de résonance à haute fréquence, et elle peut fournir des signaux enveloppe avec un rapport signal-bruit (SNR) élevé associés avec la périodicité d'un signal du roulement défectueux. Un procédé adaptatif d'annulation du bruit a été également proposé pour améliorer le spectre d'enveloppe obtenue par HFRT [27].

Toutefois, l'approche dans le domaine fréquentiel, comme toutes les autres techniques, n'est pas sans inconvénient. Elle n'est pas efficace quand il s'agit de signaux de forme d'onde non-stationnaires qui sont très fréquents dans les machines défectueuses.

II. 5. 4 TECHNIQUES DE TRAITEMENT DU SIGNAL EN DOMAINE TEMPS-FREQUENCE

Les techniques temps-fréquence étudient les signaux de forme d'onde à la fois dans le temps et dans la fréquence. Elles abordent donc le problème rencontré dans l'analyse du domaine fréquentiel lorsque les signaux sont non-stationnaires.

Les distributions temps-fréquence les plus utilisées sont la transformée de Fourier de courte durée (STFT) et la distribution de Wigner-Ville.

L'idée de la STFT consiste à diviser le signal forme d'onde en segments avec fenêtre de courte durée, puis d'appliquer la transformation de Fourier pour chaque segment. Le spectrogramme a une certaine limitation dans la résolution temps-fréquence en raison de la segmentation du signal. Il peut être appliqué à des signaux non stationnaires avec une dynamique lente.

Les transformées bilinéaires, telle que la distribution Wigner-Ville, ne sont pas fondées sur la segmentation du signal et donc peuvent surmonter ainsi la limitation de la résolution temps-fréquence du spectrogramme. Cependant, il existe un inconvénient majeur de la transformée bilinéaire en raison des interférences formées par la transformation elle-même. Ces termes d'interférence ou termes croisés rendent l'interprétation de la distribution estimée difficile [28]. Des transformées améliorées telles que la distribution Choi-Williams ont été développées pour remédier à cet inconvénient.

Toujours pour surmonter l'inconvénient de la STFT, une autre technique temps-fréquence, qui est la transformée en ondelettes, a été développée. Alors que STFT donne une résolution constante à toutes fréquences, la transformée en ondelettes utilise la technique multi-résolution dont laquelle des fréquences différentes sont analysées avec des résolutions différentes.

Mathématiquement, la transformée en ondelettes continue *CWT* de la fonction $f(t)$ peut s'écrire sous la forme

$$CWT(a, b) = \int_{-\infty}^{\infty} f(t)\psi_{a,b} dt$$

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

où a , b et $\psi_{a,b}$ sont le paramètre d'échelle, de translation et l'ondelette mère, respectivement. *CWT* fournit un outil puissant d'analyse multi-résolutions temps-fréquence pour révéler les caractéristiques transitoires des signaux non stationnaires. La *CWT* peut décomposer le signal inspecté en famille de fonctions élémentaires. Cette capacité rend l'analyse du signal inspecté plus facile.

La transformée en ondelettes discrète (*DWT*), qui est basée sur un codage en sous-bandes, est mise en œuvre pour donner un calcul rapide de la transformée en ondelettes. Les fonctions de base orthogonale utilisées dans l'analyse en ondelettes sont des familles de fonction de dilatation, $\phi(t)$ et d'ondelette associée, $\psi(t)$. La fonction de dilatation peut être représentée par l'expression mathématique suivante

$$\phi_{j,k}(t) = \sum_k H_k \phi(2^j t - k)$$

où H_k représente le coefficient de la fonction de dilatation, k , j représentent le facteur de translation et de dilatation, respectivement. De même, l'ondelette associée peut être générée en utilisant le même coefficient comme la fonction de dilatation.

$$\psi_{j,k}(t) = \sum_k (-1)^k \sqrt{2} h_{1-k} \phi(2^j t - k)$$

La fonction de dilatation est orthogonale entre elle ainsi qu'à la fonction d'ondelette comme le montre les équations ci-dessous. Ce fait est essentiel et fait partie du cadre d'analyse multi-résolution.

$$\int_{-\infty}^{\infty} \phi(2k - t)\phi(2k - t) dt = 0$$

$$\int_{-\infty}^{\infty} \psi(t)\phi(t) dt = 0$$

Par utilisation d'un procédé itératif, la fonction de dilatation et l'ondelette associée peuvent être calculées si les coefficients sont connus.

Un signal peut être décomposé en coefficients approximatifs $a_{j,k}$ à travers le produit scalaire du signal original à l'échelle j et la fonction de dilatation.

$$a_{j,k} = \int_{-\infty}^{\infty} f_j(t) \phi_{j,k}(t) dt$$

$$\phi_{j,k}(t) = 2^{-j/2} \phi(2^{-j} t - k)$$

De même, le coefficient de détail $d_{j,k}$ peut être obtenu par le produit scalaire du signal au conjugué complexe de la fonction ondelette.

$$d_{j,k} = \int_{-\infty}^{\infty} f_j(t) \psi_{j,k}(t) dt$$

$$\psi_{j,k}(t) = 2^{-j/2} \phi(2^{-j} t - k)$$

Le signal d'origine peut, cependant, être décomposé à différentes échelles comme suit

$$f(t) = \sum_{j=-\infty}^{\infty} a_{j_0,k} \phi_{j_0,k} + \sum_{j=-\infty}^{j_0} \sum_{k=-\infty}^{\infty} d_{j,k} \psi_{j,k}(t)$$

$$f[n] = \sum_{k=0}^{N-1} a_{j,k} \phi_{j,k}(t) = \sum_{k=0}^{N-1} a_{j+1,k} \phi_{j+1,k}(t) + \sum_{k=0}^{N-1} d_{j+1,k} \psi_{j+1,k}(t)$$

Le coefficient du niveau de décomposition suivant $(j + 1)$ peut être exprimé comme suit :

$$a_{j+1,k} = \sum_{k=0}^N a_{j,k} \int \phi_{j,k}(t) \phi_{j+1,k}(t) dt$$

$$d_{j+1,k} = \sum_{k=0}^N a_{j,k} \int \phi_{j,k}(t) \psi_{j+1,k}(t) dt$$

$$a_{j+1,k} = \sum_k a_{j,k} g[k] \quad \text{et} \quad d_{j+1,k} = \sum_k a_{j,k} h[k]$$

Les coefficients de décomposition peuvent être déterminés par convolution et mis en œuvre à l'aide d'un filtre. Le filtre $g[k]$ est un filtre passe-bas et $h[k]$ est un filtre passe-haut.

$$y[n] = \sum_{k=1}^N h[k] x[n - k]$$

L'étude comparative de l'efficacité et de la fiabilité de la transformée en ondelettes par rapport à d'autres techniques d'analyse de signal de vibration est présentée dans [29].

Les techniques de transformée en ondelettes ont été utilisées avec succès dans diagnostic de défaut, comme les réducteurs [30], les roulements [31].

[32] utilisent la transformée en ondelettes *CWT* pour le traitement des signaux de vibrations des roulements à des fins pronostics.

II. 6 LES METHODES ACTUELLES DE DIAGNOSTIC DE DEFAUT

Le terme *diagnostic de défaut* est généralement utilisé pour décrire un large éventail de tâches et de capacités. Bien que les définitions du terme diagnostic de défaut soient acceptées pour les différentes tâches couvertes par ce terme, elles ne sont pas encore devenues standardisées. Cependant, dans la communauté de CBM, les définitions des termes suivants sont de plus en plus acceptées [33]:

- le diagnostic de défaut concerne la détection, l'isolement et l'identification, d'une condition de défaut imminente, ou naissante dans un système à l'aide d'un raisonnement logique fondé sur un ensemble d'informations provenant d'une inspection, d'un contrôle ou d'un test. Le terme défaut implique que le système sous observation est encore opérationnel, mais ne peut pas continuer à fonctionner indéfiniment sans intervention de maintenance.
- Le diagnostic de défaillance concerne la détection, l'isolement et l'identification d'un système qui a cessé de fonctionner.

Dans les définitions données ci-dessus, la détection, l'isolement et l'identification de défaut impliquent généralement les significations suivantes [34]:

La détection de défaut (défaillance) consiste à identifier l'apparition d'un défaut, ou d'une défaillance, dans un système contrôlé, ou l'identification d'un comportement anormal qui peut être signe d'une condition de défaut.

L'isolement de défaut (défaillance) consiste à identifier quel composant/sous-système/système a un défaut, ou une défaillance.

L'identification de défaut (défaillance) consiste à déterminer la nature et l'ampleur d'un défaut du système ou d'une défaillance

Une fois le type défaut est identifié et selon la connaissance disponible sur le système, il est parfois possible de propager les effets d'un défaut sur les équipements du système afin de prédire les conséquences [3].

Pour comprendre le domaine d'application du diagnostic et du pronostic considéré, la figure II. 3, illustre la chronologie de la progression du défaut d'un composant de système typique [33]. Au début de la durée de vie du composant, il est considéré comme étant en bon état, après un certain temps de fonctionnement, un état de défaut naissant se développe. Ainsi, la sévérité de défaut augmente jusqu'à ce que le composant soit finalement défaillant. Si le système est autorisé de continuer à fonctionner, il est probable que d'autres dommages puissent être causés à d'autres composants ou systèmes secondaires.

Selon [33], le domaine d'application du diagnostic a généralement eu lieu au moment de la défaillance du composant, ou sur l'intervalle entre la défaillance du composant et la défaillance éventuelle de l'ensemble du système. Cependant, si une condition de défaut peut être détectée à un stade embryonnaire, des actions de maintenance peuvent être retardées jusqu'à ce que le défaut progresse à un état plus avancé (grave), mais toujours avant la défaillance. Cet intervalle entre la détection d'une condition de défaut naissant et la survenue d'une défaillance, nit le domaine de pronostic. Un intervalle suffisant entre la détection de défaut naissant et défaillance du système, permet un éventail de fonctions opérationnelles et des travaux de maintenance corrective qui peuvent être planifiés à l'avance, avec les ressources et le personnel nécessaires alloués selon les besoins [34].

Le positionnement relatif des processus de « détection », de « diagnostic » et de « pronostic » est schématisé sur la figure II. 4-a.

- la détection vise à identifier le mode de fonctionnement du système, son état,
- lorsqu'une défaillance est apparue, le diagnostic permet d'isoler et d'identifier le composant qui a cessé de fonctionner (propagation arrière : des effets vers les causes),
- le pronostic vise la prédiction des états futurs du système (propagation avant : des causes vers les effets).

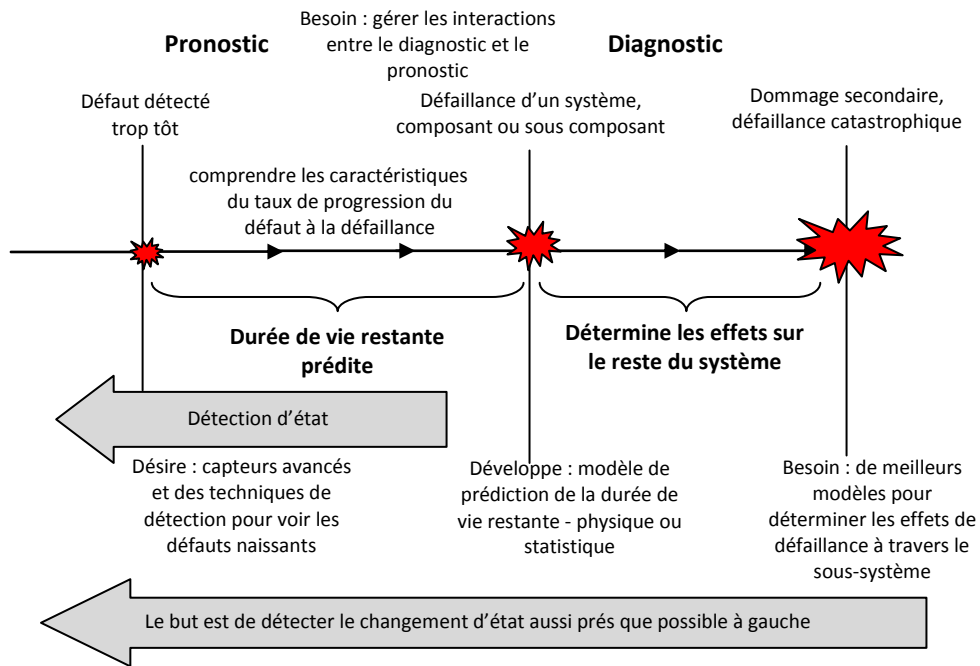
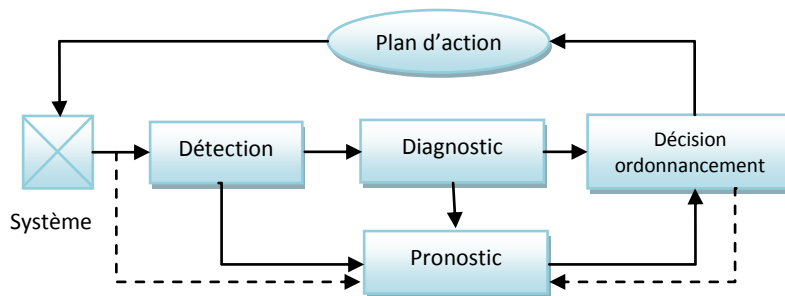
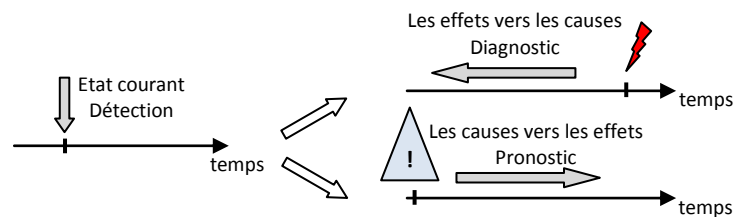


Figure II. 3: Chronologie de progression de défaillance d'après [33]



a- Détection, diagnostic, pronostic et ordonnancement dans les systèmes de maintenance intelligente



b- Détection, diagnostic et pronostic les aspects phénoménologiques
Figure II. 4 : Complémentarité des activités de détection, diagnostic, pronostic

II. 7 CLASSIFICATION DES METHODES DE DIAGNOSTIC

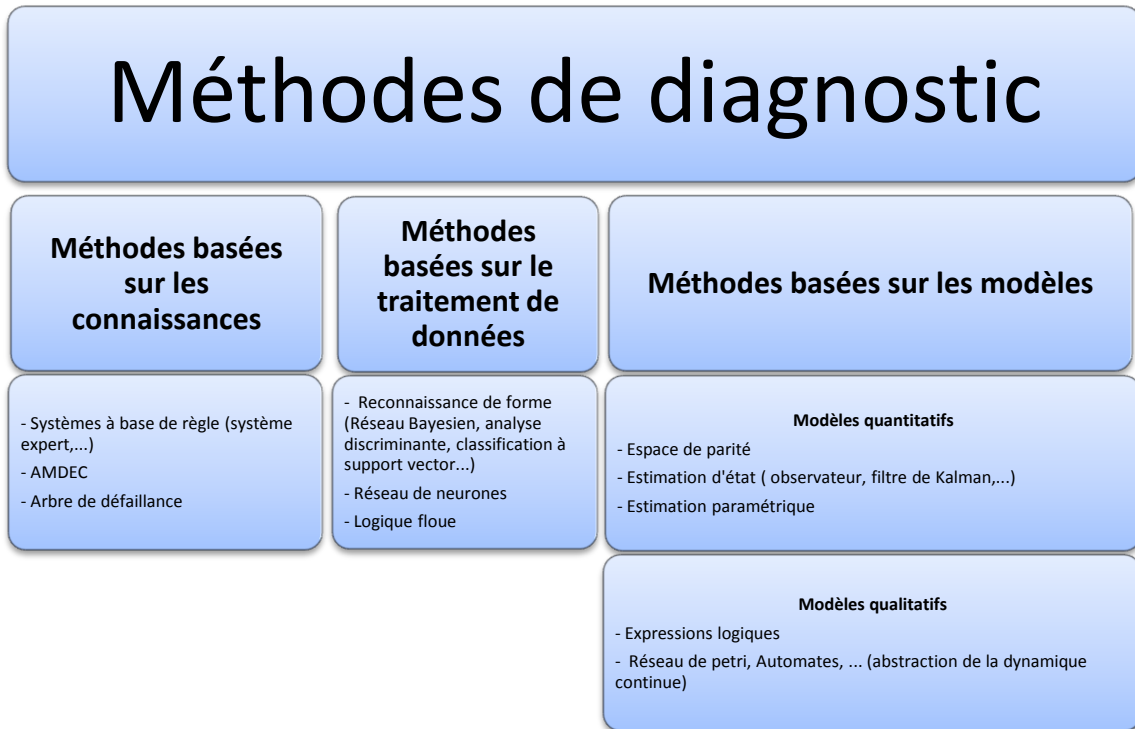


Figure II. 5 : Classification des méthodes de diagnostic

II. 7. 1 APPROCHES BASEES SUR LA CONNAISSANCE

Ce type d'approche utilise une connaissance explicite de relations causales entre les symptômes, les défauts et les défaillances. Cette approche associe directement un symptôme à un défaut qui en est la cause. La connaissance n'est pas extraite d'un modèle explicite structurel ou de comportement du système. Elle est souvent acquise durant la phase de conception et provient d'une analyse fonctionnelle et structurelle du système [3].

Cette source de connaissance peut résulter d'une analyse des modes de défaillance et de leurs effets (AMDE) ou bien d'un historique des mauvais fonctionnements du système représenté par un arbre de défaillances par exemple. L'analyse des modes de défaillance et de leurs effets et les arbres de défaillances sont utilisées pour identifier les causes des défaillances possibles d'un système. Cette connaissance, qui se présente sous la forme d'associations entre effets et causes, est dite *externe* ou de *surface*. Des approches classiques de diagnostic qui utilisent ce type de connaissance externe sont des systèmes à bases de règles comme par exemple les systèmes experts [35], [36].

II. 7. 1. 1 ANALYSE DES MODES DE DEFAILLANCE ET DE LEURS EFFETS

L'analyse des modes de défaillance et de leurs effets est une méthode largement utilisée dans de nombreux domaines industriels. C'est une méthode *inductive* qui consiste, pour un système et une défaillance donnée, à étudier de façon détaillée les effets (ou les conséquences) de cette défaillance sur le système lui-même et/ou son environnement. La base de connaissance de cette méthode est définie à partir d'une analyse structurelle et fonctionnelle du système, c'est-

à-dire d'une analyse de ses fonctions et de ses composants. L'ensemble des modes de défaillance possibles des composants doit être établi et pour chaque mode de défaillance, les causes possibles de son apparition sont recherchées. Finalement, une étude des effets sur le système est faite pour chaque combinaison (cause, mode de défaillance).

Les résultats sont présentés sous forme de tableaux comme sur la figure II. 6.

Identification du composant	Fonctions Etats	Modes de défaillance	Cause possibles	Effets	Criticité	Moyens de détection	Parades de l'opérateur humain	Observations

Figure II. 6 : AMDEC

L'analyse des modes de défaillance, de leurs effets et de leur criticité (AMDEC) est une extension de l'AMDE qui inclut une analyse de criticité des modes de défaillance. La criticité permet d'extraire les modes de défaillance les plus critiques.

Les tableaux d'AMDE(C) sont obtenus à partir d'une analyse *déductive*, à partir des causes (des défaillances), on détermine les conséquences qu'elles peuvent avoir sur le système (les effets). L'utilisation des tableaux d'AMDE(C) à des fins de diagnostic conduit à utiliser une procédure *abductive*. Supposant qu'un système est défaillant, une démarche abductive consiste à rechercher les causes pouvant expliquer les effets observés de cette défaillance. Les tableaux de l'AMDE(C) sont utilisés comme un outil d'identification des causes de défaillances à partir des effets observés. Les inconvénients d'une telle méthode sont qu'elle nécessite une longue expérience et que toute modification entraîne une réécriture du tableau [3].

II. 7. 1. 2 ARBRE DE DEFAILLANCE

Les arbres de défaillance sont créés à l'aide d'une procédure déductive optimisée qui détermine des chemins critiques dans un système. Les chemins critiques dans un système correspondent aux diverses combinaisons possibles d'événements qui entraînent la réalisation d'un événement indésirable unique (événement de défaut entraînant une panne du système).

La procédure qui utilise les arbres de défaillance à des fins de diagnostic est *abductive*, elle se focalise d'abord sur les événements indésirables pour identifier ensuite leurs causes. Un arbre de défaillance est établi sous la forme d'un diagramme logique et comporte au sommet l'événement indésirable. Les causes immédiates qui produisent cet événement sont ensuite hiérarchisées à l'aide de symboles logiques « ET » et « OU ».

Pour exécuter un diagnostic correct à partir des arbres de défaillances, ceux-ci doivent largement représenter toutes les relations causales du système, capable d'expliquer tous les scénarios de défauts possibles. Cette méthode déductive s'utilise difficilement pour les systèmes qui dépendent du temps [3].

II. 7. 1. 3 SYSTEMES EXPERTS

La technique des systèmes experts est la plus répandue pour la supervision des systèmes complexes. Les systèmes experts sont des outils de l'intelligence artificielle, utilisés lorsqu'aucune méthode algorithmique exacte n'est disponible ou possible. La propriété principale de ces systèmes est de pouvoir représenter et restituer les connaissances acquises par un expert. Dans la plupart des cas, les connaissances utilisées pour le développement d'un

système expert d'aide au diagnostic, reposent sur l'apprentissage des relations entre les causes et les effets observés pour chaque défaillance du système [37]. Un système expert est composé de deux parties indépendantes.

- Une base de connaissance qui est elle-même composée d'une base de faits qui contient les informations, les données concernant le cas traité et d'une base de règles connues qui modélisent la connaissance du domaine considéré.
- Un moteur d'inférence capable de raisonner à partir des informations contenues dans la base de connaissance et de faire des déductions.

Le moteur d'inférence utilise les données et les règles pour produire de nouvelles données. Le rôle d'un système expert est donc d'inférer des règles du type :

$$SI [A = "vrai"] ET [A implique B] ALORS [B = "vrai"].$$

Au fur et à mesure que les règles sont appliquées, des nouveaux faits se déduisent et se rajoutent à la base de faits. Le diagnostic par systèmes experts se fonde sur l'expérience disponible sur le système pour construire une table de correspondance qui associe les observations aux diagnostics correspondants [35]. Ces techniques sont efficaces au niveau du temps de calcul, elles sont peu coûteuses et leur implémentation est simple. L'expérience dont dépendent ces approches est cependant difficile à acquérir et si le système évolue, les règles sont à remettre en cause. Une nouvelle expertise est alors nécessaire.

II. 7. 2 APPROCHES BASEES SUR LE TRAITEMENT DE DONNEES

Dans ces méthodes, les seules informations disponibles sont les signaux issus des capteurs positionnés sur le système. Les capteurs sont supposés fiables et leurs valeurs correctes. L'objectif de ces méthodes est d'associer un ensemble de mesures à des états de fonctionnement connus du système. Ces approches font appel à des méthodes de reconnaissance de formes qui utilisent des techniques d'apprentissage numérique et de classification afin d'établir un modèle de référence du système fondé sur l'expérience (exploitation des données, des mesures sous la forme d'historique). Le modèle établi ne provient donc pas d'une spécification du système durant la phase de conception et ne repose pas sur une connaissance physique du système [38]. Ce modèle de référence saisit le comportement normal du système et est utilisé pour la détection et le diagnostic.

Le principe des méthodes de reconnaissance de formes est illustré par la figure II. 7.

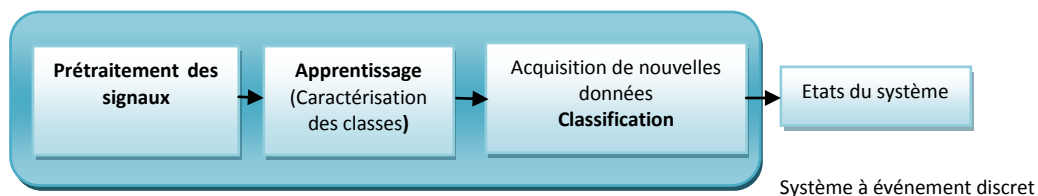


Figure II. 7: Diagnostic par reconnaissance de forme

Un prétraitement des signaux issus des capteurs permet d'établir les variables les plus pertinentes pour déterminer les états de fonctionnement du système et d'éliminer les bruits possibles à l'aide de techniques de filtrage.

Les méthodes de reconnaissance de formes consistent à reconnaître des formes parmi différentes possibilités à partir d'observations bruitées dans le but d'identifier les états du système (état de fonctionnement normal, état de défaillance, ...). Un prototype est défini par des valeurs précises de l'ensemble des paramètres (ou caractéristiques) d'une situation particulière (défaillance).

L'objectif d'une méthode de classification est d'identifier à partir de mesures du système les différentes classes en regroupant les données qui ont des valeurs de paramètres (caractéristiques) similaires. Ces classes contiennent l'information qui caractérise les états et les défaillances du système et sont représentées par leur prototype qui définit le mode de fonctionnement du système [3] :

Classe x_1 → Prototype P_1 = « Fonctionnement normal »,

Classe x_2 → Prototype P_2 = « Mode défaut 1 »,

.....

Classe x_n → Prototype P_n = « Mode défaut n ».

Une fois les classes identifiées, la méthode de classification peut évaluer la distance entre une forme particulière et son prototype.

Des techniques d'apprentissage permettent d'établir les valeurs caractéristiques des paramètres de chaque classe et d'évaluer le prototype associé aux nouvelles observations dans le but de déterminer l'état de fonctionnement du système. Elles réalisent donc un partitionnement de l'espace de représentation des données en déterminant les frontières entre les classes. Avec une technique d'apprentissage supervisée, il est possible de créer en ligne de nouvelles classes afin d'identifier des situations non prévues.

Bien qu'il existe plusieurs techniques pour la reconnaissance de formes, mais les plus couramment utilisées sont: la reconnaissance de formes par réseaux de neurones et la reconnaissance de formes par la logique floue.

II. 7. 2. 1 TECHNIQUES NEURONALES

Un réseau de neurones est un modèle de calcul dont la conception est inspirée du fonctionnement des vrais neurones. Ils sont utilisés pour la classification des données et des formes. Les réseaux de neurones sont optimisés par des méthodes d'apprentissage par expérience [39] [40]. Les poids des neurones sont ajustés lorsqu'on leur présente de nouvelles données (observations) à traiter.

II. 7. 2. 2 TECHNIQUES FLOUES

La logique floue introduite par [41] est issue de la théorie mathématique des ensembles flous qui considère des ensembles définis de manière graduelle. A l'inverse de la logique booléenne, la logique floue associe à une donnée un degré d'appartenance à un ensemble qui peut être différent d'un état booléen 0 ou 1.

Le modèle de référence utilisé par ces méthodes de classification est le résultat d'un processus d'apprentissage. Aucune garantie ne peut être donnée quant à la complétude, la cohérence et la précision du modèle. De plus, la phase d'apprentissage nécessite qu'un grand nombre de données soient disponibles (données collectées à partir des mesures réelles).

II. 7. 3 APPROCHES A BASE DE MODELES

Les approches de diagnostic à base de modèles reposent sur une connaissance physique profonde du système à diagnostiquer. Le système est représenté sous forme d'un ou plusieurs modèles qui décrivent la structure du système et son comportement nominal ou encore son comportement en présence d'un défaut. La méthode de diagnostic s'appuie sur la comparaison du comportement réel observé sur le système physique avec le comportement prédit à l'aide de modèles. La détection d'incohérences permet de conclure sur l'occurrence de défaut dans le système. Un modèle de dysfonctionnement (modèle de défaut) permet de localiser les défauts et éventuellement de les identifier. La figure II. 8 illustre les concepts de base d'une approche typique à base de modèle pour le diagnostic.

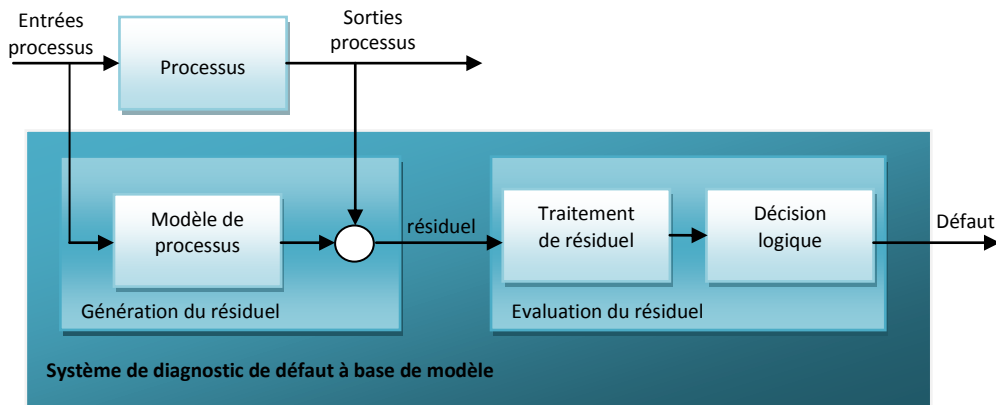


Figure II. 8 : Approche à Base de Modèles

Deux principales approches peuvent se distinguer dans les méthodes de diagnostic à base de modèles. L'approche FDI « Fault Detection and Isolation » issue de la communauté automatique utilise des modèles *quantitatifs* pour décrire le modèle de comportement du système. Des équations algèbro-différentielles permettent de représenter le comportement continu du système avec une certaine précision numérique. L'approche DX, fondée sur une théorie logique du diagnostic, provient de la communauté de l'intelligence artificielle. Elles utilisent des modèles *qualitatifs* qui permettent de représenter de manière efficace les interactions entre composants ou systèmes.

II. 7. 3. 1 APPROCHES FDI

Les approches FDI reposent sur une connaissance approfondie du fonctionnement du système. Elles utilisent des modèles de référence quantitatifs qui peuvent être obtenus à partir de lois fondamentales de la physique [42]. Ces modèles analytiques numériques décrivent le comportement normal du système par un ensemble d'équations différentielles [43].

Le diagnostic consiste à générer des indicateurs de présence de défaut dans le système que l'on appelle des *résidus*. Pour obtenir ces résidus, le modèle de référence est restreint aux entrées et sorties observables (mesurées par capteurs sur le système réel). Une valeur non nulle d'un résidu est interprétée comme une modification anormale (ou une déviation inacceptable) d'une propriété ou d'un paramètre caractéristique du système modélisé.

Les équations différentielles du modèle qui décrivent le comportement du système peuvent être linéaires ou non linéaires [44]-[45], à temps discret ou à temps continu. Le système modélisé peut également être un système hybride [46]-[48].

A cause des erreurs de modélisation et de la présence de bruit, les résidus ne sont jamais réellement nuls même s'il n'y a pas de défaut dans le système. La décision d'une détection nécessite d'évaluer les expressions des résidus obtenus en utilisant les mesures du système afin de déterminer si la différence à laquelle ils sont associés est significative.

Pour l'évaluation des résidus, des techniques de reconnaissance de formes, de logique floue, de seuillage sont utilisées ou bien encore des techniques issues de la théorie de décision statistique [49]. Le principe d'une approche FDI basée sur le calcul des résidus est illustré sur la figure II. 9.

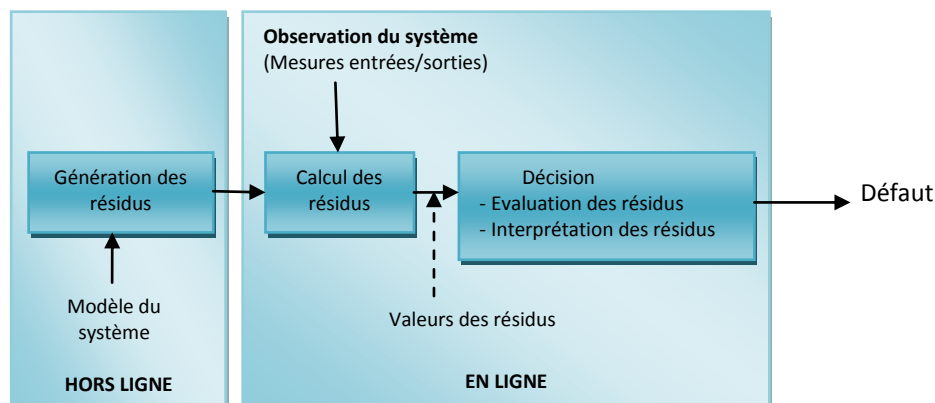


Figure II. 9: Approche FDI à base de résidus

Pour obtenir les expressions analytiques des résidus, plusieurs techniques peuvent être utilisées : des relations de parité, des observateurs, des filtres de détection. D'autres méthodes s'intéressent plutôt à l'estimation de paramètres pour la détection de défaut.

Plusieurs articles de synthèse comme [50] décrivent l'ensemble de ces méthodes utilisées dans les approches FDI.

a) ESPACE DE PARITE

Les approches de l'espace de parité reposent sur la génération de résidus [51]. Le comportement normal (nominal) d'un système est caractérisé par les valeurs nulles des résidus. Les mesures sont alors cohérentes avec les propriétés de l'espace de parité. La conception de l'espace de parité est basée sur le développement des expressions analytiques.

L'approche consiste à déterminer des redondances analytiques parmi les entrées et les sorties du système [52] [53]. Les relations de parité utilisent la redondance directe au moyen de relations algébriques liant les différentes grandeurs du modèle du système. Les relations de redondance analytique sont obtenues en éliminant les variables d'état non observables du modèle sur l'espace de parité. On obtient un modèle entrée/sortie du système dans lequel le vecteur d'état x inconnu n'apparaît pas [43]. Les équations obtenues n'impliquent donc que des variables entrées/sorties mesurables, donc connues. Les relations de redondance analytiques sont ensuite directement utilisées pour construire des indicateurs de défaut pour le diagnostic.

[54] et [55] sont des références d'applications industrielles de l'espace de parité pour la détection avec des relations statiques. Des applications industrielles sur des systèmes dynamiques sont développées dans [56] dans le domaine de l'automobile. Quand certaines sorties ne sont pas directement mesurables, des techniques d'observateurs complètent la méthode de l'espace de parité.

b) ESTIMATION D'ETAT : OBSERVATEURS

Les observateurs sont des systèmes dynamiques qui peuvent être utilisés pour la détection et l'isolation de défaut dans des systèmes linéaires ou non linéaires. La méthode des observateurs consiste à reconstruire à partir d'un modèle analytique et d'un ensemble d'observations partielles du système (entrées/sorties) les sorties non mesurables du système par une estimation de l'état du système. Les observateurs permettent d'estimer les valeurs des variables d'état. A partir de ces valeurs reconstruites et du modèle dynamique, les valeurs des sorties sont calculées.

Un vecteur de résidus est obtenu représentant la différence entre les mesures des sorties et les valeurs des sorties estimées. Plusieurs travaux traitent de la détection de défaut à l'aide d'observateurs dans le cas des systèmes linéaires [54], [57]. [44] et [58] proposent une solution pour observer des systèmes non linéaires.

c) ESTIMATION PARAMETRIQUE

L'estimation paramétrique permet d'analyser l'influence des défauts sur les paramètres structuraux du modèle du système. C'est donc une approche naturelle qui permet de détecter et d'isoler les défauts multiplicatifs sur les paramètres du modèle dynamique du système.

L'estimation paramétrique utilise des méthodes analytiques pour calculer les valeurs des paramètres du modèle dont la structure est connue. Ces paramètres structuraux sont généralement constants.

Pour estimer ces paramètres, des techniques de filtrage sont utilisées. Par exemple, le filtre de Kalman estime de manière récursive les paramètres structuraux d'un modèle linéaire sous forme de représentation d'état à partir de mesures bruitées. Pour les systèmes non linéaires mais linéarisables localement, un filtre de Kalman étendu peut être utilisé et pour les systèmes non linéaires et non linéarisable, un filtrage particulière est appliqué (simulation et technique statistiques).

La détection d'un défaut se fait en comparant les paramètres estimés avec les paramètres nominaux qui caractérisent le comportement normal du système. L'idée principale d'une telle méthode consiste à minimiser la distance entre les valeurs réelles du paramètre et les valeurs estimées. La méthode d'estimation est donc caractérisée par la définition de distance (maximum de vraisemblance, moindres carrés, ...).

Dans [59], les approches d'estimation paramétrique sont comparées à la méthode d'espace de parité.

Les techniques présentées ici sont les plus connues pour la génération de résidus. Des équivalences entre les techniques utilisant l'espace de parité, les observateurs et l'estimation paramétrique ont été établies dans [52].

d) LOCALISATION ET IDENTIFICATION DES DEFAUTS

Pour localiser les défauts, différents types de résidus sont générés : les *résidus structurés* (propriétés booléennes) et les *résidus directionnels* (propriétés géométriques).

Les résidus structurés sont conçus pour que chaque résidu soit sensible à un sous ensemble de défauts connus et insensible aux autres défauts. Lorsqu'un défaut connu apparaît, la valeur de certains résidus est nulle ou proche de zéro tandis que d'autres résidus seront différents de zéro. L'ensemble des valeurs des différents résidus représente la signature du défaut. L'ensemble des signatures pour les différents défauts connus pouvant apparaître dans le système est appelé la matrice de signatures.

Les deuxièmes types de résidus permettant la localisation des défauts sont les résidus directionnels. Ils sont représentés sous la forme d'un vecteur de résidus orienté selon une direction particulière de l'espace des résidus selon le type de défaut apparu. La connaissance sur les différents défauts ou les modes de défaillances d'un système représenté par un modèle de dysfonctionnement permet l'identification des défauts [3].

II. 7. 3. 2 APPROCHE DX

L'approche DX est une approche qualitative basée sur la cohérence qui provient du domaine de l'intelligence artificielle. La technique du diagnostic de cohérence consiste à comparer le comportement réel du système observé et son comportement attendu tel qu'il peut être prédit grâce à des modèles de bon comportement. Elle repose sur la théorie logique du diagnostic introduite par [60]. Le but du diagnostic logique est de déterminer les équipements (ou composants) du système dont le fonctionnement anormal peut expliquer les incohérences détectées entre les comportements prédits et les observations du système [61]. Il s'agit d'un diagnostic abductif, le but est de trouver l'ensemble des causes qui expliquent les observations. Le principe de cette approche logique basée sur la cohérence est illustré sur la figure II. 10.

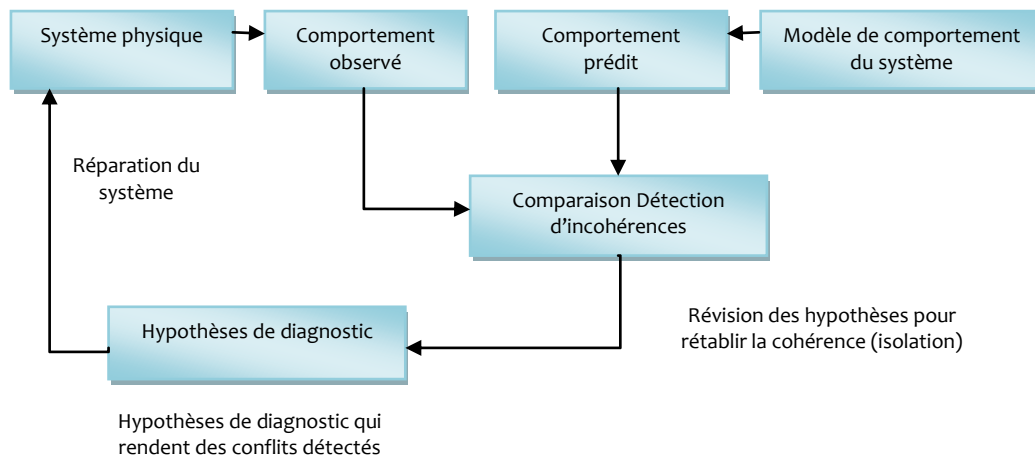


Figure II. 10: Approche logique basée sur la cohérence

L'avantage majeur des approches de diagnostic à base de modèle est la capacité de détecter des défauts non-anticipés, puisque les modèles employés sont toujours basés sur la physique de défaillance. Cela à l'encontre des approches à base de données, qui nécessitent pratiquement des exemples historiques pour chaque condition de défaut à détectée. Cependant, dans des situations réelles, il n'est pas possible d'appliquer des approches de

diagnostic à base de modèle, puisque plusieurs processus sont complexes pour en développer un modèle mathématique précis.

II. 8 PRONOSTIC ET ETAT DE L'ART DES TRAVAUX

II. 8. 1 DEFINITION :

Le terme pronostic provient du grec « progignôskein » signifiant « connaître à l'avance ». Il a ensuite été essentiellement utilisé en médecine où il porte d'une part sur l'évaluation du degré de gravité d'une pathologie, et d'autre part, sur l'estimation de l'évolution ultérieure de la maladie.

Dans le cadre de la maintenance des systèmes industriels, un grand nombre de définitions a été donné proposant des finalités différentes et/ou des informations d'entrée différentes. Cette discordance est particulièrement visible à travers la terminologie anglaise qui comporte plusieurs mots « prognosis-prognostic-prognostics » là où la langue française ne propose que pronostic [2].

Pour [62] « *le pronostic - Prognostic - est la capacité de prédire la condition futur de la machine basée sur l'état diagnostic actuel de la machine et les données de fonctionnement et de défaillance historiques disponibles* ». Cette définition met en évidence la nécessité pour le pronostic de s'intégrer dans son environnement informationnel, pour obtenir les informations (passées ou courantes) nécessaires à sa mise en œuvre.

Pour [63] : « *la fonction primaire du pronostic est de projeter l'état de santé actuel de l'équipement dans le futur en tenant compte des estimations des profils d'utilisation futurs* ». Cette définition apporte un point de vue supplémentaire puisqu'elle indique qu'il faut prendre en compte dans le pronostic l'état courant du système, mais également des connaissances disponibles sur les utilisations futurs du système.

« Prognostics » est un mot dont l'utilisation en maintenance a récemment émergé dans la littérature. Il est notamment intégré au concept de « Prognostics and Health Management » (PHM) et porte en lui une vision plus système (maintenance de systèmes et système de maintenance). Dans [64] les auteurs définissent les deux termes « Prognostics » et « Prognosis » par:

- « *-Prognostics and Health Management (PHM)- est une approche de gestion de santé utilisant les mesures, les modèles, et les logiciels pour accomplir la détection des défauts précoces, l'évaluation de la condition, et la prédiction de la progression de la défaillance* ».
- « *-Prognosis- est prédire sûrement et précisément la durée de vie restante ou le risque pour une mission planifiée complète. Suivre la tendance de la progression de la santé et produire une RUL basée sur un seuil HI prédéfini. Les limites de tolérances ou l'intervalle de confiance doivent accompagner la prédiction* ».

Ces deux définitions permettent de mesurer la différence entre *prognosis/prognostic* qui concerne le processus de prédiction de la durée de vie résiduelle d'un matériel et *prognostics* qui regroupe l'ensemble des processus permettant d'aboutir à un résultat de pronostic, c'est-à-dire les processus de mesure, de détection, de diagnostic et de pronostic proprement dit. Toujours dans le même sens, [65] définit encore: « *-Prognostics- traite l'utilisation des méthodes automatisées pour détecter, diagnostiquer, et analyser la dégradation des performances physiques d'un système, calculer la durée de vie restante dans un état fonctionnel acceptable avant une défaillance ou une dégradation inacceptable de performance* ».

Enfin, depuis 2004, la norme ISO 13381-1 [66] décrit de façon standard ce que devrait être le processus de pronostic, sans en proposer de modélisation. Le pronostic y est défini par : « *Estimation de la durée de fonctionnement avant défaillance et du risque d'existence ou d'apparition ultérieure d'un ou de plusieurs modes de défaillance* ». Si cette définition ne prend pas en compte la nécessité d'inclure un scénario d'utilisation futur, elle apporte deux aspects importants et complémentaires au pronostic. Tout d'abord elle précise que le résultat du pronostic doit être formulé sous la forme d'une durée de vie résiduelle (Remaining Useful Life : RUL). Cette définition intègre également la notion de mode de défaillance courant et de mode de défaillance potentiel (qui n'aurait de réalité physique que dans le futur) [2], [4].

Comme le pronostic est focalisé sur le futur, les aspects suivants doivent être pris en compte [66] :

- les modes de défaillance existants et les taux de détérioration,
- les critères de déclenchement de futurs modes de défaillance,
- le rôle des modes de défaillance existants dans le déclenchement de futurs modes de défaillance,
- l'influence entre les modes de défaillance existants et les modes futurs et leurs taux de détérioration,
- la sensibilité à la détection et aux modifications des modes de défaillance existants et futurs du fait des techniques de surveillance actuelles,
- la conception et les changements de stratégies de surveillance afin de s'adapter à tous les éléments ci-dessus,
- l'effet des actions de maintenance et/ou des conditions d'exploitation,
- les conditions ou les hypothèses dans lesquelles les pronostics restent valables.

II. 9 CLASSIFICATION DES APPROCHES DE PRONOSTIC

Le premier classement des approches de pronostic a été proposé par [67]. Dans leur papier, les auteurs suggèrent une classification pyramidale à trois niveaux (figure II. 11) : les approches basées sur l'expérience, les approches guidées par les données et les approches basées sur les modèles. Pour dissocier ces trois types d'approches, les critères suivants sont considérés : le coût et la complexité de mise en œuvre, la précision des résultats obtenus, et l'applicabilité des approches.

D'autres travaux ont proposé de nouveaux critères pour classer les approches, ces critères pouvant être fonction du résultat attendu, ou de la nature des informations utilisées par le pronostic. Ainsi, [7] définit une classification basée sur le résultat attendu du pronostic. Ils distinguent les travaux cherchant à obtenir une RUL et ceux cherchant à obtenir la probabilité de défaillance à une date donnée. Toutefois, les auteurs font la constatation que la seconde classe ne contient que peu de travaux de recherche et répartissent la première classe en trois sous-catégories « approches statistiques », « approches intelligentes » et « approches à base de modèle ».

Ainsi [9] propose d'utiliser la nature des entrées. Elle catégorise les modèles basés sur des événements, les modèles basés sur l'état et les modèles intégrant ces deux approches. Les événements représentent principalement les passages dans un état donné (ex. la défaillance) et la classe d'approches associées recense l'ensemble des modèles fiabilistes et stochastiques de

la dégradation. Les pronostics basés sur l'état font l'objet d'une sous classification : pronostics basés sur les modèles physiques et pronostics guidés par les données.

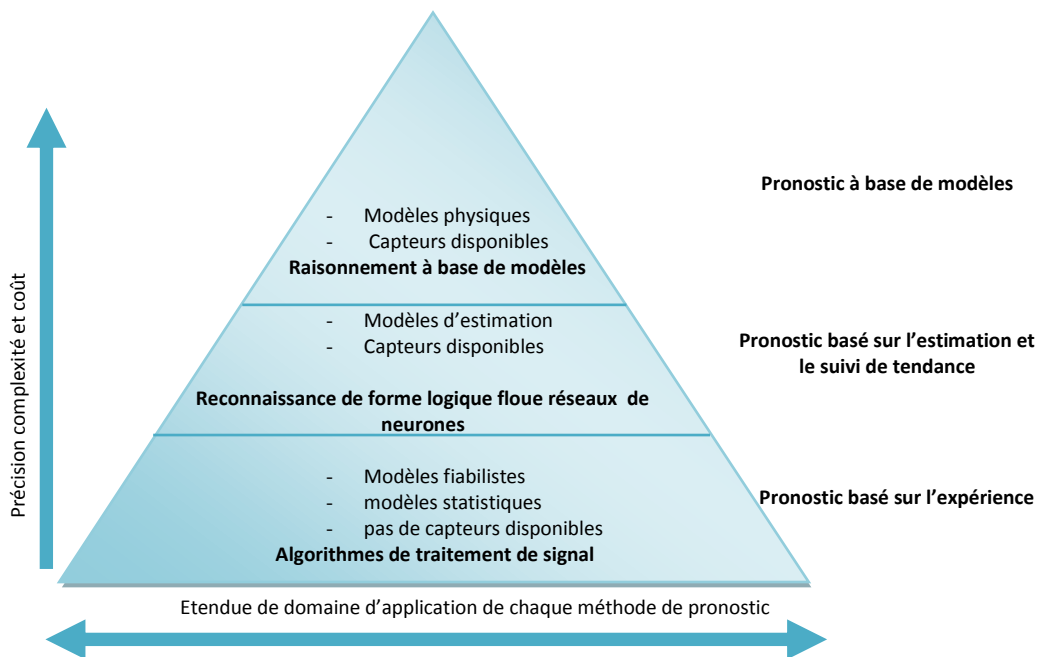


Figure II. 11: Classification des Approches de Pronostic

Nous choisissons ainsi pour la suite de distinguer les approches basées sur un modèle physique, les approches guidées par les données, et les approches basées sur l'expérience.

II. 9. 1 APPROCHES BASEES SUR UN MODELE PHYSIQUE

La mise en œuvre de ce type d'approches s'appuie généralement sur une représentation mathématique du mécanisme de dégradation. Ce qui implique des connaissances spécifiques liées à la défaillance ainsi qu'au mode de fonctionnement du système analysé. Les relations causales issues des lois de la physique sont utilisées pour modéliser les interactions entre les entités du système. Ainsi, les résidus sont employés comme instruments mathématiques. Ces résidus sont les résultats de contrôles de cohérence entre les mesures détectées du système réel et les sorties du modèle mathématique. De gros résidus impliquent la présence de défauts de fonctionnement, et des petits résidus la présence de perturbations normales comme les bruits ou les erreurs de modélisation. Les techniques statistiques sont normalement utilisées pour définir des seuils pour détecter la présence de défauts [4].

Le principe du pronostic basé sur les modèles physiques est de déterminer le niveau de dégradation courant du système (grâce aux processus de surveillance et de diagnostic) et d'évaluer le temps restant avant la défaillance en utilisant la courbe d'évolution de la dégradation en fonction de la sollicitation du système.

Les principaux avantages de cette approche sont [4] :

- la flexibilité : si une des propriétés du système ou de la dégradation change, le modèle peut être réajusté pour prendre en compte cette modification,

- l'interprétabilité : est la capacité d'intégrer la compréhension physique du système au modèle. la connaissance de la structure du modèle permet de relier la variation d'indicateurs à une modification de paramètre.

Les principaux inconvénients du pronostic basé sur les modèles physiques sont :

- le développement des modèles est extrêmement coûteux. Cela nécessite un niveau de qualification élevé et une expérience certaine,
- il est difficile de construire un modèle global pour les systèmes complexes : complexité de modélisation des interactions entre les différents mécanismes, difficultés de calcul associées à la résolution d'un système d'équations différentielles (explosion des temps de calcul pour une simulation de Monte Carlo),
- il peut être impossible de généraliser une approche basée sur un modèle physique orientée « composant » à un niveau « système »,
- il est nécessaire de connaître les mécanismes de dégradation et les facteurs influents sur celle-ci.

Les techniques employées pour cette approche sont basées sur les outils suivants : l'espace de parité, les observateurs (filtres de Kalman,...) et les estimations paramétriques.

II. 9. 1. 1 ESPACE DE PARITE

Un espace de parité est une relation de redondance analytique représentée par une équation dans laquelle toutes les variables sont connues. La génération de telles relations permet d'engendrer des résidus. En l'absence de défaut, le résidu est statistiquement nul. Lors de l'apparition d'un défaut, son amplitude évolue de manière significative.

II. 9. 1. 2 OBSERVATEURS

La génération de résidus à l'aide d'une estimation d'état consiste à reconstruire l'état ou, plus généralement, la sortie du processus à l'aide d'observateurs, et à utiliser l'erreur d'estimation comme résidu. Cette méthode s'est beaucoup développée car elle donne lieu à la conception de générateurs de résidus flexibles.

II. 9. 1. 3 ESTIMATION PARAMETRIQUE

L'approche par estimation paramétrique considère que l'influence de défauts se reflète sur les paramètres et non pas uniquement, comme c'est le cas des observateurs, sur les variables du système physique. Le principe de cette méthode consiste à estimer continuellement les paramètres du procédé en utilisant les mesures d'entrée/sortie et en évaluant la distance qui les sépare des valeurs de référence de l'état normal du procédé. L'estimation paramétrique possède l'avantage d'apporter de l'information sur l'importance des déviations. Toutefois, un des inconvénients majeurs de cette méthode réside dans la nécessité d'avoir un système physique excité en permanence. Ceci pose donc des problèmes d'ordre pratique dans le cas de procédés dangereux, coûteux ou fonctionnant en mode stationnaire.

Pratiquement, un modèle stochastique non linéaire de la dynamique des fissures par fatigue pour le calcul en temps réel du taux de dommage en fonction du temps dans les structures mécaniques, a été proposé par [68]. Cette configuration de modèle permet la construction d'un filtre pour l'estimation de l'état des dommages et la prédiction de durée de vie restante basée sur le principe du Filtre de Kalman étendu au lieu de résoudre l'équation de Kolmogorov.

Dans un article ultérieur, ces auteurs [69] ont également examiné la prévision de croissance des fissures de fatigue en utilisant des processus de Gauss-Markov qui ne nécessitent pas de solution de l'équation du filtre de Kalman étendu. Cependant, la validation du modèle a été limitée pour des données statistiques expérimentalement générés.

[70] a proposé un modèle de modification de lois de dégradation en fonction de deux types de facteurs influents : ceux qui appartiennent au composant en lui-même (état initial de dégradation, caractéristiques géométriques, chimiques) et ceux qui relèvent des conditions opérationnelles et environnementales d'utilisation (mode de fonctionnement, température, humidité, etc.).

Un modèle de choc à temps discret et à états finis peut être utilisé à des fins de modélisation des dommages cumulatifs pour un seul composant. Dans cette forme de base, ces modèles offrent un moyen de calculer la fonction de distribution cumulative du temps aléatoire nécessaire pour atteindre un état de défaillance. L'état de défaillance dans le modèle de choc correspond à un niveau prédéterminé de dommages cumulatifs qui est supposé être une fonction monotone croissante du temps.

[71] ont analysé un système dont la défaillance a été causée par l'apparition d'un choc supérieur à un certain niveau prédéfini. Une paire corrélée (X_n, Y_n) de renouvellement des séquences avec une fonction de distribution conjointe, associée à leur modèle de choc :

$$F_{X,Y}(x, y) = P\{X_n \leq x, Y_n \leq y\}, n = 0, 1, 2, \dots$$

X_n est l'amplitude du n^{eme} choc et Y_n est l'intervalle de temps entre deux chocs consécutifs.

[72] et [73] ont présenté deux modèles de propagation de défauts par l'intermédiaire du mécanisme de défaillance employé pour évaluer la RUL des roulements.

[74] font remarquer que la durée de vie du roulement peut être évaluée et prédite efficacement en surveillant le changement de la rigidité dynamique du système sur la base de mesures de vibrations en temps réel.

[75] a modélisé l'accumulation des dommages dans un système dynamique structurel comme équations différentielles non linéaires du premier/second ordre.

[76] a modélisé la dégradation comme un processus à « temps lent », qui est couplé avec un sous-système observable à « temps rapide ». Le modèle a été utilisé pour suivre la dégradation de la batterie (tension) dans un oscillateur électromécanique.

[77] ont utilisé le processus de Markov discret pour représenter le processus de rupture, de même que le processus covariable pour le calcul de la durée de vie utile résiduelle en fonction des conditions actuelles.

[78] ont démontré la capacité des méthodes de classification à base du modèle de Markov caché (HMM) pour le diagnostic et le pronostic autonomes. Le modèle de pronostic est guidé par une distribution multidimensionnelle des points de transition d'état générés par le HMM.

II. 9. 2 APPROCHES DE PRONOSTICS GUIDEES PAR LES DONNEES

Les approches guidées par les données proviennent directement des données systématiques de fonctionnement du système (par exemple, l'étalonnage, les données colorimétriques, les données spectrométriques, puissance, vibration et acoustique, la température, la pression, les débris dans l'huile, des courants et tensions...). Dans de nombreuses applications, les données entrée/sortie mesurées sont les principales sources pour obtenir une compréhension

approfondie du comportement de la dégradation du système. Ce type d'approches vient ainsi à apprendre (par les exemples) le phénomène de dégradation, et à capturer les relations subtiles entre données, même si ces relations sont inconnues ou difficile à décrire. Elles sont construites sur la base des enregistrements historiques (données d'apprentissage) et donnent une prédiction sur la base des données de surveillance conditionnelle (CM). Les approches basées sur les données reposent sur l'hypothèse que les caractéristiques statistiques des données sont relativement uniformes tant qu'un événement de dysfonctionnement dans le système n'est survenu. La mise en œuvre d'une approche « guidée par les données » est relativement simple car elle ne requiert pas la connaissance formelle des mécanismes de dégradation. Le principe étant qu'une même cause aura toujours les mêmes effets. En outre, cette approche peut être adaptée à tout type d'application disposant de données suffisantes en quantité et qualité. En revanche, c'est là également sa principale limite : l'efficacité d'une approche de pronostic guidée par les données dépend fortement de l'existence de ces données opérationnelles. Par ailleurs, le processus de surveillance conditionnelle évalue également la justesse du pronostic par sa capacité à fournir une estimation fiable et précise de l'état de santé actuel du système [4].

Dans ce qui suit, les approches basées sur les données sont classées en deux catégories: approches statistiques (analyse des séries temporelles, ...) et approche par Intelligence Artificielle (IA).

II. 9. 2. 1 TECHNIQUES STATISTIQUES

Les techniques statistiques les plus populaires sont les techniques de lissage (tel que le lissage exponentiel) et les techniques d'analyse des séries temporelles tels que les modèles de régression, notamment les modèles AR (Auto-Regressive) et ARMA (Auto-Regressive Moving Average models). Ces techniques sont aussi classées sous la catégorie analyse par tendance. Ces techniques présentent l'intérêt d'être simples à programmer. Cependant, elles supposent que la dégradation du système est monotone, et ne permettent pas d'intégrer les variations des conditions d'utilisation du système : l'évolution de la dégradation à venir doit être similaire à celle qui a été apprise. Le paragraphe suivant décrit des techniques d'analyse de séries temporelles.

II. 9. 2. 2 APPROCHES PAR ANALYSE DES SERIES TEMPORELLES

Si des quantités suffisantes de données en fonction du temps sont disponibles, les techniques d'analyse des séries temporelles sont souvent utilisées pour déterminer l'état du système en un point donné dans le futur. Ces techniques s'appuient fortement sur des données antérieures pour prédire les performances dans le futur.

II. 9. 2. 3 TECHNIQUES DE REGRESSION

L'analyse par régression utilise les données existantes et détermine les relations, le cas échéant, entre le résultat mesurable et les variables qui contribuent à ce résultat. Un modèle général de régression linéaire est donnée par,

$$Y_i = \beta_0 + \beta_1 X_{i,1} + \beta_2 X_{i,2} + \dots + \beta_{p-1} X_{i,p-1} + \varepsilon_i \quad i = 1, \dots, n$$

où Y_i est une variable aléatoire indiquant la valeur de la réponse de l' i^{eme} essai, $\beta_0, \beta_1, \beta_2, \dots, \beta_{p-1}$ sont les paramètres estimées, $X_{i,1}, X_{i,2}, \dots, X_{i,p-1}$ sont les valeurs du prédicteur et ε_i est l'erreur aléatoire de moyenne = 0, variance = σ^2 , et de covariance = 0.

L'analyse par régression cherche à estimer les paramètres de la fonction de régression, $\beta_0, \beta_1, \beta_2, \dots, \beta_{p-1}$, à l'aide de la méthode des moindres carrés pour trouver un modèle représentatif. La méthode des moindres carrés définit une variable Q , où

$$Q = \sum_{i=1}^n (Y_i - \beta_0 - \beta_1 X_{i,1} - \beta_2 X_{i,2} - \dots - \beta_{p-1} X_{i,p-1})^2$$

et tente de trouver des estimations pour $\beta_0, \beta_1, \beta_2, \dots, \beta_{p-1}$, notées $b_0, b_1, b_2, \dots, b_{p-1}$, qui minimisent Q pour les observations $(X_1, Y_1), (X_2, Y_2), \dots, (X_n, Y_n)$. La solution simultanée des équations formées en prenant la dérivée de Q par rapport à $\beta_0, \beta_1, \beta_2, \dots, \beta_{p-1}$, donne les estimations des moindres carrés, $b_0, b_1, b_2, \dots, b_{p-1}$. Estimations des moindres carrés ont une variance minimale entraînant

$$\hat{Y} = b_0 + b_1 X_1 + b_2 X_2 + \dots + b_{p-1} X_{p-1}$$

La méthode du maximum de vraisemblance peut également être utilisée pour estimer $b_0, b_1, b_2, \dots, b_{p-1}$, si la distribution de probabilité des termes d'erreur est connue.

Les auteurs [79] ont examiné une approche adaptative de pronostic où une amplitude de défaut futur de roulement a été calculée au moment $t + \Delta$ ($\Delta > 0$) connaissant la condition de fonctionnement et l'amplitude de défaut à l'instant t . Cet algorithme adaptatif, basé sur un algorithme récursif des moindres carrés appliqué à un modèle de propagation de défauts, a été ensuite utilisé pour tenir compte du comportement variable avec le temps et permet de prévoir les pannes imminentes.

[80] propose un modèle de régression logistique afin de calculer la probabilité d'une défaillance pour une condition donnée des variables, et un modèle de série temporelle autorégressif moyenne mobile (ARMA) pour étendre la condition pour une éventuelle prédiction de défaillance.

[81] fournissent beaucoup plus de détails sur les modèles de régression linéaires et non linéaires, ainsi que les dangers de l'extrapolation au-delà des données observées. Cependant, la régression n'est pas la seule méthode d'estimation d'état utilisé pour le pronostic. Certaines de ces méthodes sont décrites ci-dessous.

II. 9. 2. 4 AUTOREGRESSIF MOYENNE MOBILE INTEGREE (ARIMA)

Les séries temporelles autorégressives moyenne mobile intégrée (ARIMA) est une technique d'estimation d'état couramment utilisée dans le pronostic. Elle est également connue comme analyse par tendance. Le modèle ARIMA est une construction générique qui intègre des processus autorégressifs, processus moyenne mobile, et une capacité à rendre compte des données non stationnaires.

Soit de $\tilde{Z}_t - \mu$, un processus AR d'ordre p défini mathématiquement comme

$$\tilde{Z}_t = \phi_1 \tilde{Z}_{t-1} + \phi_2 \tilde{Z}_{t-2} + \dots + \phi_p \tilde{Z}_{t-p} + \alpha_t$$

réécrit sous la forme

$$\phi(B)\tilde{Z}_t = \alpha_t$$

où μ est la moyenne des données de séries temporelles, Z_1, Z_2, \dots, Z_n sont les observations ordonnées en temps, $\phi_1, \phi_2, \dots, \phi_p$ sont les paramètres inconnus d'un processus autorégressif, α_t est un bruit blanc et B est l'opérateur retard respectivement. Les données observées fournissent des estimations pour $\phi_1, \phi_2, \dots, \phi_p$.

Un processus à moyenne mobile d'ordre q est défini comme

$$\tilde{Z}_t = \alpha_t - \theta_1\alpha_{t-1} - \theta_2\alpha_{t-2} - \dots - \theta_q\alpha_{t-q}$$

et peut être réécrit sous la forme

$$\tilde{Z}_t = \theta(B)\alpha_t$$

où les données observées fournissent des estimations pour $\mu, \theta_1, \theta_2, \dots, \theta_q$

Un modèle non-stationnaire doit être transformé en un modèle stationnaire avant que les techniques d'auto-régression et/ou de moyenne mobile soient appliquées. Cette transformation se fait normalement par différenciation, $(I - B)^d$, mais elle peut également être effectuée en prenant le logarithme des données des séries temporelles. Par conséquent, un modèle ARIMA complet d'ordre (p, d, q) est défini mathématiquement par

$$\phi(B)(I - B)^d\tilde{Z}_t = \theta(B)\alpha_t$$

Ce modèle peut décrire à la fois la série temporelle stationnaire et non stationnaire mais nécessite une quantité importante de données pour estimer $\mu, \phi_1, \phi_2, \dots, \phi_p$ et $\theta_1, \theta_2, \dots, \theta_q$. Les auteurs [82] ont utilisé les modèles ARIMA pour prédire le temps de défaillance des machines à commande numérique (tour et fraiseuse...). Ces machines ont été surveillées en temps réel à l'aide des capteurs de son, de vibrations et de consommation d'énergie. Les auteurs ont été en mesure de prédire si les machines nécessitent un entretien dans les périodes futures.

L'auteur [83] a utilisé un processus autorégressif (AR) pour modéliser des signaux de vibration pour le pronostic. L'état de santé de l'engrenage est diagnostiqué en caractérisant le signal d'erreur entre les signaux filtrés et non filtrés en utilisant à la fois la simulation numérique et des données expérimentales. Cependant, les paramètres AR (les coefficients polynomiaux) n'ont pas de signification physique liée au système surveillé.

Les auteurs [84] passent en revue les modèles de régression non linéaires et présentent une méthode en deux étapes pour estimer les paramètres du modèle. Etape 1 : les estimations des paramètres ont été obtenues à partir de chaque trajectoire de dégradation. Ces estimations ont été transformées, si nécessaire, pour s'assurer que les estimations des paramètres proviennent d'une distribution normale multivariée. Toutes les estimations de l'étape 1 ont été ensuite combinées pour déterminer les estimations de la moyenne, la variance et la covariance utilisées par la suite pour trouver la distribution de la durée de vie.

[85] incorporent la modélisation des séries temporelles pour estimer la distribution de la probabilité de dégradation du matériau réflecteur solaire en un point donné dans le temps et la distribution de la probabilité de la durée de vie. La dégradation a été modélisée avec un processus autorégressif (AR) utilisant la dégradation quotidienne prédite sur la base des données enregistrées des années précédentes. Dans leur travail, la simulation de Monte Carlo

a fournies de nombreuses trajectoires échantillons qui ont été utilisées pour construire des fonctions de distribution empiriques de la dégradation et de la durée de vie.

II. 9. 2. 5 APPROCHES A BASE INTELLIGENCE ARTIFICIELLE (IA)

Les réseaux de neurones artificiels (RNA), les algorithmes génétiques, la logique floue et autres techniques d'apprentissage automatique constituent une classe d'approches connue sous le nom approches d'intelligences artificielles (IA). Ces techniques ont la capacité d'apprendre en utilisant l'historique passé et ensuite tenter de prédire l'état ou le résultat quand un nouveau ensemble de données d'entrée est présenté. Par conséquent, ces techniques sont fréquemment utilisées dans les procédés actuels de pronostic. Certaines des techniques d'intelligence artificielle et leur application pour le pronostic sont passées en revue dans cette section.

L'une des approches d'apprentissage-machine les plus populaires pour le pronostic est l'utilisation des RNA pour modéliser un système. Les RNA sont des modèles (en général non linéaires) qui établissent un ensemble de relations fonctionnelles interconnectées entre l'entrée stimuli et la sortie désirée où les paramètres de la relation fonctionnelle doivent être ajustés pour une performance optimale. Le fonctionnement des RNA est proche des neurones réelles du cerveau humain. Chaque neurone a des dendrites, qui sont les voies d'entrée, un soma qui traite les entrées et un axone qui chemine la sortie. Un RNA est formée d'un ensemble de ces neurones artificiels. L'architecture du RNA à trois couches est illustrée dans la figure II. 12. Ce réseau a normalement une couche d'entrée, une ou plusieurs couches cachées et une couche de sortie. Chaque couche, à l'exception de la couche d'entrée, est constituée d'un certain nombre de neurones. Les poids sont désignés pour chaque entrée (de la couche d'entrée aux neurones de la couche cachée). Compte tenu des résultats de la fonction de transfert de la couche cachée, un autre ensemble de poids est appliqué comme entrée à la couche de sortie. Les derniers résultats de la fonction de transfert de la couche de sortie forment l'ensemble des sorties du RNA.

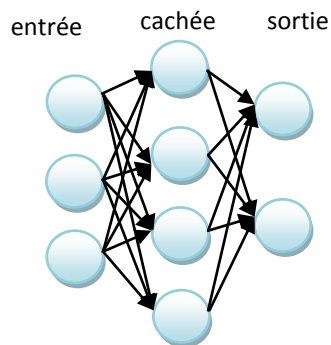


Figure II. 12 : Architecture d'un réseau de neurone à trois couches

Les techniques RNA pour l'estimation de la durée de vie utile résiduelle ont été utilisées par de nombreux chercheurs. [86] ont examiné l'utilité des réseaux neurones à retro-propagation pour prédire l'endommagement par fatigue. Leur méthodologie a été utilisée pour réduire le temps de calcul nécessaire par la méthode conventionnelle pour résoudre numériquement les équations différentielles non linéaires. en outre, les auteurs indiquent que ces réseaux de neurones peuvent éventuellement être utilisés pour l'analyse en temps réel des modèles

d'endommagement par fatigue et d'autres types de modèles de défaillance en plus de la prédiction de la durée de vie restante.

Un réseau neurones dynamique à ondelettes (DWNN) a été mis en œuvre pour transformer les données des capteurs en une évolution temporelle du défaut pour prédire la durée de vie utile résiduelle des roulements [87]. Le modèle DWNN a été entraîné en utilisant les données de vibration des roulements défectueux avec une largeur et une profondeur de fissure variable, et a été ensuite utilisé pour prédire l'évolution de la fissure jusqu'à la rupture finale.

[88] a utilisé un RNN pour effectuer une prédiction d'un pas en avant des données conditionnelles (CM) d'un réducteur. Il a démontré que le RNN sont capables d'apprendre la tendance des données temporelles non linéaires et peuvent être adaptés pour la prédiction des séries temporelles [9].

[89] ont développé deux classes de modèles de réseau neurones pour l'estimation du temps de défaillance du roulement pendant sa durée de vie. Ils ont utilisé à la fois des modèles de roulement simples et clustrés. Chaque classe a été modelée utilisant trois techniques différentes de calcul de poids: application de poids au temps de défaillance (Weight Application to Failure Times (WAFT)), application de poids aux paramètres exponentiels (Weight Application to Exponential Parameters (WAEP)), et application de poids au paramètres exponentiels - paramètres mis à jour (WAEP-UP). Les résultats ont montré que 92% des prévisions de temps de défaillance calculées utilisant des roulements de validation étaient dans les 20% de la durée de vie réelle des roulements.

[90] notent que pour les roulements à une rangée fortement accélérées, ce n'est ni pragmatique ni utile de modéliser le processus de prédiction sur tout leur cycle de vie en raison de la forte dispersion de ce dernier. Ils ont intégré l'extraction d'indicateur de dégradation du Self-Organizing Map (SOM) avec la prédiction de durée de vie résiduelle à base de réseaux neuronaux rétro-propagateurs. L'indicateur de la dégradation est extrait en utilisant trois caractéristiques temporelles et trois de fréquence avec apprentissage non supervisé du RN.

[91] ont proposé un réseau neuronal dynamique à ondelettes (DWNN) multi-entrées multi-sorties, doté d'une capacité de stockage et d'information temporelle dans ses fonctionnalités de manière à pouvoir prédire le futur, réalisant ainsi les tâches de classification des défauts et de pronostic. Un réseau de neurones à ondelettes (WNN) fonctionne comme un capteur virtuel et le DWNN comme prédicteur. Un réseau de neurones récurrent ou dynamique permet une rétro-propagation des signaux. En associant l'apprentissage renforcé avec l'algorithme génétique (GA) cela permet à l'algorithme d'interagir avec l'environnement pour améliorer sa performance et mettre à jour uniquement lorsqu'est nécessaire.

[92] prétendent que les méthodes floues fréquemment utilisées dans les équipements d'analyse de la criticité ne sont pas précises, et nécessitent de nombreuses données en question. Dans leur étude, un modèle gris et un réseau neuronal à rétro-propagation ont été appliqués à un sous-système de pompage à eau. Les résultats montrent que le procédé est faisable et efficace pour une application dans les centrales électriques.

[93] ont présenté une machine à vecteurs de support-réseau bayésien hybride « Support Vector Machine- Bayesian Network » (SVM- BN) pour prédire l'erreur thermique dans les machines-outils. Dans leur recherche, le SVM-BN a été développé pour classer les erreurs par groupes en fonction des conditions de fonctionnement et ensuite cartographier le profil de température avec l'erreur mesurée. Ce modèle est particulièrement utile dans des

environnements de production où les machines-outils sont soumises à une variété de conditions de fonctionnement.

Une autre technique IA populaire utilisée pour le pronostic est la technique logique floue. La logique floue fournit un langage (avec syntaxe et sémantiques locales) dans lequel on peut traduire les connaissances qualitatives sur le problème à résoudre. En particulier, la logique floue permet d'utiliser des variables linguistiques pour modéliser des systèmes dynamiques. Ces variables prennent des valeurs floues qui se caractérisent par une phrase et une fonction d'appartenance. La signification d'une variable linguistique peut être interprétée comme une contrainte élastique sur sa valeur. Ces contraintes sont propagées par les opérations d'inférence floue. Le mécanisme de raisonnement résultant a des propriétés d'interpolation puissantes qui à leur tour donnent à la logique floue une robustesse remarquable en ce qui concerne les variations des paramètres et des perturbations du système.

Lorsqu'elle est appliquée au pronostic, la logique floue est généralement appliquée en conjonction avec une méthode d'apprentissage machine et est utilisée pour traiter certaines des incertitudes que toutes les estimations pronostics font face.

[94], [95] emploie la technique logique floue pour une évaluation fiable de l'état de santé du système. Pour l'estimation automatique de l'état de santé, ils ont utilisé cette technique pour représenter les degrés de gravité ou de dégradation.

Un document important sur pronostic Neuro-Flous (NF) est présentée par [96]. Les deux techniques RNNs et neuro-flous (NF) ont été évaluées en utilisant des données de référence test des taches solaires et des données test en ligne d'un réducteur. Dans le cas des taches solaires, le NF sans interpolation est moins précis que les RNNs. Mais le NF avec interpolation produit des résultats plus précis que les RNNs et fournit des résultats similaires avec environ dix pour cent (10%) des époques d'apprentissage. Quand il s'agit du test en ligne du réducteur, le NF saisit le comportement dynamique du système plus rapidement et se révèle plus supérieur que les RNNs.

Une autre approche d'apprentissage automatique est l'algorithme de détection d'anomalie (connu aussi comme détection de nouveauté ou algorithme de détection des valeurs aberrantes). Ces algorithmes apprennent au modèle le comportement nominal du système et notent lorsque de nouvelles données des capteurs ne parviennent pas à correspondre au modèle, indiquant une anomalie qui pourrait être un signe de défaillance [97], [98].

[99] ont présenté une vue d'ensemble de l'intégration des outils d'intelligence artificielle dans le projet européen PROTEUS dans la e-Maintenance. Un modèle générique AI pour spécifier comment les outils d'IA peuvent être intégrés dans la plateforme a été proposé. Le document présente plusieurs exemples d'outils d'IA dans le diagnostic et le pronostic et conclut du besoin de « méta-outil » qui peut être intégré dans le modèle générique.

L'utilisation de modèles de Markov cachés (HMM) dans le pronostic des défauts a été étudiée par [100]. Dans un HMM, un système est modélisé comme un processus stochastique dans lequel les états suivants n'ont aucun lien de causalité avec les états précédents. Dans leur travail, un HMM a été entraîné pour reconnaître un type de défaut des roulements coniques et à galets sur la base des données de vibration correspondantes. Les HMM ont également été entraînés pour évaluer les états de défaut. La similitude entre état actuel et l'état de défaillance a été utilisée comme indice de dégradation du roulement, qui a été ensuite extrapolée pour estimer le temps de dépassement d'un seuil de défaillance prédéterminé.

Le HMM présente une lacune dans la représentation temporelle réelle puisque sa distribution de durée d'état se limite uniquement à une distribution exponentielle. [101] ont ainsi proposé un Modèle semi-Markov caché (HSMM) segmenté. Contrairement aux HMMs, qui génèrent une seule observation pour chaque état, les HSMMs génèrent un segment d'observations et estiment les durées à partir de données d'apprentissage.

La force des techniques orientées données réside dans leur capacité à transformer les données de grandes dimensions bruitées en données de dimensions réduites pour la prise de décision en diagnostic/pronostic. Le principal inconvénient est leur dépendance de la quantité et de la qualité des données opérationnelles du système.

II. 9. 3 APPROCHES BASEES SUR L'EXPERIENCE

Les approches de pronostic basées sur l'expérience se fondent sur la modélisation stochastique des phénomènes de dégradation (loi de fiabilité, processus markoviens, ou non markoviens). Ces modèles peuvent être construits sur la base d'expériences similaires passées [4].

Les approches basées sur l'expérience sont principalement issues des méthodes traditionnelles de modélisation et de traitement fiabilistes. Le pronostic est en ce sens assimilé à une étude de fiabilité prévisionnelle, l'objectif étant d'identifier les paramètres d'une distribution aléatoire décrivant le phénomène de dégradation ou de défaillance (loi de poisson, loi exponentielle, loi de weibull, loi log-normal). Dans cet ensemble, la loi de Weibull reste la plus répandue [102]. Les approches conventionnelles basées sur la fiabilité dédiées au pronostic peuvent être classées en deux catégories: *fiabilité basée sur la défaillance* et *fiabilité basée sur dégradation*. La fiabilité basée défaillance est utilisée pour estimer la distribution de la durée de vie et ses paramètres lorsque des données temporelles de défaillance suffisantes, complètes et/ou censurées existent. Si la connaissance préalable de la distribution de la durée de vie existe pour des composants similaires, souvent la distribution de la durée de vie est supposée suivre la même distribution d'un composant similaire.

Comparée à la fiabilité à base de défaillance, la fiabilité à base de dégradation se concentre sur l'utilisation des mesures de la dégradation du composant, et non les données de défaillance, pour évaluer la durée de vie résiduelle d'un composant. La dégradation est également connue sous le nom dommages cumulatifs [103].

De plus, l'utilisation d'un modèle fiabiliste évolutif de type ALM (Accelerated Life Model), ou modèle de risque proportionnel PHM (Proportional Hazard Model), ou la mise en œuvre d'une démarche bayésienne permettant d'actualiser les paramètres de la loi de dégradation à chaque nouvelle information disponible peut également représenter une solution. La simulation de Monte Carlo permet de combiner différents phénomènes aléatoires mais elle est confrontée au problème d'explosion du temps de simulation [4]. Les modèles (PHM) sont couramment utilisés dans la prédiction de la défaillance et de l'analyse de la fiabilité. Le PHM suppose que les risques changent proportionnellement avec les covariables et que la constante de proportionnalité est la même en tout temps.

Parmi les travaux récents dans le domaine du pronostic basé sur l'expérience, nous pouvons mentionner les suivants.

[104] a proposé le logiciel EXAKT pour optimiser le remplacement d'équipements critiques (turbine, vanne, moteur...). La méthodologie proposée combine le suivi du taux de défaillance

obtenu par un modèle de Weibull et la prévision de l'évolution des covariables suivant un processus markovien discret non homogène.

[105] a proposé la suite logicielle IBTV « Inference Bayésienne pour le Traitement du Vieillissement » d'EDF en intégrant une méthodologie originale de pronostic où la stratégie de maintenance repose sur le déclenchement d'interventions suite à la détection du vieillissement d'un composant (vanne, relais électrique, structure mécanique, par exemple). L'approche est basée sur une démarche bayésienne avec une modélisation a priori du comportement du taux de défaillance (composition loi exponentielle/Weibull). De plus, l'efficacité des actions de maintenance est prise en compte suivant l'avis d'expert.

[106] ont proposé une approche basée sur la fiabilité pour l'estimation de la politique optimale de maintenance pour minimiser le coût total de la maintenance par unité de temps. Ils ont utilisé le modèle PHM pour identifier l'importance des variables surveillées et le temps total sur un graphique de test pour trouver la politique de maintenance optimale.

[107] ont utilisé un cadre des risques proportionnels pour modéliser des systèmes réparables complexes.

[108] a présenté un modèle de prédiction de durée de vie de roulement utilisant la détermination de la durée de vie statistique.

Les modèles statistiques tels que les modèles d'intensité proportionnels « Proportional Intensity Models » (PIM) et PHM sont des outils utiles pour l'estimation de durée de vie utile et la tendance du processus de dégradation des composants [7].

Ainsi, [109] ont utilisés une estimation de durée de vie résiduelle « Residual Life Estimate » (RLE) statistique sur des roulements à rouleaux pour étudier l'influence des variations dans les mesures de vibrations et les niveaux de lubrification sur la vie des roulements. Les RLEs sont basées sur les modèles d'intensité proportionnels (PIM) et sont principalement utilisés pour les systèmes non-réparables utilisant des données historiques de défaillance et des mesures correspondantes.

[77] ont discuté l'estimation de la RUL pour un processus Markovien de défaillance temporel qui comprend un modèle commun PHM et la propriété Markovienne pour l'évolution de la covariable comme cas particulier.

II. 9. 4 TENDANCE VERS L'IDENTIFICATION DE NOUVELLES CARACTERISTIQUES DE PRONOSTIC

Au lieu d'échouer soudainement, la plus part des systèmes, surtout les systèmes mécaniques, subissent un processus de dégradation graduel jusqu'à la défaillance totale. Comme les algorithmes de pronostic basés sur la dégradation suivent la tendance d'une certaine mesure de dégradation appelée caractéristique de pronostic, l'identification d'une caractéristique de pronostic appropriée qui doit bien saisir la tendance de la progression du défaut sur toute la durée de vie du composant/système est essentielle pour la performance du modèle pronostic. Il est difficile de faire une prédiction précise si la tendance de l'indicateur de santé n'est pas évidente sur tout le cycle de vie du composant ou si la tendance se produit seulement juste avant une défaillance comme indiqué sur la figure III. 8 (RMS).

De nombreuses méthodes de sélection de caractéristiques pertinentes ont été développées pour des tâches de classification en utilisant des approches à base d'Intelligence Artificielle IA. Mais rares sont celles qui se focalisent sur l'identification de nouvelles caractéristiques de

diagnostic ou de pronostic. Dans ce contexte, les auteurs [110] ont développé deux nouvelles caractéristiques de défaut nommés TALAF et THIKAT, qui combinent des caractéristiques classiques, afin d'améliorer le diagnostic jusqu'au point où les signes ultimes d'une défaillance catastrophique sont observés, pour diagnostiquer la gravité de la dégradation des roulements à billes.

L'identification des caractéristiques de pronostic est laissée au premier lieu à l'analyse et les jugements des experts. Elles sont généralement choisies par une inspection visuelle des données disponibles. Les caractéristiques appropriées de dégradation ne doivent pas être obligatoirement mesurées directement. Nous pouvons aussi fusionner plusieurs caractéristiques dans un indicateur de santé unique qui est ensuite utilisé pour prédire la durée de vie utile restante RUL en utilisant une approche automatisée qui résulte en une caractéristique optimale, ou presque optimale. Comme approche automatisée, [111] et [112] ont proposé une méthode basée sur l'algorithme génétique pour identifier un ensemble optimal de caractéristiques de pronostic à partir d'une population de caractéristiques et ont également proposé un ensemble de mesures d'aptitude pour ces caractéristiques, à savoir mesure de monotonie, mesure de « pronosticabilité », et mesure de « tendance ». Ces paramètres (ou mesures) peuvent être utilisés avec n'importe quel algorithme d'optimisation traditionnel pour identifier la caractéristique de pronostic appropriée. Les auteurs [113] ont proposé d'utiliser une métrique basée sur la mesure de la séparabilité des segments temporels consécutifs afin d'évaluer la qualité des caractéristiques extraites à partir des signaux bruts pour la prédiction de la RUL. Une approche basée sur la programmation génétique (GP) a été aussi proposé par [114] pour identifier des caractéristiques avancées possédant une forte corrélation avec la progression de défaut afin de réaliser une meilleure prédiction de la RUL. Cependant, il est très difficile de déterminer un seuil de défaillance de la caractéristique identifiée surtout quand des différentes caractéristiques sont impliquées par la GP. Par conséquent, garder les caractéristiques originales est crucial pour déterminer ce seuil de défaillance.

II. 10 CONCLUSION

La complémentarité de la maintenance traditionnelle préventive et curative par une orientation plus réactive et proactive maintient en condition opérationnelle les systèmes industriels à moindre coût offrant une meilleure performance et compétitivité aux entreprises.

Les méthodes de diagnostic/pronostic présentées ci-avant ne constituent pas un état de l'art exhaustif des méthodes existantes. Seulement, les techniques décrites sont les plus connues et les plus utilisées. Malgré que le diagnostic se soit bien développé dans ces dernières décades, le pronostic reste confronté à plusieurs challenges.

Le choix d'une méthode de diagnostic dépend de la connaissance du système, de la présence de capteurs ou de modèles qui permettent de suivre l'état réel du composant/sous-système/système. Pour cela, nous avons limité le cadre de notre travail dans la partie diagnostic (chapitre III) à l'amélioration du taux de classification correcte des défauts du moteur à induction tels que les défauts des roulements, les défauts déséquilibre des phases statoriques et les défauts cassures de barres rotoriques. A cet égard, nous avons proposé une approche orientée données vu la disponibilité des données expérimentales d'un banc d'essai

sur ce moteur. Une approche statistique probabiliste et une autre neuronale sont proposées et ensuite comparées.

Le pronostic, quand à lui, considéré comme un processus clef dans les stratégies de maintenance, telle que la CBM, se révèle être une activité tout à fait prometteuse. L'activité de pronostic est supportée par des approches à base de modèle, des approches guidées par les données et des approches à base d'expérience vue la diversité des méthodes et les outils disposés. Sachant qu'aucune approche de pronostic n'est universelle et que le choix d'une technique adéquate dépend des contraintes classiques limitant l'applicabilité des outils : disponibilité des données et/ou connaissances et/ou expériences, dynamique et complexité des systèmes réels, contraintes d'implémentation (précision, temps de calcul, etc.), possibilités de mesures (capteurs, système SCADA, etc.), etc.

Ainsi, nous choisissons d'orienter notre travail vers le pronostic au niveau composant que nous jugeons critique: approche « locale » au lieu d'une approche « globale » niveau système. L'approche de pronostic guidée par les données est privilégiée dans notre cas pour les raisons suivantes :

- Difficulté de disposer de connaissance sur les phénomènes de dégradation complexes.
- Insuffisance d'expériences passées pour déployer des approches statistiques significatives à cause de la variabilité de conditions d'usage et de détérioration.
- Envisagement d'une stratégie de maintenance plus « dynamique » basée sur les données conditionnelles du composant au lieu d'une stratégie de maintenance préventive systématique.

II. 11 REFERENCES

- [1] NF EN 13306 “Terminologie de la Maintenance,” 6, 9, 10, 11, 2001.
- [2] P. Cochetoux “Contribution à la Maintenance Proactive par la Formalisation du Processus de Pronostic des Performances de Systèmes Industriels,” PhD thesis, Université Henri Poincaré Nancy I, 2010
- [3] P. Ribot “Vers l'Intégration Diagnostic/Pronostic pour la Maintenance des Systèmes Complexes,” PhD thesis, Institut National Polytechnique de Toulouse, 2009.
- [4] M. EL Koujok “Contribution au Pronostic Industriel : Intégration de la Confiance à Un Modèle Prédicatif Neuro-Flou,” PhD thesis, Université De Franche-Comté, 2010.
- [5] “An Operations and Maintenance Information Open System Alliance ” disponible sur site: <http://www.mimosa.org/>
- [6] M. Milfelner, F. Cus, et J. Balic, “An Overview of Data Acquisition System for Cutting Force Measuring and Optimization in Milling,” *Journal of Materials Processing Technology AMPT/AMME05 Part 2*, p. 1281-1288, 2005.
- [7] Jardine A.K.S., D. Lin et D. Banjevic. “A Review on Machinery Diagnostics and Prognostics Implementing Condition-Based Maintenance,” *Mechanical Systems and Signal Processing*, vol. 20, p. 1483-1510, 2006.
- [8] M. Kang, J.M. Kim “Reliable Fault Diagnosis of Multiple Induction Motor Defects Using a Two-Dimensional Representation of Shannon Wavelets,” *IEEE Transactions on Magnetics*, 2014.
- [9] Heng A., S. Zhang, A.C.C. Tan et J. Mathew, “Rotating Machinery Prognostics: State of the Art, Challenges and Opportunities,” *Mechanical Systems and Signal Processing*, vol. 23 (3), p. 724-739, 2009.
- [10] P. D. McFadden, “A Revised Model for the Extraction of Periodic Waveforms by Time Domain Averaging,” *Mechanical Systems and Signal Processing*, vol.1 , p. 83-95, 1987.
- [11] M. Karimi, J. Mathew, A. C. C. Tan, et B. Senadji, “Bearing Fault Detection Using Blind Deconvolution and Time Interval Averaging,” *Proceedings of the 3rd Asia Pacific Conference*, p. , 2003.
- [12] B. Widrow, R. C. Goodlin, “Adaptive Noise Cancelling: Principles and Applications,” *Proceedings of the IEEE Conference*, p. 1692-1716, 1975.
- [13] C. C. Tan et B. Dawson, “An Adaptive Noise Cancellation Technique for the Enhancement of Vibration Signatures from Rolling Element Bearings: A Feasibility Study,” *Proceedings of the 2nd IMEKO Symposium on Technical Diagnostc*, London, 1981.
- [14] J. Shiroishi, Y. Li, S. Liang, S. Danyluk, et T. Kurfess, “Vibration Analysis for Bearing Outer Race Condition Diagnostcs,” *Journal of the Brazilian Society of Mechanical Sciences* vol. 21 , p. 484-492, 1999.
- [15] J. Antoni et R. B. Randall, “Unsupervised Noise Cancellation for Vibration Signals: Part I--Evaluation of Adaptive Algorithms,” *Mechanical Systems and Signal Processing*, vol. 18, p. 89-101, 2004.

- [16] J. Antoni et R. B. Randall, "Unsupervised Noise Cancellation for Vibration Signals: Part II-A Novel Frequency-Domain Algorithm," *Mechanical Systems and Signal Processing*, vol.18, p. 103-117, 2004.
- [17] A. C. C. Tan, M. Karimi, J. Mathew, et B. Senadji, "Blind Deconvolution as a Notch Filter in Detecting Bearing Faults Corrupted by Noise," *Scientific Journal of The Vibration Institute of India*, vol.5, 2006.
- [18] H. Qiu, J. Lee, J. Lin, et G. Yu, "Robust Performance Degradation Assessment Methods for Enhanced Rolling Element Bearing Prognostics," *Advanced Engineering Informatics: Intelligent Maintenance Systems*, vol.17, p.127-140, 2003.
- [19] T. Burgess et L. Shimmel, "What is the Prognosis on Your Maintenance Program?" *E&MJ: Engineering & Mining Journal*, vol. 196, p. 32, 1995.
- [20] S. Poyhonen, P. Jover, et H. Hyotyniemi, "Signal Processing of Vibrations for Condition Monitoring of An Induction Motor," *ISCCSP: 2004 First International Symposium on Control, Communications and Signal Processing*, New York, p. 499-502, 2004.
- [21] Y. M. Zhan et A. K. S. Jardine, "Adaptive Autoregressive Modeling of Nonstationary Vibration Signals Under Distinct Gear States. Part 1: Modeling," *Journal of Sound and Vibration*, vol. 286, p. 429-450, 2005.
- [22] D. C. Baillie et J. Mathew, "A Comparison of Autoregressive Modeling Techniques for Fault Diagnosis of Rolling Element Bearings," *Mechanical Systems and Signal Processing*, vol. 10, p. 1-17, 1996.
- [23] Y. Zhan, V. Makis, A.K.S. Jardine, "Adaptive Model for Vibration Monitoring of Rotating Machinery Subject to Random Deterioration," *Journal of Quality in Maintenance Engineering*, vol.9, p.351-375, 2003.
- [24] P. D. McFadden, "Determining the Location Of A Fatigue Crack in A Gear from the Phase of the Change in the Meshing Vibration," *Mechanical Systems and Signal Processing*, vol. 2, no. 4, p. 403- 409, 1988.
- [25] D. Kocur et R. Stanko, "Order Bispectrum: A New Tool for Reciprocated Machine Condition Monitoring," *Mechanical Systems and Signal Processing*, vol. 14, p. 871-890, 2000.
- [26] J. T. Kim et R. H. Lyon, "Cepstral Analysis as a Tool for Robust Processing, Deverberation and Detection of Transients," *Mechanical Systems and Signal Processing*, vol. 6, p. 1-15, 1992.
- [27] P.D. McFadden and J.D. Smith, "Vibration Monitoring of Rolling Element Bearings by the High Frequency Resonance Technique a Review," *TRIBOLOGY international* Vol. 17, No. 1, p. 3-10, February 1984.
- [28] P. Bonato, R. Ceravolo, A. De Stefano, M. Knaflitz, "Bilinear Time-Frequency Transformations in the Analysis of Damaged Structures," *Mechanical Systems and Signal Processing*, vol.11, p. 509-527, 1997.
- [29] G. Dalpiaz et A. Rivola, "Condition Monitoring and Diagnostics in Automatic Machines: Comparison of Vibration Analysis Techniques," *Mechanical Systems and Signal Processing*, vol. 11, p. 53-73, 1997.
- [30] W. J. Wang et P. D. McFadden, "Application of Wavelets to Gearbox Vibration Signals for Fault Detection," *Journal of Sound and Vibration*, vol. 192, p. 927-939, 1996.

- [31] G. Y. Luo, D. Osypiw, et M. Irle, "On-Line Vibration Analysis With Fast Continuous Wavelet Algorithm for Condition Monitoring of Bearing," *Journal of Vibration and Control*, vol. 9, p. 931-947, 2003.
- [32] H. Qiu, J. Lee, J. Lin, et G. Yu, "Wavelet Filter-Based Weak Signature Detection Method and its Application on Rolling Element Bearing Prognostics," *Journal of Sound and Vibration*, vol. 289, p. 1066-1090, 2005.
- [33] G. Vachtsevanos, F. L. Lewis, M. Roemer, A. Hess et B. Wu, "Intelligent Fault Diagnosis and Prognosis for Engineering Systems," Wiley, 2006.
- [34] Shane Butler, "Prognostic Algorithms for Condition Monitoring and Remaining Useful Life Estimation," PhD thesis, National University of Ireland, Maynooth, 2012.
- [35] B. Buchanan et E. Shortliffe, "Rule Based Expert Systems : The Mycin Experiments of the Stanford Heuristic Programming Project," Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA, 1984.
- [36] V. Venkatasubramanian, R. Rengaswamy, S. Kavuri et K. Yin, "A Review of Process Fault Detection and Diagnosis - Part III : Process History Based Methods," *Computers and Chemical Engineering*, vol. 27, p. 327-346, 2003.
- [37] J. Aguilar-Martin, "Knowledge-Based Supervision and Diagnosis of Complex Process," *Proceedings of IEEE International Symposium on Intelligent Control, Intelligent Systems and Semiotics (ISIS'99)*, Cambridge, USA, p. 225-230, 1999.
- [38] M. Fouladirad et I. Nikiforov, "Optimal Statistical Fault Detection With Nuisance Parameters," *Proceedings of the American Control Conference*, Denver, Colorado, USA, 2003.
- [39] H. Hernandez De Leon, "Supervision et Diagnostic des Procédés de Production d'Eau Potable," PhD thesis, Institut National des Sciences Appliquées, Toulouse, 2006.
- [40] R. Fellouah, "Contribution au Diagnostic de Pannes pour les Systèmes Différentiellement Plats," PhD thesis, Institut National des Sciences Appliquées de Toulouse, 2007.
- [41] L. Zadeh, "Fuzzy Sets," *Information and Control*, vol. 8, no. 3, p. 338-353, 1965.
- [42] P. Frank, "Analytical and Qualitative Model-Based Fault Diagnosis, A Survey and Some New Results," *European Journal of Control*, vol. 2, p. 6-28, 1996.
- [43] J. Gertler, "Fault Detection and Diagnosis in Engineering Systems," Marcel Dekker, 1998.
- [44] P. Frank, "Advanced Fault Detection and Isolation Schemes Using Nonlinear and Robust Observers," *Proceedings of 10th IFAC Congress*, vol. 3, Munich, Germany, pages 63-68, 1987.
- [45] M. Staroswiecki et P. Declerck, "Analytical Redundancy in Nonlinear Interconnected Systems by Means of Structural Analysis," *Proceedings of IFAC AIPAC' 89 Symposium*, vol. 2, p. 23-27, 1989.
- [46] S. Narasimhan et G. Biswas, "Model-Based Diagnosis of Hybrid Systems," *Proceedings of the 18th International Joint Conference on Artificial Intelligence (IJCAI'03)*, Acapulco, Mexico, p. 376-381, 2003.
- [47] M. Bayouhd, L. T. Massuyès et X. Olive, "Hybrid Systems Diagnosis By Coupling Continuous and Discrete Event Techniques," *Proceedings of the 17th IFAC World Congress*, Seoul, Corée du Sud, July 2008.

- [48] V. Coquempot, T. El Mezzyani et M. Staroswiecki, "Hybrid Dynamical System Monitoring Using Structured Analytical Redundancy Relations," *Proceedings of 17ème IMACS World Congress*, Paris, France, 2005.
- [49] M. Basseville et L. Nikiforov, "Detection of Abrupt Changes," Prentice Hall, Englewood Cliffs, NJ, 1993.
- [50] R. Isermann, "Supervision, Fault-Detection and Fault-Diagnosis Methods. An Introduction," *Control Engineering Practice*, vol. 5, p. 639–652, 1997.
- [51] M. Staroswiecki et G. C. Varga, "Analytical Redundancy Relations for Fault Detection and Isolation in Algebraic Dynamic Systems," *Automatica*, vol. 37, no. 5, p. 687–699, 2001.
- [52] R. Patton et J. Chen, "A Re-Examination of the Relationship Between Parity Space and Observer-Based Approaches in Fault Diagnosis," *Revue Européenne Diagnostic et Sécurité de Fonctionnement*, vol. 1, no. 2, p. 183–200, 1991.
- [53] D. Maquin, V. Cocquempot, J. Cassar, M. Staroswiecki et J. Ragot, "Generation of Analytical Redundancy Relations for FDI Purposes," *Proceedings of IFAC Symposium on Diagnosis for Electrical Machines, Power Electronics and Drives (SDEMPED'97)*, 1997.
- [54] M. Staroswiecki, V. Cocquempot et J. Cassar, "Observer Based and Parity Space Approaches for Failure Detection and Identification," *Proceedings of IMACSIFAC International Symposium*, p. 536–541, 1991.
- [55] M. Desai et A. Ray, "A Fault Detection and Isolation Methodology," *proceedings of the 20th IEEE Conference on Decision and Control*, p. 1363–1369, 1981.
- [56] A. Schwarte, F. Kimmich et R. Isermann, "Model-Based Fault Detection of A Diesel Engine with Turbo Charger - A Case Study," *Proceedings of Conference Safe Process*, 2003, Washington D.C., USA, 2003.
- [57] J. Magni et P. Mouyon, "On Residual Generation by Observer and Parity Space Approaches," *IEEE Transactions on Automatic Control*, vol. 39, p. 441–447, 1994.
- [58] H. Hammouriand, M. Kinnaert et E. Yaagoubi, "Observer-Based Approach to Fault Detection and Isolation for Non-Linear Systems," *IEEE Transactions and Automatic Control*, vol. 44, p. 1879–1884, 1999.
- [59] J. Gertler, "Diagnosing Parametric Faults: from Parameter Estimation to Parity Relations," *Proceedings of the American Control Conference*, p. 1615–1620, 1995.
- [60] R. Reiter, "A theory of Diagnostic from First Principles," *Artificial Intelligence*, vol. 32, , pages 57–95, 1987.
- [61] J. De Kleer et B. Williams, "Diagnosing Multiple Faults," *Artificial Intelligence*, vol. 32, no. 1, p. 97–130, 1987.
- [62] Byington C.S., M.J. Roemer, G.J. Kacprzyński et T. Galie "Prognostic Enhancements to Diagnostic Systems for Improved Condition-Based Maintenance," *IEEE Aerospace Conference*, Big Sky, USA, 2002.
- [63] Lebold M. et M. Thurston, "Open Standards for Condition-Based Maintenance and Prognostic Systems," *5th Annual Maintenance and Reliability Conference*, Gatlinburg, USA (MARCON 2001).

- [64]Kalgren P.W., C.S. Byington et M.J. Roemer, "Defining PHM, A Lexical Evolution of Maintenance and Logistics," *IEEE AUTOTESTCON Conference Record*, Anaheim, California, 2006.
- [65]Peng Y., M. Dong et M.J. Zuo, "Current Status of Machine Prognostics in Condition-Based Maintenance: A Review," *International Journal of Advanced Manufacturing Technology*, vol. 50(1-4), p. 297-313, 2010.
- [66]Norme ISO 13381-1, " Surveillance et Diagnostic des Machines - Pronostic - Partie 1:Lignes Directrices Générales," *International Standard Organization ISO* (2004).
- [67]M. Lebold et M. Thurston, "Open Standards for Condition-Based Maintenance and Prognostic Systems," *5th Annual Maintenance and Reliability Conference (MARCON 2001)*, Gatlinburg, USA, 2001. 13, 14, 19, 20, 101
- [68]A. Ray, "Stochastic Modeling of Fatigue Crack Damage for Risk Analysis and Remaining Life Prediction," *Journal of Dynamic Systems, Measurement, and Control (ASME)*, vol.121, p.386-393, 1999.
- [69]A. Ray et S. Tangirala, "A Nonlinear Stochastic Model of Fatigue Crack Dynamics," *Probabilistic Engineering Mechanics*, vol. 12, p. 33-40, 1997.
- [70]W.Q Meeker et L.A Escobar, "Statistical Methods for Reliability Data," *Wiley series in probability and statistics*, John Wiley & sons, 23, 30, 1998.
- [71] J. G. Shanthikumar et U. Sumita, "General Shock Models Associated with Correlated Renewal Sequences," *Journal of Applied Probability*, vol. 20, p. 600-614, 1983.
- [72]Y. Li, S. Billington, C. Zhang, T. Kurfess, S. Danyluk et S. Liang, "Adaptive Prognostics for Rolling Element Bearing Condition," *Mechanical Systems and Signal Processing*, vol.13, p.103-113, 1999.
- [73]Y. Li, T. R. Kurfess et S. Y. Liang "Stochastic Prognostics for Rolling Element Bearing," *Mechanical Systems and Signal Processing*, vol.14, p.747-762, 2000.
- [74] J. Qiu, B. B. Set, S. Y. Liang, et C. Zhang, "Damage Mechanics Approach for Bearing Lifetime Prognotics," *Mechanical Systems and Signal Processing*, vol. 16, p. 817-829, 2002.
- [75]D. E. Adams, "Nonlinear Damage Models For Diagnosis and Pognosis in Structural Dynamic Systems," *SPIE Conference Proceeding*, vol. 4733, p. 180-191, 2002.
- [76]D. Chelidze, "Multimode Damage Tracking and Failure Prognosis in Electromechanical System," *SPIE Conference Proceeding*, vol. 4733, p. 1-12, 2002.
- [77]D. Banjevic et A. K. S. Jardine, "Calculation of Reliability Function and Remaining Useful Life for A Markov Failure Time Process," *IMA Journal of Management Mathematics*, vol. 17, p. 115-130, 2006.
- [78]R. B. Chinnam et P. Baruah, "Autonomous Diagnostics and Prognostics Through Competitive Learning Driven HMM-Based Clustering," *Proceedings of the International Joint Conference on Neural Networks*, p. 2466- 2471, 2003.
- [79] Y. Li, C. Zhang, T. R. Kurfess, S. Dahyluk, et S. Y. Liang, "Diagnostics and Prognostics of A Single Surface Defect on Roller Bearings," *Proceedings on the Institution of Mechanical Engineers*, p. 1173-1185, 2000.
- [80]J. Yan, M. Koc, et J. Lee, "A Prognostic Algorithm for Machine Performance Assessment and its Application," *Production Planning and Control Engineering Practice*, vol. 15, p. 796-801, 2004.

- [81] J. Neter, M. H. Kutner, C. J. Nachtsheim, et W. Wasserman, "Applied Linear Statistical Models," Irwin, 1996.
- [82] R. J. Goncalves, M. M. Barata, J. A. Lopes, et A. S. Garcao, "Application of Stochastic Modeling to Support Predictive Maintenance for Industrial Environments," *IEEE International Conference on Systems Man and Cybernetics*, vol. 1, p. 117-122, 1996.
- [83] W. Wang et A. Wong, "Autoregressive model based gear fault diagnosis," *Journal of Vibration and Acoustics*, vol. 124, p. 172-179, 2002.
- [84] C. J. Lu et W. Q. Meeker, "Using Degradation Measures to Estimate A Time-To-Failure Distribution," *Technometrics*, vol. 35, p. 161-174, 1993.
- [85] V. Chan et W. Q. Meeker, "Estimation of Degradation-Based Reliability in Out-Door Environments," Iowa State University.
- [86] C. Li and A. Ray, "Neural Network Representation of Fatigue Damage Dynamics," *Journal of Fatigue Damage Dynamics*, p. 126-133, 1995.
- [87] P. Wang et G. Vachtsevanos, "Fault Prognosis Using Dynamic Wavelet Neural Networks," *Maintenance and Reliability Conference. MARCON*, 1999.
- [88] R. C. M. Yam, P. W. Tse, L. Li, et P. Tu, "Intelligent Predictive Decision Support System for CBM," *The International Journal of Advanced Manufacturing Technology*, vol. 17, p. 383-391, 2001.
- [89] N. Gebraeel, M. Lawley, R. Liu, et V. Parmeshwaran, "Residual Life Predictions from Vibration-Based Degradation Signals: A Neural Network Approach," *IEEE Transactions on Industrial Electronics*, vol. 51, p. 694-700, 2004.
- [90] R. Huang, L. Xi, X. Li, C. Richard Liu, H. Qiu, et J. Lee, "Residual Life Predictions for Ball Bearings Based on Self-Organizing Map and Back Propagation Neural Network Methods," *Mechanical Systems and Signal Processing*.
- [91] P. Wang et G. Vachtsevanos, "Fault Prognostics Using Dynamic Wavelet Neural Networks," *AI EDAM*, vol. 15, p. 349-365, 2001.
- [92] Y.-L. Dong, G. Yu-Jiong, and K. Yang, "Research on the Condition Based Maintenance Decision of Equipment in Power Plant," *Machine Learning and Cybernetics*, 2004.
- [93] R. Ramesh, M. A. Mannan, A. N. Poo, et S. S. Keerthi, "Thermal Error Measurement and Modeling in Machine Tools, Part II. Hybrid Bayesian Network-Support Vector Machine Model," *International Journal of Machine Tools & Manufacture*, vol. 43, pp. 405-419, 2003.
- [94] C. S. Byington, M. Watson, et D. Edwards, "Dynamic Signal Analysis and Neural Network Modeling for Life Prediction of Flight Control Actuators," *Proceedings of the American Helicopter Society 60th Annual Forum*, Alexandria, VA, USA.
- [95] C. S. Byington, M. Watson, et D. Edwards, "Data-Driven Neural Network Methodology to Remaining Life Predictions for Aircraft Actuator Components," in *IEEE Aerospace Conference*, Big Sky, MT, USA, 2004.
- [96] W. Q. Wang, M. F. Golnaraghi, et F. Ismail, "Prognosis of Machine Health Condition Using Neuro-Fuzzy Systems," *Mechanical Systems and Signal Processing*, vol. 18, p. 813-831, 2004.
- [97] J. R. Bock, T. Brotherton, P. Grabill, D. Gass, et J. A. Keller, "On False Alarm Mitigation," in *Proceedings of the IEEE Aerospace Conference*, New York, USA.

- [98] D. Clifton, "Condition Monitoring of Gas-Turbine Engines," Department of Engineering Science, University of Oxford, 2006.
- [99] L. Déchamp, A. Dutech, T. Montroig, X. Qian, D. Racoceanu, I. Rasovska, B. Brézillon, F. Charpillat, J. Y. Jaffray, N. Moine, B. S. Morello, Müller, G. Nguengang, N. Palluat, et L. Pelissier, "On the Use of Artificial Intelligence for Prognosis and Diagnosis in the PROTEUS E-Maintenance Platform," 2005.
- [100] X. Zhang, R. Xu, C. Kwan, S. Y. Liang, Q. Xie, et L. Haynes, "An Integrated Approach to Bearing Fault Diagnostics and Prognostics," *Proceedings of American Control Conference*, Portland, OR, USA, p. 2750-2755, 2005.
- [101] M. Dong et D. He, "A Segmental Hidden Semi-Markov Model (HSMM)-Based Diagnostics and Prognostics Framework and Methodology," *Mechanical Systems and Signal Processing*, p.2248-2266, 2007.
- [102] A. Schömig, O. R. Schömig et O. Rose, "On the Suitability of the Weibull Distribution for the Approximation of Machine Failures," *Proceedings of the Industrial Engineering Research Conference*, vol. 26, Portland OR, 2003.
- [103] Hack-Eun Kim "Machine Prognostics Based on Health State Probability Estimation," PhD thesis, Queensland University of Technology, 2010
- [104] A. Jardine et D. Banjevic, "Interpretation of Inspection Data Emanating from Equipment Condition Monitoring Tools Method and Software," *Fourth International Conference on Mathematical Methods in Reliability Methodology and Practice*. Santa Fe, USA, 2004.
- [105] C. Clarotti, A. Lannoy, S. Odin et H. Procaccia, "Detection of Equipment Aging and Determination of the Efficiency of A Corrective Measure," *Reliability Engineering and System Safety*, vol. 84, p.57-64, 2004.
- [106] D. Kumar et U. Westberg, "Maintenance Scheduling Under Age Replacement Policy Using Proportional Hazards Model and TTTplotting," *European Journal of Operational Research*, vol. 99, pp. 507-515, 1997.
- [107] S. Gasmi, C. E. Love, et W. Kahle, "A General Repair, Proportional-Hazards, Framework to Model Complex Repairable Systems," *IEEE Transactions on Reliability*, vol. 52, pp. 26-32, 2003.
- [108] T. E. Tallian, "Data Fitted Bearing Life Prediction Model for Variable Operating Conditions," *Tribology Transactions*, vol. 42, p. 241, 1999.
- [109] P.-J. Vlok, M. Wnek, et M. Zygmunt, "Utilising Statistical Residual Life Estimates of Bearings to Quantify the Influence of Preventive Maintenance Actions," *Mechanical Systems and Signal Processing*, vol. 18, p. 833-847, 2004.
- [110] B. B. Sassi S. and T. M., "TALAF and THIKAT as innovative time domain indicators for tracking ball bearings," *Proceedings of the 24th Seminar on machinery vibration (M. Thomas, ed.)*, p. 404-419, Canadian Machinery Vibration Association, 2005.
- [111] J. Coble and J.W. Hines, "Identifying Optimal Prognostic Parameters from Data: A Genetic Algorithms Approach," *Annual Conference of the Prognostics and Health Management Society*, 2009.
- [112] J. Coble. (2010) "Merging Data Sources to Predict Remaining Useful Life – An Automated Method to Identify Prognostic Parameters," PhD diss., University of Tennessee. [Online]. http://trace.tennessee.edu/utk_graddiss/683

- [113] F. Camci, K. Medjaher, N. Zerhouni, and P. Nectoux, "Feature Evaluation for Effective Bearing Prognostics," *International Quality and Reliability Engineering*, 2012.
- [114] Linxia Liao, "Discovering Prognostic Features Using Genetic Programming in Remaining Useful Life Prediction," *IEEE Transactions on Industrial Electronics*, vol. 61, no. 5, p. 2464-2472. 2014.

III. 1 INTRODUCTION

Dans le but de situer et limiter le cadre de notre présent travail, nous avons présenté dans le chapitre précédent le concept de diagnostic suivi d'un état de l'art des approches diagnostic. Sachant que le domaine diagnostic s'appuie sur une connaissance plus ou moins approfondie du système à considérer (données historiques, données mesurées en ligne, modèles...) en plus qu'il n'existe pas une manière unique et universelle d'envisager le processus diagnostic, cela nous a conduit à définir le cadre de travail suivant :

Le présent chapitre est destiné essentiellement à améliorer le taux de classification correcte des défauts du moteur à induction tels que les défauts des roulements, les défauts de déséquilibre des phases statoriques et les défauts des cassures de barres rotoriques. Nous avons opté pour une approche diagnostic orientée données vu la disponibilité des données expérimentales d'un banc d'essai sur ce moteur. A cet égard, nous avons proposé une méthode statistique probabiliste à base d'un modèle d'erreur pour sélectionner un nombre optimal de caractéristiques extraites à partir du plan d'ambiguïté lissé par un noyau qui est délibérément conçu pour maximiser la séparabilité entre classes. Ce modèle dépend de deux paramètres, à savoir, le noyau de lissage utilisé pour construire le vecteur de forme et la mesure de distance choisie. Jouissant de la capacité des réseaux de neurones artificiels RNA à apprendre et à généraliser des relations fonctionnelles non linéaires entre les variables d'entrée et de sortie et afin de fournir un mécanisme souple pour l'apprentissage et la reconnaissance des défauts du système, nous avons aussi proposé des classifieurs à base de RNA avec des fonctions d'activation sigmoïdes et ondelettes et comparé leurs performances au classifieur statistique basé sur les distances Mahalanobis et euclidienne. Le reste de ce chapitre est organisé comme suit :

Après une brève présentation sur les méthodes diagnostic le choix de l'espace temps-fréquence comme domaine d'extraction est introduit à la section III. 2, la section III. 3 présente une description des étapes de l'algorithme de classification, les fondements théoriques de la RTF, la construction du noyau et l'introduction du classifieur statistique probabiliste. La section III. 4 porte sur les classifieurs à base de RNA. La section III. 5 présente l'expérimentation et les résultats commentés. Enfin, une conclusion est tirée à la section III. 6.

III. 2 ESPACE TEMPS-FREQUENCE COMME DOMAINE D'EXTRACTION DE CARACTERISTIQUES

Le domaine du diagnostic des défauts est en mouvement continu vers la mise en œuvre des systèmes de surveillance de santé des équipements plus fiables et performants. Plusieurs techniques de surveillance ont été mises au point pour détecter les défauts dans les entraînements électriques. Parmi elles, celles qui reposent sur l'analyse des composantes du spectre de fréquence du courant moteur appelée aussi analyse de signature du courant moteur. Sa facilité d'utilisation a attiré l'attention de beaucoup de chercheurs [1], [2]. D'autres techniques reposent sur l'analyse de vibration, la mesure du bruit acoustique, l'analyse du profil de couple, l'analyse de température, et l'analyse de champ magnétique [3]. Ces

techniques nécessitent des capteurs sophistiqués et coûteux, des installations électriques et mécaniques supplémentaires et une fréquente maintenance.

Récemment, de nouvelles techniques basées sur l'intelligence artificielle (IA) ont été mises en œuvre pour aider à la tâche de la détection et de la prise de décision, à l'aide de nouveaux concepts tels que les algorithmes génétiques [5], colonie de fourmis [6], les réseaux de neurones artificiels (ANN) [7] - [10], la logique floue [11], [12], les réseaux neuro-flous [13], et la transformée en ondelettes [14] - [17]. Les techniques basées sur l'intelligence artificielle IA ne sont pas uniquement destinées à la classification des défauts mais aussi permettent de mesurer la sévérité des défauts. Ces méthodes apprennent hors ligne les signatures pour chaque condition de fonctionnement du moteur et testent en ligne la signature de l'état actuel du moteur à surveiller. Un classifieur compare les signatures qui ont été déjà apprises à la signature générée en ligne afin de classer l'état de fonctionnement actuel du moteur et déterminer la sévérité du défaut.

En règle générale, le diagnostic de système utilise des signaux, soit dans le domaine temporel ou fréquentiel. Dans notre approche, il est peut-être plus informatif de joindre à la fois le domaine temporel au domaine fréquentiel [18].

L'objectif classique de la représentation temps-fréquence RTF, consiste à concevoir une fonction qui décrira la densité d'énergie d'un signal en temps et en fréquence simultanément (c.-à-d. une distribution temps-fréquence). Pour raisonner en termes de classification, il n'est pas nécessairement souhaitable de représenter avec précision la distribution d'énergie d'un signal en temps-fréquence. Il peut être avantageux de concevoir une RTF qui accentue spécifiquement les différences entre les classes (le terme classe est utilisé pour définir un groupe, arbitraire ou autre, des données « semblables »). En outre, un signal peut être représenté par un nombre infini de RTF. Cependant, une RTF optimale pour une tâche de classification s'avère nécessaire pour l'extraction des motifs (caractéristiques pertinentes).

La méthode proposée nécessite deux étapes principales : l'extraction de motif et la classification. L'extraction de motif saisit l'information pertinente des signaux à classer et fournit les motifs comme des entrées au système de classification.

Les représentations temps-fréquence peuvent être caractérisées uniquement par une fonction sous-jacente appelée noyau (kernel). Dans les recherches précédentes, les noyaux ont été obtenus afin de remplir certaines propriétés, telles que la réduction des interférences quadratiques, bien que certaines RTF résultantes peuvent offrir des avantages pour la classification de certains types de signaux. Le but de la détection sensible ou de la classification précise est rarement un objectif explicite de la conception du noyau. Les quelques méthodes qui permettent d'optimiser le noyau à des fins de classification limitent la forme du noyau à des fonctions paramétriques prédéfinies avec symétries qui ne peuvent pas être adaptées rigoureusement à la détection ou la classification [19] - [22].

Pour le processus de classification, la procédure d'optimisation de la RTF via les paramètres du noyau est une tâche à calcul prohibitif. La méthode choisie est la RTF dite « dépendante de la classe du signal (DCS) ». Celle-ci n'exige a priori aucune forme paramétrique et prédéfinie du noyau.

Comme toutes les RTF peuvent être obtenues à partir du plan d'ambiguïté, il serait préférable d'utiliser la RTF optimisée pour la classification directement dans ce plan. Aucune hypothèse

à priori n'est faite quant au lissage nécessaire pour une classification précise. Ainsi, les RTF, de second degré, lissées ne conservent que l'information qui est essentielle pour la classification [23] - [26].

III. 3 ALGORITHME CLASSIFICATION

Le diagnostic des défauts représente un champ d'application approprié pour les méthodes de classification, car son but principal est de parvenir à une cartographie optimale de l'état actuel du système surveillé parmi un ensemble prédéfini d'états du système. L'ensemble des états du système comprend l'état normal et les états défectueux. Pour cette raison, l'utilisation des techniques de classification représente un choix naturel lors de la conception d'un système de diagnostic de défaut.

Pour répondre aux questions principales découlant de la tâche de classification qui sont :

Comment générer les caractéristiques ?

Quel est le nombre optimal de caractéristiques à utiliser?

Comment est faite la conception du classifieur?

Nous avons réparti notre travail en trois parties décrites dans les paragraphes suivants :

III. 3. 1 EXTRACTION DE CARACTERISTIQUES DE DIAGNOSTIC

Dans cette section, nous présentons brièvement la théorie de la RTF. Pour plus de détails, nous recommandons au lecteur d'examiner nos travaux précédents [23], [27].

La RTF dépendante de la classe du signal [26] est définie par

$$RTF [n, k] = F_{n \rightarrow \eta}^{-1} \{ F_{\tau \rightarrow k} \{ A[\eta, \tau] \varphi[\eta, \tau] \} \} \quad (III. 1)$$

où F représente la transformée de Fourier, F^{-1} représente la transformée de Fourier inverse, η représente le décalage de fréquence discrète, τ représente le temps de retard discret, n représente l'indice d'échantillon, et k est la fréquence discrète.

De (III. 1), la fonction caractéristique pour chaque RTF est définie par $A[\eta, \tau] \varphi[\eta, \tau]$. En d'autres termes, pour un signal donné, la RTF peut être mappée de manière unique à partir d'un noyau. La RTF optimisée pour la classification peut être obtenue en lissant le plan d'ambiguïté avec un noyau approprié φ_{opt} , qui est le noyau optimal de classification. Le problème de la conception de la RTF optimale devient équivalent à la conception du noyau optimal de classification $\varphi_{opt}(\eta, \tau)$.

Par conséquent, les caractéristiques peuvent être extraites directement à partir du plan d'ambiguïté lissé $A(\eta, \tau) \varphi(\eta, \tau)$ au lieu de la RTF optimale de classification. Ce raccourci simplifie la complexité du calcul lors de l'extraction des caractéristiques en le réduisant considérablement.

III. 3. 2 CONCEPTION DU NOYAU DE CLASSIFICATION

Pour déterminer et obtenir les N emplacements distincts dans le plan d'ambiguïté, de manière à ce que les valeurs contenues dans ces emplacements soient très similaires aux signaux provenant de la même classe alors qu'ils varient d'une façon significative par rapport aux signaux provenant des classes différentes, nous avons conçu et utilisé pour cette fin le rapport discriminant de Fisher généralisé (FDR)

$$FDR_i(\eta, \tau) = \frac{(m_i[\eta, \tau] - m_{i-remain}[\eta, \tau])^2}{V_i^2[\eta, \tau] + V_{i-remain}^2[\eta, \tau]} \quad (III. 2)$$

$$m_i[\eta, \tau] = \frac{1}{N_i} \sum_{j=1}^{N_i} A_{ij}[\eta, \tau] \quad (III. 3)$$

$$m_{i-remain}[\eta, \tau] = \frac{\sum_{k=i+1}^4 \sum_{j=1}^{N_k} A_{kj}[\eta, \tau]}{\sum_{k=i+1}^4 N_k} \quad (III. 4)$$

$$V_i^2[\eta, \tau] = \frac{1}{N_i} \sum_{j=1}^{N_i} (A_{ij}[\eta, \tau] - m_i[\eta, \tau])^2 \quad (III. 5)$$

$$V_{i-remain}^2[\eta, \tau] = \frac{\sum_{k=i+1}^4 \sum_{j=1}^{N_k} (A_{kj}[\eta, \tau] - m_{i-remain}[\eta, \tau])^2}{\sum_{k=i+1}^4 N_k} \quad (III. 6)$$

où $m_i[\eta, \tau]$ et $m_{i-remain}[\eta, \tau]$ moyennes de la classe i et du reste des classes de localisation (η, τ) respectivement,

$V_i^2[\eta, \tau]$ et $V_{i-remain}^2[\eta, \tau]$ variances de la classe i et du reste des classes de localisation (η, τ) respectivement.

Un noyau $\varphi_{opt}(\eta, \tau)$ est conçu pour chaque tâche de classification. Dans notre procédure de classification, $(C - 1)$ noyaux doivent être conçus pour un système de classification à C -classes. La discrimination entre les différentes classes [25] est faite en séparant la classe « i » de toutes les classes restantes $\{i + 1, \dots, C\}$. L'avantage de la méthode réside dans la séparation optimale entre les différentes classes.

Le FDR est transformé en un noyau φ_{opt} en remplaçant un nombre optimal de N points choisis par « un » et le reste par « zéro ». Cela résulte en une matrice binaire ayant les mêmes dimensions que le plan d'ambiguïté. Par suite, les caractéristiques peuvent être extraites directement par lissage du plan d'ambiguïté d'un signal quelconque par ce noyau optimisé, $A(\eta, \tau) \circ \varphi_{opt}(\eta, \tau)$, où \circ est le produit matriciel élément par élément.

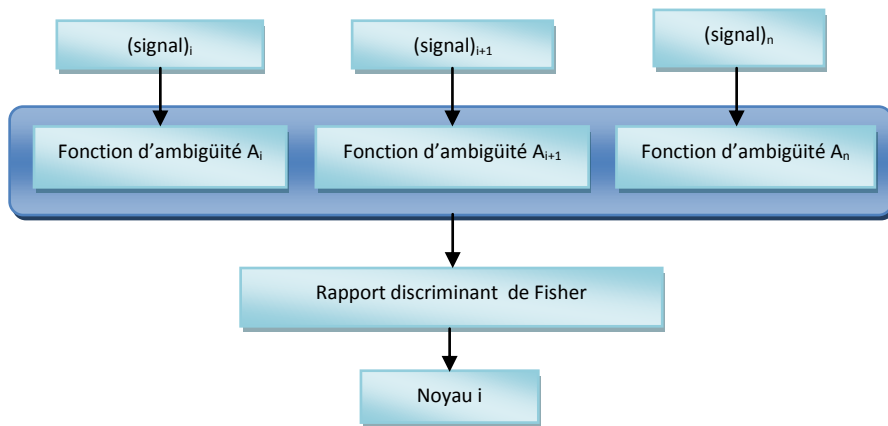


Figure III. 1: Conception du noyau

$$\varphi_{opt}^{(c)}[\eta, \tau] \circ \bar{A}^{(c)}[\eta, \tau] = \begin{cases} \bar{A}^{(c)}[\eta, \tau], & \text{si } \varphi_{opt}^{(c)}[\eta, \tau]=1 \\ 0, & \varphi_{opt}^{(c)}[\eta, \tau]=0 \end{cases} \quad (\text{III. 7})$$

où

$\varphi_{opt}^{(c)}[\eta, \tau]$, noyau optimal d'apprentissage;

$\bar{A}^{(c)}[\eta, \tau]$, classe moyenne du plan d'ambigüité.

Étant donné que le plan de l'ambigüité du signal est symétrique par rapport aux deux dimensions (Doppler retard), seuls les points sur un quart de plan sont considérés.

Les valeurs extraites des N locations sont les caractéristiques de diagnostic de ce signal. Les caractéristiques obtenues sont ensuite ordonnées dans un vecteur appelé vecteur de forme ($FV_{train}^{(c)}(k)$). Les vecteurs de forme issus de chaque classe constituent une base de données pour la phase d'apprentissage de classification.

III. 3. 3 SELECTION DES CARATERISTIQUES.

Le but de l'étape de sélection est de générer un nombre optimal N_{opt} de caractéristiques parmi 10000 positions (§ III. 5. 4) du plan l'ambigüité du signal.

En premier lieu, nous excluons les points des noyaux qui sont fortement corrélés avec des points de noyau de rang supérieur (en utilisant le FDR). Ces points fortement corrélés ne contribuent pas à améliorer la classification [29].

Cette approche sous-optimale améliore les performances de classification globale par rapport aux FDR seul.

La procédure est comme suit :

Soit x_{mk} , $m = 1, 2, \dots, m$ et $k = 1, 2, \dots, L^2/4$, représente la $k^{\text{ème}}$ caractéristique (entrée de la fonction ambigüité) du $m^{\text{ème}}$ signal issu de l'ensemble d'apprentissage. Le coefficient de cross-corrélation entre deux entrées de plan d'ambigüités est donnée par :

$$\rho_{ij} = \frac{\sum_{n=1}^N x_{ni} x_{nj}}{\sqrt{\sum_{n=1}^N x_{ni}^2 \sum_{n=1}^N x_{nj}^2}} \quad (\text{III. 8})$$

On peut remarquer que $|\rho_{ij}| \leq 1$.

Nous pouvons donc ordonner les entrées par la procédure suivante.

- 1) Calculer la matrice FDR comme mentionné plus haut et ordonner le classement des entrées dans l'ordre décroissant; nous appelons cet ordre comme $C(j)$.
- 2) Choisissez l'emplacement du premier point correspondant à la valeur maximale dans la matrice FDR comme entrée caractéristique première; appelons-la x_{i1} .
- 3) Pour sélectionner le deuxième point, on calcule le coefficient de cross-corrélation entre l'entrée x_1 et les $L^2/4$ caractéristiques restantes. Nous choisissons alors l'entrée x_{i2} selon les critères suivants:

$$i_2 = \arg \max_j \{\alpha_1 C(j) - \alpha_2 |\rho_{i_1 j}|\}, \quad i \neq j \quad (\text{III. 9})$$

Ici, α_1 et α_2 sont les facteurs de pondération qui pondèrent les termes relatifs mentionnés plus haut.

Cette procédure itérative continue jusqu'à ce que nous déterminons les x_{ik} , $k = 3, \dots, L^2/4$, tel que

$$i_k = \arg \max_j \left\{ \alpha_1 C(j) - \frac{\alpha_2}{k-1} \sum_{r=1}^{k-1} |\rho_{i_r, j}| \right\}, \quad i_r = j, \quad r = 1, 2, 3 \dots k-1 \quad (\text{III. 10})$$

En d'autres termes, nous pondérons et nous soustrayons la corrélation moyenne avec toutes les caractéristiques précédemment sélectionnées.

Les entrées qui correspondent aux endroits où le noyau a une valeur zéro sont éliminées.

**PROPOSITION D'UN MODELE DE PROBABILITE D'ERREUR POUR
L'OPTIMISATION DE LA DIMENSION DU VECTEUR DE FORME**

Le but du modèle de probabilité d'erreur proposé est d'obtenir un nombre optimal N_{opt} de caractéristiques de diagnostic (optimiser la dimension du vecteur de forme) par minimisation de ce modèle.

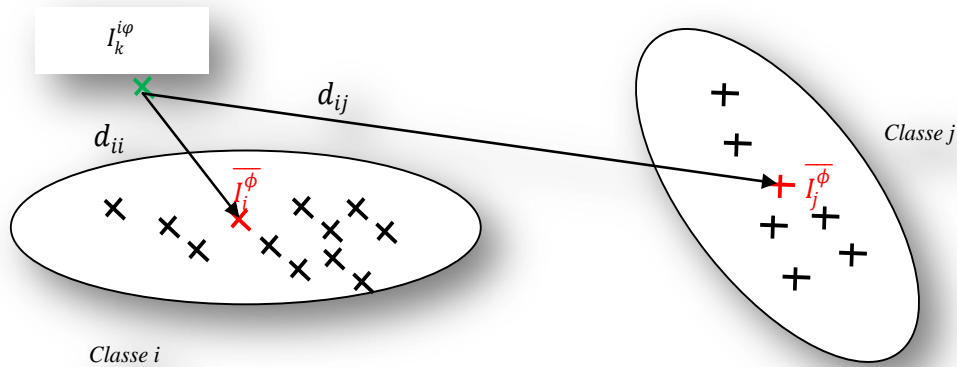


Figure III. 2: Normalement $I_k^{i\phi}$ est affecté à la classe i .

Soit $I_k^{i\phi}$ un vecteur de forme représentant un signal à classer (après calcul de son plan d'ambiguïté et son filtrage par le noyau ϕ construit dans la phase d'apprentissage).

Et soient C_i , les différentes classes, avec $i = 1, \dots, N$ (indice de classes)

Chaque classe C_i possède un ensemble de signaux d'apprentissage $X_i = \{I_1^{i\phi}, \dots, I_{n_i}^{i\phi}\}$ de dimension n_i ; c.à.d. $I_k^{i\phi}$; $k = 1, \dots, n_i$. (base de données construite).

Supposant qu'elles sont équiprobables c.à.d. $I_k^{i\phi}$ peut être affecté à n'importe quelle classe :
 \Rightarrow la probabilité de classification pour chaque classe est donc $P(i) = \frac{1}{N}$

La règle de décision est la suivante :

$I_k^{i\phi}$ serait affecté à C_i si :

$$\hat{i} = \arg \min_{i=1, \dots, N} d \left(I_k^{i\phi}, \bar{I}_i^{\phi} \right) \quad (\text{III. 11})$$

d : est une distance de mesure choisie.

\bar{I}_i^{ϕ} : est la moyenne des signaux d'apprentissage de la classe C_i

Si l'indice du résultat de classification est réellement \hat{i} , c.à.d. que la classe C_i qui l'emporte. (on commet une erreur si on affecte $I_k^{i\phi}$ à d'autres classes que C_i).

Donc, la probabilité d'erreur (qui dépend de la distance choisie et du noyau) peut être écrite sous la forme :

$$P(d, \varphi) = P[\hat{t} \neq i]$$

Le théorème de la probabilité totale nous donne [22]:

$$\begin{aligned} &= \sum_{i=1}^N P[\hat{t} \neq i|i]P(i) \\ &= \frac{1}{N} \sum_{i=1}^N P[\hat{t} \neq i|i] \end{aligned}$$

L'événement erreur $\hat{t} \neq i$ est l'union de tout évènements $\hat{t} = j$ avec $j \neq i$ avec $j \in \{1, \dots, N\}$.

Donc,

$$P[\hat{t} \neq i|i] = \sum_{\substack{j=1 \\ j \neq i}}^N P[\hat{t} = j|i]$$

Par conséquent :

$$\begin{aligned} P(d, \varphi) &= \frac{1}{N} \sum_{i=1}^N \sum_{\substack{j=1 \\ j \neq i}}^N P[\hat{t} = j|i] \\ P[\hat{t} = j|i] &= P(e < 0) \\ P(d, \varphi) &= \frac{1}{N} \sum_{i=1}^N \sum_{\substack{j=1 \\ j \neq i}}^N P(e < 0) \end{aligned} \quad (\text{III. 12})$$

Par hypothèse que l'erreur e_{ij} possède une distribution Gaussienne ;

$$\begin{aligned} e_{ij} &= d_{ij} - d_{ii} \\ e_{ij} = d_{ij} - d_{ii} &= d(I_k^{i\varphi}, \bar{I}_j^\varphi) - d(I_k^{i\varphi}, \bar{I}_i^\varphi) \end{aligned} \quad (\text{III. 13})$$

d_{ij} : c'est la distance calculée entre un signal provenant de la classe i , $I_k^{i\varphi}$, en utilisant un noyau φ et la moyenne de la classe j ; \bar{I}_j^φ .

d_{ii} : c'est la distance entre $I_k^{i\varphi}$ calculé en utilisant un noyau φ et la moyenne de la classe i , \bar{I}_i^φ .

Si $d_{ii} > d_{ij}$; c.à.d. qu'on a mal classé le signal $I_k^{i\varphi} \Rightarrow e_{ij} < 0$.

par hypothèse que l'erreur e_{ij} possède une distribution Gaussienne :

$$\begin{aligned} P(e_{ij} < 0) &= \int_{-\infty}^0 \frac{1}{\sqrt{2\pi}\sigma_{ij}} \exp\left(-\frac{1}{2}\left[\frac{u-m_{ij}}{\sigma_{ij}}\right]^2\right) du \\ &= Q\left(\frac{m_{ij}}{\sigma_{ij}}\right) \end{aligned} \quad (\text{III. 14})$$

Q : Q-fonction.

$$Q(c) = \int_c^{\infty} \frac{1}{(2\pi)^{1/2}} \exp\left(-\frac{1}{2}u^2\right) du$$

Q-fonction est reliée à la fonction d'erreur *erf* par :

$$Q(c) = \frac{1}{2} - \frac{1}{2} \operatorname{erf}\left(\frac{c}{\sqrt{2}}\right)$$

m_{ij} : c'est la moyenne de l'erreur.

$$\hat{m}_{ij} = \frac{1}{n_i} \sum_{k=1}^{n_i} \left[d\left(I_k^{i\varphi}, \bar{I}_j^{\varphi}\right) - d\left(I_k^{i\varphi}, \bar{I}_i^{\varphi}\right) \right] \quad (\text{III. 15})$$

σ_{ij}^2 : c'est la variance de l'erreur.

$$\hat{\sigma}_{ij}^2 = \frac{1}{n_i} \sum_{k=1}^{n_i} \left[d\left(I_k^{i\varphi}, \bar{I}_j^{\varphi}\right) - d\left(I_k^{i\varphi}, \bar{I}_i^{\varphi}\right) - \hat{m}_{ij} \right]^2 \quad (\text{III. 16})$$

On peut estimer ces paramètres à partir de l'ensemble d'apprentissage en utilisant les estimateurs de la moyenne et de la variance des signaux des classes *i* et *j* respectivement.

Donc, le modèle de probabilité d'erreur est exprimé par

$$\tilde{P}(d, \varphi) = \frac{1}{N} \sum_{i=1}^N \sum_{j=1}^N Q\left(\frac{\hat{m}_{ij}}{\hat{\sigma}_{ij}}\right) \quad i \neq j \quad (\text{III. 17})$$

Algorithme

- En partant du choix de la distance *d* ; exemple distance Mahalanobis ou euclidienne ou autre,
- Pour $l = 1, \dots, L^2/4$ (*l* est la dimension du vecteur de forme qu'on cherche à optimisée et *L* est la dimension du plan d'ambiguité tenant compte de la symétrie suivant l'axe doppler et l'axe retard, \Rightarrow on garde seulement le quart supérieur $\equiv L^2/4$) on calcule les distances d_{ij} et d_{ii} puis les \hat{m}_{ij} et $\hat{\sigma}_{ij}$ pour tout $i, j = 1, \dots, n_i$ avec n_i est le nombre de vecteur de forme d'une classe donnée.
- On calcule $\tilde{P}(d, \varphi)$.
- à la sortie de la boucle, on choisi *l* (c.à.d. le nombre optimal des caractéristiques) qui donne le minimum de probabilité d'erreur.

N.B. : dans notre cas le noyau n'est pas soumet à la procédure d'optimisation (car il n'est pas paramétré).

L'un des résultats obtenus par cet algorithme est donné à la figure III. 3 :

Exemple :

Modèle de probabilité pour 2 classes :

$$\tilde{P}(d, \varphi) = \frac{1}{2} \left[Q\left(\frac{\hat{m}_{12}}{\hat{\sigma}_{12}}\right) + Q\left(\frac{\hat{m}_{21}}{\hat{\sigma}_{21}}\right) \right]$$

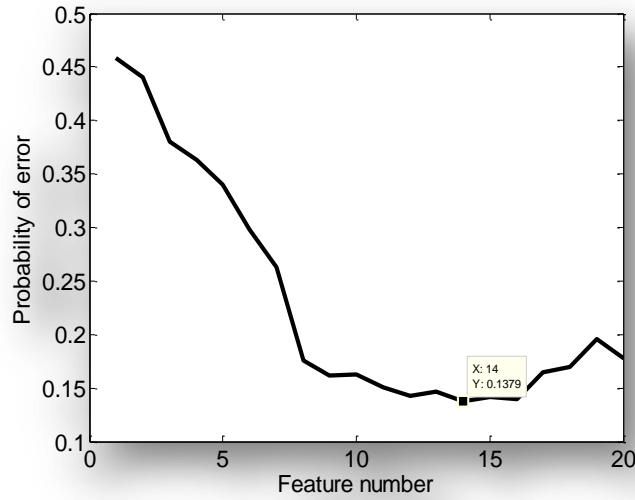


Figure III. 3: Sélection d'un nombre optimal de caractéristiques utilisant la distance Mahalanobis (noyau défaut stator)

Pour le choix d'une mesure de distance, si on suppose que les classes des vecteurs de forme $x = \varphi_{opt}^{(c)} o A_x$, sont équiprobables et normalement distribuées, on peut, alors, montrer [29] que le classifieur optimal pour les classes distribuées en tant que tel peut être exprimé comme

$$d_{quad} \left(\left(\varphi_{opt}^{(c)} o A_x \right), \left(\varphi_{opt}^{(c)} o \bar{A}^{(c)} \right) \right) = \left(\varphi_{opt}^{(c)} o A_x - \varphi_{opt}^{(c)} o \bar{A}^{(c)} \right)^T \sum_c^{-1} \left(\varphi_{opt}^{(c)} o A_x - \varphi_{opt}^{(c)} o \bar{A}^{(c)} \right) + \ln |\Sigma_c| - 2 \ln \pi_c \quad (\text{III. 18})$$

Connue sous le nom fonction discriminante quadratique, où $(\cdot)^T$ désigne la matrice transposée, et (\cdot) désigne le déterminant.

La probabilité a priori de la classe C , π_c , est supposée connue. La classe moyenne $\bar{A}^{(c)}$ et la matrice de covariance de la classe Σ_c sont estimées à partir des données d'apprentissage, ce qui peut être accompli en utilisant une approche d'estimation du maximum de vraisemblance. Si l'on suppose en outre que les classes non seulement sont équiprobables mais partagent aussi la même matrice de covariance, nous pouvons ignorer le dernier terme de (7) résultant en la distance de Mahalanobis

$$d_M \left(\left(\varphi_{opt}^{(c)} o A_x \right), \left(\varphi_{opt}^{(c)} o \bar{A}^{(c)} \right) \right) = \left(\varphi_{opt}^{(c)} o A_x - \varphi_{opt}^{(c)} o \bar{A}^{(c)} \right)^T \Sigma_c^{-1} \left(\varphi_{opt}^{(c)} o A_x - \varphi_{opt}^{(c)} o \bar{A}^{(c)} \right) \quad (\text{III. 19})$$

En outre, si la matrice de covariance est de la forme $\Sigma = \sigma^2 I$, où I est la matrice identité. Alors la mesure de la distance quadratique réduite au carré de la distance euclidienne bien connue

$$d_{eucl} \left(\left(\varphi_{opt}^{(c)} o A_x \right), \left(\varphi_{opt}^{(c)} o \bar{A}^{(c)} \right) \right) = \left\| \left(\varphi_{opt}^{(c)} o A_x - \varphi_{opt}^{(c)} o \bar{A}^{(c)} \right)^T \right\| \quad (\text{III. 20})$$

La distance choisie est calculée entre FV_{AV} et FV_{TEST} .

Par conséquent, FV_{TEST} est affecté à la classe C_i selon :

$$C_i \Leftrightarrow i = \underset{i=1, \dots, C}{\operatorname{arg\,min}} \{d(FV_{\text{TEST}}, FV_{\text{AV}})\} \quad (\text{III. 21})$$

FV_{TEST} est le vecteur de forme du signal test.

FV_{AV} est le vecteur de forme moyen du signal d'apprentissage.

III. 4 CLASSIFICATION PAR RESEAUX NEURONES ARTIFICIELS (RNA)

Les RNA, en raison de leur capacité à apprendre et à généraliser des relations fonctionnelles non linéaires entre les variables d'entrée et de sortie, ils fournissent un mécanisme souple pour l'apprentissage et la reconnaissance des défauts du système [7] - [10].

Parmi une variété d'architectures, les réseaux de neurones récurrents sont couramment utilisés dans des tâches de reconnaissance de formes. Il ya aussi de nombreux types et d'architectures de réseaux de neurones fondamentalement différents dans la manière dont ils apprennent; les détails sont bien documentés dans la littérature [7].

L'algorithme d'apprentissage fréquemment utilisé dans des problèmes de classification est l'algorithme de rétro-propagation (BP), qui est aussi utilisé dans le présent travail.

Dans ce chapitre, deux réseaux de neurones applicables à la classification des défauts de la machine à induction seront utilisés pour la conception de classifieurs, à savoir, le réseau de neurones « feedforward » (FFNN) à fonction d'activation sigmoïde et le réseau de neurone à base d'ondelettes « Wavelet Neural Network » (WNN).

III. 4. 1 RESEAU DE NEURONES « FEEDFORWARD » (FFNN)

Chacun des trois classifieurs FFNN adoptés dans notre algorithme de classification pour les trois types de défauts possède trois couches (entrée - cachée - sortie). Contrairement à l'approche statistique, de nombreuses expériences de classification ont été menées pour déterminer la structure des réseaux de neurones optimisés dans laquelle la taille du réseau doit être gardée aussi petite que possible. Ceci est imposé par la capacité de la généralisation que doit posséder le réseau.

La structure du FFNN pour discriminer le défaut statorique est 16-5-1 (nombre de nœud de la couche d'entrée- nombre de nœud de la couche cachée - nombre de nœud de la couche de sortie); celle du défaut rotorique est 15-6-1, pour le défaut roulements est 15-6-1. Les fonctions de transfert et d'apprentissage adoptées pour la FFNN sont les suivantes [23], [28] : La fonction de transfert sigmoïde tangente hyperbolique comme fonction de transfert de la couche cachée dont l'implémentation en Matlab est donnée par :

$$F(x) = \frac{2}{(1 + \exp(-2 \cdot x))} - 1$$

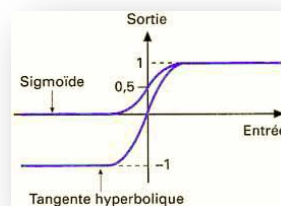


Figure III. 4: Fonction de transfert

La fonction de transfert linéaire comme fonction de transfert de la couche de sortie, l'algorithme de Levenberg-Marquardt BP en tant qu'algorithme d'apprentissage du réseau et la fonction erreur quadratique moyenne (MSE) comme fonction d'évaluation de performance. Les entrées du RNA sont les valeurs des caractéristiques normalisées, et la sortie du RNA est la décision binaire faite.

III. 4. 2 LES RESEAUX DE NEURONES A ONDELETTES « WAVELET NEURAL NETWORKS » WNN

Dans ce chapitre, un réseau de neurone feed-forward à base d'ondelettes WNN à une seule couche cachée a été conçu avec fonction d'activation à base d'ondelette nommée chapeau mexicain

$$\psi(x) = (d - \|X\|^2)e^{-\frac{\|X\|^2}{2}}$$

$\|X\|$ représente la norme euclidienne de X et $d = \dim(X)$.

Selon [31]-[32], la construction d'un WNN comporte deux étapes:

Tout d'abord, construire une bibliothèque W d'ondelettes de versions discrètement dilatées et décalées de la fonction ondelette mère ψ

$$W = \{\psi_i: \psi_i(x) = \alpha_i \psi((a_i x - t_i))\}, \quad (\text{III. 22})$$

$$\alpha_i = \left(\sum_{k=1}^N [\psi(a_i(x_k - t_i))]^2 \right)^{-\frac{1}{2}}, \quad i = 1 \dots L.$$

où x_k est l'entrée échantillonnée, L est le nombre d'ondelettes dans W , a_i et t_i sont respectivement les coefficients dilatation et translation.

Ensuite, sélectionner les M meilleures ondelettes sur la base des données d'apprentissage parmi les ondelettes de la bibliothèque W , afin de construire la régression

$$f_M(x) = \sum_{i \in I} u_i \psi_i(x), \quad (\text{III. 23})$$

où I est un sous-ensemble de M -éléments de l'ensemble d'indices $\{1, 2, \dots, L\}$ et $M \leq L$.

Deuxièmement, minimiser la fonction de coût.

$$J(I) = \min_{u_i, i \in I} \frac{1}{N} \sum_{k=1}^N \left(y_k - \sum_{i \in I} u_i \psi_i(x) \right)^2, \quad (\text{III. 24})$$

En principe, une telle sélection peut être effectuée en examinant tous les sous-ensembles M -éléments de W , mais la combinaison de tous les sous-ensembles possibles est généralement très grande, et un examen exhaustif peut ne pas être réalisable dans la pratique. Certaines solutions non optimales et heuristiques doivent être considérées. Dans ce qui suit, nous proposons d'appliquer certaines de ces procédures heuristiques.

1) SELECTION A BASE RESIDUELLE :

L'idée de cette méthode est de choisir, pour la première étape, l'ondelette qui correspondent le mieux aux données d'apprentissage \mathcal{O}_1^N puis sélectionner de manière itérative l'ondelette qui correspond le mieux à la valeur résiduelle du fitting de l'étape précédente.

Définir un résidu initial $\gamma_0(k) = y_k, k = 1, \dots, N$, avec y_k les observations de sortie dans \mathcal{O}_1^N .

Mettre $f_0(x) \equiv 0$, à l'étape $I (I = 1, \dots, M)$, rechercher, parmi W , l'ondelette ψ_j qui minimise

$$J(\psi_j) = \frac{1}{N} \sum_{k=1}^N \left(\gamma_{i-1}(k) - \sum_{i \in I} u_j \psi_j(x_k) \right)^2, \quad (\text{III. 25})$$

où

$$u_j = \left(\sum_{k=1}^N (\psi_j(x_k))^2 \right)^{-1} \sum_{k=1}^N \psi_j(x_k) \gamma_{i-1}(k), \quad (\text{III. 26})$$

et $\gamma_{i-1}(k), k = 1, \dots, N$ sont les résidus de l'étape $i - 1$.

Notez que $l_i = \arg \min_{1 \leq j \leq L} J(\psi_j)$, puis, $\psi_{l,i}$ est l'ondelette sélectionné à l'étape i .

Mettre à jour f_i et γ_i suivant :

$$\begin{aligned} f_i(x) &= f_{i-1}(x) + u_{l,i} \psi_{l,i}(x_k) \\ \gamma_i(k) &= \gamma_{i-1}(k) - u_{l,i} \psi_{l,i}(x_k), \quad k = 1 \dots N \end{aligned}$$

2) SELECTION PAS A PAS PAR ORTHOGONALISATION

La procédure ci-dessus fondée sur la sélection à base de résidus [30], [31] ne considère pas explicitement l'interaction ou la non-orthogonalité des ondelettes dans W . L'idée de cette méthode consiste à sélectionner, pour la première étape, l'ondelette dans W qui correspond le mieux aux données d'apprentissage \mathcal{O}_1^N puis sélectionner de manière itérative l'ondelette qui correspond le mieux \mathcal{O}_1^N tout en travaillant avec les ondelettes précédemment sélectionnés.

Le nombre d'ondelettes M est choisi comme étant le minimum de la dite critère d'erreur de prédiction finale de Akaike (FPE) [31]

$$J_{FPE}(\hat{f}) = \frac{1 + n_{pa}/N}{1 - n_{pa}/N} \frac{1}{2N} \sum_{k=1}^N (\hat{f}(x_k) - y_k)^2, \quad (\text{III. 27})$$

où n_{pa} est le nombre de paramètres dans l'estimateur

$$n_{pa} = h(J + 2) + d + 1.$$

h est le nombre d'ondelettes dans le réseau, et d est la dimension du vecteur d'entrée.

Après la construction du WNN initial, il est en outre entraîné par les algorithmes de descente du gradient tels que les moindres carrés pour minimiser l'erreur quadratique moyenne

$$J(w) = \frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}(w))^2, \quad (\text{III. 28})$$

où $\hat{y}(w)$ est la sortie réelle d'un WNN entraîné au vecteur de poids fixe w .

III. 5 DESCRIPTION DU ROCEDE EXPERIMENTAL

III. 5. 1 PRESENTATION DU BANC MOTEUR

Le banc d'essai est réalisé au niveau du laboratoire Ampère à l'université Lyon 1, Lyon France. Le banc est composé d'un moteur asynchrone triphasé à cage d'écureuil Leroy Somer LS132S, IP 55, classe F, T°C normalisée = 40°C, régime nominal S1. Ce moteur est caractérisé par :

- La tension nominale entre phases : 400 V
- La fréquence d'alimentation : 50 Hz
- La vitesse nominale : 1440 tr/min
- La puissance utile nominale : 5.5 kW
- Le facteur de puissance : $\cos\phi = 0.84$
- Le courant nominal : 11,4 A
- Le nombre de paires de pôles : $p = 2$
- La résistance stator par phase : 1,315 Ω
- Le nombre d'encoches au rotor $N_r = 28$
- Le nombre d'encoches au stator $N_s = 48$

Les enroulements au stator sont couplés en étoile. Le moteur est chargé par un frein à poudre. Son couple maximal (100 Nm) est atteint à la vitesse nominale. Ce frein est dimensionné pour dissiper une puissance maximale de 5 kW. La figure III. 5 présente le banc moteur.

III. 5. 2 ACQUISITIONS DES DONNEES

Les signaux enregistrés par le système d'acquisition sont :

- les trois courants d'alimentation de la machine,
- les trois tensions d'alimentation de la machine,
- la vitesse de rotation,
- les vibrations dans les directions suivantes :
 - direction verticale et perpendiculaire à l'axe du moteur, avec un accéléromètre côté opposé accouplement,
 - direction parallèle à l'axe du moteur, côté opposé accouplement,
 - direction horizontale et perpendiculaire à l'axe du moteur, côté opposé accouplement.

Le système d'acquisition utilisé pour mesurer ces signaux possède :

- Huit entrées différentielles dites voies rapides qui peuvent être échantillonnées jusqu'à 10 MHz sur 14 bits. Ces entrées ont servi à mesurer les tensions ainsi que les courants.
- Huit entrées en mode commun dites voies lentes sur lesquelles la fréquence d'échantillonnage peut atteindre 1 MHz sur 16 bits. Une de ces entrées a servi à mesurer la vitesse de rotation.
- Quatre entrées sont réservées aux accéléromètres mesurant les vibrations. Les entrées "accéléromètres" offrent aussi la possibilité d'obtenir une fréquence d'échantillonnage de 100 kHz. Les accéléromètres piézo-électriques sont alimentés par des sources de courants constants fournissant entre 1 et 10 mA sous 28 V.

La figure III. 6 présente le dispositif mis en place pour l'acquisition des signaux: ensemble machine, charge et acquisition des signaux.

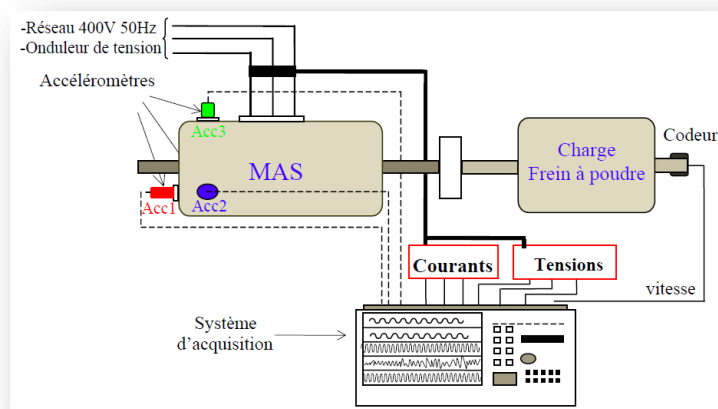


Figure III. 5: Système d'acquisition des signaux

Toutes les acquisitions seront réalisées en régime permanent sur une durée de 5 secondes, avec une fréquence d'échantillonnage de 20 kHz. Soit 100.000 points pour chacun des signaux mesurés.

III. 5. 3 MODES DE FONCTIONNEMENT ETUDIES

Les modes de fonctionnement ayant servi à valider la procédure de diagnostic sont repris dans le tableau III. 1. Les défauts ont été créés au stator et au rotor de la machine asynchrone et ceci pour plusieurs niveaux de charge.

Au niveau des rotors, les barres ont été cassées en perçant les barres des cages d'écurueil, figure III. 8. L'usure obtenue sur les roulements est une usure naturelle d'utilisation (roulements fournis par la société SECCO), figure III. 7. Le déséquilibre d'alimentation au stator est obtenu à l'aide d'un alternostat placé sur une des phases du réseau.

Une partie de ces mesures va servir à construire l'ensemble d'apprentissage ainsi que les ensembles de test.

III. 5. 4 ENSEMBLE D'APPRENTISSAGE

Pour surmonter le problème de la charge, on a proposé d'utiliser les valeurs à vide et en pleine charge pour la phase d'apprentissage, à savoir 0% et 100%, et tester les valeurs de charge intermédiaires (25%, 50% et 75%). Ce choix permet, d'un coté de tester la base d'apprentissage et de l'autre coté permettre de s'affranchir du niveau de charge. Par conséquent, chaque classe de l'ensemble d'apprentissage (les trois défauts et la machine saine) est composée de dix signaux de courant à vide et dix autres signaux de pleine charge. Par conséquent, nous avons 20 signaux d'apprentissage pour chaque classe.

Chaque signal est sous-échantillonné avec un taux de 50 (de 20KHz à 400 Hz). Seule la gamme des fréquences nécessaires est conservée. Par sous-échantillonnage, la dimension du signal a été considérablement réduite. La dimension du plan d'ambiguïté est ($200 \times 200 = 40\ 000$) points. En considérant la symétrie par rapport à l'origine, on ne retient que le quart de plan d'ambiguïté, ce qui correspond à $n = 10\ 000$.

Nous avons conçu trois noyaux: un noyau pour défaut statorique, un autre pour le défaut rotorique, et le dernier pour le défaut roulement. Les emplacements des points de plus fort

contraste de Fisher sont représentés dans le plan retard-Doppler (figure III. 9). Les caractéristiques des emplacements des points provenant de chaque signal sont rangés dans des vecteurs de forme $\{FV_1, \dots, FV_N\}$ pour la phase d'apprentissage.

Tableau III. 1: Modes de fonctionnement pour constituer l'ensemble d'apprentissage et de test [34]

	Mode de fonctionnement	Types de défauts	Taux de charge (Pour une puissance nominale : $P_u = 5.5$ kW)	Alimentation
Machine asynchrone	Machine saine	Aucun défaut	0%, 25%, 50%, 75%, 100%	Réseau et onduleur
	1 barre cassée au rotor	Défaut rotor (barres cassées)	0%, 25%, 50%, 75%, 100%	Réseau et onduleur
	3 barres cassées au rotor			
	4 barres cassées au rotor			
	Usure progressive (vieillessement naturel)	Défaut roulement	0%, 25%, 50%, 75%, 100%	Réseau et onduleur
	5% de déséquilibre sur une phase (Vb)	Déséquilibre d'alimentation au stator	0%, 25%, 50%, 75%, 100%	Réseau
	10% de déséquilibre sur une phase (Vb)			
	20% de déséquilibre sur une phase (Vb)			
	40% de déséquilibre sur une phase (Vb)			



Figure III. 6: Banc d'essai du moteur à induction.



Figure III. 7: Défauts roulements.



Figure III. 8: Rotors avec des barres cassées.

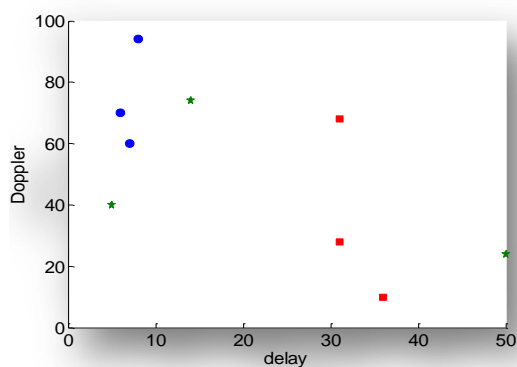


Figure III. 9: Emplacement des caractéristiques dans le plan Doppler-retard pour chacun des trois noyaux

III. 5. 5 DECISION PAR MESURE DE DISTANCE

Dans nos travaux précédents, le vecteur de forme est choisi pour avoir une dimension arbitraire. En appliquant le modèle d'erreur probabiliste proposé et pour chacune des mesures de distance choisies, la probabilité d'erreur est directement influencée par le nombre de caractéristiques. Nous tenons à maintenir le nombre de caractéristiques aussi petit que possible. À partir du tracé de la figure III. 10, nous voulons rechercher le plus petit nombre de caractéristiques qui nous donne une probabilité d'erreur acceptable. Les résultats sont reportés au tableau III. 2 pour les deux distances sélectionnées à savoir, distance euclidienne et Mahalanobis et pour chaque noyau.

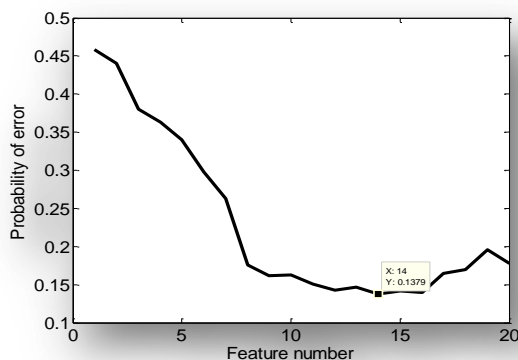


Figure III. 10: Sélection du nombre optimal de caractéristiques utilisant la distance Mahalanobis (noyau défaut stator)

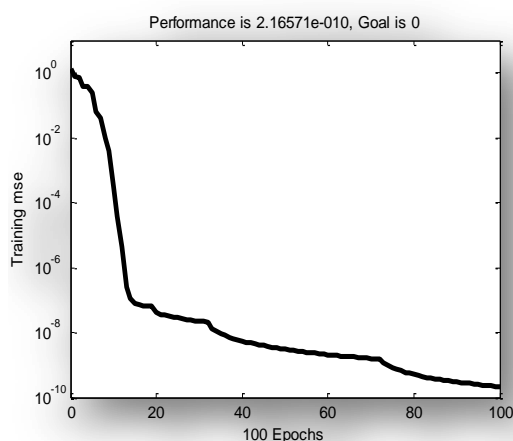


Figure III. 11: Courbe d'apprentissage pour le ANN utilisant l'algorithme Levenberg Marquardt.

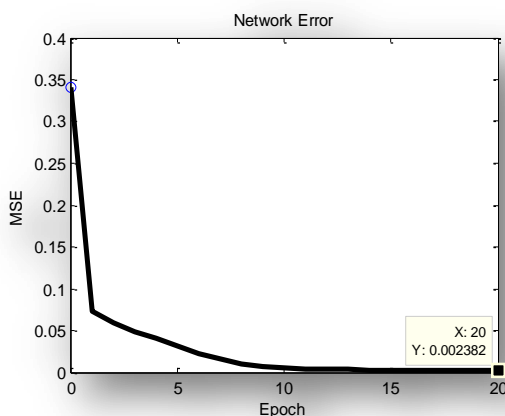


Figure III. 12: Courbe d'apprentissage pour le WNN utilisant l'algorithme gradient conjugué

La règle de décision d'affectation du signal est faite selon l'équation (III. 21). Nous avons testé les signaux qui ne sont pas utilisés dans la phase d'apprentissage pour les trois défauts: défaut roulement, défaut stator et défaut rotor pour différents niveaux de charge (25%, 50% et 75%). Cinq signaux sont employés pour la phase d'apprentissage et les dix signaux restant pour la phase de test. Cette opération est réalisée pour chaque type de défaut et à différents niveaux de charge. Ainsi, nous pouvons utiliser jusqu'à 15 signaux de test pour chaque défaut. Les résultats de performance du classifieur statistique sont présentés au tableau III. 2.

Tableau III. 2: Résultats de classification de la distance Mahalanobis (Mahal) et Euclidienne (eucl) :

	Probabilité d'Erreur Prédite		Erreur Réelle		Nombre Optimal de Caractéristiques	
	Mahal	eucl	Mahal	eucl	Mahal	eucl
FDR1	0,16	0,34	1/15	5/15	9	17
FDR2	0,0952	0,1582	3/15	6/15	18	8
FDR3	0,0591	0,442	2/15	5/15	9	8

D'après le tableau III. 2, nous pouvons voir que la distance de Mahalanobis, comme critère de décision, est plus intéressante que la distance euclidienne.

La faible différence entre la probabilité d'erreur prédite et l'erreur réelle est principalement due à la taille limitée de l'ensemble d'apprentissage et à sa de distribution non Gaussienne.

III. 5. 6 DECISION PAR RESEAUX DE NEURONES

Contrairement à l'approche statistique, de nombreuses expériences de classification ont été menées pour déterminer les structures des réseaux neurones optimisés dont la taille du réseau doit être maintenue aussi réduite que possible. Cela est imposé par la capacité de généralisation que le réseau doit posséder.

Tableau III. 3: Résultats de classification avec les différents classifieurs:

	Noyau 1	Noyau 2	Noyau 3
FFNN	0/15	1/15	1/15
WNN	3/15	5/15	6/15
Mahal (stat.)	1/15	3/15	2/15

Si le nombre de nœuds dans la couche cachée du réseau est trop faible, le WNN peut ne pas tenir compte de la relation de la fonction complexe entre les données d'entrée et la valeur de sortie. En revanche, un grand nombre peut créer un réseau complexe qui pourrait conduire à une très grande erreur de sortie causée par un sur-ajustement de l'échantillon d'apprentissage. Il a été noté que la meilleure performance a été obtenue pour l'ensemble d'apprentissage avec les modèles dont la couche cachée possède cinq neurones ou plus figure III. 12.

A l'encontre de peu d'études de la classification par WNN qui ont montré une certaine amélioration dans la performance, notre étude démontre que ce n'est pas toujours aussi évidant (tableau III. 3). Les résultats présentés montrent que l'ANN avec l'algorithme de Levenberg-Marquardt a une nette supériorité sur le WNN; cela est dû principalement à la grande taille du vecteur d'entrée (jusqu'à 15) car les WNN sont généralement limités aux problèmes de dimension d'entrée réduite.

D'autre part, les résultats de classification correcte obtenus à partir du classifieur statistique améliorée avec la mesure de distance Mahalanobis comme critère de décision, sont semblables à ceux du classifieur à base ANN entraîné avec l'algorithme de Levenberg-Marquardt et avec sigmoïde comme fonction de transfert. La différence minimale réside dans le deuxième noyau qui est utilisé pour la séparation de la classe de défaut des barres rotoriques du reste des classes ; classe défaut roulement et machine saine. Quatorze parmi les 15 signaux de test ont été correctement classés par le classifieur ANN au lieu de 12 pour l'approche statistique. En conclusion, sur la base de cette comparaison expérimentale et sur bien d'autres études, nous soulignons qu'aucun algorithme de classification ne peut être considéré comme idéal. Par conséquent, on pourrait exploiter les caractéristiques complémentaires des différents classifieurs.

III. 6 CONCLUSION

Dans ce chapitre, nous avons proposé une nouvelle approche de classification basée sur la RTF et un critère de décision statistique et neuronale pour améliorer le taux de classification correcte des défauts de la machine à induction. Nous avons opté de concevoir un classifieur directement dans le plan d'auto-ambiguïté utilisé comme espace d'extraction des caractéristiques pertinentes de diagnostic. Comme toutes les RTF peuvent être dérivées du plan d'auto-ambiguïté, aucune hypothèse à priori n'est faite concernant le lissage exigé pour la classification.

La représentation temps-fréquence employée dans cette étude est dite « dépendante de la classe du signal (DCS) » dont le plan d'ambiguïté est lissé par un noyau conçu spécialement pour de réaliser une séparation maximale entre la machine en défaut et saine. Le contraste de Fisher, basé sur la compacité et la séparabilité des classes, est utilisé pour la séparation des classes. A cet égard, nous avons conçu un noyau spécifique pour la séparation de chaque type de défaut (trois noyaux au total).

Pour la partie sélection, nous avons appliqué une approche sous-optimale pour exclure les points du noyau qui sont fortement corrélés avec les points du noyau de rang supérieur. Ces derniers ne contribuent pas à améliorer la classification. Ensuite, un nombre optimal de points de noyau a été sélectionné en minimisant un modèle probabiliste d'erreur de classification proposé pour l'approche de classification statistique. Ce dernier dépend de deux paramètres à savoir le type de noyau utilisé pour construire le vecteur de forme et une mesure de distance (distance Mahalanobis ou euclidienne).

Les résultats de classification correcte obtenus à partir du classifieur statistique optimisé avec la distance Mahalanobis comme critère de décision atteignent celles des classifieurs à base de RNA avec sigmoïde comme fonction de transfert. Les résultats obtenus à partir du classifieur à réseau de neurones montrent aussi que le RNA avec l'algorithme de Levenberg-Marquardt comme algorithme d'apprentissage a une nette supériorité sur le WNN.

Enfin, ces résultats vérifient que le système de classification proposé et testé avec des données expérimentales recueillies à partir de la mesure du courant statorique à différentes charges et à base de RNA, comme critère de décision, entraînés avec algorithme Levenberg-Marquardt, est capable de détecter et de diagnostiquer les défauts avec une précision acceptable indépendamment du niveau de charge et du type de défaut.

III. 7 REFERENCES

- [1] M. E. H. Benbouzid, "A Review of Induction Motors Signature Analysis as a Medium for Faults Detection," *IEEE Trans. Ind. Electron.*, vol. 47, no. 5, p. 984–993, Oct. 2000.
- [2] A. Belini, F. Filipetti, C. Tassoni, and G. A. Capolino, "Advances in Diagnostic Techniques for Induction Machines," *IEEE Trans. Ind. Electron.*, vol. 55, no. 12, p. 4109–4126, Dec. 2008.
- [3] A. Siddique, G. S. Yadava, and B. Singh, "A Review of Stator Fault Monitoring Techniques of Induction Motors," *IEEE Trans. Energy Convers.*, vol. 20, no. 1, p. 106–114, Mar. 2005.
- [4] S. H. Kia, H. Henao, and G. Capolino, "Some Digital Signal Processing Techniques for Induction Machines Diagnosis," in *Proc. IEEE SDEMPED*, p. 322–329, Sep. 2011.
- [5] A. K. Mahamad and T. Hiyama, "Improving Elman Network Using Genetic Algorithm for Bearing Failure Diagnosis of Induction Motor," in *Proc. IEEE SDEMPED*, p. 1–6, 2009.
- [6] A. Soualhi, G. Clerc, H. Razik, and O. Ondel, "Detection of Induction Motor Faults by an Improved Artificial Ant Clustering," in *Proc. IECON*, p. 3446–3451, Nov 2011.
- [7] B. Ayhan, M. Y. Chow, and M. H. Song, "Multiple Discriminant Analysis and Neural-Network-Based Monolith and Partition Fault-Detection Schemes for Broken Rotor Bar in Induction Motors," *IEEE Trans. Ind. Electron.*, vol. 53, no. 4, p. 1298–1308, Jun. 2006.
- [8] M. B. K. Bouzid, G. Champenois, N. M. Bellaaj, L. Signac, and K. Jelassi, "An Effective Neural Approach for the Automatic Location of Stator Interturn Faults in Induction Motor," *IEEE Trans. Ind. Electron.*, vol. 55, no. 12, p. 4277–4289, Dec. 2008.
- [9] H. Su and K. T. Chong, "Induction Machine Condition Monitoring Using Neural Network Modeling," *IEEE Trans. Ind. Electron.*, vol. 54, no. 1, p. 241–249, Feb. 2007.
- [10] J. F. Martins, V. F. Pires, and A. J. Pires, "Unsupervised Neural-Network Based Algorithm for an On-Line Diagnosis of Three-Phase Induction Motor Stator Fault," *IEEE Trans. Ind. Electron.*, vol. 54, no. 1, p. 259–264, Feb. 2007.
- [11] F. Zidani, D. Diallo, M. E. H. Benbouzid, and R. Nait-Said, "A Fuzzy Based Approach for the Diagnosis of Fault Modes in a Voltage-Fed PWM Inverter Induction Motor Drives," *IEEE Trans. Ind. Electron.*, vol. 55, no. 2, p. 586–593, Feb. 2008.
- [12] M. Sulowicz and T. J. Sobczyk, "Application of Fuzzy Inference System for Cage Induction Motors Rotor Eccentricity Diagnostic," in *Proc. IEEE SDEMPED*, p. 101–105, 2007.
- [13] M. S. Ballal, Z. J. Khan, H. M. Suryawanshi, and R. L. Sonolikar, "Adaptive Neural Fuzzy Inference System for the Detection of Inter-Turn Insulation and Bearing Wear Faults in Induction Motor," *IEEE Trans. Ind. Electron.*, vol. 54, no. 1, p. 250–258, Feb. 2007.
- [14] Z.-L. Gaing, "Wavelet-Based Neural Network for Power Disturbance Recognition and Classification," *IEEE Trans. Power Del.*, vol. 19, no. 4, p. 1560–1568, Oct. 2004.
- [15] T. W. S. Chow and S. Hai, "Induction Machine Fault Diagnostic Analysis with Wavelet Technique," *IEEE Trans. Ind. Electron.*, vol. 51, no. 3, p. 558–565, Jun. 2004.
- [16] J. Cusido, L. Romeral, J. A. Ortega, J. A. Rosero, and A. G. Espinosa, "Fault Detection in Induction Machines Using Power Spectral Density in Wavelet Decomposition," *IEEE Trans. Ind. Electron.*, vol. 55, no. 2, p. 633–643, Feb. 2008.

- [17] J. Pons-Llinares, J. A. Antonino-Daviu, M. Riera-Guasp, M. Pineda-Sanchez, and V. Climente-Alarcon, "Induction Motor Fault Diagnosis Based on Analytic Wavelet Transform Via Frequency B-splines," in *Proc. IEEE SDEMPED*, p. 1–7, 2009.
- [18] M. Delgado, A. Garcia, and J. A. Ortega, "Evaluation of Feature Calculation Methods for Electromechanical System Diagnosis," in *Proc. IEEE SDEMPED*, p. 495–502, 2011.
- [19] B. Yazici and G. B. Kliman, "An Adaptive Statistical Time–Frequency Method for Detection of Broken Bars and Bearing Faults in Motors Using Stator Current," *IEEE Trans. Ind. Appl.*, vol. 35, no. 2, p. 442–452, Mar./Apr. 1999.
- [20] M. Davy and C. Doncarli, "Optimal Kernels of Time–Frequency Representations for Signal Classification," in *Proc. IEEE-SP Int. Symp. Time-Freq. Time-Scale Anal.*, p. 581–584, 1998.
- [21] T. Boukra, A. Lebaroud, "Classification of Induction Machine Faults," *7th International Multi-Conference on Systems, Signals and Devices*, Amman, Jordan, 2010.
- [22] M. Davy, C. Doncarli, and G. F. Bourdreaux-Bartels, "Improved Optimization of Time–Frequency-Based Signal Classifiers," *IEEE Signal Process. Lett.*, vol. 8, no. 2, p. 52–57, Feb. 2001.
- [23] Tahar Boukra, Abdesselam Lebaroud, Guy Clerc, "Statistical and Neural-Network Approaches for the Classification of Induction Machine Faults Using the Ambiguity Plane Representation," *IEEE Transactions on Industrial Electronics*, vol. 60, no. 9, p.4034-4042, 2013.
- [24] S.-H. Cho, G. Jang, and S.-H. Kwon, "Time–Frequency Analysis of Power Quality Disturbances via the Gabor–Wigner Transform," *IEEE Trans. Power Del.*, vol. 25, no. 1, p. 494–499, Jan. 2010.
- [25] B. W. Gillespie and L. Atlas, "Optimizing Time–Frequency Kernels for Classification," *IEEE Trans. Signal Process.*, vol. 49, no. 3, p. 485–496, Mar. 2001.
- [26] M. Wang, G. I. Rowe, and A. V. Mamishev, "Classification of Power Quality Events Using Optimal Time–Frequency Representations—Part 2: Application," *IEEE Trans. Power Del.*, vol. 19, no. 3, p. 1496–1503, Jul. 2004.
- [27] A. Lebaroud and G. Clerc, "Classification of induction Machine Faults by Optimal Time Frequency Representations," *IEEE Trans. Ind. Electron.*, vol. 55, no. 12, p. 4290–4298, Dec. 2008.
- [28] T. Boukra, A. Lebaroud, G. Clerc, "Neural-Network Approaches for Classification of Induction Machine Faults Using Optimal Time-Frequency Representations," *10th International Conference on Environment and Electrical Engineering (EEEIC)*, 2011
- [29] S. Theodoridis and K. Koutroumbas, "Pattern Recognition," Amsterdam, The Netherlands: Elsevier, 2006.
- [30] Q. Zhang and A. Benveniste, "Wavelet Networks," *IEEE Trans. Neural Netw.*, vol. 3, no. 6, p. 889–898, Nov. 1992.
- [31] Q. Zhang, "Using Wavelet Network In Nonparametric Estimation," *IEEE Trans. Neural Netw.*, vol. 8, no. 2, p. 227–236, Mar. 1997.
- [32] T. Boukra, A. Lebaroud, A. Medoued, "Wavelet-Network for Classification of Induction Machine Faults Using Optimal Time-Frequency Representations," *7th International Conference on Electrical and Electronics Engineering (ELECO)*, p. I-358 - I-362, 2011.

- [33] C.-H. Lu, "Wavelet Fuzzy Neural Networks for Identification and Predictive Control of Dynamic Systems," *IEEE Trans. Ind. Electron.*, vol. 58, no. 7, p. 3046–3058, Jul. 2011.
- [34] Olivier Ondel, "Diagnostic par Reconnaissance des Formes : Application à un Ensemble Convertisseur – Machine Asynchrone," Thèse de doctorat, École Doctorale Électronique, Électrotechnique, Automatique de Lyon, Oct. 2006.

IV. 1 INTRODUCTION

L'intégration dans le système de supervision une fonction de pronostic permet d'anticiper l'apparition d'un phénomène qui vient de se manifester (défaillance) plutôt que de comprendre a posteriori ce phénomène (fonction diagnostic) afin de recourir en conséquence à des actions protectrices.

Généralement les méthodes de diagnostic utilisent une connaissance sur le comportement du système alors que les méthodes de pronostic reposent plutôt sur une connaissance de l'état de santé (état de dégradation) du système.

Dans ce chapitre, nous choisissons d'orienter notre travail vers le pronostic au niveau composant que nous jugeons critique: approche « locale » au lieu d'une approche « globale » niveau système. L'approche de pronostic guidée par les données est privilégiée dans notre cas à cause de la difficulté de disposer de connaissance sur les phénomènes de dégradation complexes.

Les tendances pronostic actuelles traitées dans les chapitres précédents concernant l'identification de nouvelles caractéristiques de pronostic nous ont conduits à poser le problème de pronostic de la manière suivante :

IV. 1. 1 POSITIONNEMENT DU PROBLEME

Dans une situation où des données de défaut jusqu'à la défaillance sont rares et le développement des modèles de dégradation s'avère difficile, le problème est de savoir s'il est possible de faire une bonne prédiction de la RUL en l'absence de caractéristiques de pronostic pouvant être identifiées à représenter la progression de défaut (figure IV. 8). Nous devons donc être adressés à l'identification de nouvelles caractéristiques de pronostic possédant les qualités mentionnées ci-dessus. Nous proposons pour cette tâche une technique d'extraction avancée en utilisant l'algorithme d'optimisation par essaim de particules (PSO) pour révéler la corrélation entre les caractéristiques de pronostic et la dégradation du système (composant) en utilisant l'ensemble des métriques proposées dans [6] et le modèle semi Markov (HSMM) caché comme algorithme de pronostic pour la prédiction de la durée de vie utile restante RUL.

IV. 1. 2 ORGANISATION DU CHAPITRE

Le reste de ce chapitre est organisé comme suit. La méthodologie proposée est donnée à la section IV. 2, suivie d'une explication détaillée de l'algorithme d'optimisation par essaim particulière dans section IV. 3. Après l'identification des caractéristiques de pronostic, la prédiction de la durée de vie utile résiduelle RUL en utilisant le modèle semi-Markov caché est présentée à la section IV. 4. Dans cette section, on présente au premier lieu la théorie du modèle de Markov caché et celle du modèle semi-Markov caché avant d'adapter ce dernier à la prédiction de la RUL à la section IV. 5. 2. La description du dispositif expérimental et les commentaires sur les résultats sont présentés dans les sections IV. 6 et IV. 7 respectivement. Enfin, une conclusion est tirée à la section IV. 9.

IV. 2 METHODOLOGIE PROPOSEE POUR LE PRONOSTIC

La figure IV. 1 représente les étapes d'un système typique de surveillance de l'état de fonctionnement.

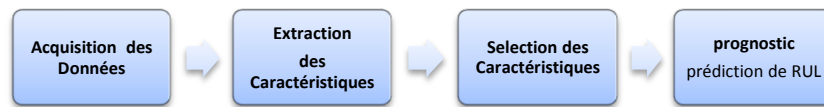


Figure IV. 1 : Système typique de surveillance

Le choix des roulements comme échantillon test peut être expliqué par le fait que ces éléments sont considérés comme des éléments mécaniques couramment utilisés dans l'industrie et sont présents dans la quasi-totalité des procédés industriels, en particulier dans ceux qui utilisent des éléments rotatifs et des machines. Les vibrations occupent une position privilégiée parmi les données des capteurs à considérer lors de la surveillance de l'état des roulements sans affecter leur fonctionnement.

L'étape d'extraction des caractéristiques de pronostic vise à cartographier les données acquises dans un espace de caractéristiques pertinentes de l'état de santé de l'équipement en utilisant diverses techniques de traitement du signal. Les signaux de vibration peuvent être traités dans le domaine temporel, le domaine fréquentiel ou le domaine temps-fréquence [10].

L'analyse temporelle est la méthode la plus simple pour la surveillance conditionnelle « CM » à partir des données de vibrations. Les caractéristiques statistiques les plus couramment utilisées pour le diagnostic ou le pronostic des roulements sont la valeur de crête (peak value) (Pv), valeur moyenne quadratique (root mean square value) (RMS), la valeur de Kurtosis - Kurtosis value (Kv), facteur de crête (crest factor) (Crf), facteur de dégagement (clearance factor) (Clf), facteur d'impulsion (impulse factor) (Imf) et le facteur de forme (shape factor) (Shf).

En domaine fréquentiel, la FFT (Fast Fourier Transform) décompose un signal de forme d'onde en son spectre de fréquences qui contient les composantes et leurs amplitudes. Les énergies (définies comme la somme des carrés des amplitudes) sur des bandes de fréquences centrées sur des fréquences spécifiques (par exemple fréquence de rotation et ses harmoniques et/ou fréquences des défauts roulement) sont prises également comme caractéristiques pertinentes.

D'avantage de méthodes de traitement du signal peuvent être utilisées pour extraire bien d'autres caractéristiques pertinentes dans cette étape [11].

Dans l'étape suivante, qui est l'étape de sélection des caractéristiques, la démarche optimisation par essaim de particules est proposée pour identifier les caractéristiques avancées.

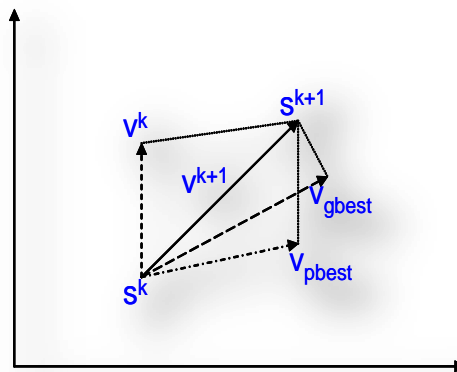
IV. 3 ALGORITHME D'OPTIMISATION PAR ESSAIM DE PARTICULES (PSO)

L'optimisation par essaim de particules (PSO) est une technique d'optimisation stochastique à base de population développée par le Dr Eberhart et le Dr Kennedy en 1995, inspiré par le comportement social des oiseaux en quête de nourriture dans une zone définie [12] - [16]. Le PSO adapte ce comportement et cherche la meilleure solution, vecteur, dans l'espace de recherche. Une solution unique est appelé particule. Chaque particule possède une valeur de fitness/coût qui est évaluée par la fonction à minimiser, et chaque particule possède aussi une

vitesse qui dirige le "vole" des particules. Les particules explorent l'espace de recherche en suivant les particules optimales. L'algorithme est initialisé avec des particules à des positions aléatoires, par suite, il explore l'espace de recherche pour trouver de meilleures solutions. Dans chaque itération, chaque particule ajuste sa vitesse pour suivre deux meilleures solutions. La première est la partie cognitive, où la particule suit sa propre meilleure solution trouvée à ce moment, celle qui produit le coût le plus bas (le fitness maximal). Cette valeur est appelée $pBest$ (meilleure particule). L'autre meilleure valeur est la meilleure solution actuelle de l'essaim, à savoir, la meilleure solution pour toutes particules de l'essaim. Cette valeur est appelée $gBest$ (meilleur global). Ensuite, chaque particule ajuste sa vitesse et sa position selon les équations suivantes:

$$\begin{aligned} v^{k+1} &= v^k + c_1 \cdot r_1 \cdot (pBest - s^k) + c_2 \cdot r_2 \cdot (gBest - s^k) \\ s^{k+1} &= s^k + v^{k+1} \end{aligned} \quad (IV. 1)$$

v^k est la vitesse actuelle, v^{k+1} est la nouvelle vitesse, s^k la position actuelle, s^{k+1} la nouvelle position, $pBest$ et $gBest$ comme indiqués ci-dessus, r_1 et r_2 sont des nombres aléatoirement distribués dans l'intervalle $[0, 1]$, et c_1 et c_2 sont des coefficients d'accélération. Où c_1 est le facteur qui influence le comportement cognitif, à savoir, de combien la particule suivra sa propre meilleure solution et c_2 est le facteur de comportement social, à savoir, de combien la particule suivra la meilleure solution de l'essaim.



v_{pbest} : vitesse basée sur $pBest$.

v_{gbest} : vitesse basée sur $gBest$.

L'algorithme peut être décrit comme suit:

1. initialiser chaque particule avec une vitesse et une position aléatoires. (la position dans notre cas sont les poids pour l'ensemble des caractéristiques de pronostic élues)
2. Calculer le coût pour chaque particule. Si le coût actuel est inférieur à la meilleure valeur à ce moment, appelez cette position $pBest$.
3. Choisissez la particule avec le coût le plus bas (fitness maximal) parmi toutes les particules. La position de cette particule est $gBest$.
4. On calcule, pour chaque particule, la nouvelle vitesse et position selon les équations (IV. 1).
5. Répétez les étapes 2-4 tant que l'itération maximale ou les critères d'erreur minimale ne soient pas atteints.

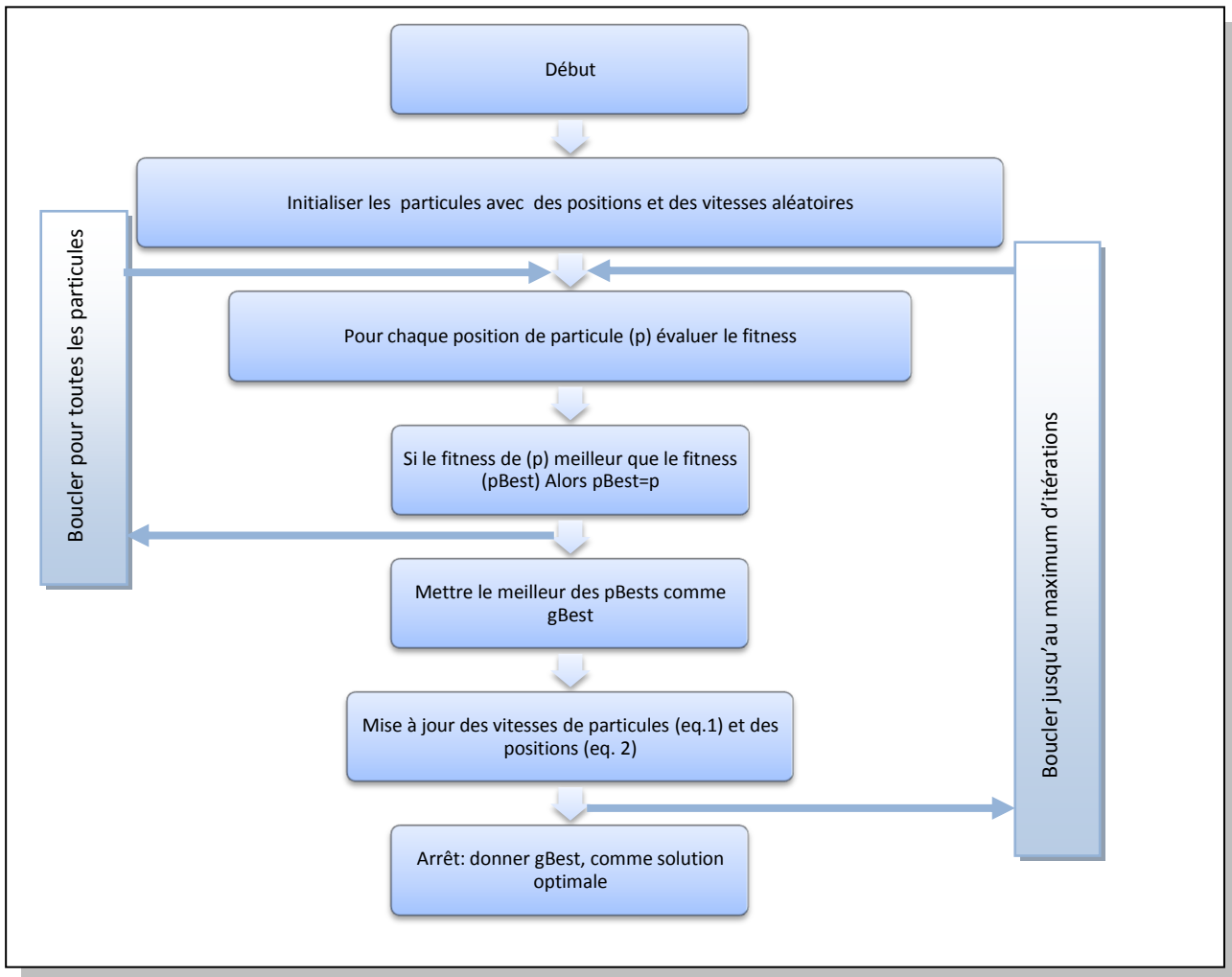


Figure IV. 2 : Organigramme de la méthode PSO.

IV. 3. 1 CONCEPTION DE FONCTION FITNESS

La fonction fitness est la somme pondérée des trois mesures (méthode de la somme pondérée des fonctions objectives (SWO)): [6]

$$fitness = w_m \times monotonicity + w_p \times prognosability + w_t \times trendability \quad (1)$$

La Monotonicité d'une population de caractéristiques est donnée par la moyenne de la différence des fractions des dérivées positives et négatives pour chaque trajet.

$$Monotonicity = mean \left(\left| \frac{\#pos \frac{d}{dx}}{n-1} - \frac{\#neg \frac{d}{dx}}{n-1} \right| \right) \quad (IV. 2)$$

« Pronosticabilité » est calculée comme étant l'écart entre les valeurs finales de défaillance de chaque trajet, divisé par l'intervalle moyen du trajet. Le tout est pondéré exponentiellement pour donner le zéro désiré à l'échelle 1:

$$Prognosability = \exp \left(- \frac{std(failurevalues)}{mean(|failurevalue - startingvalue|)} \right) \quad (IV. 3)$$

Les paramètres de pronostic sont ré-échantillonnés par rapport à la fraction de durée de vie totale ce qui en résulte sur 100 observations pour chaque caractéristique pronostic.

« Mesure de tendance » est donnée par la plus petite corrélation absolue entre caractéristiques pronostic :

$$Trendability = \min(|corrcoef_{ij}|) \quad (IV.4)$$

Il est important de lisser de manière adéquate les données pour éliminer l'effet de bruit et donner des estimations précises pour le calcul des dérivées.

IV. 4 UTILISATION DU HMM ET HSMM POUR PREDICTION DE LA DUREE DE VIE UTILE RESTANTE

IV. 4. 1 BACKGROUND THEORIQUE

Les mécanismes de défaillance des systèmes mécaniques impliquent généralement plusieurs états de santé dégradés. Par exemple, un diagramme illustrant les transitions des états de santé du roulement de l'état normal à la défaillance est illustré sur la figure IV. 3. Ce processus peut être idéalement décrit par un modèle mathématique appelé modèle de Markov caché (Hidden Markov Model - HMM) puisqu'il peut être utilisé pour estimer les états de santé inobservés en utilisant des signaux observables issus des capteurs. En d'autres termes, les états HMM se manifestent par un comportement probabiliste. Les flèches représentent les chemins de transition d'un état à un autre. Il y a habituellement un numéro indiquant la probabilité de transition qui lui est associée. On peut prétendre qu'il peut y avoir plusieurs mécanismes de rupture à une défaillance du roulement. Cela peut être vrai. Cependant, il faut également noter que les signaux issus de chaque étape du processus de dégradation peuvent avoir des caractéristiques similaires. Par exemple, un mécanicien expérimenté peut distinguer une bonne transmission d'une mauvaise en écoutant les émissions acoustiques de cette transmission. Le HMM peut saisir exactement les caractéristiques de chaque étape du processus de défaillance, qui est la base de l'utilisation du HMM pour le diagnostic et le pronostic des défauts [17].

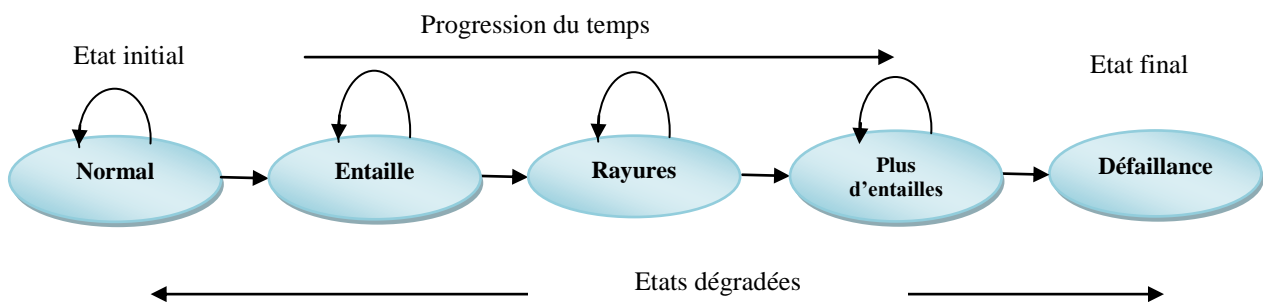


Figure IV. 3 : Transitions des états du roulement de l'état normal à la défaillance

IV. 4. 2 ELEMENTS D'UN HMM

Une chaîne de Markov est une séquence de variables aléatoires S_1, S_2, S_3, \dots avec la propriété de Markov suivante:

« étant donné l'état actuel, les états futurs et passés sont indépendants » [18],

$$P(S_{n+1} = s | S_1 = s_1, S_2 = s_2, \dots, S_n = s_n) = P(S_{n+1} = s | S_n = s_n)$$

La probabilité d'une séquence $s = s_1, s_2, \dots, s_n$ peut être calculée par

$$P(s_1, s_2, \dots, s_n) = P(s_1)P(s_2|s_1) \dots P(s_{n-1}|s_{n-2})P(s_n|s_{n-1})$$

Dans le modèle de Markov, les états sont directement visibles à l'observateur. Les modèles de Markov cachés étendent les chaînes de Markov pour inclure le cas où l'observation est une fonction probabiliste de l'état. Maintenant, les états ne sont pas directement visibles (ils sont cachés), mais les symboles d'observation de sortie sont visibles. Précisément, « le HMM est un processus stochastique double dont un processus stochastique sous-jacent qui n'est pas observable, mais peut être observé à travers un autre processus stochastique qui produit la séquence d'observations » [22].

Soit N , le nombre d'états du modèle. Nous notons les états individuels par $S = \{S_1, S_2, \dots, S_N\}$ et l'état au temps t par q_t .

M , le nombre de symboles d'observation distincts par état. Nous notons les symboles individuels par $V = \{v_1, v_2, \dots, v_M\}$.

La distribution des probabilités de transition d'état, $A = \{a_{ij}\}$, où

$$a_{ij} = P(q_{t+1} = S_j | q_t = S_i), 1 \leq i, j \leq N$$

La distribution de probabilité de symbole d'observation (probabilité d'émission) à l'état j , $B = \{b_j(k)\}$, où

$$b_j(k) = P(v_k \text{ à } t | q_t = S_j), i \leq j \leq N, 1 \leq k \leq M$$

La distribution d'état initial $\pi = \{\pi_i\}$ où

$$\pi_i = P(q_1 = S_i) \quad 1 \leq i \leq N$$

Etant donné les valeurs des termes N, M, A, B et π ci-dessus, le modèle de Markov caché peut générer une séquence d'observation $O = O_1 O_2 \dots O_T$ (où O_t désigne le symbole observé au temps t et T est la quantité des observations de la séquence.) par la procédure suivante :

1. Choisissez un état initial $q_1 = S_i$ selon la probabilité d'état initial π_i et mettre $t = 1$.
2. Choisissez $O_t = v_k$ selon la probabilité de symbole observation $b_i(k)$.
3. Transiter vers l'état $q_{t+1} = S_j$ selon la probabilité de transition d'état a_{ij} .
4. Mettre $t = t + 1$, si $t < T$, passez à l'étape 2, sinon arrêter la procédure.

Un HMM peut traiter trois problèmes séparément [22] :

Problème 1 (problème d'évaluation): Compte tenu de la séquence d'observation $O = O_1 O_2 \dots O_T$ et du modèle $\lambda = (A, B, \pi)$, comment calculer efficacement la probabilité de la séquence d'observation, c'est-à-dire $P(O|\lambda)$?

Problème 2 (problème de décodage): Compte tenu de la séquence d'observation $O = O_1 O_2 \dots O_T$ et du modèle $\lambda = (A, B, \pi)$, comment choisir une séquence d'état correspondante $Q = q_1 q_2 \dots q_T$ « expliquant » d'une façon optimale ces observations ?

Problème 3 (problème d'apprentissage): Comment estimer les paramètres du modèle $\lambda = (A, B, \pi)$ qui maximiseraient $P(O|\lambda)$?

Nous présentons des solutions mathématiques à chacun des trois problèmes de base du HMM en annexe A. Dans la Section A, nous introduisons les algorithmes avant - arrière du HMM standard, pour le problème 1. L'algorithme de Baum-Welch du HMM standard est présenté à

la Section B, pour résoudre le problème 3. Ensuite, l'algorithme de Viterbi est introduit dans la section C, pour résoudre le problème 2.

IV. 4. 3 ADAPTATION DU HMM POUR LA PREDICTION DE L'ETAT FUTURE.

Principe du pronostic basé sur l'HMM présenté dans [20]:

Supposons que M HMMs sont entraînés avec succès pour reconnaître M états distincts d'un mécanisme de dégradation. La présentation des séquences d'observations ordonnées temporellement d'un tel mécanisme devrait résulter sur des trajectoires de vraisemblances illustrées dans la figure IV. 4. Si le HMM $_j$ résulte en la plus grande vraisemblance pour une séquence d'observations O_i donnée, alors, on peut déclarer que le processus est dans l'état j .

Quand l'ensemble des HMMs entraîné (HMM $_1$, HMM $_2$, ..., HMM $_M$) est présenté avec une séquence d'observations ordonnée temporellement à partir d'une unité (figure IV. 4), il devrait résulter en un vecteur de $M - 1$ temps de transition d'état notés $S = [t_{S_1 \rightarrow S_2}, t_{S_2 \rightarrow S_3}, \dots, t_{S_{M-1} \rightarrow S_M}]$ (ce sont les coordonnées des intersections des profils des vraisemblances des différents HMMs le long de l'axe d'usage/vie). Supposons que les séquences d'observation sont disponibles à partir de R unités similaires dans le but de développer des modèles de prédiction d'état. Cela se traduira par R vecteurs estimés de points de transition d'état, $T = [S_1, S_2, \dots, S_R]$. L'hypothèse fondamentale est que T contient toutes les informations nécessaires pour fournir la prédiction de l'état futur.

La procédure est la suivante. Supposant que T suit une distribution multivariée. Une fois cette distribution est évaluée, la distribution de probabilité conditionnelle pour un état de transition donné, $t_{S_i \rightarrow S_{i+1}}$, étant donné les points de transition des états précédents ($S_1 \rightarrow S_2, \dots, S_{i-1} \rightarrow S_i$) d'une unité individuelle quelconque sous investigation, peut être estimée. Le processus peut être itéré de manière récursive pour faire des prédictions concernant plusieurs temps de transitions d'état séquentiels [20].

Après la détermination de la distribution des points de transition supposée généralement Gaussienne dont on pourrait vérifier cette supposition en appliquant le test de normalité chi-square, il est possible de déterminer les distributions conjointes de tous les temps de transitions d'états $[t_{S_1 \rightarrow S_2}, t_{S_2 \rightarrow S_3}, \dots, t_{S_{M-1} \rightarrow S_M}]$ qui sont nécessaires pour l'estimation des paramètres des distributions conditionnelles, $f(t_{S_{M-1} \rightarrow S_M} | t_{S_1 \rightarrow S_2}, t_{S_2 \rightarrow S_3}, \dots)$. Ces dernières peuvent être utilisées pour la prédiction de l'état futur.

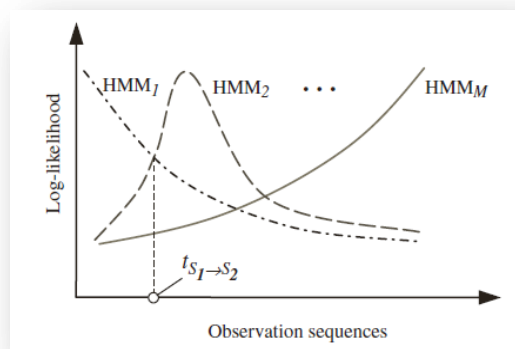


Figure IV. 4 : Les vraisemblances de chaque état du composant en fonction des séquences d'observations

L'inconvénient du HMM standard - limitant celui-ci à utiliser seulement la distribution géométrique pour modéliser sa distribution de durée - pourrait lourdement entraver son application dans de nombreux domaines. Le HMM avec durée (HMMD), présenté dans la section suivante, surmonte cette lacune en intégrant dans son modèle la distribution de la durée explicitement. Ainsi, le HMMD peut modéliser directement toute distribution. Par conséquent, l'utilisation du HMMD au lieu du HMM devrait offrir un avantage significatif dans de nombreuses applications,.

IV. 5 HSMM POUR PREDICTION DE LA DUREE DE VIE UTILE RESTANTE

Dans le HMM standard, lorsqu'un état i est accédé, celui-ci est occupé pendant une période de temps, via des self-transitions, jusqu'au transit vers un autre état j . Si l'intervalle de temps d de l'état est donné, l'expression de distribution de probabilité de la durée du HMM standard est implicitement donnée par :

$$p_i(d) = \text{probabilité de } d \text{ observations consécutives dans l'état } S_i.$$

$$= a_{ii}^{d-1} (1 - a_{ii}) \quad (\text{IV. 5})$$

Où a_{ii} est la probabilité du self-transition de l'état i . Cette distribution géométrique ne convient pas dans bien des cas. Il faudrait chercher un autre paradigme - HMM avec durée (HMMD) - qui pourrait incorporer directement la version de $p_i(d)$ qui modélise la distribution de la durée réelle de l'état i dans le modèle de Markov caché. Si la connaissance explicite de la durée des états HMM est codée, alors le HMMD généralisé peut être illustré dans la figure IV. 5.

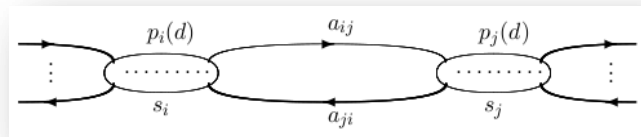


Figure IV. 5 : HMM avec durée (HMMD)

Lors de l'accès à l'état s_i , cet état va avoir une durée d selon sa densité de probabilité $p_i(d)$, il transite alors vers un autre état s_j selon la probabilité de transition d'état a_{ij} (les self-transitions, a_{ii} , ne sont pas autorisées dans ce formalisme). Bien que la transition entre les différents états a_{ij} reste inchangée - toujours obéit à la propriété de Markov (premier ordre), les self-transitions a_{ii} sont remplacées maintenant par les distributions $p_i(d)$ explicites. Étant donné que la propriété de Markov soit violée, le nouveau modèle est souvent appelé comme « Hidden Semi-Markov Model (HSMM) ». Il est facile de voir que le HMMD va se transformer en un HMM si $p_i(d)$ est remplacée par une distribution géométrique présentée dans équation (IV. 5).

Une méthode plus intuitive consiste à diviser chaque état en sous-états basés sur la valeur de la durée maximale autorisée D pour cet état. Plus grand soit D , plus de sous-états devront être générés et vice versa. La conséquence immédiate est un D^2 de coût de calcul supplémentaire.

Une idée commune consiste à remplacer la fonction de densité de probabilité de durée avec certaines fonctions de probabilité choisies proches des distributions de durée des applications réelles. Parmi elles, la distribution gaussienne ou la distribution normale qui est une distribution de probabilité continue décrivant les données d'un groupe autour de la moyenne. La distribution gaussienne est peut-être la distribution la plus couramment utilisée par les variables aléatoires comme générateur de données dans de nombreux domaines scientifiques. C'est le choix opté, dans notre cas, pour l'émission d'observations et pour la distribution de la durée.

Il serait avantageux d'utiliser les HSMMs avec des densités d'observations continues. Dans cette recherche, une mixture de distribution gaussienne est utilisée. La représentation générale de la fonction de densité de probabilité PDF d'une mixture finie est donnée sous la forme :

$$b_j(O) = \sum_{m=1}^{M_j} c_{jm} \cdot \eta[O, \mu_{jm}, \Sigma_{jm}] = \sum_{m=1}^{M_j} P(M_m = m | s_t = j) \cdot \eta[O, \mu_{jm}, \Sigma_{jm}], \quad 1 \leq j \leq N, \quad (IV.6)$$

où η représente une distribution gaussienne avec vecteur moyenne μ_{jm} et une matrice covariance Σ_{jm} pour le $m^{\text{ème}}$ composant de mixture dans l'état j , O est le vecteur d'observations en cours de modélisation, M_j est le nombre de composant gaussien η à l'état j , $c_{jm} = P(M_m = m | s_t = j)$ est le poids conditionnels pour le $m^{\text{ème}}$ composant de mixture dans l'état j . Les gains de mixture c_{jm} satisfont la contrainte stochastique suivante :

$$\sum_{m=1}^{M_j} c_{jm} = 1, \quad 1 \leq j \leq N,$$

$$c_{jm} \geq 0, \quad 1 \leq j \leq N, \quad 1 \leq m \leq M_j,$$

pour que la PDF soit correctement normalisée, c'est-à-dire,

$$\int_{-\infty}^{\infty} b_j(x) dx = 1, \quad 1 \leq j \leq N,$$

IV. 5. 1 ÉLEMENTS D'UN HSMM

Dans la notation introduite par Rabiner [22]. Voici les éléments du HSMM (λ)

N , nombre d'états,

M , nombre d'observations distincts,

D , longueur maximale de la durée,

a_{ij} , probabilité de transition d'état,

$b_j(k)$, probabilité d'émission : probabilité d'observer des v_k dans l'état i ,

π_i , probabilité de l'état initial.

$p_i(d)$, densité de probabilité de durée de l'état i : la probabilité d'obtenir exactement d observations consécutives à partir de l'état i après l'accès à cet état.

Étant donné le modèle ci-dessus (λ), une séquence d'observation $O = O_1 O_2 \dots O_T$ peut être générée comme suit

1. Choisissez un état initial $q_1 = S_i$ selon la distribution d'état initial π_i et mettre $t = 1$.
2. une durée d est choisie en fonction de la densité de probabilité de durée $p_i(d)$.

3. Choisissez $O_t O_{t+1} \dots O_{t+d-1}$ selon la distribution de probabilité conjointe d'émission à l'état S_i , c'est-à-dire, de $b_i(O_t) b_i(O_{t+1}) \dots b_i(O_{t+d-1})$. Nous supposons que chaque séquence d'observations est indépendante de toute autre séquence d'observations. Dans cette recherche, une mixture de distribution gaussienne est utilisée sous la forme décrite à l'équation (IV. 6).

4. transiter vers un nouveau état $q_{t+d} = S_j$ selon la distribution de probabilité de transition à l'état S_i , c'est-à-dire a_{ij} .

La probabilité de transition serait seulement utilisée au moment où la durée du séjour dans un état est terminée et l'état, donc, va changer. La matrice de transition, A , ne possède, alors, qu'un seul élément différent de zéro, ce qui indique le passage obligatoire à l'état suivant. La matrice de transition de l'HSMM dans ce cas peut être écrite comme suit [23]-[24].

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & 0 & \dots & 0 & 0 \\ \vdots & & \ddots & & & \vdots & \\ 0 & 0 & 0 & 0 & \dots & 0 & 1 \\ 0 & 0 & 0 & 0 & \dots & 0 & 1 \end{bmatrix} \quad (IV.7)$$

où son élément à la $i^{\text{ème}}$ ligne et la $j^{\text{ème}}$ colonne, a_{ij} , est la probabilité de transition de l' $i^{\text{ème}}$ état au $j^{\text{ème}}$ état.

5. mettre $t = t + d$, si $t < T$ Retournez à l'étape 2, sinon mettre fin à la procédure.

Les formules de ré-estimation du maximum de vraisemblance pour le HMMD (algorithme de Baum-Welch) sont présentées dans l'annexe B.

L'ensemble des paramètres définis pour la mise en œuvre de notre HSMM seraient les suivants,

$$\lambda = \underbrace{\{\pi_0, \mu_1, \dots, \mu_M, \Sigma_1, \dots, \Sigma_M\}}_{\text{Paramètres HMM}} \underbrace{\{\mu d_1, \dots, \mu d_M, \sigma d_1, \dots, \sigma d_M, d_{max}\}}_{\text{Paramètres de distribution de durée}}$$

où π_0 est la probabilité antérieure de l'état de santé initial, μ_i et Σ_i sont respectivement la moyenne et la matrice de covariance utilisées dans les distributions gaussiennes pour calculer la probabilité d'émission au temps t compte tenu du fait que $S_t = i$, μd_i et σd_i sont les paramètres de la distribution gaussienne représentant la variable de durée à l' $i^{\text{ème}}$ état notée d_i .

Pour la mise en œuvre, le terme (d_{max}) doit nécessairement être défini comme durée maximale possible des états.

La phase d'entraînement vise à ajuster les paramètres du modèle semi-Markov cachés afin de maximiser $P(O|\lambda)$ à l'aide d'un algorithme Expectation Maximisation EM.

Il y a eu beaucoup d'efforts pour améliorer le formalisme et l'efficacité de calcul du HSMM. Les travaux les plus notables incluant également les implémentations des algorithmes efficaces avant-arrière des modèles semi-Markov cachés [25], [26]. Des Formules de ré-estimation des paramètres HSMM se trouvent aussi dans [21].

IV. 5. 2 ADAPTATION DU HSMM AU PRONOSTIC

En utilisant les formules de ré-estimation, la durée d'état $D(S_n)$ après la phase d'apprentissage est donnée par :

$$D(S_n) = \mu(S_n) + \rho\sigma^2(S_n) \quad (IV. 8)$$

$$\rho = (T - \sum_{n=1}^M \mu(S_n)) / \sum_{n=1}^M \sigma^2(S_n) \quad (IV. 9)$$

La durée de vie utile RUL est donnée par [28]:

$$RUL = \hat{t}_c + \sum_{current\ state + 1}^M D(S_n) \quad (IV. 10)$$

\hat{t}_c est la durée estimée au point-changement d'état suivant (temps à la prochaine transition) [28] donnée par :

$$\hat{t}_c = \sum_{t=1}^T (t \times [p(s_{t-1} = i|O) - p(s_t = i|O)]), \quad (IV. 11)$$

où $p(s_t = i|O_{1,\dots,T})$ est la probabilité postérieure qu'un point de données appartient à l'état i , compte tenu de toutes les observations.

$p(s_{t-1} = i|O)$ et $p(s_t = i|O)$ peuvent être calculées efficacement par l'algorithme avant-arrière.

IV. 6 MISE EN ŒUVRE EXPÉRIMENTALE

Le choix des roulements peut être expliqué par le fait que ces éléments sont considérés comme des éléments mécaniques couramment utilisés dans l'industrie et sont présents dans la quasi-totalité des procédés industriels, en particulier dans ceux qui utilisent des éléments rotatifs et des machines. En outre, les roulements sont les principaux composants qui échouent le plus souvent dans les machines tournantes.

Afin de tenir compte de la dégradation naturelle des roulements, des tests de vieillissement sur des roulements conduisant à la rupture ont été effectués sur un banc d'essai conçu avec le soutien de Rexnord Corp.

Les données expérimentales ont été téléchargés à partir du Centre de pronostics d'excellence [[http://ti.arc.reférentiel.nasa.gov/tech/dash/pcoe/prognosticdata- /](http://ti.arc.reférentiel.nasa.gov/tech/dash/pcoe/prognosticdata/)].

Le banc d'essai est composé de quatre roulements Rexnord ZA-2115 à deux rangées installés sur un arbre comme le montre la figure IV. 6.

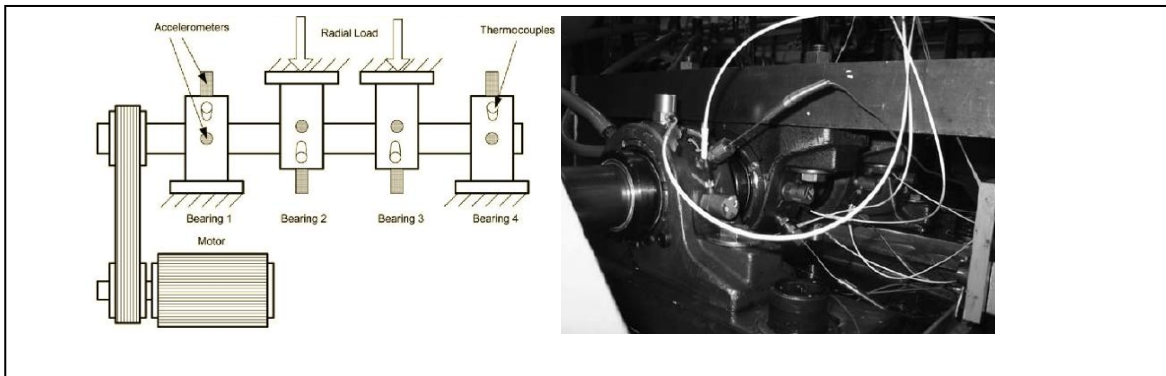


Figure IV. 6 : Banc d'essai de roulements.

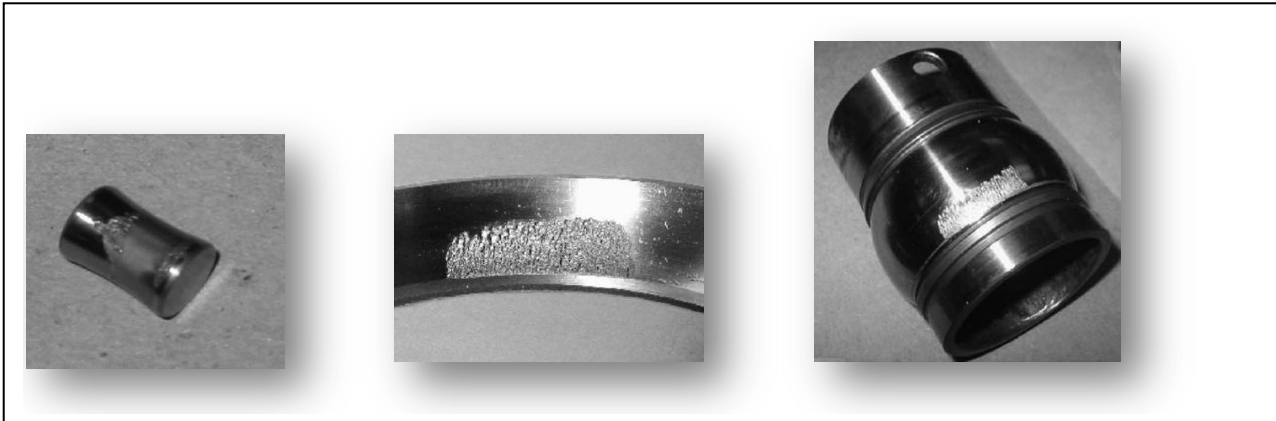


Figure IV. 7 : Défauts des roulements

(a) défaut dans élément roulant roulement 4 (b) défaut bague extérieure roulement 4. (c) bague intérieure roulement 3.

La vitesse de rotation est maintenue constante à 2000 tours par minute avec une charge radiale de 6000 lb placée sur l'arbre et les roulements par un mécanisme à ressort. Un bouchon magnétique installé dans le tuyau de retour d'huile recueilli les débris de l'huile comme une preuve de roulements en dégradations. Un accéléromètre à haute sensibilité PCB 353B33 en quartz ICP a été installé sur l'axe horizontal X et l'axe vertical Y pour chaque roulement. Les données de vibration ont été enregistrées toutes les 10 minutes avec un taux d'échantillonnage fixe de 20 kHz.

Les caractéristiques de ces roulements et les quatre fréquences de défaut à savoir la fréquence fondamentale de train (FTF), la fréquence de bille passant sur la bague intérieure en défaut (BPFI), la fréquence de bille passant sur la bague extérieure en défaut (BPFO), la fréquence de rotation de bille en défaut (BSF) sont détaillés dans le tableau I.

Tableau IV. 1: Caractéristiques des roulements étudiés

Diamètre de l'Elément Roulant (in)	0.331
Nombre de Rouleaux/Rangée	16
Diamètre Primitif (in)	2.82
Angle de Contact ϕ ($^{\circ}$)	15.17
$FTF = \frac{f_r}{2} \left(1 - \frac{B_d}{P_d} \cos\phi\right)$	14.77522
$BPFI = \frac{N_b}{2} f_r \left(1 + \frac{B_d}{P_d} \cos\phi\right)$	296.9299
$BPFO = \frac{N_b}{2} f_r \left(1 - \frac{B_d}{P_d} \cos\phi\right)$	236.4035
$BSF = \frac{P_d}{2B_d} f_r \left(1 - \left(\frac{B_d}{P_d}\right)^2 (\cos\phi)^2\right)$	139.9167

où f_r est la fréquence de rotation, N_b nombre de billes, B_d le diamètre des billes (rouleau), P_d diamètre primitif, ϕ l'angle de contact.

IV. 7 EXTRACTION DES CARACTERISTIQUES ET EXPLORATION

Il existe 2156 fichiers. Chaque fichier est nommé avec son horodatage, et contient une matrice 8×20480 d'une durée d'une seconde.

Les caractéristiques statistiques qui peuvent caractériser la forme des données, peuvent nous informer à quel point les données correspondent à une distribution normale [31].

Le RMS de la figure IV. 8 est représentatif de l'énergie de vibration, le facteur de dissymétrie dans la figure IV. 9, qui est une sorte de mesure de l'asymétrie, indique si la majorité des points des données sont supérieures ou inférieures à la valeur moyenne.

Le coefficient d'aplatissement kurtosis, figure IV. 10, indique à quelle point est pointue ou aplatis une courbe correspondante aux données. Sa valeur diminue lorsque le défaut atteint un état avancé de dégradation. Par conséquent, il est plus efficace pour détecter les stades précoces de défaillances des roulements. Les quantiles (ou parfois appelés les fractiles) qui peuvent aussi nous informer à quel point les données correspondent à une distribution normale, sont présentés dans la figure IV. 11. Les données sont réparties en quatre sous groupes égaux. Les valeurs limites de ces quatre ensembles constituent cinq caractéristiques. Les valeurs maximales et minimales de l'ensemble de données sont toujours la première et la dernière caractéristique, tandis que la médiane de l'ensemble de données est la troisième caractéristique. La deuxième et la quatrième peuvent être considérées comme les médianes de la moitié inférieure; (des données entre le minimum et le point médian), et la partie supérieure; (des données entre la médiane et le maximum) de l'ensemble de données.

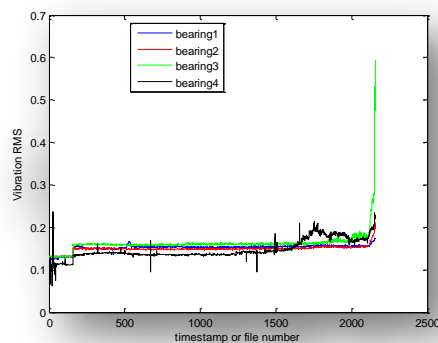


Figure IV. 8 : Caractéristiques RMS de vibration des roulements (1, 2, 3 et 4)

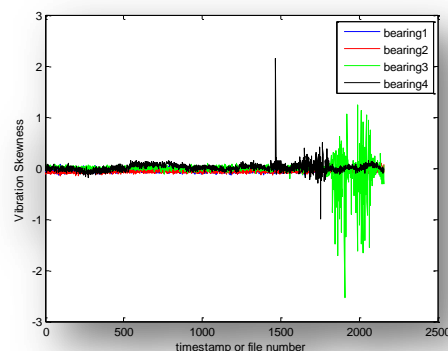


Figure IV. 9 : Caractéristiques de dissymétrie de vibration des roulements (1, 2, 3 et 4)

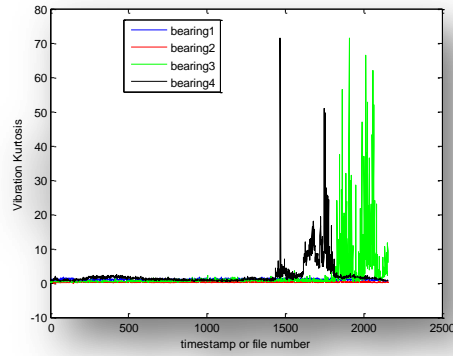


Figure IV. 10 : Caractéristiques Kurtosis de vibration des roulements (1, 2, 3 et 4)

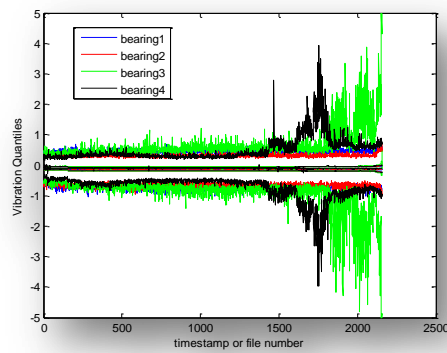


Figure IV. 11 : Caractéristiques Quantiles de vibration des roulements (1, 2, 3 et 4)

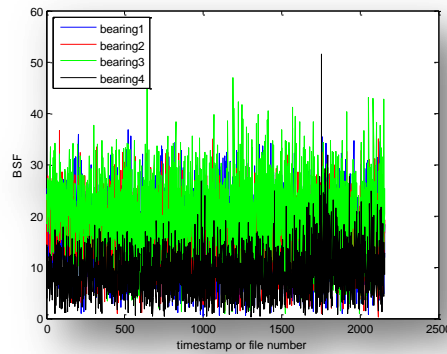


Figure IV. 12 : Amplitude BSF (Ball Spin Frequency) des roulements (1, 2, 3 et 4)

La courbe du BPFi en fonction du numéro de fichier pour le roulement 3, figure IV. 13, montre une augmentation du BPFi, qui confirme bien notre connaissance sur la défaillance du roulement 3 ; défaut de bague intérieure. La figure IV. 13 montre également une tendance à la hausse vers la fin de l'expérience due à l'augmentation des vibrations globale par l'accumulation de débris.

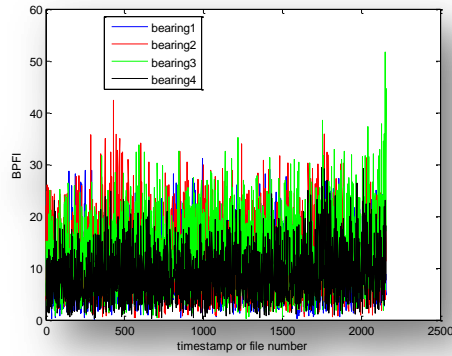


Figure IV. 13 : Amplitude BPFI (Ball Passing Frequency Inner-race) des roulements (1, 2, 3 et 4)

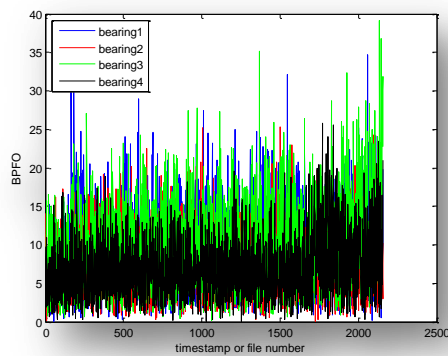


Figure IV. 14 : Amplitude BPFO (Ball Passing Frequency Outer-race) des roulements (1, 2, 3 et 4)

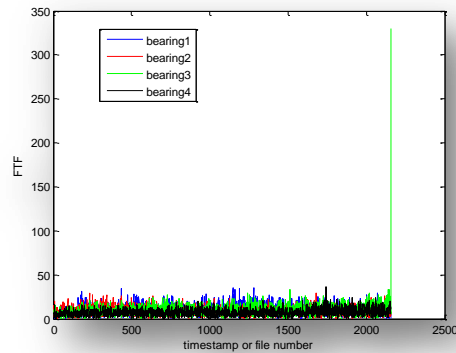


Figure IV. 15 : Amplitude FTF (Foundation Train Frequency) des roulements (1, 2, 3 et 4)

Il ya quelques pics importants dans la puissance BSF figure IV. 12, autour de l'échantillon 1725, et diminuent à nouveau avant la fin de l'expérience. Il a été diagnostiqué d'une défaillance de l'élément roulant.

Le roulement 1 présente certains niveaux élevés au début de l'expérience, mais elle atteint un état stable par la suite.

Les roulements 3 et 4 montrent clairement un intervalle de vibration accrue autour des temps soupçonnés de leur défaillance. Cela confirme notre conviction que le roulement 4, en fait, a échoué autour de l'échantillon 1725.

Le roulement 2 augmente également dans l'intervalle juste avant la fin de l'expérience, mais avec un niveau beaucoup plus faible que pour les roulements 3 et 4. Ceci est tout à fait susceptible d'être le début d'un défaut.

Le spectre de fréquence a été divisé en quatre bandes à savoir la bande basse fréquence (LF) avec moins de 1,25 kHz, la bande des fréquences moyennes (MF) entre 1,25 et 2,6 kHz, la bande des hautes fréquences (HF) entre 2,6 kHz et 6 kHz et finalement la bande des très hautes fréquences (VHF) entre 6 kHz et 40 kHz. Les puissances de ces bandes sont calculées comme caractéristiques.

La totalité des fréquences clés principales sont localisées dans de bande des basses fréquences figure IV. 20, celles-ci prennent en charge les points de défaillance présumés des roulements 3 et 4. La bande des hautes fréquences, figure IV. 18, est plus intéressante, car elle montre des changements progressifs dans les niveaux de puissance tout le long du cycle de vie du roulement 1, tandis que les autres courbes se correspondent étroitement avec leurs homologues de basses fréquences.

La bande des fréquences moyennes, figure IV. 19, est similaire à la bande hautes fréquences, mais la bande de très hautes fréquences figure IV. 16 est plus intéressante. Les roulements 3 et 4 ont atteint un stade avancé de défaillance, alors que des roulements 1 et 2 ne le sont pas encore. Ils peuvent y avoir des problèmes de brassage.

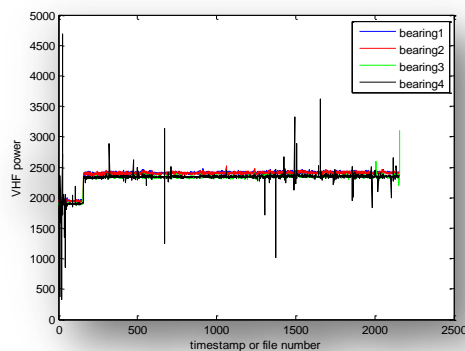


Figure IV. 16 : Amplitude de puissance VHF des roulements (1, 2, 3 et 4)

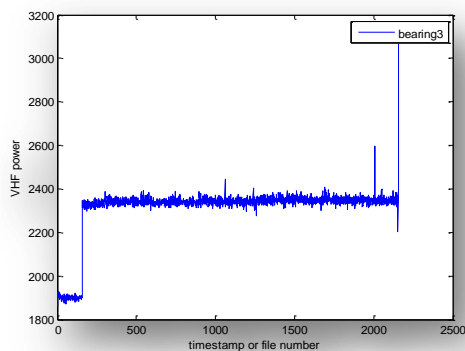


Figure IV. 17 : Amplitude de puissance VHF du roulement (3)

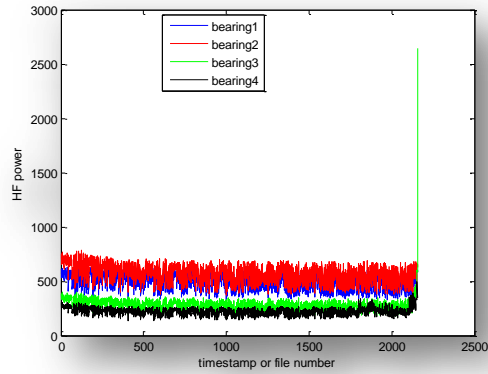


Figure IV. 18 : Amplitude de puissance HF des roulements (1, 2, 3 et 4)

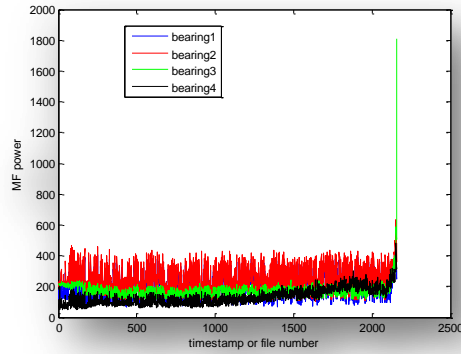


Figure IV. 19 : Amplitude de puissance MF des roulements (1, 2, 3 et 4)

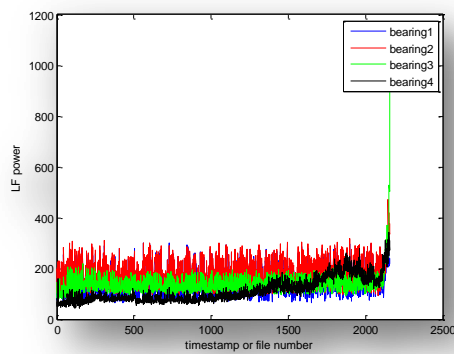


Figure IV. 20 : Amplitude de puissance LF des roulements (1, 2, 3 et 4)

Nous pouvons maintenant voir les candidats des caractéristiques extraites pour la phase d'essaimage. Cela se fait d'abord par une inspection visuelle à la recherche de caractéristiques qui ont une tendance évidente sur tout leur cycle de vie. Les caractéristiques élues par inspection visuelle (I.V.) sont présentés au tableau IV. 2.

Après l'essai, le meilleur point final (coefficients de pondération) correspondant aux caractéristiques élues est présenté dans le tableau IV. 2:

Tableau IV. 2: Résultats après l'essai

Caractéristiques élues par inspection visuelle (I.V.)	Meilleur point final
asymétrie	1.1165
Kurtosis	0.58516
RMS	0.63503
VHF	0.79954
HF	0.10644
MF	0.18937
LF	0.68993

Les caractéristiques identifiées pour les quatre roulements sont montrées dans figure IV. 21. Il est clair qu'elles sont affectées par la forme de la caractéristique VHF. Ils présentent une légère tendance monotone croissante. Cet effet est indésirable pour l'étape de prédiction.

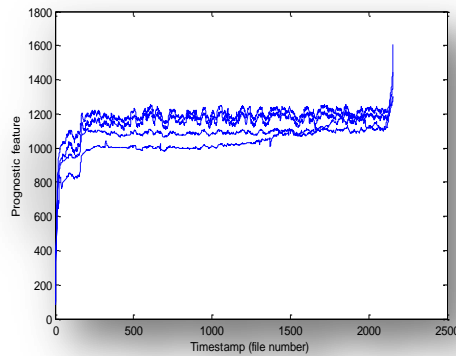


Figure IV. 21 : Caractéristiques de pronostic identifiées pour les roulements (1, 2, 3 et 4)

IV. 8 PRÉDICTION DE LA RUL

Le HSMM a été adopté pour la prédiction de la RUL. Les courbes des caractéristiques identifiées peuvent être divisées en trois grandes parties, donc trois états. L'état du milieu occupe une plage importante de l'espace du cycle de vie.

Selon la théorie énoncée précédemment pour le HSMM et après la phase d'apprentissage, les paramètres HSMM estimés sont:

L'état antérieur: $[1, 0, 0]$,

La matrice de transition:

$$\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix},$$

La durée moyenne d'état $[96.496 \ 563 \ 28.548]$,

La durée de variance de d'état $[5.8149 \ 392.0078 \ 0.5770]$,

$\rho = 3.68$,

$D_{S1} = 117.894832$,

$$D_{S_2} = 2005.588704,$$

$$D_{S_3} = 30.67136,$$

La durée moyenne et la variance de chaque état sont utilisées pour calculer la RUL suivant l'équation (IV. 10).

La conséquence du peu d'états est la dégradation de la qualité de la prédiction de la RUL.

IV.9 CONCLUSION

L'approche présentée dans ce chapitre est basée sur l'identification des caractéristiques de pronostic avancées susceptibles d'améliorer la prédiction de la durée de vie utile restante (RUL) en utilisant l'algorithme d'optimisation par essaim particulaire (PSO) et le modèle semi-markov caché (HSMM) pour estimer cette prédiction.

Dans notre approche, nous avons proposé un cadre unifié basé sur un HSMM pour le pronostic. L'inconvénient majeur du HSMM est la complexité de calcul qui peut augmenter pour les procédures d'inférence et d'estimations des paramètres. À cet égard, la modélisation de la durée par des distributions de probabilité paramétriques, gaussiennes dans notre cas, a été utilisée pour alléger les calculs.

Cette approche a été validée par un ensemble de données recueillies à partir des tests de vieillissements accélérés sur des roulements. Les résultats des caractéristiques identifiées présentent une légère amélioration de la qualité de tendance; cela est dû principalement au fait que la caractéristique fortement pondérée, parmi toutes les caractéristiques candidates, garde son empreinte sur la caractéristique identifiée. Pour élever cet inconvénient, de nouvelles mesures d'aptitude devront être utilisées.

IV. 10 REFERENCES

- [1] A. Heng, S. Zhang, A. C. Tan, and J. Mathew, "Rotating Machinery Prognostics: State of the Art, Challenges and Opportunities," *Mechanical Systems and Signal Processing*, vol. 23, no. 3, p. 724 – 739, 2009.
- [2] J.Z. Sikorska, M. Hodkiewicz, L. Mac, "Prognostic Modelling Options for Remaining Useful Life Estimation by Industry," *Mechanical Systems and Signal Processing*, vol. 25, p 1803–1836, 2011.
- [3] J.B. Yu, "Local and Nonlocal Preserving Projection for Bearing Defect Classification and Performance Assessment," *IEEE Transactions on Industrial Electronics*, vol. 59, no. 5, 2012.
- [4] Tahar Boukra, Abdesselam Lebaroud, Guy Clerc, "Statistical and Neural-Network Approaches for the Classification of Induction Machine Faults Using the Ambiguity Plane Representation," *IEEE Transactions on Industrial Electronics*, vol. 60, no. 9, p.4034-4042, 2013.
- [5] B. B. Sassi S. and T. M., "TALAF and THIKAT as innovative time domain indicators for tracking ball bearings," *Proceedings of the 24nd Seminar on machinery vibration (M. Thomas, ed.)*, p. 404–419, Canadian Machinery Vibration Association, 2005.
- [6] J. Coble and J.W. Hines, "Identifying Optimal Prognostic Parameters from Data: A Genetic Algorithms Approach," *Annual Conference of the Prognostics and Health Management Society*, 2009.
- [7] J. Coble. (2010) "Merging Data Sources to Predict Remaining Useful Life – An Automated Method to Identify Prognostic Parameters," PhD diss., University of Tennessee. [Online]. http://trace.tennessee.edu/utk_graddiss/683
- [8] F. Camci, K. Medjaher, N. Zerhouni, and P. Nectoux, "Feature Evaluation for Effective Bearing Prognostics," *International Quality and Reliability Engineering*, 2012.
- [9] Linxia Liao, "Discovering Prognostic Features Using Genetic Programming in Remaining Useful Life Prediction," *IEEE Transactions on Industrial Electronics*, vol. 61, no. 5, p. 2464-2472. 2014.
- [10] A. K. Jardine, D. Lin, and D. Banjevic, "A Review On Machinery Diagnostics And Prognostics Implementing Condition-Based Maintenance," *Mechanical Systems and Signal Processing*, vol. 20, no. 7, p. 1483 – 1510, 2006.
- [11] H. Qiu, J. Lee, J. Lin, and G. Yu, "Wavelet filter-based weak signature detection method and its application on rolling element bearing prognostics," *Journal of Sound and Vibration.*, vol. 289, no. 45, p. 1066 – 1090, Feb. 2006.
- [12] Eberhart, R.C., J. Kennedy. "A New Optimizer Using Particle Swarm Theory," *Proceedings of the Sixth International Symposium on Micro Machine and Human Science*. Nagoya, Japan, 1995.
- [13] Razik, H. ; Oumaamar, M.E.K. "An Application Of Particle Swarm Optimization To The Tracking Of Faulty Lines For The Monitoring Of Outer-Race Bearing Faults," *International Conference on Condition Monitoring and Diagnosis (CMD)*, p. 710- 713, 2012.
- [14] N. Bracikowski, M. Hecquet, P. Brochet, Sergey V. Shirinskii "Multiphysics Modeling of a Permanent Magnet Synchronous Machine by Using Lumped Models," *IEEE Transactions on Industrial Electronics*, vol. 59, no. 6, June 2012.
- [15] Ming Yu , Danwei Wang , Qijun Chen "Prediction of multiple failures for a mobile robot steering system," *International Symposium on Industrial Electronics IEEE*, p. 1240-1245, 2012.

- [16] Pasquale Arpaia, Member, IEEE, Carlo Manna, Giuseppe Montenero, and Giovanni D'Addio "In-Time Prognosis Based on Swarm Intelligence for Home-Care Monitoring: A Case Study on Pulmonary Disease," *IEEE Sensors Journal*, vol. 12, no. 3, 2012.
- [17] C. Kwan, X. Zhang, R. Xu, L. Haynes, "A novel approach to fault diagnostics and prognostics," *Proceedings of the 2003 IEEE International Conference on Robotics and Automation*, Taipei, Taiwan, vol. 1, p. 604–609, 2003.
- [18] Andrey Andreyevich Markov, "Theory of Algorithms," Academy of Sciences of the USSR, 1954.
- [19] K. P. Zhua, G. S. Hong, and Y. S. Wong, "A Comparative Study of Feature Selection for Hidden Markov Model-Based Micro-Milling Tool Wear Monitoring," *Machining Science and Technology*, vol. 12, no. 3, p. 348-369, 2008.
- [20] P. Baruah, R.B. Chinnam, "HMMs for diagnostics and prognostics in machining processes," *International Journal of Production Research*, vol. 43, p. 1275–1293, 2005.
- [21] M. Dong, and D. He, "A segmental hidden semi-Markov Model (HSMM)-based diagnostics and prognostics framework and methodology," *Mechanical Systems and Signal Processing*, vol. 21, p. 2248-2266, 2007.
- [22] L. R. Rabiner, "A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition," *Proceedings of the IEEE*, vol. 77, no. 2, p. 257–286, 1989.
- [23] Geramifard, O., Jian-Xin Xu , Jun-Hong Zhou , Xiang Li, "Multimodal Hidden Markov Model-Based Approach for Tool Wear Monitoring," *IEEE Transactions on Industrial Electronics*, vol. 61 , no. 6, p. 2900- 2911, 2014.
- [24] Omid Geramifard, Jian-Xin Xu, Jun-Hong Zhou, and Xiang Li, "A Physically Segmented Hidden Markov Model Approach for Continuous Tool Condition Monitoring: Diagnostics and Prognostics," *IEEE Transactions on Industrial Informatics*, vol. 8, no. 4, 2012.
- [25] S.Z. Yu, "Hidden Semi Markov Models", *Artificial Intelligence*, vol. 174, no. 2, p. 215-243, 2010.
- [26] S-Z. Yu and H. Kobayashi, "An Efficient Forward-Backward Algorithm for an Explicit Duration Hidden Markov Model," *IEEE Signal Processing Letters*, vol. 10, no. 1, p. 11-14, 2003.
- [27] Strangas, E.G. ; Aviyente, S. , Neely, J.D. , Zaidi, S.S.H. "The Effect of Failure Prognosis and Mitigation on the Reliability of Permanent-Magnet AC Motor Drives," *IEEE Transactions on Industrial Electronics*, vol. 60 , no. 8, p. 3519 – 3528, 2013.
- [28] Xianping Ge, Padhraic Smyth "Hidden Markov Models for Endpoint Detection in Plasma Etch Processes" *Technical Report UCI-ICS 01-54 Department of Information and Computer Science University of California*, Irvine, September, 2001.
- [29] G. Soules L. E. Baum, T. Petrie and N. Weiss, "A Maximization Technique Occurring in The Statistical Analysis of Probabilistic Functions of Markov Chains," *The Annals of Mathematical Statistics*, vol. 41, p.164-171, 1970.
- [30] Krogh A. Durbin R., Eddy S. and Mitchison, "Biological Sequence Analysis," Cambridge University Press, Cambridge, MA, 1998.
- [31] Tahar Boukra, Abdesselam, "Identifying New Prognostic Features for Remaining Useful Life Prediction," 16th International Power Electronics and Motion Control Conference and Exposition (PEMC 2014) Antalya, Turkey 21-24, p:1437-1443, Sept 2014.

V CONCLUSIONS ET TRAVAIL FUTUR

Cette conclusion résume dans la première section les contributions de ce travail comme réponses aux défis soulevés. Des suggestions pour le travail futur sont ébauchées dans la deuxième section.

V.1 CONCLUSIONS

Les travaux présentés dans cette thèse traitent l'intégration du diagnostic et du pronostic, qui sont des processus clés de la CMB, dans un cadre d'un PHM intégral pour optimiser la maintenance des composants de la machine. Le diagnostic permet de déterminer les composants de la machine en défaut qui doivent être réparés. Le pronostic permet, quand à lui, d'anticiper les défaillances du système en fournissant des informations sur les états futurs des composants à partir desquelles des actions de maintenance préventive peuvent être envisagées.

Nous avons montré, au Chapitre II, l'intérêt d'une stratégie de maintenance préventive et nous avons dégagé dans un premier temps les modèles et les méthodes qui s'appuient sur une connaissance plus ou moins approfondie des composants à travers une revue extensive de la littérature sur le diagnostic et le pronostic des machines. Ce chapitre nous a permis, entre autre, de limiter le cadre de notre travail et de l'orienter. Pour cela, nous avons restreint notre travail pour la partie diagnostic à l'amélioration du taux de classification correcte des défauts du moteur à induction tels que les défauts des roulements, les défauts déséquilibre des phases statoriques et les défauts cassures de barres rotoriques.

A cette fin, nous avons proposé, au chapitre III, une nouvelle approche de classification basée sur la RTF et un critère de décision. Nous avons opté de concevoir un classifieur directement dans le plan d'auto-ambiguïté utilisé comme espace d'extraction des caractéristiques pertinentes. Comme toutes les RTF peuvent être dérivées du plan d'auto-ambiguïté, aucune hypothèse a priori n'est faite concernant le lissage exigé pour la classification. Le contraste de Fisher, basé sur la compacité et la séparabilité des classes, est utilisé pour la séparation des classes. A cet égard, nous avons conçu un noyau spécifique pour la séparation de chaque type de défaut (trois noyaux au total).

Pour la partie sélection, nous avons appliqué en premier lieu une approche sous-optimale pour exclure les points du noyau fortement corrélés avec les points du noyau de rang supérieur car ils ne contribuent pas à améliorer la classification. Ensuite, un nombre optimal de points de noyau a été sélectionné en minimisant un modèle probabiliste d'erreur de classification proposé pour l'approche statistique. Ce dernier dépend de deux paramètres à savoir le type de noyau utilisé pour construire le vecteur de forme et une mesure de distance (Mahalanobis ou euclidienne).

Les résultats de classification correcte obtenus à partir du classifieur statistique optimisé avec la distance Mahalanobis comme critère de décision atteignent celles des classifieurs à base de RNA avec sigmoïde comme fonction de transfert et l'algorithme de Levenberg-Marquardt comme algorithme d'apprentissage. Ces derniers ont une nette supériorité sur le WNN.

Enfin, ces résultats vérifient que le système de classification, proposé et testé sur des données expérimentales recueillies à partir de la mesure du courant statorique à différentes charges, à

base de RNA entraînés avec l'algorithme Levenberg-Marquardt comme critère de décision a une capacité diagnostique plus précise indépendamment du niveau de charge et du type de défaut.

Pour la partie pronostic, nous avons orienté le cadre de notre travail, au chapitre IV, vers le pronostic au niveau composant que nous jugeons critique: approche « locale » au lieu d'une approche « globale » niveau système. L'approche de pronostic guidée par les données est privilégiée dans notre cas pour les raisons suivantes :

- Difficulté de disposer de connaissance sur les phénomènes de dégradation complexes.
- Insuffisance d'expériences passées pour déployer des approches statistiques significatives à cause de la variabilité de conditions d'usage et de détérioration.
- Envisagement d'une stratégie de maintenance plus « dynamique » basée sur les données conditionnelles du composant au lieu d'une stratégie de maintenance préventive systématique.

L'approche que nous avons présentée dans ce chapitre est basée sur l'identification des caractéristiques de pronostic avancées susceptibles d'améliorer la prédiction de la durée de vie utile restante (RUL) en utilisant l'algorithme d'optimisation par essaim particulaire (PSO) et le modèle semi-Markov caché (HSMM) pour estimer cette prédiction.

Nous avons abordé le problème du pronostic en effectuant une analyse de données sur des tests de vieillissement accélérés des roulements en utilisant des méthodes statistiques en domaine temporel et fréquentiel. Cette analyse nous a permis d'identifier des indicateurs de vieillissement par inspection visuelle susceptibles d'être candidats pour la phase d'essaimage.

Les résultats des caractéristiques identifiées après la phase d'essaimage présentent une légère amélioration de la qualité de tendance; cela est dû principalement au fait que la caractéristique fortement pondérée, parmi toutes les caractéristiques candidates, garde son empreinte sur la caractéristique identifiée. Pour élever cet inconvénient, de nouvelles mesures d'aptitude devraient être utilisées.

V.2 TRAVAIL FUTUR

Si le diagnostic a atteint une certaine maturité, le pronostic est, quant à lui, une discipline émergente. Plusieurs opportunités pour continuer les recherches au-delà de l'étendue de cette thèse.

Quelques idées formulées sont présentées comme suit :

- L'idée est de tenir compte de l'hétérogénéité des composants pour une approche globale orientée système au lieu d'une approche orientée composant.
- Tenir compte de des modes de fonctionnements (modes opérationnels) car Les conditions opérationnelles et environnementales impactent sur la dynamique d'évolutions des modes de dégradation/défaillance et des déviations. Ces effets doivent être modélisés pendant la phase de projection.

- Employer des critères de performance unifiés pour évaluer les résultats de la prédiction.
- Proposition de nouvelles mesures d'aptitude pour la fonction fitness afin d'améliorer la tendance des caractéristiques de pronostic identifiées.
- D'avantage de validation des modèles proposés sur les défaillances des machines en domaine industriel.
- Finalement, en vue d'une application en monde réel de supervision, il est nécessaire de développer un outil logiciel de gestion de santé intégré qui peut être utilisé dans la détection, le diagnostic et le pronostic des défauts des composants de la machine.

ANNEXE A

A. ALGORITHME AVANT-ARRIERE

Problème 1 est un problème d'évaluation. Etant donné un modèle λ , nous voulons calculer la probabilité que la séquence $O = O_1 O_2 \dots O_T$ a été produite par ce modèle. Le fait de considérer toutes les séquences d'état possibles et ensuite les additionner est irréaliste, puisque le nombre de chemins possibles augmente de façon exponentielle avec la longueur de la séquence T . Les algorithmes avant et arrière [22] décrits dans cette section résolvent efficacement ce problème.

En premier lieu, considérant l'algorithme avant. Définissant la variable avant $\alpha_t(i)$ comme

$$\alpha_t(i) = P(O_1 O_2 \dots O_t, q_t = S_i | \lambda)$$

c'est-à-dire, la probabilité de la séquence d'observation partielle, $O_1 O_2 \dots O_t$ et l'état S_i au temps t , étant donné les paramètres du modèle λ . Nous pouvons ainsi calculer $\alpha_t(i)$ par induction :

1. initialisation :

$$\alpha_1(i) = \pi_i b_i(O_1), \quad 1 \leq i \leq N$$

2. induction :

$$\begin{aligned} \alpha_t(i) &= P(O_1 O_2 \dots O_t, q_t = S_j | \lambda) \\ &= \sum_{i=1}^N P(O_1 O_2 \dots O_t, q_{t-1} = S_i, q_t = S_j | \lambda) \\ &= \left[\sum_{i=1}^N \alpha_{t-1}(i) a_{ij} \right] b_j(O_t), \quad t = 2, 3, \dots, T, \quad 1 \leq i \leq N \end{aligned}$$

3. Finalisation :

$$P(O | \lambda) = \sum_{i=1}^N \alpha_T(i)$$

Étape (1) Initialiser les probabilités « avant » par multiplication de la probabilité de l'état initial S_i et la probabilité d'observation O_1 à S_i . L'étape d'induction (2) calcule $\alpha_t(i)$ en additionnant tous les N états possibles S_i au temps t qui sont accessibles à l'état S_i au temps $t + 1$ via les transitions d'état, puis en la multipliant par la probabilité d'observation $b_j(O_t)$. L'étape de terminaison (3) donne le calcul résultant de $P(O | \lambda)$ comme la somme de tous les N variables « avant » terminaux $\alpha_T(i)$, puisque par définition,

$$\alpha_T(i) = P(O_1 O_2 \dots O_T, q_T = S_i | \lambda)$$

$P(O | \lambda)$ est basé sur la somme de $\alpha_T(i)$ sur i . Il est facile de voir que l'algorithme avant a une complexité de calcul de $O(TNN)$.

L'algorithme arrière est présenté d'une façon similaire. Définir la variable arrière $\beta_t(i)$ comme probabilité de la séquence d'observation partielle de $t + 1$ à la fin de la séquence, compte tenu de l'état S_i au temps t et les paramètres du modèle λ ,

$$\beta_t(i) = P(O_{t+1} O_{t+2} \dots O_T | q_t = S_i, \lambda)$$

Encore une fois, nous pouvons calculer $\beta_t(i)$ par induction :

Initialisation :

$$\beta_T(i) = 1, \quad 1 \leq i \leq N$$

Induction :

$$\begin{aligned} \beta_t(i) &= P(O_{t+1}O_{t+2} \dots O_T | q_t = S_i, \lambda) \\ &= \sum_{j=1}^N P(O_{t+1}O_{t+2} \dots O_T | q_t = S_i, q_{t+1} = S_j, \lambda) \\ &= \sum_{j=1}^N \beta_{t+1}(j) a_{ij} b_j(O_{t+1}), \quad t = T-1, T-2, \dots, 1, \quad 1 \leq i \leq N \end{aligned}$$

3. Finalisation :

$$P(O|\lambda) = \sum_{i=1}^N \pi_i b_i(O_1) \beta_1(i)$$

Étape (1) Initialise arbitrairement les probabilités « arrières » toutes à 1. L'étape d'induction (2) calcule $\beta_t(i)$ en additionnant sur tous les N états possibles S_j au temps $t + 1$ qui sont accessibles depuis l'état S_i au temps t via la transitions d'état, tenant compte de la probabilité d'observation $b_j(O_{t+1})$ respectivement. Encore une fois, l'étape de finalisation (3) donne le résultat de calcul de $P(O|\lambda)$, puisque par définition,

$$\beta_1(i) = P(O_2 \dots O_T | q_1 = S_i, \lambda)$$

$P(O|\lambda)$ est basé sur la somme de $\beta_1(i)$, tenant compte de l'état initial et de la probabilité d'observation $b_i(O_1)$ respectivement. La complexité de l'algorithme en arrière est également $O(TNN)$.

Pour éviter l'erreur de « soupassement » lorsque les algorithmes sont implémentés sur ordinateur, on multiplie généralement $\alpha_t(i)$ par un facteur d'échelle c_t à chaque indice de temps t , afin qu'il reste dans un intervalle numérique gérable [30]. Nous définissons généralement c_t comme

$$c_t = \frac{1}{\sum_{i=1}^N \alpha_t(i)}$$

Donc, à chaque indice t , nous calculons tout d'abord

$$\alpha_t(j) = \left[\sum_{i=1}^N \hat{\alpha}_{t-1}(i) a_{ij} \right] b_j(O_t)$$

Puis la version normalisée $\hat{\alpha}_t(j)$ est calculée à partir

$$\hat{\alpha}_t(j) = \frac{[\sum_{i=1}^N \hat{\alpha}_{t-1}(i) a_{ij}] b_j(O_t)}{\sum_{j=1}^N \{[\sum_{i=1}^N \hat{\alpha}_{t-1}(i) a_{ij}] b_j(O_t)\}}$$

La variable arrière doit être normalisée avec le même coefficient d'échelle à chaque index de temps, c'est-à-dire

$$\hat{\beta}_t(j) = c_t \beta_t(j)$$

Étant donné que les amplitudes de la variable avant et arrière sont comparables, utiliser le même facteur d'échelle c_t à chaque index de temps est un moyen efficace pour garder leurs valeurs dans des limites facilement gérables.

B. ALGORITHME DE BAUM-WELCH

Problème 3 est un problème d'estimation de paramètres. On veut optimiser les paramètres du modèle afin qu'ils « donnent plus de sens » à la séquence d'observation donnée générée par ce modèle. Nous appelons la séquence observation utilisée pour ajuster les paramètres de modèle la séquence « d'apprentissage ». Dans cette section, nous présentons une méthode couramment utilisée, appelée algorithme de Baum-Welch [29] qui ajustera les paramètres du modèle de telle façon que $P(O | \lambda)$ est maximisée localement.

L'algorithme de Baum-Welch est un genre d'algorithme Expectation Maximisation (EM), qui est un algorithme général pour l'estimation de la vraisemblance maximale (ML) avec « données manquantes ».

Pour les HMMs, les données manquantes sont la séquence d'état sous-jacente inconnue, puisque nous pouvons connaître les observations, mais pas la séquence d'état sous-jacente qui les génère.

Voici l'algorithme de Baum-Welch :

1. initialisation : initialiser les paramètres du modèle, c'est-à-dire λ .
2. Récurrence:
 - Calculer les $\alpha_t(i)$ et $\beta_t(i)$ à l'aide de l'algorithme avant et arrière.
 - Calculer :

$$a_{ij}^{new} = \frac{\text{nombre attendu de transitions de l'état } S_i \text{ à l'état } S_j}{\text{nombre de transitions attendus de l'état } S_i}$$

$$= \frac{\sum_{t=1}^N \alpha_t(i) a_{ij} b_j(O_{t+1}) \beta_{t+1}(j)}{\sum_{j=1}^N \sum_{t=1}^N \alpha_t(i) a_{ij} b_j(O_{t+1}) \beta_{t+1}(j)}$$

$$b_j^{new}(k) = \frac{\text{nombre de fois dans l'état } S_j \text{ et observant symbol } v_k}{\text{nombre de fois attendus en état } S_j}$$

$$= \frac{\sum_{t=1}^T \alpha_t(j) \beta_t(j)}{\sum_{k=1}^M \sum_{t=1}^T \alpha_t(j) \beta_t(j)}$$

$$\pi_i^{new} = \text{nombre de fois en état } S_i \text{ au temps } t = 1$$

$$= \frac{\pi_i b_i(O_1) \beta_1(i)}{P(O|\lambda)}$$

3. Finalisation : Arrêter si le changement dans le log vraisemblance est inférieur à un seuil prédéfini ou dépassement du nombre maximal d'itérations.

Étape (1) donne des valeurs initiales raisonnables pour les paramètres du modèle. À l'étape de récurrence (2), les $\alpha_t(i)$ et $\beta_t(i)$ sont calculées en utilisant l'algorithme avant et arrière.

$P(q_t = S_i, q_{t+1} = S_j | O, \lambda) = \frac{b_j(O_{t+1}) \beta_{t+1}(j) \alpha_t(i) a_{ij}}{P(O|\lambda)}$ est la probabilité d'être à l'état S_i au temps t , ou S_j au temps $t + 1$, étant donné le modèle λ et la séquence d'observation.

$P(q_t = S_j | O, \lambda) = \frac{\alpha_t(j)\beta_t(j)}{P(O|\lambda)}$ est la probabilité d'être en état S_j au temps t , étant donné le modèle λ et la séquence d'observation. Chaque récursion aide à la convergence de $P(O|\lambda)$ vers un maximum local [30]. L'étape de finalisation (3) arrête l'algorithme lorsque le changement dans le log vraisemblance est suffisamment petit, ou le nombre prédéfini d'itérations est atteint.

Comme indiqué, les algorithmes avant et arrière utilisent la procédure de mise à l'échelle pour éviter le problème de soupassement. Puisque les mêmes coefficients apparaissent dans le numérateur et le dénominateur dans les formules ci-dessus, ils s'annulent tout à fait, les formules restent inchangées lorsque nous utilisons la version de $\alpha_t(j)$ et $\beta_t(j)$ mise à échelle. Pour une démonstration détaillée, veuillez consulter [22].

C. ALGORITHME DE VITERBI

Problème 2 est un problème de décodage. Nous allons essayer de découvrir la partie cachée du modèle. Nous voulons trouver la séquence d'état sous-jacente la plus « raisonnable » pour une séquence d'observation. Il y a plusieurs façons d'y parvenir et ceci selon la définition des « critères d'optimalité ». Dans cette section, nous allons décrire la plus commune, connue sous le nom d'algorithme de Viterbi, qui est un algorithme de programmation dynamique. L'algorithme de Viterbi trouve une seule « meilleure » séquence d'état (avec la probabilité la plus élevée), $Q = \{q_1 q_2 \dots q_T\}$ pour une séquence donnée d'observations $O = \{O_1 O_2 \dots O_T\}$. Nous définissons la variable Viterbi comme

$$\delta_t(i) = \max_{q_1, q_2, \dots, q_{t-1}} P(q_1 q_2 \dots q_t = i, O_1 O_2 \dots O_t | \lambda)$$

c'est-à-dire, $\delta_t(i)$ est le chemin d'accès le plus probable se terminant à l'état i au temps t avec l'observation O_t . Nous définissons également $\psi_t(i)$ pour mémoriser l'état qui transite à l'état i par le chemin d'accès le plus probable, afin de garder une trace des chemins les plus probables. L'algorithme de Viterbi peut être aussi calculé par induction comme suit

1. Initialisation:

$$\begin{aligned} \delta_1(i) &= \pi_i b_i(O_1), 1 \leq i \leq N \\ \psi(i) &= 0, 1 \leq i \leq N. \end{aligned}$$

2. Induction:

$$\delta_t(i) = b_i(O_t) \max_{j=1}^N (\delta_{t-1}(j) a_{ij}), 2 \leq t \leq T, 1 \leq i \leq N:$$

$$\psi_t(i) = \operatorname{argmax}_{j=1}^N (\delta_{t-1}(j) a_{ij})$$

3. Finalisation :

$$\begin{aligned} P^* &= \max_{i=1}^N (\delta_T(i)) \\ q_T^* &= \operatorname{argmax}_{i=1}^N (\delta_T(i)) \end{aligned}$$

4. Chemin de retour en arrière:

$$q_t^* = \psi_{t+1}(q_{t+1}^*), \quad t = T - 1, T - 2, \dots, 1$$

L'algorithme de Viterbi est similaire à l'algorithme avant. La principale différence est que l'opération somme est maintenant remplacée par une opération de maximisation. Pour éviter l'erreur de soupassement pendant le calcul, généralement Viterbi est exécuté en espace log, donc les étapes ci-dessus deviennent :

Initialisation :

$$\begin{aligned}\delta_1(i) &= \log(\pi_i) + \log(b_i(O_1)), 1 \leq i \leq N \\ \psi(i) &= 0, 1 \leq i \leq N.\end{aligned}$$

Induction :

$$\begin{aligned}\delta_t(i) &= \log(b_i(O_t)) \max_{i=1}^N (\delta_{t-1}(i) + \log(a_{ij})), 2 \leq t \leq T, 1 \leq j \leq N: \\ \psi_t(i) &= \operatorname{argmax}_{i=1}^N (\delta_{t-1}(i) + \log(a_{ij}))\end{aligned}$$

Finalisation :

$$\begin{aligned}P^* &= \max_{i=1}^N (\delta_T(i)) \\ q_T^* &= \operatorname{argmax}_{i=1}^N (\delta_T(i))\end{aligned}$$

Chemin de retour en arrière :

$$q_t^* = \psi_{t+1}(q_{t+1}^*), \quad t = T - 1, T - 2, \dots, 1$$

Les variables avant-arrière suivantes sont utilisées par HMMD

- $\alpha_t(i) = P(O_1 O_2 \dots O_t, S_i \text{ termine à } t | \lambda)$
- $\beta_t(i) = P(O_{t+1} O_{t+2} \dots O_T | S_i \text{ termine à } t, \lambda)$
- $\alpha_t^*(i) = P(O_1 O_2 \dots O_t, S_i \text{ débute à } t + 1 | \lambda)$
- $\beta_t^*(i) = P(O_{t+1} O_{t+2} \dots O_T | S_i \text{ débute à } t + 1, \lambda)$

où $\alpha_t(i)$ peut être calculée par

$$\alpha_t(i) = \sum_{j=1}^N \sum_{d=1}^D \alpha_{t-d}(j) a_{ji} p_i(d) \prod_{s=t-d+1}^t b_i(O_s)$$

D'autres peuvent être calculés de la même façon. Les relations entre les α , α^* , β^* et β sont

$$\begin{aligned}\alpha_t^*(i) &= \sum_{j=1}^N \alpha_t(j) a_{ij} \\ \beta_t^*(i) &= \sum_{d=1}^D \beta_{t+d}(i) p_i(d) \prod_{s=t+1}^{t+d} b_i(O_s) \\ \alpha_t(i) &= \sum_{d=1}^D \alpha_{t-d}^*(i) p_i(d) \prod_{s=t-d+1}^t b_i(O_s) \\ \beta_t(i) &= \sum_{j=1}^N a_{ij} \beta_t^*(j)\end{aligned}$$

ANNEXE B

Rabiner donne les formules de ré-estimation des maximum vraisemblances suivantes (algorithm de Baum-Welch) pour HMMD

$$\begin{aligned}\pi_i^{\text{new}} &= \frac{\pi_i \beta_0^*(i)}{P(O|\lambda)} \\ a_{ij}^{\text{new}} &= \frac{\sum_{t=1}^N \alpha_t(i) a_{ij} \beta_t^*(j)}{\sum_{j=1}^N \sum_{t=1}^N \alpha_t(i) a_{ij} \beta_t^*(j)} \\ b_i^{\text{new}}(k) &= \frac{\sum_{t=1}^T \sum_{s.t.O_t=k} \{\sum_{r<t} \alpha_r^*(j) \beta_r^*(j) - \sum_{r<t} \alpha_r(j) \beta_r(j)\}}{\sum_{k=1}^M \sum_{t=1}^T \sum_{s.t.O_t=k} \{\sum_{r<t} \alpha_r^*(j) \beta_r^*(j) - \sum_{r<t} \alpha_r(j) \beta_r(j)\}} \\ p_i^{\text{new}}(d) &= \frac{\sum_{t=1}^T \alpha_t^*(i) p_i(d) \beta_{t+d}(i) \prod_{s=t+1}^{t+d} b_i(O_s)}{\sum_{d=1}^D \sum_{t=1}^T \alpha_t^*(i) p_i(d) \beta_{t+d}(i) \prod_{s=t+1}^{t+d} b_i(O_s)}\end{aligned}$$

A. ALGORITHME DE BAUM-WELCH POUR HMMD

Définir la variable avant

$$\begin{aligned}\alpha_t(i, d) &= \begin{cases} \left[\sum_{j=1, j \neq i}^N \hat{\alpha}_{t-1}(j) a_{ij} \right] b_i(O_t) & \text{if } d = 1 \\ \alpha_{t-1}(i, d-1) b_i(O_t) & \text{if } 2 \leq d \leq D \end{cases} \\ &= \begin{cases} \check{\alpha}_{t-1}(i) b_i(O_t) & \text{if } d = 1 \\ \alpha_{t-1}(i, d-1) b_i(O_t) & \text{if } 2 \leq d \leq D \end{cases} \quad (\text{A.1})\end{aligned}$$

où

$$\hat{\alpha}_t(i) = \sum_{d=1}^D \alpha_t(i, d) p_i(d) \quad (\text{A.2})$$

$$\check{\alpha}_t(i) = \sum_{j=1, j \neq i}^D \hat{\alpha}_t(j) a_{ij} \quad (\text{A.3})$$

Il est facile de voir la relation suivante

$$\alpha_t(i, d) p_i(d) = P(O_1 O_2 \dots O_t, S_i \text{ termine à } t \text{ avec durée } d | \lambda)$$

Une transition vers l'état i avec durée d a lieu soit du même état i avec durée $d - 1$ dans l'index de temps précédent, si $d > 1$, ou de tous les autres états qui sont autorisés à transiter à l'état i , si $d = 1$. Pour enregistrer les calculs, la probabilité de durée $p_i(d)$ est liée à l'état quand l'état va pouvoir transiter vers d'autres états, comme le montre éq.(A.2), où $\hat{\alpha}_t(i) = P(O_1 O_2 \dots O_t, S_i \text{ termine à } t | \lambda)$ de même on défini la variable arrière

$$\beta_t(i, d) = \begin{cases} \left[\sum_{j=1, j \neq i}^N \hat{\beta}_{t+1}(j) a_{ij} \right] b_i(O_t) & \text{if } d = 1 \\ \beta_{t+1}(i, d-1) b_i(O_t) & \text{if } 2 \leq d \leq D \end{cases}$$

$$= \begin{cases} \check{\beta}_{t+1}(i)b_i(O_t) & \text{if } d = 1 \\ \beta_{t+1}(i, d-1)b_i(O_t) & \text{if } 2 \leq d \leq D \end{cases} \quad (\text{A.4})$$

où

$$\hat{\beta}_t(i) = \sum_{d=1}^D \beta_t(i, d)p_i(d) \quad (\text{A.5})$$

$$\check{\beta}_t(i) = \sum_{j=1, j \neq i}^D \hat{\beta}_t(j)a_{ij} \quad (\text{A.6})$$

et nous avons la relation suivante

$$\beta_t(i, d)p_i(d) = P(O_1 O_2 \dots O_t, S_i \text{ a une durée restante } d \text{ à } t, \lambda)$$

Un état i d'une durée de vie restante d à l'index de temps actuel restera au même état i avec une durée de vie restante $d - 1$ à l'index de temps suivant, si sa valeur de durée $d > 1$. Sinon, si la durée $d = 1$ à l'index actuel de temps, il doit transiter vers d'autres états admissibles à l'index de temps suivant, comme indiqué dans éq.(A.6).

Étant donné les variables avant et arrière ci-dessus, α , α^* , β^* et β dans les formules de la Rabiner peuvent être exprimées par

$$\alpha_t^*(i) = \check{\alpha}_t(i)$$

$$\beta_t^*(i) = \hat{\beta}_{t+1}(i)$$

$$\alpha_t(i) = \hat{\alpha}_t(i)$$

$$\beta_t(i) = \check{\beta}_{t+1}(i)$$

Pour plus de commodité dans ce qui suit, nous définissons maintenant

$$\begin{aligned} \omega(t, i, d) &= P(O_1 O_2 \dots O_T, q_{t-1} \neq S_i, q_t \dots q_{t+d-1} = S_i, q_{t+d} = S_i | \lambda) \\ &= \check{\alpha}_{t-1}(i)\beta_t(i, d)p_i(d) \end{aligned} \quad (\text{A.7})$$

$$\begin{aligned} \mu_t(i, j) &= P(O_1 O_2 \dots O_T, q_t = S_i, q_{t+1} = S_j | \lambda) \\ &= \hat{\alpha}_t(i)\hat{\beta}_{t+1}(j)a_{ij} \end{aligned} \quad (\text{A.8})$$

$$\varphi(i, j) = \sum_{t=1}^{T-1} \mu_t(i, j) \quad (\text{A.9})$$

$$v_t(i) = P(O_1 O_2 \dots O_T, q_t = S_i | \lambda)$$

Dans ce qui suit on fait usage de la relation suivante [26]

$$\begin{aligned} &P(O_1 O_2 \dots O_T, q_t = S_i, q_{t+1} = S_i | \lambda) \\ &= P(O_1 O_2 \dots O_T, q_t = S_i | \lambda) - P(O_1 O_2 \dots O_T, q_t = S_i, q_{t+1} \neq S_i | \lambda) \\ &= P(O_1 O_2 \dots O_T, q_{t+1} = S_i | \lambda) - P(O_1 O_2 \dots O_T, q_t \neq S_i, q_{t+1} = S_i | \lambda) \end{aligned}$$

duquel nous obtenons la formule de récurrence suivante

$$v_t(i) = \begin{cases} \hat{\alpha}_T(i) & \text{if } i = T \\ v_{t+1} + \sum_{j \neq i}^N (\mu_t(i, j) - \mu_t(j, j)) & \text{if } 1 \leq t \leq T - 1 \end{cases} \quad (\text{A. 10})$$

En utilisant la formulation ci-dessus, les équations de ré-estimation de Rabiner peuvent être exprimées par

$$\pi_i^{\text{new}} = \frac{\pi_i \hat{\beta}_1(i)}{P(O|\lambda)} \quad (\text{A. 11})$$

$$a_{ij}^{\text{new}} = \frac{\varphi(i, j)}{\sum_{j=1}^N \varphi(i, j)} \quad (\text{A. 12})$$

$$b_i^{\text{new}}(k) = \frac{\sum_{\substack{t=1 \\ s.t. O_t=k}}^T v_t(i)}{\sum_{k=1}^M \sum_{\substack{t=1 \\ s.t. O_t=k}}^T v_t(i)} \quad (\text{A. 13})$$

$$p_i^{\text{new}}(d) = \frac{\sum_{t=1}^T \omega(t, i, d)}{\sum_{d=1}^D \sum_{t=1}^T \omega(t, i, d)} \quad (\text{A. 14})$$

Voici un résumé de l'algorithme d'apprentissage de Baum-Welch

1. initialiser les éléments (λ) de HMMD.
2. calculer $\alpha_t(i, d)$ à l'aide des éqs. (A. 1) - (A. 3).
(enregistrez deux tables: $\hat{\alpha}_t(i)$ et $\check{\alpha}_t(i)$)
3. calculer $\beta_t(i)$ à l'aide des éqs. (A. 4)-(A. 6).
4. ré-estimation des éléments (λ) de HMMD à l'aide des éqs. (A. 11) - (A. 14).
5. terminer si la condition d'arrêt est satisfaite, sinon passez à l'étape 2.

B. ALGORITHME DE VITERBI POUR HMMD

définir

$$\delta_t(i, d) = \begin{cases} \left[\max_{j=1, j \neq i} \{V_{t-1}(j) a_{ij}\} \right] b_i(O_t) & \text{if } d = 1 \\ \delta_{t-1}(i, d-1) b_i(O_t) & \text{if } 2 \leq d \leq D \end{cases} \quad (\text{A. 15})$$

Où

$$V_t(i) = \max_{d=1}^D \{\delta_t(i, d) p_i(d)\}$$

pour lesquels nous avons une définition récursive utile

$\delta_t(i, d) p_i(d)$ = la probabilité du chemin le plus probable qui se termine à l'état i avec la durée d au temps t

Le but de l'algorithme est de trouver

$$\operatorname{argmax}_{i,d} \{\delta_T(i, d) p_i(d)\}$$

Comme une mise à l'échelle logarithmique pouvant être appliquée pour les termes a_{ij} , $b_i(k)$ et $p_i(d)$, les équations ci-dessus deviennent

$$\delta_t(i, d) = \begin{cases} \max_{j=1, j \neq i}^N \{V_{t-1}(j) + \log a_{ij}\} + \log b_i(O_t) & \text{if } d = 1 \\ \delta_{t-1}(i, d-1) + \log b_i(O_t) & \text{if } 2 \leq d \leq D \end{cases}$$

où

$$V_t(i) = \max_{d=1}^D \{\delta_t(i, d) + \log p_i(d)\}$$

L'objectif se simplifie à

$$\operatorname{argmax}_{i,d} \{\delta_T(i, d) + \log p_i(d)\}$$

La complexité de l'algorithme de Viterbi pour HMMD est $O(TNN + TND)$. Parce que la mise à l'échelle logarithmique peut être effectuée à l'avance, la procédure de Viterbi est composée seulement des additions, ce qui donne un calcul très rapide.

□□