

الجمهورية الجزائرية الديمقراطية الشعبية
وزارة التعليم العالي والبحث العلمي

UNIVERSITÉ BADJI MOKHTAR - ANNABA
BADJI MOKHTAR – ANNABA UNIVERSITY



جامعة باجي مختار – عنابة

Faculté : Sciences de L'ingéniorat
Département : Electronique
Domaine : Sciences et Techniques
Filière : Génie Électrique
Spécialité : Automatique et Système

Mémoire

Présenté en vue de l'obtention du Diplôme de Master
Thème :

**Classification des Défauts de Roulements d'une Machine
Asynchrone par Séparateurs à Vaste Marge**

Présenté par : *Delhoum Belkis*

Encadrant : *Arbaoui Fayçal*

Grade professeur

Université

Jury de Soutenance :

Saadi Med Nacer	M.C.B	U BMA	Président
Arbaoui Fayçal	Professeur	U BMA	Encadrant
Saidi Med Larbi	Professeur	U BMA	Examineur

Année Universitaire : 2020/2021

Dédicaces

Afin d'être reconnaissante envers ceux qui m'ont appuyée et encouragée à réaliser ce travail de recherche, je dédie ce mémoire.

À ma chère Maman **GOURI NAIMA** Qui était toujours derrière moi.

Je prie notre Seigneur **Allah Le Tout-Puissant, Le Clément et Le Miséricordieux**, pour qu'Il lui accorde la félicité éternelle en l'accueillant dans sa vaste demeure lui et tous les défunts membres de ma famille.

Tous les mots ne sauraient exprimer ma gratitude envers mon mari **LARBI ABDERRAHIME**, pour ses encouragements réguliers. Que DIEU te garde pour moi.

A ceux qui m'ont souhaitée de la chance : mon Père, mes Frères et mes Sœurs.

A ceux qui m'ont aidée, encouragée, appréciée mon effort et ont créé le milieu favorable, l'ambiance joyeuse et l'atmosphère joviale.

(Qu'**Allah Le Tout-Puissant**, les protège et les garde).

Remerciements

Au terme de ce travail, je remercie Allah le Tout-Puissant

الله سبحانه وتعالى

De m'avoir donné la volonté et le courage qui m'ont permis de réaliser ce travail, puisse-t-Il me guider dans le droit chemin. Louange à L'Éternel.

الحمد لله

Ce travail a été réalisé au Laboratoire d'Automatique et Signaux d'Annaba (LASA) du Département de L'électronique, Université Badji Mokhtar.

Mes premiers remerciements vont au **Dr. Arbaoui Fayçal**, qui m'a encadrée quotidiennement. Merci d'avoir eu confiance en mes capacités en me proposant ce sujet ô combien intéressant, ouvrant beaucoup de perspectives et durant lequel j'ai acquis de nombreuses compétences. Merci pour votre soutien, pour votre disponibilité, pour la pertinence de vos remarques et votre expérience... Ceux-ci ont beaucoup contribué à la réalisation de ce travail.

Je veux remercier les membres du jury :

Saidi Med Larbi, Professeur de l'université de Badji Mokhtar Annaba.

Saadi Med Nacer, Professeur de l'université de Badji Mokhtar Annaba.

Pour avoir pris le temps de soigneusement étudier ce manuscrit et qui m'ont fait l'honneur d'évaluer ce travail.

Enfin, je remercie mon amie **BOUMENTEN INES** qui a toujours été là pour moi. Ce sont son soutien inconditionnel et ses encouragements qui m'ont été d'une grande aide.

ملخص

العنوان: تصنيف عيوب المدرجة الكروية للآلة غير المتزامنة باستعمال تحليل المكونات الرئيسية وفاصل الهوامش الكبير

الخلل الناتج عن العيوب مثل تآكل أو تلف في المدرجة الكروية هو أحد الأسباب الرئيسية لسوء الأداء في الآلات غير المتزامنة، الذي يؤدي إلى خسائر مادية معتبرة.

الغرض من هذه المذكرة هو التشخيص الآلي لعيوب المدرجة الكروية على أساس التحليل الاهتزازي.

الشعاع المكون من المعلمات (الخصائص أو المؤشرات)، أو ما يسمى الشعاع الشكلي، تستخرج مؤشرات الإحصائية والطيفية لنتائج أمثل من الإشارات الاهتزازية الأصلية ومن الإشارات الاهتزازية التي تم معالجتها باستخدام تحليل النموذجي التجريبي.

أجريت الدراسة باستعمال قاعدة بيانات جامعة: *Case Western Reserve University* ، حيث تم الحصول على الإشارات الاهتزازية في أوضاع وظروف تشغيل مختلفة (في الحالة السليمة وفي حالة وجود العيب، تغيير في السرعة، تغيير في الحمولة وتغيير في شدة العيب).

علاوة على ذلك تمت إجراء دراسة مقارنة مع نوع آخر من المصنفات: الشبكة العصبية الاصطناعية. وقد أثبتت الطريقة المنتقاة أي فاصل الهوامش الكبير متعدد الأصناف، فعالية أفضل من حيث وقت ومعدل تصنيف بالمقارنة.

في مرحلة التصنيف، تم الاحتفاظ بفاصل الهوامش الكبير متعدد الأصناف.

الكلمات الدالة: فاصل الهوامش الكبيرة، المعلمات الإحصائية، التحليل النموذجي التجريبي، إشارة اهتزازية، عيوب المدرجة الكروية.

Résumé

Titre : Multi-Classification des Défaits de Roulements d'une Machine Asynchrone par Combinaison de la décomposition des signaux par la méthode EMD et des Séparateurs à Vaste Marges

Le défaut de roulement est l'une des causes dominantes de dysfonctionnements dans les machines asynchrones, entraînant une perte économique significative.

L'objectif de ce mémoire est le diagnostic automatisé des défauts de roulements basé sur la décomposition empirique modale.

Un vecteur de paramètres (Indicateurs ou caractéristiques), appelé aussi vecteur forme, est extrait de chacun des signaux de vibration bruts et spectres, il se compose d'indicateurs statistiques et spectraux pour une performance optimale dans un premier temps, ensuite les paramètres sont extraits des signaux traités en utilisant la décomposition des signaux par la méthode EMD.

L'étude a été menée sur la base de données de la Case Western Reserve University, où les signaux ont été acquis dans différents modes et conditions de fonctionnement (sain et avec défauts, variations de vitesse, variation de charges et variation de la sévérité du défaut).

Dans la phase de classification, les séparateurs à vaste marges (SVM) multi-classes ont été retenus.

Ensuite une étude comparative a été effectuée avec les classificateurs à base Réseaux de Neurones Artificiels (RNA). On a constaté une meilleure efficacité en termes de temps et taux de classification du multi-SVM par rapport aux RNA.

Mots clés : Machines à Vecteurs de Supports (SVM), Paramètres statistiques, décomposition empirique modale (EMD), défaut roulement, signal vibratoire.

Abstract

Title: Induction Machine bearings defects Multi-Classification by Empirical Mode Decomposition and Support Vector Machines

The rolling defect is one of the main causes of dysfunctions in induction machines, resulting in a significant economic loss.

The main purpose of this study is an automated diagnosis of bearing defects based on Empirical Mode Decomposition.

A vector of parameters (Indicators or characteristics), also known as vector form and composed of statistical and spectral indicators for optimal performance, is extracted from the original vibration signals and from the signals processed using Empirical Mode Decomposition EMD .

The study was carried out on the *Case Western Reserve University* database, where the signals were acquired in different modes and operating conditions (healthy and with defaults, speed variations, load variation and variation in defect severity).

In the classification phase, the Multi-Class Support Vector Machine (SVM) was chosen.

A comparison of the results of the temporal domain with and without Empirical Mode Decomposition was performed; the results obtained show the effectiveness of this approach.

As a result, a comparative study was achieved on another type of classifiers: Artificial Neural Network (ANN). A better efficiency in terms of time and classification rate goes to the multi-class SVM compared to the ANN.

Key words: Support Vector Machines (SVM), Statistical parameters, Empirical Mode Decomposition (EMD) rolling failure, vibration signal.

Liste des Figures

Chapitre 1	
1.1 Constitution d'un roulement	14
1.2 Roulement à contact oblique.....	15
1.3 (a) Ecaillage d'une bague intérieure, (b) Criques corrosion.....	17
1.4 Photos représentant les différents défauts de roulements. (a) Corrosion, (b) : Cratère bille, (c) : Cratère bague intérieure, (d) : Cratère bague extérieure, (e) : mécanismes généraux d'usure en présence de contaminant	18
1.5 Spectre théorique d'un défaut	18
1.6 Fréquence caractéristique d'un défaut de type écaillage affectant la bague extérieure	19
Chapitre 2	
2.1 Techniques d'extraction des caractéristiques (W. Yan, 2008).....	24
Chapitre 3	
3.1 Séparation de deux ensembles de points par un Hyperplan H.....	36
3.2 Vecteurs de support	37
3.3 Hyperplan optimal, marge et vecteurs de support.....	37
3.4 Maximisation de la marge	38
3.5 Illustration du choix de la marge avec un hyperplan optimal	38
3.6 Schéma de changement des représentations non linéaires vers un nouvel espace.....	41
3.7 Classification SVM un contre tous (One Vs All)	42
3.8 Schéma des hyperplans linéaires pour la classification un contre tous (One vs all) / n=4	42
3.9 Classification SVM un contre un (One Vs One).....	43

Chapitre 4	
4.1 Le Banc d'essai de la CWRU.....	47
4.2 Signal original cas sain.....	48
4.3 Autocorrélation du signal original cas sain.....	48
4.4 Variation des paramètres temporels et fréquentielles du signal sans défauts (sain).....	48
4.5 Signal original avec défaut bille.....	49
4.6 Autocorrélation du signal avec défaut bille.....	49
4.7 Variation des paramètres temporels et fréquentielles du signal avec défaut bille.....	49
4.8 Signal original avec défaut bague interne.....	50
4.9 Autocorrélation du signal avec défaut bague interne.....	50
4.10 Variation des paramètres temporels et fréquentielles du signal avec défaut bague interne.....	50
4.11 Signal original avec défaut bague externe.....	51
4.12 Autocorrélation du signal avec défaut bague externe.....	51
4.13 Variation des paramètres temporels et fréquentielles du signal avec défaut bague externe.....	51
4.14 Extraction des 3IMFs les plus pertinents cas sain.....	52
4.15 Spectre du signal cas sain.....	52
4.16 Extraction des 3 IMFs les plus pertinentes du signal avec défaut bille.....	53
4.17 Spectre du signal avec défaut bille.....	53
4.18 Extraction des 3 IMFs les plus pertinentes du signal avec défaut Bague interne.....	54
4.19 Spectre du signal avec défaut bague interne.....	54
4.20 Extraction des IMFs les plus pertinentes du signal avec défaut bague externe.....	55
4.21 Spectre du signal avec défaut bague externe.....	55

Table de matière

Dédicaces.....	2
Remerciements.....	3
Résumé / Abstract / ملخص.....	4
Liste des figures.....	Error!
Bookmark not defined.	
Introduction générale.....	13
<i>Chapitre 1 Défauts de Roulements</i>	
1.1. Introduction.....	14
1.2. Constitution d'un roulement	14
1.3. Fréquences caractéristiques d'un roulement.....	15
1.4. Types de défauts d'enroulements.....	16
1.4.1. Les défauts de montages	16
1.4.2. Les défauts de fabrication.....	17
1.5. Fréquences caractéristiques et défauts.....	17
Références Bibliographiques.....	20
<i>Chapitre 2 : Analyse Vibratoire et Décomposition Empirique Modale</i>	
2.1. Analyse Vibratoire.....	22
2.1.1. Introduction.....	22
2.1.2. Définition d'une vibration	22
2.1.3. Les Méthodes d'analyse Vibratoire	23
2.1.4. Extraction des caractéristiques.....	23
2.1.5. Sélection et réduction des caractéristiques.....	24
2.2. Analyse vibratoire par la transformée de Hilbert Huang.....	27
2.2.1. Principe de la décomposition	27
2.2.2. Critère d'arrêt du processus de tamisage.....	28
2.2.3. Les différents Types d'EMD	29
2.2.4. Bruit et Valeurs aberrantes.....	29
2.2.5. Les nouvelles approches pour la sélection des IMFs	31
Conclusion	32
References Bibliographies.....	32
<i>Chapitre 3 Multi classification par les Séparateurs à Vaste Marge</i>	

3.1. Introduction.....	35
3.2. Historique.....	35
3.3. Application des SVM à la classification binaire	35
3.3.1. Définition.....	35
3.3.2. Notion d'apprentissage.....	36
3.3.3. Principe de fonctionnement général et notion de base.....	36
3.3.4. Maximisation de la marge	37
3.3.5. Séparateur à vaste marge dans le cas linéaire.....	38
3.3.6. Séparateur à vaste marge dans le cas non linéaire.....	40
3.3.7. SVM multi-classes à l'origine.....	41
Conclusion.....	44
Références Bibliographiques.....	45
<i>Chapitre 4 Résultats et Discussion</i>	
4.1. Matériels et méthodes	47
4.1.1. Présentation du Banc d'Essai de la CWRU	47
4.2. Discussion des résultats des classifications.....	58
Conclusion Générale.....	62

Liste des tableaux	
Tableau 4.1 Données vibratoires des défauts avec une fréquence d'échantillonnage de 12K HZ / CWRU.....	49
Tableau 4.2 Récapitulatif des résultats	58

Introduction Générale

Introduction générale

Les machines industrielles automatisées sont composées de divers sous-systèmes, tels que le système de commande électrique, unités de commande(API), les actionneurs (vérins, moteurs asynchrones, distributeurs) ces composants sont impliqués dans l'exécution des opérations désirées des machines.

Les composants structurels d'une machine tels que les roulements à billes, sont soumis à l'usure au fil du temps, la surveillance de ces composants est donc nécessaire pour réduire régulièrement le risque de défaillances et de pannes par un diagnostic anticipatif ou précoce des défauts et assurer la disponibilité et l'utilisation efficace et optimale des machines qui sont les facteurs clés de l'économie de l'entreprise manufacturière.

L'objectif principal de ce mémoire est de mettre en œuvre un algorithme de surveillance de l'état des roulements à l'aide des machines à vecteurs de support.

Le succès de cette méthode est justifié par les solides bases théoriques qui la soutiennent. Il existe en effet un lien direct entre la théorie de l'apprentissage statistique et l'algorithme d'apprentissage du SVM.

La méthode SVM est une technique de classification particulièrement bien adaptée pour la reconnaissance des formes et traite des données très diversifiées ayant plusieurs dimensions telles que les textes, les images, la voix, les diagnostics médicaux, la bio-informatique pour la reconnaissance de gènes ...etc. Le champ d'application des SVM est large et représente une méthode de classification intéressante.

Ce mémoire comporte principalement 4 chapitres. Le premier est consacré à la présentation des différents types de défauts des roulements ainsi que l'analyse vibratoire. Le deuxième chapitre, introduit la théorie de la décomposition des signaux par la méthode EMD et son utilisation en diagnostic. Dans le troisième chapitre, nous détaillons la classification supervisée par les séparateurs à vaste marges.

Le dernier chapitre présente l'application expérimentale du système de diagnostic sur les données du banc d'essai de la CWRU avec deux techniques de classification : les réseaux de neurones artificiels (RNA) et les séparateurs à vaste marge (SVM). Une comparaison des performances est donnée et les résultats obtenus sont interprétés et discutés.

On achèvera, enfin par une conclusion générale et des perspectives.

Chapitre

1

Défauts de Roulements

1.1 Introduction

Les roulements sont des éléments clés de toute machines tournantes, ils sont souvent sujets à des dégradations qui peuvent provoquer l'arrêt du processus de production. Dans certains cas ils peuvent causer des accidents de fonctionnement graves. Les vibrations de roulements génèrent souvent des signaux relativement complexes qui varient en fonction de plusieurs facteurs : la charge, la présence de défauts et la géométrie du roulement. Quand des défauts ponctuels surviennent, des chocs sont générés et créent des vibrations qui peuvent être soit analysées dans le domaine temporel (niveau crête, efficace, facteur de crête, Kurtosis, Cepstre) ou bien dans le domaine fréquentiel (analyse spectrale ou d'enveloppe, transformées de Hilbert, etc.)

1.2 Constitution d'un roulement :

Un roulement est composé de 4 éléments principaux :

- 1) la bague extérieure, à monter dans un logement qui est, soit fixe, soit tournant.
- 2) la bague intérieure, à monter sur un arbre tournant, ou fixe.
- 3) les éléments roulants (billes, rouleaux ou aiguilles) qui permettent la rotation d'une bague par rapport à l'autre.
- 4) la cage qui maintient en place les billes, les rouleaux ou les aiguilles et les empêche de frotter l'un contre l'autre.

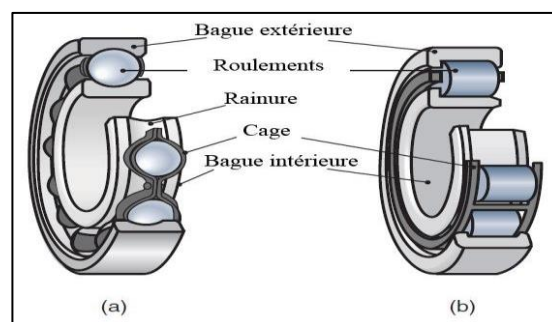


Figure 1.1 : Constitution d'un roulement

1.3 Fréquences caractéristiques d'un roulement :

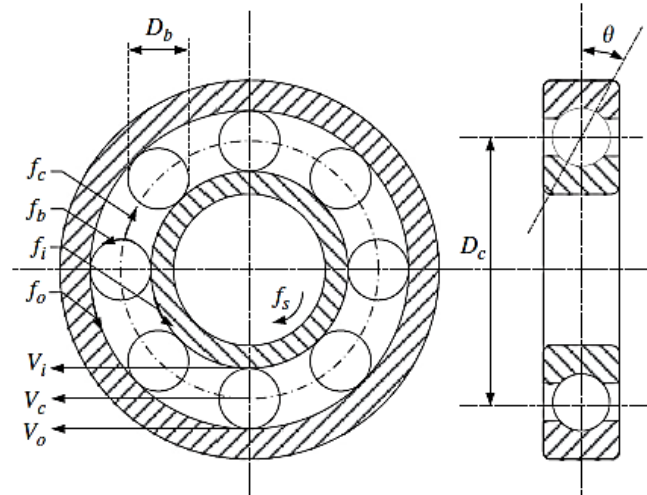


Figure 1.2 : Roulement à contact oblique

- N Le nombre de billes,
- D_b le diamètre de bille
- D_p le diamètre moyen du roulement
- θ l'angle de contact

$$f_{ou}(HZ) = f_0 \left(\frac{N}{2} \right) \left(1 - \frac{D_b}{D_p} \cos \theta \right)$$

$$f_i(HZ) = f_0 \left(\frac{N}{2} \right) \left(1 + \frac{D_b}{D_p} \cos \theta \right)$$

$$f_c(HZ) = f_0 \left(\frac{1}{2} \right) \left(1 - \frac{D_b}{D_p} \cos \theta \right)$$

$$f_b(HZ) = f_0 \left(\frac{D_b}{2D_p} \right) \left(1 - \frac{D_b}{D_p} \cos \theta \right)$$

(1.1)

f_b (Ball Frequency [BF]) : La fréquence de passage d'un défaut de bille ;

f_{ou} (Ball Passing Frequency Outer Race [BPFO]) : La fréquence de passage d'un élément roulant sur un défaut de bague extérieure ;

f_c (Cage Frequency [CF]) : La fréquence de passage d'un défaut de cage ;

f_0 La fréquence de rotation mécanique.

Il existe des exceptions telles que : roulements sans cage, ou sans bague intérieure, ou sans bague extérieure, ou simple cage à billes, à rouleaux, sans les deux bagues, etc.

1.4 Types de défauts d'enroulements

Les incidents de roulements se classent en deux catégories principales suivantes qu'ils sont dus à des défauts de fabrication, des erreurs de montage ou une lubrification insuffisante.

1.4.1 Les défauts de montages

— La propreté :

Toute intrusion de particules étrangères entraîne des détériorations qui évoluent inéluctablement vers l'écaillage.

— Les chocs :

Les roulements sont très sensibles au choc. Une bague mal engagée sur l'arbre ne doit pas recevoir de coups pour sortir. Un rapide passage au séchoir permettra de la sortir aisément.

— La géométrie :

Il faut éviter des ovalisations, les conicités, les désalignements, les jeux trop faibles ou trop forts, etc.

— **La vaseline de protection :**

L'épaisse couche de vaseline qui couvre certains roulements doit être éliminée. Selon leurs applications, et la criticité des machines sur lesquelles ils sont montés, les roulements nécessitent plus ou moins une surveillance et un entretien préventif.

1.4.2. Les défauts de fabrication :

— **Écaillage :**

Se caractérise par des traces de fissuration et d'arrachement de fragments de matière.

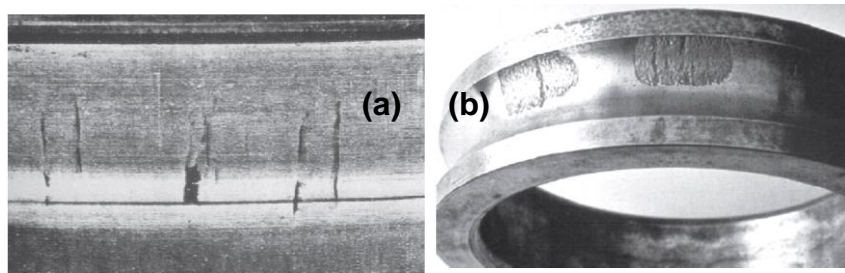


Figure 1.3 : (a) Ecaillage d'une bague intérieure, (b) Criques de corrosion

— **Grippage :**

Des traces brunes dues à l'échauffement, des déformations de corps roulants, et des micro-fusions peuvent apparaître.

— **Coups, fissures, cassures :**

Se manifestant avec des traces visibles d'outils, et des ruptures des bagues.

— Détérioration des cages :

Se manifestant sous plusieurs formes : déformation ; usure ou rupture des cages.

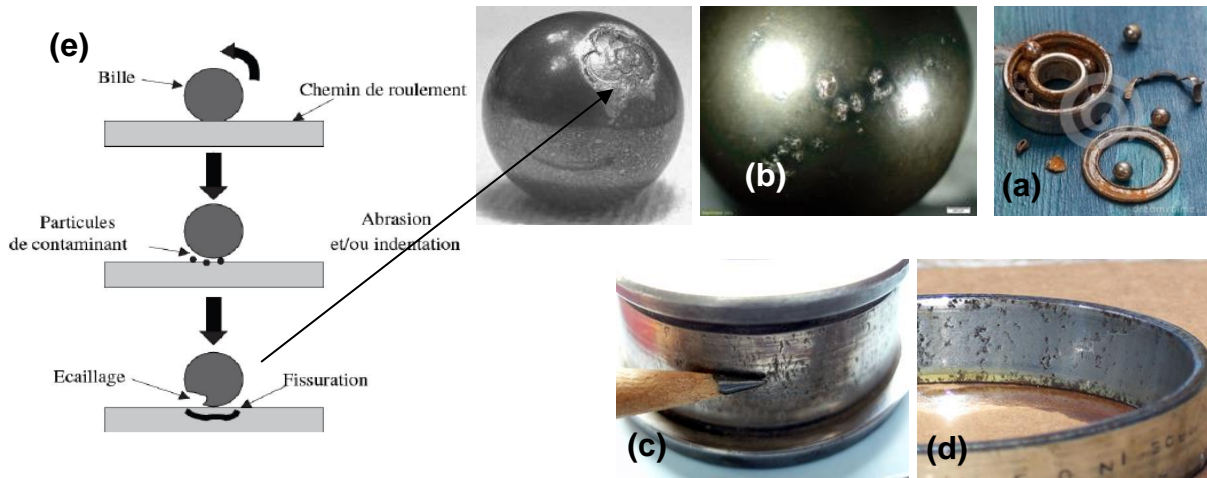


Figure 1.4 : Photos représentant les différents défauts de roulements. (a) Corrosion, (b) : Cratère bille, (c) : Cratère bague intérieure, (d) : Cratère bague extérieure, (e) : mécanismes généraux d'usure en présence de contaminant

2.5 Fréquences caractéristiques et défauts

La présence d'un défaut au niveau d'un roulement a un impact sur le spectre de son signal vibratoire tel qu'une augmentation de l'amplitude de la vibration aux fréquences du roulement. Aussi, il y a la présence de vibrations aux harmoniques de ces fréquences ainsi qu'à des fréquences de modulation et le spectre va faire apparaître un peigne de raie à la fréquence du défaut et des bandes latérales à la fréquence de rotation de l'arbre (F_0).

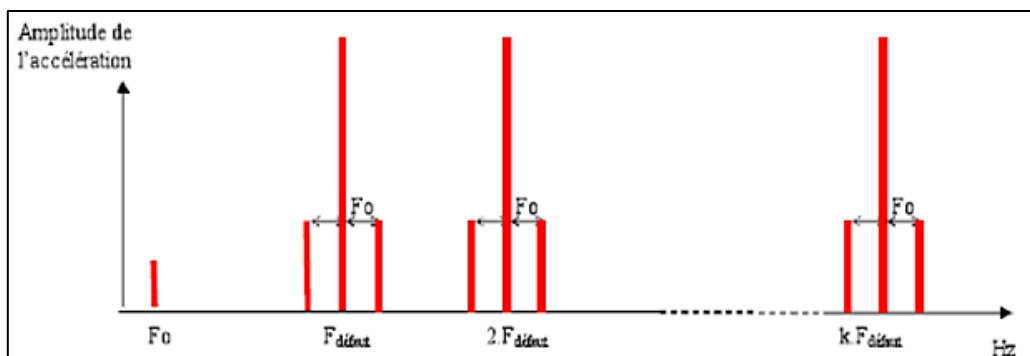


Figure 1.5 : Spectre théorique d'un défaut

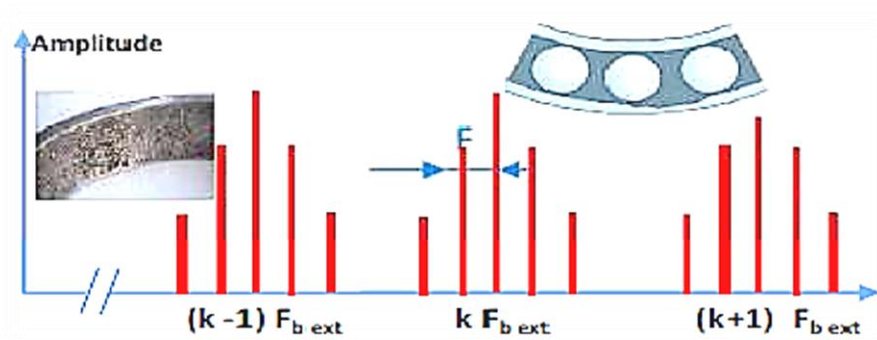


Figure 1.6 : Fréquence caractéristique d'un défaut de type écaillage affectant la bague externe

Différentes fréquences sont obtenues pour des défauts. Les quatre fréquences caractéristiques sont les suivantes :

- f_{be} est la fréquence de bague externe,
- f_{bi} est la fréquence de bague interne,
- f_c est la fréquence de cage,
- f_{bille} est la fréquence de bille,
- f_0 est la fréquence de rotation mécanique,

En cas de défaut se présentant sur une partie du roulement, des vibrations à la fréquence caractéristique associée sont engendrées.

Des harmoniques supplémentaires, issus de combinaisons entre les fréquences caractéristiques peuvent apparaître dans les spectres vibratoires.

Références Bibliographiques

Ilyes, KHELFI. (2014) *Diagnostic des machines tournantes par les techniques de l'intelligence artificielle*. Thèse de doctorat. Université de Badji Mokhtar Annaba.

Ishak T. (2012) *Extraction d'indicateurs robustes pour le diagnostic des défauts mécaniques : Comparaison de L'EMD et des ondelettes (WT)*. Thèse de doctorat. Université Ferhat Abbas de Sétif 1.

McInerny S.A. and Dai Y. (2003), Basic vibration signal processing for bearing fault detection, IEEE Transactions On Education, Vol.46, No.1.

Ming, Y., Chen, J., & Dong, G. (2011). Weak fault feature extraction of rolling bearing based on cyclic Wiener filter and envelope spectrum. *Mechanical Systems and Signal Processing*, 25(5), 1773-1785.

Morel, J. (1992). *Vibrations des machines et diagnostic de leur état mécanique*. Collection de la Direction des études et recherches d'électricité de France.

Chapitre

2

*Analyse Vibratoire et Décomposition
Empirique Modale*

2.1. Analyse Vibratoire

2.1.1. Introduction

Les vibrations des machines sont considérées comme la réponse de la fonction de transfert de leurs structures aux différents efforts auxquels elles sont sollicitées. La modification de la vibration d'une machine constitue souvent la première manifestation physique d'une anomalie, cause potentielle de dégradations, voire de pannes par effet de fatigue au sein de la matière ou de fretting entre deux surfaces accolées, entraîneront la fissuration puis la rupture et dans d'autres cas, les vibrations transporteront avec elles une énergie acoustiquement gênante (Tang et al, 2014).

La surveillance des vibrations de ces structures peut permettre donc de déceler l'apparition d'efforts perturbateurs à un stade précoce, la facilité relative de son installation rend l'analyse des signaux vibratoires une des techniques les plus utilisées pour la surveillance et le diagnostic des défauts des machines tournantes, cela dit il existe d'autres techniques de surveillance tel que :

- La thermographie infrarouge
- L'analyse des mesures ultrasonore

2.1.2. Définition d'une vibration :

La vibration d'une machine soumise à une force périodique peut être décrite en termes de déplacement, de vitesse ou d'accélération. La vitesse du mouvement vibratoire correspond à la variation de son déplacement pour une unité de temps. L'accélération représente une variation de la vitesse par unité de temps.

Le déplacement exprimé en [μm], la vitesse en [mm/s] et l'accélération en [m/s^2].

A ces grandeurs sont associées une période temporelle et une phase. Ainsi, dans le cas d'une machine tournante, une vibration élémentaire peut se mettre sous la forme :

$$x(t) = X. \sin(2\pi. f. t) \quad (2.1)$$

Une propriété fondamentale relie les trois représentations déplacement D , vitesse V et accélération a :

$$D = \frac{V}{2\pi f} = \frac{a}{(2\pi f)^2} \quad (2.2)$$

2.1.3 Les Méthodes d'analyse Vibratoire :

Les techniques du domaine temps-fréquence sont principalement utilisées pour traiter des signaux vibratoires non stationnaires. On y distingue la transformée de Fourier courte, la distribution de Wigner-Ville, l'analyse en ondelettes notamment la décomposition en paquets d'ondelettes (Wavelet Packet Decomposition : WPD), la décomposition modale empirique et la transformée de Hilbert-Huang. La transformée de Fourier courte segmente le signal en fenêtres temporelles et applique ensuite la transformée de Fourier sur les fenêtres obtenues, cependant, la taille des fenêtres est difficile à choisir.

Concernant la distribution de Wigner-Ville, elle conduit parfois à l'apparition de termes croisés, ce qui peut conduire à des analyses erronées.

La décomposition en paquets d'ondelettes (WPD) consiste à utiliser une paire de filtres passe-haut et passe-bas. Ces filtres permettent de réaliser une division du contenu fréquentiel du signal brut en deux composantes (ou paquets), l'une contient les hautes fréquences et l'autre les basses fréquences. Puis, ces paquets sont successivement injectés dans les filtres pour séparer de nouveau en deux l'espace fréquentiel. Cette opération permet d'atteindre des niveaux de décomposition assez fins qui permettent de suivre l'évolution des fréquences, tout en gardant une bonne résolution pour les phénomènes rapides et lents grâce à l'utilisation d'une ondelette mère et des paramètres d'échelle et de translation. La WPD peut être utilisée pour la détection et le suivi de dégradations, mais son efficacité dépend fortement de la qualité des signaux traités.

L'une des Méthodes la plus efficace pour l'analyse des signaux non stationnaires est la transformée de Hilbert-Huang. Elle s'appuie sur deux techniques : l'EMD et la transformée de Hilbert. L'EMD décompose le signal en fonctions modales intrinsèques (Intrinsic Mode Functions : IMFs) représentant les modes oscillatoires du signal. Ensuite, la transformée de Hilbert est appliquée sur chaque IMF pour extraire les fréquences instantanées et les amplitudes instantanées. La transformée de Hilbert-Huang permet ainsi une représentation tridimensionnelle (amplitude, fréquence, temps) de chaque mode oscillatoire du signal.

2.1.4 Extraction des caractéristiques :

Les paramètres statistiques quantifient les informations sur la forme et l'énergie du signal, elles peuvent refléter l'état de santé mécanique sous différents aspects et peuvent se compenser mutuellement à différents niveaux de défaut.

L'extraction des caractéristiques repose essentiellement sur des techniques issues du traitement du signal. Ces techniques peuvent être regroupées dans deux catégories principales selon que le signal à traiter est stationnaire ou non stationnaire. La figure ci-dessous donne un aperçu des techniques utilisées pour l'extraction des caractéristiques.

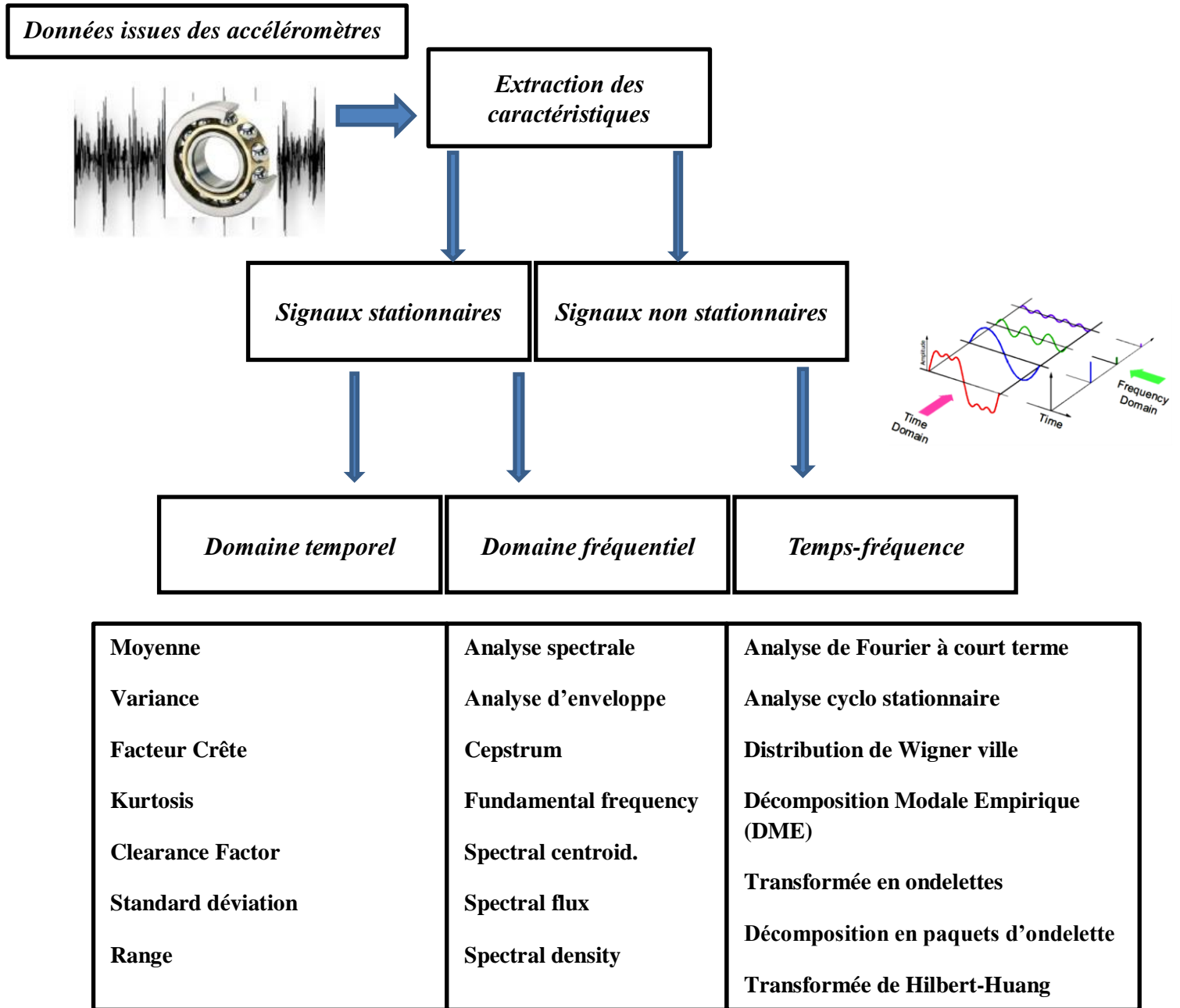


Figure2.1 : Techniques d'extraction des caractéristiques (W. Yan, 2008)

2.1.5 Sélection et réduction des caractéristiques :

➤ Caractéristiques ou paramètres temporelles

$$\text{Moyenne : } x_m = \sum_{n=1}^N \frac{x(n)}{N} \quad (2.3)$$

$$\text{Variance : } x_{\text{var}} = \frac{\sum_{n=1}^N (x(n) - x_m)^2}{N} \quad (2.4)$$

$$\text{Kurtosis } x_{\text{kur}} = \frac{\sum_{n=1}^N (x(n) - x_m)^4}{(N - 1)\sigma^4} \quad (2.5)$$

$$\text{Clearance factor : CLF} = \frac{x_{\text{max}}}{\left(\frac{1}{N} \sum_{n=1}^N \sqrt{|x(n)|}\right)^2} \quad (2.6)$$

$$\text{Mean absolute deviation : MAD} = \sum_{n=1}^N \frac{|\sum_{n=1}^N (x(n) - x_m)|}{N} \quad (2.7)$$

$$\text{Range : } x_{\text{range}} = |x_{\text{max}} - x_{\text{min}}| \quad (2.8)$$

$$\text{Standarddeviation : STD} = \sqrt{\frac{\sum_{n=1}^N (x(n) - x_m)^2}{N}} \quad (2.9)$$

$$\text{PeakmagnitudetoRMSratio: Crest factor} = \frac{x_{\text{max}}}{\text{RMS}} \quad (2.10)$$

➤ **Caractéristiques ou paramètres fréquentielles :**

L'analyse fréquentielle consiste à visualiser l'énergie du signal à différentes fréquences. En connaissant les conditions critiques de fonctionnement du composant (vitesse, charge, géométrie des éléments constitutifs, ...etc.). On peut extraire des caractéristiques permettant d'isoler les défauts. Ces caractéristiques sont aussi appelées signatures fréquentielles.

La fréquence moyenne :

$$\text{Mean Frequency : } f_{\text{mean}} = \frac{\sum_{n=1}^N I_n \cdot f_n}{\sum_{n=1}^N I_n} \quad (2.11)$$

Où f_n et I_n sont les fréquences et les amplitudes au point n du spectre du signal

$$\text{Mean power frequency: MNF} = C_s = \frac{\sum_{n=1}^N P_n \cdot f_n}{\sum_{n=1}^N P_n} \quad (2.12)$$

Où P_n est la puissance spectrale du signal au point n.

$$\text{Densité spectrale de puissance DSP} = \frac{\sum_{n=1}^N P_n}{N} \quad (2.13)$$

$$\text{Spectral Entropy } E_{n_s} = - \sum_{n=1}^N P_n \log P_n \quad (2.14)$$

$$\text{Short time Energy} = \sum_{n=1}^N [x[m] W[n - m]]^2 \quad (2.15)$$

2.2. Analyse vibratoire par la transformée de Hilbert Huang

La méthode EMD est définie par un processus appelé tamisage (sifting) permettant de décomposer un signal en contributions de base appelées modes empiriques ou IMF (Intrinsic mode functions) qui sont des signaux de type AM-FM mono-composante (au sens large), chacune de moyenne nulle. L'extraction des IMFs est non-linéaire, mais leur recombinaison pour la reconstruction exacte du signal est linéaire. En se basant essentiellement sur les variations ou oscillations naturelles du signal, l'EMD permet une interprétation des phénomènes physiques présents.

En plus de sa simplicité de mise en œuvre et de sa capacité à décrire ponctuellement et de manière instantanée les phénomènes fréquentiels non résolus par l'analyse de Fourier, l'EMD est bien adaptée à l'étude des signaux non-stationnaires et/ou générés par des systèmes non-linéaires.

2.2.1. Principe de la décomposition

Le principe de l'EMD repose sur une décomposition adaptée en décrivant localement le signal comme une succession de contributions d'oscillations rapides (hautes fréquences) sur des oscillations plus lentes (basses fréquences). Pour illustrer la décomposition, nous considérons l'exemple simple d'un signal formé de deux contributions : une oscillation rapide et une autre lente :

$$x(t) = d(t) + m(t) \quad (2.16)$$

$x(t)$: *signal*

$d(t)$: *oscillation (a)*

$m(t)$: *tendance (b)*

L'oscillation (a) est une oscillation rapide et la tendance (b) est une oscillation lente. La reconstruction du signal $s(t)$ est réalisée en sommant les deux courbes point par point. La tendance du signal contient localement des oscillations de plus basse fréquence que celle de l'oscillation rapide. Ce résultat peut être généralisé à toutes les contributions comme suit : chacune contient localement des oscillations de plus basse fréquence que celle extraite précédemment. Conceptuellement, la mise en place de la méthode permettant de décrire cette décomposition est assez simple. Il suffit de considérer le signal à l'échelle de ses oscillations

locales, de soustraire l'oscillation la plus rapide du signal et de réitérer le processus sur le résidu de manière à écrire le signal $x(t)$ comme une combinaison finie d'oscillations.

Le signal $x(t)$ peut alors s'écrire :

$$x(t) = \sum_{j=1}^N IMF_j(t) + r(t) \quad (2.17)$$

où j est la j ème oscillation, $r(t)$ est le résidu de la décomposition et N est le nombre d'IMFs. Même si la notion d'IMFs repose plus sur une intuition que sur une définition très rigoureuse, cette fonction doit respecter certains critères comme précisé ci-dessous.

Une IMF (ou mode empirique) est une fonction telle que :

$$IMF \left\{ \begin{array}{l} R \rightarrow R \\ t \rightarrow IMF_j(t) \end{array} \right\} \quad (2.18)$$

Qui vérifie les conditions suivantes :

- a. Elle est de moyenne nulle.
- b. Les nombres d'extrema et de passages à zéro diffèrent au plus de un (en d'autres termes, cela signifie qu'entre un minimum et un maximum successif, une IMF passe par zéro).
- c. Elle suit une loi de modulation en amplitude et en fréquence (comportement oscillant) naturellement de type mono-composante.

2.2.2. Critère d'arrêt du processus de tamisage

Comme indiqué précédemment, l'EMD est composée de deux boucles : l'une permet d'extraire l'IMF (qui détermine le niveau de profondeur de la décomposition) et l'autre conditionne la fonction IMF $j(t)$ de manière à valider la définition d'une IMF.

La boucle de tamisage qui permet d'extraire des IMF correctement conditionnées est pilotée par un critère note $SD(i)$ de manière générale, les définitions du critère d'arrêt du tamisage reposent toutes sur le même principe : $h_{j,i}(t)$ est de moyenne nulle et symétrique.

Ainsi dans ; les auteurs proposent un critère d'arrêt $SD(i)$ reposant sur la déviation standard et est défini par :

$$SD(i) = \sum_{t=0}^T \frac{|h_{i-1}(t) - h_i(t)|^2}{[h_{i-1}(t)]^2} \quad (2.19)$$

2.2.3. Les différents Types d'EMD :

La conception d'un système de diagnostic performant n'est pas une tâche facile à réaliser, puisque la majorité des systèmes réels sont de nature non-linéaire et non-stationnaire souvent corrélées et bruitées.

Les performances du système de diagnostic se basent sur La robustesse de la méthode face à la présence de bruit, perturbation et erreur de modélisation. La manipulation du prétraitement des données peut améliorer les performances du diagnostic. de nouvelles méthodes offrent la possibilité de réduire le bruit et d'extraire seulement l'information pertinente.

L'EEMD :

Est une méthode d'analyse de données bruitées, Noise Assisted Data Analysis (NADA).

L'EEMD consiste à "tamiser" un bruit blanc ajouté au signal.

Le bruit blanc est calculé en moyenne avec un nombre suffisant d'essais, la seule partie persistante qui survit au processus de calcul de la moyenne est le signal, qui est ensuite traité comme la vraie et plus significative réponse physique. En outre, L'EEMD représente une amélioration substantielle par rapport à l'EMD d'origine.

CEEMDAN :

Complete Ensemble EMD with Adaptive Noise est une variation de l'algorithme EEMD qui fournit une reconstruction exacte du signal d'origine et une meilleure séparation spectrale des IMFs avec moins de bruits et plus d'information physiques.

2.2.4. Bruit et Valeurs aberrantes :

En pratique un signal que l'on cherche à mesurer est toujours accompagné de signaux perturbateurs, que l'on peut appeler des bruits :

- Le bruit électronique issu des composants électroniques peut être considéré comme blanc et peu énergétique. On appelle bruit blanc un signal aléatoire dont l'énergie spectrale est la même pour toutes les fréquences.
- Le bruit de quantification résultant de la conversion des données analogiques en données numériques.

Le rapport signal sur bruit est défini par :

$$\text{SNR} = 10 \log \left(\frac{\sum_{i=1}^n y^2(i)}{\sum_{i=1}^n e^2(i)} \right) \quad (2.20)$$

$$\text{FNR} = 10 \log_{10} \frac{\sigma_f^2}{\sigma_v^2} \quad (2.21)$$

Les signaux peuvent également contenir des données erronées ou fausses. La manipulation et la détection des valeurs aberrantes est nécessaire à toute modélisation car leurs présences détériorent l'élaboration d'un bon modèle.

- Parmi les techniques de détection des valeurs aberrantes dans les bases de données on cite la distance de Mahalanobis et l'analyse de résidu de la régression PLS avec les réseaux de neurones dans les travaux publiés par Hugo and Sbarbaro en 2011.

2.2.5. Les nouvelles approches pour la sélection des IMFs :

L'EMD est utilisé comme outil de prétraitement qui est particulièrement adaptée pour traiter des signaux non stationnaires.

Plusieurs IMFS sont obtenues dont on choisit les plus significatives pour extraire l'enveloppe qui contient l'information requise.

- Sélection des IMFs par la densité spectrale de Hilbert, les IMFs retenues sont celles qui maximisent la densité spectrale de Hilbert autour des fréquences caractéristiques des défauts. Pour chaque fréquence caractéristique, on regarde parmi toutes les IMFs celle qui maximise la densité spectrale de Hilbert sur cette fréquence et on la garde. On obtient ainsi une seule IMF pour une fréquence donnée. Cela permet d'avoir au minimum une IMF (la même pour toutes les fréquences caractéristiques) et au maximum trois IMFs différentes (une pour chaque fréquence). La valeur de chaque IMF retenue est prise

comme valeur de l'indicateur de santé, ce qui conduit au total à trois indicateurs (un par fréquence caractéristique) (Medjaher, K. 2014).

- L'EMD est une procédure numérique sujette aux erreurs pouvant persister dans la décomposition en IMFs en tant que supplément additionné.

Dans les travaux publiés par De Souza en 2014, un nouvel algorithme est proposé pour sélectionner les IMFs pertinents, basé sur l'approche énergie- bruit par comparaison avec la technique traditionnelle qui utilise la corrélation.

- Dans un d'autres travaux publiés par Md. Burhan Uddin en 2016 un algorithme de sélection adaptatif des IMFs qui exploite le Caractéristiques du seuillage basé sur l'énergie et du potentiel empirique de SVM pour la classification des IMFs sans bruit : les IMFs dominants avec bruit sélectionnés de manière adaptative sont désactivés à travers le filtre SGF et Les IMFs sans bruit sont ensuite reconstruits. Une conclusion ultérieure est tirée ; Les résultats obtenus par la comparaison du SNR (Signal-to-Noise Ratio), PRD et Max Error.

Conclusion

Dans ce chapitre, nous avons étudié le principe de la décomposition par la méthode EMD et notamment le processus de tamisage permettant d'imposer des contraintes sur les IMFs. Lors de cette étude des problématiques ont été identifiées comme par exemple l'interpolation de la décomposition.

La décomposition par la méthode EMD a été comparée à la transformée en ondelettes. La transformée en ondelettes apparait comme une méthode paramétrique tandis que l'EMD est plutôt une approche non-paramétrique. La transformée en ondelettes décompose un signal en partant des plus basses fréquences vers les plus hautes. Par contre la méthode EMD aborde la décomposition en débutant par les hautes fréquences pour terminer vers les plus basses. Les erreurs d'estimation des différentes composantes sont globalement plus importantes pour l'approche basée sur les ondelettes que celle basée sur l'EMD.

Enfin, nous avons examiné à l'aide d'un simple exemple de simulation l'inconvénient majeur de cette méthode qui est le mixage de modes.

Références Bibliographiques

A. Boashash, (2003): Time-Frequency Signal Analysis and Processing: A Comprehensive Reference. Elsevier Ltd., Oxford, UK.

B. Boashash, (1992): Time-Frequency Signal Analysis-Methods and Applications. Longman-Cheshire, Melbourne and John Wiley Halsted Press, New York.

De Souza, D. B., Chanussot, J., & Favre, A. C. (2014, May). On selecting relevant intrinsic mode functions in empirical mode decomposition: An energy-based approach. In *Acoustics, Speech and Signal Processing (ICASSP), 2014 IEEE International Conference on* (pp. 325-329). IEEE.

Garces, H., & Sbarbaro, D. (2011). Outliers detection in environmental monitoring databases. *Engineering Applications of Artificial Intelligence*, 24(2), 341-349.

Guyon, I., & Elisseeff, A. (2003). An introduction to variable and feature selection. *Journal of machine learning research*, 3(Mar), 1157-1182.

Jean-Christophe Cexus, (2005) Analyse des signaux non-stationnaires par transformation de Huang, Operateur de Teager-Kaiser, et Transformation de Huang-Teager (THT). Acoustics. Université Rennes 1. French. <Tel-00012141>.

Lin, S. W., Lee, Z. J., Chen, S. C., & Tseng, T. Y. (2008). Parameter determination of support vector machine and feature selection using simulated annealing approach. *Applied soft computing*, 8(4), 1505-1512.

M. G. Rilling, (2007). Décompositions Modales Empiriques Contribution à la théorie, l'algorithme et l'analyse de performances. Thèse de doctorat. Université de Lyon.

Management (PHM)-a Survey. Technical report, Air Force Research Laboratory.

Medjaher, K. (2014). *Contribution au pronostic de défaillances guidé par des données* (Doctoral dissertation, Université de Franche-Comté).

Morel, J. (1992). Vibrations des machines et diagnostic de leur état mécanique. *Collection de la Direction des Études et recherches de l'Électricité de France*.

N.E. Huang, Z. Shen, S.R. Long, M.C. Wu, H.H. Shih, Q. Zheng, N.C. Yen, C.C. Tung et H.H. Liu, (1998): The empirical mode decomposition and the Hilbert spectrum for Nonlinear and non-stationary time series analysis. *Proceedings of the Royal Society of London*, 454:903-995.

Oskoei, M. A., & Hu, H. (2008). Support vector machine-based classification scheme for myoelectric control applied to upper limb. *IEEE transactions on biomedical engineering*, 55(8), 1956-1965.

Samanta, B., Al-Balushi, K. R., & Al-Araimi, S. A. (2003). Artificial neural networks and support vector machines with genetic algorithm for bearing fault detection. *Engineering Applications of Artificial Intelligence*, 16(7), 657-665.

Tang, D., Wei, F., Qin, B., Liu, T., & Zhou, M. (2014, August). Coooolll: A deep learning system for twitter sentiment classification. In *Proceedings of the 8th International Workshop on Semantic Evaluation (SemEval 2014)* (pp. 208-212).

Uddin, M. B., Uddin, J., Sultana, R., & Islam, S. (2016, December). A new machine learning approach to select adaptive IMFs of EMD. In *Electrical, Computer & Telecommunication Engineering (ICECTE), International Conference on* (pp. 1-4). IEEE.

W. Yan, H. Qiu, and N. Iyer, (2008). Feature extraction for bearing prognostics and health

Chapitre

3

Multi classification par les Séparateurs à Vaste Marge

3.1. Introduction

Le classifieur SVM (Séparateurs à Vaste Marge, Machine à Vecteurs de Support, ou Support Vector Machines en anglais) est un algorithme qui a été développé par V.Vapnik. Initialement, il a été conçu pour traiter le problème de la classification binaire supervisé et très vite les SVM se sont étendus pour des problèmes de natures diverses. Parmi leurs avantages, nous citerons la possibilité de traiter des problèmes dont les exemples sont non homogènes, la convergence vers une solution unique (pas de problème de minimum local comme le cas des réseaux de neurones). De nombreux travaux ont été achevés dans le domaine de la reconnaissance de la parole en appliquant les SVMs et plusieurs noyaux sont développés pour cet objectif. Nous présentons dans ce chapitre une description des SVMs.

3.2. Historique

Les Séparateurs à Vaste Marge sont des nouvelles techniques discriminantes dans la théorie de l'apprentissage statistique. Elles ont été proposées en 1995 par V. Vapnik. Elles permettent d'aborder plusieurs problèmes divers et variés comme la régression, la classification, la fusion etc. En 1982, Vapnik propose un premier classificateur basé sur la minimisation du risque structurel baptisé SVM. Ce modèle était toutefois linéaire et l'on ne connaissait pas encore le moyen d'induire des frontières de décision non linéaires. En 1992, Boser et autres collaborateur proposent d'introduire des noyaux non-linéaires pour étendre le SVM au cas non-linéaire. En 1995, Cortes et al. proposent une version régularisée du SVM qui tolère les erreurs d'apprentissage tout en les pénalisant. Depuis, les SVMs n'ont cessé de susciter l'intérêt de plusieurs communautés de chercheurs de différents domaines d'expertise.

3.3. Application des SVM à la classification binaire

3.3.1. Définition

Cette méthode repose sur l'existence d'un classificateur linéaire dans un espace approprié. Puisque c'est un problème de classification à deux classes, cette méthode fait appel à un jeu de données d'apprentissage pour apprendre les paramètres du modèle. Elle est basée sur l'utilisation de fonction dites noyau (kernel) qui permettent une séparation optimale des données. Dans la présentation des principes de fonctionnements, nous schématiserons les données par des « points » dans un plan.

3.3.2. Notion d'apprentissage

L'apprentissage par induction permet d'arriver à des conclusions par l'examen d'exemples particuliers. Il se divise en apprentissage supervisé et non supervisé. Le cas qui concerne les SVM est l'apprentissage supervisé. Les exemples particuliers sont représentés par un ensemble de couples d'entrée/sortie. Le but est d'apprendre une fonction qui correspond aux

exemples vus et qui prédit les sorties pour les entrées qui n'ont pas encore été vues. Les entrées peuvent être des descriptions d'objets et les sorties la classe des objets donnés en entrée.

3.3.3. Principe de fonctionnement général et notion de base :

Hyperplan, marge et support vecteur

Le but des SVMs est de trouver un classificateur qui va séparer les données et maximiser la distance entre ces deux classes. Avec les SVMs, ce classificateur est linéaire appelé « hyperplan ». Dans le schéma qui suit, on détermine un hyperplan qui sépare les deux ensembles de points.

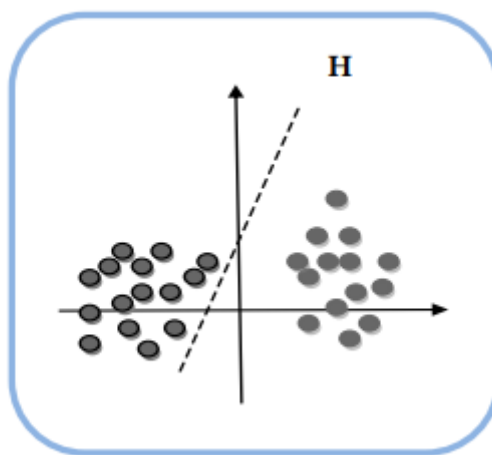


Figure 3.1 Séparation de deux ensembles de points par un Hyperplan H

Les points les plus proches, qui seuls sont utilisés pour la détermination de l'hyperplan, sont appelés vecteurs de support. Il est évident qu'il existe une multitude d'hyperplans valides mais la propriété remarquable des SVMs est que cet hyperplan doit être optimal. Nous allons donc en plus chercher parmi les hyperplans valides, celui qui passe « au milieu » des points des deux classes d'exemples. Cela revient à chercher un hyperplan dont la distance minimale aux exemples d'apprentissage est maximale. On appelle cette distance « marge » entre l'hyperplan et les exemples. L'hyperplan séparateur optimal est celui qui maximise la marge. Comme on cherche à maximiser cette marge, on parlera des séparateurs à vaste marge.

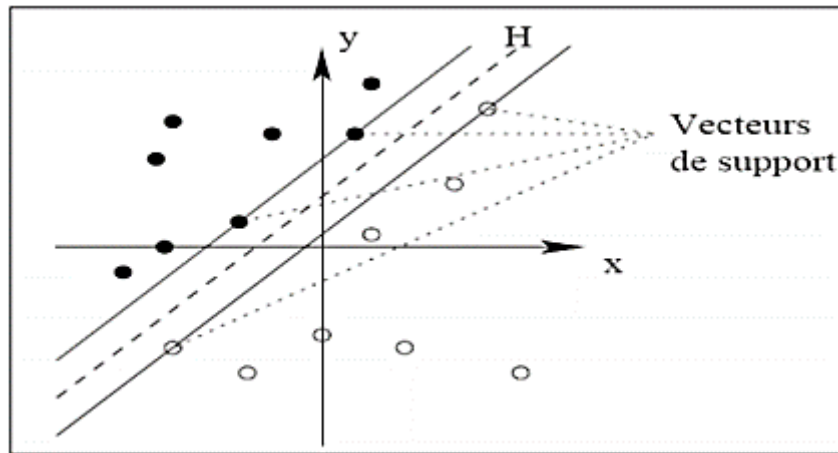


Figure 3.2 Vecteurs de support

Dans le schéma qui suit, on détermine un hyperplan qui sépare les deux ensembles de points.

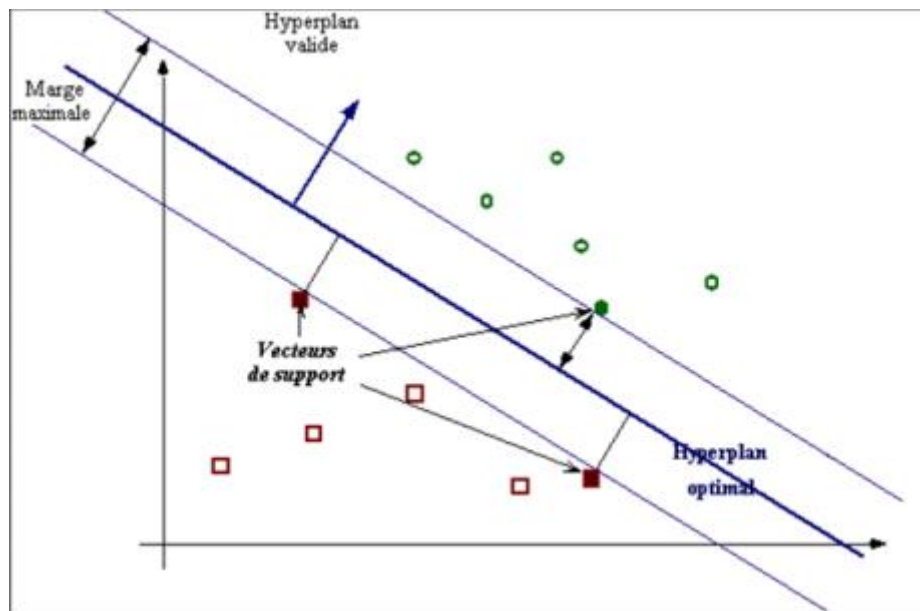


Figure 3.3 Hyperplan optimal, marge et vecteurs de support

3.3.4. Maximisation de la marge

Intuitivement, le fait d'avoir une marge plus large procure plus de sécurité lorsque l'on classe un nouvel exemple. De plus, si l'on trouve le classificateur qui se comporte le mieux vis-à-vis des données d'apprentissage, il est clair qu'il sera aussi celui qui permettra au mieux de classer les nouveaux exemples. Dans le schéma qui suit, la partie droite nous montre qu'avec un hyperplan optimal, un nouvel exemple reste bien classé alors qu'il tombe dans la marge.

On constate sur la partie gauche qu'avec une plus petite marge, l'exemple se voit mal classé.

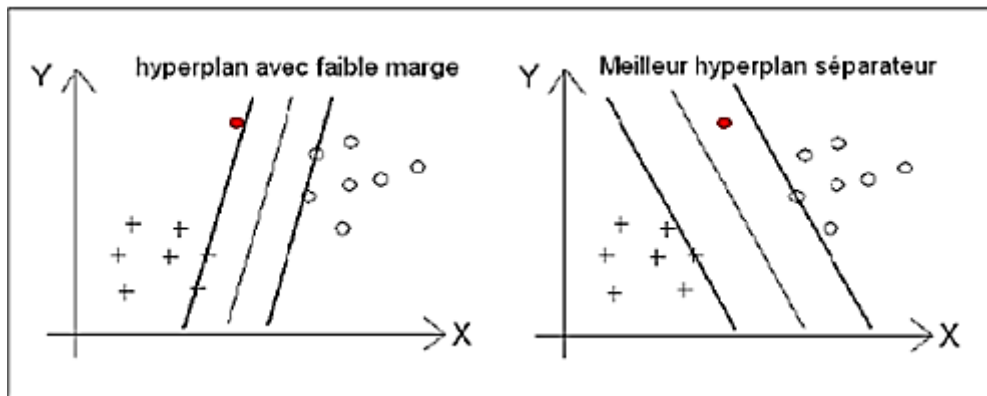


Figure 3.4 Maximisation de la marge

3.3.5. Séparateur à vaste marge dans le cas linéaire :

Le fait d'avoir une marge plus large procure plus de sécurité lorsque l'on classe un nouvel exemple. De plus, si l'on trouve le classificateur qui se comporte le mieux vis-à-vis des données d'apprentissage, il est clair qu'il sera aussi celui qui permettra au mieux de classer les nouveaux exemples. Dans le schéma qui suit, la partie (b) nous montre qu'avec un hyperplan optimal, un nouvel exemple reste bien classé alors qu'il tombe dans la marge. On constate sur la partie (a) qu'avec une plus petite marge, l'exemple se voit mal classé.

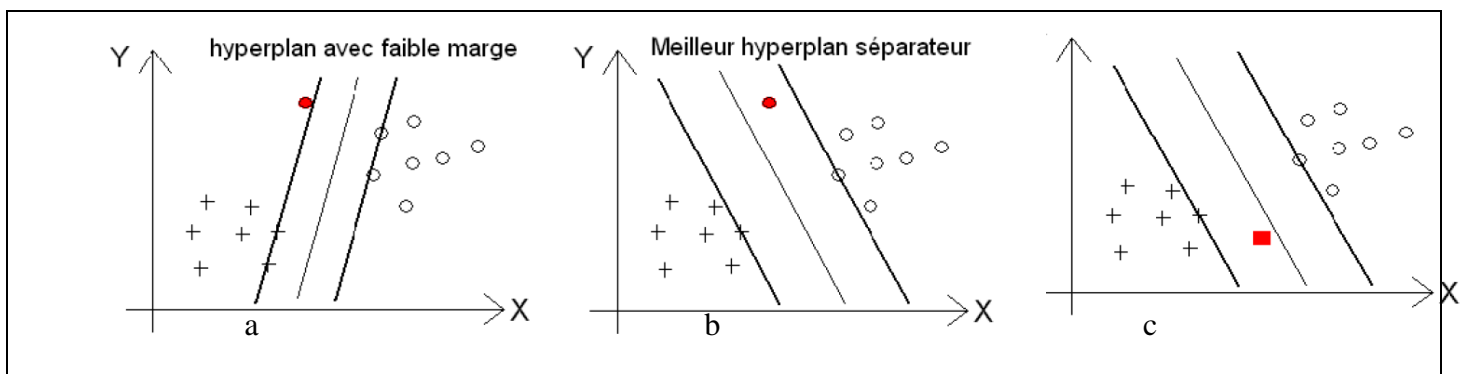


Figure 3.5 Illustration du choix de la marge avec un hyperplan optimal

En général, la classification d'un nouvel exemple inconnu est donnée par sa position par rapport à l'hyperplan optimal. Dans le schéma (c), le nouvel élément sera classé dans la catégorie des « + ».

Calcul de la marge :

Rappelons qu'en deux dimensions la distance entre un point (x_0, y_0) et une droite

$Ax + By + c = 0$, est donnée par la relation suivante :

$$\frac{|Ax_0 + By_0 + c|}{\sqrt{A^2 + B^2}} \quad (3.1)$$

De façon similaire, dans un modèle linéaire (voir figure (3.1)), la distance entre un point et l'hyperplan séparateur qui a pour équation $wx + b = 0$ est donnée par :

$$d(x) = \frac{|wx + b|}{\|w\|} \quad (3.2)$$

Calcul de l'Hyperplan optimal :

L'hyperplan optimal est celui pour lequel la distance aux points les plus proches (Marge) est maximale. L'équation générale de l'hyperplan est $H : wx + b = 0$ où :

- w est le vecteur orthogonal à l'hyperplan.
- x un vecteur de l'hyperplan.
- $b/\|w\|$ est la distance entre l'hyperplan et l'origine.

Cette équation doit satisfaire les conditions suivantes pour une classification binaire :

$$\begin{cases} \text{Si } y_i = +1 ; w x_i + b \geq 1 \\ \text{Si } y_i = -1 ; w x_i + b \leq -1 \\ \text{Pour tous } \{x_i, y_i\} y_i (w x_i + b) \geq 1 \end{cases} \quad (3.3)$$

Soient $f(x) = w \cdot x + b$ tel que x_1 et x_2 deux points de classes différentes :

$$(f(x_1) = +1 \text{ et } f(x_2) = -1).$$

$$(w \cdot x_1) + b = +1 \text{ et } (w \cdot x_2) + b = -1 \text{ donc } (w \cdot (x_1 - x_2)) = 2$$

$$\text{D'où : } (w \cdot (x_1 - x_2) / \|w\|) = 2 / \|w\|.$$

On peut donc en déduire que maximiser la marge revient à minimiser son inverse $\|w\|/2$ sous la contrainte $y_i(w \cdot x + b) \geq +1$.

$$\min \frac{1}{2} \|w\| = \min \frac{1}{2} w^T w \quad (3.4)$$

$$\text{On a } \begin{cases} \min \frac{1}{2} w^T w \\ \text{sous la contrainte} \\ y_i (w \cdot x + b) \geq +1 \end{cases} \quad (3.5)$$

Ce genre de problème d'optimisation peut être résolu en associant un multiplicateur de Lagrange $\alpha_i \geq 0$.

Le Lagrangien est donné par :

$$L(w, s, \alpha) = \frac{1}{2} w^t w - \sum_{i=1}^N \alpha_i y_i (w \cdot x_i + b) + \sum_{i=1}^N \alpha_i \quad (3.6)$$

Pour trouver la solution du problème primal, il faut minimiser le lagrangien par rapport à w et le maximiser par rapport aux α_i . Pour cette solution, les dérivées partielles du lagrangien s'annulent, selon les conditions de Kuhn-Tucker, on obtient :

$$\frac{\partial L(w,b,\alpha)}{\partial w} = 0, \frac{\partial L(w,b,\alpha)}{\partial b} = 0 \text{ et } \alpha_i \geq 0 \quad (3.7)$$

Cela revient à trouver les α_i et w qui annulent ses dérivées partielles :

On trouve :

$$w = \sum_{i=1}^N \alpha_i y_i x_i \text{ et } \sum_{i=1}^N \alpha_i y_i = 0 \quad (3.8)$$

Et en les réinjectant dans le Lagrangien on obtient le Lagrangien dual qui doit être maximisé

$$L(w, b, \alpha) = \sum_{i=1}^N \alpha_i - \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N \alpha_i \alpha_j y_i y_j x_i x_j \quad (3.9)$$

Si les α_i^* sont solutions de ce problème c.à.d. les multiplicateurs optimaux correspondant aux points les plus proches et sont non-nuls. Alors les points x_i sont appelés **vecteurs de supports**.

Et on a le w^* optimal :

$$w^* = \sum_{i=1}^n \alpha_i^* y_i x_i \quad (3.10)$$

En le remplaçant dans l'équation de l'hyperplan séparateur on obtient l'hyperplan optimal

$$h(x) = \sum_{i=1}^n \alpha_i^* y_i x_i \cdot x + b \quad (3.11)$$

La fonction de décision associée est donc :

$$f(x) = \text{sign}(\sum_{i=1}^n \alpha_i^* y_i x_i \cdot x + b) \quad (3.12)$$

3.3.6. Séparateur à vaste marge dans le cas non linéaire

Pour surmonter les inconvénients des cas non linéairement séparables, l'idée des SVMs est d'utiliser des fonctions noyaux et de projeter les données de l'espace d'entrée (non linéairement séparables) dans un espace de plus grande dimension, de façon à ce que les données deviennent linéairement séparables.

Plus la dimension du nouvel espace est grande, plus la probabilité de pouvoir trouver un hyperplan séparateur entre les classes est élevée, ceci est illustré dans la figure (3.3).

En pratique, il revient à l'utilisateur du SVM d'effectuer des tests pour déterminer la fonction noyau qui convient le mieux pour son application.

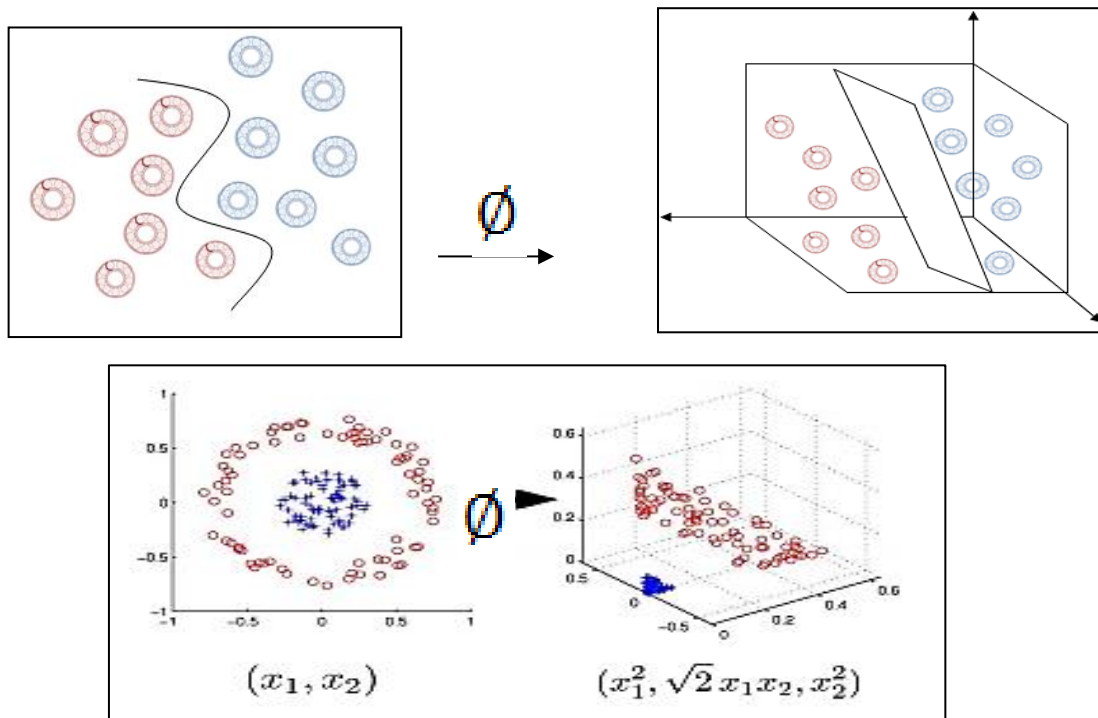


Figure 3. 6 Schéma de changement des représentations non linéaires vers un nouvel espace

3.3.7. SVM multi-classes à l'origine

Les SVM ont été conçus essentiellement pour les problèmes à 2 classes, cependant plusieurs approches permettant d'étendre cet algorithme aux cas à N classes ont été proposées. Pour pouvoir traiter plus de deux classes il convient d'apporter les modifications nécessaires. La généralisation dans le cas multi-classes peut se faire de deux façons différentes.

a. Un-Contre-Tous

L'approche la plus naturelle est d'utiliser cette méthode de discrimination binaire et d'apprendre N fonctions de décision, f_m ($m = 1, \dots, N$) permettant de faire la discrimination entre chaque classe de toutes les autres (chaque classe est opposée à toutes les autres). Il faut donc poser N problèmes binaires. Le k ième classificateur sépare les données de la classe k de tout le reste des données d'apprentissage. Chaque classificateur renvoie 1 si la forme à reconnaître appartient à la classe, -1 sinon. Alors pour reconnaître une forme il faut la soumettre à tous les autres classificateurs. Pour prendre une décision de classification, on garde la classe qui a eu la valeur maximale de toutes les fonctions de décision. Donc l'affectation d'un nouveau point x à une classe C_i se fait par la relation :

$$i = \arg \max_{m=1 \dots N} f_m(x) \quad (2.17)$$

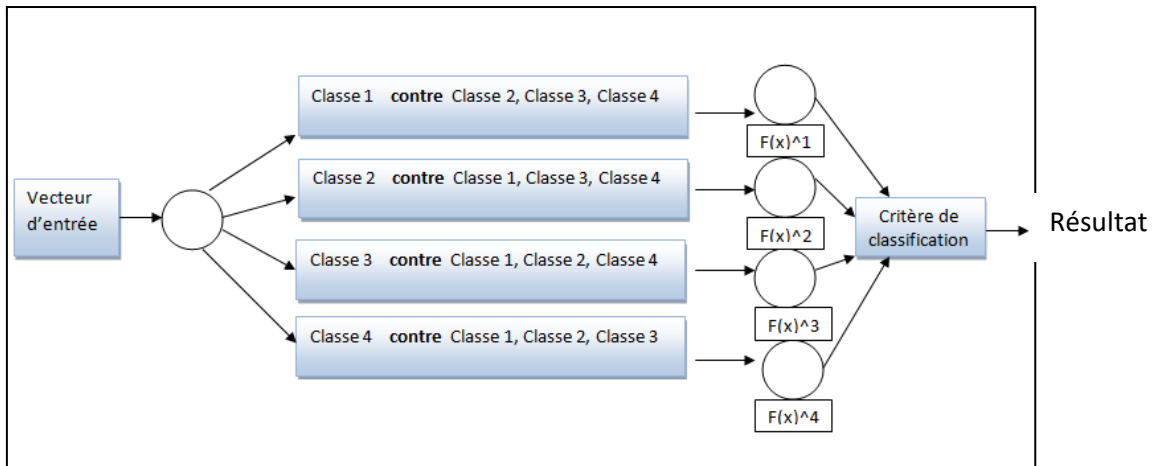


Figure 3.7 Classification SVM un contre tous (One Vs All)

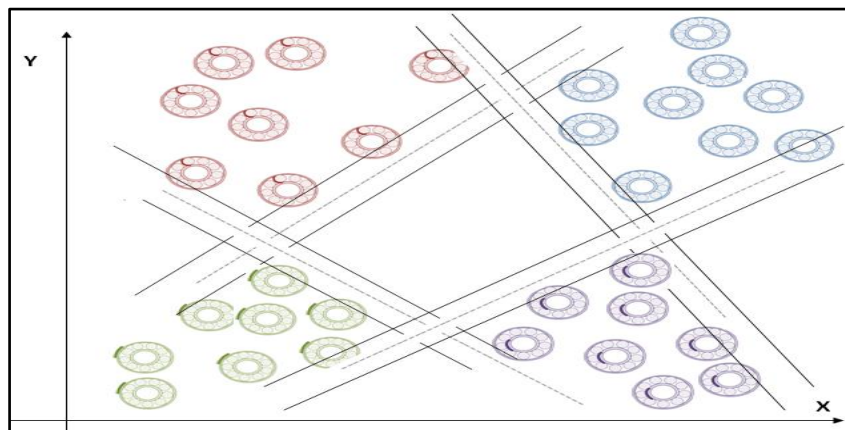


Figure 3.8 Schéma des hyperplans linéaires pour la classification un contre tous (One vs all) /n=4

b. Un-Contre-Un

La deuxième méthode est une méthode dite de un contre un. Au lieu d'apprendre N fonctions de décisions, ici chaque classe est discriminée d'une autre. Ainsi, $N(N - 1)/2$ fonctions de décisions (classificateurs) sont apprises et chacune d'entre elles effectue un vote pour l'affectation d'un nouveau point x . La classe de ce point x a le plus grand nombre de votes devient ensuite la classe majoritaire.

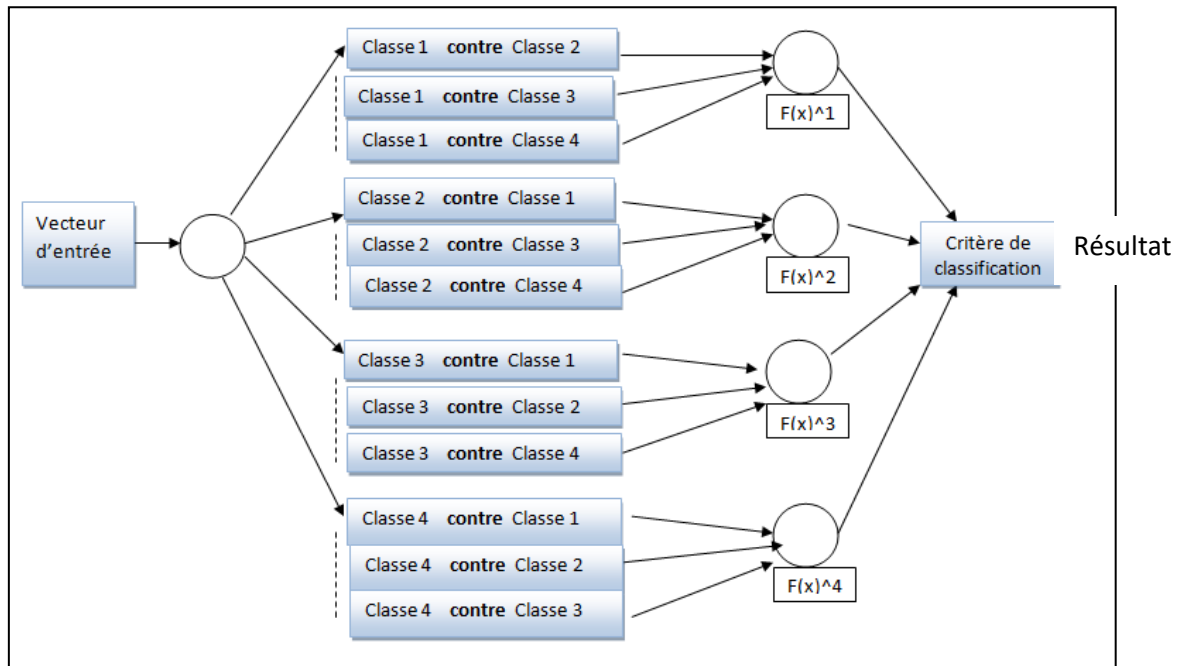


Figure 3.9 Classification SVM un contre un (One Vs One)

Conclusion

Les séparateurs à vaste marges proposent une méthode de classification qui montre de bonnes performances dans la résolution de problèmes variés. Cette méthode a montré son efficacité dans de nombreux domaines d'applications tels que le traitement d'image, la catégorisation de textes ou le diagnostics médicales et ce même sur des ensembles de données de très grandes dimensions. La réalisation d'un programme d'apprentissage par SVM se ramène à résoudre un problème d'optimisation impliquant un système de résolution dans un espace de dimension conséquente. L'utilisation de ces programmes revient surtout à sélectionner une bonne famille de fonctions noyau et à régler les paramètres de ces fonctions.

Ces choix sont le plus souvent faits par une technique de validation croisée, dans laquelle on estime la performance du système en la mesurant sur des exemples n'ayant pas été utilisés en cours d'apprentissage.

Références Bibliographiques

De Souza, D. B., Chanussot, J., & Favre, A. C. (2014, May). On selecting relevant intrinsic mode functions in empirical mode decomposition: An energy-based approach. In *Acoustics, Speech and Signal Processing (ICASSP), 2014 IEEE International Conference on* (pp. 325-329). IEEE.

Guyon, I., & Elisseeff, A. (2003). An introduction to variable and feature selection. *Journal of machine learning research*, 3(Mar), 1157-1182.

Jean-Christophe Cexus, (2005) Analyse des signaux non-stationnaires par transformation de Huang, Operateur de Teager-Kaiser, et Transformation de Huang-Teager (THT). Acoustics. Université Rennes 1. French. <tel-00012141>.

M. G.Rilling. , (2007). Décompositions Modales Empiriques Contribution à la théorie, l'algorithme et l'analyse de performances. Thèse de Doctorat. Université de Lyon.

N.E. Huang, Z. Shen, S.R. Long, M.C. Wu, H.H. Shih, Q. Zheng, N.C. Yen, C.C.Tung et H.H. Liu, (1998): The empirical mode decomposition and the Hilbert spectrum for Nonlinear and non-stationary time series analysis. *Proceedings of the Royal Society of London*, 454:903–995.

Chapitre

4

Résultats et Discussion

4.1. Matériels et méthodes :

4.1.1. Présentation du Banc d'Essai de la CWRU :

Comme représenté sur la figure 4.1, le banc d'essai se compose principalement d'un moteur (à gauche), d'un accouplement « transducteur/encodeur » (au centre), d'un dynamomètre (à droite) et d'un module d'acquisition des données.

Les signaux ont été enregistrés à l'aide des accéléromètres, qui ont été fixés à la cage du moteur par aimant. Pour différents couples résistants (donc différentes vitesses de rotations).

Les couples sont respectivement 0, 1, 2, 3 Hp dont des vitesses respectives sont 1797, 1772, 1750, 1730 tr/min.

Les capteurs ont été placés à la position 12 heures verticale des deux côtés accouplement et ventilateur de la cage du moteur en plus de deux autres capteurs à 6 et 3 heures.

La fréquence d'échantillonnage est $f_e = 12$ KHz (ou 48 KHz) et le nombre de points total est de $N_p = 243938$ pts.

Les données ont été enregistrées pour :

- Les roulements sans défaut.
- Les défauts uniques du cotés accouplement (12.000 et 48.000 points/second).
- Les défauts uniques du cotés ventilateur (12.000 points/second.)

Dans ce travail, les roulements que nous avons considérés, sont ceux du côté accouplement (12.000 points/second, pour les données) pour l'apprentissage et les tests. Les fréquences caractéristiques sont :

- Fréquence de la bague intérieure : $f_{bi} = 5.415 \times f_r$
- Fréquence de la bague extérieure : $f_{be} = 3.584 \times f_r$
- Fréquence de la cage : $f_c = 0.3982 \times f_r$
- Fréquence de la bille : $f_b = 4.7135 \times f_r$

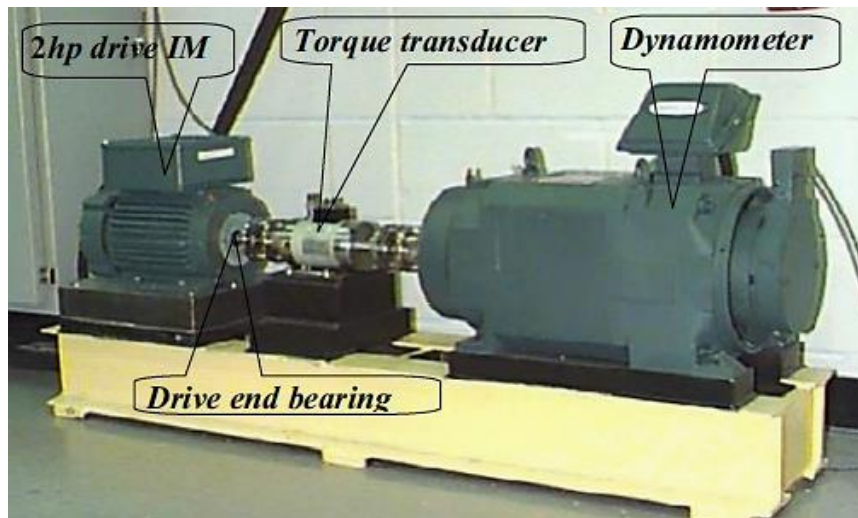


Figure 4.1 : Le Banc d'essai de la CWRU

Tableau 4.1 : Données vibratoires des défauts avec une fréquence d'échantillonnage de 12K HZ / CWRU

Fault Diameter (inches)	Motor Speed (rpm)	Inner Race	Ball	Outer Race Position Relative to Load Zone (Load Zone Centered at 6:00)		
				Centered @6:00	Orthogonal @3:00	Opposite @12:00
0.007"	1797	I07_0	B07_0	O07@6_0	O07@3_0	O07@12_0
	1772	I07_1	B07_1	O07@6_1	O07@3_1	O07@12_1
	1750	I07_2	B07_2	O07@6_2	O07@3_2	O07@12_2
	1730	I07_3	B07_3	O07@6_3	O07@3_3	O07@12_3
0.014"	1797	I14_0	B14_0	O14@6_0		
	1772	I14_1	B14_1	O14@6_1		
	1750	I14_2	B14_2	O14@6_2		
	1730	I14_3	B14_3	O14@6_3		
0.021"	1797	I21_0	B21_0	O21@6_0	O21@3_0	O21@12_0
	1772	I21_1	B21_1	O21@6_1	O21@3_1	O21@12_1
	1750	I21_2	B21_2	O21@6_2	O21@3_2	O21@12_2
	1730	I21_3	B21_3	O21@6_3	O21@3_3	O21@12_3
0.028"	1797	I28_0	B28_0			
	1772	I28_1	B28_1			

Figure 4.2 : Signal original cas sain

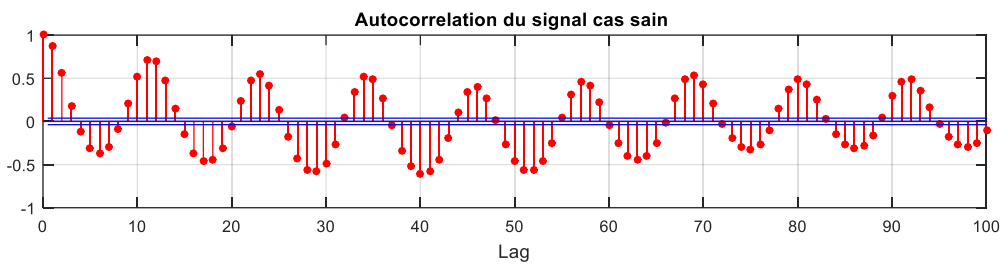
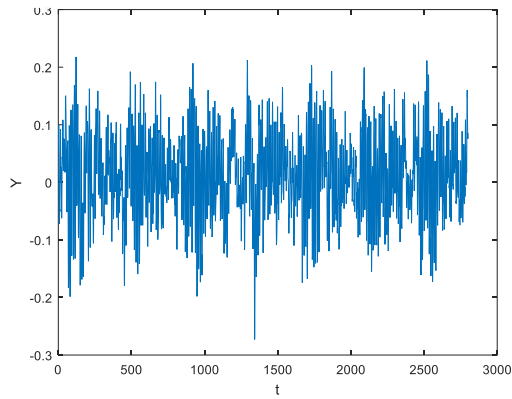


Figure 4.3 : Auto corrélation du signal original cas sain

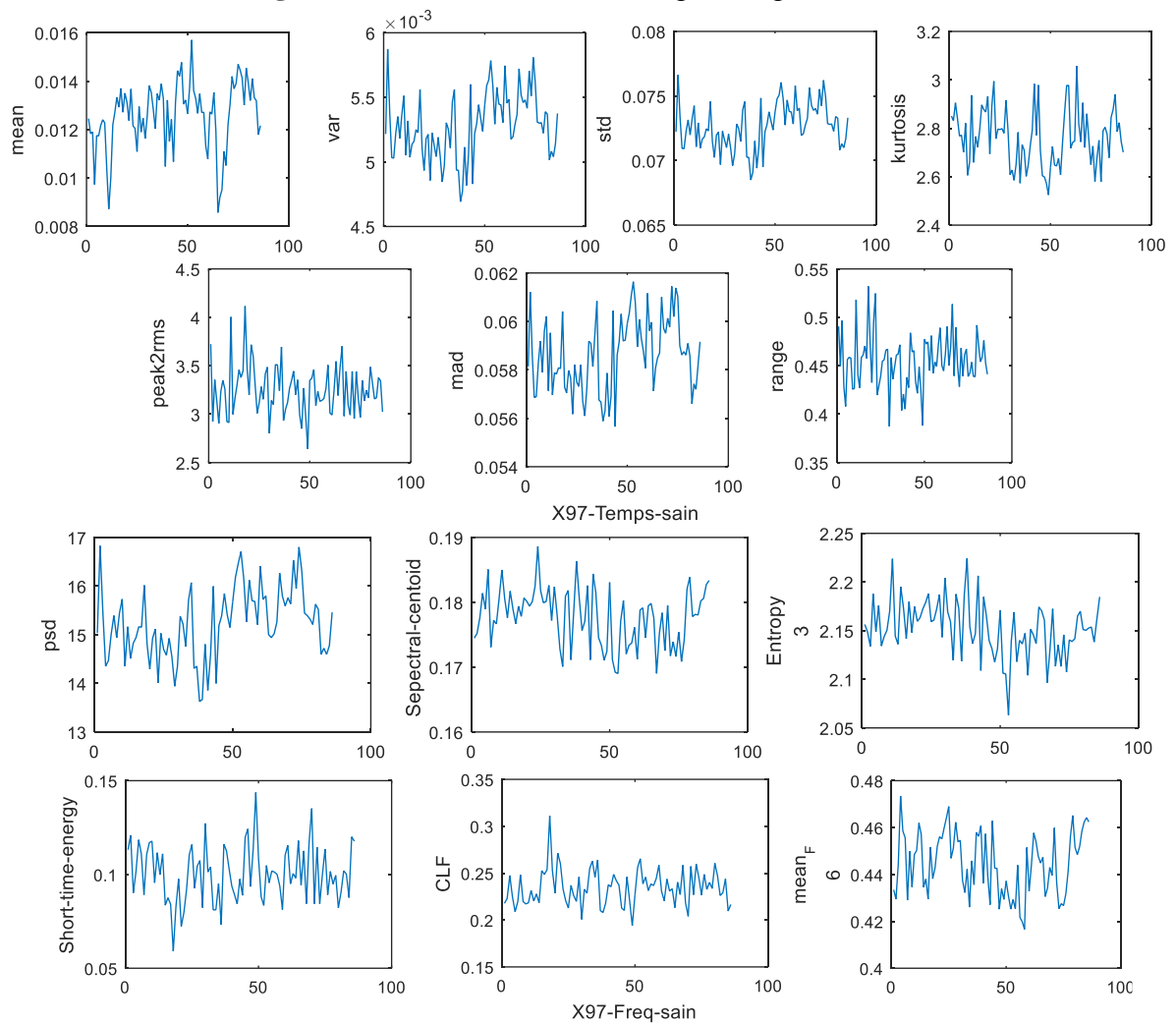


Figure 4.4 : Variation des paramètres temporels et fréquentiels du signal sans défauts (sain)

Figure 4.5 : Signal original avec défaut bille

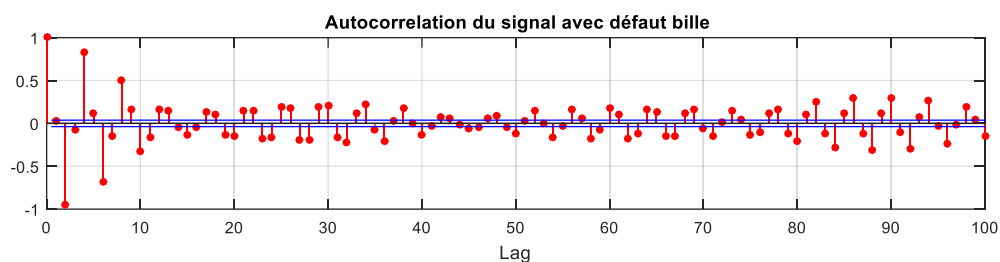
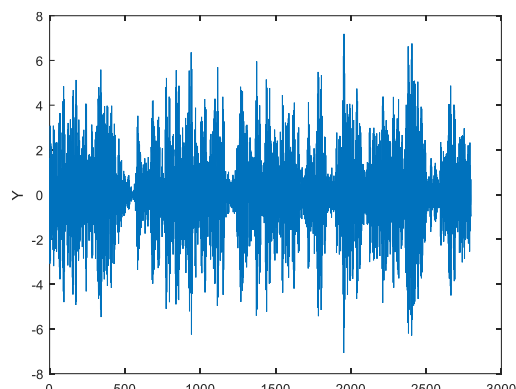


Figure 4.6 : Auto corrélation du signal ave défaut bille

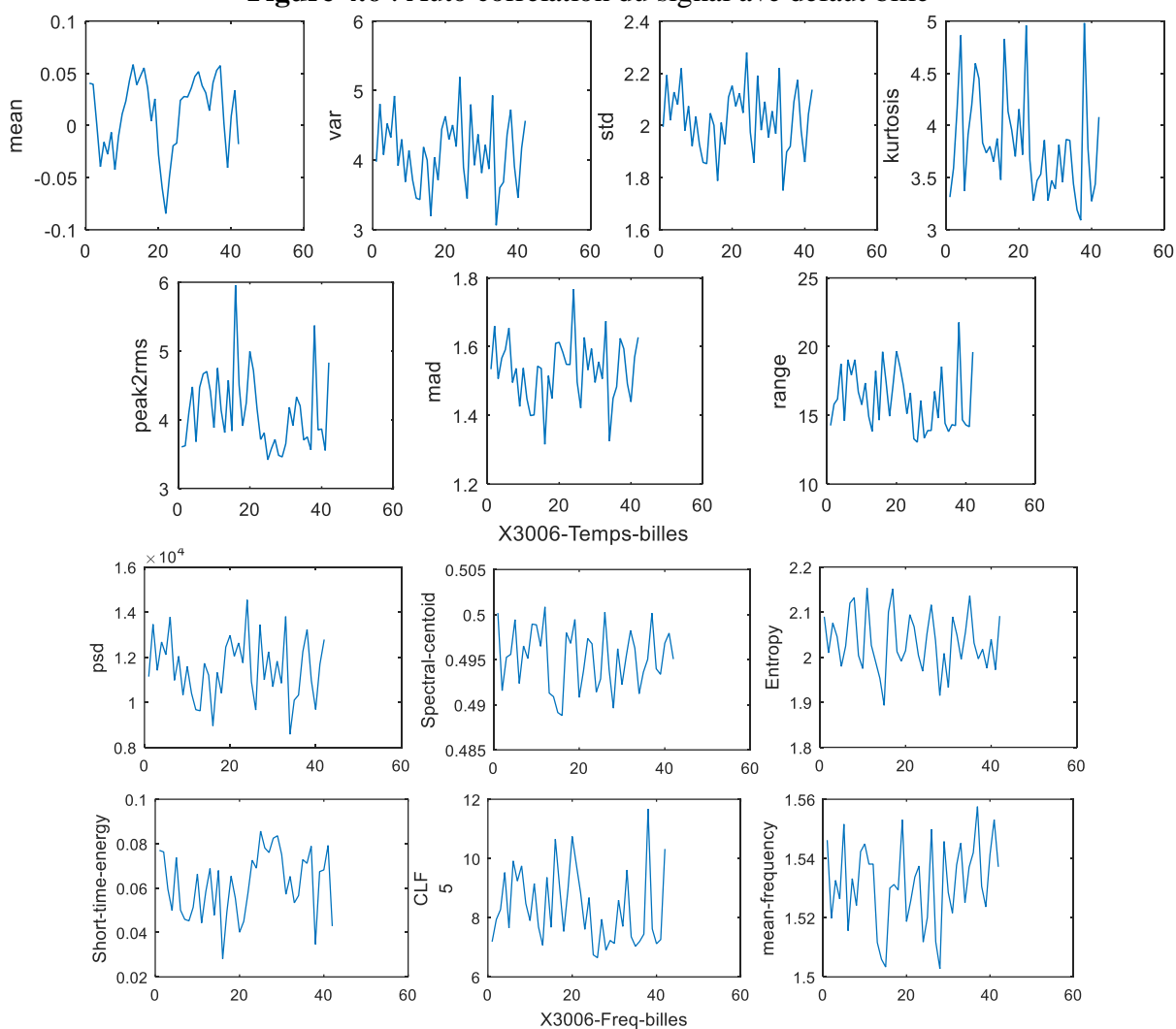


Figure 4.7 : Variation des paramètres temporelles et fréquentielles du signal avec défaut bille

Figure 4.8 : Signal original avec défaut bague interne

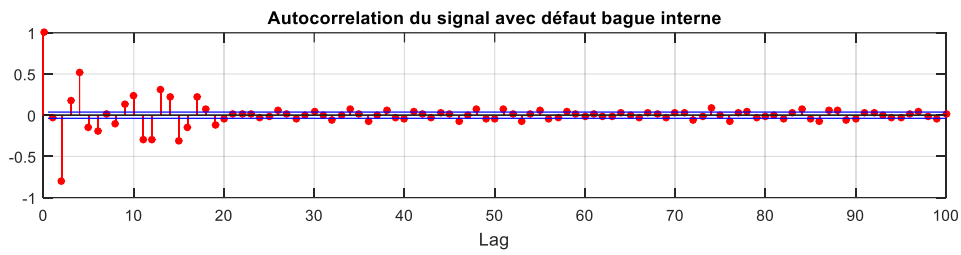
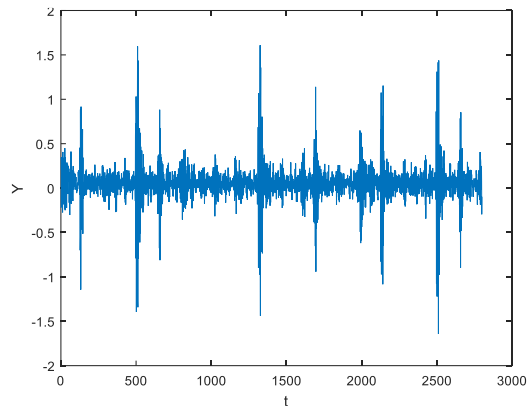


Figure 4.9 : Auto corrélation du signal avec défaut bague interne

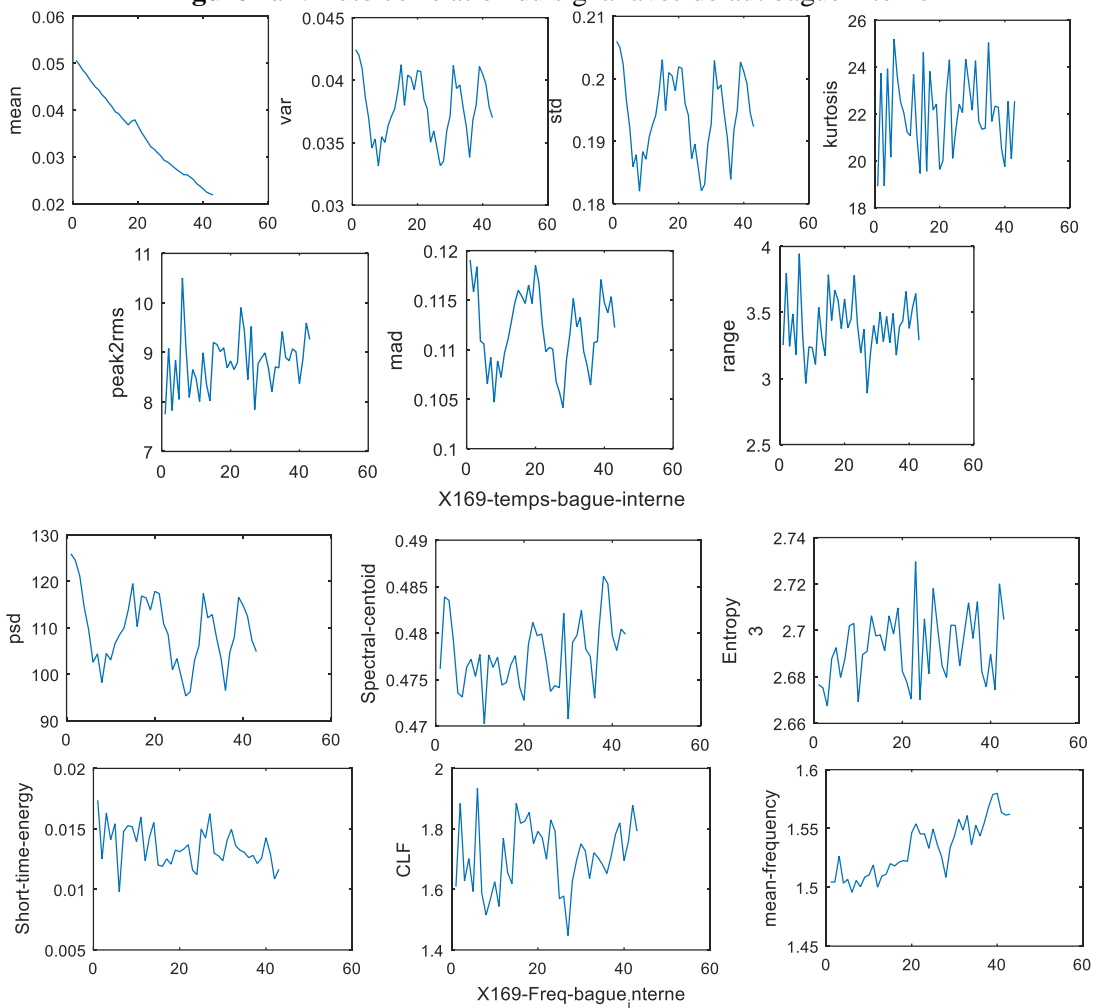


Figure 4.10 : Variation des paramètres temporels et fréquentiels du signal avec défaut bague interne.

Figure 4.11 : Signal original avec défaut bague externe

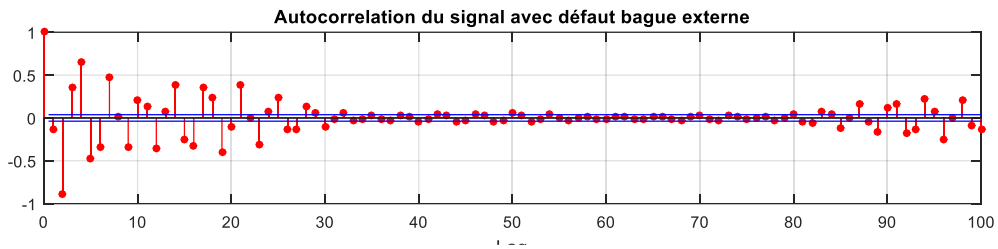
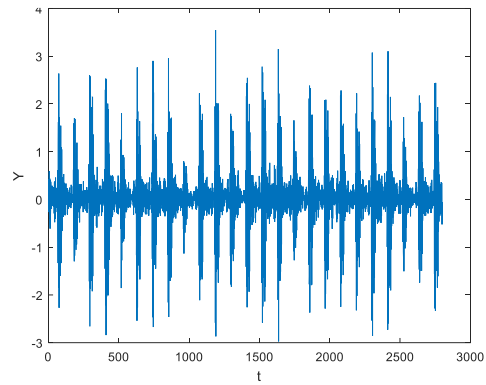


Figure 4.12 : Auto corrélation du signal avec défaut bague externe.

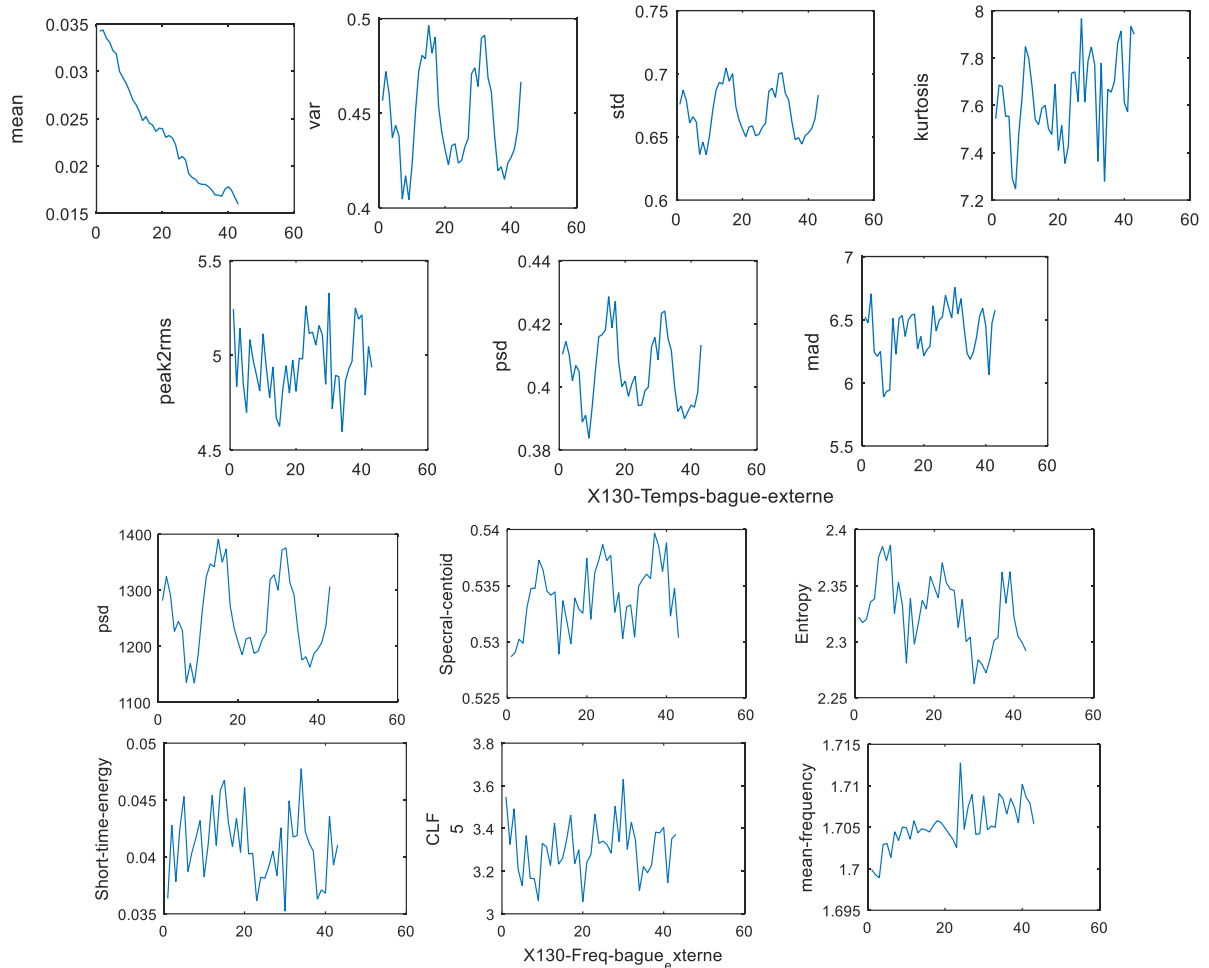


Figure 4.13 : Variation des paramètres temporels et fréquentielles du signal avec défaut bague externe

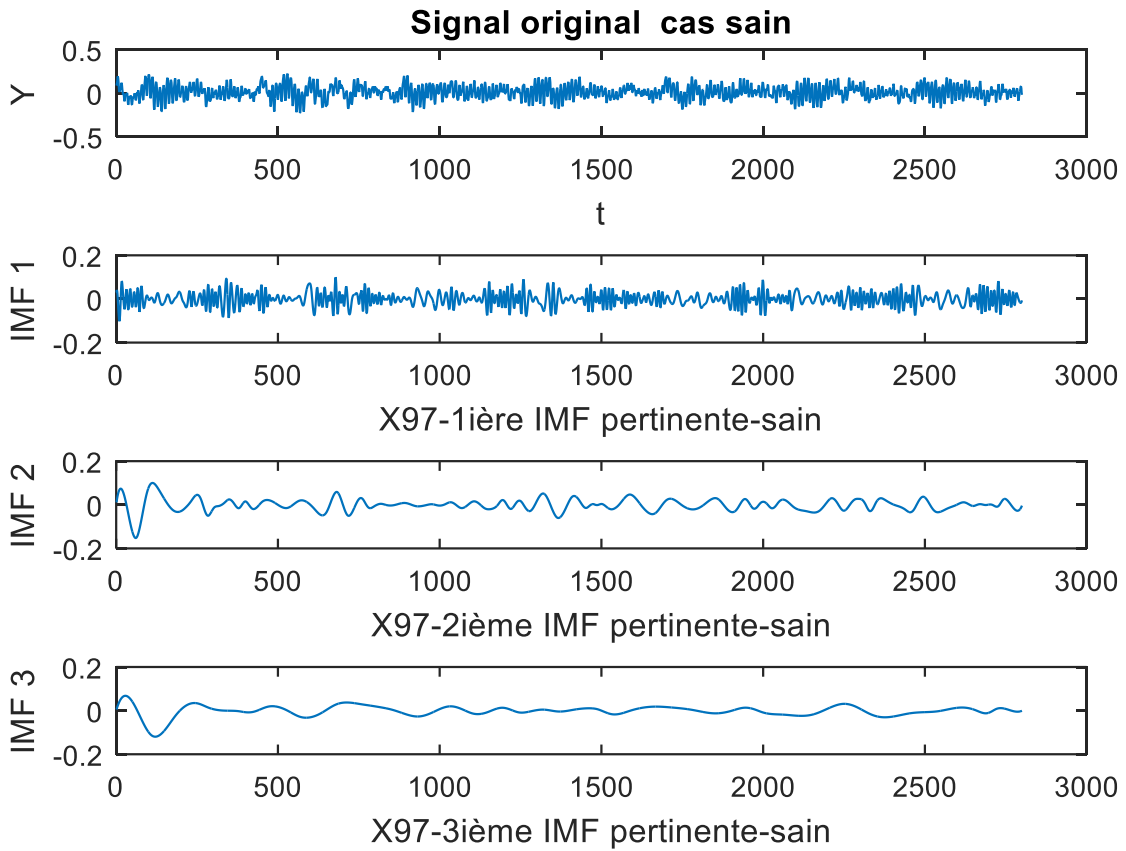


Figure 4.14 : Extraction des 3IMFs les plus pertinentes cas sain

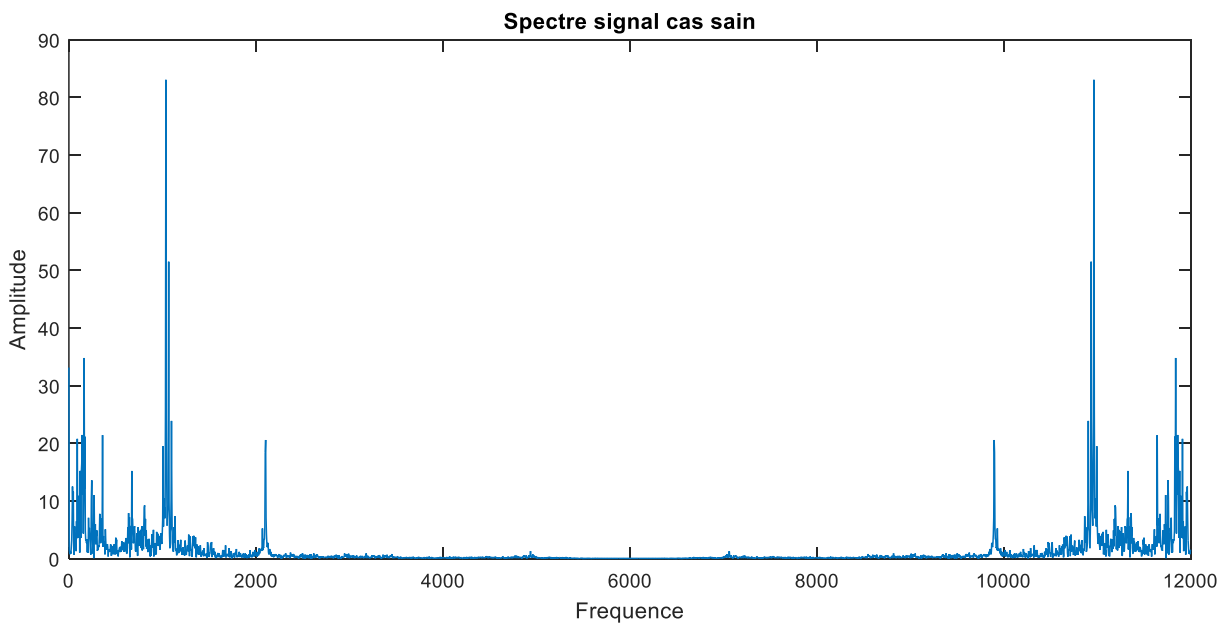


Figure 4.15 : Spectre du signal cas sain

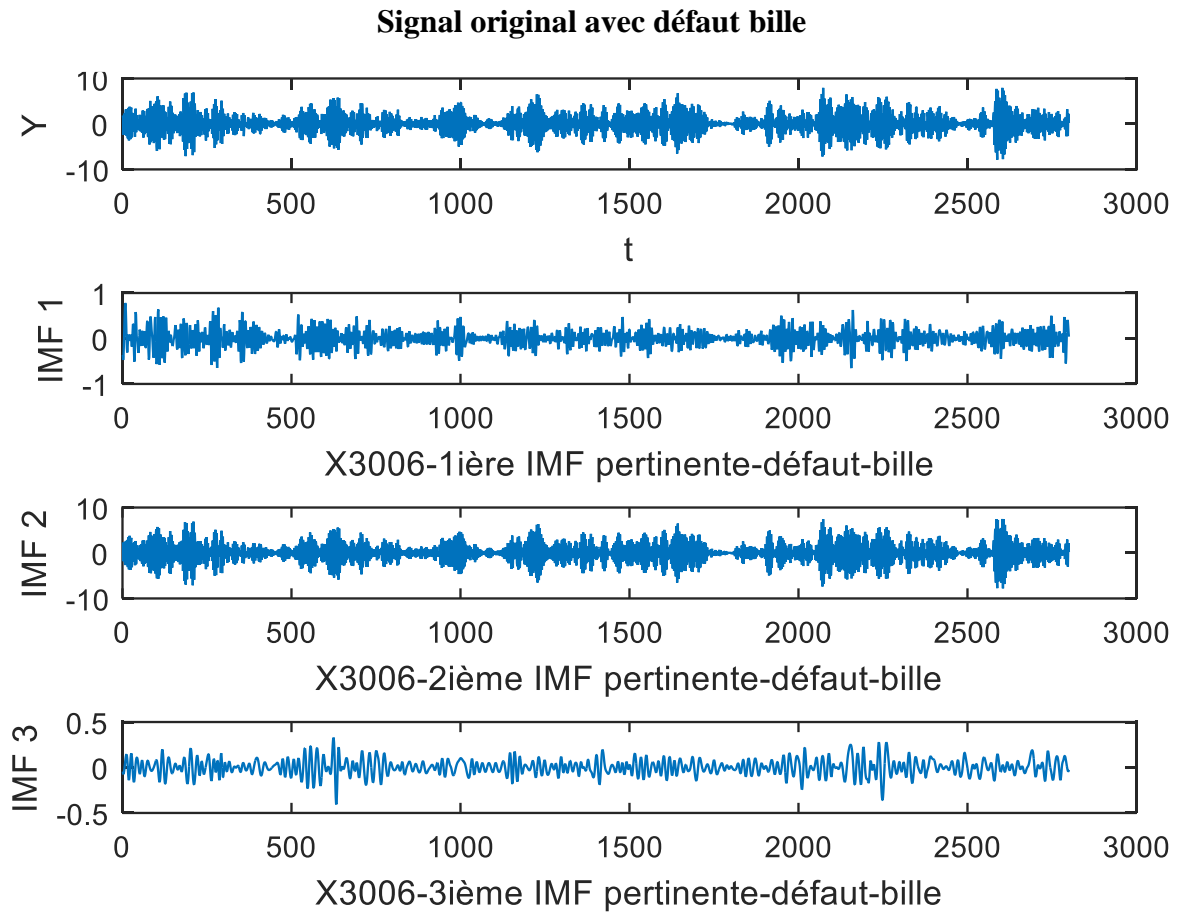


Figure 4.16 : Extraction des 3 IMFs les plus pertinentes du signal avec défaut bille

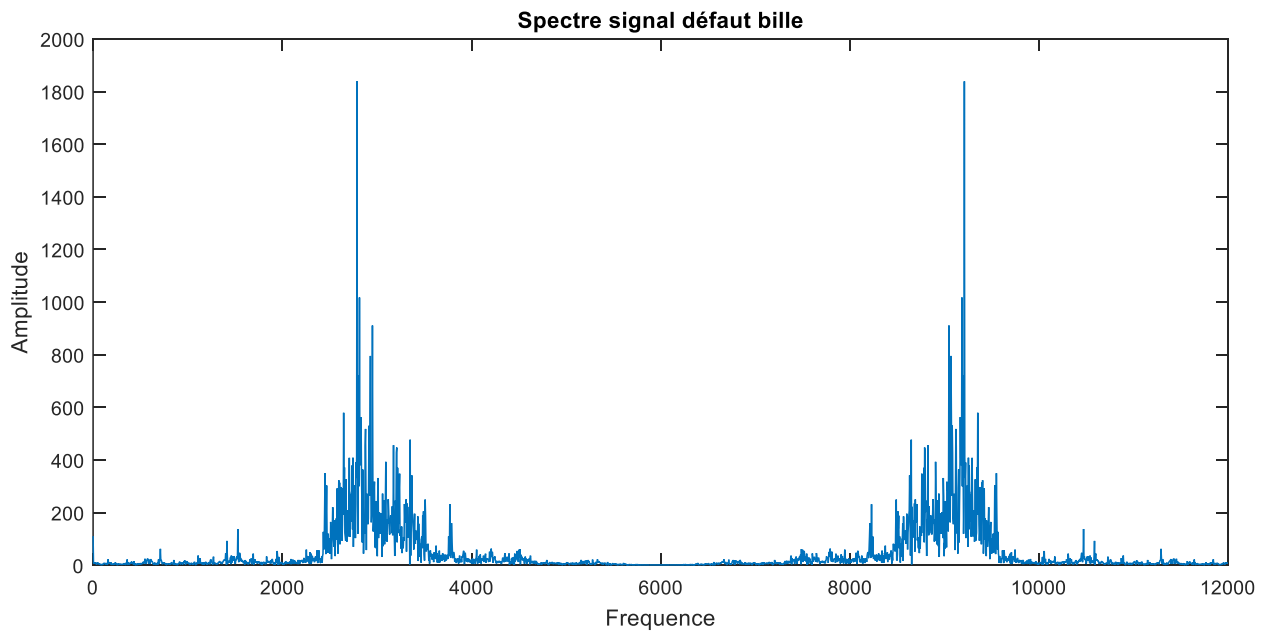


Figure 4.17 : Spectre du signal avec défaut bille

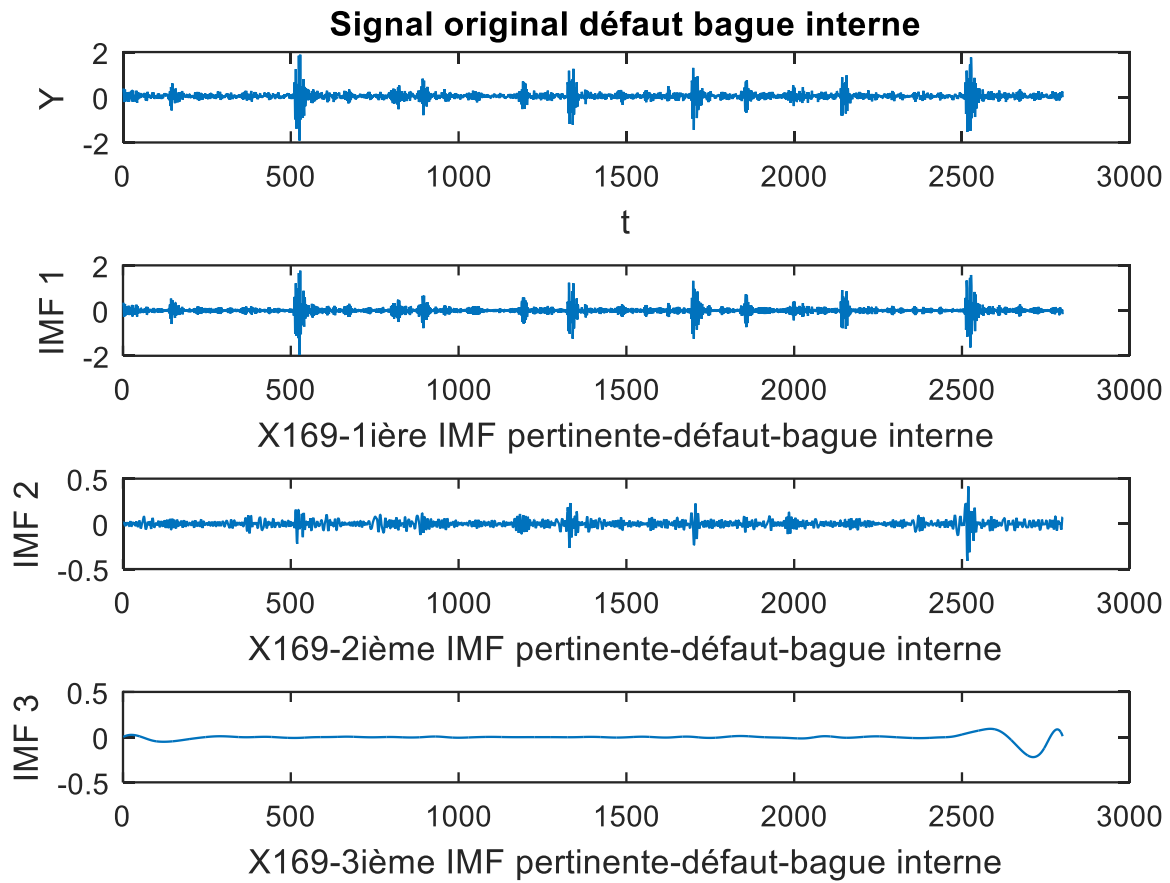


Figure 4.18 : Extraction des 3 IMfs les plus pertinentes du signal avec défaut Bague interne

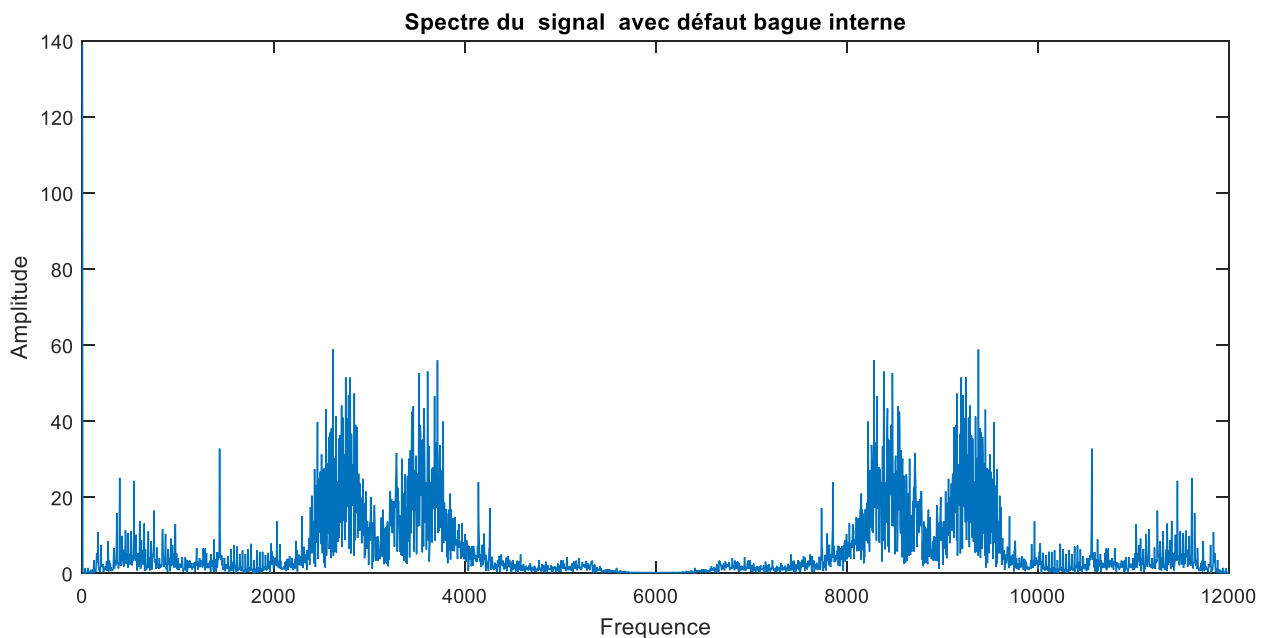


Figure 4.19 : Spectre du signal avec défaut bague interne

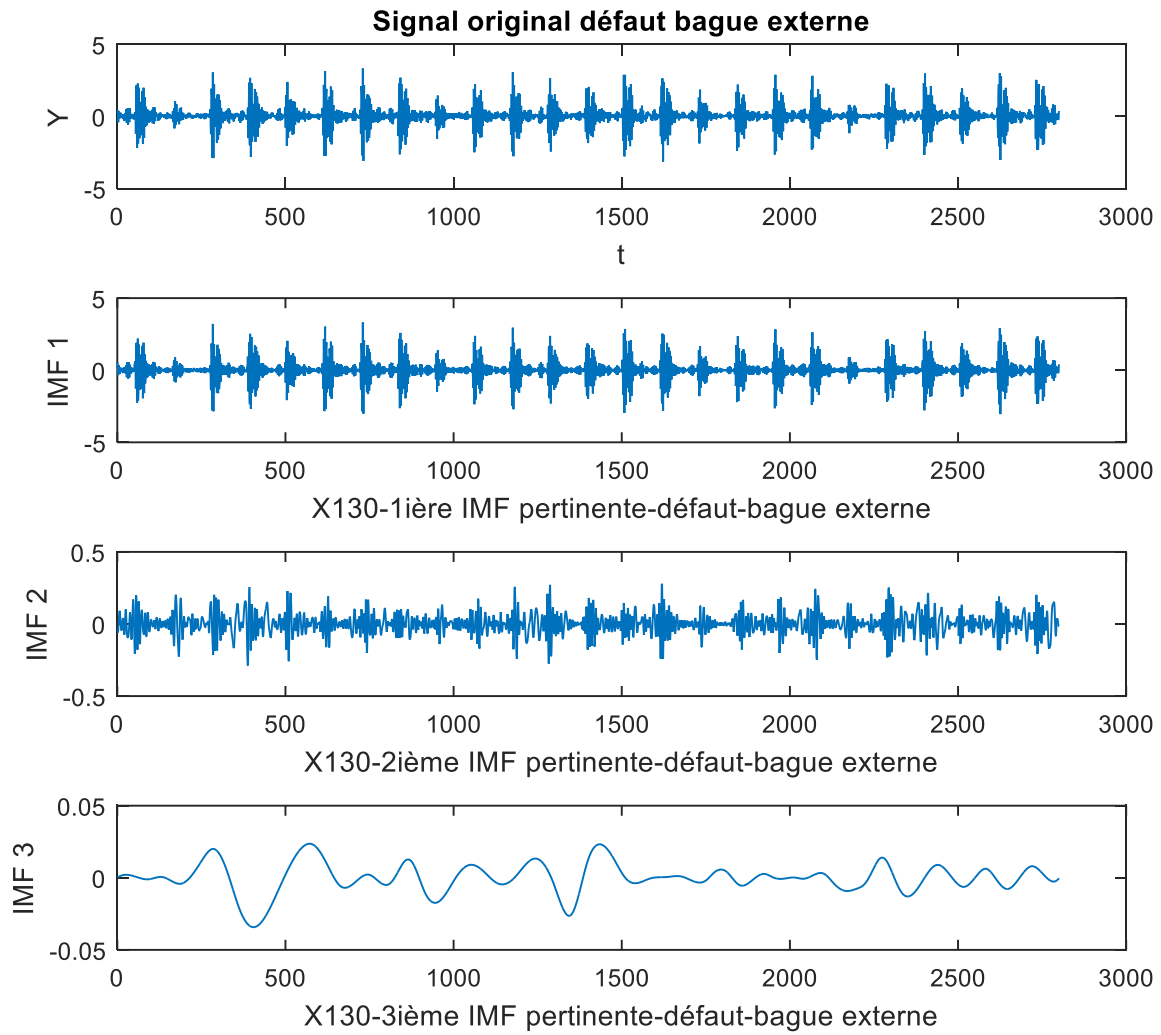


Figure 4.20 : Extraction des IMFs les plus pertinentes du signal avec défaut bague externe

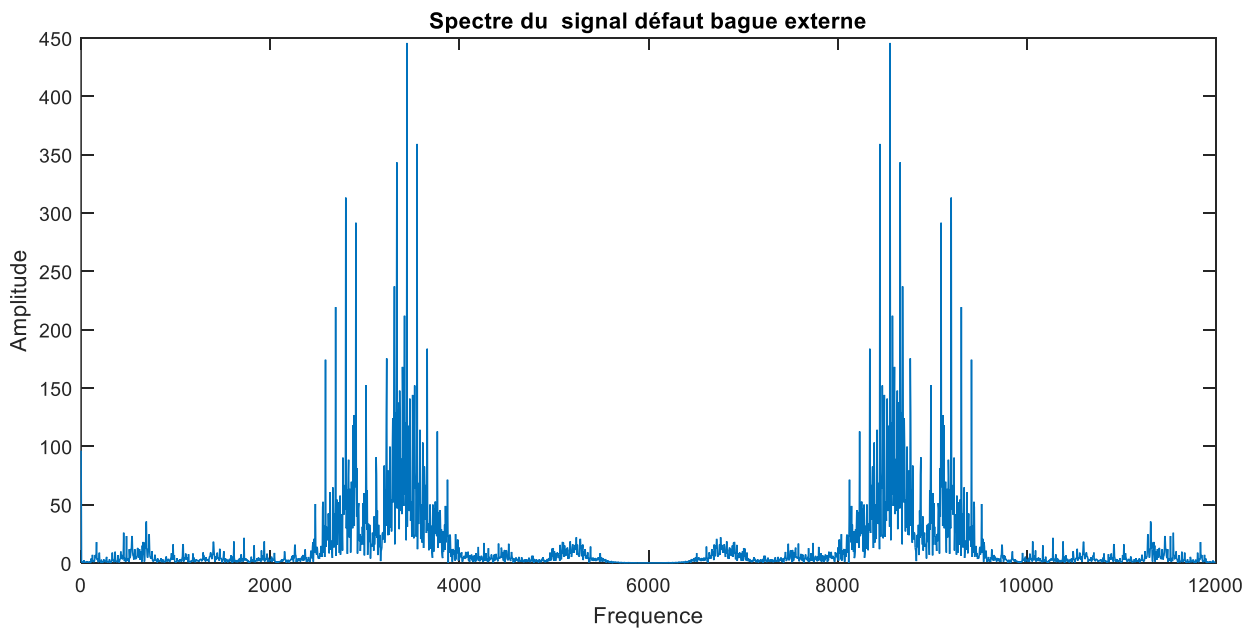


Figure 4.21 : Spectre du signal avec défaut bague externe

	Nbre Attributs Statistiques	Signal d'origine (SVM)		Réseaux de Neurones		EMD (03 IMFs) + SVM		Réseaux de Neurones	
		Taux bonne Classification	Temps Classification	C %	T _c	C %	T _c	C %	T _c
		C %	T _c (Sec)						
Fenêtrage 28000 pts	09	100/24	0.390	100	1.12	100/24	0.141	98/24	1.124
	10	/	/			/	/	/	/
	11	/	/			93.75/4	0.030	75/4	0.875
Fenêtra ge 2800 pts	09	95/24	0.135	93/24	0.397	99/24	0.046	92/24	0.182
	10	92/12	0.077	87/12	0.25	97/12	0.032	95/12	0.13

Tableau 4.2 : Récapitulatif des résultats

N.B. :

Le réseau de neurones utilisé pour la classification (pour des fins de comparaison avec la classification SVM) est un réseau multicouche possédant une couche cachée à 20 neurones et trois sorties binaires (0 0 0 pour reconnaître le cas sain, 1 0 0, 0 1 0 et 0 0 1 pour les défauts bague interne, bague externe et défauts billes respectivement).

$$C\% = \text{Taux de bonne classification} = \frac{\text{Nombre d'échantillons bien classifiés}}{\text{Nombre d'échantillons total}} \times 100$$

4, 12, 24 \longrightarrow Nombre d'échantillons utilisé pour le test

Exemple : Nombre d'échantillons X fenêtrage = 28000 X 24 = 672000 échantillons utilisés à une fréquence de 12KHz

Exemple de Matrice de Confusion

$$\begin{array}{l} S \\ Bin \\ Bext \\ B \end{array} \begin{bmatrix} 24 & 0 & 0 & 0 \\ 0 & 24 & 0 & 0 \\ 0 & 0 & 24 & 0 \\ 0 & 0 & 0 & 24 \end{bmatrix} \text{ Pour Fenêtrage (24 échantillons) du signal 28000, 09 attributs, et}$$

$$C\% = 100$$

$$\begin{array}{l} S \\ Bin \\ Bext \\ B \end{array} \begin{bmatrix} 4 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 \\ 0 & 2 & 2 & 0 \\ 0 & 0 & 0 & 4 \end{bmatrix} \text{ Pour Fenêtrage (04 échantillons) du signal 28000, 11 attributs, et}$$

$$C\% = 93.75$$

4.2. Discussion des résultats des classifications

Nous avons appliqué la phase d'apprentissage et la phase de test à la base de données de la CWRU préalablement détaillé. Le tableau (3) montre les taux de bonne classification ainsi que les temps de classification. Des exemples de matrices de confusion sont donnés en dessous.

Le taux de classification obtenu sans utiliser la DME, c'est-à-dire avec calcul des attributs statistiques directement sur les signaux vibratoires bruts, est de 100%. Lorsque les attributs statistiques sont calculés sur les 3 IMFs les plus pertinentes obtenues par l'application de la DME ensuite par la classification des IMFs ayant le facteur le plus élevé de corrélation avec le signal original, le taux de classification atteint 96.88% mais pour un temps de classification meilleur.

On remarque qu'en travaillant avec trois IMFs pertinentes au lieu du signal original, plus le nombre d'échantillons augmente plus le taux de bonne classification augmente mais au prix d'un temps de classification supérieur mais qui reste très appréciable et meilleur que celui par classification par réseaux de neurones.

Ces résultats montrent bien l'intérêt de la procédure de sélection des IMFs et des paramètres statistiques ainsi que la nécessité de combiner avec la DME pour la rapidité de la convergence.

L'efficacité du séparateur à vaste marge par rapport à la méthode des réseaux de neurones pour la classification des défauts roulements est plus qu'évidente.

Conclusion Générale

Conclusion Générale

En reconnaissance de formes, plusieurs méthodes ont été introduites pour le diagnostic mécanique des pannes, parmi lesquelles et les plus couramment utilisées sont les réseaux neuronaux artificiels et, les séparateurs à vaste marge.

La formation des réseaux de neurones artificiels nécessite toujours un grand nombre d'échantillons, alors que les séparateurs vastes marges qui sont basés sur la théorie de l'apprentissage statistique sont spécifiques pour un nombre d'échantillons réduit et ont une meilleure généralisation que les réseaux de neurones artificiels et c'est l'utilisation de noyaux qui permet cette généralisation. Cependant la précision d'une machine vectorielle de support est hautement déterminée par la sélection des paramètres optimaux.

La plus grande différence entre le SVM et le RNA est dans la méthode d'optimisation des paramètres du modèle, le RNA utilise la retro propagation du gradient qui ne garantit pas l'obtention d'une solution optimale par contre l'implémentation du SVM trouvera les paramètres optimaux.

Cependant, ils peuvent être complémentaire avec des combinaisons SVM et RNA. Par exemple, on peut utiliser plusieurs couches RNA et avoir la dernière classification via SVM à la couche de sortie.

L'algorithme développé dans cette étude peut être étendu à plusieurs types de défauts (composants, actionneurs, capteurs, ...etc.) sur les machines jusqu'à la conception d'un système d'avertissement adapté dans lequel l'apparition des défauts est rapportée à l'utilisateur sur l'écran de commande de la machine pour un diagnostic en temps réel, ou bien sur une manette portative pour un diagnostic à la demande.

Une autre idée est d'utiliser les informations existantes sur les machines et les exploiter avec cet algorithme pour en tirer un diagnostic fiable, sans pour autant déployer un matériel supplémentaire (capteurs et interfaces de calculs).