



HAL
open science

Détection de nouveauté pour le monitoring vibratoire des structures de génie civil: Approches chaotique et statistique de l'extraction d'indicateurs

Antoine Clément

► To cite this version:

Antoine Clément. Détection de nouveauté pour le monitoring vibratoire des structures de génie civil: Approches chaotique et statistique de l'extraction d'indicateurs. Vibrations [physics.class-ph]. Université Paul Sabatier - Toulouse III, 2011. Français. NNT: . tel-00687065

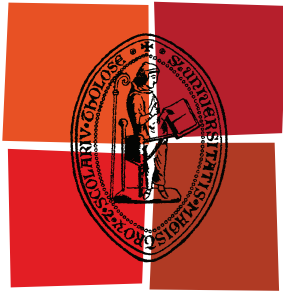
HAL Id: tel-00687065

<https://theses.hal.science/tel-00687065>

Submitted on 12 Apr 2012

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Université
de Toulouse

THÈSE

En vue de l'obtention du
DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE

Délivré par :

Université Toulouse III Paul Sabatier (UT3 Paul Sabatier)

Discipline ou spécialité :

Génie Civil

Présentée et soutenue par :

CLEMENT Antoine

le : 21 Novembre 2011

Titre :

Détection de nouveauté pour le monitoring vibratoire des structures de génie civil : approches chaotique et statistique de l'extraction d'indicateurs

Ecole doctorale :

Mécanique, Energétique, Génie civil et Procédés (MEGeP)

Unité de recherche :

Laboratoire Matériaux et Durabilité des Constructions

Directeur(s) de Thèse :

LAURENS Stéphane

Rapporteurs :

C.CREMONA (SETRA)

A.DERAEMAER (Université Libre de Bruxelles)

Membre(s) du jury :

J-P.BALAYSSAC (Université de Toulouse III)

E.RINGOT (Université de Toulouse III)

F.SCHOEFS (Université de Nantes)

Titre

Détection de nouveauté pour le monitoring vibratoire des structures de génie civil : Approches chaotique et statistique de l'extraction d'indicateurs

Résumé

Le suivi vibratoire de l'état des ouvrages de génie civil vise à anticiper une défaillance structurale par la détection précoce d'endommagement. Dans ce contexte, la détection de nouveauté constitue une approche particulièrement adaptée à l'analyse des signaux compte tenu des difficultés à modéliser une structure unique et soumise à de nombreux facteurs extérieurs influant sur la dynamique vibratoire. Une telle approche présente un double intérêt consistant à éviter de formuler des hypothèses a priori sur le comportement dynamique et à intégrer tous les facteurs de variabilité. Ce travail de thèse poursuit ainsi deux objectifs. Le premier objectif consiste à observer dans quelle mesure la détection de nouveauté parvient à détecter un endommagement dans un contexte fortement perturbé par des variations environnementales d'une part, et par une excitation de nature impulsionnelle, d'autre part. Le deuxième objectif est de proposer et d'étudier un nouvel indicateur vectoriel, désigné par JFV (pour Jacobian Feature Vector). Le calcul du JFV s'appuie sur la reconstruction de la trajectoire du système dynamique observé dans son espace des phases. Cette approche exploite les développements scientifiques récents réalisés en théorie des systèmes dynamiques non linéaires, parfois qualifiée de théorie du chaos. Le JFV est comparé aux coefficients de modèles auto-régressifs (AR), couramment utilisés en analyse des séries temporelles. Pour réaliser ce travail de thèse, plusieurs cas d'études expérimentaux sont utilisés dont notamment une maquette de structure en bois sur laquelle l'excitation est contrôlée et des variations environnementales sévères sont imposées.

Les indicateurs AR et JFV sont extraits des signaux vibratoires relatifs aux différents cas d'études et normalisés par le biais du concept de distance de Mahalanobis. Les résultats expérimentaux montrent que, pour les deux indicateurs vectoriels, la détection de l'endommagement est favorisée par une sollicitation comportant une composante de bruit. Une excitation purement instationnaire, constituée de séquences aléatoires d'impulsions, dégrade de façon significative les performances de détection. Les variations environnementales génèrent une forte variabilité des indicateurs, rendant difficile l'ajustement d'un modèle statistique robuste dédié à la discrimination des dégradations. Seuls les niveaux d'endommagement extrêmes sont repérés dans la configuration d'essai la plus pénalisante. L'analyse comparée des coefficients AR et du JFV met en évidence une dispersion beaucoup plus grande des composantes de ce dernier, conduisant à une sensibilité plus faible. Une étude paramétrique montre cependant que la sensibilité du JFV peut être améliorée par une optimisation des méthodes de sélection des paramètres de reconstruction de l'espace des phases. Face aux performances limitées des indicateurs AR et JFV dans certains cas très défavorables, un autre indicateur est proposé, basé sur la corrélation croisée des informations portées par une paire de capteurs. Cet indicateur présente une performance intéressante sur un cas d'étude complexe combinant variations environnementales fortes et sollicitation purement instationnaire. Une discussion est également proposée sur la façon de répartir les mesures de référence dans

les différentes bases de données nécessaires à l'application de la démarche de détection de nouveauté. Enfin, différentes approches de modélisation statistique des indicateurs normalisés sont mises en œuvre dans le but de comparer leurs aptitudes respectives à la définition d'un seuil de classification robuste.

Mots-clés

détection de nouveauté, structure, analyse vibratoire, détection de l'endommagement, théorie des systèmes dynamiques non-linéaires, espace des phases, attracteur, variations environnementales, sollicitation instationnaire

Title

Novelty detection for vibration-based structural health monitoring of Civil structures : damage sensitive feature extraction by chaotic and statistical approaches

Abstract

The aim of structural health monitoring of civil structure is the early detection of damage to prevent structure failure. But modelling the behaviour of such structure is a very challenging task due to its uniqueness and to the effect of environmental parameters on the dynamic. In this context, the novelty detection approach appears to be well adapted since it avoids the need of prior hypothesis on the nature of the dynamical behaviour, and integrates all variability factors. The work of this thesis has two principal aims. The first one is to quantify the ability of novelty detection to discriminate damage under strong environmental variations and impulse excitation. The second one is to introduce a new damage sensitive feature, referred as *Jacobian Feature Vector* (JFV). The JFV calculation is based on the reconstructed state space which exploits the progress achieved in the theory of non-linear dynamical systems, also known as chaos theory. The comparison between AR parameters, widely used for time series analysis, and the JFV is carried out on several case studies. One of them is a three storey wooden laboratory structure subjected to strong environmental variations and controlled excitation.

Both the JFV and AR parameters are estimated on vibration measurements and normalized by the *Mahalanobis distance* concept. The experimental results show that damage detection is improved if the excitation contains a part of noise in comparison to a random sequence of pulses. Environmental variations generate an important variability of the feature vectors, making difficult the fitting of any statistical model needed for damage discrimination. Only the extremes damage levels are detected in the worst test configuration. The comparison between the feature vectors reveals a higher dispersion of the JFV components, leading to a lower sensitivity to damage, as well as to the effect of temperature. However, a parametric investigation shows that damage detection can be improved if the selection process of the state-space reconstruction parameters is optimized. Given the limited performances of AR and JFV in the most difficult configurations, another damage sensitive feature based on cross-correlation of sensor pairs is proposed. In a complex case study which combines environmental variations and impulse excitation, the performances of this feature vector are promising. A discussion about the more rigorous way to divide the undamaged data to apply novelty detection procedure is also proposed. Finally, since the last step of novelty detection is decision making based on statistical modelling of the normalized damage sensitive features, the robustness of several approaches for the setting of the classification threshold is investigated.

Keywords

novelty detection, vibration analysis, damage detection, nonlinear dynamical systems theory, state-space, attractor, environmental variations, instationary excitation

REMERCIEMENTS

JE voudrais tout d'abord exprimer ma gratitude envers l'état Français. C'est assez rare, mais c'est grâce à lui que j'ai pu étudier pendant toutes ces années dans des conditions privilégiées. D'abord en classes prépas qui ont renforcées mon goût pour la science et la réflexion. Puis j'ai eu la chance d'étudier à l'ENS avec un statut me permettant de me consacrer seulement à ma formation, dans un environnement pédagogique et scientifique rare. Enfin, en thèse, où j'ai eu la possibilité d'organiser et d'orienter mon travail en toute liberté.

Je suis très reconnaissant envers les membres du jury et plus particulièrement mes rapporteurs M.Cremona et M.Deraemaeker pour leurs remarques qui me seront précieuses dans la poursuite de mes travaux.

Mes pensées vont ensuite envers toutes les personnes qui m'ont aidé à mener à bien mes travaux. Je pense aux personnels techniques du laboratoire (Sylvain, Frédéric, Marc, Guislaine), à Ginette Arliguie pour avoir su trouver les ressources nécessaires, à Bertrand Combes pour ses conseils et son soutien matériel, à Jean-Paul Balayssac pour sa bienveillance et plus généralement au laboratoire qui m'a accueilli.

La réussite humaine d'une thèse dépend en grande partie de ceux que l'on côtoie quotidiennement, à savoir ses co-doctorants. Et c'est une grande réussite pour moi. Un grand merci donc aux habitués du bureau 210 (Paco, Pauline, Thomas), 316 (Youssef) pour les moments partagés à la recherche des bons mots. Merci aux *Gossip 319*, aux *Unusual Suspects*, ou dernièrement aux *Big LMDC Theory*, c'est à dire Sam, Camille, Rackel, Muazzam, Steeves, Hugo et Marie, d'avoir supporté quotidiennement mes crises d'hypoglycémies, mes goûts musicaux (Katy ?) et plus généralement mon caractère. Ce fut un plaisir !

Certains co-doctorants un peu lointains ont été également d'un grand soutien pendant les périodes de doutes caractéristiques d'une thèse, je pense à mes collègues Kchanais de Lyon, Sylvain, Aline, Emilien, Jérôme, de Nantes, Raf et de Paris, Mimine. Merci pour nos échanges et pour les grands bols d'airs pris ensemble et ceux à venir...

Merci à mon "*chef*" Stéphane de m'avoir fait vivre une thèse à *l'ancienne*. Je te suis reconnaissant de ta confiance, de ton soutien indéfectible et de tes encouragements dans les nombreux moments de remise en question dont tu as été le premier observateur. Merci également pour l'image que tu m'as donnée de la recherche.

Merci à tous ceux qui ont montré un intérêt pour mon travail, mes amis de toujours, ma seconde famille bouillarguaise et enfin ma Famille.

J'ai oublié un soutien essentiel, celui de l'équipe technique de Beaucaire constituée par mon père qui a apporté une importante contribution à mon travail. Plus généralement, je ne remercierai jamais assez mes parents à qui je dois ce que je suis aujourd'hui.

Enfin, merci à *Micheline*, d'avoir vécu avec moi cette étape de plus d'un long parcours commencé il y a bien longtemps...

LISTE DES ABRÉVIATIONS

ACP	<i>Analyse en Composante Principale</i>
AIC	<i>Akaike Information Criterion</i>
ALAVR	<i>Average Local Attractor Variance Ratio</i>
AR	<i>Auto-Régressif</i>
ARMA	<i>Auto-Régressif à Moyenne Mobile</i>
BA	<i>Base d'Apprentissage</i>
BC	<i>Base de Classification</i>
BT	<i>Base de Test</i>
CAAD	<i>Chaotic Amplification of Attractor Distortion</i>
EL	<i>Exposants de Lyapunov</i>
EVS	<i>Extreme Value Statistics</i>
FA	<i>Fausse Alarme</i>
FAC	<i>Fonction d'Auto-Corrélation</i>
FIM	<i>Fonction d'Information Mutuelle</i>
FRF	<i>Fonction de Réponse en Fréquence</i>
FV	<i>Faux Voisins</i>
GMM	<i>Gaussian Mixture Model</i>
GDP	<i>Generalized Pareto Distribution</i>
IM	<i>Information Mutuelle</i>
JFV	<i>Jacobian Feature Vector</i>
LANL	<i>Los Alamos National Laboratory</i>
MA	<i>Moyenne Mobile</i>
MB	<i>Maquette en Bois</i>
MD	<i>Distance de Mahalanobis</i>
MM	<i>Maquette Métallique</i>
MR	<i>Masse/Ressort</i>
MSP	<i>Maîtrise Statistique des Processus</i>
PACF	<i>Partial Auto-Correlation Function</i>
PE	<i>Prediction Error</i>
RBF	<i>Radial Basis Function</i>
RFS	<i>Reconnaissance de formes Statistiques</i>
RNA	<i>Réseau de Neurones Artificiels</i>
RRF	<i>Relative Response Function</i>
SHM	<i>Structural Health Monitoring</i>
SPRT	<i>Sequential Probability Ration Test</i>
SVM	<i>Support Vector Machine</i>
UCL	<i>Upper Confidence Limit</i>
VD	<i>Vraie Détection</i>

TABLE DES MATIÈRES

TABLE DES MATIÈRES	xii
Introduction Générale	1
I Éléments contextuels et bibliographiques	5
1 DÉTECTION DE NOUVEAUTÉ EN CONTEXTE DE SHM	7
INTRODUCTION	8
1.1 PROBLÉMATIQUE DES STRUCTURES DE GÉNIE CIVIL	9
1.1.1 Définitions	9
1.1.2 Particularités des structures de génie civil	10
1.1.3 Bilan	13
1.2 PARADIGME DE LA RECONNAISSANCE DE FORMES STATISTIQUES (RFS)	14
1.3 DÉTECTION DE NOUVEAUTÉ	16
1.3.1 Indicateurs	16
1.3.2 Normalisation	21
1.3.3 Seuil de rejet	33
BILAN	42
2 POTENTIEL DE L'APPROCHE CHAOTIQUE	43
INTRODUCTION	44
2.1 REPRÉSENTATION DANS L'ESPACE DES PHASES	45
2.1.1 Définition	45
2.1.2 Reconstruction de l'espace des phases	45
2.2 EXTRAPOLATION AUX SYSTÈMES STOCHASTIQUES	47
2.3 INDICATEURS POTENTIELS	48
2.3.1 Dimension de l'attracteur	48
2.3.2 Les Exposants de Lyapunov	50
2.3.3 L'erreur de prédiction non-linéaire	52
2.3.4 Caractérisation géométrique	54
BILAN	55

II	Méthodes de calcul et cas d'études expérimentaux	57
3	MÉTHODES DE CALCUL	59
	INTRODUCTION	60
3.1	INDICATEURS	61
3.1.1	Coefficients auto-régressifs	61
3.1.2	Reconstruction de l'espace des phases	63
3.1.3	Exposants de Lyapunov	66
3.1.4	Indicateur Jacobien	69
3.2	NORMALISATION PAR LA DISTANCE DE MAHALANOBIS	70
3.3	MODÉLISATION STATISTIQUE	71
3.3.1	Test sur les valeurs individuelles : $N_e = 1$	71
3.3.2	Test sur un échantillon : $N_e > 1$	73
4	CAS D'ÉTUDES EXPÉRIMENTAUX	77
	INTRODUCTION	78
4.1	SYSTÈME MASSE/RESSORT AMORTI	79
4.2	MAQUETTE MÉTALLIQUE	80
4.3	POUTRE EN BÉTON FIBRÉ	81
4.4	MAQUETTE EN BOIS	84
III	Applications de la détection de nouveauté	91
	INTRODUCTION	93
5	SENSIBILITÉ À L'ENDOMMAGEMENT	95
	INTRODUCTION	96
5.1	ANALYSE COMPARÉE DES PERFORMANCES DES EXPOSANTS DE LYAPUNOV ET DU JFV	97
5.1.1	Étude paramétrique des algorithmes de calcul des EL et du JFV	97
5.1.2	Paramètres de reconstruction de l'espace des phases	97
5.1.3	Exploitation des données expérimentales	99
5.1.4	Résultats	99
5.1.5	Conclusion	104
5.2	COMPARAISON ENTRE COEFFICIENTS AR ET JFV	104
5.2.1	Données analysées	104
5.2.2	Paramètres de calcul	105
5.2.3	Résultats	108
5.2.4	Conclusions	114
	BILAN	115
6	SENSIBILITÉ AUX VARIATIONS ENVIRONNEMENTALES	117

INTRODUCTION	118
6.1 DONNÉES ÉTUDIÉES	119
6.2 PARAMÈTRES DE CALCUL	119
6.2.1 Maquette en bois	119
6.2.2 Poutre en béton fibré	120
6.3 OBSERVATION DES INDICATEURS	122
6.3.1 Effet des variations environnementales	122
6.3.2 Répartition dans l'espace des indicateurs vectoriels	123
6.4 DIVISION DE LA BASE DE DONNÉES COLLECTÉES SUR STRUC- TURE SAINTE	125
6.5 DÉTECTION DE L'ENDOMMAGEMENT	128
6.5.1 Poutre en béton fibré	130
6.5.2 Maquette en bois	131
BILAN	135
7 INFLUENCE D'UNE EXCITATION INSTATIONNAIRE	137
INTRODUCTION	138
7.1 DONNÉES ÉTUDIÉES	139
7.2 PARAMÈTRES DE CALCUL	139
7.2.1 Modèle AR	139
7.2.2 Paramètres de reconstruction	140
7.3 OBSERVATION DES INDICATEURS	141
7.3.1 Coefficients AR	141
7.3.2 JFV	142
7.4 DÉTECTION DE L'ENDOMMAGEMENT	143
7.4.1 Conditions environnementales stables	143
7.4.2 Conditions environnementales variables	146
7.4.3 Augmentation de la dimension	146
7.4.4 Autre méthode : corrélation croisée	149
BILAN	153
8 MODÉLISATION STATISTIQUE ET SEUIL DE CLASSIFICATION	155
INTRODUCTION	156
8.1 LES DIFFÉRENTES MÉTHODES TESTÉES	157
8.2 DONNÉES UTILISÉES	158
8.3 RÉSULTATS	159
BILAN	165
Conclusion générale	167
PERSPECTIVES	171
A ANNEXES	173

A.1	TABLEAU CHARGEMENT DE LA POUTRE EN BÉTON FIBRÉ . .	174
A.2	RÉSULTATS MAQUETTE EN BOIS	175
A.2.1	Coefficients AR	175
A.2.2	Composantes JFV	181
A.2.3	Corrélation croisée	187
A.3	MODÉLISATION STATISTIQUE	189

BIBLIOGRAPHIE	191
---------------	-----

INTRODUCTION GÉNÉRALE

DANS le contexte du suivi vibratoire de l'état de santé des structures de génie civil, la détection de l'endommagement est confrontée à trois problèmes principaux :

- le caractère unique de l'ouvrage,
- les effets des variations des conditions opérationnelles et environnementales sur la dynamique vibratoire,
- une sollicitation dynamique complexe, inconnue, non mesurable et extrêmement variable d'une mesure à l'autre.

Pour faire face à ces difficultés, la démarche de *reconnaissance de formes statistiques* est de plus en plus mise en avant. Son principe général consiste à modéliser les données, et non le comportement dynamique de la structure physique. Une telle approche revêt un double intérêt :

- éviter de formuler des hypothèses a priori sur le comportement dynamique, hypothèses résultant de modèles plus ou moins fidèles à la réalité de l'ouvrage,
- intégrer tous les facteurs de variabilité.

Parmi les différentes approches de reconnaissance de formes statistiques, la *détection de nouveauté* est plus particulièrement adaptée à un apprentissage non supervisé, inévitable sur un ouvrage au caractère unique pour lequel la signature vibratoire dans un état endommagé est très difficile, voire impossible à appréhender a priori.

La détection de nouveauté s'appuie tout d'abord sur la définition d'*indicateurs* estimés sur chaque mesure. Un indicateur condense l'information véhiculée par une mesure (signal vibratoire ou, de façon plus générique, série temporelle) en un vecteur de dimension réduite. Un indicateur idéal serait sensible à l'endommagement de la structure et insensible aux variations environnementales et opérationnelles (conditions normales variables). Hélas, un indicateur révèle un comportement dynamique de façon générale et sa sensibilité aux modifications de cette dynamique, si elle existe, traduit, de fait, les effets d'un endommagement et des variations environnementales. La détection de nouveauté passe alors par une étape dite de *normalisation* consistant à situer et comparer chaque indicateur nouveau (à tester) par rapport à une base de référence constituée d'indicateurs relatifs à l'état sain de l'ouvrage. Enfin, selon le résultat de l'étape de normalisation, un modèle statistique permet de décider, avec un niveau de confiance donné, de la classe d'appartenance de la mesure à tester : *classe associée à la structure saine* (la mesure ressemble aux mesures de la base de référence) ou *classe associée à la structure endommagée* (la mesure apparaît comme nouvelle). Il s'agit là de l'étape de *classification*.

En marge des développements importants réalisés dans le domaine de l'analyse modale pour le suivi des propriétés dynamiques classiques, de nouveaux types d'indicateurs sont apparus sur la base d'*approches sans modèle*¹ de l'interprétation des signaux en contexte de monitoring vibratoire. Ces nouveaux types d'indicateurs sont fondés sur des paradigmes divers. En raison de sa simplicité et de son efficacité, l'indicateur vectoriel le plus répandu est formé par les coefficients de modèles auto-régressifs ajustés sur les mesures. Par ailleurs, la théorie des systèmes dynamiques non linéaires apporte de nouveaux outils pour l'analyse des systèmes complexes au travers des signaux qu'ils génèrent. Elle s'appuie notamment sur la reconstruction de la trajectoire du système dans son espace des phases et laisse le champ libre à la définition d'indicateurs variés et originaux.

De nombreuses études mettent en œuvre avec succès la détection de nouveauté pour déceler l'occurrence de dégradations sur des dispositifs expérimentaux. Cependant, le transfert de ces résultats du laboratoire aux ouvrages in situ pose encore problème. La principale raison réside dans la prise en compte insuffisante des variations environnementales et de la nature potentiellement instationnaire et extrêmement variable de la source d'excitation dynamique.

Le premier objectif du travail présenté dans ce mémoire de thèse consiste à observer dans quelle mesure la capacité à détecter des dégradations est altérée par de fortes variations environnementales, mais également par une excitation de nature impulsionnelle. Pour cela, plusieurs dispositifs expérimentaux sont utilisés. Deux de ces dispositifs sont originaux et garantissent, à la fois, le contrôle de l'excitation et la possibilité de faire varier de façon importante les conditions ambiantes.

Le deuxième objectif est de proposer et d'étudier un indicateur original (*JFV* pour *Jacobian Feature Vector*) calculé dans l'espace des phases reconstruit. Pour juger de ses performances en matière de caractérisation de l'état de la structure, il est comparé aux coefficients de modèles auto-régressifs.

STRUCTURE DU DOCUMENT

Ce mémoire est segmenté selon *3 parties* et *8 chapitres*. La première partie, constituée des deux premiers chapitres, présente les éléments contextuels et bibliographiques associés à ces travaux doctoraux.

Le *premier chapitre* du document dresse un état de l'art non exhaustif sur le contexte général du suivi vibratoire de l'état de santé des structures et sur les difficultés rencontrées par les approches classiques. Le but de ce chapitre est de montrer l'intérêt que représente la détection de nouveauté, dont le principe et les différentes étapes sont énoncés.

Constituant le socle théorique associé à la définition de l'indicateur original proposé (*JFV*), les concepts élémentaires de la théorie des systèmes dynamiques non linéaires sont introduits dans le *chapitre 2*, suivis de quelques exemples d'utilisation en matière de contrôle des structures.

Les détails techniques et pratiques des développements proposés sont regroupés dans la partie II, constituée des *chapitres 3 et 4*. Les détails des

¹*Model-free approach*

calculs relatifs aux différents objets conceptuels et méthodes abordés (indicateurs, choix des paramètres, normalisation, modèles statistiques) sont présentés dans le *chapitre 3*. Le *chapitre 4* décrit précisément l'ensemble des cas d'études expérimentaux dont sont extraites les données analysées.

La troisième et dernière partie du document, constituée des *chapitres 5 à 8*, s'attache à montrer, pour des niveaux de difficulté croissants, les capacités et les limites des indicateurs et méthodes choisis en matière de détection d'endommagement.

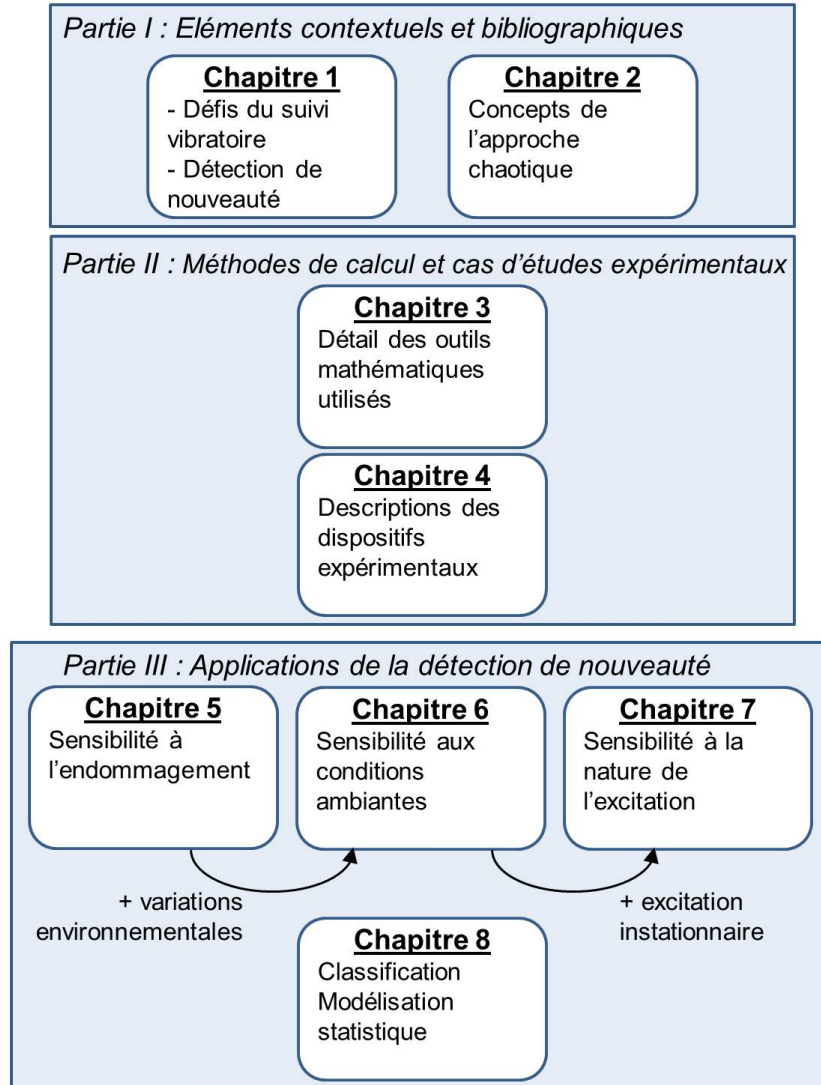
Dans le *chapitre 5*, l'excitation est stationnaire et les conditions ambiantes constantes. Par le biais de données simulées issues d'un système masse/ressort amorti et de données réelles provenant d'une maquette métallique, une étude préliminaire permet d'établir l'intérêt du nouvel indicateur proposé et d'en fixer les modalités de calcul. Le JFV est ensuite comparé à un indicateur plus conventionnel (coefficients d'un modèle auto-régressif) sur des données collectées sur une maquette de structure en bois à trois niveaux, subissant différentes séquences d'endommagement progressif et réversible.

L'effet des variations environnementales sur la capacité de détection d'endommagement est abordé dans le *chapitre 6* au travers de deux essais originaux conduits sur une poutre en béton fibré endommagée par des cycles de chargement mécanique et sur la maquette en bois précédemment citée. L'impact des conditions climatiques sur la variabilité des indicateurs et la capacité des méthodes à détecter l'endommagement sous variations environnementales sont étudiées. Ce chapitre aborde également le problème, souvent éludé, de la répartition des données de référence pour la constitution des différentes bases nécessaires à la mise en œuvre de la détection de nouveauté en contexte très variable.

L'effet de la nature de l'excitation est analysé dans le *chapitre 7*. Trois excitations différentes sont utilisées dans le cadre d'expérimentations réalisées sur la maquette en bois : un bruit de bande de fréquence limitée, une séquence aléatoire d'impulsions et une forme d'excitation hybride combinant les deux formes précédentes. Les conséquences de ces formes de sollicitation sur la sensibilité à l'endommagement sont observées tout d'abord séparément, puis combinées aux effets des variations environnementales,

Le 8^{ème} et dernier chapitre se focalise sur la détermination du seuil de classification, fixé jusqu'ici de façon arbitraire. Les méthodes statistiques permettant de guider le choix de cette limite de classification sont multiples. Dans ce chapitre, sept méthodes sont testées. Elles sont basées sur des modèles statistiques ajustés soit sur les valeurs individuelles des données, soit sur des échantillons de données. Pour réaliser la comparaison de ces méthodes, les taux de fausses alarmes et de vraies détections sont calculés sur 5 jeux de données issus des essais décrits dans les chapitres précédents.

Enfin, ce document s'achève par un chapitre de conclusion synthétisant les principaux résultats issus de ces travaux et définissant diverses perspectives de recherche envisageables. Parmi ces perspectives, un projet de recherche collaborative est proposé par le biais d'un benchmark s'appuyant sur les nombreux essais réalisés sur la maquette en bois sous variations environnementales sévères et sous sollicitations de natures diverses.

FIG. 1 *Structure du mémoire.*

Première partie

Éléments contextuels et
bibliographiques

DÉTECTION DE NOUVEAUTÉ EN CONTEXTE DE SHM



SOMMAIRE	
INTRODUCTION	8
1.1 PROBLÉMATIQUE DES STRUCTURES DE GÉNIE CIVIL	9
1.1.1 Définitions	9
1.1.2 Particularités des structures de génie civil	10
1.1.3 Bilan	13
1.2 PARADIGME DE LA RECONNAISSANCE DE FORMES STATIS- TIQUES (RFS)	14
1.3 DÉTECTION DE NOUVEAUTÉ	16
1.3.1 Indicateurs	16
Modèles auto-régressifs et auto-régressifs à moyenne mobile	17
Modèles AR-ARX	18
Modèles AR non linéaires	19
Indicateurs modaux	19
La transformation en ondelettes	20
Caractérisation dans l'espace des phases	21
1.3.2 Normalisation	21
Les estimateurs de densité	22
Méthodes de reconstruction	24
Méthodes frontières	28
Introduction des données environnementales	32
1.3.3 Seuil de rejet	33
Cartes de contrôle	33
Test d'hypothèse	35
Statistique des valeurs extrêmes	36
Test séquentiel du rapport des vraisemblances	39
BILAN	42

INTRODUCTION

LE suivi de l'état de santé des structures de génie civil en conditions réelles représente toujours un vrai défi scientifique. En effet, la connaissance imparfaite des caractéristiques physiques et géométriques des ouvrages existants, les effets des variations climatiques et la nature des sollicitations ambiantes sont autant de difficultés pour la mise au point d'outils robustes d'analyse des mesures vibratoires.

L'objet de ce chapitre est de montrer comment la méthode de détection de nouveauté peut apporter des éléments de réponse face à ces problèmes. Pour cela, le contexte, les difficultés et les objectifs du suivi de l'état de santé des ouvrages sont rapidement présentés avant d'introduire la démarche de reconnaissance de formes statistiques dont fait partie la détection de nouveauté.

1.1 PROBLÉMATIQUE DES STRUCTURES DE GÉNIE CIVIL

1.1.1 Définitions

Un ouvrage de génie civil présente une durée de vie de 50 à 100 ans selon son importance. Durant cette période, son niveau de *performance* doit lui permettre d'assurer sa fonction en garantissant la sécurité des usagers.

Cependant, son comportement peut être altéré au cours du temps par l'apparition de dégradations, qui peuvent avoir plusieurs origines.

- L'environnement extérieur modifie progressivement les caractéristiques physico-chimiques des matériaux. L'exemple le plus courant est celui du béton armé qui présente souvent des réductions de section des armatures et des épaufrures consécutives au développement de la corrosion de l'acier. La corrosion résulte soit de la carbonatation du béton d'enrobage liée à la pénétration du CO₂ atmosphérique, soit de la diffusion des ions chlorures. Ce dernier cas concerne particulièrement les ouvrages proches des façades maritimes ou situés dans les pays froids très consommateurs de sels de déverglaçage. L'altération du matériau peut avoir également une origine endogène comme la réaction alcali-granulat.
- Des actions accidentelles comme un impact de véhicule sur une pile de pont ou un évènement naturel important (tempête, séisme) sont susceptibles de modifier l'intégrité de la structure.
- L'usure et le dysfonctionnement de composants tels que les appareils d'appuis ou les joints de dilatation peuvent affecter le comportement global de l'ouvrage.

Toutes les dégradations n'ont pas les mêmes conséquences. [Worden et Dulieu-Barton \(2004\)](#) proposent trois définitions pour hiérarchiser les niveaux de dégradation :

- *Défaillance*¹ : la structure ne peut plus fonctionner de façon satisfaisante. Si la qualité est définie comme l'aptitude de l'ouvrage à remplir sa fonction ou à satisfaire les besoins des usagers, la défaillance se définit alors comme un changement qui produit une réduction inacceptable de la qualité.
- *Endommagement* : la structure ne fonctionne plus dans des conditions idéales, mais peut toujours assurer un service satisfaisant d'une façon sous-optimale.
- *Défaut* : il est inhérent aux matériaux et n'empêche pas un fonctionnement optimal.

Pour des raisons socio-économiques évidentes, il convient donc de disposer d'*indicateurs* de performances capables de détecter l'apparition d'un endommagement avant qu'il n'évolue en défaillance. Ces indicateurs peuvent provenir de la fusion d'informations issues de sources variées allant des mesures dynamiques (vitesse, accélération, chargement) ou quasi-statiques (déformations, inclinaison) aux simples relevés d'inspection visuelle.

L'ensemble de ces indicateurs doit permettre de caractériser la *santé* de l'ouvrage ([Crémona 2009](#)) à un instant donné. La comparaison avec des

¹Fault

niveaux limites, conjuguée à un regard d'expert, permet alors d'appréhender le niveau de performance de la structure et d'en planifier la maintenance.

Le *suivi de l'état de santé des structures* (SHM²) est défini par l'ensemble du processus suivant :

- définition des objectifs du système de suivi,
- détermination des caractéristiques du système,
- collecte d'informations,
- extraction d'indicateurs pertinents sur l'état de santé,
- diagnostic (évaluation de l'état présent),
- pronostic (prédiction de l'état futur).

Il existe différents niveaux d'exigence sur les performances du dispositif de surveillance. Ces niveaux sont numérotés de 1 à 5 par Rytter (1993) en fonction des objectifs visés :

1. détecter l'apparition de l'endommagement,
2. localiser l'endommagement,
3. déterminer le type d'endommagement,
4. quantifier la sévérité de l'endommagement,
5. estimer la durée de vie résiduelle.

Après cette présentation rapide de la notion générale de suivi de l'état de santé des structures, la suite de la discussion est axée sur l'analyse des mesures dynamiques et plus particulièrement des mesures vibratoires.

1.1.2 Particularités des structures de génie civil

L'approche la plus répandue d'extraction d'indicateurs des mesures vibratoires est *l'identification de systèmes*. Son objectif est d'identifier le *modèle modal* de la structure, composé essentiellement des fréquences propres, des modes propres, des taux d'amortissement et des facteurs de participation modale. La recherche de l'endommagement passe donc par la détection soit de changements dans les caractéristiques du modèle, soit du moment où le modèle ne permet plus de représenter le comportement vibratoire observé. Les méthodes de mise à jour de modèles éléments finis permettent, par analyse inverse, de localiser et quantifier l'endommagement. La littérature est très abondante sur le sujet. Un aperçu des différents indicateurs basés sur les modèles modaux est présenté dans Doebling et al. (1996; 1997) et Carden et Fanning (2004).

Cependant, malgré les progrès considérables réalisés dans le perfectionnement des méthodes d'identification de ces modèles (Basseville et al. 2001, Peeters et De Roeck 2001, Alvandi 2003), celles-ci se heurtent toujours aux particularités des structures de génie civil lorsqu'elles sont transposées des laboratoires aux ouvrages réels.

Variations des conditions environnementales

Contrairement à de nombreux systèmes mécaniques, les structures de génie civil subissent des variations climatiques à la fois quotidiennement et

²Structural Health Monitoring

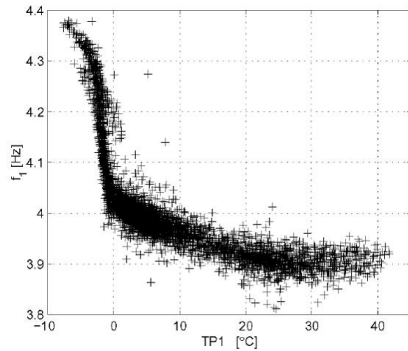


FIG. 1.1 Corrélation entre la 1^{re} fréquence propre et la température sur le pont Z-24 (Peeters et De Roeck 2001)

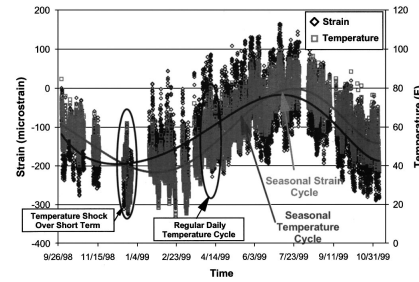


FIG. 1.2 Relevé des déformations d'un tablier au cours d'une année (Catbas et Aktan 2002)

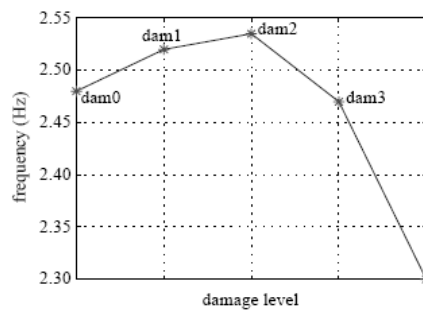


FIG. 1.3 Variation de la fréquence fondamentale en fonction de la profondeur de l'entaille, pont I-40 Farrar et al. (2000)

annuellement. Dans le cas des ponts, les contrastes thermiques entre la face supérieure du tablier et la face inférieure peuvent atteindre 30°C et affecter de façon significative le comportement vibratoire. Bien que marginales, les structures en bois sont en plus très sensibles aux variations hygrométriques.

Des études portant sur la détection d'endommagement sur des ponts réels ont montré que les caractéristiques modales sont très sensibles aux variations climatiques. Le suivi du pont d'Alamosa Canyon sur 24 heures a montré des variations des fréquences propres de l'ordre de 5% (Farrar et al. 2000). Des écarts de 3% ont été mesurés sur le pont haubanné Tianjin Yonghe en Chine sur une durée d'écoute de deux semaines (Li et al. 2009). Le projet du pont Z-24 en Suisse a vu certaines fréquences propres changer de 15% au cours d'une année (Peeters et De Roeck 2001) avec une relation bilinéaire entre la température et la première fréquence propre (Fig.1.1). Cette relation est expliquée par un changement du module d'élasticité de l'asphalte au-dessous de 0°C. Ces variations peuvent également être expliquées par une modification des conditions de liaisons consécutive aux variations dimensionnelles du tablier lorsque les joints de dilatation ne fonctionnent pas correctement.

Or, ces variations sont bien souvent supérieures à celles engendrées par l'apparition d'un endommagement. Le pont I-40 (Farrar et al. 2000) a été endommagé par une entaille de profondeur croissante réalisée sur une poutre. La figure 1.3 montre la variation de la première fréquence propre en fonction de la profondeur de l'entaille. L'évolution observée reflète davantage les

variations des conditions environnementales que de la présence de l'entaille. Le problème du filtrage de ces variations apparaît alors comme crucial. Il sera abordé dans le chapitre 6.

Algorithme de calcul et nature de l'excitation

Les premières méthodes d'analyse modale expérimentale exploitent les fonctions de réponse en fréquence (FRF) qui nécessitent la connaissance de l'excitation. Les méthodes ERA³ (Juang et Phan 2001) et PDT⁴ se basent sur la réponse impulsionnelle de la structure. Cependant, la mise en place de systèmes de monitoring vibratoire sur des ouvrages réels impose des outils d'analyse capables de fonctionner en présence d'une excitation aléatoire et non mesurée.

Il existe aujourd'hui de nombreuses méthodes de calcul des caractéristiques modales fonctionnant uniquement à partir de l'information de sortie (réponse). Les plus répandues, en raison de leur simplicité, sont les algorithmes de *relevé des pics*⁵ et de la CMIF⁶. Ils sont tous les deux basés sur le calcul de la matrice du spectre de puissance et requièrent un relevé manuel des fréquences propres, ce qui constitue un frein à leur utilisation dans un processus de surveillance automatisé.

Pour s'accommoder du caractère aléatoire de l'excitation, les méthodes du type *stochastiques sous-espaces* (SSI) basées soit sur la matrice de covariance des signaux mesurés, soit directement sur les signaux, sont assez populaires (Basseville et al. 2001, Peeters 2000).

L'étude menée sur le pont d'*Alamosa Canyon* par Farrar et al. (2000) a montré que les écarts dans les paramètres identifiés selon différentes méthodes de calcul sont faibles mais peuvent être de l'ordre de grandeur des variations engendrées par un endommagement.

Une autre difficulté réside dans l'automatisation de ces algorithmes pour l'analyse de grandes quantités de mesures (Giraldo et al. 2006). L'article de Peeters et De Roeck (2001) compare les méthodes pré-citées sur une tour d'antenne. Il en ressort que les algorithmes SSI permettent une meilleure identification en présence de bruit et de modes proches. Néanmoins, les modes identifiés n'ont pas toujours de sens physique. Il faut alors avoir recours à des diagrammes de stabilisation, des critères de sélection (matrice MAC⁷) et l'avis d'un expert pour sélectionner les vrais modes physiques. Ces inconvénients constituent un problème majeur dans un contexte de suivi continu automatique.

Par ailleurs, la plupart de ces algorithmes supposent une excitation stationnaire, ce qui ne représente pas la réalité d'un ouvrage soumis à des sollicitations ambiantes. Dans le cas des ponts, la circulation des véhicules et le vent génèrent une excitation très irrégulière formée de trains d'impulsions aléatoires. La figure 1.4 présente un accélérogramme collecté sur

³ *Eigen Realization Algorithm*

⁴ *Polyreference Time Domain*

⁵ *Peak Picking*

⁶ *Complex Mode Indicator Function*

⁷ *Modal Assurance Criterion*

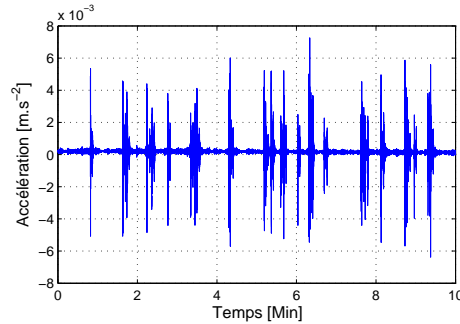


FIG. 1.4 Exemple de mesure d'accélération sur le pont Z-24 (Krämer et al. 1999)

le pont Z-24. S'il est possible d'estimer les fréquences propres, ce type de signal rend très difficile l'estimation des modes et des amortissements.

Enfin, en règle générale, seuls les premiers modes (moins de 5) peuvent être identifiés avec une précision correcte. Les premiers modes capturent en effet l'essentiel de l'énergie vibratoire et plus les fréquences propres augmentent, moins les modes associés sont visibles. Or, l'apparition d'endommagement à l'échelle locale affecte d'avantage les modes de hautes fréquences que les premiers modes plus représentatifs du comportement global de la structure.

Masse variable

Les caractéristiques modales d'un système dynamique dépendent en partie de sa masse. Or, en fonction de la circulation, le chargement imposé à un ouvrage varie au cours du temps. Le modèle identifié aux heures de pointe est alors différent de celui identifié en pleine nuit. Les écarts ainsi produits sont une fois de plus, susceptibles de masquer l'effet d'un endommagement. Kim et al. (2003) montrent que sur un pont mixte de 46 mètres de portée, les fréquences propres présentent une diminution de 5.4% aux heures de pointe. Cet effet semble se réduire avec l'augmentation de la portée.

1.1.3 Bilan

La majorité des points abordés dans cette partie sont extraits du travail bibliographique de Sohn (2007). Il apparaît que la détection d'endommagement par l'identification des caractéristiques modales des structures de génie civil demeure un problème non résolu pour des raisons multiples : absence de connaissances de tous les facteurs environnementaux influençant l'ouvrage, impossibilité de mesurer l'excitation et difficultés à modéliser des structures aussi complexes.

Face à ces difficultés, la question de la nécessité de recourir à un modèle de la structure peut se poser. Le travail de Haritos et Owen (2004) compare deux approches pour détecter l'endommagement sur des ponts de tailles réduites réalisés en laboratoire. La première approche met en œuvre une technique de mise à jour d'un modèle éléments finis de la structure. Il est estimé à partir des fonctions de réponse en fréquence normalisées par rapport

à un capteur de référence (RRF⁸). La seconde ne modélise pas la structure. Elle s'appuie sur le paradigme de la reconnaissance des formes statistiques (qui sera présenté dans la partie suivante) en utilisant les RRF en données d'entrée. L'étude montre que les deux approches sont complémentaires.

La première permet de localiser et quantifier l'endommagement, mais au prix d'une quantité importante de mesures de bonne qualité. En effet, pour situer l'endommagement un maillage de capteurs assez dense est nécessaire. Elle nécessite également une grande expertise pour désigner avec justesse les paramètres pertinents à mettre à jour.

La seconde, est seulement capable de détecter l'apparition de l'endommagement, mais avec assez peu de mesures. De plus, aucune hypothèse n'est formulée a priori sur l'ouvrage et la méthode peut être rapidement transposée à l'étude d'une autre structure.

Pour ces dernières raisons, l'étude présentée dans ce rapport porte sur la mise en œuvre de la reconnaissance des formes statistiques.

1.2 PARADIGME DE LA RECONNAISSANCE DE FORMES STATISTIQUES (RFS)

Du fait de la quasi-impossibilité de prendre en compte tous les facteurs influençant les indicateurs estimés, la détection de l'endommagement dans un ouvrage de génie civil se pose de plus en plus comme un problème de reconnaissance de formes (ou motifs) statistiques⁹. Le terme *forme* désigne ici la structure cachée des données analysées. Cette approche ne représente qu'une façon différente d'appréhender les données en s'appuyant sur de nombreux outils mathématiques dont certains sont présentés dans les parties suivantes.

D'après [Farrar et Worden \(2007\)](#), le suivi de l'état de santé des structures par la RFS comporte 4 grandes étapes :

Etape 1 : L'évaluation opérationnelle (réflexion sur l'objectif de l'instrumentation, les contraintes et les moyens à mettre en œuvre),

Etape 2 : L'acquisition des données, la normalisation et le nettoyage,

Etape 3 : La réduction des données (étape centrale qui donne lieu au calcul des indicateurs de l'endommagement),

Etape 4 : La classification (modèles statistiques ajustés sur les indicateurs et permettant de déterminer l'état de la structure avec un niveau de confiance donné).

La première étape relève davantage des compétences du gestionnaire d'ouvrage qui doit choisir, en fonction de ses impératifs économiques, les objectifs et les moyens de son système de surveillance. La seconde est liée au domaine du traitement du signal (suppression des mesures aberrantes, réduction du bruit de mesure par filtrage, etc.). Les deux dernières étapes associées à l'analyse et à l'exploitation des données constituent le sujet de ce travail de thèse.

⁸ *Relative Response Function*

⁹ *Statistical Pattern Recognition*

La réduction des données a pour principal objectif de condenser l'information contenue dans les mesures. Elle aboutit, pour chaque mesure, au calcul d'un indicateur vectoriel de dimension faible représentant un *individu*. L'étape suivante est la classification. Elle associe à chaque *individu* (indicateur vectoriel) une *classe* correspondant à un état de santé structurale avec un niveau de confiance donné. Le problème mathématique et l'algorithme de classification adéquat diffèrent selon le degré de connaissance sur l'état endommagé de l'ouvrage (mesures sur les différentes classes) (Hodge et Austin 2004).

Aucune classe n'est connue. Les individus doivent être classés en plusieurs groupes relatifs à différents états structuraux sans information préalable. Le problème est alors du type *regroupement*¹⁰ non-supervisé. La difficulté est de garantir que les groupes formés soient réellement liés à différents niveaux d'endommagement, et non aux conditions environnementales. Cury et al. (2010) montrent l'aptitude de plusieurs algorithmes de regroupement *non-supervisés* à séparer correctement en trois classes les données vibratoires collectées sur un pont rail (avant, pendant et après renforcement de la structure), et par conséquent, à suivre avec succès une série de modifications structurales.

Les classes sont connues. Pour chaque classe, ou scénario d'endommagement, des individus sont disponibles. Il s'agit alors d'un problème de *régression* ou de *classification supervisée*. Après une phase d'apprentissage fondée sur les données disponibles, l'algorithme entraîné doit être en mesure d'associer une classe à chaque nouvel individu présenté en entrée. Hormis la difficulté liée à la disponibilité des données, la phase d'apprentissage est assez délicate et conditionne l'efficacité de la classification (Cury et Crémona 2010). En cas de sur-apprentissage, l'algorithme perd sa capacité de généralisation à de nouvelles données, on parle alors d'*apprentissage par cœur*, sans compréhension de la structure des données.

Une seule classe est connue. Seuls des individus associés à la structure saine sont disponibles. L'objectif est d'apprendre à les reconnaître. Comme les données d'une seule classe (*structure non endommagée*) sont utilisées pour l'apprentissage, celui-ci est dit *semi-supervisé*. Le problème est qualifié de *détection de nouveauté* puisqu'il consiste à rechercher les individus qui diffèrent notablement de ceux composant la base d'apprentissage (Worden et al. 2000).

Ce dernier cas de figure correspond à la problématique du monitoring des structures de génie civil par la RFS. En effet, en dehors d'une approche de modélisation, il est impossible de disposer a priori de mesures du comportement endommagé. En revanche, il est facile d'enregistrer le comportement non-endommagé d'un ouvrage sur une certaine période, suffisamment longue pour intégrer l'ensemble des variations des conditions environnementales et opérationnelles en fonctionnement normal.

L'article de Worden et al. (2007) propose 7 axiomes concernant le suivi de l'état de santé des structures, basés sur l'analyse des dernières décennies de recherche dans ce domaine. Les auteurs estiment qu'il est possible de détecter l'apparition d'endommagement et sa position par une approche non-supervisée (ou semi-supervisée), mais que le type et la sévérité de l'en-

¹⁰ *clustering*

dommagement ne peuvent être obtenus qu’avec des méthodes supervisées. Puisque que l’approche testée dans ce travail de thèse exclut le recours à toute modélisation du comportement de l’ouvrage, seuls les objectifs 1 et 2 de Rytter (1993) (détection et localisation) peuvent être envisagés.

Parmi les différentes approches constituant la RFS, la *détection de nouveauté* semble la plus adaptée pour détecter et, éventuellement, localiser l’apparition de l’endommagement.

1.3 DÉTECTION DE NOUVEAUTÉ

La détection de nouveauté regroupe de nombreux algorithmes d’apprentissage utilisés dans le domaine de l’intelligence artificielle. Loin de se limiter au SHM, ses champs d’applications sont multiples et en progression constante : reconnaissance de forme (écriture, voix, imagerie médicale), détection de comportement anormaux (télécommunications, banques, informatique), etc. Plusieurs articles passent en revue les différents outils répondant à ces problèmes. Chandola et al. (2009) présentent une synthèse bibliographique insistant sur la spécificité de chaque domaine d’application. Hodge et Austin (2004) préfèrent une présentation en fonction du problème mathématique posé, et un classement basé sur les outils opérationnels, à l’instar de Markou et Singh (2003b;a).

Le processus général mis en œuvre systématiquement dans ces travaux de thèse pour détecter l’endommagement à partir des mesures vibratoires comporte 3 étapes :

- calcul des indicateurs,
- normalisation,
- fixation du seuil de rejet.

Les outils mathématiques relatifs à chacune de ces 3 étapes sont très nombreux. Dans la suite de cette partie, seuls les outils les plus employés sont succinctement présentés et illustrés. Les outils spécifiquement exploités dans ce travail de thèse sont détaillés dans le chapitre §3.

1.3.1 Indicateurs

Les bases de données se présentent sous forme d’ensembles de séries temporelles (une série par capteur, généralement des accélérogrammes en monitoring vibratoire) collectées à différents instants. Les quantités de mesures collectées génèrent des bases de données de grandes dimensions qu’il convient de réduire tout en limitant la perte d’information sur le comportement de la structure. Pour cela, différents indicateurs scalaires ou vectoriels peuvent être extraits de chaque individu de la base de données. Un individu est constitué soit par une série temporelle, soit par l’ensemble des séries collectées par différents capteurs au même instant. L’indicateur scalaire ou vectoriel est supposé résumer l’information relative à l’individu dont il est extrait. Idéalement, un indicateur doit varier de façon significative avec un endommagement croissant.

La méthode générale d’extraction d’indicateurs consiste tout d’abord à ajuster un modèle sur les séries temporelles de la base de données de référence. Il est alors possible de proposer différents indicateurs caractérisant un

individu tels que l'erreur de prédiction ou encore simplement les paramètres du modèle ajusté sur l'individu.

Modèles auto-régressifs et auto-régressifs à moyenne mobile

Ces modèles, très utilisés dans le domaine de l'analyse des séries temporelles, permettent une prédiction des valeurs futures d'une série à l'aide d'une combinaison linéaire de ses valeurs passées. Les séries sont considérées comme des processus stochastiques linéaires, c'est à dire qu'elles représentent un bruit blanc filtré par un système linéaire. Le comportement de la structure doit théoriquement être linéaire et la vibration stationnaire. Malgré tout, ces modèles tolèrent dans une certaine mesure la violation de ces hypothèses.

Par ailleurs, il est possible, à partir des paramètres d'un modèle ARMA¹¹ de calculer les fréquences et modes propres (Peeters et De Roeck 2001).

Un modèle auto-régressif (AR) à moyenne mobile (MA) ARMA(n_x, n_e) est de la forme :

$$x_i = \sum_{j=1}^{n_x} a_j \cdot x_{i-j} + \sum_{j=1}^{n_e} b_j \cdot \epsilon_{i-j} + \epsilon_i \quad (1.1)$$

où x_i est la valeur estimée du $i^{\text{ème}}$ terme de la série x , a_j sont les coefficients AR, n_x est l'ordre auto-régressif (nombre de points passés utilisés), b_j sont les coefficients MA, n_e est l'ordre de la moyenne mobile et ϵ_i est le terme d'erreur supposé aléatoire, de moyenne nulle et de variance constante. Ce type de représentation permet de dissiper rapidement les effets de l'erreur (vue comme une perturbation aléatoire du système) au travers de la moyenne mobile, ou d'en conserver une trace plus longtemps grâce à la partie auto-régressive.

Les coefficients AR sont calculés simplement par une méthode de moindres carrés alors que les coefficients MA nécessitent des algorithmes d'optimisation non linéaire (Brockwell et Davis 1991, méthode de Box-Jenkins). C'est probablement pour des raisons de simplicité que les modèles simplement auto-régressifs sont majoritairement utilisés ($n_e = 0$) pour des applications de suivi de l'état de santé des structures.

$$x_i = \sum_{j=1}^{n_x} a_j \cdot x_{i-j} + \epsilon_i \quad (1.2)$$

Certains travaux exploitent directement les variations des coefficients AR (a_j) ajustés sur chaque série comme indicateurs (Garcia et Osegueda 2000, Nair et al. 2006). Dans Gul et Catbas (2009), une poutre sur deux appuis est étudiée. Un modèle AR d'ordre 3 est ajusté sur la réponse impulsionnelle estimée par la méthode du décrément aléatoire. Figueiredo et al. (2009) montrent que l'amplitude des coefficients AR diminue avec l'apparition de non linéarité qui est souvent un signe d'endommagement de la structure (Fig.1.5).

Il est également possible d'étudier l'évolution de l'erreur de prédiction du modèle AR. En effet, un modèle AR ajusté sur des données associées à un

¹¹Auto-Regressive Moving Average

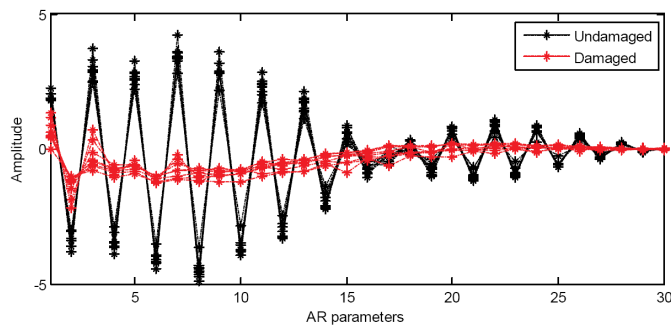


FIG. 1.5 Coefficients AR calculés sur les accélérations d'une maquette métallique dans laquelle l'endommagement est introduit par des chocs (Figueiredo et al. 2009)

état structural sain devrait produire une erreur forte sur la simulation d'une mesure relative à un état structural dégradé. Fugate et al. (2000) utilisent les outils de la maîtrise statistique des procédés¹² pour suivre les variations de l'erreur de prédiction commise sur des mesures issues d'une colonne en béton armé endommagée par des charges statiques. Cependant, lorsque le comportement sain de la structure est enregistré sous différentes conditions environnementales et opérationnelles pour former la base de données de référence, il n'y a pas unicité des coefficients AR. Il est alors difficile de choisir la série dont les coefficients seront utilisés pour prédire les nouvelles mesures. Sohn et Farrar (2001) proposent d'utiliser la série la plus proche dans l'espace des coefficients AR. Même si les résultats sont concluants, le choix de la série la plus proche doit être comparé à un choix aléatoire pour prouver le bénéfice réel de la méthode.

La plupart des études proposées montrent que l'apparition de l'endommagement est correctement détectée grâce à ce type d'indicateurs. Cependant, rares sont celles ayant porté sur des données vibratoires de pont en conditions ambiantes. Seuls Cheung et al. (2008) utilisent les coefficients AR lors de l'étude du pont Z-24 qui a subi différents scénarios d'endommagement. Mais dans ce cas également, il est difficile de savoir si les changements détectés ont pour origine les dégradations ou les variations environnementales et/ou opérationnelles qui ne sont pas mesurées.

Modèles AR-ARX

Les modèles ARX sont constitués d'une partie auto-régressive et d'une partie dite *exogène* correspondant à une combinaison linéaire des termes de l'excitation. La forme générale de ces modèles est donnée par l'équation 1.3, dans laquelle y_{i-j} sont les termes de la série sollicitant la structure pondérés par les coefficients c_j et utilisés pour prédire le terme x_i .

$$x_i = \sum_{j=1}^{n_x} a_j \cdot x_{i-j} + \sum_{j=1}^{n_y} c_j \cdot y_{i-j} + \epsilon_i \quad (1.3)$$

Cependant, dans le cadre d'une surveillance sous sollicitations ambiantes, l'excitation n'est pas connue. Dans plusieurs travaux de recherche,

¹²Statistical Process Control

l'erreur ϵ d'un modèle AR est supposée être liée à l'excitation. Ainsi, après avoir ajusté un modèle AR sur une série, les auteurs utilisent cette erreur comme excitation ($y_{ARX} = \epsilon_{AR}$) pour générer un modèle ARX sur chacune des mesures. Une nouvelle mesure (état de la structure inconnu) est ensuite prédite à l'aide des coefficients ARX d'une des séries de la base de référence. L'erreur de prédiction est utilisée comme indicateur. La difficulté du choix de la série à prendre comme référence est traitée par différentes méthodes présentées dans la suite du rapport (ACP à noyau, [Oh et Sohn \(2009\)](#); réseau de neurones auto-associatifs, [Sohn et al. \(2002; 2005\)](#); distance euclidienne, [Sohn et Farrar \(2001\)](#)). Tous ces travaux portent sur un système masse-ressort à 8 degrés de liberté dans lequel une butée génère de la non-linéarité synonyme d'endommagement.

Modèles AR non linéaires

Dans le cas de processus non linéaires, les modèles précédents ne sont plus appropriés. [Bornn et al. \(2009\)](#) ajustent un modèle AR non linéaire en utilisant la théorie des Machines à Vecteurs de Support (AR-SVM) issue de l'intelligence artificielle. Les SVM sont alors utilisés à des fins régressives ([Smola et Schölkopf 2004](#)).

La combinaison linéaire réalisée dans un modèle AR est transformée en une combinaison linéaire de fonctions noyaux (fonction symétrique positive), ce qui permet de représenter des relations non linéaires entre les points de la série temporelle.

$$k(x, y) = \exp\left(-\frac{\|x - y\|^2}{2\sigma^2}\right) \quad (1.4)$$

$$x_i = f(x_{i-p:i-1}) = \sum_{j=p+1}^{t_0} \beta_j \cdot k(x_{j-p:j-1}, x_{i-p:i-1}) \quad (1.5)$$

avec le vecteur $\{x_{i-p}, \dots, x_{i-1}\}$ noté $x_{i-p:i-1}$ et $\{x_{p+1}, \dots, x_{t_0}\}$ la partie du signal sur laquelle est ajusté le modèle AR-SVM d'ordre p .

Pour simplifier, la fonction noyau définit un produit scalaire (Eq.1.4). L'optimisation du modèle AR-SVM, formulée comme la minimisation d'une fonction quadratique, aboutit à la sélection d'un certain nombre de segments du signal pour lesquels β_j est non nul, ce sont les vecteurs supports. La prédiction d'une valeur x_i est la somme pondérée des produits scalaires des p valeurs précédentes avec les vecteurs supports.

La méthode est appliquée sur une structure de laboratoire à 4 niveaux ne subissant pas de variations environnementales ([Bornn et al. 2009](#)). Un modèle *non endommagé* est généré pour chaque série et l'apparition de l'endommagement augmente l'erreur de prédiction de ce modèle. Cependant, le gain par rapport à un modèle AR n'est pas démontré.

Indicateurs modaux

Bien que non développées dans ce travail de thèse, les caractéristiques modales et leurs variantes peuvent naturellement constituer des indicateurs à part entière dans une démarche de reconnaissance de formes statistiques.

Ainsi, les fréquences et modes propres peuvent être utilisés pour caractériser les mesures, mais sont très sensibles aux variations environnementales et opérationnelles (§1.1). Des variantes ont été développées parmi lesquelles la courbure des déformées propres, la matrice de flexibilité (Giraldo et al. 2006, Hsu et Loh 2010), la courbure de flexibilité et l'énergie de déformation (Garcia et Osegueda 2000). Alvandi (2003) compare certaines de ces approches dans ses travaux doctoraux. Une revue détaillée de l'utilisation des indicateurs modaux est également proposée par Carden et Fanning (2004) et Doebling et al. (1996; 1997).

La transformation en ondelettes

La transformation en ondelettes est une technique de traitement du signal apparue dans les années 1980 jouissant actuellement d'une forte popularité dans la communauté scientifique. Cet outil conceptuel et ses variantes offrent de multiples possibilités de développement qui ne peuvent être décrites exhaustivement en quelques lignes. Cette partie se limitera ainsi à une présentation rapide du principe de la technique.

La transformation en ondelettes peut se concevoir comme une extension de la transformation de Fourier sur une fenêtre ajustable en position et en taille. Elle permet ainsi une analyse temps-fréquence du signal en le décomposant en une somme de fonctions issues de la translation (information en temps) et de la dilatation (information en fréquence) d'une même fonction oscillante dite *ondelette mère*. Il est alors possible de déceler les irrégularités du signal grâce à la bonne résolution à la fois en temps et en fréquence¹³ ainsi obtenue, ce qui n'est pas le cas avec l'analyse de Fourier à court terme (ou *locale*). Les coefficients d'ondelette se calculent selon l'expression :

$$C(a, b) = \frac{1}{\sqrt{|a|}} \int_{-\infty}^{+\infty} x(t) \Psi^* \left(\frac{t-b}{a} \right) dt \quad (1.6)$$

où $C(a, b)$ sont les coefficients d'ondelettes, a le paramètre d'échelle, b le paramètre de translation dans le temps et Ψ l'ondelette mère.

Son succès est lié en grande partie au développement de l'*analyse multi-résolution*. Après plusieurs décompositions successives (correspondant à une

¹³Ici, la fréquence, n'apparaissant pas explicitement, est remplacée par la notion d'échelle.

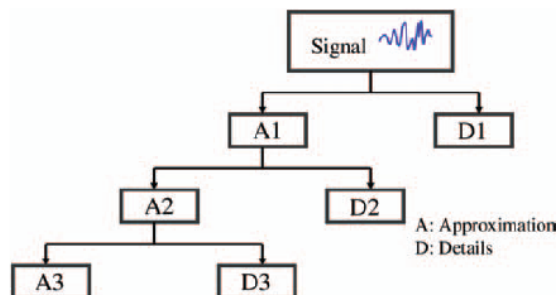


FIG. 1.6 Décomposition d'un signal en une partie approximation (A) et une partie détails (D) lors de l'analyse multi-résolution (Taha et al. 2006).

diminution du paramètre d'échelle), le signal est scindé en une version approximée (ou lissée), et plusieurs fonctions de détails contenant l'information associée aux hautes fréquences (Fig.1.6). Les capacités de cette technique à compresser les données et à repérer les changements brutaux à partir de l'observation des fonctions de détails, en ont fait un outil très usité dans les domaines applicatifs de l'analyse du signal.

Parmi les différents moyens d'exploiter la transformation en ondelette pour la détection d'endommagement, la plus immédiate est l'observation du diagramme temps/échelle des coefficients d'ondelettes (scalogramme), à la recherche de changements significatifs (Figueiredo et al. 2009, Alvandi et al. 2009). Les non linéarités liées à l'apparition de chocs (assemblages desserrés, fissures) sont également observables sur les fonctions de détails. Yen et Lin (2000) proposent un indicateur basé sur l'énergie de ces fonctions. Taha et al. (2006) présentent une synthèse des principales méthodes de suivi de l'état de santé des structures mettant en œuvre les différentes possibilités offertes par les ondelettes. Ils insistent également sur l'association des techniques basées sur les ondelettes avec certaines techniques de reconnaissance de formes (réseaux de neurones, algorithmes génétiques, systèmes flous).

Caractérisation dans l'espace des phases

La théorie des systèmes dynamiques non linéaires propose de nouveaux outils conceptuels permettant d'étudier la dynamique sous-jacente d'un système générant des séries temporelles observables. Le corpus scientifique associé à cette science jeune, toujours en construction aujourd'hui, offre de multiples perspectives en terme de définition de nouveaux indicateurs permettant de caractériser les signaux. Ces indicateurs sont définis et calculés dans l'espace des phases du système. Composante substantielle de ce travail de thèse, les notions générales de la théorie des systèmes dynamiques non linéaires et les indicateurs calculés dans l'espace des phases sont détaillés dans le chapitre §2.

1.3.2 Normalisation

L'axiome IVb, énoncé dans l'article de Worden et al. (2007), déclare que plus un indicateur est sensible à l'endommagement, plus il le sera également aux variations environnementales et opérationnelles. Aussi, l'approche de détection de nouveauté doit être en mesure de décider si les variations observées sont normales, ou bien la manifestation d'un comportement nouveau.

L'étape précédente a permis de caractériser chaque mesure, ou individu (au sens statistique du terme) par un indicateur vectoriel. La détection de nouveauté compare chaque nouvel individu à une population de référence (ensemble de mesures réalisées sur une durée assez longue) et fournit un indice lié à sa probabilité d'appartenance à la population de référence.

La normalisation est généralement définie comme l'étape de réduction de l'effet des variations environnementales. Cette étape est souvent réalisée à l'aide de la création d'une régression ou d'une analyse en composante principale. Dans tous les cas, une base de référence est nécessaire soit pour

ajuster la régression, soit pour déterminer les directions principales. Dans ce mémoire, la notion de normalisation est élargie et définie comme la phase de comparaison entre un individu et la population de référence. Ainsi, tous les outils mathématiques permettant de réaliser cette comparaison peuvent être employés (y compris l'analyse en composante principale ou les méthodes régressives).

La thèse de [Tax \(2001\)](#) propose une liste très complète des méthodes de classification à une classe constituant un problème similaire à celui de la détection de nouveauté. Elles peuvent être segmentées en trois types :

- les estimateurs de densité : plus l'individu se situe dans une zone faiblement peuplée (faible densité), moins il a de chances d'appartenir à la population de référence,
- méthodes de reconstruction : un algorithme est entraîné pour prédire les indicateurs, une forte erreur de prédiction est alors le signe d'un nouvel individu,
- les méthodes de frontière : une frontière est créée pour délimiter les zones peuplées.

Les estimateurs de densité

En fonction des hypothèses formulées sur la distribution des individus, les estimateurs sont dits paramétriques (forme de distribution supposée), semi-paramétriques ou non-paramétriques (pas d'hypothèse).

Méthode paramétrique : la Distance de Mahalanobis (MD)

Parmi les méthodes paramétriques, la distance de Mahalanobis est probablement la plus simple et la plus populaire. Couramment employée dans le champ de la statistique pour la détection de points aberrants, elle consiste à calculer la distance entre un point et le centre d'un nuage de point en tenant compte des corrélations linéaires entre les variables de l'espace (Fig.1.7). Elle suppose alors que la distribution des points (représentant les individus de la population de référence) suit une loi gaussienne multidimensionnelle, les paramètres à ajuster étant ceux de la loi, c'est-à-dire la matrice de covariance et le vecteur moyen (3.2).

[Worden et al. \(2000\)](#) utilisent la distance de Mahalanobis pour détecter avec succès l'endommagement sur un système masse-ressort simulé numériquement et sur l'arbre d'une machine. Une application en laboratoire sur un treillis métallique est également proposée par [Gul et Catbas \(2009\)](#).

[Worden et al. \(2002b\)](#) proposent une démarche originale consistant à rendre la matrice de covariance et le vecteur moyen dépendants des facteurs environnementaux. Leur étude consiste à simuler un système masse-ressort intégrant une dépendance à la température. La matrice de covariance et le vecteur moyen sont évalués en tenant compte de la température par interpolation ou régression. L'idée fonctionne, mais nécessite l'application à des cas réels dont les dépendances aux facteurs environnementaux sont souvent multiples et couplées.

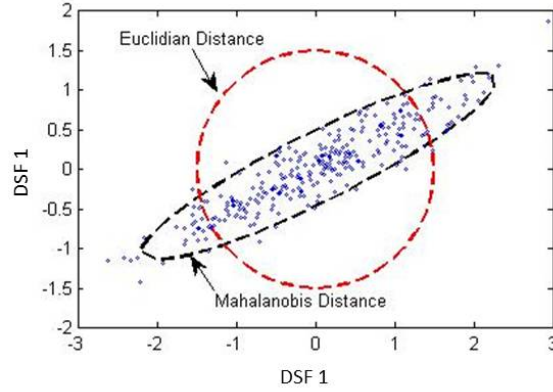


FIG. 1.7 Représentation d'une ligne d'égal distance de Mahalanobis (en noir) et d'une ligne d'égal distance euclidienne (en rouge) par rapport au centre du nuage de points sur une distribution bidimensionnelle de points

Méthode semi-paramétrique : Modèles de Mélanges de Gaussiennes

L'hypothèse d'une distribution gaussienne des individus étant très forte, un modèle basé sur des mélanges de gaussiennes permet de considérer des distributions multimodales. La densité est estimée par une combinaison linéaire de distributions gaussiennes (Eq.1.7).

$$p_{GMM}(X) = \frac{1}{N} \sum_{j=1}^N \alpha_j p_{\mathcal{N}}(X; \mu_j, \Sigma_j) \quad (1.7)$$

avec X l'indicateur vectoriel, N le nombre de gaussiennes et μ_j et σ_j respectivement le vecteur moyen et la matrice de covariance de chaque distribution gaussienne individuelle. Le nombre de paramètres à ajuster peut rapidement devenir important, c'est pourquoi les matrices de covariances peuvent être limitées à des matrices diagonales, surtout en cas de quantité de données limitée.

Le nombre de gaussiennes N peut être estimé en appliquant un algorithme de regroupement non-supervisé.

Méthode non-paramétrique : Estimation par noyau

Introduite par Parzen (1962), cette méthode constitue un équivalent continu de l'estimation par histogrammes. Pour schématiser, alors que les histogrammes peuvent être vus comme une somme de fonctions rectangulaires placées sur chaque individu, les estimateurs par noyau placent une courbe en cloche sur chacun d'eux (Fugate et al. 2000). En effet, la densité est estimée comme une somme de fonctions noyaux identiques, centrées sur chacun des individus de la population de référence (Eq.1.8). La matrice de covariance est généralement choisie diagonale, $\Sigma = h\mathcal{I}$.

$$p_h(X) = \frac{1}{N_{ref}} \sum_{j=1}^{N_{ref}} p_{\mathcal{N}}(X; X_j, h\mathcal{I}) = \frac{1}{N_{ref}} \sum_{j=1}^{N_{ref}} k(X, X_j) \quad (1.8)$$

$$k(X, X_j) = \frac{1}{(\sqrt{2\pi}h)^d} \exp\left(-\frac{\|x - y\|^2}{2h^2}\right) \quad (1.9)$$

avec N_{ref} le nombre d'individus dans la population de référence, d la dimension d'un indicateur vectoriel, et h un paramètre de lissage qui représente la largeur de la fonction noyau associée à chaque individu. La détermination de ce paramètre n'est pas simple. La valeur de $h = N_{ref}^{-\frac{1}{d+4}}$ est souvent recommandée (Scott et Sain 2005).

La fonction noyau étant souvent choisie comme une fonction gaussienne de moyenne nulle et de variance unitaire, cet estimateur est sensible à l'effet d'échelle entre les composantes de l'indicateur vectoriel. Il faut alors prendre soin d'uniformiser les données par centrage et réduction. L'estimation de la densité est également coûteuse en calcul puisqu'elle nécessite le stockage de tous les indicateurs de la population de référence et le calcul, pour chaque individu, de sa distance avec chaque élément de la population de référence.

Toivola et al. (2010) comparent ces trois méthodes et celle du $k^{ème}$ plus proche voisin (kNN , voir §1.3.2) pour la détection d'endommagement sur une maquette de pont en bois. L'endommagement est simulé par des ajouts de masses ponctuelles. L'étude montre que la méthode des plus proches voisins est celle qui donne les meilleurs résultats de classification pour la plus large gamme de dimensions de vecteurs. En effet, l'estimation de la densité se dégrade avec l'augmentation de la dimension en raison de la dispersion des points dans l'espace (Fugate et al. 2000).

Méthodes de reconstruction

Les méthodes de reconstruction regroupent les approches permettant de reproduire les données grâce à une représentation condensée de celles-ci, l'objectif étant toujours, de faire apparaître la structure de corrélation cachée de la population de référence.

Analyse en composantes principales

L'analyse en composantes principales (ACP) est probablement la méthode la plus répandue de compression de données. Elle commence par la diagonalisation de la matrice de covariance C_1 des variables de la base de référence (composantes de l'indicateur vectoriel), puis recherche les directions de plus fortes variances dans l'espace des variables (axes principaux ou axes factoriels). Il est alors possible de transformer un certain nombre de variables corrélées en un nombre plus faible de nouvelles variables linéairement non corrélées en ne considérant que les composantes principales de plus fortes variances.

En notant $\{X_i, i = 1, \dots, m\}$ l'ensemble des m réalisations de l'indicateur vectoriel X , de dimension d , représentant les individus de la population de référence, la matrice de covariance s'écrit,

$$C_1 = \frac{1}{m} \sum_{j=1}^m X_j X_j^T = U \Sigma^2 U. \quad (1.10)$$

La diagonalisation de C_1 permet d'introduire la matrice orthogonale U contenant les vecteurs propres, et la matrice diagonale Σ les valeurs propres.

Généralement, les plus grandes valeurs propres sont considérées comme celles qui traduisent le plus les variations environnementales subies par la structure, alors que les dernières sont probablement liées au bruit de mesure ou à des facteurs marginaux. En réarrangeant U et Σ par ordre décroissant des valeurs propres, la matrice de transformation T contenant les d' premiers vecteurs propres (de dimension $d \times d'$, $d' < d$) est créée. Elle permet d'obtenir une représentation réduite des individus. La perte d'information induite par cette transformation s'exprime par l'erreur e_i de l'équation 1.13.

$$\hat{Y}_i = X_i T \quad (1.11)$$

$$\hat{X}_i = \hat{Y}_i T^T \quad (1.12)$$

$$e_i = \|X_i - \hat{X}_i\| = \|X_i - X_i T T^T\| \quad (1.13)$$

Un exemple d'utilisation de cette approche est proposé par [Yan et al. \(2005\)](#) sur un modèle numérique de pont soumis à un gradient de température et un pont réel en bois sur lequel des masses ponctuelles sont ajoutées pour simuler un endommagement. Les résultats sont positifs, mais un fort gradient de température fait apparaître une relation non-linéaire entre les indicateurs (fréquences propres) rendant l'approche moins sensible. Cela peut être en partie compensé par une augmentation du nombre de vecteurs propres conservés. Les auteurs soulignent également l'importance de se contenter de centrer les données sans les réduire.

ACP à noyau

Elle constitue l'équivalent non linéaire de l'ACP. Elle exploite l'*astuce du noyau* qui consiste à appliquer une transformation non linéaire (Eq.1.14) sur l'espace des composantes de l'indicateur vectoriel, augmentant au passage la dimension. Une méthode linéaire (ici, l'ACP) est ensuite mise en œuvre dans le nouvel espace, ce qui équivaut à utiliser une méthode non-linéaire dans l'espace initial.

$$\begin{aligned} \Phi : \mathcal{R}^d &\rightarrow \mathcal{F} \\ X &\rightarrow \Phi(X) \end{aligned} \quad (1.14)$$

La matrice de covariance des composantes de l'indicateur vectoriel centrées dans \mathcal{F} s'écrit

$$C_2 = \frac{1}{m} \sum_{j=1}^m \Phi(X_j) \Phi(X_j)^T \quad (1.15)$$

et le problème de recherche des valeurs propres λ , et des vecteurs propres V ,

$$\lambda V = C_2 V. \quad (1.16)$$

En multipliant de part et d'autre l'équation 1.16 par $\Phi(X_k)$ et en exprimant V par une combinaison linéaire des $\{X_i, i = 1, \dots, m\}$, le problème des valeurs propres précédent se réécrit (Schölkopf et al. 1998)

$$m\lambda\alpha = K\alpha \quad (1.17)$$

$$K_{ij} = \Phi(X_i)^T \Phi(X_j) = K(X_i, X_j) \quad (1.18)$$

avec α représentant un vecteur propre V normalisé, et $K(X, Y)$ une fonction noyau.

La dimension importante de K ($m \times m$) et la transformation non linéaire permettent de mettre en évidence des corrélations non linéaires dans les données et ainsi de réaliser des classifications plus robustes. La forme des corrélations représentées dépend des fonctions noyaux utilisées. Si les données ont tendance à être regroupées par paquets, des fonctions à base radiale (RBF) sont préférables, alors que si les corrélations sont plutôt courbes, des noyaux polynomiaux sont utilisés.

Dans le cadre de l'analyse des données statiques de tension des câbles du pont suspendu de *Yeongjong* en Corée du Sud, Oh et al. (2009) appliquent une ACP à noyau sur les vecteurs formés par les déformations issues de différents capteurs. L'objectif n'est pas ici de reconstruire les mesures initiales, mais seulement de réduire la dimension des vecteurs en un nombre limité de composantes principales non linéaires. Ensuite, une méthode de détection de nouveauté basée sur la distance est appliquée. Les auteurs mettent ainsi en évidence une corrélation initialement invisible entre la première composante principale et la température.

Réseaux de neurones auto-associatifs

Les réseaux de neurones artificiels (RNA) sont des outils permettant de modéliser tous types de relations fonctionnelles entre différentes variables, et notamment les relations non linéaires. Développés depuis les années 1950, ils avaient pour vocation originelle de modéliser le fonctionnement du cerveau dans lequel des ensembles de neurones sont inter-connectés pour permettre des apprentissages complexes. Ils se sont ensuite imposés comme de puissants outils de modélisation statistique non-paramétrique aux applications multiples.

Un neurone artificiel ou formel est la base d'un réseau de neurones. Il s'agit d'une fonction paramétrée non linéaire à valeurs bornées. L'association de plusieurs neurones formels est appelée réseau de neurones artificiels. Les neurones d'un réseau sont organisés en couches successives comme illustré, par exemple, sur la figure 1.9 qui en comporte 4. Le neurone j de la couche k est représenté sur la figure 1.8 où S_{jk} est la sortie du neurone, e_{ik} sont les variables d'entrée, ω_{ij} et θ_{jk} sont les paramètres à ajuster et σ est la fonction d'activation (ou fonction de transfert). La fonction sigmoïde est souvent utilisée comme fonction d'activation

$$\sigma(x) = \frac{1}{1 + e^{-x}}. \quad (1.19)$$

Cybenko (1989) a démontré que l'utilisation d'un RNA comportant une couche, dite *cachée*, de fonctions sigmoïdes, et une couche de sortie linéaire,

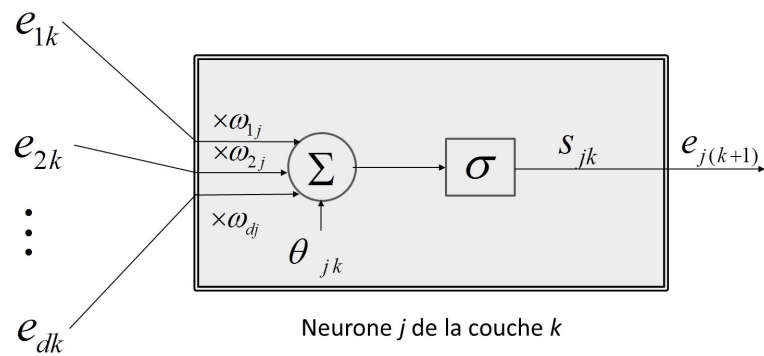


FIG. 1.8 Fonctionnement d'un neurone

est en mesure de représenter n'importe quelle fonction continue. Soit G une fonction continue non linéaire, telle que $Y = G(X)$, alors chaque composante y_l de Y peut être calculée par

$$y_l = \sum_{j=1}^{N_1} \omega_{ji2} \cdot \sigma \left(\sum_{i=1}^{N_2} \omega_{ij1} x_i + \theta_j \right) \quad (1.20)$$

avec 1 et 2 représentant les indices des deux couches comportant respectivement N_1 et N_2 nœuds. Après une phase d'apprentissage qui permet l'ajustement des poids et des biais, cette forme de RNA est capable de reproduire n'importe quelle relation entre variables. Cette faculté a conduit au développement de l'utilisation des RNA face à des problèmes de régression et de classification supervisée. De nombreuses applications en monitoring des structures ont été proposées lorsque des données sont mesurées sur la structure endommagée (Cury et Crémona 2010, Worden et Manson 2007) ou simulées (Sun 2002).

Le cas d'un apprentissage non supervisé est géré en créant un réseau dont la sortie doit être identique à l'entrée. On parle alors d'*auto-encodeur* ou de *réseau de neurones auto-associatif* (AANN¹⁴) (Kramer 1991). Dans ce cas, une première partie du réseau projette le vecteur d'entrée dans un espace de dimension inférieure via une fonction G ($Y_i = G(X_i)$), puis une deuxième partie le renvoie dans l'espace initial via, une fonction H ($\bar{X}_i = H(Y_i)$). La sortie de la fonction G constitue l'entrée de H (Fig.1.9).

Pour forcer une représentation condensée de l'information, la couche de sortie de la projection doit comporter moins de composantes que les vecteurs d'entrée. Ces composantes représentent des variables cachées dont dépendent les vecteurs d'entrée, comme par exemple la température. Les auto-encodeurs peuvent également être considérés comme un moyen de réaliser une ACP non linéaire en récupérant les variables de sortie de la couche de projection.

Dans les travaux de Hsu et Loh (2010), un pont est simulé numériquement. Ses différents éléments constitutifs présentent une dépendance non-linéaire à la température. La procédure d'identification de la matrice de

¹⁴Auto-Associative Neural Network

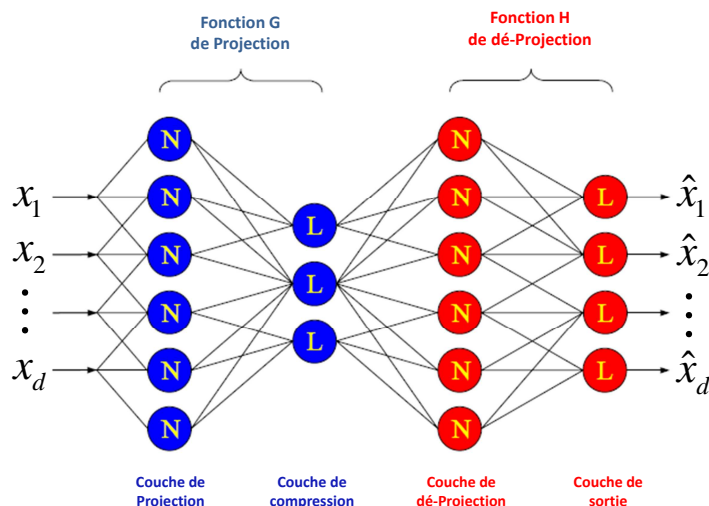


FIG. 1.9 Composition d'un auto-encodeur, L et N signifient que la fonction d'activation est respectivement linéaire et non-linéaire (Sohn et al. 2002)

rigidité est mise en œuvre et sert d'entrée à un AANN qui parvient à détecter correctement des réductions de raideur.

Sohn et al. (2002) présentent une autre application basée sur la détection de non-linéarités sur un système masse-ressorts à 8 degrés de liberté. Les variations environnementales sont simulées par une excitation d'intensité variable. Les coefficients d'un modèle AR-ARX sont utilisés pour l'apprentissage. Chaque nouvelle mesure est alors prédite avec les coefficients obtenus par le AANN et l'erreur de prédiction est quantifiée. La même étude est également réalisée selon une ACP à noyaux (Oh et Sohn 2009). Les auteurs insistent sur la supériorité de l'ACP à noyau qui :

- est moins sensible au risque du sur-apprentissage,
- permet une plus grande souplesse dans le choix du nombre de composantes principales à retenir, alors que l'AANN impose un réseau différent (et donc un nouvel apprentissage) lorsque le nombre de composantes change (l'ACP à noyau permet de les calculer toutes simultanément),
- présente une solution globale à un problème aux valeurs propres alors que l'AANN nécessite la résolution d'un problème d'optimisation non linéaire sujette aux effets néfastes des optimums locaux.

Méthodes frontières

Les méthodes frontières cherchent à déterminer directement une limite permettant de regrouper les individus de l'unique classe disponible pour l'apprentissage. Parmi les différentes approches existantes, les *Machines à Vecteurs de Supports* (SVM¹⁵) constituent une méthode frontière très populaire depuis les années 1990. Moins exploitée en SHM, la méthode du $k^{\text{ème}}$ plus proche voisin est brièvement présentée.

¹⁵Support Vector Machine

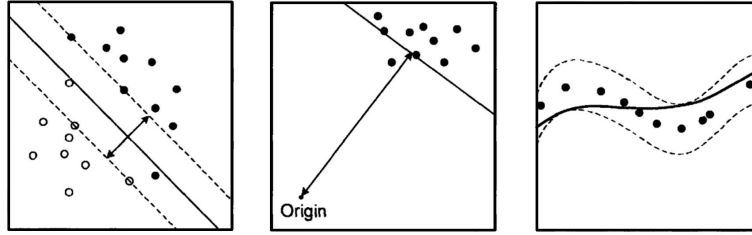


FIG. 1.10 *Interprétation géométrique des Machines à Vecteurs de Support pour des applications de classification multi-classes, de détection de nouveauté et de régression (respectivement de gauche à droite) (Tsou 2007)*

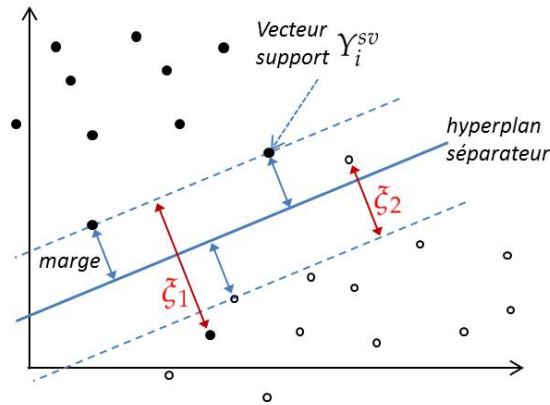


FIG. 1.11 *Principe de la classification de deux populations par les SVM*

Machines à Vecteurs de Support

Les SVM s'appliquent à la classification, à la détection de nouveauté ou encore à la régression (Fig.1.10). Dans tous les cas, les SVM sont basées sur deux idées principales.

La première idée réside dans la transformation des données vers un espace de dimension supérieure par le biais des fonctions noyaux comme présenté précédemment. La seconde consiste à rechercher un hyperplan de cet espace séparant les classes d'individus.

L'hyperplan qui permet la meilleure généralisation de la classification est celui que maximise les marges, c'est-à-dire les distances entre les individus des classes d'apprentissage et l'hyperplan (Fig.1.11). Les points de distance minimale sont qualifiés de *vecteurs supports* (VS).

Avec les mêmes notations que dans l'équation 1.14, soit $Y_i = \Phi(X_i)$, la transformation de X_i dans l'espace de dimension supérieure. Dans cet espace, l'équation de l'hyperplan de séparation \mathcal{H} a pour paramètres b et un vecteur normal w . L'ensemble des points Y de l'espace appartenant à \mathcal{H} vérifient :

$$h(Y) = w \cdot Y + b = 0. \quad (1.21)$$

Puisque l'hyperplan doit classer les points suivant leur appartenance à deux classes possibles, il est plus simple de considérer que les points de la classe 1 sont au-dessus de \mathcal{H} ($h(Y)$ positif) et ceux de la classe 2 en-dessous

($h(Y)$ négatif). L'hyperplan doit donc vérifier :

$$\begin{cases} w \cdot Y_i^1 + b \geq 0 \\ w \cdot Y_i^2 + b \leq 0 \end{cases} \quad (1.22)$$

En normalisant w et b de sorte que $h(Y_i^{sv})$ soit égal à ± 1 pour tous les vecteurs supports, l'équation 1.22 se réécrit,

$$\begin{cases} w \cdot Y_i^1 + b \geq 1 \\ w \cdot Y_i^2 + b \leq -1 \end{cases} \quad (1.23)$$

La marge, définie comme le minimum de la distance entre les vecteurs Y_i et \mathcal{H} , s'écrit

$$\min_k \left(\frac{|w \cdot Y_k + b|}{\|w\|} \right). \quad (1.24)$$

Étant donné que les points les plus proches sont les supports et que $h(Y_i^{sv}) = \pm 1$ pour ces points, la marge vaut $\frac{1}{\|w\|}$. La maximisation de la marge passe donc par la minimisation de la norme de w .

Le problème de classification se résume donc à la recherche de w et b vérifiant :

$$\min_w \left(\frac{1}{2} \|w\|^2 \right) \quad (1.25)$$

sous les contraintes de l'équation 1.23.

Il est rare les points soient tous séparables par un hyperplan. Pour ajouter un peu de souplesse à ces contraintes et ainsi permettre au classifieur une meilleure généralisation, une *variable ressort* ξ_i est introduite. Elle autorise le mauvais classement de certains points ($\xi_i \neq 0$, Fig.1.11) en contre-partie d'une pénalisation de la minimisation.

Le problème assoupli s'écrit alors

$$\min_{w, \xi_i} \left(\frac{1}{2} \|w\|^2 + C \sum_i \xi_i \right) \quad (1.26)$$

sous les contraintes

$$\begin{cases} w \cdot Y_i^1 + b \geq 1 - \xi_i \\ w \cdot Y_i^2 + b \leq \xi_i - 1 \\ \xi_i \geq 0 \end{cases} \quad (1.27)$$

Le paramètre C pilote le compromis entre un nombre d'erreurs de classification important (faibles valeurs de C) et une faible marge. Il conditionne donc la capacité de généralisation du classifieur.

A l'aide du multiplicateur de Lagrange, le problème posé devient quadratique et peut ainsi être facilement résolu. En notant S_i ($i = 1, \dots, n_{sv}$) les vecteurs supports obtenus dans l'espace initial, w s'exprime alors comme la combinaison linéaire :

$$w = \sum_{i=1}^{n_{sv}} \alpha_i l_i \Phi(S_i) \quad (1.28)$$

avec $l_i = 1$ si les vecteurs S_i appartiennent à la classe 1, sinon $l_i = -1$. L'expression de h permet de faire apparaître le produit scalaire des vecteurs dans \mathcal{F} et ainsi d'utiliser une fonction noyau K :

$$\begin{aligned} h(X_i) &= \sum_{k=1}^{n_{sv}} \alpha_k l_k \Phi(S_k) \cdot \Phi(X_i) + b \\ &= \sum_{k=1}^{n_{sv}} \alpha_k l_k K(S_k, X_i) + b \end{aligned} \quad (1.29)$$

Toute l'information nécessaire à la classification est contenue dans les vecteurs supports. Lorsque les paramètres ne sont pas adaptés (mauvaise fonction noyau, valeur de C trop importante), le nombre de vecteurs supports peut être trop important, caractérisant ainsi un sur-apprentissage. De plus amples détails sur les étapes de calculs présentées ci-dessus sont disponibles dans le tutoriel de [Burges \(1998\)](#) et l'ouvrage de [Schölkopf et Smola \(2002\)](#).

Après cette introduction au principe de la classification par les SVM, son application à la détection de nouveauté s'appuie sur le même raisonnement. Cependant, dans la mesure où une seule classe d'individus est disponible, [Schölkopf et al. \(2000\)](#) propose de maximiser la marge entre l'origine du repère de l'espace et l'hyperplan (Fig.1.10). Le problème devient alors :

$$\min_{w, \xi_i} \left(\frac{1}{2} \|w\|^2 - b + \frac{1}{\nu m} \sum_i \xi_i \right) \quad (1.30)$$

sous les contraintes

$$w \cdot Y_i \geq b - \xi_i \quad \text{et} \quad \xi_i \geq 0, \quad (1.31)$$

avec m le nombre d'individus dans la population d'apprentissage et ν un scalaire compris entre 0 et 1. La solution du problème quadratique fournit l'expression de h :

$$h(X_i) = \sum_{k=1}^{n_{sv}} \alpha_k K(X_i, S_k) - b. \quad (1.32)$$

Le paramètre ν contrôle le compromis entre un grand nombre de points tolérés en dehors de la frontière (*outliers*) et le nombre de vecteurs supports. Si $\nu \rightarrow 0$ alors il n'y a quasiment pas d'outliers. En revanche, si $\nu \rightarrow 1$, tous les points sont des vecteurs supports. Ce cas est équivalent à l'utilisation d'un estimateur non paramétrique par noyau (§1.3.2). Hormis certaines recommandations dans le cas des régressions ([Cherkassky et Ma 2004](#)), il n'y a pas de consensus sur une méthode rigoureuse de sélection des paramètres des SVM. L'approche courante consiste en une validation croisée testant plusieurs combinaisons des paramètres à déterminer.

Hormis de nombreuses applications dans des cas de détection d'endommagement en mode supervisé ([Cury et Crémona 2010](#), [Worden et Manson](#)

2007, Mita et Hagiwara 2003), rares sont celles qui exploitent les SVM à des fins de détection de nouveauté. Dans la thèse de Tsou (2007), l'endommagement est recherché sur deux modèles numériques de structures (dont le benchmark ASCE) en exploitant directement les accélérations en données d'entrée. Das et al. (2007) exploitent l'information extraite de l'analyse fréquentielle d'ondes ultrasonores pour entraîner un SVM à une classe dans le but de détecter l'endommagement d'une plaque en matériau composite. Cette approche permet de détecter correctement une délamination dans la plaque.

k^{ème} plus proche voisin

De façon non probabiliste, la méthode du *k^{ème} plus proche voisin* (*kNN*¹⁶) offre un moyen simple de classification à une classe. Pour cela, le *kNN* du nouvel individu X , noté $k_{NN}(X)$, est recherché dans la population de référence. A son tour, le *kNN* de $k_{NN}(X)$ est sélectionné, il est désigné par $k_{NN}(k_{NN}(X))$. Si le quotient $f(X)$ de l'équation 1.33 est inférieur à 1, alors le point est considéré comme appartenant à la population de référence (Tax 2001).

$$f(X) = \frac{\|X - k_{NN}(X)\|}{\|k_{NN}(X) - k_{NN}(k_{NN}(X))\|} \quad (1.33)$$

Introduction des données environnementales

Une possibilité, non évoquée précédemment, consiste à exploiter les mesures des facteurs environnementaux, si elles existent, pour les intégrer dans le processus de normalisation. L'approche la plus intuitive consiste à créer une fonction reliant par exemple les informations climatiques et les indicateurs. La plupart des outils de régression semblent alors adaptés.

La première difficulté réside dans la sélection des variables environnementales affectant le plus les indicateurs. La température de l'air risque de présenter des variations beaucoup plus importantes que celle de la structure en raison de l'inertie de cette dernière. De plus, la température interne de l'ouvrage présente de fortes hétérogénéités en fonction de l'exposition. Farrar et Doebling (1997) montrent des contrastes de température variables au cours d'une journée et pouvant atteindre 20°C au sein du tablier du pont I-40.

Peeters et De Roeck (2001) cherchent à expliquer les variations des fréquences propres du pont Z-24 suivi pendant un an. Parmi les 49 capteurs installés sur l'ouvrage, les auteurs sélectionnent celui mesurant la température de la couche de roulement car celle-ci a un effet significatif sur la rigidité du tablier lorsque la température devient négative. Un modèle de type *boîte noire* est ensuite ajusté avec les fréquences propres, rendant possible la détection d'endommagement par l'erreur de prédiction.

¹⁶k-Nearest Neighbors

Pour comparer les indicateurs obtenus dans les mêmes conditions de température, [Worden et al. \(2002b\)](#) paramètrent le vecteur moyen et la matrice de covariance (utilisés dans le calcul de la distance de Mahalanobis) avec la température à l'aide d'une régression et d'une interpolation. Le test est réalisé sur un modèle masse/ressort à deux degrés de liberté et dépendant de la température. La détection de l'endommagement s'en trouve améliorée par rapport à l'approche sans prise en compte de la variable de température. Cependant, l'utilisation de la régression ne garantit pas le caractère semi-défini positif de la matrice de covariance, ce qui conduit à des distances potentiellement négatives. L'interpolation ne présente pas ce problème.

Cependant, les variables affectant le comportement d'un ouvrage sont souvent multiples et agissent en interaction. Chercher à modéliser des relations fonctionnelles avec ces facteurs extérieurs peut rendre encore plus complexe l'analyse de données déjà très variables. Dans ce travail de thèse, seules les données liées au comportement de l'ouvrage sont utilisées, dans un souci de développement d'outils facilement transposables à tous types de structures.

1.3.3 Seuil de rejet

Une fois les données normalisées, chaque individu est caractérisé par une valeur scalaire (ex : distance de Mahalanobis) ou un vecteur (ex : erreur de reconstruction ou de prédiction). Malgré l'étape de normalisation, les valeurs obtenues présentent des variations dues aux facteurs environnementaux, aux algorithmes de calculs, aux aléas de la mesure, etc. Ces valeurs sont alors considérées comme des variables aléatoires.

Il faut ensuite décider si la nouvelle observation peut être expliquée par la distribution de la population de référence, ou bien si elle est issue d'une nouvelle distribution. Dans ce contexte, cette partie présente une série d'outils liés à la maîtrise statistique des processus et aux tests d'hypothèses. Les statistiques des valeurs extrêmes et le test séquentiel du rapport des vraisemblances sont également abordés.

Cartes de contrôle

Les cartes de contrôle sont des outils employés en maîtrise statistique des processus (MSP). Leurs principes généraux sont présentés ci-après, mais de plus amples détails sont disponibles dans les ouvrages de [Montgomery \(2005\)](#) et [Pillet \(2008\)](#).

Une carte de contrôle ([Fig.1.12](#)) représente le tracé d'une caractéristique de la qualité d'un processus (production industrielle) en fonction du numéro de l'échantillon (individu) examiné. Sur ce graphique sont également tracées la ligne correspondant à la valeur moyenne de la caractéristique (CL), et les limites de contrôle supérieures et inférieures (UCL-LCL). Tant que les valeurs obtenues restent entre ces limites, le processus est considéré *sous contrôle*. Dans ce cas, il ne subit que les effets des *causes communes* (facteurs de variabilité normaux). En revanche, si un dépassement est constaté, alors une *cause spéciale* agit sur le processus. Il devient *hors de contrôle* et un opérateur doit intervenir pour identifier le problème et le supprimer.

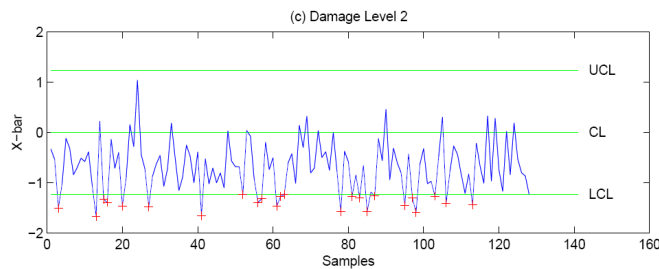


FIG. 1.12 Exemple de carte \bar{X} (Sohn et Farrar 2000)

Bien que le vocabulaire ne soit pas exactement identique, l'approche semble bien adaptée à la détection de nouveauté pour des structures de génie civil. Les indicateurs sont tracés chronologiquement et comparés à des limites correspondant aux valeurs extrêmes obtenues dans la base de référence.

L'objectif des cartes de contrôle est de détecter à la fois une déviation du processus par rapport à sa valeur moyenne, mais également des changements de son écart-type. Or, la prise en compte d'un seul individu à la fois ne permet pas de détecter les changements de l'écart-type. C'est pourquoi il est courant de travailler sur des échantillons composés de n individus successifs, obtenus avec ou sans chevauchement des échantillons. Cette démarche offre plusieurs avantages. Tout d'abord, les variations dans la mesure de la moyenne sont réduites, ce qui améliore la détection de faibles écarts. De plus, si la distribution de la population de référence n'est pas normale, la moyenne calculée sur un échantillon tend vers une distribution normale quand n augmente.

La désignation de la carte de contrôle dépend de la caractéristique suivie :

- carte \bar{X}* → moyenne
- carte \bar{X}* → moyenne sur un échantillon
- carte \bar{S}* → écart-type sur un échantillon

En fonction du degré de connaissance de la distribution de la population de référence, l'expression des limites de contrôle diffère. En MSP, dans un souci de simplicité, ces limites sont généralement exprimées en fonction d'un coefficient tabulé et de l'estimation non biaisée de la moyenne et de l'écart-type du processus lorsqu'il est sous contrôle. Par habitude, elles sont placées à 3 écarts-types (3σ) de la moyenne, correspondant à une probabilité de 0,0027 d'obtenir une valeur *hors de contrôle* sous la double hypothèse qu'aucune perturbation n'est présente (erreur de type I) et que la distribution est normale. Ce choix relève d'une habitude liée à la fois à la simplicité d'un coefficient entier et à une efficacité constatée du processus de contrôle.

En se plaçant dans une démarche de *test d'hypothèse*, les limites de contrôle peuvent être directement choisies en fonction de l'*erreur de type I* tolérée, notée α . Il n'existe pas de distinction claire entre les cartes de contrôle et le *test d'hypothèse*, suscitant ainsi un débat même au sein des communautés de spécialistes concernées (Woodall 2000).

Dans la partie suivante, la démarche générale du *test d'hypothèse* est décrite, ainsi que les expressions courantes des limites de contrôle.

Test d'hypothèse

Pour les applications qui concernent le suivi des structures, l'objectif du *test d'hypothèse* est de vérifier si un échantillon de valeurs provient d'une population donnée (population de référence), ce qui constitue l'hypothèse, ou bien d'une autre distribution.

Le test comprend plusieurs étapes :

1. Définir l'*hypothèse nulle* H_0 à vérifier et l'*hypothèse alternative* H_1 .
Ex : H_0 = les valeurs proviennent d'une variable aléatoire de moyenne μ_0 ; H_1 = la moyenne est différente.
2. Choisir une *variable de décision* permettant le contrôle de l'hypothèse.
Ex : la moyenne des échantillons est la même que celle de la distribution, $\mu = \mu_0$.
3. Déterminer la *distribution de la statistique* si H_0 est vraie.
Ex : la distribution de la moyenne d'un échantillon de n individus, \bar{X} , est normale de moyenne μ_0 et d'écart-type $\frac{\sigma_0}{\sqrt{n}}$.
4. Définir la *région de rejet* de l'hypothèse en fonction de la probabilité d'erreur de type I acceptée (fausse alarme), notée α .
Ex : H_0 n'est pas rejetée si $\mu_0 - Z_{\alpha/2} \frac{\sigma_0}{\sqrt{n}} \leq \bar{X} \leq \mu_0 + Z_{\alpha/2} \frac{\sigma_0}{\sqrt{n}}$, avec $Z_{\alpha/2}$ le quantile de la loi normale centrée réduite associé à la probabilité $1 - \alpha/2$.
5. Calculer avec les échantillons disponibles la valeur de la variable de décision.
6. Rejeter ou non H_0 , en fonction de la valeur de la variable de décision par rapport à la zone de rejet établie.

La MSP peut être vue comme un test d'hypothèse sur chaque valeur caractéristique du processus, dont H_0 suppose qu'il est sous contrôle, le test statistique portant sur la moyenne ou l'écart-type. Toutes les valeurs en dehors des limites sont *hors de contrôle* et rejettent l'hypothèse nulle (Montgomery 2005).

Le tableau 1.1 détaille les principaux tests utilisés pour la détection de nouveauté et les limites qui fixent le domaine d'acceptation de l'hypothèse nulle. Le cas multidimensionnel est traité par le test *Hotelling* T^2 , qui est une généralisation du test de *Student* (2^{me} ligne du tableau 1.1). En considérant maintenant que X est une variable aléatoire de dimension p , si sa matrice de covariance est inconnue, le test s'écrit :

$$H_0 : \boldsymbol{\mu} = \boldsymbol{\mu}_0 \quad H_1 : \boldsymbol{\mu} \neq \boldsymbol{\mu}_0 \quad (1.34)$$

$$\text{Statistique de test : } T^2 = n (\bar{X} - \boldsymbol{\mu}_0)^T W^{-1} (\bar{X} - \boldsymbol{\mu}_0) \quad (1.35)$$

$$\text{Domaine de non rejet : } 0 < T^2 < \frac{p(n-1)}{n-p} F_{(\alpha, p, n-p)} \quad (1.36)$$

avec $F_{(\alpha, p, n-p)}$, le quantile correspondant à α de la loi de *Fisher*, de paramètres p et $n-p$, et W , la matrice de covariance estimée sur la population de référence :

$$W = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^T (X_i - \bar{X}). \quad (1.37)$$

De très nombreux exemples d'utilisation sont présents dans la littérature. [Bornn et al. \(2010\)](#) estiment l'erreur de prédiction de deux mesures obtenues à partir du même modèle. Les distributions sont comparées à l'aide du *F-test*.

Tous ces tests reposent sur l'hypothèse que la distribution des valeurs de la population de référence est normale. Or la nature précise de cette distribution n'est pas connue a priori. Par exemple, l'hypothèse d'une distribution normale, sur une loi log-normale sous-estime de façon importante la limite supérieure engendrant de nombreuses fausses alarmes ([Sohn et al. 2005](#)).

Dans la mesure où la détection de nouveauté s'intéresse plus particulièrement aux valeurs présentes dans les queues des distributions, il est essentiel d'être en mesure de modéliser ces zones avec le plus de précision possible dans le but de fixer des limites de contrôle optimales. Or, lorsque des cartes \bar{X} et S sont tracées, les moyennes et écarts-types sont estimés sur un échantillon de plusieurs valeurs, réduisant ainsi la sensibilité aux changements dus aux valeurs extrêmes du processus. Une réponse possible à ce problème est apportée par la *statistique des valeurs extrêmes*.

Statistique des valeurs extrêmes

La *statistique des valeurs extrêmes* (EVS) est une branche de la statistique d'ordre. Elles est très employée dans le domaine météorologique et permet d'associer une loi de probabilité à la réalisation d'évènements extrêmes. Pour cela, deux approches sont possibles ([Girard 2004](#)) :

- le raisonnement sur les *maxima* (et minima)
- le raisonnement sur les *excès*.

Soit X , une variable aléatoire de fonction de répartition $F(x) = P(X < x)$. X correspond, par exemple, aux indicateurs calculés sur la base de référence dont la distribution est inconnue. La détermination du domaine de rejet né-

Hypothèses	Remarque	domaine de non rejet de H_0
$H_0 : \mu = \mu_0$	σ connu ou n grand	$\mu_0 - Z_{\alpha/2} \frac{\sigma}{\sqrt{n}} < \bar{X} < \mu_0 + Z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$
$H_1 : \mu \neq \mu_0$	σ inconnu	$\mu_0 - t_{(\alpha/2;n-1)} \frac{S}{\sqrt{n}} < \bar{X} < \mu_0 + t_{(\alpha/2;n-1)} \frac{S}{\sqrt{n}}$
$H_0 : \sigma^2 = \sigma_0^2$ $H_1 : \sigma^2 \neq \sigma_0^2$	μ_0 inconnue	$\frac{\chi_{(1-\alpha/2;n-1)}^2 \sigma_0^2}{n+1} < S^2 < \frac{\chi_{(\alpha/2;n-1)}^2 \sigma_0^2}{n+1}$
$H_0 : \sigma_1^2 = \sigma_2^2$ $H_1 : \sigma_1^2 \neq \sigma_2^2$	$F_0 = \frac{S_1^2}{S_2^2}$ loi de Fisher	$F_{(1-\alpha/2;n_1-1;n_2-1)} < F_0 < F_{(\alpha/2;n_1-1;n_2-1)}$

TAB. 1.1 *Cas courants de tests statistiques et limites correspondantes*; $\bar{X} = \frac{\sum_{k=1}^n X_k}{n}$ et $S_i^2 = \frac{\sum_{k=1}^n (X_k - \bar{X})^2}{n-1}$ sont les estimateurs non biaisés de la moyenne et de la variance d'un échantillon; $Z_{\alpha/2}$ est le quantile à $(1 - \alpha/2)100\%$ de la loi normale; $t_{\alpha/2;n-1}$ est le quantile à $(1 - \alpha/2)100\%$ de la loi de Student à $n - 1$ ddl; $\chi_{(\alpha/2;n-1)}^2$ est le quantile à $(1 - \alpha/2)100\%$ de la loi χ^2 à $n - 1$ ddl.

cessite la connaissance de $F(x)$. Pour en modéliser correctement les queues, il faut s'intéresser plus particulièrement aux grandes (et petites) valeurs de X . Compte tenu du domaine d'application relatif à ce travail de thèse, seul le cas des grandes valeurs est présenté. Pour les petites valeurs, la démarche est cependant analogue (Castillo et al. 2005).

Maxima

Soit un échantillon de n réalisations de X , rangées dans l'ordre croissant. La $i^{\text{ème}}$ plus grande valeur de l'échantillon est notée $X_{n,i}$.

⇒ Si n est suffisamment grand, alors la répartition des maxima suit une *loi des valeurs extrêmes* (EVD¹⁷). Avec $F_{max}(x) = P(X_{n,n} < x)$, la forme généralisée de la distribution des valeurs extrêmes s'écrit :

$$F_{max}(x) = \exp \left\{ - \left[1 + \gamma \left(\frac{x - \mu}{\sigma} \right) \right]^{\frac{-1}{\gamma}} \right\} \quad (1.38)$$

avec $-\sigma - \gamma(x - \mu) \leq 0$ et $\sigma \geq 0$.

Selon les valeurs du paramètre de forme γ , la distribution appartient à l'un des trois domaines d'attraction suivants :

- $\gamma > 0$: domaine de *Fréchet*, loi à *queue lourde*, tendant vers 1 comme une puissance,
- $\gamma = 0$: domaine de *Gumbel*, loi à *queue légère*, tendant vers 1 comme une exponentielle,
- $\gamma < 0$: domaine de *Weibull*, loi à *queue finie*, égale à 1 à partir d'un point donné.

Les méthodes classiques d'inférence statistique sont utilisées pour déterminer les paramètres.

- Moindres carrés : minimisation de l'erreur quadratique entre les fonctions de répartition estimée et empirique. La méthode nécessite un algorithme d'optimisation quadratique (Park et Sohn 2006).
- Moments statistiques pondérés : l'expression analytique des moments statistiques dépend des paramètres de la distribution. L'inversion de ces expressions permet le calcul des paramètres. Le domaine de validité de cette méthode est limité à $|\gamma| \leq 1/2$ (Hosking et al. 1985).
- Maximum de vraisemblance : la fonction de maximum de vraisemblance correspond à la probabilité d'obtenir les données d'apprentissage, sachant le jeu de paramètres à optimiser. Il s'agit donc de maximiser cette fonction pour estimer les paramètres les plus probables.

Kotz et Nadarajah (2000) proposent une discussion sur la formulation et l'utilisation des deux dernières méthodes.

Une fois la distribution identifiée, deux approches sont possibles. La carte de contrôle est construite :

- soit sur les maxima d'échantillons, auquel cas

$$UCL = F_{max}^{-1}(1 - \alpha) \quad (1.39)$$

¹⁷Extreme Value Distribution

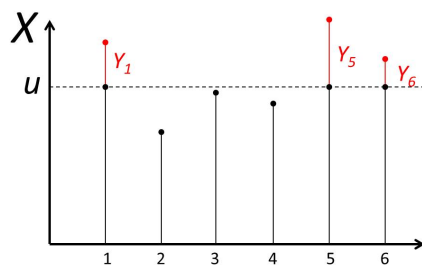


FIG. 1.13 Représentation des excès Y_i de la variable X , par rapport au seuil u .

- soit sur les valeurs individuelles de X ; dans ce cas, il faut revenir à l'estimation de F .

$$\begin{aligned}
 F_{max}(x) &= P(X_{n,n} < x) \\
 &= \prod_{i=1}^n P(X_{n,i} < x) \\
 &= F^n(x)
 \end{aligned} \tag{1.40}$$

Et :

$$UCL = F_{max}^{-1}((1 - \alpha)^n). \tag{1.41}$$

Des exemples d'application des cartes avec maxima d'échantillons sont présentés par [Oh et al. \(2009\)](#) sur des données réelles de monitoring d'un pont suspendu. [Sohn et al. \(2005\)](#) démontrent l'intérêt des EVS sur des données simulées. Cependant, lorsque cette méthode est appliquée sur une erreur de prédiction, l'avantage n'est pas significatif. En effet, la distribution d'une erreur est théoriquement normale si le modèle est bien calé.

Une remarque importante concerne la limite inférieure. Dans le cas où des données d'erreur de prédiction sont analysées, une augmentation de l'erreur avec l'apparition d'endommagement est supposée. Ainsi, seule une limite supérieure est nécessaire pour la surveillance. Si aucune supposition ne peut être formulée sur l'évolution de l'indicateur, les deux limites sont requises. Pour fixer la limite de contrôle inférieure d'une carte de maxima, il faut ajuster une EVD *de minima* sur les maxima des échantillons ([Worden et al. 2002a](#)).

Excès

Travailler avec les maxima nécessite beaucoup de valeurs disponibles pour pouvoir diviser les données en échantillons. Une alternative consiste à utiliser les excès qui représentent seulement les valeurs supérieures à un certain seuil u (Fig.1.13).

L'excès est la variable aléatoire définie par $Y = X - u$ quand $X > u$. Si le seuil est choisi suffisamment loin dans la queue de distribution de X , alors Y suit une loi de *Pareto généralisée* (GPD) dont la fonction de répartition est notée $F_u(y) = P(Y < y)$:

$$F_u(y) = \begin{cases} 1 - (1 + \gamma \frac{y}{\sigma})^{-\frac{1}{\gamma}} & \text{si } \gamma \neq 0 \\ 1 - \exp(-\frac{y}{\sigma}) & \text{si } \gamma = 0 \end{cases} \tag{1.42}$$

Il faut ensuite faire le lien entre la distribution des excès et la distribution de X :

$$\begin{aligned}
 F_u(y) &= P(Y < y) \\
 &= P(X < x | X > u) \\
 &= \frac{P(X < x \cup X > u)}{P(X > u)} \\
 &= \frac{F(x) - F(u)}{1 - F(u)} \tag{1.43}
 \end{aligned}$$

$$\text{d'où } F(x) = F_u(x - u) [1 - F(u)] + F(u) \tag{1.44}$$

Une fois que les paramètres intervenant dans l'équation 1.44 sont déterminés, $UCL = F^{-1}(1 - \alpha)$.

Il y a une équivalence rigoureuse entre la convergence des excès vers une GPD et celle des maxima vers une EVD. Les deux approches mettent en jeu 3 paramètres à ajuster pour extrapoler $F(x)$ dans ses valeurs extrêmes.

Les paramètres d'une GPD sont plus simples à déterminer que ceux d'une EVD. Sachant que u (ou $F(u)$) est une valeur choisie par l'utilisateur, il faut déterminer $F(u)$ (ou u), γ et σ . La méthode du maximum de vraisemblance et celle des moments pondérés sont appropriées pour le calcul des paramètres. Le problème est largement simplifié si l'utilisateur choisit une valeur de γ égale à 0 pour utiliser une distribution exponentielle. La variance de l'estimation de la limite supérieure est alors grandement réduite, mais le biais augmenté.

Zapico-Valle et al. (2011) optent pour ce choix concernant l'étude du benchmark UNIOVI, consistant en une structure métallique de 7,3 m de hauteur dont certains écrous sont dévissés. Les auteurs considèrent l'endommagement comme détecté si le seuil fixé par une distribution exponentielle est dépassé 1, 2 ou 3 fois consécutivement. Ce choix de considérer plusieurs valeurs permet de réduire le nombre de fausses alarmes.

Test séquentiel du rapport des vraisemblances

Contrairement à un test d'hypothèse classique dans lequel la taille de l'échantillon est fixe, un test séquentiel utilise un échantillon formé de l'accumulation de valeurs collectées au fil du temps. L'hypothèse est testée après l'ajout d'une nouvelle mesure dans l'échantillon. L'information nouvellement apportée est donc intégrée à celle issue des mesures précédentes pour être comparée aux limites du test. La finalité est d'obtenir une décision plus rapide en réduisant la taille de l'échantillon.

Une très bonne description du *test séquentiel du rapport des vraisemblances* (SPRT) est proposée par Sohn et al. (2003). La présentation suivante en reprend la formulation.

Soit à l'instant n , l'échantillon $\mathbf{X}_n = (X_1, \dots, X_n)$ issu de la variable aléatoire X , dont la distribution a pour paramètre θ . Le SPRT est posé de la façon suivante :

$$\text{Hypothèses } H_0 : \theta = \theta_0, \quad H_1 : \theta = \theta_1, \quad \theta_0 \neq \theta_1 \quad (1.45)$$

$$(\text{ ou } H_0 : \theta \leq \theta_0, \quad H_1 : \theta \geq \theta_1, \quad \theta_0 < \theta_1) \quad (1.46)$$

A chaque test, la variable de décision Z_n est calculée. Elle correspond à la probabilité d'obtenir l'échantillon \mathbf{X}_n , sachant que H_1 est vraie, divisée par la probabilité quand H_0 est vérifiée (Eq.1.47).

$$Z_n = \ln \left(\frac{f(\mathbf{X}_n | \theta_1)}{f(\mathbf{X}_n | \theta_0)} \right) \quad (1.47)$$

$$= Z_{n-1} + z_n \quad (1.48)$$

$$\text{avec } z_n = \ln \left(\frac{f(X_n | \theta_1)}{f(X_n | \theta_0)} \right) \quad (1.49)$$

Cette somme correspond également à la somme cumulée des rapports de *vraisemblance* de chaque individu de l'échantillon (Eq.1.49). Puisque le test est séquentiel, des lois d'arrêt doivent être formulées.

1. H_0 est acceptée, si $Z_n \leq b$
2. H_1 est acceptée, si $Z_n \geq a$
3. le test continue, si $b \leq Z_n \leq a$.

Si la variable de décision se situe entre les deux limites, aucune décision ne peut être prise, alors de nouvelles mesures sont requises. Pour cette raison, la taille de l'échantillon n'est pas constante.

Les limites a et b sont fixées à la fois en fonction de la probabilité d'erreur de type I (α , fausse alarme), mais également de la probabilité d'erreur de type II (β , non-détection). Dans ce test, la probabilité β peut-être calculée, car les caractéristiques de l'hypothèse alternative, θ_1 , sont connues.

$$b \simeq \ln \left(\frac{\beta}{1 - \alpha} \right) \quad \text{et} \quad a \simeq \ln \left(\frac{1 - \beta}{\alpha} \right) \quad (1.50)$$

Sohn et al. (2003) utilisent le SPRT avec des modèles de distributions de valeurs extrêmes. La première application proposée compare plusieurs approches sur des distributions simulées de lois connues. La seconde analyse les données d'une structure métallique à 3 niveaux de 1,5 m de hauteur. Au lieu de travailler sur les accélérations, les auteurs choisissent d'utiliser les différences de mesure entre deux capteurs placés de part et d'autre d'un assemblage. Un modèle AR est ajusté sur les signaux de la structure non endommagée. Enfin, l'erreur de prédiction obtenue par ce modèle sur des mesures à l'état endommagé est exploitée à l'aide d'un SPRT. Un premier test est mené en ajustant un modèle de loi normale sur l'erreur de prédiction. Le second test ajuste un modèle avec une loi extrême sur les maxima. Puisque l'erreur de prédiction doit tendre vers la normalité, les deux méthodes fonctionnent correctement. Cette étude présente de façon très complète la façon dont ces outils statistiques peuvent être manipulés. Mais leur transposition à des cas réels de suivi de structures pose problème. Tout d'abord, le fait que les échantillons soient obtenus par des divisions d'une même mesure, et

non par des groupes de mesures, ne permet pas de gain de temps dans la détection d'endommagement. Ensuite, le problème principal réside dans le fait que la valeur de θ_1 , paramètre lié à la distribution des individus dans un état dégradé, doit être connue à l'avance. La détection de nouveauté devient alors plus un problème de classification supervisée, ce qui est difficilement envisageable pour un ouvrage de génie civil.

BILAN

Ce chapitre montre que l'approche traditionnelle d'identification des caractéristiques modales des structures n'est pas en mesure, à elle seule, de gérer les fluctuations induites par la variabilité des facteurs environnementaux et opérationnels.

L'approche globale de reconnaissance de formes statistiques permet la prise en compte de ces fluctuations, sans recourir à une modélisation physique de l'ouvrage étudié. Dans le cas des structures de génie civil, pour lesquelles il n'est pas possible de disposer de mesures du comportement endommagé, le problème se pose en terme de détection de nouveauté. Le comportement sain est statistiquement caractérisé à l'aide d'une base de mesures de référence, puis les mesures classées comme *différentes*, ou *nouvelles*, sont associées à l'apparition de dégradations.

La majorité des travaux de recherche présentés dans cette partie s'appuie sur le paradigme de la détection de nouveauté pour la détection d'endommagement. Les cas d'études proposés se divisent en trois principaux groupes :

- les modèles numériques, qui ne peuvent pas rendre compte de toute la complexité du système et du bruit expérimental,
- les montages expérimentaux de laboratoire, qui présentent des mesures et des scénarios d'endommagement réalistes, mais souvent dans des conditions environnementales stables,
- les structures réelles, qui nécessitent la mesure du comportement courant sous plusieurs conditions environnementales et opérationnelles pour avoir l'assurance que les changements observés soient liés à l'endommagement. Or, cette longue phase de suivi préalable est rarement réalisée.

L'objectif de ce travail de thèse est de mettre en œuvre des outils de détection de nouveauté :

- basés seulement sur les mesures vibratoires issues de la structure,
- automatisables,
- transposables à tous types d'ouvrages,
- sur des données de monitoring en conditions ambiantes (excitations instationnaires, variations environnementales sévères).

POTENTIEL DE L'APPROCHE CHAOTIQUE

2

SOMMAIRE

INTRODUCTION	44
2.1 REPRÉSENTATION DANS L'ESPACE DES PHASES	45
2.1.1 Définition	45
2.1.2 Reconstruction de l'espace des phases	45
2.2 EXTRAPOLATION AUX SYSTÈMES STOCHASTIQUES	47
2.3 INDICATEURS POTENTIELS	48
2.3.1 Dimension de l'attracteur	48
2.3.2 Les Exposants de Lyapunov	50
2.3.3 L'erreur de prédiction non-linéaire	52
2.3.4 Caractérisation géométrique	54
BILAN	55

INTRODUCTION

LE chapitre précédent montre que concevoir la détection d'endommagement comme un problème de détection de nouveauté nécessite la détermination d'indicateurs robustes caractérisant l'état de la structure. Dans ce but, la théorie des systèmes dynamiques non-linéaires, parfois qualifiée de *théorie du chaos* et déjà utilisée dans d'autres domaines pour l'analyse de signaux complexes (Ghafari et al. 2008, Hively et al. 1999, LeBaron 1994), offre des perspectives attrayantes.

En effet, la représentation de l'évolution du système dans son espace des phases permet le calcul de nouveaux indices sensibles à des changements d'état, autrement dit à des changements de la dynamique. De plus, certaines structures peuvent présenter un comportement non linéaire. C'est le cas des ouvrages formés par de nombreux assemblages et interfaces entre matériaux, ou encore de ceux équipés de systèmes de surveillance après l'apparition de dégradations (Farrar et al. 2007, Bornn et al. 2010). Un paradigme non linéaire semble donc en mesure de détecter correctement l'apparition et/ou l'évolution de ces non-linéarités.

Dans un premier temps, ce chapitre présente succinctement le paradigme des systèmes dynamiques non linéaires, et notamment la notion fondamentale de reconstruction de l'espace des phases. Ensuite, un point est fait sur les outils de caractérisation utilisés dans le cadre de cette approche.

2.1 REPRÉSENTATION DANS L'ESPACE DES PHASES

2.1.1 Définition

Un système dynamique est considéré ici au sens mathématique, c'est à dire un système qui évolue au cours du temps de façon causale et déterministe. Ainsi, son état futur ne dépend que de son état passé, et, à une condition initiale donnée ne correspond qu'un seul état futur. Cette définition exclut donc a priori les systèmes stochastiques, dont font partie les structures soumises à une excitation aléatoire.

Le comportement d'un système dynamique est gouverné par l'équation d'évolution :

$$\dot{X} = F(X, t; p) \quad (2.1)$$

avec $X = (x_1(t), x_2(t), \dots, x_n(t))$ le vecteur formé par les différents degrés de liberté (ddl) du système (qualifiés également de variables d'état), n le nombre de ddl, F un champ de vecteurs potentiellement non linéaire et $p = (p_1, p_2, \dots, p_m)$ l'ensemble des paramètres.

La façon la plus commune de représenter l'évolution d'un système dynamique consiste à utiliser l'espace des phases (qualifié également d'espace des états), espace abstrait dont les axes sont formés par les différents degrés de libertés x_1, x_2, \dots, x_n . Chaque état occupé par le système à un instant donné représente un point dans l'espace des phases. L'ensemble de ces points figuratifs de l'état du système au cours du temps dessine la trajectoire du système dynamique dont l'évolution asymptotique est qualifiée d'attracteur. L'attracteur peut présenter différentes formes :

- dans le cas d'un régime amorti et non excité, il est représenté par un point puisque tout les ddl tendent vers zéro,
- dans le cas d'un régime oscillant entretenu, il forme une courbe fermée possédant autant de boucles que de pseudo-périodes,
- dans le cas d'un régime chaotique, il prend la forme d'un objet complexe à la structure fractale, on parle alors d'attracteur étrange.

Bien que relatif à un système parfaitement déterministe, un régime chaotique est caractérisé par un comportement d'apparence stochastique résultant d'une très forte sensibilité aux conditions initiales et rendant toute prédiction à long terme très difficile.

Dans certains cas, la modification d'un paramètre de la loi d'évolution (dit *paramètre de contrôle*) peut entraîner un changement radical de régime dynamique, on parle ici de bifurcation. Suivant l'évolution de leurs paramètres, les systèmes dynamiques sont alors susceptibles d'alterner entre régime périodique, pseudo-périodique et chaotique.

2.1.2 Reconstruction de l'espace des phases

Dans le domaine des sciences expérimentales, il est très rare de connaître a priori la nature et le nombre réel des variables d'état (ddl) d'un système observé expérimentalement et/ou de disposer de mesures pour chacune d'elles (leur nombre pouvant être potentiellement infini). Dans le pire des cas, l'information expérimentale accessible se limite à la mesure d'une seule série temporelle scalaire. Ce constat aurait pu limiter l'application de la théo-

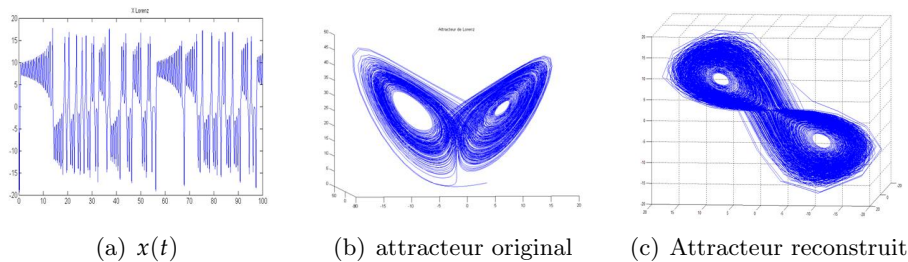


FIG. 2.1 *Reconstruction de l'attracteur du système dynamique de Lorenz à partir de sa variable $x(t)$ (Lorenz 1963)*

rie des systèmes dynamiques non linéaires (TSNL) au cas minoritaire des systèmes parfaitement définis a priori.

Cependant, l'idée introduite par Takens (1980) a considérablement élargi le champ d'application de la théorie à l'ensemble des systèmes physiques observables et a connu un succès retentissant dans le domaine des sciences expérimentales. Takens (1980) a démontré que pour des évolutions suffisamment douces, la mesure au cours du temps d'un seul degré de liberté véhicule assez d'information sur le système pour permettre la reconstruction d'un attracteur équivalent, conservant certaines propriétés fondamentales (les invariants) de l'attracteur original sous-jacent. Ainsi, la mesure des invariants de l'attracteur sous-jacent (*dimension fractale, exposants de Lyapunov,...*) est rendue accessible par un processus de reconstruction adéquat.

Le processus de reconstruction passe par la création d'un ensemble de vecteurs d'état à partir de la série temporelle disponible, $\mathbf{x} = \{x(n)|n = 1, \dots, N\}$. Ils sont de la forme :

$$\mathbf{X}(n) = [x(n), x(n - \tau), \dots, x(n - (m - 1)\tau)] \quad (2.2)$$

avec τ et m , deux nombres entiers représentant respectivement le décalage temporel (ou retard) et la dimension de reconstruction.

Takens démontre que la condition $m \geq 2d$ (d la dimension de l'attracteur original §2.3.1) doit être vérifiée pour garantir que la complexité de l'attracteur soit totalement révélée. Le retard τ peut théoriquement prendre n'importe quelle valeur entière si la longueur de la série temporelle est infinie. Cependant, en pratique, certaines méthodes sont employées pour déterminer les valeurs optimales de ces 2 paramètres de reconstruction.

Le choix du décalage temporel est lié à la corrélation entre les termes de la série. Si le décalage est trop faible, les vecteurs successifs sont presque identiques et l'information est alors trop redondante. En revanche, s'il est trop fort, les différentes coordonnées sont complètement décorréelées et l'influence de la loi d'évolution sous-jacente est perdue. Deux méthodes sont principalement employées pour déterminer le meilleur compromis :

- choisir le décalage correspondant au premier zéro de la fonction d'auto-corrélation de \mathbf{x} ,
- choisir le premier minimum de la fonction d'information mutuelle de la série (Fraser et Swinney 1986). Cette fonction, basée sur la théorie de l'information, est liée à la dépendance statistique des termes de la série.

La dimension de reconstruction doit être suffisante pour révéler toute

la complexité de l'attracteur sans pour autant alourdir le traitement des données. Deux techniques sont préconisées (Brown et al. 1991) :

- *la méthode des faux voisins* : elle se base sur l'idée que lorsqu'un objet en dimension m est projeté en dimension $m - 1$, certains points initialement éloignés deviennent voisins. On parle alors de *faux voisins*. Inversement, en partant de la dimension 1, le nombre de faux voisins à chaque passage à la dimension supérieure devrait diminuer, jusqu'à devenir nul lorsque la topologie de l'attracteur est entièrement dévoilée.
- l'analyse des valeurs singulières. La série est transformée en une série de vecteurs de dimension importante qui constitue la matrice des données. L'analyse des valeurs singulières de cette matrice ou des valeurs propres de sa matrice de covariance, fournit des informations sur la dimension optimale de reconstruction.

La méthode des faux voisins est cependant la plus employée.

Pecora et al. (2002) généralisent la procédure de reconstruction à partir de plusieurs séries temporelles simultanément traduisant les évolutions de différentes variables du système. L'intérêt est d'intégrer davantage d'informations concernant l'attracteur. Cependant, le résultat produit un espace des phases de dimension importante dans lequel les calculs sont plus lourds.

Une fois l'attracteur reconstruit, l'étude ne se fait plus sur une série d'apparence aléatoire, mais sur une figure géométrique structurée (Fig.2.1).

2.2 EXTRAPOLATION AUX SYSTÈMES STOCHASTIQUES

L'emploi de la reconstruction de l'espace des phases à partir du théorème de *Takens* et les indicateurs présentés dans la partie suivante (§2.3) supposent que le comportement de la structure est déterministe. Rien ne garantit pourtant que les mesures relevées sur les ouvrages présentent ce type de comportement. La plupart du temps, leur comportement est stochastique linéaire, dans certains cas non-linéaires, mais rarement chaotique. Or la sensibilité des systèmes chaotiques aux faibles perturbations représente une propriété intéressante vis à vis de la détection de l'endommagement.

Pour bénéficier de cette sensibilité, une possibilité consiste à apporter du chaos dans le comportement dynamique de la structure. Si la forme de la sollicitation appliquée à l'ouvrage est chaotique, la réponse présentera des caractéristiques de séries chaotiques. Dans de nombreux travaux, une série issue du système convectif de *Lorenz* est choisie pour former la sollicitation (Todd et al. 2001, Nichols et al. 2003c, Todd et al. 2002; 2004, Nichols et al. 2003b; 2006, Moniz et al. 2005). Il faut alors s'assurer que l'ensemble structure/sollicitation possède une dimension réduite, ce qui est réalisé en jouant sur les caractéristiques du système chaotique appliqué. Cependant, cette approche relève d'un contrôle actif des structures et n'est donc pas adaptée à une démarche de suivi en conditions ambiantes.

Dans le cas du contrôle passif, la tentation est forte d'extrapoler aux systèmes stochastiques l'utilisation des outils développés pour des systèmes déterministes, et plus particulièrement pour les systèmes chaotiques.

Utilisant l'erreur de prédiction comme indicateur (§2.3.3), Nichols et al.

(2003a) comparent les résultats obtenus avec des sollicitations chaotiques à ceux produits à l'aide de sollicitations aléatoires (bruits de différentes bandes de fréquence). Les résultats montrent que, malgré le caractère aléatoire de la sollicitation, si elle est de bande de fréquence limitée, la réponse de la structure est telle qu'elle peut être considérée comme un système de faible dimension. L'erreur de prédiction est ici employée avec succès pour déceler le desserrage des boulons sur une plaque métallique, bien que les résultats soient meilleurs avec les signaux chaotiques. Overbey et al. (2007) étudient également la possibilité d'utiliser des bruits de différentes bandes de fréquence pour exciter un portique en aluminium et un système masse-ressort. Ils montrent que la réponse de la structure est suffisamment corrélée pour pouvoir reconstruire des attracteurs de dimension réduite et exploiter les outils de la théorie du chaos.

Il est également possible d'avoir recours à une excitation impulsionnelle. Cependant, dans ce cas, la réponse n'est pas stationnaire. Pour résoudre ce problème, Casciati et Casciati (2006) ne sélectionnent que la partie la plus stable de la réponse et l'additionne à d'autres morceaux pour créer une grande série. Cette méthode est appliquée à une façade maçonnée endommagée par un séisme pour en extraire les exposants de *Lyapunov* (§2.3.2). La réponse impulsionnelle est également fréquemment utilisée dans le domaine du diagnostic par ultrasons. Dans ce contexte, Lu et Michaels (2005) font intervenir le chaos en convoluant la réponse impulsionnelle et une série chaotique. Le résultat du produit représente la réponse de la structure si elle était soumise à une force excitatrice chaotique, sous réserve que son comportement soit linéaire.

Ces travaux montrent qu'en dépit d'un comportement stochastique, l'utilisation de la reconstruction de l'espace des phases peut présenter un intérêt en terme de détection de modifications structurales.

2.3 INDICATEURS POTENTIELS

Il existe de nombreux outils qui permettent de caractériser un système dynamique en exploitant sa trajectoire dans l'espace des phases reconstruit. Les plus répandus sont présentés ci-après, accompagnés d'exemples d'utilisation dans le domaine de la détection d'endommagement. Une description plus complète est disponible dans l'article de (Abarbanel et al. 1993) ou dans l'ouvrage de Kantz et Schreiber (2004).

2.3.1 Dimension de l'attracteur

La dimension de l'attracteur, notée d_f , est une caractéristique topologique décrivant la façon dont celui-ci occupe l'espace qui le contient. Par exemple, l'attracteur d'un pendule à 2 ddl en régime périodique est une ligne fermée. On parle de *cycle limite* dont la dimension d_f est égale à 1 (dimension euclidienne classique associée à une ligne). Pour un système amorti libre, la trajectoire tend asymptotiquement vers un attracteur de type *point fixe* pour lequel la dimension d_f est égale à 0.

Dans le cas des systèmes chaotiques, l'attracteur est un objet topologique complexe présentant une structure fractale caractérisée par la pro-

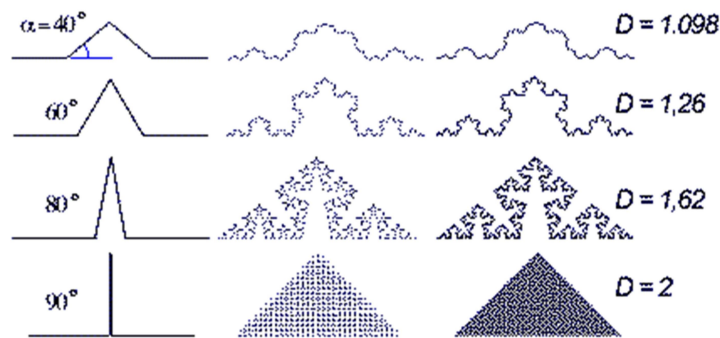


FIG. 2.2 Exemple d'objet fractal, plus l'objet est proche de la ligne droite, plus sa dimension est proche de 1

priété d'auto-similarité (ou invariance d'échelle) et par une dimension fractionnaire (*dimension fractale*). La géométrie fractale est née sensiblement en même temps que la théorie du chaos dans les années 1960 (Mandelbrot 1983) et ces deux sciences se sont tout d'abord développées indépendamment l'une de l'autre pour se rejoindre quelques années plus tard. Un attracteur chaotique remplit l'espace des phases d'une façon très irrégulière en apparence et cette irrégularité est reproduite à l'identique à toutes les échelles. Cependant, cette irrégularité apparente est parfaitement traduite par la notion de dimension fractale. Ainsi, un objet dont la dimension fractale est comprise entre 1 et 2 n'est ni une ligne, ni un plan. A titre d'exemple, la dimension fractale de l'objet *trajectoire brownienne d'une particule de pollen à la surface d'un liquide* tend vers 2 dans la mesure où la trajectoire (ligne brisée) tend à occuper tout l'espace disponible (en l'occurrence un plan). La figure 2.2 montre quatre objets fractals dont la dimension augmente avec le degré de complexité et atteint des valeurs non entières, comprises en 1 et 2.

Un système dynamique à N ddl est ainsi représenté par un attracteur de dimension inférieure à N dans la mesure où un attracteur ne peut en aucun cas remplir complètement l'espace des phases.

La dimension de corrélation, notée d_2 pour *dimension généralisée d'ordre 2*, est la méthode la plus répandue pour estimer numériquement la dimension fractale de l'attracteur. Sa détermination est basée sur l'hypothèse selon laquelle le nombre de points situés à l'intérieur d'une hyper-sphère, centrée sur un point donné de l'attracteur, est proportionnel à son rayon (ϵ) élevé à une puissance correspondant à la dimension de l'attracteur. Grassberger et Procaccia (1983) reformulent cette hypothèse pour proposer un algorithme de calcul rapide de la dimension de corrélation basé sur la probabilité C que deux points de l'attracteur soient distants d'une quantité inférieure ou égale à ϵ :

$$C(\epsilon) = \frac{2}{(N - n_{min})(N - n_{min} - 1)} \sum_{i=1}^N \sum_{j=i+1+n_{min}}^N \Theta(\epsilon - \|X_i - X_j\|) \quad (2.3)$$

avec N , le nombre de points de l'attracteur, $\Theta(x)$ la fonction de *Heaviside*, X_i les points de l'attracteur et n_{min} la taille (en pas de temps) de la fenêtre d'exclusion des points corrélés temporellement à chaque X_i (Theiler 1986). La fonction $C(\epsilon)$, ainsi obtenue numériquement, est définie comme l'intégrale de corrélation. Puisque $C(\epsilon) \propto \epsilon^{d_2}$, il suffit d'ajuster une droite

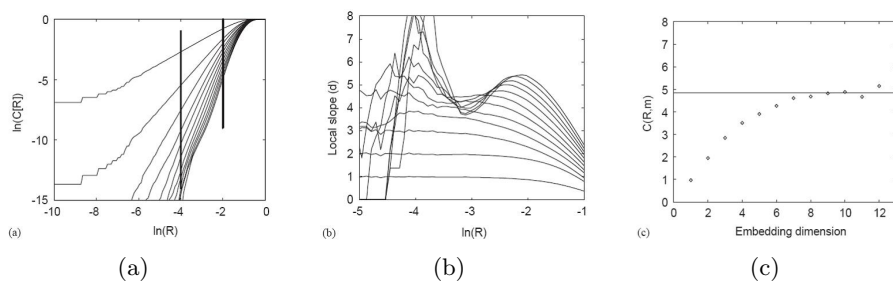


FIG. 2.3 Détermination de d_2 a) Tracé des intégrales de corrélation pour plusieurs dimensions de reconstruction avec la délimitation de la zone linéaire des courbes - b) Tracé des pentes locales et recherche d'un palier - c) Estimation de d_2 pour chaque dimension de reconstruction et recherche d'une stabilisation (Nichols et al. 2003c)

sur la courbe $\log(\epsilon) - \log(C(\epsilon))$ pour estimer la dimension de corrélation. En réalisant l'opération pour chaque dimension de reconstruction, la dimension estimée doit se stabiliser lorsque l'attracteur est pleinement déployé dans un espace de dimension suffisante, fournissant ainsi la valeur de d_2 .

La figure 2.3 montre l'application de la méthode sur les données issues de Nichols et al. (2003c) et illustre les imprécisions pouvant affecter l'estimation de la dimension de corrélation. En effet, la zone de plateau qui la détermine n'apparaît pas toujours clairement et le bruit peut polluer les courbes obtenues rendant leur interprétation difficile. Le calcul de la dimension de corrélation est par conséquent difficilement automatisable.

La dimension de corrélation a été beaucoup étudiée dans le cadre du suivi de systèmes mécaniques tournants. Cette caractéristique topologique est utilisée dans ce cas comme indicateur d'endommagement des engrenages ou des roulements (Craig et al. 2000, Ghafari et al. 2008, Yang et al. 2007). Logan et Mathew (1996) étudient l'influence des différents paramètres de reconstruction de l'espace des phases sur le calcul de la dimension de corrélation. Dans cette étude, l'endommagement est introduit sur un roulement en réalisant des entailles sur une bague. Les auteurs montrent que, dans certains cas, la dimension de corrélation est délicate à déterminer. Ces difficultés sont également relayées par Nichols et al. (2003c) dans une étude portant sur une poutre cantilever en aluminium excitée par une sollicitation chaotique, les changements dynamiques étant induits par la diminution de la force de serrage au niveau d'un encastrement.

Sa capacité à caractériser les systèmes dynamiques est néanmoins démontrée par les travaux de Kacimi et Laurens (2009) dans lesquels des sources ultrasonores d'origines différentes sont correctement distinguées par la dimension de corrélation.

2.3.2 Les Exposants de Lyapunov

Les exposants de Lyapunov représentent le taux moyen avec lequel une petite perturbation, appliquée au système, est amplifiée dans chaque direction de l'espace des phases. Soit une petite perturbation $\delta_{i,0}$ appliquée au temps $t = 0$ dans la direction i . Au temps t , elle vaudra :

$$\delta_{i,t} = \delta_{i,0} \cdot e^{\lambda_i t} \quad (2.4)$$

avec λ_i l'exposant de *Lyapunov* dans la direction i .

Le signe des exposants est particulièrement important. Si tous les exposants de *Lyapunov* sont négatifs, les perturbations sont amorties dans toutes les directions. C'est le cas des systèmes dissipatifs libres. Un des exposants est généralement nul. Il représente la direction tangente à la trajectoire, dans laquelle la perturbation introduite n'est pas modifiée. Enfin, un régime est chaotique si au moins un exposant est positif, indiquant une direction d'instabilité dans laquelle la perturbation est amplifiée.

L'intérêt porté aux exposants de Lyapunov provient également du fait qu'ils constituent une mesure de la prédictibilité du système. Plus un exposant positif est grand, plus le système est sensible aux petites perturbations qui sont alors amplifiées rapidement, le rendant par conséquent imprévisible.

Un autre moyen de concevoir les exposants de Lyapunov est de considérer une hyper-sphère autour d'un point de la trajectoire. En suivant son évolution le long de la trajectoire, elle se transforme en un hyper-ellipsoïde. La moyenne logarithmique des taux d'expansion des ses axes principaux forme les exposants de Lyapunov (Geist et al. 1990).

Les différentes méthodes utilisées pour les estimer sont présentées dans l'ouvrage de Kantz et Schreiber (2004) et l'article de Brown et al. (1991). L'algorithme pratique de calcul à partir de données expérimentales est détaillé dans le chapitre §3.1.3. Il se base sur l'estimation locale de la matrice jacobienne de la dynamique à partir d'un point initial et tout au long d'une trajectoire issue de ce point. Pour obtenir une estimation fiable des exposants, la matrice jacobienne doit être suivie sur plusieurs milliers de pas de temps, représentant un coût très important en temps de calcul. De plus, ce calcul doit être réitéré à partir d'autres points initiaux.

Le spectre des exposants est étroitement lié à la dimension fractale puisqu'ils indiquent quelles sont les directions de l'espace qui sont instables. La conjecture de *Kaplan Yorke* (Kaplan et Yorke 1979) permet d'estimer la dimension fractale de l'attracteur à partir de son spectre de *Lyapunov*.

$$d_L = M + \frac{\sum_{n=1}^M \lambda_n}{-\lambda_{M+1}} \quad (2.5)$$

Lorsque les exposants sont classés dans l'ordre décroissant, M est le nombre maximum d'exposants qu'il faut additionner avant que la somme ne devienne négative.

Quelques études exploitent les exposants de *Lyapunov* comme indicateurs d'endommagement. Une d'entre elles est réalisée par Livingston et al. (2001) sur un modèle éléments finis de pont. Outre le recours aux exposants de Lyapunov pour déceler l'endommagement, une originalité de ce travail réside dans la création d'un outil de simulation de l'excitation relative au trafic routier. Deux fissures sont introduites dans les modèles pour représenter l'état endommagé de la structure. Pour chaque ddl, un attracteur est reconstruit à partir des données de déplacement, et l'exposant de *Lyapunov* maximal est calculé. Le tracé de la variation de cet exposant sur toute la longueur fait apparaître deux pics situés au droit des fissures, illustrant le potentiel de la méthode dans le cas de l'instrumentation dense d'un ouvrage.

Casciati et Casciati (2006) étudient la reproduction à l'échelle 1/2 d'une façade de bâtiment historique instrumentée par un ensemble d'accéléromètres. Après avoir été endommagée par une sollicitation sismique (état initial), la façade a été réparée (renforcement et colmatage des fissures). Chaque capteur est ici considéré comme un ddl utilisé pour la représentation dans l'espace des phases. Les séries temporelles analysées résultent de la concaténation de morceaux *pseudo-stationnaires* des mesures originales, obtenues sous bruit ambiant ou marteau d'impact. Les résultats montrent que, parmi les composantes du spectre de Lyapunov et la dimension associée (d_L), seul l'exposant maximal est sensible à la réparation.

2.3.3 L'erreur de prédiction non-linéaire

L'exploitation des trajectoires de l'attracteur pour prévoir les valeurs futures du système permet de créer un modèle prédictif non linéaire entièrement guidé par les données recueillies (*data driven models*), sans identifier de paramètres, ni formuler d'hypothèses a priori, à l'exception de l'hypothèse de stationnarité du processus étudié. Dans le cadre du SHM, le but sera principalement de déterminer avec quelle précision une série de référence relative à un système peut prédire une autre série générée ultérieurement par ce même système, ou, dans le cas présent, avec quelle précision un attracteur de référence peut prédire une trajectoire. Si le système n'a pas observé de changement dynamique majeur, l'attracteur de référence est supposé pouvoir prédire de façon satisfaisante une série collectée ultérieurement. L'erreur de prédiction attendue dans ce cas est faible. Par contre, si une modification de la dynamique est intervenue entre temps (due par exemple à l'initiation d'un endommagement), l'attracteur de référence ne sera plus approprié pour prédire la nouvelle série et l'erreur de prédiction devrait croître de façon significative.

A partir de la série de référence $x(1), \dots, x(N_x)$, les points \mathbf{X} de l'attracteur de référence sont reconstruits, $\mathbf{X} = \{X_1, \dots, X_{N_x - (m-1)\tau}\}$. Le même protocole est appliqué à la série à tester $y(1), \dots, y(N_y)$ afin d'établir sa trajectoire réelle $\mathbf{Y} = \{Y_1, \dots, Y_{N_y - (m-1)\tau}\}$ dans l'espace des phases reconstruit. On cherche alors à prévoir la trajectoire \mathbf{Y} à partir de l'attracteur de référence. Un point Y_n est choisi et, conformément à l'équation 2.6, son évolution après Δn pas de temps est prédite par l'évolution moyenne $\hat{Y}_{n+\Delta n}$ du voisinage constitué par tous les points X_i de l'attracteur de référence situés à une certaine distance ϵ de Y_n (voisinage noté $\mathcal{V}_\epsilon(Y_n)$). L'erreur de prédiction est alors égale à la différence entre l'évolution réelle du point Y_n et l'évolution moyenne de son voisinage après Δn pas de temps. La figure 2.4 illustre le calcul de l'erreur de prédiction.

$$\hat{Y}_{n+\Delta n} = \frac{1}{|\mathcal{V}_\epsilon(Y_n)|} \sum_{X_i \in \mathcal{V}_\epsilon(Y_n)} X_{i+\Delta n} \quad (2.6)$$

avec $|\mathcal{V}_\epsilon(Y_n)|$ le nombre de points contenus dans le voisinage.

Le calcul de l'erreur de prédiction est réalisé sur un grand nombre de points et les résultats sont moyennés afin d'obtenir un indicateur valable sur

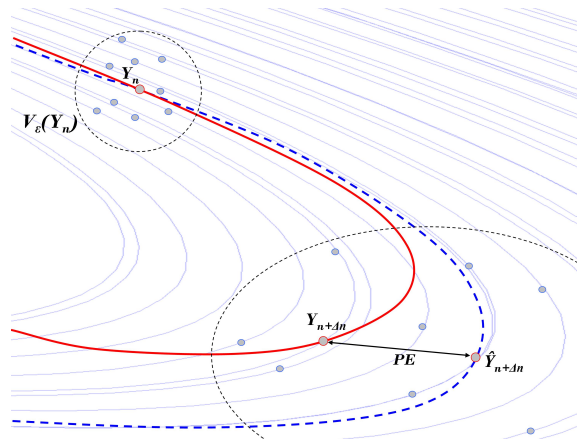


FIG. 2.4 Les traits bleus représentent les trajectoires dans l'attracteur de référence. Le point Y_n est prédit à l'aide de l'évolution du voisinage de points de l'attracteur de référence situés autour de lui. La ligne rouge matérialise l'évolution réelle de Y_n et la ligne bleue discontinue l'évolution moyenne du voisinage. L'erreur de prédiction est égale à la différence entre ces deux trajectoires.

l'ensemble de l'attracteur. Il est généralement recommandé que le nombre de points dans le voisinage soit compris entre $10^{-4}N$ et $10^{-3}N$ et de réaliser $0,05N$ prédictions sur l'attracteur (Nichols 2003).

L'erreur de prédiction (PE) est très utilisée dans la littérature. Cependant, les études proposées, dont quelques exemples sont décrits ci-après, se limitent à des dispositifs de laboratoire sous conditions environnementales peu variables.

Todd et al. (2004) exploitent l'erreur de prédiction croisée entre les capteurs. Les auteurs supposent que la réduction du couple de serrage dans un assemblage aura pour conséquence la diminution de la capacité d'un capteur à prédire son voisin s'ils sont placés de part et d'autre de l'assemblage (Fig.2.5). La structure est sollicitée par un signal chaotique. La PE permet de détecter une diminution du serrage avant l'apparition d'un jeu et se montre plus sensible que l'erreur de prédiction obtenue avec un modèle auto-régressif. Cependant, dans certains cas, les auteurs remarquent de façon contre-intuitive une diminution de l'erreur de prédiction.

Des variantes sont proposées par Olson et al. (2005). La structure est la même que dans l'étude précédente, mais sollicitée par des bruits de bandes de fréquences différentes et des signaux chaotiques. Outre l'exploitation de la correspondance géométrique entre la trajectoire à prédire et l'attracteur *prédicatif*, les auteurs étudient également une correspondance temporelle (à la condition que les signaux soient mesurés au même instant). Ils proposent également la prédiction de la trajectoire moyenne d'un ensemble de points à la place de la trajectoire d'un point isolé. En présence de bruit, cette dernière approche associée à la correspondance géométrique se montre la plus efficace.

Enfin, une étude paramétrique très complète est menée par Overbey et al. (2007). Les auteurs évaluent les effets de la méthode de reconstruction, de la taille des voisinages, du pas de prédiction (Δn) et la nature de l'excitation. Il en ressort que, malgré la sensibilité de la PE à la qualité de

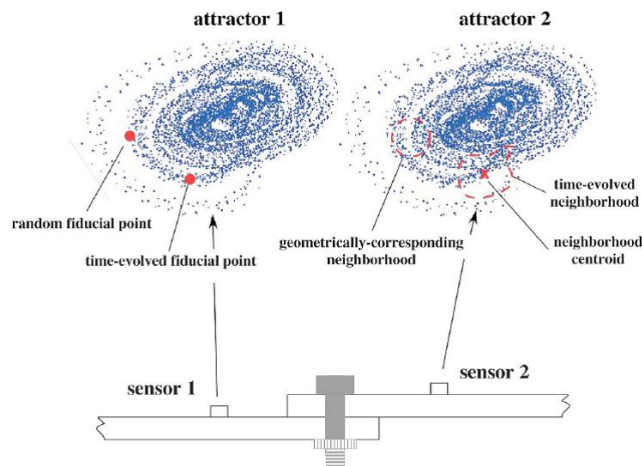


FIG. 2.5 Méthode de l'erreur de prédiction croisée ((Todd et al. 2004))

la reconstruction, il est possible d'utiliser une dimension inférieure à celle indiquée par la méthode des faux voisins.

D'autres exemples d'applications sont présentés par Nichols et al. (2003b;a), Nichols (2003).

2.3.4 Caractérisation géométrique

Le travail dans l'espace des phases donne la liberté de créer de nouveaux indicateurs caractérisant la déformation de l'attracteur liée à l'endommagement. Olson et al. (2005) proposent, par exemple, un indice basé sur l'évolution des angles formés par les segments reliant un point de référence et ses plus proches voisins.

L'article de Todd et al. (2001) introduit l'indice *ALAVR*, pour *Average Local Attractor Variance Ratio*. Il traduit le rapport de variance entre un voisinage de l'attracteur de référence (capteur ou excitation en contrôle actif, par exemple) et son équivalent temporel sur l'attracteur testé. L'approche est appliquée sur un modèle masse/ressort simulé. Les raideurs présentent une distribution gaussienne, pour tenir compte des variations environnementales, et du bruit est ajouté aux mesures. La détection de la réduction de raideur d'un ressort est meilleure avec le *ALAVR* qu'avec les caractéristiques modales courantes.

Un autre exemple s'intéresse à la divergence des trajectoires. Il est gênant de parler d'exposants de *Lyapunov*, porteurs d'un sens physique, lorsque la nature des séries étudiées est plutôt stochastique. Le CAAD (*Chaotic Amplification of Attractor Distortion*) reflète la façon dont des trajectoires initialement voisines divergent sans pour autant évoquer les exposants de Lyapunov (Moniz et al. 2005). Cet indice est utilisé pour comparer deux attracteurs. Pour cela, un point de l'attracteur de référence et son plus proche voisin dans l'attracteur à tester sont sélectionnés. Soit T , le temps que mettent en moyenne les trajectoires de l'attracteur de référence pour diverger, le CAAD est la distance entre ces points après un temps T . Ce calcul est réalisé sur un très grand nombre de points pour avoir une vision significative de l'ensemble de l'attracteur.

BILAN

Le paradigme des systèmes dynamiques non linéaires permet la représentation des séries scalaires dans un espace à plusieurs dimensions. Cette représentation offre la possibilité de calculer des invariants dynamiques tels que la *dimension de corrélation* ou les *exposants de Lyapunov*. Cependant, ces quantités présentent moins de sens physique lorsqu'elles sont calculées sur des systèmes stochastiques. De plus, l'estimation de la *dimension de corrélation* est difficilement automatisable, et celle du *spectre de Lyapunov* requiert un temps de calcul très important.

Les structures de génie civil soumises à des sollicitations ambiantes étant des systèmes dynamiques stochastiques, il est préférable d'exploiter l'espace des phases reconstruit par le biais d'outils de caractérisation géométrique de la répartition des points, tels que l'erreur de prédiction non linéaire, l'ALAVR ou le CAAD. De tels outils n'ont encore jamais été testés sur des données réelles intégrant des variations environnementales.

Les chapitres suivants introduisent un nouvel outil de caractérisation d'attracteur proposé dans le cadre de ces travaux de thèse et inspiré de l'algorithme de calcul des *exposants de Lyapunov*. Pour réellement mesurer l'efficacité de cet indicateur original en terme de détection d'endommagement, il faut être en mesure de générer des données intégrant les deux difficultés propres aux ouvrages de génie civil : variations environnementales et excitation instationnaire liée au trafic, au vent, etc. A cet effet, des dispositifs expérimentaux reproduisant ces contraintes sont développés et l'indicateur proposé est également confronté aux exposants de Lyapunov et aux coefficients auto-régressifs.

Deuxième partie

Méthodes de calcul et cas d'études expérimentaux

MÉTHODES DE CALCUL

3

SOMMAIRE	
INTRODUCTION	60
3.1 INDICATEURS	61
3.1.1 Coefficients auto-régressifs	61
3.1.2 Reconstruction de l'espace des phases	63
3.1.3 Exposants de Lyapunov	66
3.1.4 Indicateur Jacobien	69
3.2 NORMALISATION PAR LA DISTANCE DE MAHALANOBIS	70
3.3 MODÉLISATION STATISTIQUE	71
3.3.1 Test sur les valeurs individuelles : $N_e = 1$	71
3.3.2 Test sur un échantillon : $N_e > 1$	73

INTRODUCTION

Ce chapitre décrit de façon détaillée les concepts et méthodes de calcul spécifiques associés aux développements scientifiques présentés dans la suite du document. La présentation est segmentée en trois parties selon le processus général de la détection de nouveauté introduit dans le premier chapitre :

- extraction des indicateurs,
- normalisation,
- modélisation statistique pour la fixation du seuil de rejet.

Une série temporelle sera exprimée soit sous la forme d'un ensemble de valeurs scalaires $\{x_1, \dots, x_N\}$, soit sous la forme d'un ensemble de vecteurs de dimension m reconstruits dans l'espace des phases, $\{X_1, \dots, X_N\}$.

3.1 INDICATEURS

3.1.1 Coefficients auto-régressifs

Un modèle auto-régressif (AR) représente un processus stochastique linéaire stationnaire. Il vise à prédire les valeurs futures d'une série temporelle à partir de ses valeurs passées. Ce type de processus correspond au passage d'un bruit blanc à travers un filtre linéaire. Dans un modèle AR d'ordre p , $AR(p)$, chaque valeur est une combinaison linéaire des p valeurs précédentes à laquelle un terme d'erreur non nul est ajouté.

$$x_i = \sum_{j=1}^p a_j x_{i-j} + \epsilon_i \quad (3.1)$$

$$\hat{x}_i = \sum_{j=1}^p a_j x_{i-j} \quad (3.2)$$

avec a_j les coefficients auto-régressifs, ϵ_i les termes d'erreur qui doivent être indépendants, de moyenne nulle et normalement distribués, et \hat{x}_i la prédiction de x_i .

Deux méthodes sont couramment employées pour l'estimation des paramètres a_j : la méthode de *Yule Walker* et la méthode des *moindres carrés* (Box et Jenkins 1994). Les résultats sont proches pour une estimation sur un grand nombre de données. Les développements proposés dans ce travail de thèse (Partie III) s'appuient sur la méthode des moindres carrés. Celle-ci nécessite l'inversion de la matrice formée par les estimateurs non-biaisés c_{ij} de la fonction d'auto-corrélation de x au décalage $i - j$ (Eq. 3.3 et 3.4).

$$\begin{pmatrix} c_{11} & c_{12} & \cdots & c_{1p} \\ c_{21} & c_{22} & \cdots & c_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ c_{p1} & c_{p2} & \cdots & c_{pp} \end{pmatrix} \begin{pmatrix} \hat{a}_1 \\ \hat{a}_2 \\ \cdots \\ \hat{a}_p \end{pmatrix} = \begin{pmatrix} c_{01} \\ c_{02} \\ \vdots \\ c_{0p} \end{pmatrix} \quad (3.3)$$

avec

$$c_{ij} = \frac{1}{N-p} \sum_{t=p+1}^N x_{t-i} x_{t-j} \quad (3.4)$$

Le point le plus délicat réside dans la détermination de l'ordre du modèle. Il faut parvenir à déterminer une valeur de p capable de capter la dynamique de la structure, sans pour autant modéliser le bruit.

La *fonction d'auto-corrélation partielle* (PACF¹) mesure la dépendance entre les valeurs d'un processus pour différents décalages temporels. Ainsi, le terme Θ_k de cette fonction représente la corrélation entre les termes x_t et x_{t-k} en retirant l'influence des variables $x_{t-1}, \dots, x_{t-k+1}$. D'un point de vue pratique, le terme Θ_k correspond également au dernier coefficient du modèle $AR(k)$ ajusté sur la série temporelle. La PACF est donc estimée en

¹Partial Auto-Correlation Function

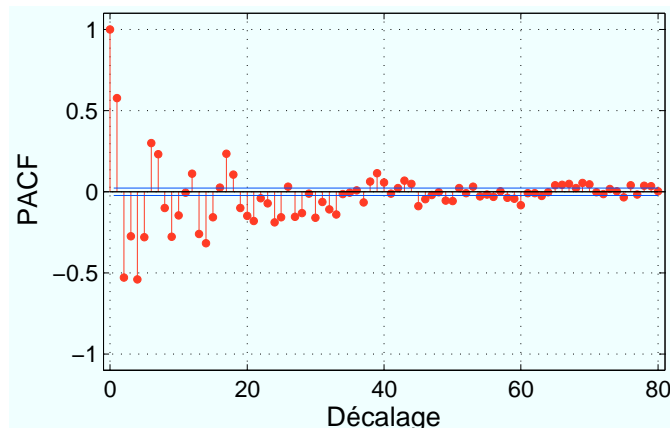


FIG. 3.1 Fonction d'auto-corrélation partielle d'une mesure d'accélération sous bruit blanc.

ajustant des modèles auto-régressifs d'ordre croissant et en conservant le dernier terme à chaque étape.

Il a été démontré que, pour un $AR(p)$, Θ_k est nul pour $k > p$. Sur des mesures réelles, ces valeurs ne sont pas nulles, mais présentent une distribution aléatoire. Cette propriété est exploitée pour choisir p . Dans l'exemple de la figure 3.1, les coefficients semblent devenir aléatoires après 45-50. L'ordre optimal est donc inférieur à 50.

Pour confirmer le choix de l'ordre du modèle AR, il est préférable de coupler la méthode précédente à d'autres critères de sélection de p . Le *critère d'information d'Akaike* (AIC) se base sur la théorie de l'information (Eq.3.5). La minimisation de ce critère traduit un compromis entre la complexité du modèle et sa précision, l'objectif étant d'éviter le sur-apprentissage.

$$AIC(p) = (N - p) \ln(\epsilon_{rms}) + 2p \quad (3.5)$$

$$\text{avec } \epsilon_{rms} = \frac{1}{(N - p)} \sum_{i=p+1}^N (x_i - \hat{x}_i)^2 \quad (3.6)$$

Une fois l'ordre choisi et le modèle ajusté, il convient d'en vérifier la pertinence par le biais de son erreur. Le caractère indépendant des termes ϵ_i est contrôlé en calculant leur densité spectrale de puissance qui doit être constante. Le tracé de leur distribution et du modèle gaussien de même moyenne et écart-type permet de s'assurer de la normalité de la distribution.

Figureiredo et al. (2010) proposent une discussion sur la sélection de l'ordre d'un modèle AR en comparant les différentes approches existantes. Même en utilisant l'ordre le plus faible indiqué par ces approches, ils parviennent à détecter l'endommagement introduit dans une structure de laboratoire.

Une fois le modèle ajusté, deux types d'exploitation peuvent être envisagés pour la détection de l'endommagement dans une structure (§1.3.1) :

- formation du vecteur composé des coefficients auto-régressifs a_i comme indicateur caractérisant la mesure,

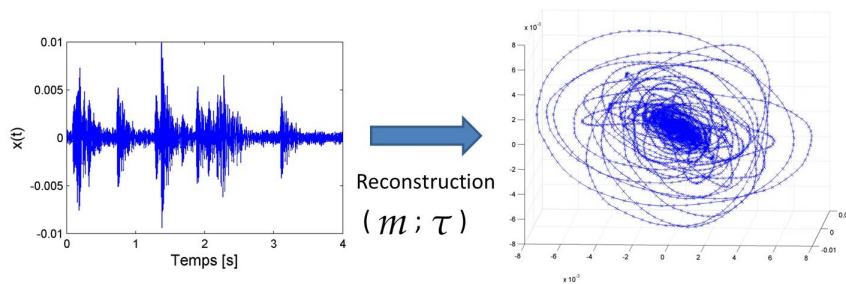


FIG. 3.2 Représentation d'une mesure d'accélération dans un espace des phases reconstruit.

- utilisation du modèle ajusté sur une mesure pour prédire une autre mesure (*AR croisé*) ; dans ce cas, l'erreur de prédiction constitue l'indicateur extrait pour l'étape de normalisation.

3.1.2 Reconstruction de l'espace des phases

La reconstruction de l'espace des phases vise à sonder l'attracteur sous-jacent du système ayant généré la série temporelle observée. Elle consiste à représenter cette série scalaire, $\{x_1, \dots, x_N\}$, dans un espace à plusieurs dimensions par la formation de vecteurs reconstruits, $\{X_1, \dots, X_{N'}\}$ (Fig.3.2). Nichols et Nichols (2001) présentent la démarche de façon concrète. De plus amples informations sont fournies dans l'article de Abarbanel et al. (1993).

La reconstruction des vecteurs d'états fait apparaître deux paramètres : la dimension de reconstruction m (parfois appelée *dimension de plongement* ou *d'immersion*), et le décalage temporel τ (également qualifié de *délai* ou *retard*) (Eq.2.2). D'un point de vue mathématique, pour un système sans bruit, les attracteurs reconstruits en dimension m avec des délais τ_1 et τ_2 sont équivalents. En pratique, les mesures sont systématiquement bruitées, voire issues d'un système stochastique. Dans ce cas, une valeur optimale de τ doit être déterminée.

Délai

Si le délai est trop faible, les composantes des vecteurs reconstruits sont trop corrélées. Les points sont alors regroupés le long d'une droite passant par l'origine et le point de coordonnées $(1; \dots; 1)$ de l'espace \mathcal{R}^m . En revanche, si τ est trop grand, alors les vecteurs d'état sont quasiment indépendants et les points forment un nuage homogène dans l'espace des phases reconstruit.

Il est donc nécessaire de trouver un compromis. La première approche intuitive s'appuie sur la fonction d'auto-corrélation (FAC) (Eq.3.7) de la mesure. Le choix du délai optimal τ est alors défini lorsque la FAC devient inférieure à un seuil limite. Ce seuil est le plus souvent fixé à zéro et le délai est donc associé à la première racine de la FAC. Si la série x est de moyenne nulle, un estimateur non biaisé de la FAC est donné par :

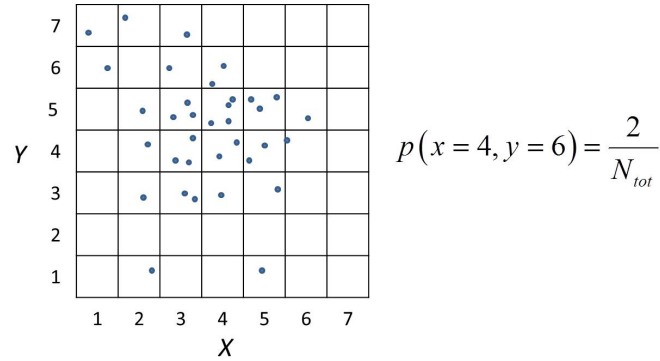


FIG. 3.3 Estimation des probabilités marginales et jointes (Nichols et Nichols 2001)

$$\hat{R}(\tau) = \frac{1}{N - \tau} \sum_{i=1}^{N-\tau} x_i x_{i+\tau}. \quad (3.7)$$

Cependant, un grand nombre des systèmes étudiés sont susceptibles de présenter des corrélations non-linéaires qui ne sont pas prises en compte par la FAC. Pour pallier ce problème, l'utilisation de la *fonction d'information mutuelle* (FIM) est proposée par Fraser et Swinney (1986). Développée dans le cadre de la théorie de l'information, elle indique (en *bits*) la quantité d'information disponible sur $x_{i+\tau}$, en connaissant x_i . En notant $p(x_i)$ et $p(x_i, x_j)$, respectivement probabilité marginale et probabilité jointe, la FIM s'écrit :

$$IM(\tau) = \sum_k p(x_k, x_{k+\tau}) \log \left(\frac{p(x_k, x_{k+\tau})}{p(x_k)p(x_{k+\tau})} \right). \quad (3.8)$$

L'estimation des probabilités marginales et jointes est réalisée à l'aide d'histogrammes à une et deux dimensions (Fig.3.3).

La taille des boîtes qui divisent l'espace est constante. Une optimisation est possible en considérant des tailles variables en fonction de la densité de l'espace (Fraser et Swinney 1986). Cependant, l'objectif n'est pas d'avoir une connaissance absolue de la FIM, mais plutôt relative par rapport à τ . Dans ce cas, la taille des divisions n'a que peu d'influence sur la forme de la FIM. Les séries temporelles étudiées dans la partie III comptent 8192 points. Pour l'estimation des probabilités conjointes nécessaires au calcul de la FIM, les axes des histogrammes sont divisés en 20 segments ($\simeq N^{1/3}$).

Du point de vue de la reconstruction, le délai optimal est associé au premier minimum de la FIM, bien qu'aucune démonstration rigoureuse ne permette de valider cette pratique. Si aucun minimum n'apparaît, le choix se porte sur la valeur de τ obtenue lorsque le ratio $IM(\tau)/IM(1)$ passe en dessous d'un seuil prédéfini.

Dimension de reconstruction

La dimension de reconstruction optimale est la dimension permettant de révéler toute la complexité de l'attracteur. Une première approche consiste

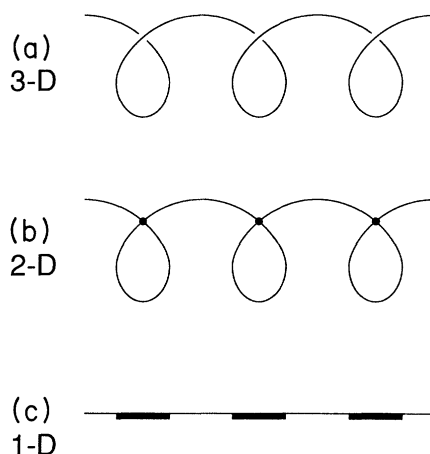


FIG. 3.4 Représentation dans un espace à 3 dimensions d'un ressort et ses projections en dimensions 2 et 1 (Brown et al. 1991).

à procéder à des reconstructions successives selon des dimensions croissantes et à calculer systématiquement un invariant du système dynamique (exposants de *Lyapunov*, *dimension de corrélation*). La stabilisation de la valeur de l'invariant indique la dimension minimale adéquate pour révéler la dynamique sous-jacente.

Cependant, la méthode des *faux voisins* (FV) (Brown et al. 1991) est probablement la plus fréquemment utilisée. Elle consiste également en des reconstructions successives selon des dimensions croissantes. Lorsque la dimension augmente de m à $m + 1$, si la topologie de l'attracteur n'est pas encore totalement révélée en dimension m , des points initialement voisins dans \mathcal{R}^m se trouvent éloignés dans \mathcal{R}^{m+1} . La figure 3.4 illustre cette idée. Le passage de 1D à 2D révèle beaucoup de faux voisins, alors que seuls les 3 nœuds sont identifiés lors du passage de 2D et 3D. Enfin, si la dimension augmente encore, le nombre de faux voisins sera nul dans la mesure où la géométrie du ressort est parfaitement révélée en dimension 3.

De façon plus formelle, la distance euclidienne entre chaque point X_k et son plus proche voisin dans \mathcal{R}^m est calculée. Elle est notée R_k^m , $k = 1, \dots, N'$. Lors du passage à la dimension supérieure, le comportement de cette paire de points est caractérisé par un coefficient d'étirement :

$$S_k = \frac{\sqrt{(R_k^{m+1})^2 - (R_k^m)^2}}{R_k^m}. \quad (3.9)$$

Les points de la paire sont alors classés comme faux voisins si S_k dépasse un seuil déterminé empiriquement, compris entre 10 et 20. Il suffit ensuite de tracer le taux de faux voisins en fonction de la dimension et de rechercher la dimension optimale indiquée par une forte diminution de cette courbe.

Dans le cas de données bruitées, la courbe peut présenter une allure exponentielle décroissante, ne présentant pas de diminution franche indiquant la dimension optimale. Le choix d'une dimension plus ou moins *arbitraire* est alors guidé par :

- la forme de la courbe,

- la nécessité d'une dimension suffisamment grande pour révéler au maximum la dynamique sous-jacente,
- les contraintes liées aux moyens informatiques qui définissent une dimension maximale exploitable en terme de temps de calcul raisonnable pour l'extraction des indicateurs dans l'espace des phases.

3.1.3 Exposants de Lyapunov

Les *exposants de Lyapunov*, λ_i , reflètent la façon dont une petite perturbation δ_0 du système est amplifiée dans l'espace des phases lorsque $t \rightarrow +\infty$. L'exposant maximal s'exprime par la relation

$$\lambda_{max} = \lim_{t \rightarrow +\infty} \frac{1}{t} \ln \left(\frac{\delta(t)}{\delta_0} \right). \quad (3.10)$$

Pour interpréter l'ensemble des exposants de *Lyapunov*, il faut imaginer une hypersphère de rayon r_0 qui se déforme au cours du temps pour devenir un hyperboloïde. A chaque axe principal du volume est associé un exposant indiquant son taux de dilatation ou de contraction. En notant $r_i(t)$ la longueur du $i^{\text{ème}}$ axe principal, il vient

$$\lambda_i = \lim_{t \rightarrow +\infty} \frac{1}{t} \ln \left(\frac{r_i(t)}{r_0} \right). \quad (3.11)$$

L'estimation du spectre de *Lyapunov* nécessite le calcul de l'évolution d'une petite perturbation appliquée au système. Or, l'évolution d'une perturbation infinitésimale est complètement décrite dans l'espace tangent à la trajectoire, c'est à dire par la dynamique linéarisée (Abarbanel et al. 1993). C'est à l'aide de cette linéarisation que Brown et al. (1991) extraient les exposants de *Lyapunov*.

Pour un système dynamique déterministe représenté par l'opérateur d'évolution F , le mouvement du point $X_k \in \mathbb{R}^m$ le long d'une trajectoire obéit à la relation

$$X_{k+1} = F(X_k). \quad (3.12)$$

Le suivi d'une perturbation est approché par l'évolution de l'écart entre deux trajectoires très proches. Soit $Y_k \in \mathbb{R}^m$, un point proche de X_k . L'écart initial entre ces deux points est noté $\delta X_0 = X_k - Y_k$ et son évolution après i pas de temps est conditionnée par le comportement des deux trajectoires : $\delta X_i = X_{k+i} - Y_{k+i}$. La linéarisation de F autour de X_k est réalisée par son développement de Taylor au premier ordre :

$$\begin{aligned} \delta X_1 &= X_{k+1} - Y_{k+1} \\ &= F(X_k) - F(Y_k) \\ &= J(X_k) [X_k - Y_k] + o \|X_k - Y_k\|^2 \\ &= J(X_k) \delta X_0 + o \|\delta X_0\|^2 \end{aligned} \quad (3.13)$$

où J est la matrice *jacobienne* de F . Lorsque X et Y sont suivis pendant L pas de temps, l'expression de l'écart en fonction de l'écart initial fait intervenir

le produit successif des matrices *jacobiennes* estimées à chaque itération (Eq.3.14).

$$\begin{aligned}
\delta X_2 &= J(X_{k+1})\delta X_1 \\
&= J(X_{k+1})J(X_k)\delta X_0 \\
\delta X_L &= J(X_k) \cdot J(X_{k+1}) \dots J(X_{k+L-1})\delta X_0 \\
&= \left[\prod_{j=0}^{L-1} J(X_{k+j}) \right] \delta X_0 \\
&= J^L(X_k)\delta X_0
\end{aligned} \tag{3.14}$$

Soient Λ_i , $\{i = 1, \dots, m\}$, les valeurs propres de la matrice $J^L(X_k) \cdot [J^L(X_k)]^T$. Le *théorème multiplicatif ergodique d'Oseledec* (Oseledec 1968) démontre que la matrice limite

$$\lim_{L \rightarrow +\infty} \left[J^L(X_k) \cdot [J^L(X_k)]^T \right]^{\frac{1}{2L}} \tag{3.15}$$

existe et que ses valeurs propres sont égales à e^{λ_i} . Les exposants de *Lyapunov* s'obtiennent alors par

$$\lambda_i = \lim_{L \rightarrow \infty} \frac{1}{2L} \ln(\Lambda_i). \tag{3.16}$$

Le théorème démontre que les valeurs ainsi obtenues des exposants ne dépendent pas du point initial X_k . Elle sont invariantes sur l'attracteur.

Dans le contexte de l'analyse des séries temporelles, l'expression de F n'est pas connue a priori. La difficulté porte alors sur l'estimation des matrices jacobiennes successives, uniquement à partir des trajectoires reconstruites dans l'espace des phases. Pour cela, F est estimée localement, sur plusieurs voisinages de points, par l'ajustement d'une fonction, par exemple polynomiale. La matrice jacobienne est ensuite calculée sur la base de cette fonction locale.

Un point X_k est choisi de façon aléatoire dans l'espace des phases. Sa distance avec les autres points de l'espace est estimée et le $i^{\text{ème}}$ point le plus proche est noté $V_{X_k}^i(0)$. Un voisinage $V_{X_k}(0)$ de r points à proximité de X_k est constitué. Pour ne considérer que les aspects géométriques de l'attracteur, il faut éviter d'intégrer dans le voisinage les points corrélés temporellement avec X_k . Ainsi, les points situés à moins de f pas de temps sont exclus (Theiler 1986).

$$\begin{aligned}
V_{X_k}(0) &= \left\{ X_i, i \in [1; N'] \mid \|X_k - X_i\| \leq \|X_k - V_{X_k}^i(0)\| \right. \\
&\quad \left. \text{et } |i - k| > f \right\}
\end{aligned} \tag{3.17}$$

$$= \left\{ V_{X_k}^i(0), i = 1, \dots, r \right\} \tag{3.18}$$

Dans la mesure où l'objectif consiste à suivre de petites perturbations, la fonction n'est pas ajustée sur l'évolution des points, mais sur l'évolution des écarts entre les points. L'ensemble des écarts entre X_k et ses plus proches voisins est noté :

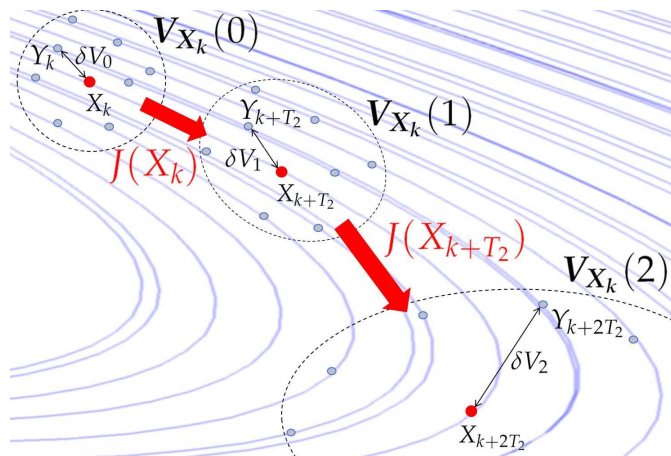


FIG. 3.5 Principe de l'estimation de la matrice jacobienne tous les T_2 pas de temps, au fil de l'évolution du voisinage centré autour de X_k .

$$\delta V_{X_k}(0) = \left\{ \delta V_{X_k}^i(0) = X_k - X_i \mid X_i \in V_{X_k}(0) \right\} \quad (3.19)$$

Le meilleur opérateur T , qui transforme les écarts du voisinage initial, $V_{X_k}(0)$, pour former les écarts du voisinage après T_2 pas de temps, $V_{X_k}(T_2)$ est recherché.

$$\delta V_{X_k}^i(0) = T \left(\delta V_{X_k}^i(T_2) \right), \quad i = 1, \dots, r \quad (3.20)$$

Initialement, T est une transformation linéaire, mais [Sano et Sawada \(1985\)](#) montrent qu'une transformation polynomiale permet une meilleure estimation des exposants négatifs. Ses paramètres sont déterminés par une méthode de moindres carrés. Le nombre de points composant le voisinage est directement lié au nombre de paramètres inconnus à ajuster. Il est recommandé de constituer des voisinages contenant deux fois plus de points que de paramètres à ajuster, ce qui représente 40 points pour un polynôme de degré 2 en dimension 5, et 110 points pour un polynôme de degré 3. La taille du voisinage doit cependant rester modeste, de façon à conserver le caractère local de la dynamique (< 50), limitant ainsi le degré de la transformation polynomiale à ajuster.

L'estimation de la matrice jacobienne est donc obtenue soit en recherchant directement une transformation linéaire, soit en conservant les coefficients du premier degré d'une transformation polynomiale. L'opération est répétée pour chaque itération, formant ainsi les matrices $J(X_{k+iT_2})$, $i = 1, \dots, L$ (Fig. 3.5).

Conformément aux équations 3.15 et 3.16, le calcul des exposants de *Lya-punov* nécessite l'estimation des valeurs propres du produit des jacobiniennes. Pour assurer la convergence de la limite de l'équation 3.16, le nombre d'itérations L doit être supérieur à 2000.

La matrice *produit* est alors très mal conditionnée. Pour calculer ses valeurs propres, une décomposition QR est nécessaire ([Eckmann et al. 1986](#)).

$$J(X_{k+iT_2})Q(i) = Q(i+1)R(i+1) \quad i = 0, \dots, L-1 \quad (3.21)$$

Puisque $Q(i)$ est une matrice orthogonale ($Q(0) = Id$), et $R(i)$ une matrice triangulaire supérieure, le produit des matrices jacobiennes s'écrit :

$$J^L(X_k) = \prod_{i=0}^{L-1} J(X_{k+iT_2}) = Q(L) \prod_{i=1}^L R(i). \quad (3.22)$$

Finalement, les exposants de *Lyapunov* sont exprimés par :

$$\lambda_i = \frac{1}{dtLT_2} \sum_{j=1}^L \ln(R_{jj}) \quad (3.23)$$

avec dt , la période d'échantillonnage de la mesure (en secondes).

Cet algorithme est répété pour plusieurs centaines de points pris aléatoirement sur l'attracteur. Les moyennes de chaque exposant constituent l'estimation du spectre de Lyapunov. Bien qu'il soit automatisable, beaucoup de paramètres doivent être préalablement définis par l'opérateur (T_2 , n , r , L), et aucune méthode systématique ne permet de les prédéfinir rigoureusement. De plus, son coût en terme de temps de calcul est potentiellement prohibitif (2000 itérations \times plusieurs centaines de points \times plusieurs dizaines de mesures).

Sur la base de ces constats, la section suivante présente un indicateur original inspiré par le concept des exposants de Lyapunov et proposé dans le cadre de ces travaux de recherche. Cet indicateur explore plutôt la dynamique locale de l'attracteur, permettant ainsi une réduction considérable des temps de calcul comparativement au temps nécessaire à l'extraction du spectre de Lyapunov.

3.1.4 Indicateur Jacobien

L'algorithme présenté ci-dessus présente plusieurs inconvénients en vue d'une utilisation pour la détection d'endommagement en contexte de SHM.

Comme évoqué, son coût en terme de temps de calcul est très important. Les résultats présentés dans la section 5.1 de ce mémoire attestent de cette réalité. Dans une optique de suivi en continu d'ouvrages, la durée d'analyse des données doit être réduite au maximum.

De plus, l'estimation des valeurs propres du produit des matrices jacobiennes constitue un problème délicat. Son mauvais conditionnement implique une incertitude assez forte sur l'estimation des exposants de *Lyapunov*.

Enfin, les exposants de *Lyapunov* sont supposés capter un changement de la dynamique globale. En effet, le produit des jacobiennes calculées sur des milliers d'itérations peut masquer d'éventuelles distorsions locales de la trajectoire pouvant, par exemple, résulter de non-linéarités de type *impacts*.

Afin de révéler ces irrégularités locales et de réduire les temps de calcul, il semble plus adapté de définir un indicateur directement basé sur le suivi de la matrice jacobienne sur des durées courtes.

Ainsi, un nouvel indicateur est proposé et noté *JFV*, pour *Jacobian Feature Vector*. Cet indicateur vectoriel est formé de toutes les composantes de

la matrice jacobienne obtenue après un faible nombre d'itérations ($L \leq 5$) :

$$JFV = \begin{pmatrix} J(X_{k+LT_2})_{1,1} \\ \vdots \\ J(X_{k+LT_2})_{1,m} \\ J(X_{k+LT_2})_{2,1} \\ \vdots \\ J(X_{k+LT_2})_{2,m} \\ \vdots \\ J(X_{k+LT_2})_{m,1} \\ \vdots \\ J(X_{k+LT_2})_{m,m} \end{pmatrix} \quad \text{avec } L \leq 5. \quad (3.24)$$

L'étude des paramètres influant sur la formation du JFV est proposée dans la section 5.1. Compte tenu de la dimension de cet indicateur vectoriel, m^2 , la dimension de reconstruction doit rester faible. Dans le cas contraire, l'étape de normalisation nécessiterait une base de référence contenant beaucoup d'individus, problème bien connu en apprentissage automatique sous le nom de *fléau de la dimension* ².

3.2 NORMALISATION PAR LA DISTANCE DE MAHALANOBIS

Chaque mesure $x^k = \{x_1^k, \dots, x_N^k\}$ est caractérisée par un indicateur vectoriel V_k de dimension p . L'étape de normalisation a pour objectif de comparer une mesure individuelle, représentée par V^* , dont la classe (structure endommagée / structure saine) est inconnue, à une population de mesures de référence de N_r individus, $\mathbf{V}^r = \{V_1, \dots, V_{N_r}\}$, dont la classe est *structure saine*.

La distance de *Mahalanobis* est définie comme la distance entre un point et le centre d'un nuage de points en tenant compte de la corrélation linéaire entre les variables de l'espace. La distance entre V^* et \mathbf{V}^r s'écrit :

$$D^* = \sqrt{(V^* - \bar{V}^r)^T \Sigma_{\mathbf{V}^r}^{-1} (V^* - \bar{V}^r)} \quad (3.25)$$

avec \bar{V}^r la moyenne calculée sur la population de référence

$$\bar{V}^r = \frac{1}{N_r} \sum_{i=1}^{N_r} V_i \quad (3.26)$$

et $\Sigma_{\mathbf{V}^r}$ la matrice de covariance de \mathbf{V}^r

$$\Sigma_{\mathbf{V}^r} = \frac{1}{N_r - 1} \sum_{i=1}^{N_r} (V_i - \bar{V}^r) (V_i - \bar{V}^r)^T. \quad (3.27)$$

L'inversion de $\Sigma_{\mathbf{V}^r}$ impose que celle-ci ne soit pas trop mal conditionnée. Pour cela, il est nécessaire que le nombre d'individus dans la population

²Curse of dimensionality

de référence soit au moins 2 fois supérieur au nombre de composantes de l'indicateur vectoriel ($N_r \geq 2p$).

Si la distribution de la population de référence est supposée gaussienne multidimensionnelle, alors la distance de Mahalanobis de V^* est inversement proportionnelle à la densité de V^r autour de V^* .

3.3 MODÉLISATION STATISTIQUE

Une fois les indicateurs normalisés, leur distribution est modélisée dans le but de prendre une décision quant à leur appartenance ou non à la population de référence (structure saine). Que le problème soit posé en terme de test d'hypothèse ou de contrôle de processus, il requiert dans les deux cas la détermination de limites permettant respectivement de rejeter l'hypothèse nulle ou de déclarer le processus comme *hors de contrôle*.

Dans le chapitre 8, quelques-unes des approches possibles sont comparées. La méthode de normalisation employée étant principalement basée sur la notion de distance à la population de référence, seule une augmentation de la distance peut révéler un endommagement. Dans ce cas, aucune borne inférieure n'est fixée et les tests sont unilatéraux.

$$H_0 : \theta = \theta_0 \quad ; \quad H_1 : \theta > \theta_0 \quad (3.28)$$

Il existe une différence entre les méthodes appliquées à un seul individu à la fois, et celles travaillant sur un échantillon de plusieurs individus. Dans tous les cas, les paramètres de la distribution de la population de référence sont inconnus et doivent être estimés.

Dans la suite du document, les indicateurs sont notés I après normalisation. L'échantillon à tester est noté $\mathbf{I}^* = \{I_1^*, \dots, I_{N_e}^*\}$, et la population de référence $\mathbf{I}^r = \{I_1, \dots, I_{N_r}\}$. La moyenne et la variance de \mathbf{I}^r sont estimées par :

$$\bar{I} = \frac{1}{N_r} \sum_{i=1}^{N_r} I_i \quad (3.29)$$

$$R^2 = \frac{1}{N_r - 1} \sum_{i=1}^{N_r} (I_i - \bar{I})^2 \quad (3.30)$$

Le pourcentage d'erreurs de type I, c'est à dire de fausses alarmes est fixé à 1% ($\alpha = 0.01$). Il est donc attendu que sur 100 individus appartenant à la population de référence, un seul dépasse la limite de classification.

3.3.1 Test sur les valeurs individuelles : $N_e = 1$

Après chaque mesure, la valeur de l'indicateur obtenue est comparée au seuil de classification préétabli. Cette démarche correspond plus au formalisme du contrôle de processus, qu'à celui du test d'hypothèse. La section 1.3.3 montre que la fixation du seuil peut être conduite selon un modèle gaussien ou une *loi de Pareto Généralisée*.

Méthode 1 : Modèle Gaussien

La distribution de \mathbf{I}^r est supposée gaussienne, de moyenne \bar{I} et de variance R^2 estimées sur un grand échantillon ($N_r > 70$). La limite supérieure de classification est :

$$UCL = \bar{I} + Z_{1-\alpha}R \quad \text{avec } Z_{0,99} = 2.3263 \quad (3.31)$$

Méthode 2 : Extrapolation par la loi des valeurs extrêmes

Pour modéliser avec plus de précision la queue de la distribution, ou dans le cas où celle-ci n'est pas gaussienne, l'utilisation des excès est conseillée (§1.3.3).

En notant les excès $\mathbf{U}_r = \{I_i - u \mid I_i > u\} = \{U_1, \dots, U_{N_u}\}$, les valeurs U_i suivent une loi de *Pareto généralisée* (GPD) lorsque u se trouve dans la queue de la distribution. La loi comporte 3 paramètres, σ , γ et u (Eq.1.42).

Pour estimer σ et γ , la méthode des moments donne des résultats satisfaisants (Hosking et Wallis 1987). Cependant, pour diminuer la variance de l'estimation de la GPD, surtout en présence d'un faible nombre de valeurs ($15 \leq N_u \leq 30$), γ est considéré comme nul, de façon à se placer dans le cas d'une distribution exponentielle. Dans ce cas, l'estimateur du maximum de vraisemblance de σ donne :

$$\hat{\sigma} = \frac{1}{N_u} \sum_{i=1}^{N_u} U_i. \quad (3.32)$$

Au lieu de fixer un seuil u , le nombre de points k utilisés pour l'ajustement du modèle est choisi par l'utilisateur. Les valeurs de \mathbf{I}^r étant rangées dans l'ordre croissants $\{I_{N_r:1}, \dots, I_{N_r:N_r}\}$, u est la $(N_r - k)^{\text{ème}}$ valeur, $u = I_{N_r-k:N_r}$ (Piera-Martinez 2008). Enfin, $F(u)$ est estimé par sa probabilité empirique :

$$F(u) = \frac{N_r - k}{N_r + 1}. \quad (3.33)$$

En utilisant les équations 1.42 et 1.44, la limite supérieure de classification s'écrit :

$$UCL = u - \hat{\sigma} \ln \left(\frac{\alpha}{1 - F(u)} \right). \quad (3.34)$$

Méthode 3 : Test d'Hotelling T^2

Le test d'Hotelling T^2 est l'équivalent multivarié du test de *Student* univarié pour la moyenne. En effet, si x est une variable aléatoire gaussienne de moyenne μ , et que sa variance est estimée à partir d'un échantillon, la variable

$$\left[(x - \mu) \frac{n}{S^2} (x - \mu) \right]^{\frac{1}{2}} \sim t \quad (3.35)$$

suit une loi de *Student*. L'équivalent multivarié de cette expression adapté à ces travaux de recherche est

$$N_r (I_i - \bar{I})^T W^{-1} (I_i - \bar{I}) = T^2. \quad (3.36)$$

$$\text{avec } W = \frac{1}{N_r - 1} \sum_{i=1}^{N_r} (I_i - \bar{I})^T (I_i - \bar{I}). \quad (3.37)$$

Sur chaque vecteur à tester, T^{2*} est estimé (Eq.3.37) et comparé à la limite de contrôle (Eq.3.39).

$$(I_i^* - \bar{I})^T W^{-1} (I_i^* - \bar{I}) = T^{2*}. \quad (3.38)$$

$$UCL = \frac{p(N_r - 1)}{(N_r - p)} F_{(1-\alpha, p, N_r - p)}. \quad (3.39)$$

Ce test ne pourra être appliqué qu'aux valeurs individuelles extraites des données analysées dans le chapitre 8. En effet, dans le cas d'échantillons ($N_e > 1$), il est nécessaire d'estimer W sur chacun d'eux. Compte tenu de la disponibilité des données, N_e ne dépasse pas 5, alors que la dimension de I est supérieure à 30. Il est donc impossible de faire une quelconque estimation de la matrice de covariance.

Séquence de valeurs

Pour réduire le nombre de fausses alarmes, une variante consiste à considérer des mesures comme nouvelles seulement lorsque que s valeurs individuelles (I^*) dépassent consécutivement le seuil de classification. La probabilité de fausse alarme restant inchangée, la valeur du seuil doit être abaissée pour tenir compte de cette méthode de décision. Ainsi, en supposant que les I^* sont indépendants, les seuils sont toujours déterminés par les équations 3.31 et 3.34, en remplaçant α par $\alpha^{1/s}$.

3.3.2 Test sur un échantillon : $N_e > 1$

Le travail sur un échantillon permet de réduire la variance des résultats, mais nécessite davantage de données. De la même façon que précédemment, un seuil est déterminé selon une distribution gaussienne, et un autre selon une *distribution de valeurs extrêmes*.

Méthodes 4, 5 et 6 : Modèle Gaussien

Si l'estimation de la moyenne ne pose pas de difficultés, la variance peut être obtenue de trois façons différentes.

→ Méthode 4 : A partir de la population mère

La distribution de I^r est supposée gaussienne, de moyenne \bar{I} et de variance R^2 estimée sur tous les points de la base de référence. La limite supérieure de classification pour un échantillon de taille N_e est par conséquent :

$$UCL = \bar{I} + Z_{(1-\alpha)} \frac{R}{\sqrt{N_e}} \quad \text{avec } Z_{0.99} = 2.3263 \quad (3.40)$$

→ Méthode 5 : Carte de contrôle X

Sur chaque échantillon, la moyenne et la variance sont calculées. Elles sont respectivement notées I_i^e et R_i^e ($i = 1, \dots, N_r^e$) lorsque l'échantillon fait partie de la population de référence, et I^{e*} et R^{e*} s'il s'agit de l'échantillon à tester.

La limite de contrôle est calculée à partir de la moyenne des moyennes et de la moyenne des variances :

$$UCL = \bar{I}^e + Z_{(1-\alpha)} \frac{\bar{R}^e}{\sqrt{N_e}} \quad \text{avec } Z_{0.99} = 2.3263 \quad (3.41)$$

$$\text{avec } \bar{I}^e = \sum_{i=1}^{N_r^e} I_i^e \quad \text{et} \quad \bar{R}^e = \sum_{i=1}^{N_r^e} R_i^e \quad (3.42)$$

→ Méthode 6 : A partir des moyennes des échantillons

La population formée par les moyennes I_i^e de chaque échantillon doit suivre également une loi gaussienne. La limite correspondant à l'équation 3.31 reste valable.

$$UCL = \bar{I}^e + Z_{1-\alpha} R^e \quad (3.43)$$

avec

$$R^{e2} = \frac{1}{N_r^e - 1} \sum_{i=1}^{N_r^e} (I_i^e - \bar{I}^e)^2 \quad (3.44)$$

En présence d'une population mère réellement gaussienne, l'estimation de la limite de contrôle par la méthode 4 présente la plus faible variance. A partir d'une taille d'échantillon de 5 individus ($N_e = 5$), les méthodes 5 et 6 sont équivalentes. En dessous, la variance associée à la méthode 6 est jusqu'à deux fois supérieure à celle relative à la méthode 5 ($N_e = 2$).

Méthode 7 : Modèle de valeurs extrêmes

Sur chaque échantillon est prélevée la valeur maximale. Quand N_e augmente, la distribution des maxima tend vers une distribution des *valeurs extrêmes* dont la forme généralisée est donnée dans l'équation 1.38.

Concernant les applications proposées dans le chapitre 8, la méthode du *maximum de vraisemblance* est retenue pour l'ajustement des paramètres de l'EVD. Certes, elle présente un biais non nul, mais compte tenu du faible nombre de valeurs disponibles, sa variance réduite est privilégiée. Les calculs sont réalisés à l'aide de la fonction `gevfit` de Matlab®.

L'expression du seuil de classification sur les maxima de chaque échantillon s'obtient par inversion de $F_{max}(x)$ (Eq.1.38) :

$$UCL = \mu + \frac{\sigma}{\gamma} \left([-\ln(1 - \alpha)]^{-\gamma} - 1 \right). \quad (3.45)$$

CAS D'ÉTUDES EXPÉRIMENTAUX

4

SOMMAIRE	
INTRODUCTION	78
4.1 SYSTÈME MASSE/RESSORT AMORTI	79
4.2 MAQUETTE MÉTALLIQUE	80
4.3 POUTRE EN BÉTON FIBRÉ	81
4.4 MAQUETTE EN BOIS	84

INTRODUCTION

LA validation de la démarche de détection de nouveauté au moyen des indicateurs proposés dans le chapitre précédent nécessite de disposer de données relatives à un système dynamique avant et après introduction d'endommagement. Rares sont les cas d'études réels remplissant ces conditions.

A titre d'exemple, le pont *I-40* (Farrar et al. 1994) a été endommagé par la création d'une entaille dans une poutre métallique. Le suivi sur 24 heures de ses caractéristiques modales a été réalisé, cependant, seuls les spectres de puissance ont été conservés.

L'étude du pont *Z-24* en Suisse représente un travail unique par l'ampleur des mesures collectées (Peeters et De Roeck 2000). Pendant une année, la réponse de ce pont sous sollicitations ambiantes a été enregistrée toutes les heures. Puis, différents scénarios d'endommagement ont été imposés (tassement d'appui, rupture de câble de précontrainte...). Dans ce cas, la détection de nouveauté est cependant difficilement applicable dans la mesure où l'instrumentation a été complètement modifiée entre la phase de suivi initiale (6 accéléromètres) et la phase d'endommagement progressif (plusieurs centaines d'accéléromètres). Le fait de changer les capteurs, et également leurs positions, engendre de la nouveauté dans les mesures. Les données acquises pendant l'année de suivi ne peuvent donc pas être utilisées comme référence.

A défaut de mesures collectées sur des structures réelles dans leur environnement, ce chapitre présente les caractéristiques de quatre cas d'études de laboratoire sur lesquels sont testés les indicateurs et les méthodes de détection de nouveauté définis dans le chapitre précédent. Ces cas d'études se divisent comme suit :

- analyse de données vibratoires simulées numériquement par le biais d'un système masse/ressort amorti,
- analyse de données vibratoires expérimentales provenant d'une maquette métallique (essais réalisés par le *Los Alamos National Laboratory* (LANL 2000)),
- analyse de données vibratoires expérimentales provenant d'une poutre en béton fibré endommagée progressivement et soumise à des sollicitations thermiques,
- analyse de données vibratoires expérimentales provenant d'une maquette en bois avec sollicitations thermiques et hydriques et scénarios d'endommagement multiples.

Les développements applicatifs sont ensuite présentés dans la partie III.

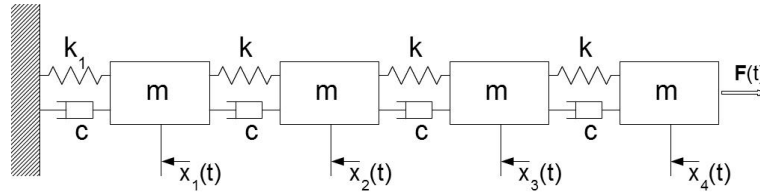


FIG. 4.1 *Système masse/ressort/amortisseur utilisé pour la génération de données vibratoires simulées.*

4.1 SYSTÈME MASSE/RESSORT AMORTI

Cette première expérimentation numérique vise à étudier la sensibilité des indicateurs chaotiques : *exposants de Lyapunov* et *JFV*. Ces données simulées sont bien contrôlées, non bruitées et sans variations environnementales.

La réponse d'un système masse/ressort amorti à quatre degrés de liberté est calculée numériquement (Fig.4.1). Le comportement dynamique du système est régi par l'équation :

$$M\ddot{\mathbf{x}} + C\dot{\mathbf{x}} + K\mathbf{x} = \mathbf{F}(t) \quad (4.1)$$

dans laquelle, M , C , et K représentent respectivement les matrices de masse, d'amortissement et de raideur. Les paramètres du système sont choisis de façon à reproduire les ordres de grandeurs relatifs aux caractéristiques d'un bâtiment réel : $k = 2,4 \cdot 10^8 \text{ N.m}^{-1}$, $c = 1,02 \cdot 10^6 \text{ N.m}^{-1} \cdot \text{s}$ et $m = 70000 \text{ kg}$. Les fréquences propres sont de 3,23 Hz, 9,31 Hz, 14,3 Hz et 17,5 Hz.

L'excitation $\mathbf{F}(t)$ est un bruit blanc appliqué à la l'extrémité libre du système. Le calcul de la réponse est réalisé à l'aide d'un schéma d'intégration de *Runge-Kutta* d'ordre 4 avec un pas de temps constant de 0,017s.

L'endommagement est simulé par l'introduction d'une raideur bilinéaire au niveau du ressort k_1 . La raideur en traction est diminuée par le biais d'un facteur α ajustable. La raideur en compression est invariante. La différence des raideurs en traction et en compression induit un comportement non linéaire, simulant l'effet d'une fissure dans la dynamique vibratoire d'une poutre (Eq.4.2).

$$k_1 = \begin{cases} k & \text{if } x_1 \geq 0 \\ (1 - \alpha)k & \text{if } x_1 < 0 \end{cases} \quad (4.2)$$

L'accélération associée à la variable $x_1(t)$ est utilisée dans la démarche de détection de nouveauté. Chaque mesure est ré-échantillonnée à 30 Hz et compte 8192 points. Au total, 120 simulations (mesures) sont réalisées dans le cas d'un comportement sain du système ($\alpha = 0$) et constituent la base de référence. Ensuite, 9 niveaux d'endommagement progressifs sont simulés par le biais du paramètre α et 10 simulations sont réalisées à chaque niveau (Tab.4.1).

Nombre de mesures	α
120	0
10	0,1
10	0,2
10	0,3
10	0,4
10	0,5
10	0,6
10	0,7
10	0,8
10	0,9

TAB. 4.1 Composition de la base de données.



FIG. 4.2 Maquette métallique réalisée et testée au LANL ; Endommagement introduit sur l'assemblage 1C.

4.2 MAQUETTE MÉTALLIQUE

Cette structure de laboratoire, réalisée et testée au sein du *Los Alamos National Laboratory* (LANL), permet de vérifier la sensibilité des indicateurs chaotiques sur des mesures réelles. La description précise du dispositif expérimental et les données collectées sont disponibles sur le site Internet du LANL (LANL 2000). Bien que l'intensité de l'excitation soit variable, les conditions environnementales demeurent constantes.

La maquette est composée de quatre profilés métalliques (longueur : 1,5 m) et de trois étages constitués par des plateaux métalliques rectangulaires (Fig.4.2 et 4.3). Les liaisons montants/plateaux sont assurées par des assemblages boulonnés. Le socle, sur lequel sont fixés les quatre montants, est isolé du sol par des isolateurs à air autorisant les déplacements horizontaux.

Un pot vibrant exerce une excitation sous la forme d'un bruit blanc de

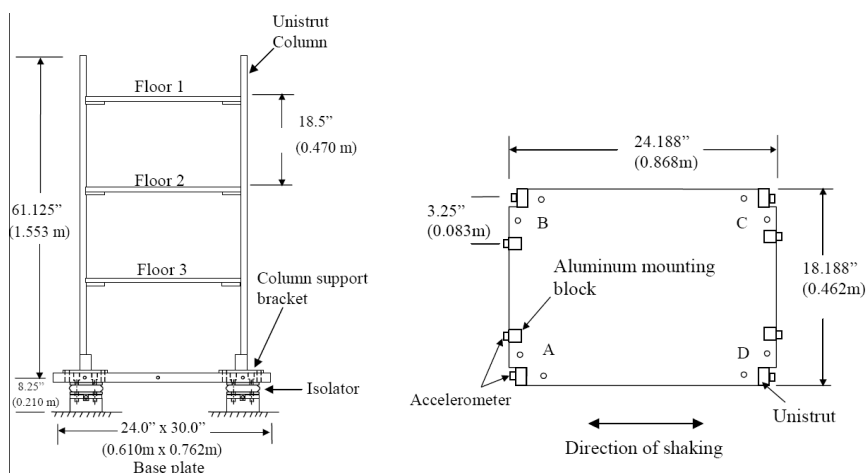


FIG. 4.3 Schéma et dimensions de la maquette métallique.

Cas	Description	Couple de serrage [N.m]	Nombre de mesures
0	Assemblage sain	6,77	150
1	Couple réduit	1,128	15
2	Couple réduit	0,564	15
3	Serré à la main	—	10
4	Boulon enlevé	—	10
5	Équerre enlevée	—	10

TAB. 4.2 Description des niveaux d'endommagement et nombre de mesures associées.

bande de fréquence comprise entre 800 et 3200 Hz, et d'intensités variables (2, 5 ou 8 V_{rms}). Ces plages de variations sont introduites volontairement par les concepteurs du dispositif expérimental afin de simuler une excitation dont les caractéristiques sont variables. La sollicitation est appliquée sur le socle au niveau de l'angle D de façon à solliciter les modes de flexion et de torsion.

Les vibrations sont mesurées dans l'axe de l'excitation par 24 accéléromètres de sensibilité $1V.g^{-1}$, placés de part et d'autre de chaque assemblage. Une mesure compte 8192 points, échantillonnés à 1600Hz.

L'endommagement est introduit en réduisant progressivement le couple de serrage des vis des assemblages. Parmi les scénarios réalisés, seul celui concernant l'assemblage 1C est étudié ici et les mesures de l'accéléromètre correspondant, placé sur le plateau, sont analysées. Le tableau 4.2 présente les différents niveaux de serrage et le nombre de mesures pour chacun d'eux.

4.3 POUTRE EN BÉTON FIBRÉ

Ce dispositif expérimental s'appuie sur une poutre en béton fibré soumise à des sollicitations thermiques variables. Il vise à mesurer le comportement dynamique de la poutre en présence :

- d'un endommagement progressif,
- de variations environnementales affectant de façon significative les mesures vibratoires.

Géométrie

La poutre présente une longueur de 2 mètres pour une section de $12 \times 8,8 \text{ cm}^2$. L'élanement élevé permet l'utilisation d'un pot vibrant de faible capacité (18 N maximum) pour la sollicitation dynamique.

Afin d'amplifier les effets des variations environnementales sur la dynamique vibratoire, une plaque métallique fixe est accolée à chaque extrémité de la poutre de façon à approcher des conditions d'encastrement et ainsi générer une légère précontrainte lors des phases de montée en température (Fig.4.5(c)). Le schéma et les dimensions du dispositif expérimental sont reportés dans la figure 4.4.

Instrumentation

Un pot vibrant (*TIRA*[®] TV50018, force 18 N, course 5 mm) placé à mi-portée applique un effort vertical au niveau de la fibre inférieure de la poutre. Quatre accéléromètres à électronique intégrée (*Bruel & Kjaer*[®] 4507B), de sensibilité $10 \text{ mV} \cdot \text{g}^{-1}$, sont positionnés sur la face supérieure de la poutre pour mesurer les accélérations verticales. Ils sont numérotés de 1 à 4 de gauche à droite.

Le système d'acquisition *PULSE 3560C* 5 voies de *Bruel & Kjaer*[®] est utilisé. Il s'accompagne du programme de pilotage *LabShop* interfacé avec *Matlab*[®] et contrôlé lors des essais par l'intermédiaire de scripts développés spécifiquement dans le cadre de la thèse. L'intérêt est ici d'automatiser entièrement les séquences de mesures, du pilotage du pot vibrant jusqu'à l'enregistrement des données.

Deux thermocouples sont noyés au niveau des premier et second tiers de la longueur et enregistrent les variations de température dans le volume.

Excitation

L'excitation est un bruit blanc de bande de fréquence de 0 – 800 Hz et d'intensité $0,7 V_{rms}$ généré par *LabShop*.

Sollicitation thermique

Afin d'appliquer une sollicitation thermique significative, six lampes de 175 W à rayonnement infra-rouge éclairent une face de la poutre (Eq.4.4). Le chargement thermique résultant est inhomogène, reflétant ainsi la réalité des structures soumises au rayonnement solaire.

Dans une approche de détection de nouveauté, il est essentiel que la base de référence englobe l'ensemble des conditions *normales* relatives à la vie de l'ouvrage. Dans un cas de structures réelles, la base de référence doit donc être constituée par un suivi continu sur une année au minimum. Afin de simuler cette période, un cycle de chauffage/refroidissement (20°C - 60°C - 20°C) est appliqué à la poutre dans son état *non endommagé* pour générer la base de données de référence. Durant la phase d'endommagement, la sollicitation thermique est exercée de façon aléatoire. La plage de tempé-

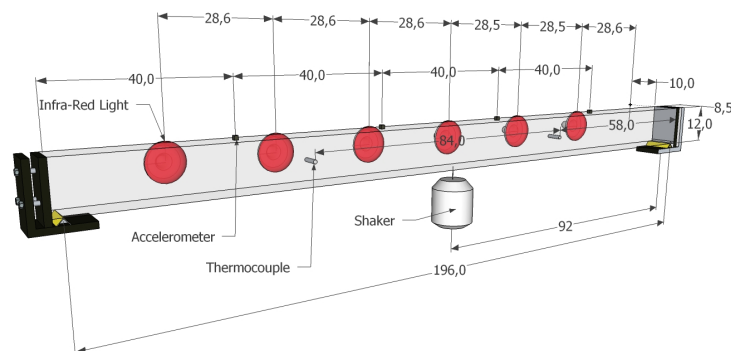


FIG. 4.4 Montage expérimental



FIG. 4.5 Photographies du montage expérimental

ratures inférieures à 20°C n'a pu être explorée pour des raisons pratiques liées à l'encombrement du dispositif. Les températures relevées par le biais des thermocouples noyés sont consignées pendant toute la durée de l'essai et présentées dans la figure 4.6. Les deux courbes correspondent aux mesures des deux thermocouples. Le premier cycle chauffage/refroidissement associé à la construction de la base de référence est clairement identifiable, suivi des sollicitations aléatoires exercées durant la phase d'endommagement.

Endommagement

Pour engendrer un endommagement progressif, des cycles de chargement statique en flexion quatre points sont imposés à la poutre (Fig.4.5(d)). Plus de 160 cycles de charge croissante (0.8 à 2.4 kN) sont réalisés jusqu'à rupture. Le tableau de l'annexe A.1 résume le protocole d'endommagement : séquence des cycles et charges statiques associées, séquences de mesures correspondantes.

Séquences et paramètres de mesures

Durant l'acquisition de la base de référence, une séquence de mesure vibratoire est effectuée toutes les 10 minutes en phase de chauffage, et toutes 25 minutes en phase de refroidissement. Lors de l'endommagement progres-

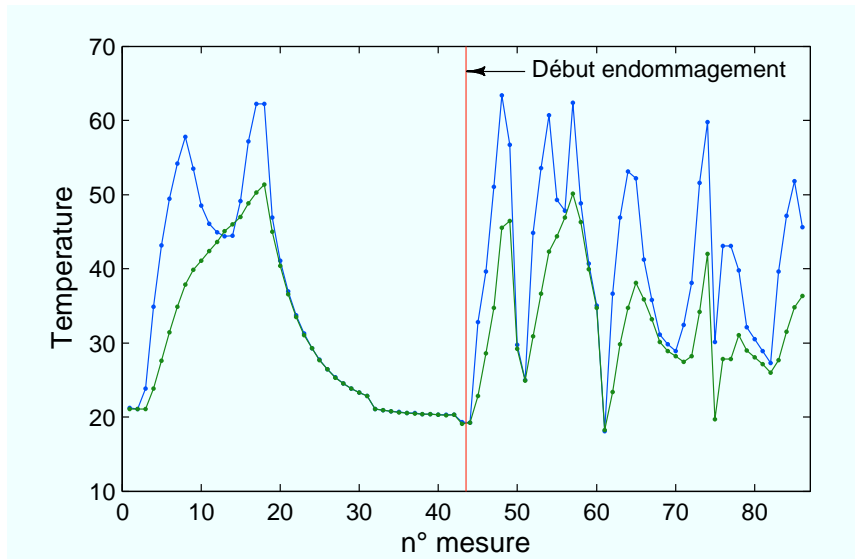


FIG. 4.6 Variations de la température en fonction de la séquence de mesure

sif, les séquences sont séparées par plusieurs cycles de chargement (entre 1 et 10).

Chaque séquence comporte 5 mesures de 4 secondes échantillonnées à 8192Hz. Pour l'analyse, elles sont rééchantillonnées à 2048Hz de façon à travailler sur des séries temporelles de 8192 points.

La base de référence est composée d'un total de 41 séquences (210 mesures) et 41 séquences supplémentaires sont réalisées durant la phase d'endommagement.

4.4 MAQUETTE EN BOIS

Le développement de ce dispositif expérimental s'inscrit dans l'objectif de réaliser des scénarios d'endommagement multiples et réversibles. Le dispositif s'appuie sur une maquette de structure en bois à 3 étages dont la géométrie est proche de celle de la maquette du LANL décrite plus-haut. Le choix du bois comme matériau constitutif est justifié par sa forte sensibilité à l'humidité relative de l'air ambiant. Cette particularité permet ainsi d'élargir le spectre des influences environnementales en exerçant des sollicitations à la fois thermiques et hydriques.

Un autre défi du suivi des ouvrages de génie civil réside dans la nature fortement instationnaire de l'excitation. Dans ce contexte, plusieurs types de sollicitations sont mis en œuvre.

La figure 4.7 présente une photographie et un schéma du dispositif.

Géométrie

Comme le montre le dessin de la figure 4.8, la maquette comporte trois niveaux carrés de $30 \times 30\text{cm}^2$ supportés par quatre montants de 90cm . La rigidité dans une direction est obtenue par le biais d'équerres métalliques qui assurent les assemblages montants/plateaux (4 boulons par équerre)

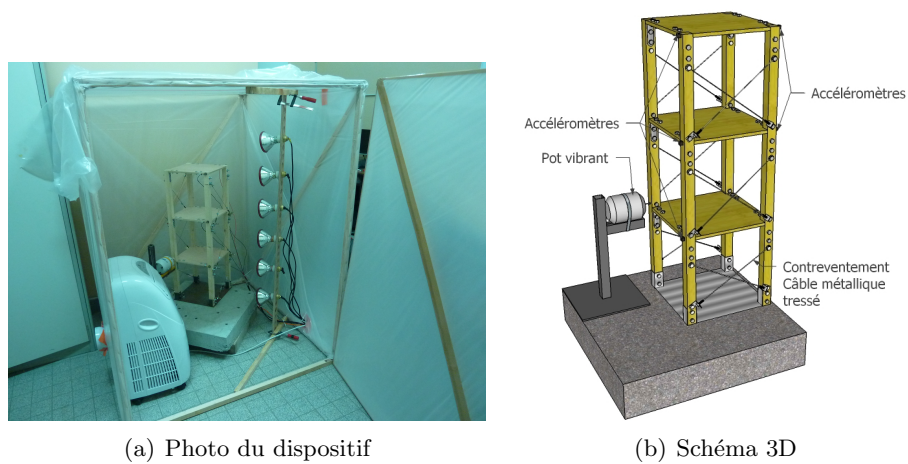


FIG. 4.7 Photographie et schéma 3D du dispositif Maquette en bois

(Fig.4.10). Dans la direction orthogonale, des câbles métalliques tendus constituent le contreventement (Fig.4.7(b)). La tension des câbles peut être modulée par un écrou de serrage (Fig.4.9).

Chaque pied de montant en bois est assemblé à une plaque métallique par 2 boulons, elle-même fixée par 4 autres boulons sur une dalle en béton, qui supporte également la potence du pot vibrant. Afin d'atténuer la transmission de vibrations externes, la dalle en béton est posée sur des plots en mousse.

Instrumentation

Ces expérimentations reprennent le matériel de mesure décrit dans la partie 4.3 (pot vibrant, 5 accéléromètres, système d'acquisition et de contrôle) et utilisé lors de l'essai sur la poutre.

Cinq accéléromètres de 10 mV.g^{-1} mesurent les accélérations dans la direction x (Fig.4.11), correspondant à la direction de contreventement par les câbles. Il sont placés sur les montants, au niveau des assemblages 1,2,3,4

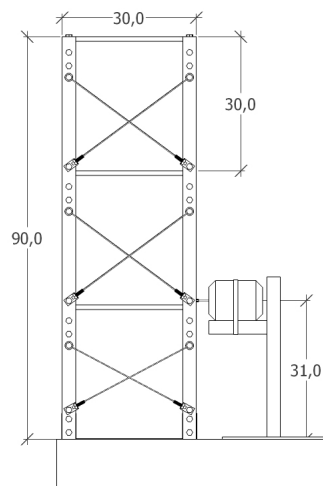


FIG. 4.8 Dimensions principales [cm]

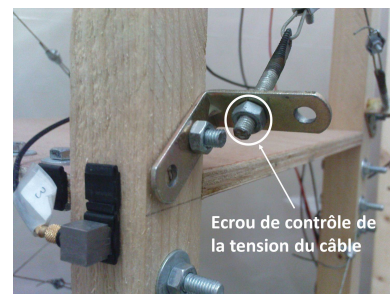


FIG. 4.9 Équerre d'attache des câbles, l'écrou de droite contrôle la tension.

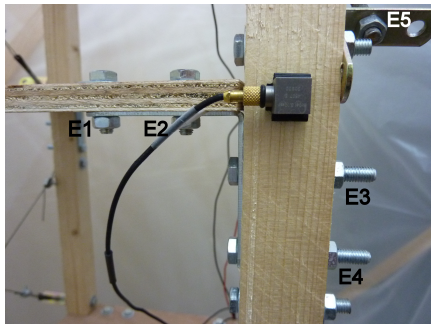


FIG. 4.10 Numérotation des écrous d'un assemblage.

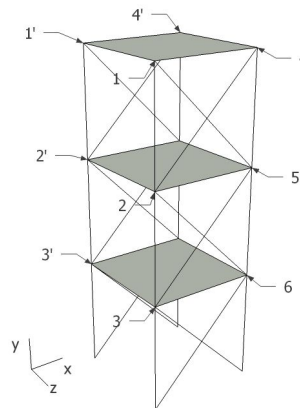


FIG. 4.11 Numérotation des assemblages et position du repère.

et 5. Dans la suite du document, le numéro de l'accéléromètre fait référence au nœud sur lequel il est placé.

La température et l'humidité relative de l'air sont enregistrées toutes les minutes grâce à un boîtier KIMO[®] KH200.

Excitation

L'excitation est appliquée dans la direction x au niveau du nœud 3' (Fig.4.11). La liaison entre le pot vibrant et la structure est réalisée sur le montant en bois.

Le pot vibrant de capacité maximale de 18N est contrôlé par le système PULSE de *Bruel & Kjaer*[®] introduit dans la section 4.3. Pour chaque séquence de mesure, trois excitations différentes sont exercées consécutivement (Fig.4.12) :

Exc1 : bruit blanc de bande de fréquence 800Hz,

Exc2 : bruit blanc (BF 800Hz) d'intensité faible + impulsions ; chaque signal comporte en moyenne 5 impulsions par seconde réparties aléatoirement sur toute sa durée (10s) ; l'amplitude des impulsions est également aléatoire.

Exc3 : impulsions uniquement (avec les mêmes caractéristiques que Exc2).

Ces 3 sources d'excitation sont générées par des scripts *Matlab*[®] dédiés et transmises automatiquement au système de contrôle du pot vibrant.

Ces différents types d'excitations visent à s'approcher des sollicitations rencontrées sur des ouvrages en service. En effet, la circulation et les rafales de vent produisent une excitation dynamique constituée des successions irrégulières d'efforts brefs et de forte intensité. Des exemples de signaux collectés sous sollicitations ambiantes sur le pont Z – 24 visibles sur la figure 4.13 illustrent ce caractère fortement instationnaire.

Enfin, toujours dans un souci de s'approcher de la réalité phénoménologique des ouvrages *in situ*, un nouveau signal d'excitation est généré à chaque mesure.

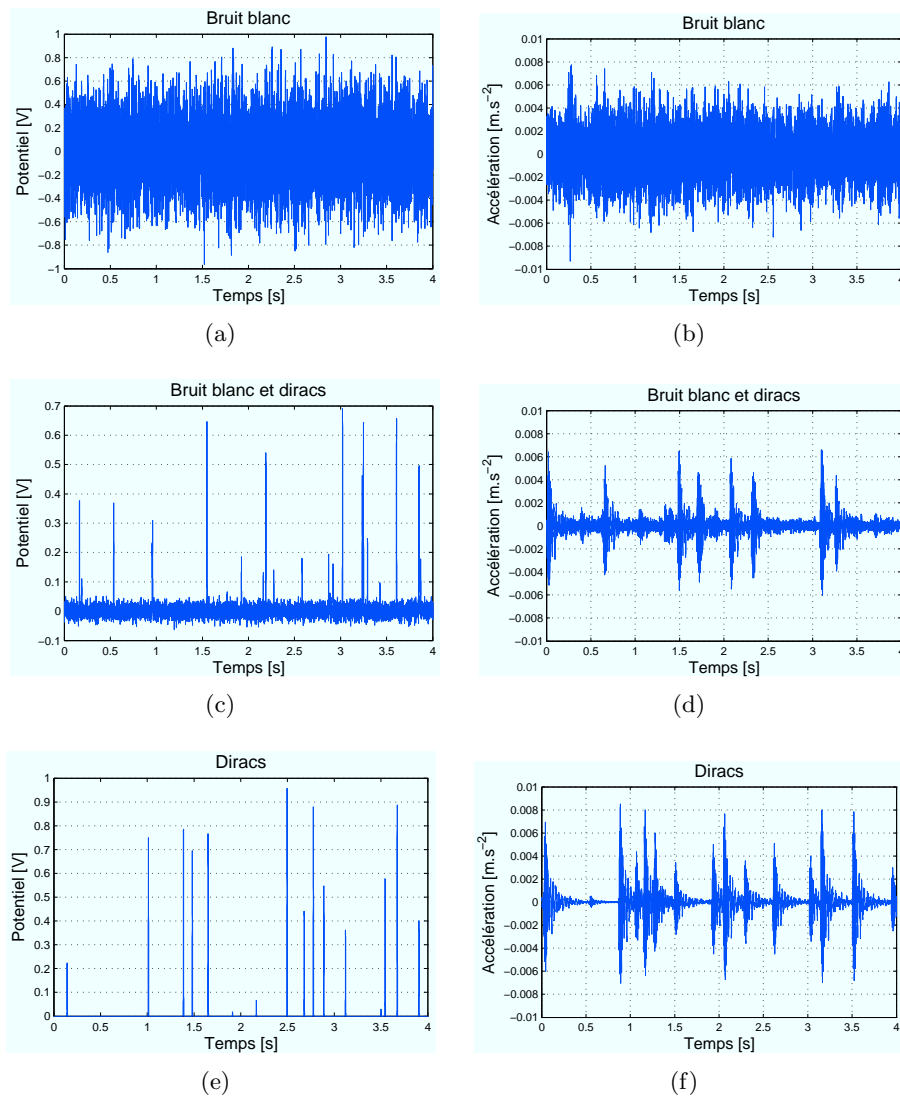


FIG. 4.12 Formes des trois types de signaux à l'entrée du système de contrôle du pot vibrant (à gauche) et réponses associées obtenues via l'accéléromètre 1 (à droite).

Conditions environnementales

Les variations environnementales sont générées par des changements de température et d'humidité.

Six lampes à rayonnement infra-rouges sont disposées verticalement afin de chauffer la structure de façon hétérogène. De plus, un climatiseur mobile permet de faire chuter rapidement la température en alternant les fonctions *ventilation* et *climatisation*. Le taux de saturation en eau du bois est modifié en pulvérisant directement de l'eau sur la maquette. Afin d'éviter l'éventuelle compensation des effets de l'humidité par ceux de la température, il convient de réaliser la pulvérisation à la fin d'une période de chauffage. L'ensemble du dispositif est placé à l'intérieur d'une enceinte climatique, dont les parois sont formées de deux couches de film plastique séparées par une lame d'air.

A l'instar de l'essai sur la poutre en béton, une année de suivi d'ouvrage est simulée par un cycle de variations extrêmes des conditions environnementales, comprenant :

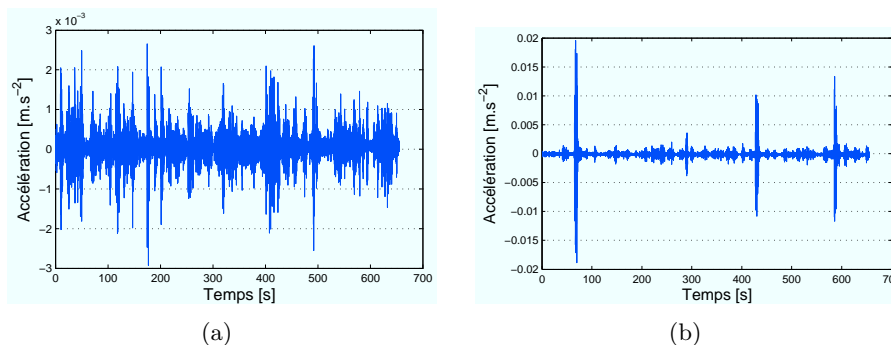


FIG. 4.13 Exemples de deux signaux mesurés pendant la campagne de mesures vibratoires sous sollicitations ambiantes du pont Z – 24.

- Refroidissement ($\simeq 43^{\circ}\text{C} \rightarrow 17^{\circ}\text{C}$)
 - 15 min ventilateur
 - 15 min climatisation
 - 15 min ventilateur
 - 15 min climatisation
- Chauffage : 1h lampes infra-rouges ($\simeq 17^{\circ}\text{C} \rightarrow 43^{\circ}\text{C}$)
- Humidification : pulvérisation

Avant chaque essai, la maquette subit un pré-conditionnement consistant à chauffer pendant 1h45 pour sécher le bois, puis à humidifier afin d'obtenir des conditions initiales proches d'un essai à l'autre. Un exemple de relevé de température et d'humidité relative pendant un essai, accompagné de l'annotation des différentes phases, est représenté dans la figure 4.14.

Endommagement

Tous les écrous de la structure sont initialement serrés à la main. L'endommagement est introduit ici de deux façons :

- en desserrant un par un la totalité des écrous d'un assemblage, en commençant par $E1$ jusqu'à $E4$ (Fig.4.10). Si l'assemblage du nœud i est supprimé, on notera A_i .
- en réduisant progressivement la tension d'un câble de contreventement par l'intermédiaire de l'écrou de serrage. Si la diagonale entre les nœuds 2 et 4 est affectée, on notera D_{24} .

Un essai se termine toujours par un retour de la structure à son état initial. Ce retour est réalisé en suivant l'ordre inverse de celui appliqué pendant la phase d'endommagement (ex : $E4 \rightarrow E1$). Cette précaution permet, dans

n° essai	N_s	N_e	Loc. Endom.	VE
22	182	73	A_2	non
24	180	128	D_{24}	non
25	180	85	A_1	non
34	400	122	D_{24}	oui
35	425	85	A_3	oui
36	425	80	A_1	oui

TAB. 4.3 Caractéristiques de chaque essai ; N_s = nombre de mesures saines ; N_e = nombre de séries endommagées ; VE = Variations environnementales.

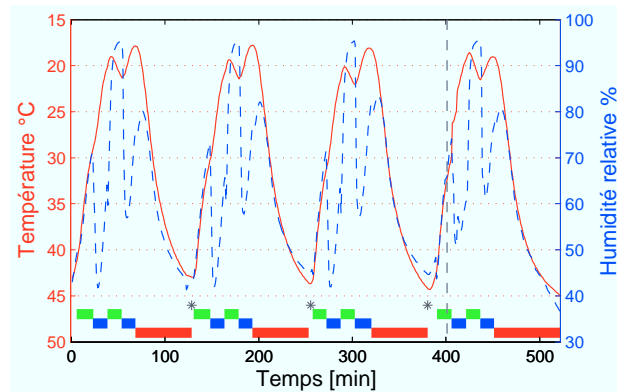


FIG. 4.14 Exemple de variations de température et d'humidité enregistrées lors d'un essai, le trait vertical à 400 min indique le début de la phase d'endommagement ; — Température ; - - - Humidité relative ; * humidification ; ■ ventilation ; ■ climatisation ; ■ chauffage.

un contexte de fortes variations environnementales, de mieux discriminer un changement causé par une modification structurale et un changement dû à une fluctuation extérieure.

Protocole et paramètres de mesure

Pour chacune des 3 excitations, une mesure est réalisée toutes les minutes pendant toute la durée de l'essai. Initialement échantillonnée à 8192Hz, chaque mesure est rééchantillonnée à 2048Hz et compte 8192 points.

En moyenne, dix enregistrements sont collectés par incrément d'endommagement. La base de référence couvre au minimum 3 cycles de variations environnementales, soit environ 240 mesures. L'endommagement est introduit au cours du cycle suivant (en général le 4^{ème}).

Afin de tester seulement la sensibilité à l'endommagement, certains essais sont réalisés sans variations environnementales. Le tableau 4.3 résume les informations relatives aux différents essais exploités dans les développements présentés ci-après.

BILAN

Ce chapitre introduit les différents cas d'études expérimentaux sur lesquels s'appuient les développements scientifiques proposés dans les chapitres 5 à 8. Chacun permet de tester, avec un degré croissant de difficulté, le potentiel des indicateurs et des approches de détection de nouveauté décrits au chapitre précédent.

Un modèle masse/ressort amorti, une maquette métallique et une maquette de structure en bois sont exploités pour évaluer la sensibilité à l'endommagement des indicateurs proposés en l'absence de variations environnementales.

L'efficacité de ces indicateurs, en présence de facteurs extérieurs affectant la dynamique vibratoire de la structure, est évaluée à l'aide d'une poutre en béton fibré sous sollicitations thermiques et de la maquette en bois précédente sous sollicitations thermiques et hydriques.

Enfin, la maquette de structure en bois permet également d'appréhender l'influence d'une sollicitation dynamique instationnaire, conjuguée aux variations environnementales, sur la capacité des approches proposées à détecter la nouveauté induite par l'endommagement.

Troisième partie

**Applications de la détection
de nouveauté**

INTRODUCTION

AVANT de tester l'approche de détection de nouveauté dans les conditions les plus défavorables, la réalisation d'un test dans un cas simple est nécessaire. Par cas simple, il faut comprendre l'analyse de mesures obtenues sous une excitation stationnaire et avec l'endommagement comme seul facteur influant sur la dynamique vibratoire de la structure.

L'objectif du chapitre 5 est, de vérifier sur le système masse/ressort amorti et sur la maquette métallique, la pertinence du nouvel indicateur basé sur l'estimation de la matrice jacobienne dans l'espace des phases (JFV), au regard du spectre des exposants de *Lyapunov*. Le JFV est ensuite comparé aux coefficients de modèles auto-régressifs par le biais d'expériences réalisées sur la maquette en bois.

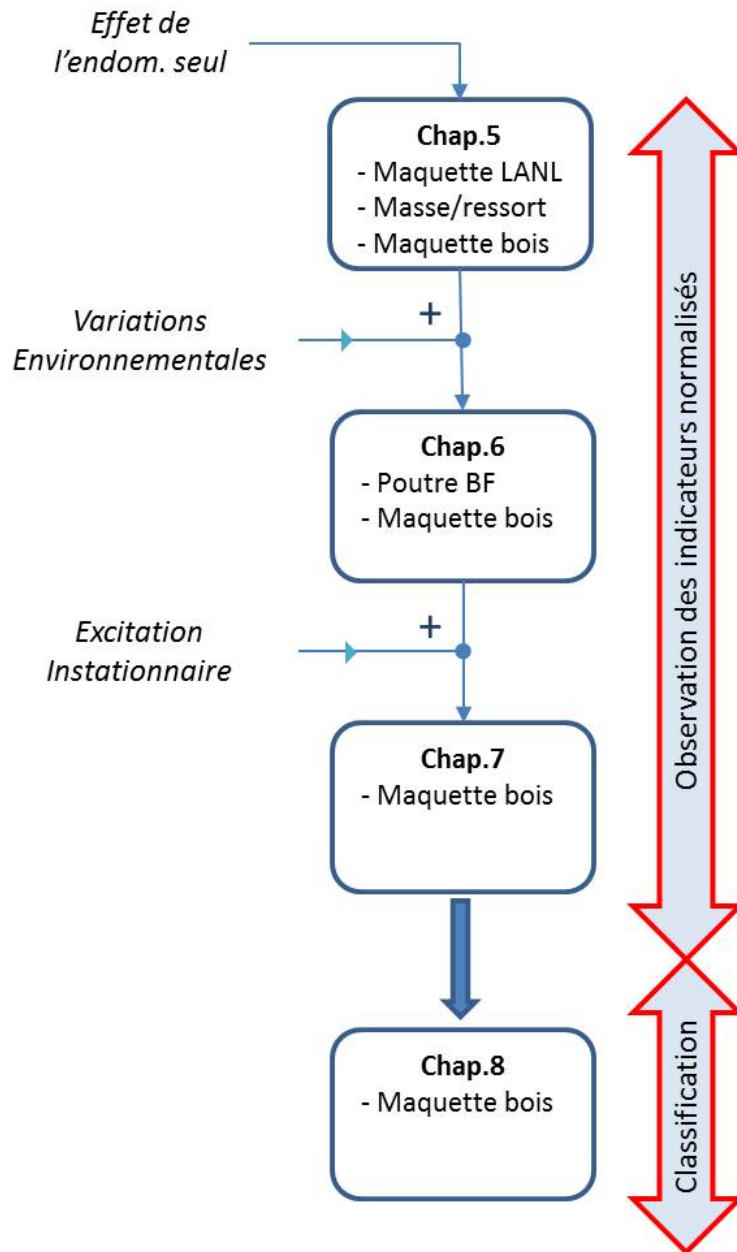
Dans le chapitre 6, les variations environnementales sont introduites dans le cadre des expérimentations réalisées sur la poutre en béton fibré et sur la maquette en bois. Le JFV et les paramètres de modèles AR ajustés sur les mesures sont utilisés pour caractériser le comportement des deux structures. Les questions abordées sont multiples : sensibilités comparées des indicateurs choisis vis à vis des modifications structurales, possibilité de localiser et/ou quantifier l'endommagement, problème de la division de la base de référence en une première partie dédiée à l'apprentissage, une deuxième à la classification, et une troisième au test de performance.

L'influence du caractère instationnaire de l'excitation est abordée dans le chapitre 7 dans le cadre de l'étude de la structure en bois, avec et sans variations environnementales.

Dans ces chapitres, le problème de la détermination du seuil de classification n'est pas évoqué. La variabilité des indicateurs normalisés est d'abord observée, ne nécessitant pas de modélisation statistique. Une limite générale est choisie, de façon indicative, à 3 écarts-types au dessus de la moyenne des indicateurs de la base de classification.

Enfin, le chapitre 8 est intégralement dédié au problème du choix de la limite de décision. Les différentes approches présentées dans la partie 3.3 (individuelles, par échantillons, valeurs extrêmes...) sont comparées sur quelques essais qui se caractérisent par une frontière floue entre mesures de la classe *structure endommagée* et mesures de la classe *structure saine*.

L'organigramme présenté sur la page suivante illustre la structuration des développements scientifiques proposées ci-après.



SENSIBILITÉ À L'ENDOMMAGEMENT

5

SOMMAIRE	
INTRODUCTION	96
5.1 ANALYSE COMPARÉE DES PERFORMANCES DES EXPOSANTS DE LYAPUNOV ET DU JFV	97
5.1.1 Étude paramétrique des algorithmes de calcul des EL et du JFV	97
5.1.2 Paramètres de reconstruction de l'espace des phases	97
5.1.3 Exploitation des données expérimentales	99
5.1.4 Résultats	99
5.1.5 Conclusion	104
5.2 COMPARAISON ENTRE COEFFICIENTS AR ET JFV	104
5.2.1 Données analysées	104
5.2.2 Paramètres de calcul	105
Ordre du modèle auto-régressif	105
Paramètres de reconstruction	107
5.2.3 Résultats	108
5.2.4 Conclusions	114
BILAN	115

INTRODUCTION

LE JFV, indicateur original proposé dans la section 3.1.4, est basé sur l'algorithme de calcul des exposants de *Lyapunov*. L'introduction du JFV comme solution alternative aux exposants de *Lyapunov* vise un double objectif de réduction des temps de calcul et d'accroissement de la sensibilité à un changement du comportement dynamique étudié. La première étape de ce travail de thèse consiste donc naturellement à comparer leurs performances sur des cas d'étude simples : le système masse/ressort amorti (MR) et la maquette métallique (MM) testée au *LANL*. Le travail réalisé ici cherche également à évaluer l'effet des différents paramètres des algorithmes de calcul des deux indicateurs.

Dans un second temps, le JFV est comparé aux coefficients de modèles auto-régressifs (AR) très populaires en matière de caractérisation des séries temporelles. Le support de cette comparaison est constitué par des expérimentations conduites sur la maquette en bois (MB) selon différents scénarios d'endommagement.

5.1 ANALYSE COMPARÉE DES PERFORMANCES DES EXPOSANTS DE LYAPUNOV ET DU JFV

5.1.1 Étude paramétrique des algorithmes de calcul des EL et du JFV

L'estimation des exposants de *Lyapunov* et du JFV nécessite l'optimisation de plusieurs paramètres.

Le nombre d'itérations L , pendant lesquelles la trajectoire est suivie, doit assurer la convergence des EL. Cependant, en raison de la longueur limitée de la mesure, il est difficile d'aller au-delà de $L = 2000$. La partie transitoire de la trajectoire reconstruite dans l'espace des phases représente environ 500 itérations qui sont écartées du calcul de la matrice produit. Dans le cas du JFV, la valeur de L est limitée volontairement à 5 et l'étude paramétrique ci-après présente l'effet de ce paramètre sur une plage variant de 1 à 5.

Il en est de même pour T_2 , paramètre définissant le nombre de pas de temps sur lequel agit la transformation ($X_{k+T_2} = T(X_k)$). L'estimation du JFV est réalisée pour $T_2 = 1, \dots, 5$. Pour les EL, le paramètre T_2 est limité à 1 et 2, la longueur des mesures (8192 points) empêchant d'explorer des valeurs supérieures. En effet, la quantité $k + T_2L$ doit demeurer inférieure à 8192 pour être en mesure de suivre chaque point suffisamment longtemps le long de sa trajectoire. Avec $T_2 = 3$, il ne reste que 2192 points utilisables pour l'estimation, ce qui est insuffisant.

Pour une évaluation correcte des coefficients de la transformation T par la méthode des moindres carrés, il est nécessaire que le nombre de points sur lesquels elle est ajustée soit au moins supérieur au nombre d'inconnues. La taille d'un voisinage est ainsi choisie de façon à contenir deux fois plus de points que de coefficients à ajuster.

La transformation étant choisie polynomiale, les calculs sont menés avec des polynômes de degrés variant de 1 à 3. Un degré supérieur implique des voisinages trop grands. Le tableau 5.1 synthétise les valeurs explorées dans le cadre de l'étude paramétrique.

5.1.2 Paramètres de reconstruction de l'espace des phases

La reconstruction de l'espace des phases exige la détermination de la dimension de reconstruction m , et du décalage temporel τ .

Indicateur	T_2	L	Degré
Exposants de <i>Lyapunov</i>	1	2000	1,2,3
	2	2000	1,2
JFV	2	2	1,2,3
	1,2,3	1	1
	1,2,3,4,5	2	1
	2	1,2,3,4,5	1

TAB. 5.1 Valeurs des paramètres étudiés pour l'estimation des exposants de Lyapunov et du JFV.

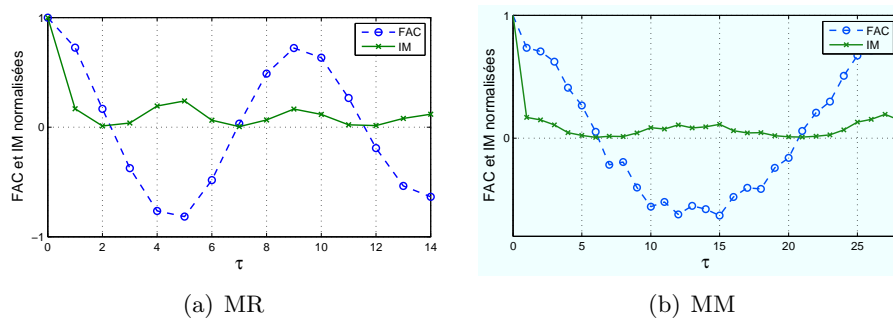


FIG. 5.1 Fonctions d'auto-corrélation et d'information mutuelle de signaux choisis au hasard pour le système masse ressort (MR) et la maquette métallique (MM).

Les figures 5.1(a) et 5.1(b) présentent les fonctions d'auto-corrélation et d'information mutuelle construites sur la base de signaux pris au hasard dans les bases de référence relatives au système masse/ressort et à la maquette métallique. Les allures de ces fonctions sont similaires pour les autres séries temporelles des bases de référence. Comme attendu, le premier minimum de la fonction d'information mutuelle et la première racine de la fonction d'auto-corrélation s'accordent sur la même valeur de retard optimal τ dans les deux cas étudiés. Ce retard est égal à 2 pour le système MR, et à 6 pour la MM. Cependant, selon les facteurs de l'étude paramétrique présentés plus haut, T_2 varie entre 1 et 5. Choisir des valeurs identiques pour τ et T_2 équivaut à rechercher un opérateur qui transforme un voisinage en un autre voisinage qui ne diffère que d'une seule composante. Pour éviter ce problème, τ est pris égal à 6 pour les deux essais.

Le choix de la dimension de reconstruction est également sujet à discussion. Le calcul du pourcentage de faux voisins en fonction de la dimension (Fig. 5.2) ne présente pas de diminution clairement identifiée. La raison émane principalement du caractère non déterministe des mesures dû à l'excitation stochastique (bruit blanc). En se plaçant néanmoins sur la zone de plus grande pente, et compte tenu du coût des calculs lié à une dimension importante, la valeur de 7 est choisie pour les deux structures.

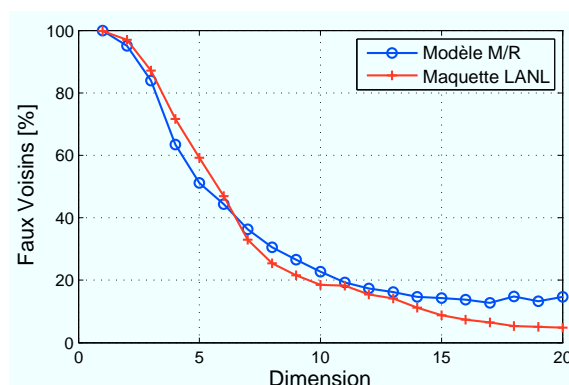


FIG. 5.2 Pourcentage de faux voisins en fonction de la dimension de reconstruction pour MR et MM.

T_2	Degré	% Points mal classés	
		MR	MM
1	1	15	41
1	2	23	53
1	3	24	62
2	1	25	61
2	2	31	60

TAB. 5.2 Pourcentage de points mal classés par les exposants de Lyapunov pour chaque jeu de paramètres testé (système masse/ressort et maquette métallique).

5.1.3 Exploitation des données expérimentales

La distance de Mahalanobis (MD) est utilisée ici afin de comparer les indicateurs vectoriels (§3.2). Pour cela, la base de données collectée sur la structure non endommagée est divisée aléatoirement en trois parties :

- 50% des données composent la base d'apprentissage dédiée au calcul de la distance de Mahalanobis,
- 25% des données composent la base de classification, utilisée pour ajuster le modèle statistique ; il est considéré comme gaussien dans cette partie et la limite de décision est fixée à 3 écarts-types au-dessus de la moyenne.
- 25% des données forment la base de test, qui permet d'estimer le taux de fausses alarmes.

Cette répartition des données est discutée dans le chapitre 6, et la pertinence du seuil dans le chapitre 8. Néanmoins, dans une optique de comparaison objective et rapide, les performances des 2 indicateurs sont appréhendées par le biais du pourcentage de points mal classés (fausse alarme et non détection).

5.1.4 Résultats

Les tableaux 5.2 et 5.3 détaillent, pour les 2 indicateurs, les pourcentages de points mal classés selon le jeu de paramètres testé. Il est à noter que ces points sont essentiellement composés de *non détections* car les fausses alarmes sont rares, comme le montrent les figures 5.4, présentées ci-après.

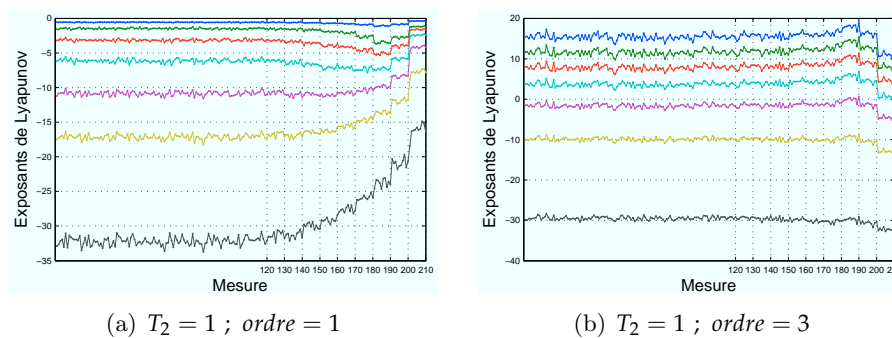


FIG. 5.3 Exposants de Lyapunov en fonction des mesures réalisées sur le système MR.

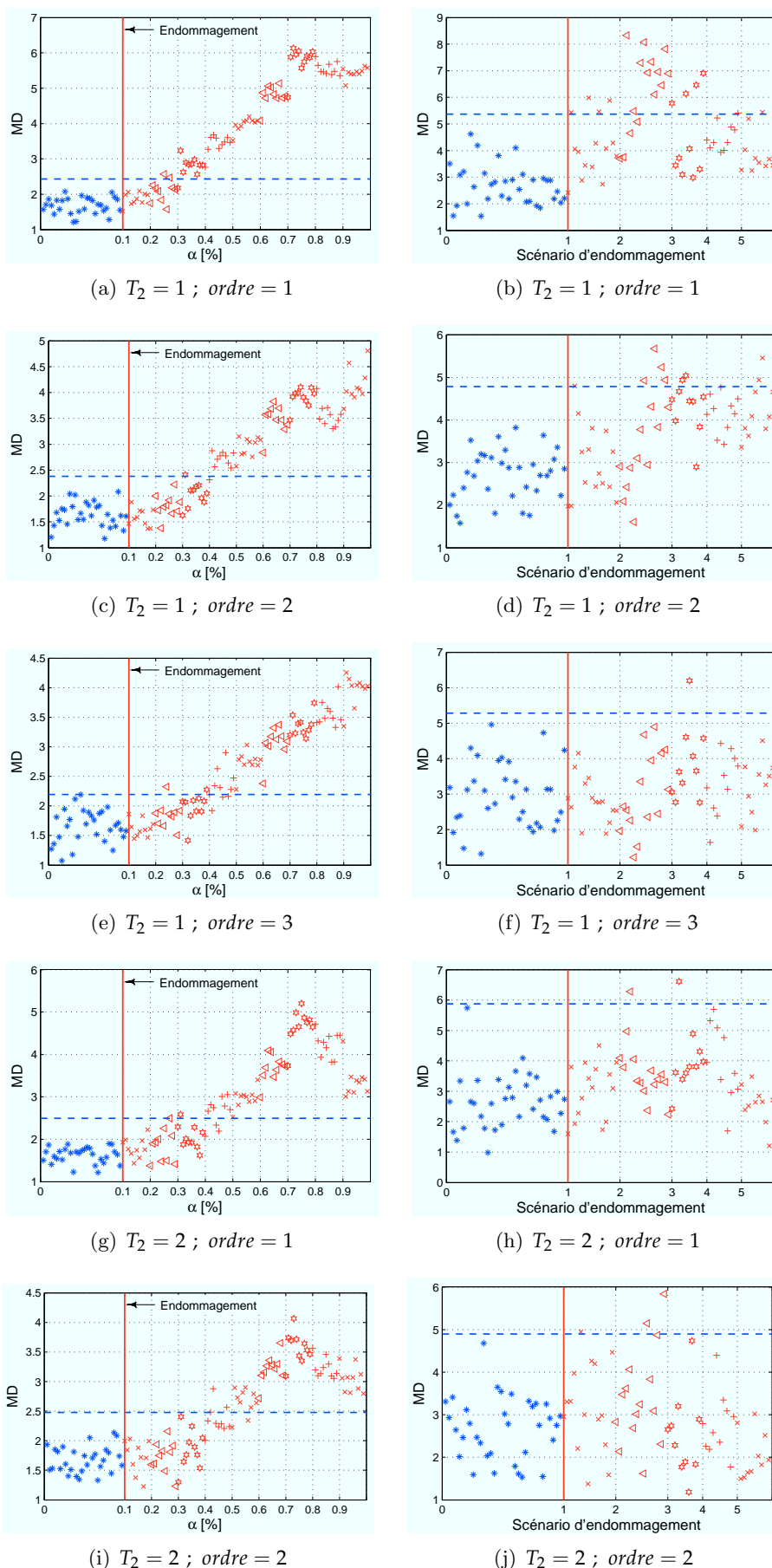


FIG. 5.4 Distance de Mahalanobis calculée sur spectre de Lyapunov en fonction de l'endommagement pour les 5 jeux de paramètres de calcul (MR à gauche, MM à droite).

Exposants de Lyapunov

Le tableau 5.2 synthétise les performances de classification des EL sur les données MR et MM. Une aptitude assez moyenne à la classification des données entre *structure saine* et *structure endommagée* est observée. Elle est même médiocre sur les données MM. Utiliser un degré supérieur pour la transformation polynomiale permet en théorie une meilleure estimation des EL. Cependant, les résultats expérimentaux montrent que l'augmentation du degré réduit la sensibilité à l'endommagement, puisque le nombre de points mal classés a tendance à augmenter pour MR et MM. La figure 5.3 présente l'évolution du spectre de Lyapunov pour les degrés 1 et 3 en fonction des mesures réalisées sur les données MR (structure saine pour les mesures 1 à 120 et endommagement croissant de la mesure 121 à la mesure 210). Sur les graphiques présentés, chaque courbe représente l'évolution d'un exposant, l'ensemble des exposants constituant le spectre de Lyapunov. Il apparaît que l'augmentation du degré polynomial engendre des exposants positifs, témoignant qu'une composante aléatoire des mesures est probablement exprimée par la transformation¹. De plus, le nombre de points composant l'espace des phases n'est pas suffisant pour tirer profit d'un degré plus important.

Les figures 5.4 tracent, pour les données MR et MM, l'évolution la distance de Mahalanobis (MD) calculée sur le spectre de Lyapunov en fonction du niveau d'endommagement. Chaque point bleu représente une mesure réalisée sur structure saine (base de test) et chaque point rouge, une mesure collectée sur structure endommagée. Les lignes verticales en trait pointillé délimitent les différents niveaux d'endommagement éprouvés par la structure et décrits dans le chapitre 4. La ligne rouge verticale marque le début de la séquence d'endommagement et la ligne pointillée horizontale matérialise la limite de classification fixée à 3 écarts-types au dessus de la moyenne calculée sur la base de classification. Enfin, chaque graphique est associé à un jeu de paramètres de l'algorithme de calcul des EL.

Il apparaît sur les figures 5.4 qu'un pas de transformation T_2 unitaire réduit la dispersion des points figuratifs des distances de Mahalanobis des mesures, amenant ainsi un indicateur plus fiable. Un pas T_2 égal à 2 présente environ 10% de points mal classés en plus du fait de l'augmentation de la dispersion, bien que ce pourcentage dépende de la limite de classification fixée.

Ces observations sont surtout valables pour le système Masse/Ressort, car les EL ne sont quasiment pas sensibles à l'endommagement induit dans la maquette métallique comme en témoigne la colonne de graphiques à droite dans la figure 5.4(b) (sauf cas (b)). Les données MM étant issues de mesures expérimentales, elles présentent un niveau de bruit potentiellement responsable de ce manque de sensibilité.

Le JFV

Le tableau 5.3 présente les performances de classification du JFV pour l'ensemble des jeux de paramètres testés. De façon générale, l'aptitude à la classification est assez homogène d'un jeu de paramètres à l'autre, même

¹Le bruit étant par définition de dimension infinie et totalement imprévisible, ses exposants de Lyapunov sont positifs.

Cas	T_2	L	Ordre	% points mal classés	
				MR	MM
1			1	27	0
2	2	2	2	34	1
3			3	34	5
4	1			51	14
5	2	1	1	43	13
6	3			36	16
7	1			30	0
8	2			25	0
9	3	2	1	42	9
10	4			28	1
11	5			46	1
12		1		46	13
13		2		32	0
14	2	3	1	49	4
15		4		30	11
16		5		21	0

TAB. 5.3 *Pourcentage de points mal classés par le JFV pour chaque jeu de paramètres testé (système masse/ressort et maquette métallique).*

si quelques observations particulières peuvent être formulées. La sensibilité à l'endommagement est par conséquent assez uniforme pour l'ensemble des cas testés.

Pour le système MR, le pire cas de figure compte 51% de points mal classés, contre 21% pour le meilleur. Concernant les essais relatifs à la MM, le pire résultat du test se porte à 16% de points mal classés, plusieurs jeux conduisant à une classification parfaite (0 %). Cette différence s'explique probablement par un endommagement plus progressif, et donc moins détectable dans le cas du système MR, comparativement à la MM. Aucun jeu de paramètre ne permet de déceler une diminution de raideur du système MR inférieure à 40%. La figure 5.5 représente l'évolution des distances de Mahalanobis, associées aux meilleurs et pires performances de classification, en fonction de l'endommagement pour les 2 structures testées. De très bonnes performances sont obtenues sur la maquette métallique.

L'observation plus fine des résultats révèle cependant des tendances dans l'influence des paramètres. Le tableau 5.3 confirme que l'augmentation du degré polynomial n'améliore pas la sensibilité. La transformation linéaire est la plus performante.

L'influence de T_2 n'est pas claire. Quand $L = 1$ (cas 4, 5 et 6), un pas de transformation important améliore la sensibilité pour le système MR, mais ce constat n'est pas confirmé sur la MM. Alors que, si $L = 2$, hormis le cas 9 (qui est expliqué par la suite), la détection reste excellente sur la MM pour toutes les valeurs de T_2 . Cette observation montre que T_2 n'a pas d'effet direct sur la sensibilité à un changement de comportement. Son impact doit être considéré en interaction avec L .

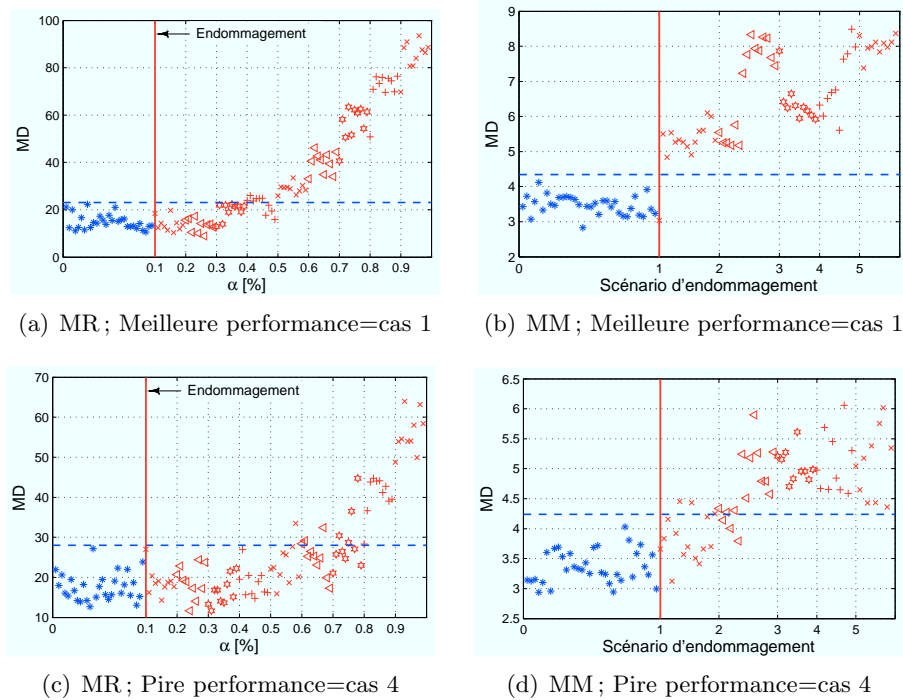


FIG. 5.5 Distance de Mahalanobis calculée sur JFV en fonction de l'endommagement pour les jeux de paramètres fournissant les meilleurs et les pires résultats de classification.

Par ailleurs, des variations irrégulières sont observées lorsque L change. Pour les deux structures, aucune tendance ne se dessine lorsque la valeur L augmente. En revanche pour certains cas, le nombre de points mal classés est anormalement grand :

- Cas 4-5-6-12 $\Rightarrow L = 1$
- Cas 9 $\Rightarrow L = 2$ et $T_2 = 3$
- Cas 14 $\Rightarrow L = 3$ et $T_2 = 2$
- Cas 15 $\Rightarrow L = 4$ et $T_2 = 2$.

Indépendamment des valeurs des autres paramètres, lorsque $L = 1$, le JFV ne donne jamais de résultats satisfaisants. Cette valeur semble par conséquent à proscrire lors du paramétrage de l'algorithme de calcul. En admettant cette conclusion, les mauvaises performances concernant les cas 9, 14, et 15 s'expliquent simplement en observant la façon dont sont construits les vecteurs (Fig.5.6). Lorsque $L = 4$ et $T_2 = 2$, la matrice jacobienne est estimée entre les vecteurs X_7 et X_9 qui ne diffèrent que par une composante des vecteurs X_1 et X_3 , ce qui équivaut à un cas où $L = 1$. Par le même raisonnement, le jeu ($L = 3$; $T_2 = 2$) envoie X_5 vers X_7 qui sont proches de X_1 et X_5 , correspondant au cas $L = 1$ et $T_2 = 4$. L'explication est identique pour $L = 2$ et $T_2 = 3$.

Chaque fois que la matrice jacobienne utilisée pour former le JFV inclut le voisinage initial, ou une variante à une ou deux composantes près (selon la dimension), la sensibilité à l'endommagement est réduite. Sans avoir de certitudes, il est possible que le voisinage nécessite d'occuper une zone plus importante de l'espace des phases pour que la matrice jacobienne puisse mieux capter les changements de la dynamique.

X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9	X_{10}
x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}
x_7	x_8	x_9	x_{10}	x_{11}	x_{12}	x_{13}	x_{14}	x_{15}	x_{16}
x_{13}	x_{14}	x_{15}	x_{16}	x_{17}	x_{18}	x_{19}	x_{20}	x_{21}	x_{22}
x_{19}	x_{20}	x_{21}	x_{22}	x_{23}	x_{24}	x_{25}	x_{26}	x_{27}	x_{29}
x_{25}	x_{26}	x_{27}	x_{28}	x_{29}	x_{30}	x_{31}	x_{32}	x_{33}	x_{34}
x_{31}	x_{32}	x_{33}	x_{34}	x_{35}	x_{36}	x_{37}	x_{38}	x_{39}	x_{40}
x_{37}	x_{38}	x_{39}	x_{40}	x_{41}	x_{42}	x_{43}	x_{44}	x_{45}	x_{46}

FIG. 5.6 Dix premiers vecteurs reconstruits ($m = 6$ et $\tau = 6$).

5.1.5 Conclusion

Sur les données résultant du système masse/ressort, les exposants de Lyapunov détectent plus rapidement la réduction de raideur (dès 20% contre 40% pour le JFV). Cependant, le JFV se montre capable de discriminer la totalité des mesures relatives à la structure endommagée dans le cadre des données réelles collectées sur la maquette du *LANL*. Dans ce dernier cas, la meilleure performance atteinte par les EL est égale à 41 % de points mal classés.

Les temps de calcul propres à la détermination des 2 indicateurs sont significativement différents. L'extraction du spectre de *Lyapunov* caractérisant une seule mesure nécessite entre 170 et 3700 secondes selon le degré de la transformation polynomiale (processeur Intel® Xeon cadencé à 3 GHz). En revanche, selon les paramètres choisis, l'extraction du JFV ne nécessite que 0,2 à 6 secondes.

Cependant, pour assurer au JFV une sensibilité optimale, il faut éviter d'exploiter la matrice jacobienne estimée sur le voisinage initial. Cette condition est vérifiée si :

$$T_2 \cdot L \neq k\tau \text{ et} \quad (5.1)$$

$$T_2 \cdot (L - 1) \neq k\tau \text{ avec } k \in \mathbb{N} \text{ et } L > 1. \quad (5.2)$$

Sur la base des résultats discutés ci-dessus, les valeurs recommandées pour le calcul du JFV sont :

- $degr = 1$ (transformation linéaire),
- $L = 2$,
- $T_2 = 2$ (ou 1).

La meilleure configuration ($T_2 = 2$) empêche donc de choisir les valeurs 2 et 4 pour le délai de reconstruction τ .

Cette étude préliminaire permet de démontrer l'intérêt potentiel de l'indicateur original proposé (JFV) et de guider le choix de ses paramètres de calcul. Les bons résultats du JFV en matière de classification se doivent d'être confirmés sur d'autres structures expérimentales et, surtout, d'être comparés avec les performances d'indicateurs plus traditionnels.

5.2 COMPARAISON ENTRE COEFFICIENTS AR ET JFV

5.2.1 Données analysées

Les données analysées sont issues de la maquette en bois excitée par un bruit blanc de bande de fréquence de 800Hz, sans variations environnementales. Le tableau 5.4 regroupe les informations relatives aux essais. Pour

Essai	N_b	N_e	Endom.	Fin Endom.
22	182	73	A_2	231
24	180	128	D_{24}	260
25	180	85	A_1	230

TAB. 5.4 *Récapitulatif des mesures analysées, avec de gauche à droite, le numéro de l'essai, le nombre de mesures saines N_b , le nombre de mesures endommagées N_e , la localisation de l'endommagement et le numéro de la mesure marquant le début du retour aux conditions initiales.*

mémoire, chaque mesure est échantillonnée à 2048Hz et compte 8192 points. Enfin, tous les calculs sont réalisés après avoir centré individuellement les signaux.

5.2.2 Paramètres de calcul

Un échantillon de signaux est prélevé pour chaque accéléromètre afin d'estimer les paramètres de calcul des indicateurs. Les propriétés des signaux sont supposées invariantes d'un essai à l'autre. Ces paramètres sont utilisés pour tous les autres essais réalisés avec la même forme de sollicitation.

Ordre du modèle auto-régressif

La fonction d'*auto-corrélation partielle* et le *critère d'Akaike* sont estimés sur 50 mesures de l'essai 24 (10 mesures par accéléromètre). Pour chacune d'elles, l'ordre du modèle est choisi lorsque les valeurs de la fonction d'auto-corrélation semblent présenter une distribution aléatoire, et lorsque le critère atteint son minimum. En pratique, ce critère ne présente cependant pas de minimum et il convient alors de rechercher la valeur à partir de laquelle il se stabilise. Malgré un effort de rigueur, le choix demeure subjectif. La figure 5.7 montre, pour une mesure, les courbes obtenues pour chaque accéléromètre et l'ordre retenu associé.

L'ordre optimal se situe entre 31 et 42, comme le montre le tableau 5.5. Les valeurs de l'erreur de prédiction de chaque signal doivent alors être indépendantes. Pour vérifier si cette propriété est assurée, la densité spectrale de puissance (PSD) de l'erreur est tracée sur la figure 5.8 pour un $AR(15)$, un $AR(40)$ et un $AR(90)$. Il apparait clairement que la propriété n'est pas vérifiée pour un ordre inférieur à 90. Cependant, il n'est pas envisageable de choisir cette valeur. Plus la taille de l'indicateur vectoriel est grande, plus la

accéléromètre	min	moy	max
1	40	40	40
2	36	36.7	39
3	30	31.7	32
4	22	33.7	45
5	27	41.5	50

TAB. 5.5 *Ordres moyen, minimum et maximum obtenus pour chaque accéléromètre (échantillon de 10 mesures par accéléromètre).*

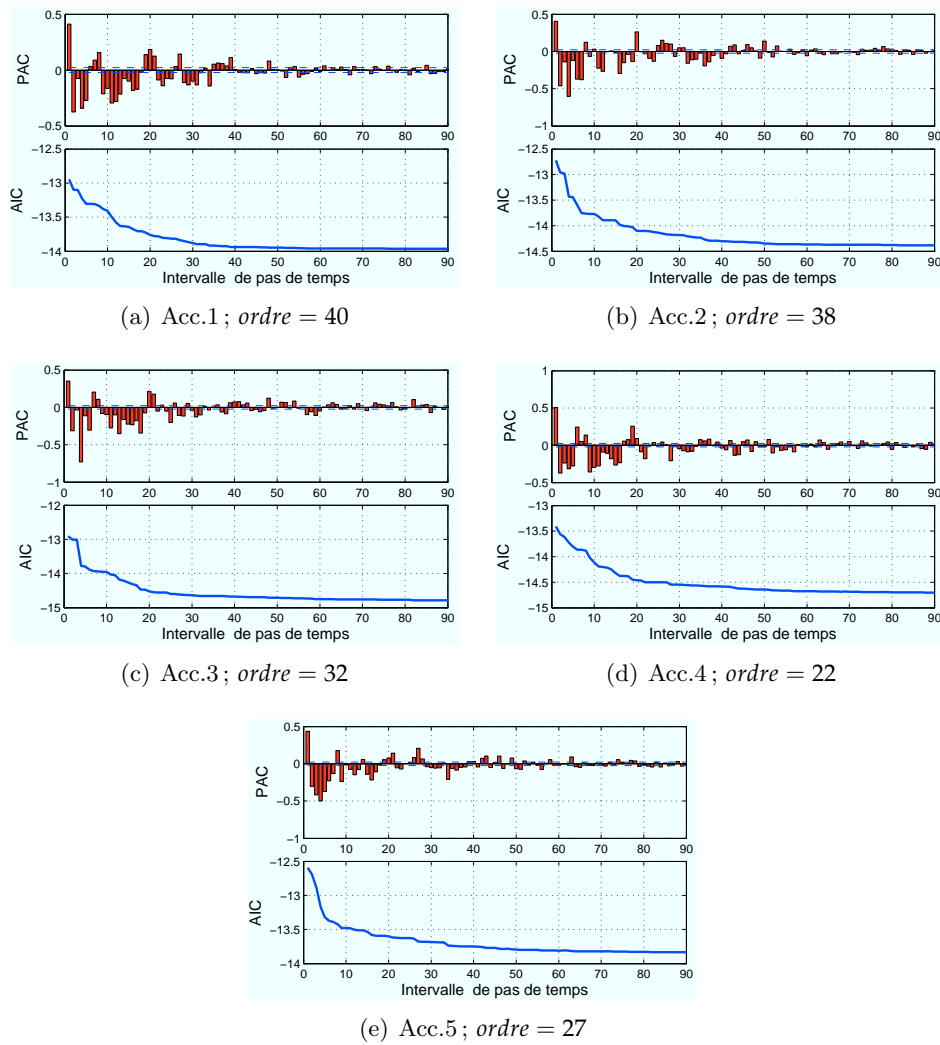


FIG. 5.7 Fonction d'auto-corrélation partielle (PAC) et critère d'Akaike (AIC) estimés sur une mesure pour chaque accéléromètre.

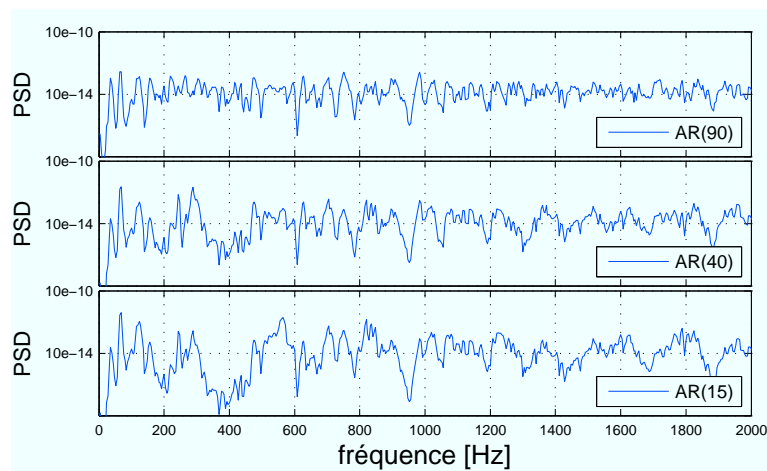


FIG. 5.8 Densité spectrale de puissance de l'erreur de prédiction (essai 24, Acc.4, mes 97)

quantité de données nécessaire pour caractériser la population de référence est importante (*fléau de la dimension*).

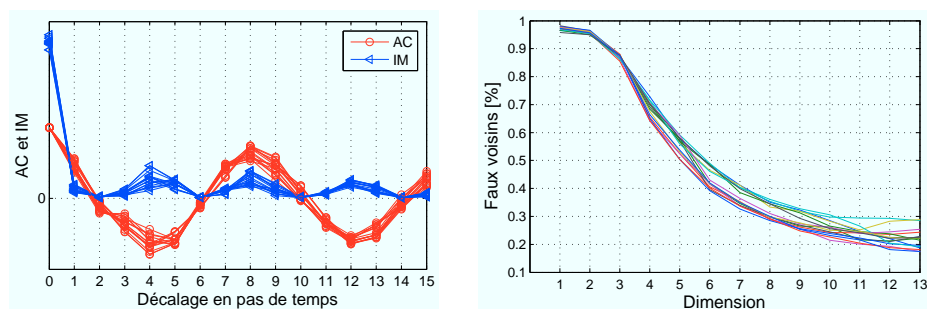
L'ordre idéal (**90**) n'étant pas envisageable en pratique, la valeur de **40** est retenue pour tous les accéléromètres. Elle correspond à la limite supérieure des ordres constatées sur les fonctions d'auto-corrélation partielle (Tab.5.5). Les conséquences de ce choix sont évaluées en testant un AR(20) et un AR(60) sur les données de l'essai 24 à la fin de ce chapitre.

Paramètres de reconstruction

Pour chaque accéléromètre, trois mesures sont choisies au hasard parmi les 182 mesures de la base de référence de l'essai 24, soit 15 signaux sélectionnés au total.

Les fonctions d'information mutuelle et d'auto-corrélation sont calculées pour les trois séries temporelles de chaque accéléromètre. Le graphique de la figure 5.9(a) identifie clairement un retard optimal de 2 pas de temps, avec peu de dispersion entre les courbes relatives aux 15 signaux. Cependant, pour obtenir la meilleure sensibilité du JFV, l'étude précédente montre que les valeurs 2 et 4 sont à proscrire pour le paramètre de décalage temporel. Le choix se porte alors sur la valeur $\tau = 6$ représentant le second minimum de la fonction d'information mutuelle et la deuxième racine de la fonction d'auto-corrélation.

Une nouvelle fois, le tracé des pourcentages de faux voisins en fonction de la dimension de reconstruction ne permet pas de trancher sur une dimension optimale (Fig.5.9(b)). Au-delà de 10, le nombre de faux-voisins tend à se stabiliser. Cependant, pour une dimension $p = 10$, le JFV compte 100 composantes, ce qui n'est pas envisageable. La dimension de reconstruction est donc fixée à **6**, dimension correspondant à une réduction de 50% des faux voisins. Ainsi, le nombre de composantes du JFV est proche du nombre de paramètres du modèle AR(40), plaçant les 2 indicateurs sur un pied d'égalité pour la comparaison.



(a) Fonctions d'information mutuelle (IM) et d'auto-corrélation (AC) des 15 signaux (b) Pourcentage de faux voisins en fonction de la dimension de reconstruction pour les 15 signaux.

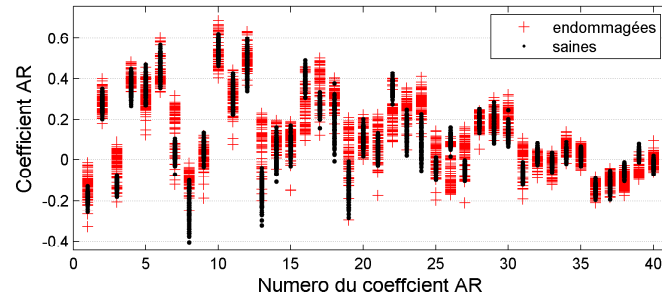
FIG. 5.9 Détermination des paramètres de reconstruction sur la base des 15 signaux sélectionnés (3 signaux par accéléromètre).

5.2.3 Résultats

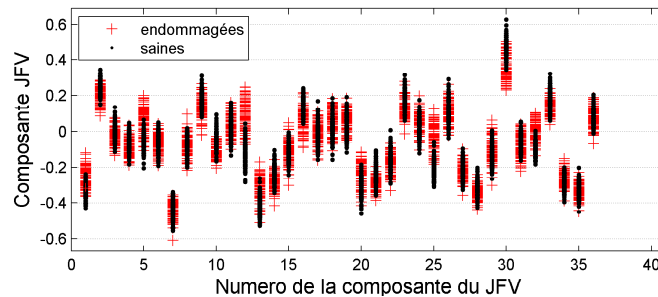
Comportements typiques des composantes des 2 indicateurs

Dans le cas d'un indicateur vectoriel idéal, toutes les composantes seraient sensibles à l'endommagement, variant de façon monotone avec la progression de celui-ci. En pratique, un tel indicateur n'existe pas à ce jour. Concernant les 2 indicateurs vectoriels testés ici, toutes les composantes ne sont pas sensibles à l'endommagement. La figure 5.10 trace les gammes de variation de chaque composante des 2 indicateurs vectoriels sur l'ensemble des mesures de l'accéléromètre 1 lors de l'essai 22 (structure saine en noir et structure endommagée en rouge). Par exemple, les coefficients 11, 36, 15 du modèle AR ne semblent pas affectés par le desserrage des écrous de l'assemblage 2, puisque les points noirs (état sain) et les points rouges (état endommagé) occupent sensiblement la même plage de variation. La même remarque peut être formulée au sujet des composantes 4, 11 et 78 du JFV. De façon générale, le nombre de composantes, pour lesquelles on observe une superposition des plages de variations associées aux états sain et endommagé, est plus important pour le JFV que pour les coefficients AR, laissant présumer une meilleure sensibilité de ces derniers à l'endommagement.

L'observation des composantes prises individuellement fait apparaître différents types de comportements lors des essais. La figure 5.11 présente des exemples de variations de deux composantes de chaque indicateur lors de l'essai 22 (Accéléromètre 1). Les points situés à gauche du trait pointillé vertical représentent les mesures réalisées sur la structure saine et les points situés à droite représentent les séries collectées durant la séquence d'endommagement réversible. Dans le cas du modèle AR, les sensibilités



(a) AR



(b) JFV

FIG. 5.10 *Gammes de variation de chaque composante des 2 indicateurs vectoriels sur l'ensemble des mesures relatives à l'essai 22 (Accéléromètre 1).*

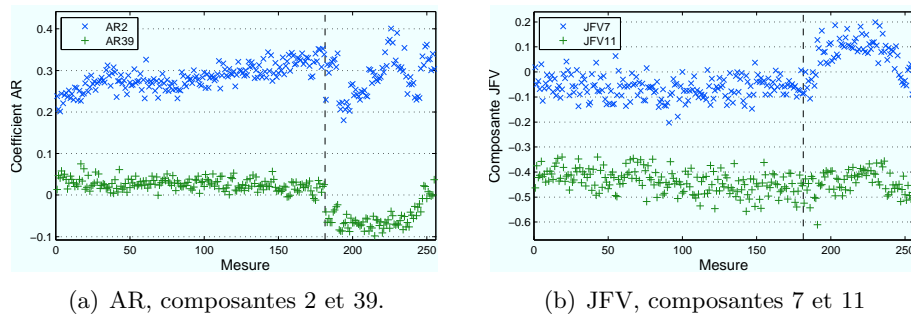


FIG. 5.11 Exemples de variations de deux composantes de chaque indicateur en fonction de la mesure lors de l'essai 22 (Accéléromètre 1) : avant endommagement (à gauche du trait pointillé vertical) et durant la séquence d'endommagement réversible (à droite du trait pointillé vertical)

des composantes sont diverses (Fig.5.11(a)). La composante AR39 se comporte de façon optimale en suivant l'endommagement de façon monotone, y compris dans sa phase de réversibilité, alors que la composante AR2 semble sensible, mais de façon désordonnée. Concernant le JFV (Fig.5.11(b)), on retrouve, par exemple pour la composante JFV7, un comportement proche de celui du coefficient AR39, suivant l'endommagement et sa phase de réversibilité (à la fin de l'essai, la composante retrouve une valeur proche de celle obtenue sur structure saine). À l'inverse, la composante JFV11 ne manifeste aucune sensibilité au cycle d'endommagement. La richesse contenue dans la réponse individuelle à un endommagement de chaque composante des indicateurs vectoriels est assurément un facteur clé du succès de démarche de détection de nouveauté.

Détection de l'endommagement

Les figures 5.12, 5.13 et 5.14 synthétisent les résultats en termes de détection et de suivi d'endommagement pour les essais 22, 24 et 25. Elles présentent, pour chaque essai et pour chaque indicateur, la distance de Mahalanobis (indicateur normalisé et réduit à une grandeur scalaire) en fonction de la mesure. Le trait pointillé vertical marque ici encore le début de la séquence d'endommagement réversible et la ligne pointillée horizontale matérialise la limite de classification fixée à 3 écarts-types au-dessus de la moyenne, ces 2 moments statistiques étant calculés sur la base de classification. Il apparaît tout d'abord que les coefficients AR manifestent une excellente sensibilité aux modifications imposées à la structure. Pour chaque essai, et pour chaque accéléromètre, la distance de Mahalanobis croît de façon suffisamment significative pour marquer une différence avec les mesures de la base de test. En revanche, ce comportement n'est pas observé systématiquement pour le JFV. Sur 15 cas analysés (5 accéléromètres sur 3 essais), 8 cas discriminent sans ambiguïté les données acquises sur structure dégradée, 3 cas présentent une variation faible de la distance de Mahalanobis, mais exploitable à la condition de disposer d'une limite de classification fixée avec précision, et enfin 4 cas n'indiquent aucun changement lors du cycle d'endommagement. Sur ces 4 derniers cas, l'accéléromètre 3 est systématiquement impliqué (3 cas sur 4). Ce constat, difficilement explicable à ce stade, peut être rapproché de la performance plus faible des coefficients AR extraits de l'accéléromètre 3 lors des essais 22 et 25 (Fig.5.12 et Fig.5.14).

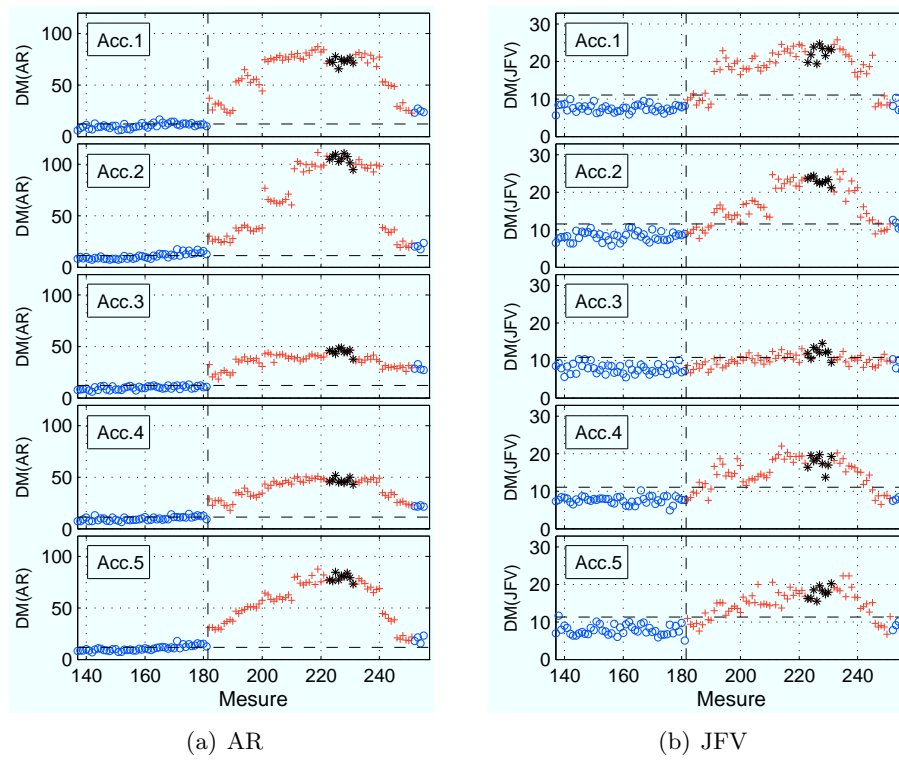


FIG. 5.12 Distance de Mahalanobis en fonction de la mesure, *essai 22*, endommagement A_2 ; \circ base de test; $+$ structure endommagée; $*$ endommagement maximum.

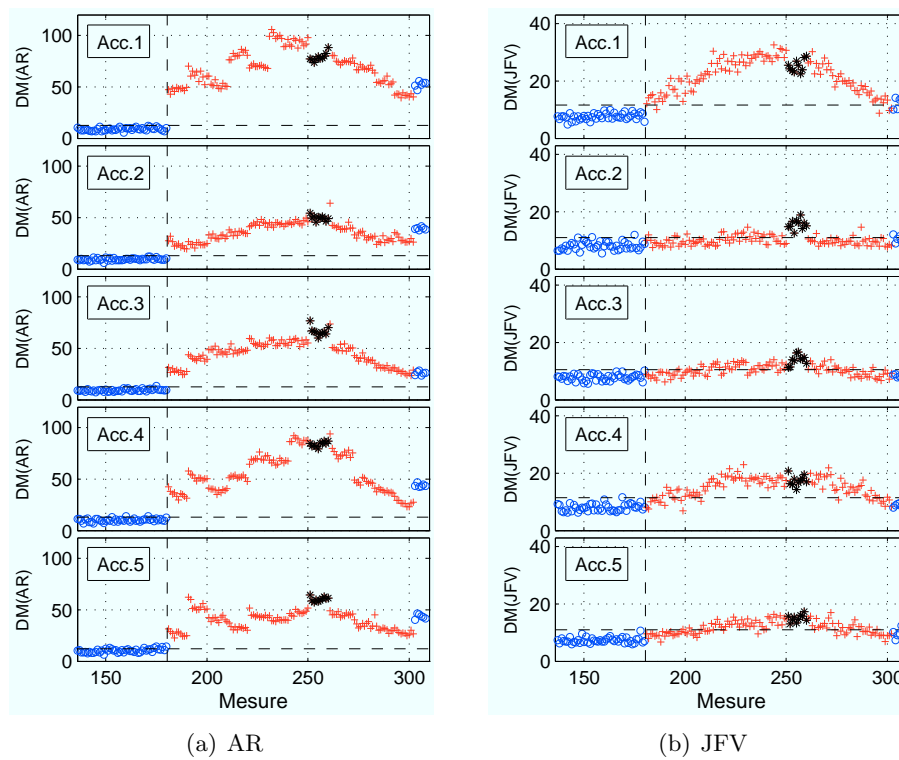


FIG. 5.13 Distance de Mahalanobis en fonction de la mesure, *essai 24*, endommagement D_{24} ; \circ base de test; $+$ structure endommagée; $*$ endommagement maximum.

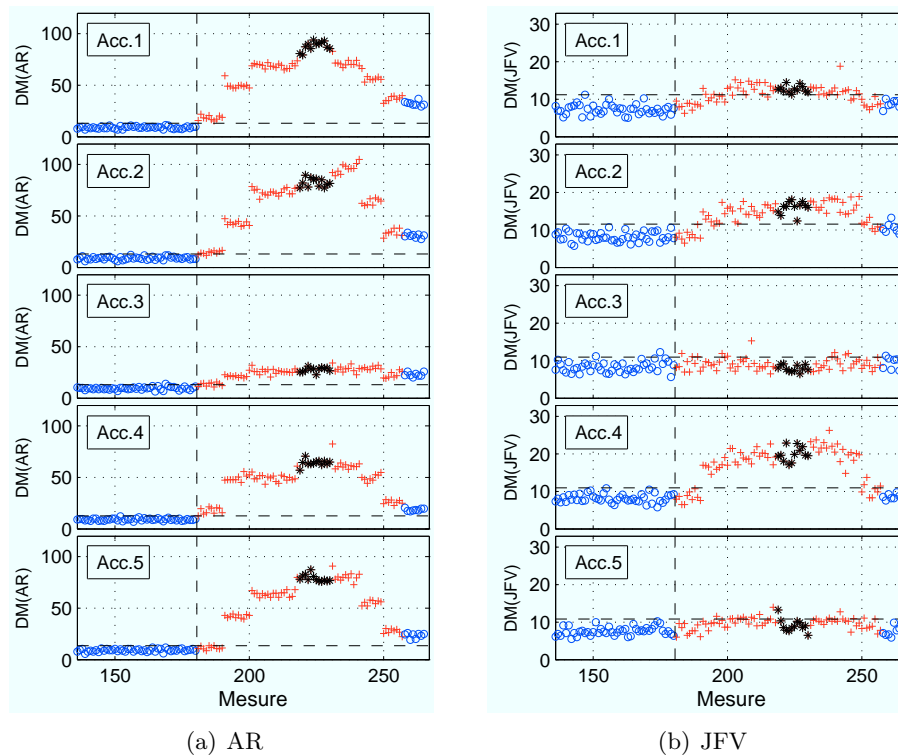


FIG. 5.14 Distance de Mahalanobis en fonction de la mesure, *essai 25*, endommagement A_1 ; \circ base de test; $+$ structure endommagée; $*$ endommagement maximum.

De façon quasi-systématique, dans les cas positifs de détection, la distance de Mahalanobis tend à varier dans le même sens que l'endommagement : augmentation lorsque l'endommagement est initié et évolue, maximum lorsque l'endommagement est au plus haut, diminution lorsque le retour à l'état initial est réalisé par suppression progressive de l'endommagement. Ce comportement tend à montrer qu'une information quantitative sur le degré de dégradation est présente dans les indicateurs, autorisant la possibilité de sa quantification relative. Contrairement au JFV, les coefficients AR détectent le premier niveau d'endommagement dans 13 cas sur 15, soulignant ainsi une très grande sensibilité. En effet, pour les essais 22 et 25, le desserrage d'un seul écrou engendre une modification légère de l'assemblage, comparativement au niveau d'endommagement suivant, qui désolidarise l'équerre du plateau. En revanche, bien que la distance de Mahalanobis décroisse significativement après être passée par son maximum, la diminution n'est en général pas suffisante pour franchir dans l'autre sens le seuil de détection. Ainsi, les dernières mesures de chaque essai, qui correspondent au retour à l'état initial de la maquette, sont systématiquement détectées comme nouvelles. Malgré les précautions prises durant les essais, il est probable que le serrage final des assemblages ne soit pas parfaitement fidèle au serrage initial. La très grande sensibilité des AR risque par ailleurs d'engendrer un nombre important de fausses alarmes lorsque des variations environnementales sont introduites. Dans la mesure où il est presque impossible d'intégrer *la totalité* des conditions normales de fonctionnement dans une base d'apprentissage, l'outil doit pouvoir se généraliser à des condi-

tions légèrement différentes de celles déjà rencontrées. Dans ce cas, le verrou scientifique majeur porte sur la définition d'un compromis entre sensibilité à l'endommagement et capacité de généralisation.

Dans les cas où la détection est précise, le JFV place systématiquement (sauf pour l'essai 24 Fig.5.13(b)) le premier niveau d'endommagement et le retour à l'état initial en dessous du seuil de détection. Cet indicateur présente donc une sensibilité moins importante, mais manifeste une meilleure capacité de généralisation. Il semble donc plus adapté aux cas pour lesquels les fausses alarmes sont très pénalisantes.

Localisation de l'endommagement

La question de la localisation de l'endommagement est également récurrente en SHM. Les résultats produits par les coefficients AR lors de l'essai 22 (Fig.5.12(a)) laissent penser que la localisation est possible, en observant le capteur présentant la plus grande distance de Mahalanobis. Sur cet essai, il apparaît en effet que la distance de Mahalanobis maximale est observée pour les données fournies par le capteur le plus proche de l'endommagement. Cependant, cette hypothèse est infirmée par l'essai 25 (Fig.5.14(a)). Ce dernier présente une distance maximale pour les données de l'accéléromètre 2 alors que l'endommagement se situe au niveau de l'assemblage 1. L'essai 24, qui consiste à réduire la tension dans la diagonale reliant les nœuds 2 et 4, engendre un changement affectant assez globalement le comportement dynamique de la maquette. Or, dans le cas du JFV, le capteur 1 indique un changement beaucoup plus significatif que les autres accéléromètres (Fig.5.13(b)). En conclusion, lors de ces essais, aucune piste n'est clairement révélée pour remédier au problème de la localisation par les indicateurs vectoriels testés.

Ce manque d'information relative à la localisation peut éventuellement être mis en relation avec une dimension trop faible des indicateurs vectoriels. En effet, dans la phase de détermination des paramètres, le choix de la dimension des indicateurs (ordre du modèle et dimension de reconstruction) est dicté par des contraintes pratiques liées aux temps de calcul et à la quantité de données disponibles, et non par des critères objectifs maximisant l'information véhiculée par les indicateurs.

Test d'indicateurs de dimension supérieure

La figure 5.15 présente l'aptitude à la détection d'endommagement du JFV sur des reconstructions en dimensions 3 et 9 (essai 22). Pour ces valeurs de dimension, la dispersion des points est accrue. Dans le cas de l'accéléromètre 4, la détection est même dégradée. Le choix de la dimension 6 est donc conforté et conservé pour la partie suivante.

Un test similaire est réalisé sur l'impact de l'ordre du modèle AR. Le test est réalisé en ajustant un AR(15) et un AR(70) sur les mesures de l'essai 22 (Fig.5.16). L'augmentation de l'ordre amplifie les écarts. L'effet est significatif sur les accéléromètres 3 et 4 qui caractérisent mal l'endommagement avec un AR(15). Cependant, en dépit d'un ordre de 15, les performances obtenues sont proches de celles du JFV avec une dimension de 6. La tendance consisterait donc d'avantage à diminuer l'ordre du modèle, qu'à l'augmenter.

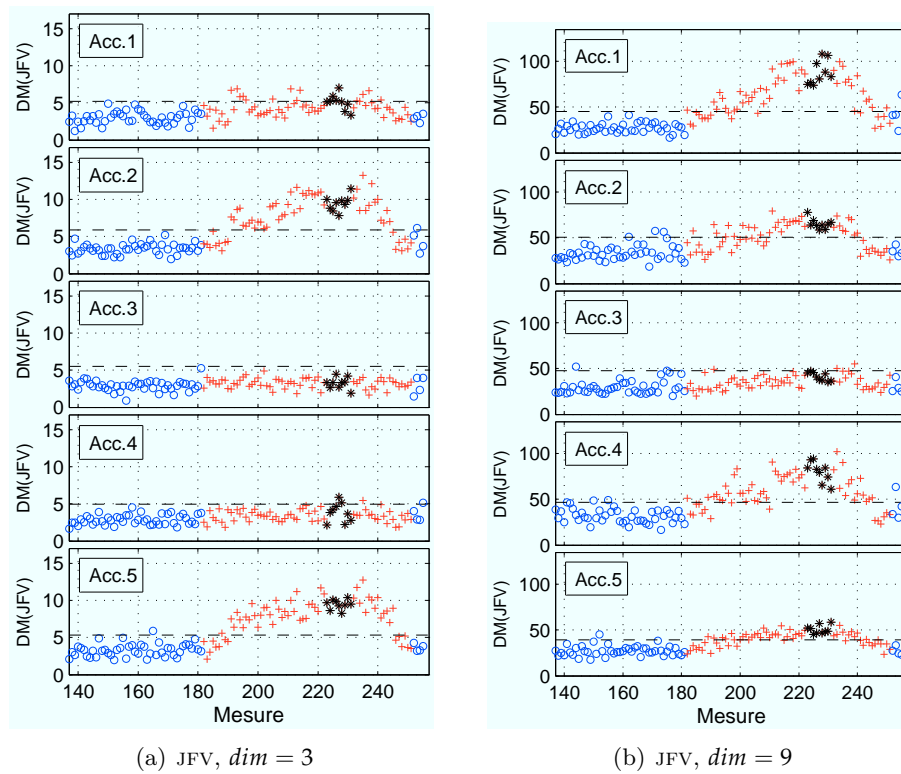


FIG. 5.15 Distance de Mahalanobis calculée sur le JFV en fonction de la mesure pour des reconstructions en dimensions 3 et 9, *essai 22*, endommagement A_2 ; \circ base de test; $+$ structure endommagée; $*$ endommagement maximum.

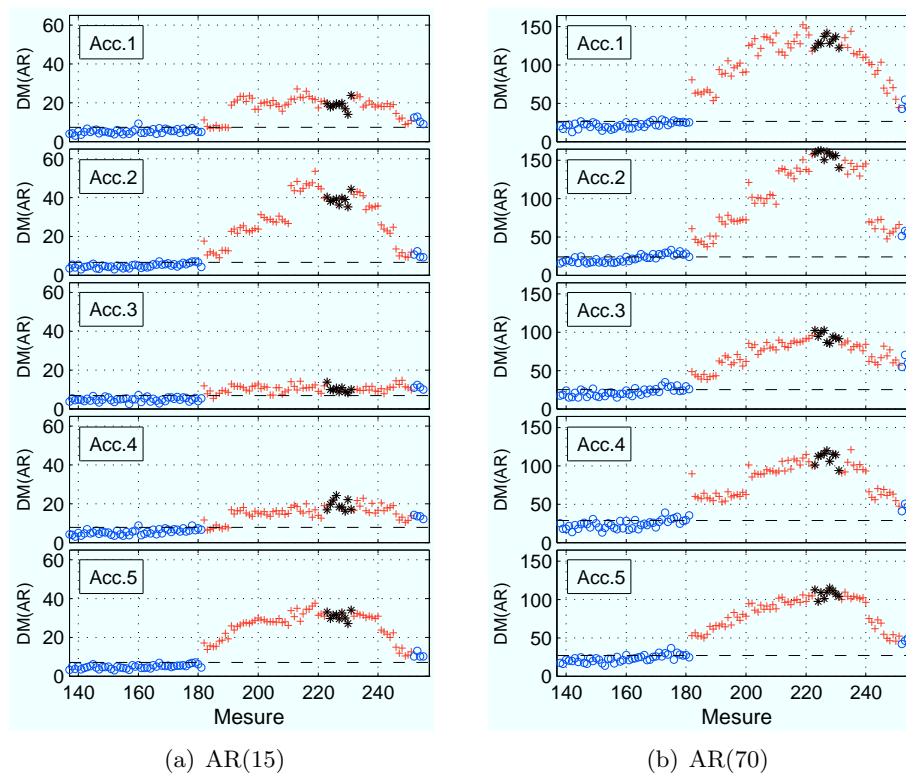


FIG. 5.16 Distance de Mahalanobis en fonction de la mesure pour un AR(15) et un AR(70), *essai 22*, endommagement A_2 ; \circ base de test; $+$ structure endommagée; $*$ endommagement maximum.

5.2.4 Conclusions

Les coefficients AR et le JFV constituent deux indicateurs en mesure de détecter l'endommagement et de suivre son évolution de façon relative. Pour chaque essai, plusieurs capteurs sont sensibles à l'apparition de la dégradation, sans pouvoir cependant les corrélérer avec sa localisation. Le niveau 2 des objectifs de *Rytter* ne semble donc pas accessible en l'état.

Dans la plupart des cas, les coefficients AR présentent une plus forte sensibilité aux changements de comportement dynamique que le JFV. Ce résultat est lié à la plus faible dispersion de chaque coefficient, qui limite cependant la capacité de généralisation hors de la base de référence. Le JFV semble donc plus adapté à des applications pour lesquelles la tolérance aux fausses alarmes est faible.

Enfin, augmenter la dimension des indicateurs ne semble pas améliorer leurs performances. Pour le JFV, la dimension 6 constitue même un optimum sur l'*essai 22*. Par ailleurs, pour les modèles AR, une forte réduction de l'ordre (15 au lieu de 40) permet tout de même de détecter l'endommagement.

BILAN

La comparaison entre les Exposants de *Lyapunov* et le JFV montre que seul ce dernier est en mesure de détecter correctement l'endommagement sur la maquette métallique, alors que les EL sont légèrement plus performants sur les données simulées sur la base du modèle masse/ressort amorti. Par ailleurs, le faible coût associé au temps de calcul du JFV constitue un argument supplémentaire en sa faveur au regard du spectre de Lyapunov.

Le calcul de ces indicateurs requiert des algorithmes spécifiques qui doivent être préalablement paramétrés de façon optimale. Les données étant issues de systèmes plus stochastiques que déterministes, les méthodes classiques de détermination de ces paramètres fournissent des plages indicatives de paramètres à défaut de valeurs déterminées. En définitive, le critère prépondérant étant la quantité de mesures disponibles et leur longueur, il apparaît que les paramètres optimaux pour le calcul du JFV sont : $T_2 = 2$ et $L = 2$. Du fait de la méthode de construction des vecteurs de l'espace des phases, cette valeur interdit les valeurs 2 et 4 pour le délai de reconstruction (τ).

Sur la base de ces recommandations, l'utilisation du JFV sur la maquette en bois, sans variations environnementales, permet de détecter correctement l'apparition de l'endommagement et d'en suivre la progression. La confrontation de ses performances avec celles des coefficients d'un modèle AR révèle une plus grande sensibilité des AR, mais une meilleure aptitude à la généralisation du JFV. Cette dernière caractéristique pourrait être exploitée en fonction de la tolérance aux fausses alarmes de la structure sous surveillance.

SENSIBILITÉ AUX VARIATIONS ENVIRONNEMENTALES

6

SOMMAIRE	
INTRODUCTION	118
6.1 DONNÉES ÉTUDIÉES	119
6.2 PARAMÈTRES DE CALCUL	119
6.2.1 Maquette en bois	119
6.2.2 Poutre en béton fibré	120
6.3 OBSERVATION DES INDICATEURS	122
6.3.1 Effet des variations environnementales	122
6.3.2 Répartition dans l'espace des indicateurs vectoriels	123
6.4 DIVISION DE LA BASE DE DONNÉES COLLECTÉES SUR STRUC- TURE SAINTE	125
6.5 DÉTECTION DE L'ENDOMMAGEMENT	128
6.5.1 Poutre en béton fibré	130
6.5.2 Maquette en bois	131
BILAN	135

INTRODUCTION

LE chapitre 5 montre que les coefficients AR sont très sensibles à tout changement du comportement de la structure. Cette extrême sensibilité engendre la classification de signaux comme nouveaux, même s'ils ne diffèrent que légèrement de ceux de la base de référence. Or, dans un contexte de suivi des structures réelles, les variations environnementales ne se reproduisent pas exactement d'une année sur l'autre. Pour éviter un trop grand nombre de fausses alarmes, l'étape de normalisation doit donc offrir un certain niveau de souplesse, favorisant l'aptitude à la généralisation.

L'objet de ce chapitre est d'éprouver l'approche de détection de nouveauté, basée sur le JFV et les coefficients AR, sur des données influencées par des variations environnementales réelles. Les mesures analysées ici proviennent des tests réalisés sur la poutre en béton fibré (PB) et sur la maquette en bois (MB).

Après avoir présenté le résultat du processus de choix des paramètres, la variabilité des indicateurs en fonction des modifications climatiques est analysée. L'introduction des variations environnementales dans les essais donne lieu à une discussion sur un problème connu, mais souvent évité dans la littérature, portant sur la répartition des données acquises sur structure saine dans les différentes bases nécessaires à la mise en œuvre de la détection de nouveauté : base d'apprentissage, base de classification et base de test de la performance. Enfin, la capacité de l'approche proposée à discriminer l'endommagement en présence de variations environnementales est présentée.

6.1 DONNÉES ÉTUDIÉES

Les mesures analysées sont issues des quatre accéléromètres placés sur la poutre en béton fibré (PB) et des 5 accéléromètres installés sur la maquette en bois (MB). Ces deux structures sont soumises à des cycles de chauffage et de refroidissement (§4). Le tableau 6.1 regroupe les principales informations utiles sur les essais analysés.

La comparaison des résultats obtenus sur les deux structures n'est cependant pas directement envisageable. En effet, lors de l'essai sur la poutre, la source d'excitation (bruit blanc en entrée du pot vibrant) est identique pour chaque mesure. Cette régularité de l'excitation biaise considérablement une approche de détection basée sur la comparaison de signaux par rapport à une base de référence. Cet écueil lié à une utilisation des fonctions élémentaires du système de mesure et de contrôle du pot vibrant (système *PULSE*[®] de *Bruel & Kjaer*), est corrigé sur les essais réalisés sur la maquette en bois, pour lesquels la source d'excitation stochastique est réinitialisée à chaque mesure. Ainsi, les analyses conduites dans ce chapitre insistent davantage sur les résultats obtenus sur la maquette en bois. Des observations intéressantes sont cependant formulées à partir de l'essai sur la poutre.

D'autres scénarios d'endommagement ont été réalisés sur la MB. Cependant, en raison de l'incapacité de l'approche proposée à le localiser, leur étude apporte des conclusions similaires à celles obtenues avec les essais présentés dans cette partie, essais jugés représentatifs.

Chaque mesure est centrée avant d'être analysée.

6.2 PARAMÈTRES DE CALCUL

6.2.1 Maquette en bois

Concernant la maquette de structure en bois, la structure et l'excitation sont identiques au cas traité dans la section 5.2. Par conséquent, les mêmes paramètres sont utilisés :

- $AR(40)$
- $dim = 6$
- $\tau = 6$
- $T_2 = 2$ et $L = 2$.

Essai	Struct.	N_b	N_e	Endom.	Fin Endom.	f_e
MB34		400	122	D_{24}	231	
MB35	Maquette bois	425	85	A_2	260	8192 Hz
MB36		425	80	A_1	230	
PB	Poutre	210	210	Fissuration		2048 Hz

TAB. 6.1 *Récapitulatif des mesures analysées, avec de gauche à droite, le numéro de l'essai, le nombre de mesures sur structure saine N_b , le nombre de mesures sur structure endommagée N_e , la localisation de l'endommagement, le numéro de la mesure marquant le début du retour aux conditions initiales et la fréquence d'échantillonnage f_e .*

Accéléromètre	Moyenne	Min-Max	Choix
1	53	50-65	
2	54	50-60	
3	52	50-90	50
4	42	40-43	

TAB. 6.2 Sélection de l'ordre du modèle AR sur la poutre (sur 5 séries par accéléromètre).

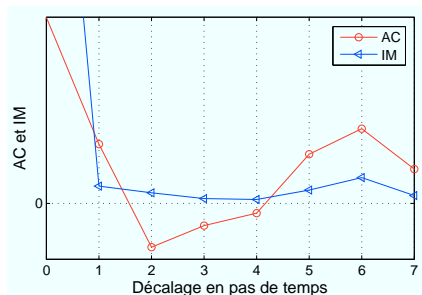


FIG. 6.1 Fonctions d'information mutuelle et d'auto-corrélation extraites d'une mesure de l'accéléromètre 3 enregistrée sur PB.

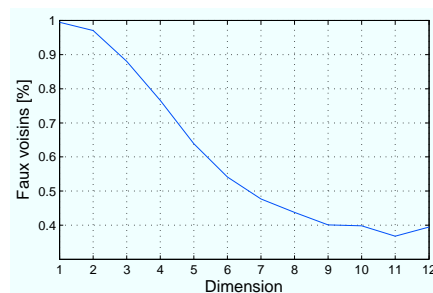


FIG. 6.2 Taux de faux voisins en fonction de la dimension de reconstruction pour une mesure de l'accéléromètre 3 enregistrée sur PB.

6.2.2 Poutre en béton fibré

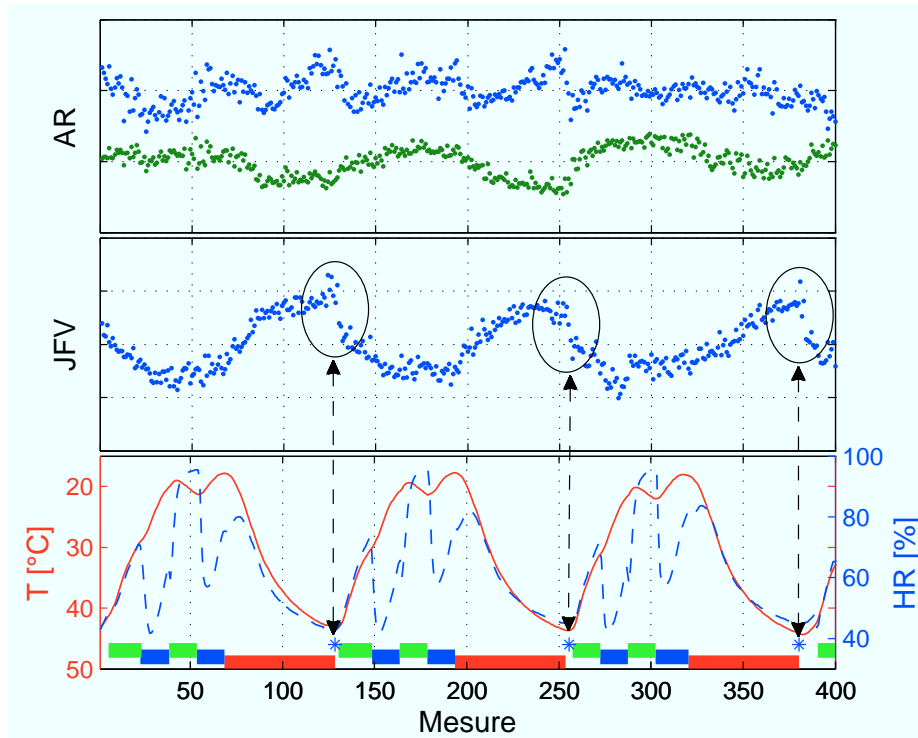
Coefficients AR

La fonction d'auto-corrélation partielle et le critère d'Akaike sont analysés de la même façon que dans la section 5.2. Les calculs sont réalisés sur 5 séries choisies au hasard pour chaque accéléromètre (soit un total de 20 séries). Les résultats indiquent un **ordre de 50** (Tab.6.2). Bien que cette valeur ne reflète pas précisément l'ordre réel du processus (probablement supérieur), le chapitre précédent montre qu'il est néanmoins suffisant pour bien caractériser le comportement de la structure.

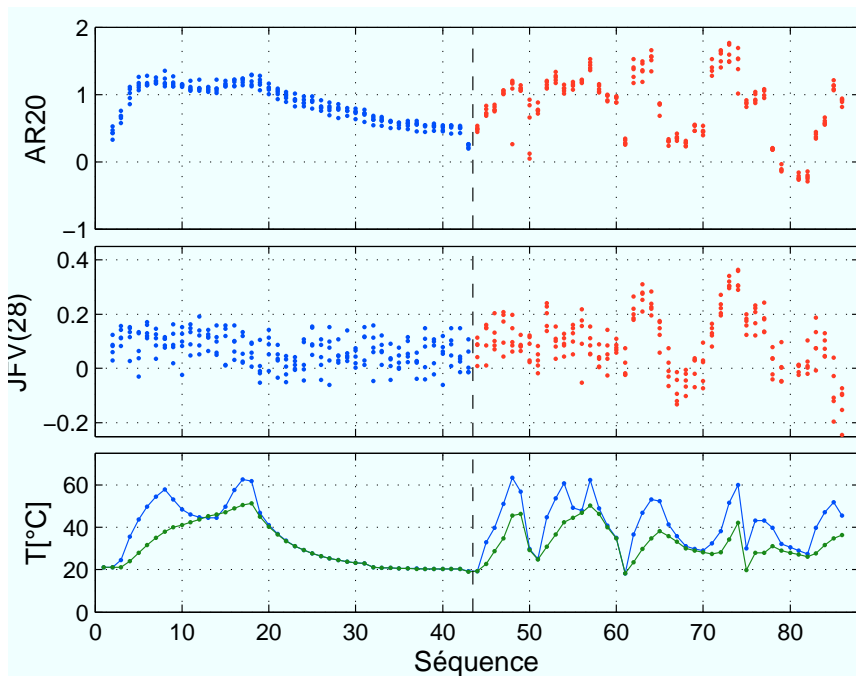
Paramètres de reconstruction

La figure 6.1 représente un exemple des fonctions d'information mutuelle et d'auto-corrélation estimées sur l'accéléromètre 3. Contrairement à celles obtenues avec la MB, les courbes de l'information mutuelle et de l'auto-corrélation n'indiquent pas les mêmes optimums. L'échantillonnage est probablement insuffisant pour dessiner correctement ces courbes. La fonction d'auto-corrélation indique une valeur de τ égale à 1 ou 2, et l'information mutuelle à 3 ou 4. Le choix de $\tau = 3$ est cependant retenu en raison de l'incompatibilité des valeurs 2 et 4 avec l'algorithme de calcul du JFV.

Le taux de faux voisins en fonction de la dimension de reconstruction est représenté dans la figure 6.2 pour l'accéléromètre 3. On note un ralentissement de sa diminution à partir d'une dimension de reconstruction de 6, et une stabilisation à partir de la dimension 9. Toujours dans le souci de limiter la dimension des indicateurs vectoriels et de permettre la comparaison avec les coefficients AR (en terme de nombre de composantes), la valeur 7 est retenue comme dimension de reconstruction.



(a) AR : CP8 (en haut) et CP2 (en bas); JFV : CP1 (MB, essai 34, Acc.1); — Température; - - - Humidité relative; * humidification; ■ ventilation; ■ climatisation; ■ chauffage.



(b) AR : coefficient 20; JFV : composante 28 (Poutre, Acc.3)

FIG. 6.3 Variations des composantes principales des coefficients AR et des composantes du JFV en regard des variations des facteurs environnementaux.

6.3 OBSERVATION DES INDICATEURS

6.3.1 Effet des variations environnementales

Lors des deux essais (PB et MB), il apparaît que la température affecte significativement les composantes des indicateurs vectoriels.

De plus, dans le cas de la maquette en bois, les variations de température se conjuguent aux fluctuations de l'humidité de l'air. Une analyse en composantes principales peut être réalisée afin d'appréhender l'effet des fluctuations environnementales sur l'ensemble des composantes des indicateurs (coefficients AR et JFV). Toutes les composantes ne sont pas affectées de la même façon. Certaines ne réagissent pas face aux fluctuations extérieures, tandis que d'autres subissent des changements qu'il est facile de corrélérer aux variations environnementales.

La figure 6.3(a) présente, selon la chronologie des 400 mesures constituant la base de référence de l'essai 34 réalisé sur la MB :

- les composantes principales 2 et 8 (*CP2* et *CP8*) calculées sur la base des coefficients AR,
- la première composante principale (*CP1*) calculée sur la base de données du JFV,
- l'évolution de la température et de l'humidité relative dans l'enceinte au cours de la séquence de mesure.

L'observation conjointe de ces graphiques montre, par exemple, que la *CP8* affiche des oscillations plus rapides que les cycles de température, elle semble ainsi plus corrélée aux variations de l'humidité. En revanche, la *CP2* montre une dépendance forte aux changements de température. Le même comportement est observé sur la *CP1* du JFV.

Dans le cas de la poutre, le 20^{ème} coefficient AR, la composante 28 du JFV et la température mesurée par les deux thermocouples sont représentés pour chaque séquence de mesure sur les 3 graphiques de la figure 6.3(b). La composante 28 du JFV apparaît peu sensible aux effets de la température, jusqu'à la séquence 61, qui marque le début d'importantes variations (trait pointillé vertical). Cette séquence correspond à l'initiation de la macro-fissuration de la poutre qui semble amplifier les effets des variations thermiques. Concernant l'indicateur AR, le 20^{ème} coefficient présente une corrélation forte avec les changements de température.

Sur la base de ces constats, il apparaît envisageable de normaliser les indicateurs par le biais d'une régression entre les variables environnementales et les composantes sur la base de données de référence. Pour observer les corrélations existantes, la figure 6.4 représente certaines composantes principales des indicateurs vectoriels en fonction de la température ou de l'humidité relative. Même si la corrélation semble visuellement significative, la figure 6.4(a) montre qu'aucune régression ne peut être créée entre la *CP2* des coefficients AR et la température sans que l'erreur du modèle ne soit très importante. La température n'est pas la seule variable à prendre en compte. Dans la figure 6.3(a), des discontinuités sont visibles sur le profil de la *CP1* du JFV. Elles correspondent aux phases d'humidification par pulvérisation de la maquette (notée par des *). En considérant également que seul un

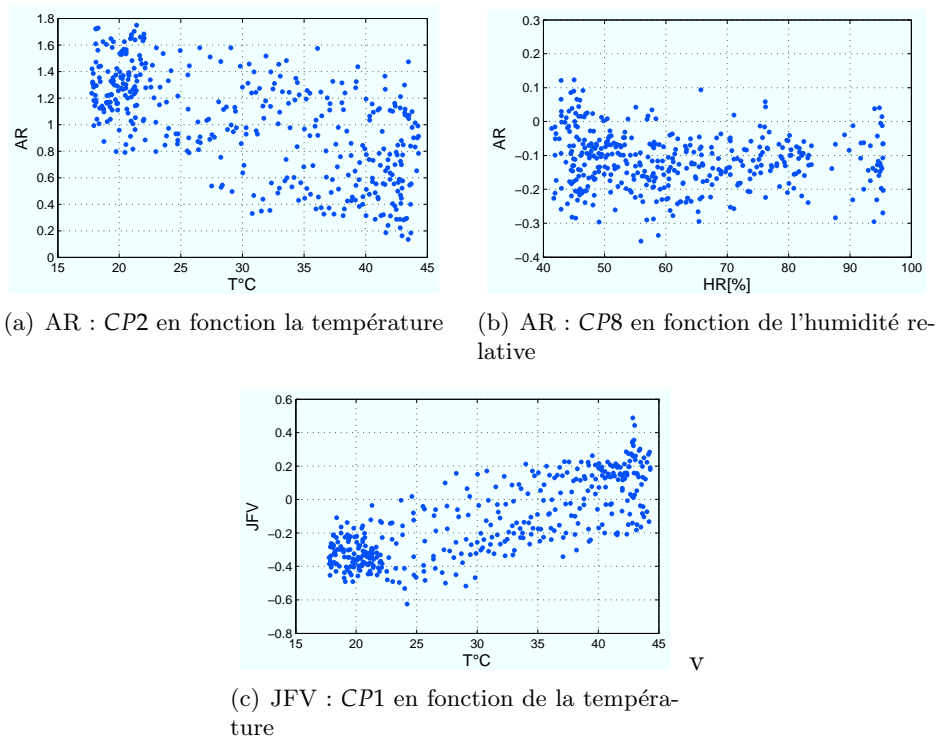


FIG. 6.4 *Corrélation des composantes principales des coefficients AR et des composantes du JFV avec les variables environnementales (MB, essai 34, ACC.1).*

nombre restreint de composantes affiche des variations aussi identifiables, l'approche régressive n'apparaît pas adaptée.

6.3.2 Répartition dans l'espace des indicateurs vectoriels

Avant de chercher directement à détecter l'endommagement induit dans les deux structures étudiées, il est important d'observer comment se distribuent les composantes de chaque indicateur dans leur espace. La question centrale est de savoir si les points relatifs aux mesures sur structure saine et ceux correspondant aux mesures sur structure endommagée occupent les mêmes zones de l'espace de l'indicateur vectoriel. Si tel est le cas, la discrimination des données collectées sur structure endommagée peut être difficile, voire impossible.

Les gammes de variations de chaque composante des coefficients AR et du JFV avant et après endommagement, sont représentées respectivement sur les figures 6.5(a) et 6.5(b). Les points noirs indiquent les mesures réalisées sur structure saine et les croix rouges celles acquises sur structure endommagée. En l'absence de variations environnementales, le chapitre précédent montre que ces domaines sont quasiment disjoints pour de nombreuses composantes (plus particulièrement pour les coefficients AR, Fig.5.10), facilitant la détection de l'endommagement.

En présence de variations environnementales, le chevauchement des plages de variations des composantes est beaucoup plus important. Seuls quelques points figuratifs de la structure endommagée se distinguent, correspondant probablement au niveau maximum de dégradation. Pour la plupart

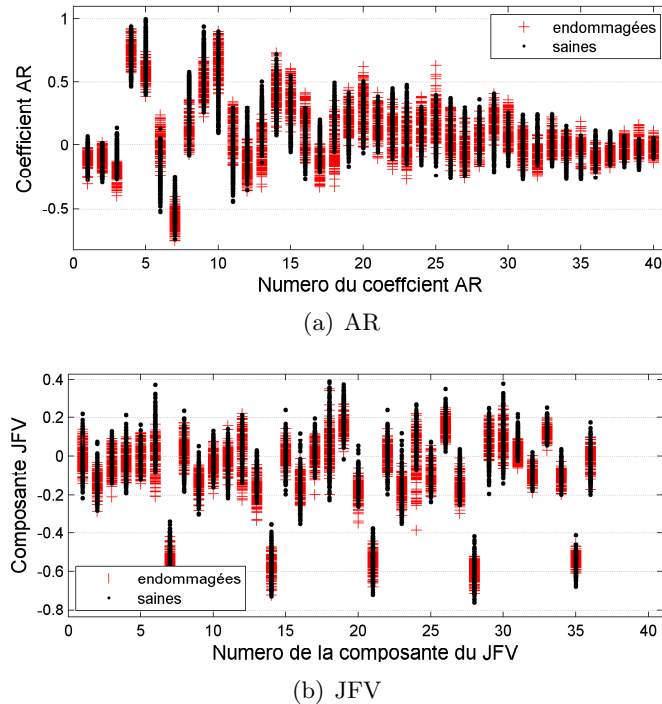


FIG. 6.5 Gamme de variations de chaque composante des 2 indicateurs vectoriels (MB, essai 34, Acc.1).

des composantes, la plage de variation n'est pas modifiée. Aucune différence n'est constatée entre les premiers coefficients AR et les derniers.

Pour avoir une idée de leur répartition dans l'espace de l'indicateur vectoriel et des corrélations existantes entre les composantes, la figure 6.6 trace les projections de l'ensemble des indicateurs dans les plans relatifs aux 5 premiers coefficients AR. Les points associés à l'endommagement sont entourés de blanc et de couleur d'autant plus sombre que la dégradation est importante. Les points associés à la structure saine sont de couleur gris clair et cerclés de noir.

Il est clair que les premiers niveaux d'endommagement sont confondus avec les points sains sur toutes les figures. Quand l'endommagement augmente, les points se situent plutôt au bord de l'ensemble (sauf AR3/4). Cependant, dans aucun des cas, une séparation n'est possible entre les deux classes, sans faire un nombre important de mauvaises classifications.

La forme des nuages de points n'est pas réellement complexe. Ils sont globalement constitués d'un nuage central s'étalant différemment dans chaque direction. Pour justifier de l'utilisation d'outils de classification différents de la distance de Mahalanobis, les points doivent présenter une répartition plus irrégulière. Par exemple, dans le cas d'une distribution à plusieurs centres, des outils d'estimation de densité semi ou non-paramétriques sont plus adaptés (*Estimateur de densité de Parzen*, ou *mélanges de Gaussiennes*). Si le nuage présente une forme bien définie, mais irrégulière (forme en *banane*), indiquant une corrélation non-linéaire, les méthodes à noyaux fournissent de bons résultats.

Le recours à des outils de classification non-linéaire n'améliorerait la classification qu'à la condition de procéder à un sur-apprentissage. La géné-

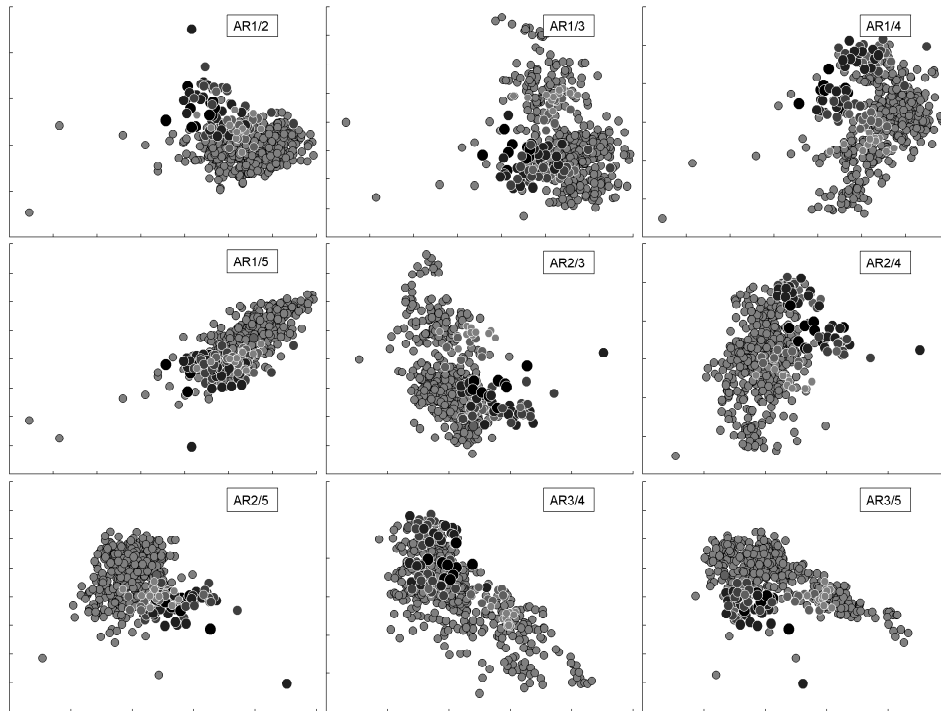


FIG. 6.6 *Projections des indicateurs vectoriels mesurés dans les plans relatifs aux 5 premiers coefficients AR ; Les points de la base de référence (structure saine) sont de couleur gris clair et cerclés de noir, les points associés à l'endommagement sont entourés de blanc et de couleur grise, d'autant plus sombre que l'endommagement augmente (MB, essai 35, Acc.1).*

realisation serait alors médiocre, augmentant le nombre de fausses alarmes. Ainsi, malgré sa simplicité, la distance de Mahalanobis apparaît bien adaptée aux cas d'études expérimentaux traités dans cette partie.

6.4 DIVISION DE LA BASE DE DONNÉES COLLECTÉES SUR STRUCTURE SAINE

La division de la base de données relative à la structure saine est rapidement évoquée dans la partie 5.1.3. Une discussion plus complète est proposée ci-après.

Pour mettre en œuvre la démarche de détection de nouveauté, la base de données de référence doit être divisée en 3 parties.

La base d'apprentissage (BA)

La base d'apprentissage constitue le support de normalisation des indicateurs. Elle est utilisée ici comme référence pour le calcul de la distance de Mahalanobis, c'est-à-dire pour l'estimation de la matrice de covariance et du vecteur moyen. S'agissant de la référence, il est essentiel qu'elle soit composée de mesures acquises sur l'ensemble des conditions environnementales et opérationnelles que rencontre la structure (conditions normales). Pour toutes les applications présentées dans ce mémoire, elle est formée par 50% des mesures collectées sur structure saine.

La base de classification (BC)

La base de classification est utilisée pour la détermination du seuil de classification. Chacun de ses individus est normalisé par rapport à la BA, puis, en fonction de l'erreur de Type I tolérée, elle est exploitée pour ajuster le modèle statistique fixant le seuil de classification. La BC est composée ici par 25% des données acquises sur structure saine.

La base de test (BT)

Le reste des données relatives à la structure saine (25%) constitue la base de test, qui permet de vérifier si le seuil de classification est correctement fixé, au regard des fausses alarmes.

Dans la plupart des exemples de la littérature, la répartition des données dans ces bases est réalisée de façon aléatoire. Cette pratique ne pose pas réellement de problème lorsque les variations environnementales subies par la structure sont faibles (§5).

Dans d'autres cas, les variations climatiques ou opérationnelles sont simulées par des ajouts de masses locales, ou par le remplacement d'une barre par une autre d'épaisseur différente, modifiant ainsi la rigidité. Ces changements des conditions de la structures sont discrets. Les données mesurées sur structure saine dans ces nouvelles conditions sont ajoutées à la base d'apprentissage. Celle-ci intègre alors de façon satisfaisante les différents états normaux de fonctionnement de la structure, et l'endommagement est détecté sans difficulté.

Il convient cependant de s'interroger sur les conditions pratiques de mise en place d'un système de suivi sur un ouvrage réel. Le comportement de la structure est enregistré sur une période d'une année au minimum pour constituer la base de données relative à l'état sain de la structure. A la suite de cette étape, le suivi peut effectivement commencer. Les conditions environnementales et opérationnelles des années suivantes sont probablement sensiblement différentes de celles rencontrées durant la période de référence. L'enjeu consiste alors à être en mesure de détecter un endommagement dans ces conditions environnementales et opérationnelles sensiblement différentes. Dans ce cas, la répartition des données collectées pendant la première année de monitoring dans les 3 bases BA, BC et BT est susceptible d'impacter la performance de l'outil de suivi.

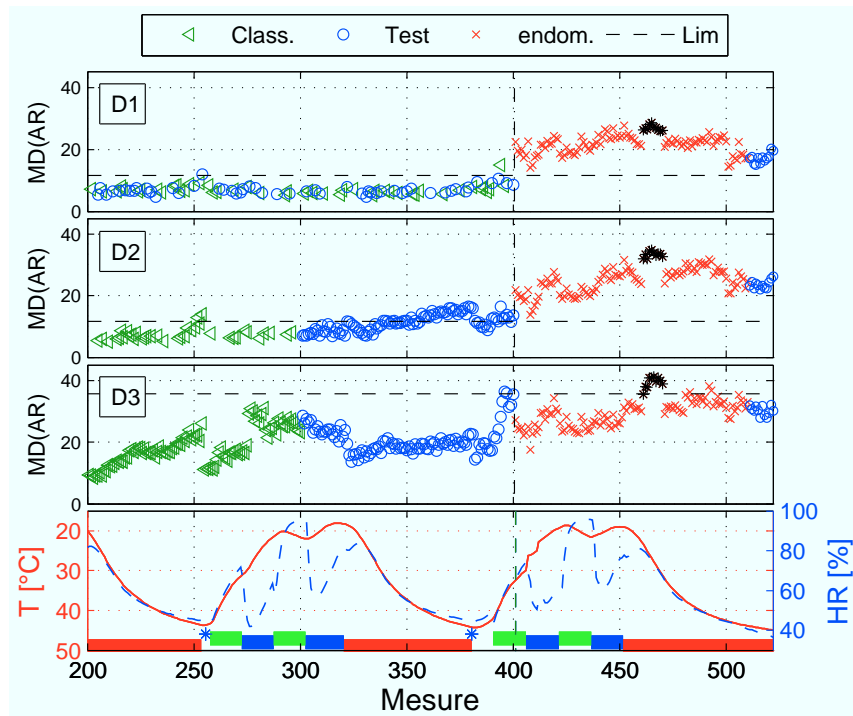
Trois divisions différentes de la base de données associée à la structure saine sont comparées pour illustrer le biais induit par une répartition aléatoire, par rapport au problème réel. Pour 100 mesures enregistrées sur structure saine, les 3 divisions sont :

D1 : les mesures sont réparties de façon aléatoire dans les 3 bases.

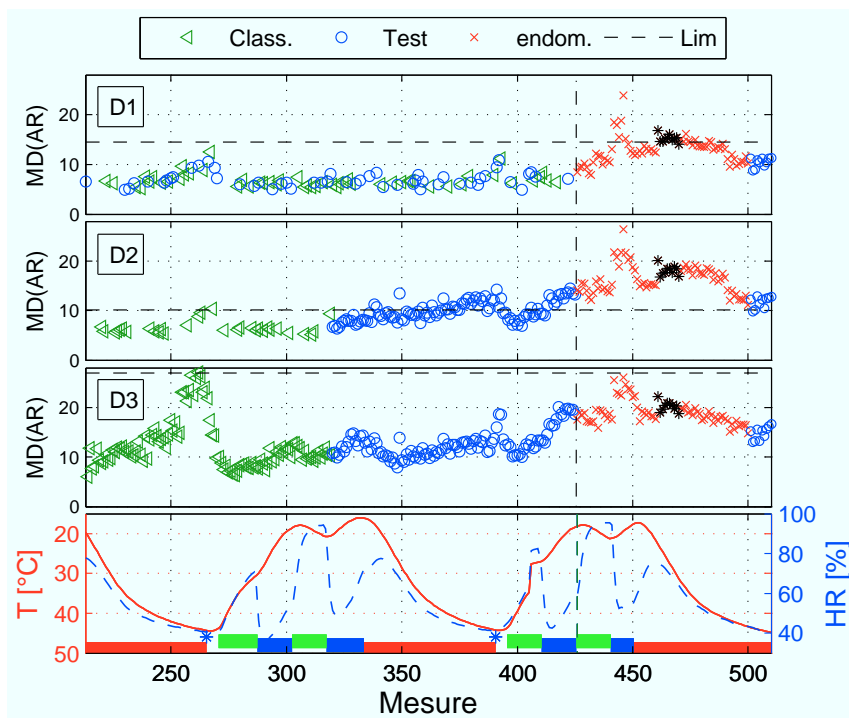
D2 : la base de test est formée par les 25 dernières mesures, et les 75 autres sont réparties aléatoirement entre BA et BC.

D3 : la chronologie des enregistrements est conservée (les 50 premières mesures forment la BA, les 25 suivantes la BC et les 25 dernières la BT).

Les figures 6.7(a) et 6.7(b) présentent les distances de Mahalanobis de chaque mesure (dans l'ordre chronologique) obtenues selon les trois distributions pour les essais 34 (Acc.4) et 35 (Acc.1), respectivement. Afin d'identifier clairement les mesures composant les différentes bases, les mesures



(a) MB, essai 34, Acc.4



(b) MB, essai 35, Acc.1

FIG. 6.7 Détection de l'endommagement selon trois répartitions différentes des données associées à la structure saine ; de haut en bas : D1, D2 et D3 ; — Température ; - - - Humidité relative ; * humidification ; ■ ventilation ; ■ climatisation ; ■ chauffage.

affectées à la BT sont représentées par des cercles, et celles affectées à la base de classification par des triangles. Les variations environnementales sont également introduites dans la figure.

En utilisant la distribution D1, les données de la base de test et de classification sont mélangées à la base d'apprentissage. Il en résulte une distance généralement plus faible que celle obtenue sur les nouvelles mesures. Les points endommagés présentant une distance plus importante apparaissent alors naturellement comme *nouveaux* et les variations environnementales sont correctement filtrées. Pour réellement tester la présence de fausses alarmes, il est préférable de séparer la base de test de celle d'apprentissage afin de la placer dans la même configuration que les nouvelles mesures. Cela correspond à la distribution D2. Dans ce cas, les points endommagés sont toujours correctement classés. Cependant des fluctuations liées aux variations environnementales apparaissent. Une augmentation progressive des distances est également observée sur la base de test conduisant à de nombreuses fausses alarmes. Cette augmentation correspond à l'éloignement temporel progressif des bases d'apprentissage et de test. Pour réduire le nombre de fausses alarmes, la base de classification doit elle aussi être séparée de la base d'apprentissage conduisant à la distribution D3. Dans ces conditions, les fluctuations sont très marquées et empêchent de discerner les variations liées à l'endommagement de celles provoquées par l'effet de la température.

De façon générale, plus la mesure s'éloigne dans le temps de la base d'apprentissage, plus la distance de Mahalanobis tend à augmenter même en l'absence d'endommagement et à passer au-dessus de la limite de classification. Ce comportement résulte de l'impossibilité de reproduire à l'identique des conditions environnementales passées. Cet effet pernicieux est naturellement masqué dans les essais lorsque les données sont affectées de façon totalement aléatoire dans les 3 bases. En pratique, l'effet existe systématiquement et il est nécessaire d'y remédier, car il est alors difficile de faire la part de choses entre un accroissement de la distance de Mahalanobis dû aux effets environnementaux et un accroissement dû à l'endommagement.

Dans un souci de rigueur, la validation d'un algorithme de détection de nouveauté en vue de son utilisation sur un ouvrage en service, impose que les bases de test, de classification et les nouvelles mesures testées possèdent la même *proximité temporelle* avec la base d'apprentissage.

En marge de cette discussion, il faut également noter que les distances successives sont désormais fortement corrélées. La limite ne peut dans ce cas pas être fixée par l'un des modèles statistiques présentés, puisque les distances ne sont plus indépendantes.

6.5 DÉTECTION DE L'ENDOMMAGEMENT

En tenant compte des remarques formulées dans la partie précédente, la répartition des données dans les bases d'apprentissage, de classification et de test est réalisée selon la distribution D3. La distance de Mahalanobis est calculée sur le JFV et sur les coefficients AR.

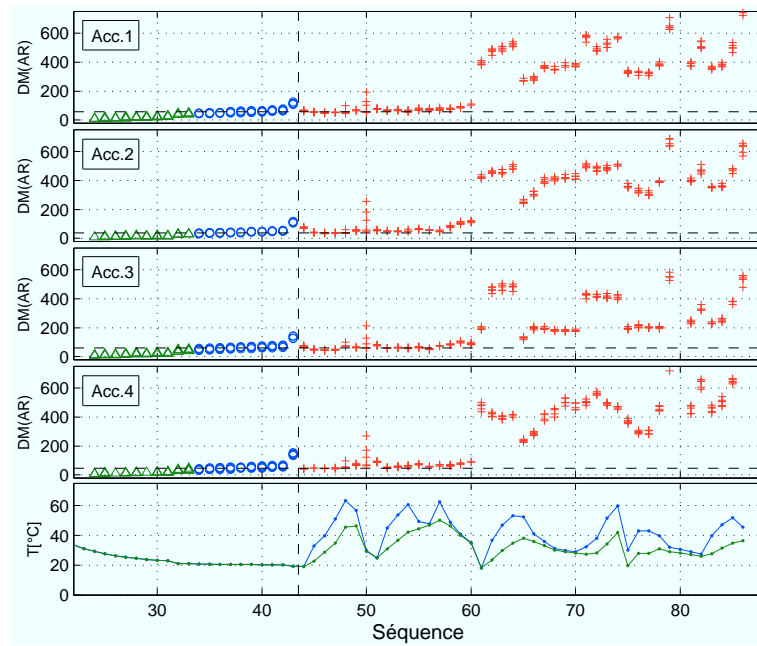


FIG. 6.8 Distance de Mahalanobis calculée sur les coefficients AR en fonction de la mesure, **Poutre**; \triangleleft base de classification; \circ base de test; $+$ données endommagées.

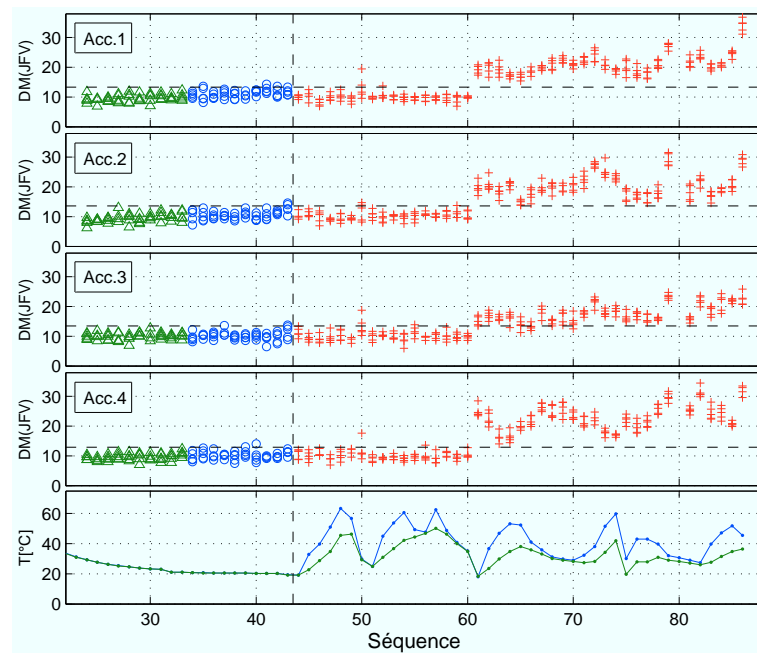


FIG. 6.9 Distance de Mahalanobis calculée sur le JFV en fonction de la mesure, **Poutre**; \triangleleft base de classification; \circ base de test; $+$ données endommagées.

6.5.1 Poutre en béton fibré

Les figures 6.8 et 6.9 présentent respectivement les résultats de détection de nouveauté pour les coefficients AR et le JFV obtenus sur les quatre accéléromètres. La distance de Mahalanobis de chaque mesure est tracée. Les points de la base de test se situent avant la séquence 43, et ceux associés à l'endommagement, à partir de la séquence 44. L'évolution de la température est également indiquée.

En terme de sensibilité à l'endommagement, les mesures sur structure endommagée commencent à se différencier nettement de celles de la base de test à partir de séquence 61. Cette séquence coïncide avec l'initiation d'une macro-fissure, qui engendre une brusque augmentation de la distance pour les deux indicateurs. La macro-fissuration semble de plus s'accompagner d'une amplification des effets de la température sur la distance, résultat observé mais inexpliqué à ce stade. Il apparaît enfin, pour les 2 indicateurs, une tendance modérée mais réelle à l'augmentation de la distance de Mahalanobis avec la progression de la macro-fissure.

Le JFV affiche des résultats plus homogènes que les coefficients AR. Les distances présentent en effet une distribution réellement constante jusqu'à la séquence 61, donnant un sens à la limite de classification fixée. En revanche, avec les coefficients AR, les distances calculées sur la base de classification, puis sur la base test, augmentent continûment du fait de l'éloignement temporel progressif de la base d'apprentissage. Durant cette période, pourtant, la température varie peu. Difficilement visible en raison de l'échelle, l'agrandissement de la figure 6.8 montre que la distance de Mahalanobis estimée sur les coefficients AR suit les variations de la température, même durant la phase de micro-fissuration (Fig.6.10) précédant l'ouverture de la macro-fissure. En résumé, les deux indicateurs parviennent à détecter correctement l'apparition de la macro-fissure malgré les effets des variations environnementales. Il est toutefois nécessaire de rappeler ici que ces résultats positifs sont probablement favorisés par le caractère invariant de la source d'excitation d'une mesure à l'autre.

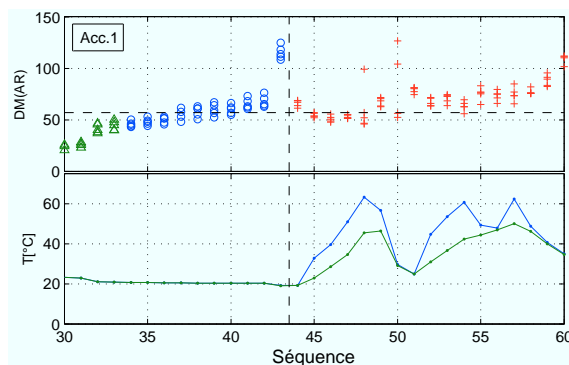


FIG. 6.10 Agrandissement de la figure 6.8

6.5.2 Maquette en bois

Les résultats, en terme de détection de nouveauté, obtenus sur la maquette en bois sont présentés dans les figures 6.11, 6.12 et 6.13, associées respectivement aux essais 36, 35 et 34. Chacune d'elles comporte 10 graphiques (1 par accéléromètre et par indicateur vectoriel) traçant la distance de Mahalanobis de chaque mesure de la base de classification, de la base de test et collectée pendant la séquence d'endommagement (identifiées respectivement par un triangle, un cercle et une croix). Les marqueurs plus foncés indiquent les mesures pour lesquelles l'endommagement est maximum.

Coefficients AR

Pour chaque essai, la distance de Mahalanobis présente d'importantes fluctuations sur la base de test. La forme de ces fluctuations invalide l'hypothèse d'indépendance des valeurs et impose en toute rigueur une méthode différente de détermination de la limite de classification. Cependant, il est toujours possible d'observer les variations induites par l'endommagement.

Dans le cas de l'essai 36 (Fig.6.12), la détection de la suppression de l'assemblage 1 échoue systématiquement. Pour tous les accéléromètres, les fluctuations de la distance sur les données de test englobent celles des mesures associées au cycle d'endommagement. La seule exception concerne les coefficients AR extraits des mesures de l'*accéléromètre 1* lors de la dégradation maximale (graphique en haut à gauche). En effet, les points associés à l'endommagement maximal franchissent la limite de classification. De plus,

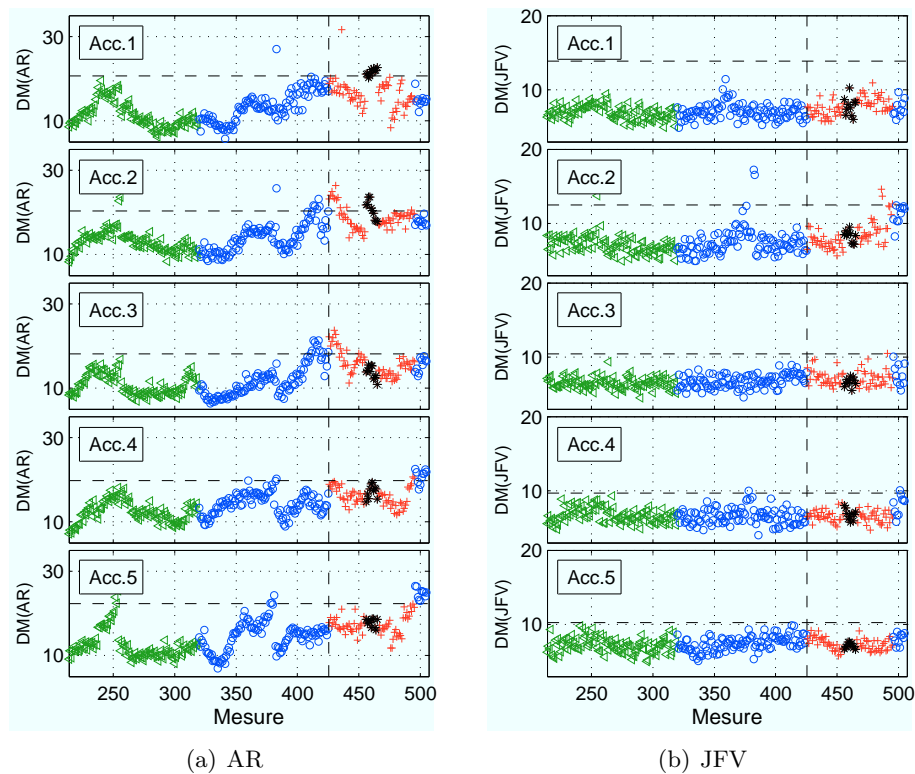


FIG. 6.11 Distance de Mahalanobis en fonction de la mesure, *essai 36*, *endommagement A₁*; \triangleleft base de classification; \circ base de test; $+$ données endommagées; $*$ endommagement maximum.

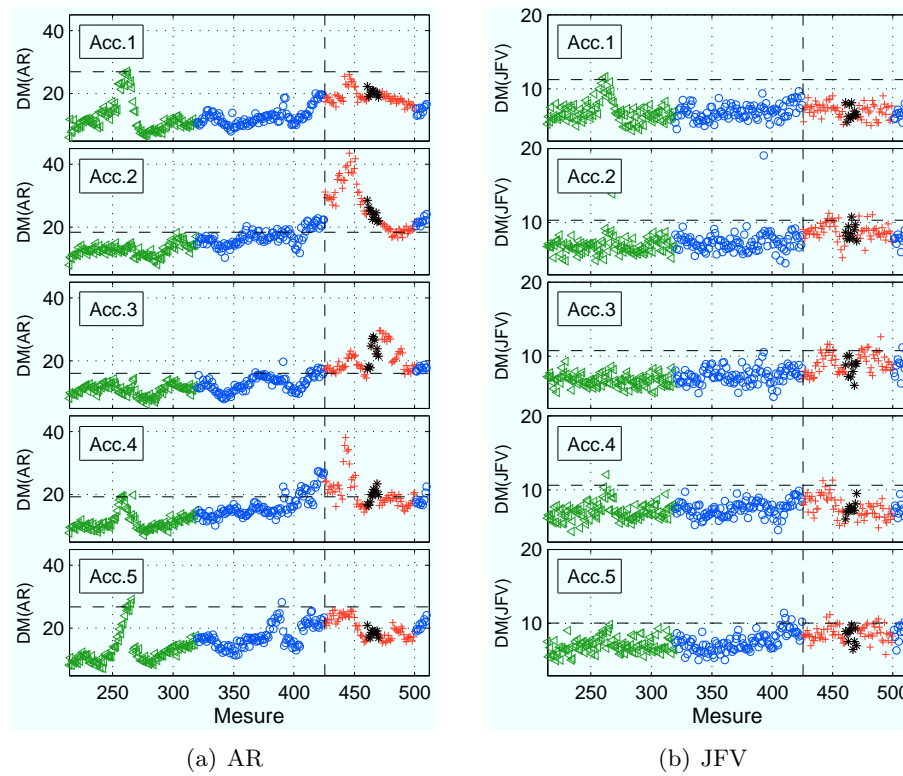


FIG. 6.12 Distance de Mahalanobis en fonction de la mesure, *essai 35*, endommagement A_3 ; \triangleleft base de classification; \circ base de test; $+$ données endommagées; $*$ endommagement maximum.

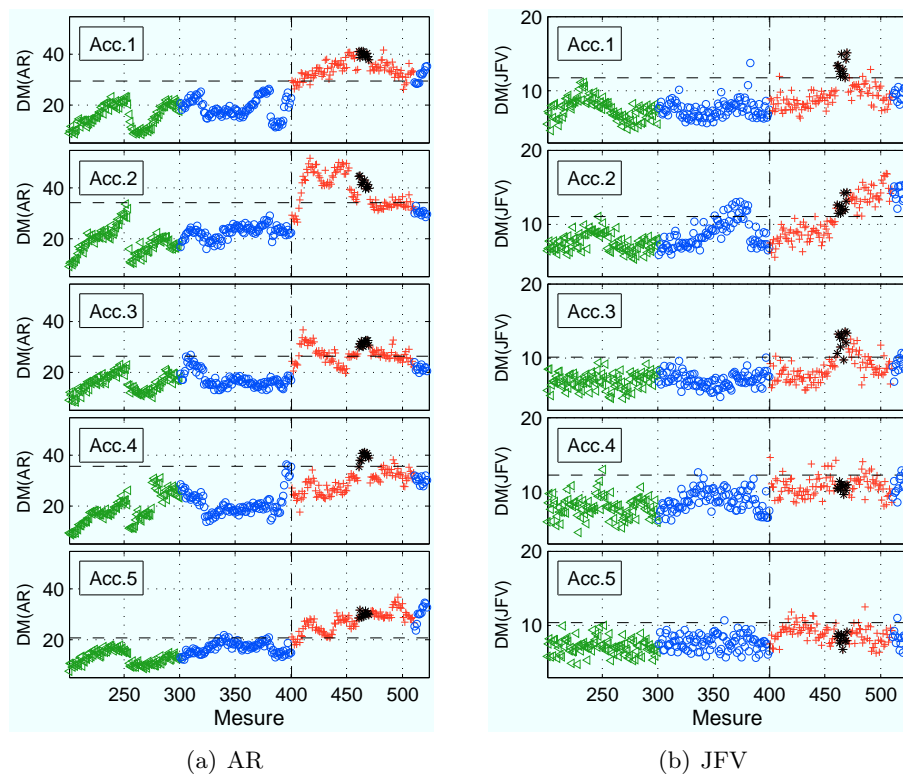


FIG. 6.13 Distance de Mahalanobis en fonction de la mesure, *essai 34*, endommagement D_{24} ; \triangleleft base de classification; \circ base de test; $+$ données endommagées; $*$ endommagement maximum.

l'accéléromètre 1 est le plus proche de l'endommagement. Cependant, il est difficile de déterminer avec certitude si l'augmentation de la distance ne résulte pas des variations environnementales. La dégradation se situant ici au troisième et dernier niveau de la structure, son impact sur le comportement dynamique est probablement plus faible comparativement aux 2 autres essais présentés ci-après, pouvant ainsi expliquer l'absence de sensibilité.

Les résultats de l'essai 35 manifestent davantage de sensibilité (Fig.6.12). Si les accéléromètres 1, 4 et 5 n'indiquent aucun changement, la distance augmente de façon significative pour les mesures relatives aux accéléromètres 2 et 3, situés sur le montant concerné par l'endommagement. Cette augmentation peut traduire une combinaison des effets des variations environnementales (début d'une période de chauffage à la mesure 450) et de l'endommagement. Cependant, les fluctuations des distances associées aux accéléromètres 2 et 3 sont trop significatives pour douter de l'impact de l'endommagement. Ici encore se pose le problème du critère de classification.

Concernant l'essai 34 (Fig.6.13), à l'exception de l'accéléromètre 4, tous les autres indiquent clairement un changement dans le comportement de la structure. La réponse de l'accéléromètre 1 reproduit même assez nettement la séquence d'endommagement (évolution réversible de la tension du câble 2-4). Ces bons résultats sont liés au fait que la suppression d'une diagonale du contreventement affecte la dynamique globale de la structure de façon plus conséquente qu'un assemblage. La réponse de l'accéléromètre 2 est assez particulière car elle présente une augmentation très brusque, puis une diminution également rapide, toutes deux non corrélées à l'évolution de l'endommagement. La première coïncide avec le déclenchement de la climatisation (mesure 406), et la seconde avec le démarrage des lampes à rayonnement IR (mesure 451). Cependant, ces mêmes opérations n'engendrent pas les mêmes réponses dans la base test. Des couplages entre les variations environnementales et l'endommagement expliquent ces évolutions irrégulières, néanmoins significativement nouvelles.

Peut-être est-il possible de réduire la sensibilité aux changements de comportement de la structure en réduisant l'ordre du modèle AR. Un modèle AR(15) est ajusté sur l'ensemble des mesures de l'essai 35. Les résultats, illustrés dans la figure 6.14, conduisent aux mêmes conclusions que celles obtenues avec un modèle AR(40).

JFV

Les performances du JFV sont considérablement dégradées par l'introduction des variations environnementales dans les données relatives à ces essais.

Bien que les distances de Mahalanobis semblent globalement indépendantes, permettant une détermination correcte d'un seuil de classification, le JFV est bien moins affecté par l'endommagement que les coefficients AR. Seul l'essai 34 (Fig.6.13(b)), associé à la plus grande perturbation du comportement dynamique de la maquette, montre une légère aptitude à la détection de l'endommagement lorsque celui-ci est à son plus haut niveau, sur les accéléromètres 1, 2 et 3. Les deux oscillations visibles sur les points de l'accéléromètre 2, coïncident avec deux phases de chauffage (mesures 320 → 380 et 541 → 522). La seconde est cependant amplifiée par la présence de l'en-

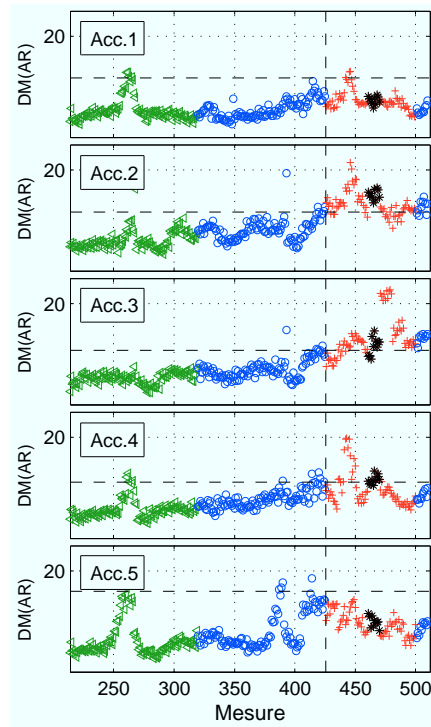


FIG. 6.14 *Distance de Mahalanobis calculée sur un $AR(15)$ en fonction de la mesure, essai 35, endommagement A_3 ; \triangleleft base de classification; \circ base de test; $+$ données endommagées; $*$ endommagement maximum.*

dommagement, ceci n'expliquant cependant pas le non retour à l'état initial à la fin de l'essai.

BILAN

Détecter l'apparition d'endommagement en présence de fortes variations environnementales impose de prendre certaines précautions. Il est nécessaire de mettre sur un pied d'égalité les données de classification et de test, et celles qui doivent être classées. En effet, le mélange aléatoire des données acquises sur structure saine pour former les bases d'apprentissage et de test réduit artificiellement la distance de Mahalanobis des mesures de test.

Afin de prendre en compte cette observation, les différentes bases de données sont constituées par des mesures consécutives, limitant ainsi la corrélation temporelle avec les données d'apprentissage, à l'instar des pratiques de surveillance sur un ouvrage réel en service.

Les résultats obtenus sur la poutre en béton fibré montrent que les deux indicateurs détectent sans ambiguïté l'initiation et la progression d'une macro-fissure. Le JFV présente une sensibilité aux variations de température plus limitée que les coefficients AR.

A l'exception de l'essai 35 réalisé sur la maquette en bois, dans lequel seuls les coefficients AR se montrent sensibles à la dégradation, les deux indicateurs vectoriels détectent généralement l'endommagement sur les mêmes capteurs et les mêmes essais, mais pas nécessairement aux mêmes moments. Les coefficients AR se montrent extrêmement sensibles, aussi bien aux variations environnementales qu'aux changements structuraux. Il en résulte de fortes fluctuations dans les valeurs de la distance de Mahalanobis, rendant difficile la détermination d'un seuil de classification et la différenciation entre variations normales et anormales. A l'inverse, les distances basées sur le JFV sont plus dispersées, mais de moyenne presque constante. Ainsi, la limite de contrôle est plus simple à fixer, mais dans ce cas, c'est la sensibilité de l'indicateur qui fait défaut, à l'exception des fortes modifications structurales.

INFLUENCE D'UNE EXCITATION INSTATIONNAIRE

7

SOMMAIRE	
INTRODUCTION	138
7.1 DONNÉES ÉTUDIÉES	139
7.2 PARAMÈTRES DE CALCUL	139
7.2.1 Modèle AR	139
7.2.2 Paramètres de reconstruction	140
7.3 OBSERVATION DES INDICATEURS	141
7.3.1 Coefficients AR	141
7.3.2 JFV	142
7.4 DÉTECTION DE L'ENDOMMAGEMENT	143
7.4.1 Conditions environnementales stables	143
7.4.2 Conditions environnementales variables	146
7.4.3 Augmentation de la dimension	146
7.4.4 Autre méthode : corrélation croisée	149
BILAN	153

INTRODUCTION

LES variations environnementales ne constituent pas le seul facteur susceptible de perturber la détection de l'endommagement. Dans un contexte de suivi des ouvrages en service, la sollicitation, inconnue et non mesurable, peut également être génératrice de nouveauté dans les mesures.

La plupart des travaux réalisés en laboratoire s'appuient sur des sollicitations correspondant à des bruits blancs de bandes de fréquences limitées. Cependant, la circulation routière et les rafales de vent peuvent générer une excitation de nature très différente, correspondant davantage à une succession aléatoire d'impulsions. Ce chapitre traite de l'effet de ce type d'excitation sur la performance des indicateurs proposés.

Pour répondre à cette question, les essais étudiés dans le chapitre 6 sont de nouveau réalisés, mais pour deux types de sollicitation dynamique : un mélange de bruit blanc et d'impulsions aléatoires tout d'abord, puis des impulsions aléatoires seules. Dans un premier temps, une sollicitation instationnaire seule est exercée sur la structure (avec et sans bruit blanc), puis, dans un second temps, la sollicitation instationnaire est combinée aux variations environnementales.

7.1 DONNÉES ÉTUDIÉES

Les données exploitées dans ce chapitre sont extraites des essais réalisés sur la maquette en bois (Tab.4.3). Les mesures des essais 22, 24, 25 (sans variations environnementales), et 33, 34 et 35 (avec variations environnementales) sont collectées pour les trois formes d'excitation présentées dans la section 4.4 :

Exc.1 : constituée d'un bruit blanc de bande de fréquence de 800Hz,

Exc.2 : constituée d'un mélange d'un bruit blanc de bande de fréquence de 800Hz et d'impulsions aléatoires en temps et en amplitude ; l'amplitude RMS du bruit est inférieure au 1/10 de l'amplitude de l'impulsion maximale (Fig.4.12),

Exc.3 : constituée uniquement par la composante impulsionnelle de l'excitation 2.

7.2 PARAMÈTRES DE CALCUL

Dans la mesure où la nature de l'excitation est modifiée, les caractéristiques de la réponse le sont également. Il convient donc de redéfinir les paramètres de calcul des indicateurs pour chacune des deux nouvelles formes de sollicitation. Des mesures de l'essai 24 sont extraites de la base de données de référence (structure saine) pour conduire les calculs présentés dans cette partie.

7.2.1 Modèle AR

La fonction d'auto-corrélation partielle et le critère d'Akaike sont estimés sur 25 mesures (5 par accéléromètre).

Le chapitre précédent montre que des écarts d'une dizaine d'unités dans l'ordre du modèle AR n'ont pas d'effet significatif sur la capacité de détection de l'endommagement lorsqu'il est supérieur à 20. La sensibilité est d'avantage impactée si l'ordre est doublé ou divisé par plus de deux. Le tableau 7.1 indique une valeur minimale de l'ordre de 20, sans différence significative entre les excitations.

Le choix se porte alors sur un modèle **AR(30)** pour tous les accéléromètres et pour les deux types d'excitations.

Acc.	Exc.2			Exc.3		
	Moy.	Min	Max	Moy.	Min	Max
1	25.0	23	27	29.4	26	32
2	31.6	26	35	31.8	26	37
3	23.6	22	24	28.8	22	38
4	24.4	23	27	28.6	26	30
5	30.8	24	39	30.8	26	35

TAB. 7.1 Ordres moyen, minimum et maximum obtenus pour chaque accéléromètre (5 signaux par accéléromètre), *essai 24*.

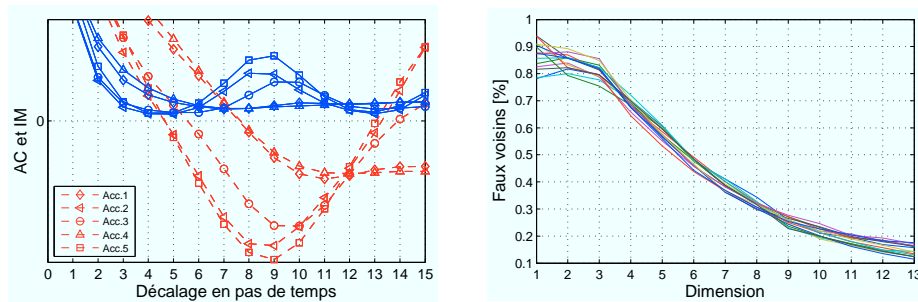
Excitation	Acc.1	Acc.2	Acc.3	Acc.4	Acc.5
2	7	5	5	7	5
3	9	6	7	9	6

TAB. 7.2 Valeurs du retard τ retenues pour chaque accéléromètre en vue de la reconstruction de l'espace des phases.

7.2.2 Paramètres de reconstruction

Les figures 7.1(a) et 7.2(a) représentent, respectivement pour les sollicitations **Exc.2** et **Exc.3**, les fonctions d'auto-corrélation et d'information mutuelle estimées sur une mesure de l'essai 24 (pour les cinq accéléromètres). Les valeurs indiquées pour le décalage temporel étant homogènes pour un même accéléromètre (constatation réalisée sur 3 mesures), seule une courbe est tracée pour chaque capteur. Les courbes montrent que les fonctions d'AC et d'IM fournissent globalement une information concordante sur le choix du retard optimal. La première s'annule lorsque la seconde atteint un minimum local. Cependant, il n'est pas possible de définir une valeur commune à tous les capteurs. Le choix du retard optimal est donc propre à chaque capteur.

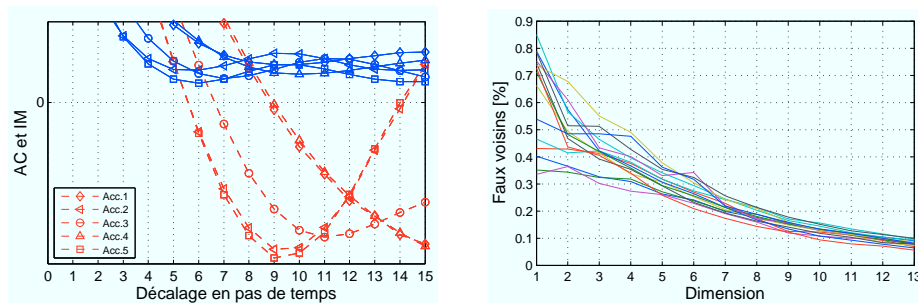
En conservant à l'esprit que les valeurs de retard 2 et 4 sont à proscrire



(a) Choix du retard (une courbe par accéléromètre); — Information mutuelle; - - Auto-corrélation.

(b) Taux de faux voisins en fonction de la dimension de reconstruction (trois courbes par accéléromètre).

FIG. 7.1 Courbes d'aide à la sélection des paramètres de reconstruction estimées sur trois mesures de l'essai 24, **Exc.2**.



(a) Choix du retard (une courbe par accéléromètre); — Information mutuelle; - - Auto-corrélation.

(b) Taux de faux voisins en fonction de la dimension de reconstruction (trois courbes par accéléromètre).

FIG. 7.2 Courbes d'aide à la sélection des paramètres de reconstruction estimées sur trois mesures de l'essai 24, **Exc.3**.

pour le calcul le JFV, les valeurs sélectionnées sont fixées et regroupées dans le tableau 7.2.

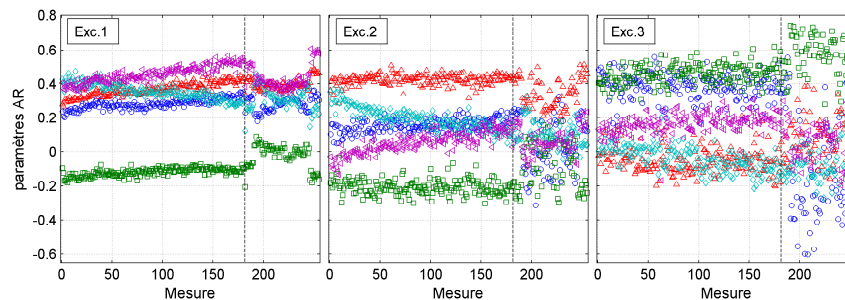
Les figures 7.1(b) et 7.2(b) représentent, respectivement pour les sollicitations **Exc.2** et **Exc.3**, le taux de faux voisins en fonction de la dimension de plongement pour des mesures collectées lors de l'essai 24 (3 mesures par capteur, soit 15 courbes). Une fois de plus, ces courbes ne permettent pas d'identifier de façon rigoureuse la dimension optimale. Une dimension arbitraire égale à 6 est conservée, toujours dans l'objectif de limiter la taille du JFV et de disposer d'un nombre de composantes proche de l'ordre du modèle AR.

7.3 OBSERVATION DES INDICATEURS

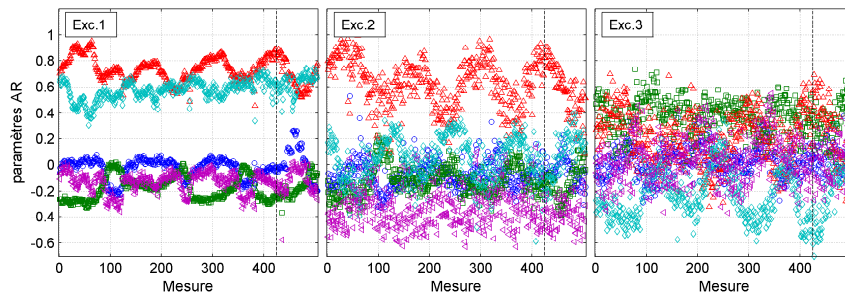
Afin d'observer les conséquences de la diminution de la part de bruit stationnaire dans l'excitation au profit d'impulsions aléatoires, le tracé brut des premières composantes de chaque indicateur vectoriel est représenté dans les figures 7.3 et 7.4 correspondant respectivement aux coefficients AR et au JFV. Sur chaque figure apparaissent les résultats relatifs à l'accéléromètre 1, pour les trois excitations (de gauche à droite) lors des essais 22 (en haut) et 36 (en bas). La ligne pointillée verticale marque le début de la séquence d'endommagement.

7.3.1 Coefficients AR

En l'absence de variations environnementales, et sous bruit blanc, les coefficients AR se situent dans une bande assez étroite. A mesure que le bruit



(a) essai 22, Acc.1, sans variations environnementales



(b) essai 36, Acc.1, avec variations environnementales

FIG. 7.3 Tracé des 5 premiers coefficients AR en fonction de la mesure relativement aux trois excitations.

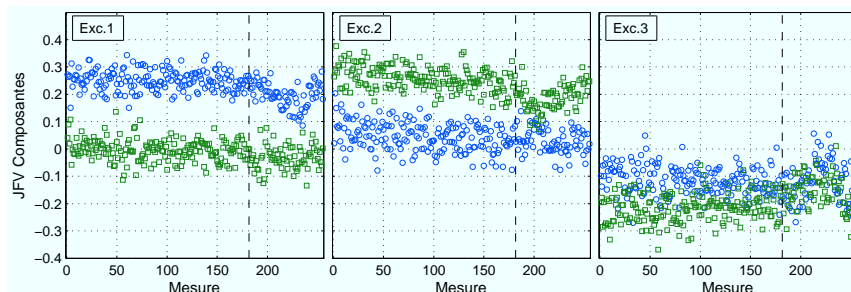
est remplacé par des impulsions, cette bande s'élargit considérablement. La variance des coefficients associés à la structure saine, représentés sur la figure 7.3(a), est multipliée par 1,5 en moyenne lors du passage à l'excitation 2, et par 3 pour l'excitation 3.

Dans le contexte favorable de conditions environnementales constantes, l'endommagement induit sur la maquette en bois est toujours observable malgré l'augmentation de la dispersion consécutive à l'introduction de la sollicitation instationnaire. Il est cependant certain que sa détection sera plus tardive, voir impossible pour de plus faibles dégradations.

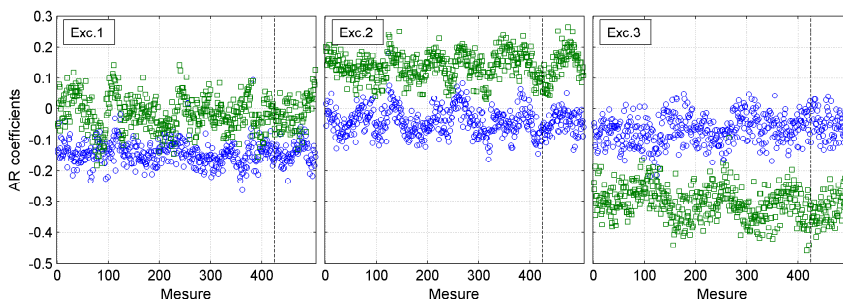
Ces observations sont également valables lorsque les variations environnementales sont introduites. Les fluctuations qu'elles imposent sont en partie masquées par l'augmentation de la dispersion (Fig.7.3(b)), réduisant ainsi la corrélation entre les valeurs successives de la distance. Le décrochement lié à l'apparition de l'endommagement n'est plus visible.

7.3.2 JFV

Les effets de l'accroissement de la part instationnaire de l'excitation sont moins visibles sur le JFV (7.4(a)). La variance des deux composantes représentées sur la figure augmente de 55% en moyenne (10% et 100%) entre l'excitation 1 et l'excitation 2, et reste égale entre la 2 et la 3. Cependant, la déviation provoquée par l'endommagement étant faible, cette augmentation peut suffire à réduire la sensibilité à l'endommagement. La dispersion augmentant peu, la sensibilité aux changements des conditions climatiques est peu modifiée (Fig.7.4(b)).



(a) essai 22, Acc.1, sans variations environnementales



(b) essai 36, Acc.1, avec variations environnementales

FIG. 7.4 Tracé des 2 premières composantes du JFV en fonction de la mesure relativement aux trois excitations.

7.4 DÉTECTION DE L'ENDOMMAGEMENT

L'ensemble des résultats en terme de détection de l'endommagement relatifs à tous les essais et à toutes les excitations sont regroupés dans l'annexe A.2. Afin d'alléger cette section, seul un échantillon de ces résultats est présenté ci-après.

Chacune des figures 7.5 à 7.9 est composée de 10 graphiques relatifs aux 5 accéléromètres (de haut en bas) et aux excitations 2 (à gauche) et 3 (à droite). Sur chaque graphique est tracée la distance de Mahalanobis des mesures de la base de classification, de la base de test et de la base associée à la structure endommagée, les points figuratifs de ces 3 bases étant respectivement représentés par un triangle vert, un cercle bleu et une croix rouge. Les points noirs indiquent le niveau d'endommagement maximum. La ligne pointillée verticale marque le début de la séquence d'endommagement et la ligne pointillée horizontale indique la limite de classification fixée à 3 écarts-types au dessus de la moyenne (moments statistiques relatifs à la base de classification).

7.4.1 Conditions environnementales stables

Les observations précédentes sont confirmées par la figure 7.5 qui représente la distance de Mahalanobis calculée sur les coefficients AR en fonction de la mesure pour l'essai 25, le plus affecté par l'excitation instationnaire.

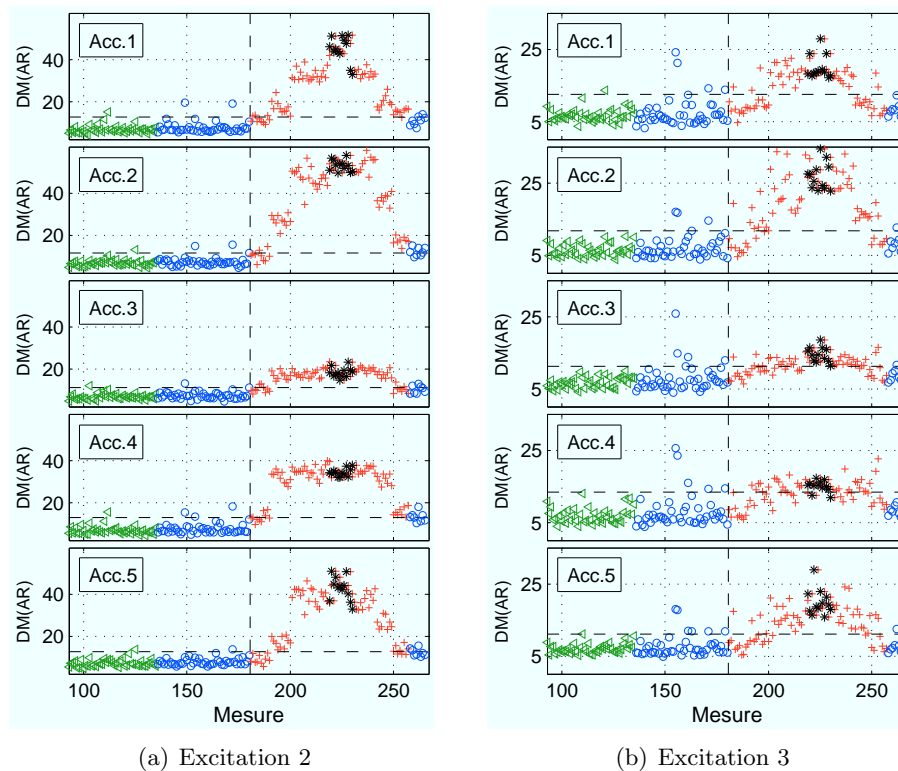


FIG. 7.5 Distance de Mahalanobis calculée sur les **coefficients AR** en fonction de la mesure, **essai 25, endommagement A₁**; \triangleleft base de classification; \circ base de test; $+$ données mesurées sur structure endommagée; $*$ endommagement maximum.

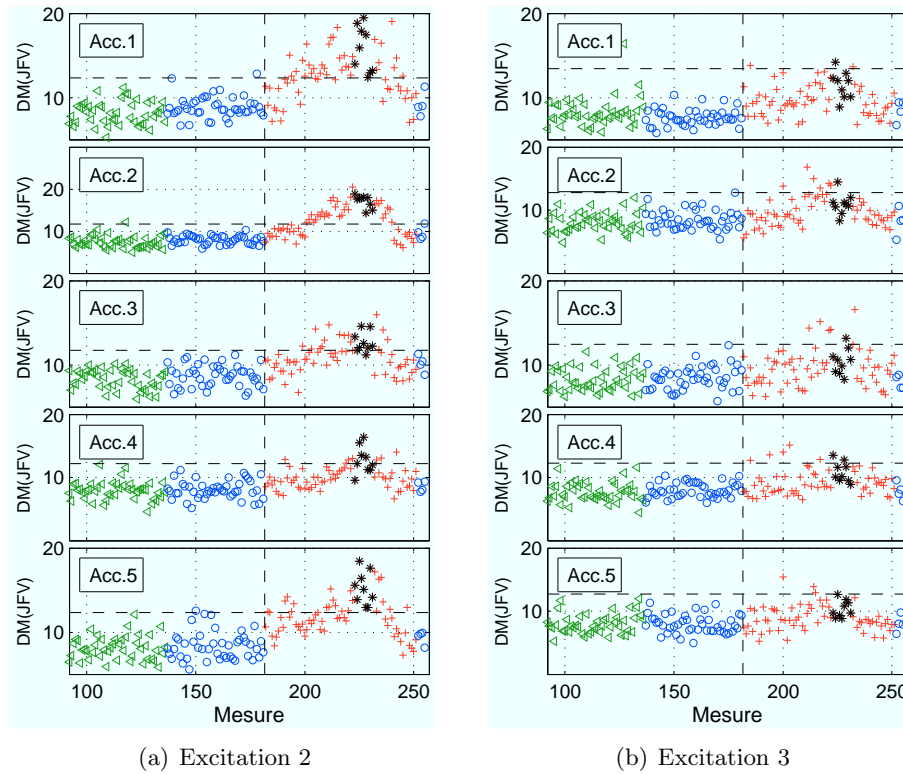


FIG. 7.6 Distance de Mahalanobis calculée sur les **composantes du JFV** en fonction de la mesure, **essai 22, endommagement A₂**; \triangleleft base de classification; \circ base de test; $+$ données mesurées sur structure endommagée; $*$ endommagement maximum.

Certes la détection de l'endommagement devient moins précise, mais elle présente toujours des performances satisfaisantes, et ce, également pour les essais 22 et 24 (Fig.A.1 et Fig.A.2).

Sur l'essai 25, le desserrage du premier écrou est clairement identifié sous un bruit blanc, ce qui n'est plus le cas sous l'excitation 2, puis la 3. Dans ces deux derniers cas, les premiers points rouges demeurent en effet sous la limite de classification, traduisant une diminution de la sensibilité.

Les résultats obtenus avec le JFV sont contrastés. L'essai 24, qui impose la plus grande dégradation, conserve le pouvoir discriminant des accéléromètres 1,3 et 4, pour toutes les sollicitations (Fig.A.8). Pour les autres essais, dont l'essai 22 (Fig.7.6), la sensibilité à l'endommagement est similaire sous les excitations 1 et 2. Une amélioration est même observée sur le capteur 5 de l'essai 25 qui peut trouver son origine dans une meilleure optimisation des paramètres de reconstruction. Ce résultat souligne la difficulté de détermination de ces paramètres et leur impact potentiel sur la performance de l'indicateur. Ce point fera l'objet d'une discussion dans la section 7.4.3. Enfin, l'excitation 3 induit une capacité de détection nulle sur l'essai 25 et limitée sur l'essai 22 (Fig.7.6(b)).

L'absence d'une composante de bruit dans la sollicitation dynamique réduit donc de façon significative la capacité du JFV à indiquer de faibles endommagements. Les coefficients AR demeurent efficaces, mais présentent une sensibilité réduite.

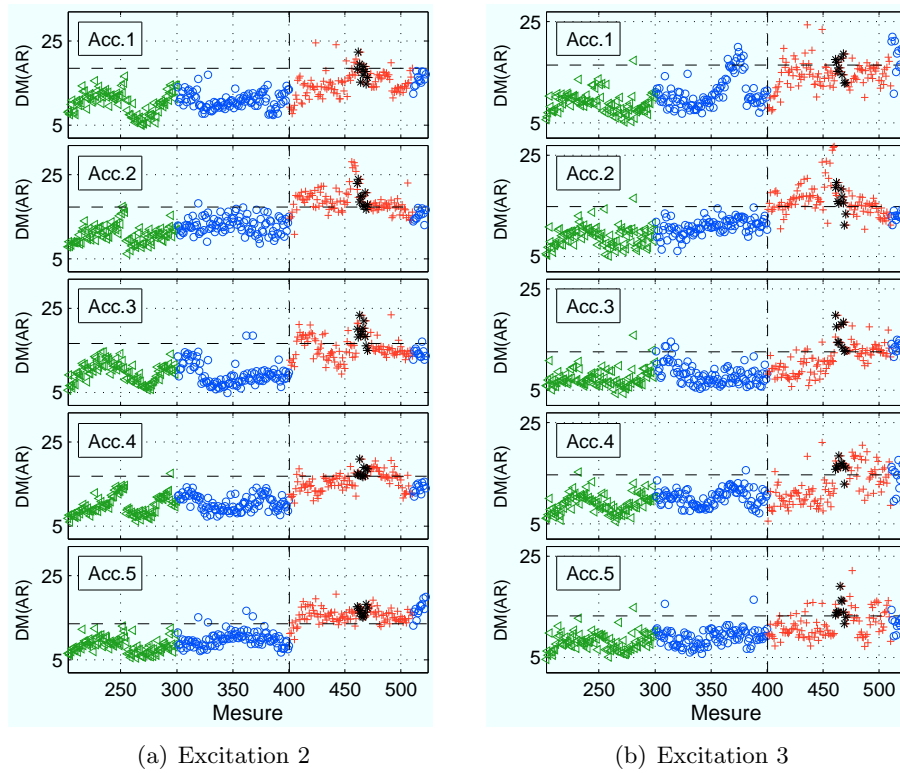


FIG. 7.7 Distance de Mahalanobis calculée sur les *coefficients AR*, en fonction de la mesure, *essai 34, Endom. D₂₄*; \triangleleft base de classification; \circ base de test; $+$ données mesurées sur structure endommagée; $*$ endommagement maximum.

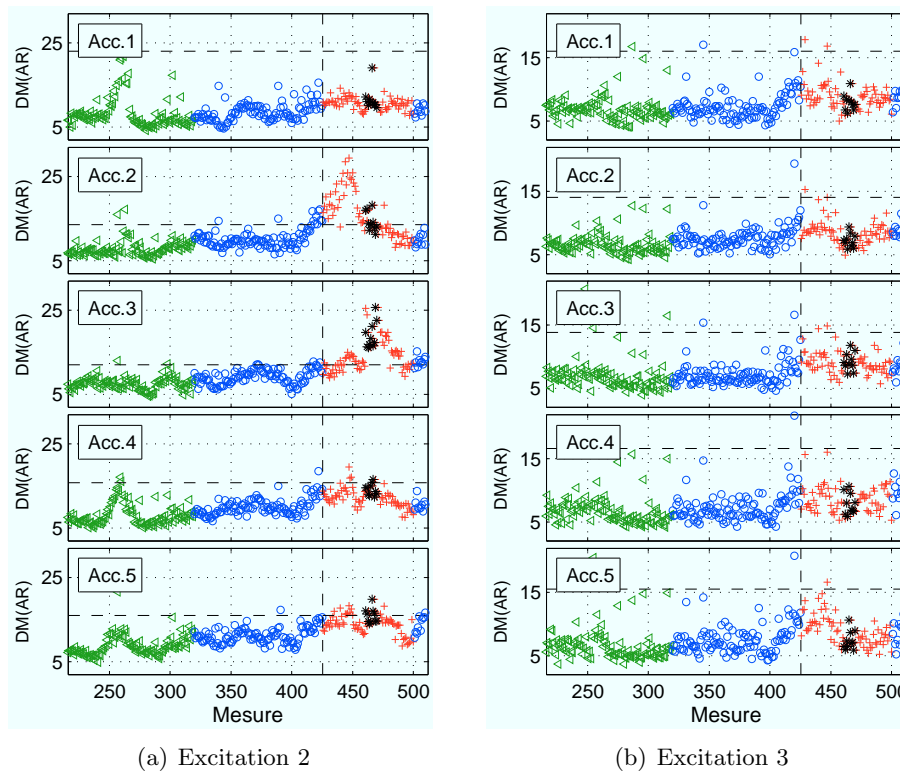


FIG. 7.8 Distance de Mahalanobis calculée sur les *coefficients AR* en fonction de la mesure, *essai 35, Endom. A₃*; \triangleleft base de classification; \circ base de test; $+$ données mesurées sur structure endommagée; $*$ endommagement maximum.

7.4.2 Conditions environnementales variables

Coefficients AR

Dans le cas des coefficients AR, l'ajout des variations environnementales confirme le fait que la dégradation des résultats est plus significative lors du passage de l'excitation 2 à l'excitation 3 que de l'excitation 1 à l'excitation 2.

En effet, les courbes des distances de Mahalanobis en fonction de la mesure sont très similaires pour les deux premiers types d'excitations (exemple : Fig.6.12(a) et Fig.7.8(a)), n'altérant pas davantage la détection de la dégradation. En revanche, sous la dernière forme de sollicitation, les fluctuations liées aux facteurs environnementaux sont considérablement réduites en raison de la dispersion accrue des valeurs de la distance. C'est le cas notamment des accéléromètres 2 et 3 durant l'essai 35 (Fig.7.8).

Sous excitation purement impulsionnelle (Exc.3), seul l'essai impliquant la suppression de la diagonale 2-4 permet encore de discriminer l'endommagement par l'accéléromètre 2, et dans une moindre mesure par les accéléromètres 3, 4 et 5 (Fig.7.7).

JFV

Considérées séparément, les variations environnementales et l'introduction d'impulsions aléatoires dans la sollicitation dégradent totalement la capacité de détection de la suppression d'un assemblage par le JFV. Il est donc attendu que leur combinaison ne remporte pas plus de succès, comme observé lors de l'essai 35 (Fig.A.11) et 36 (Fig.A.12).

Seule la suppression de la diagonale mise en œuvre lors de l'essai 34 est décelée dans certains cas (Fig.7.9). Sous l'excitation 2, les niveaux maximums d'endommagement sont encore détectés sur les accéléromètres 1, 3 et 5 (Fig.7.9). En présence de sollicitations purement impulsionnelles, la dispersion est trop importante pour que les variations causées par l'endommagement soient significatives, mis à part pour l'accéléromètre 3 qui reste le seul en mesure d'indiquer la présence de la dégradation.

7.4.3 Augmentation de la dimension

Face au choix arbitraire de la dimension de reconstruction et à l'efficacité limitée du JFV dans les cas de suivi les plus défavorables (variations environnementales et sollicitations instationnaires), se pose la question de la pertinence des paramètres retenus pour le calcul. Afin de fournir des éléments de réponse, une étude paramétrique est réalisée sur la dimension de reconstruction et le décalage temporel.

Sur la base des courbes de l'auto-corrélation et des taux de faux voisins associées à chaque signal, les paramètres suivants sont utilisés :

- Exc.2 $\Rightarrow \tau = 3 - 6 - 13$; $dim = 4 - 6 - 8$,
- Exc.3 $\Rightarrow \tau = 3 - 6 - 8 - 13 - 20$; $dim = 4 - 6 - 8$.

Les données analysées sont celles de l'essai 34, sous l'excitation 3 au niveau de l'accéléromètre 2. La partie précédente montre que cet essai constitue le seul cas, sous l'excitation 3, pour lequel l'endommagement est encore correctement détecté par les paramètres AR et non par le JFV (Fig.7.9(b) et

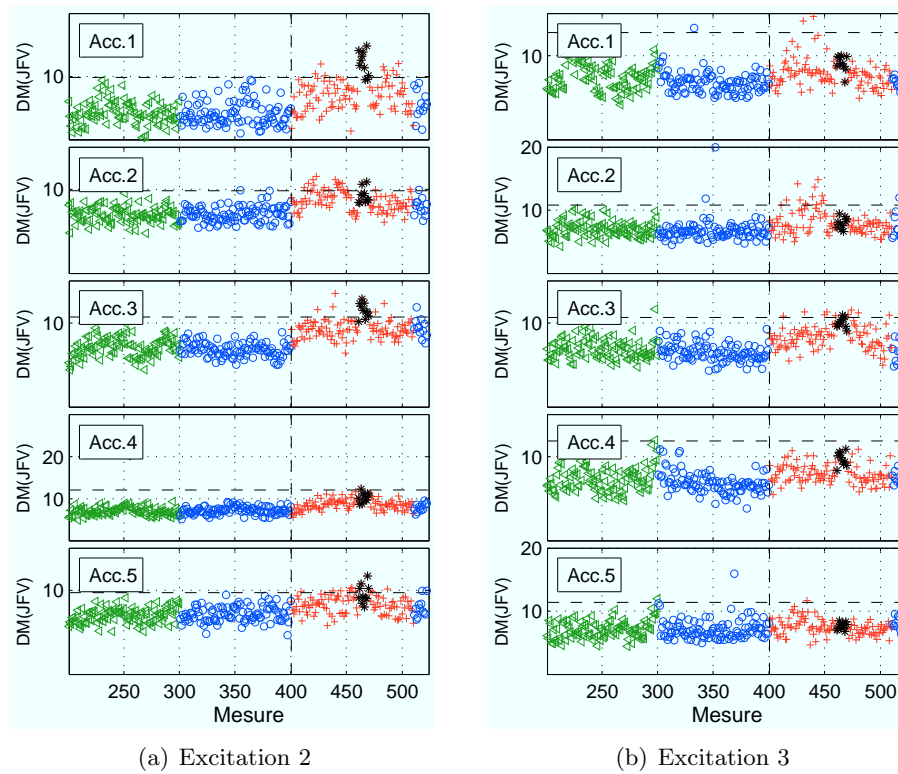


FIG. 7.9 Distance de Mahalanobis calculée sur les **composantes du JFV** en fonction de la mesure, **essai 34, endommagement D₂₄**; \triangleleft base de classification; \circ base de test; $+$ données mesurées sur structure endommagée; $*$ endommagement maximum.

Fig.7.8(b)). L'étude paramétrique cherche à vérifier s'il est possible d'égaliser la sensibilité obtenue par les coefficients AR avec des paramètres de reconstruction différents.

La qualité d'un jeu de paramètres est jugée seulement au travers de sa capacité à faire émerger les mesures collectées sur structure endommagée, tout en conservant une distance de Mahalanobis constante sur celles non endommagées.

La figure 7.10 regroupe les résultats obtenus lors de l'essai 34, sous l'excitation 3 et pour le capteur 2. La figure forme un tableau de 15 graphiques. Les lignes sont associées à une valeur constante de τ spécifiée à droite de la ligne, et les colonnes à une valeur fixe de la dimension de plongement spécifiée au-dessus. Comme précédemment, chaque graphique représente la distance de Mahalanobis de la base de classification, de la base de test et de la base associée à la structure endommagée.

La sensibilité à l'endommagement du JFV se révèle ici très influencée par les deux paramètres de reconstruction. Il apparaît que pour de faibles valeurs de τ (3 et 6), la dimension joue un rôle secondaire. A l'inverse, pour des valeurs de τ supérieures, l'augmentation de la dimension de plongement impacte de façon très positive la sensibilité de l'indicateur à l'endommagement.

Un jeu de paramètres optimaux se dessine. Les résultats présentés sur la figure plaident en effet pour une reconstruction en dimension 8 selon un délai

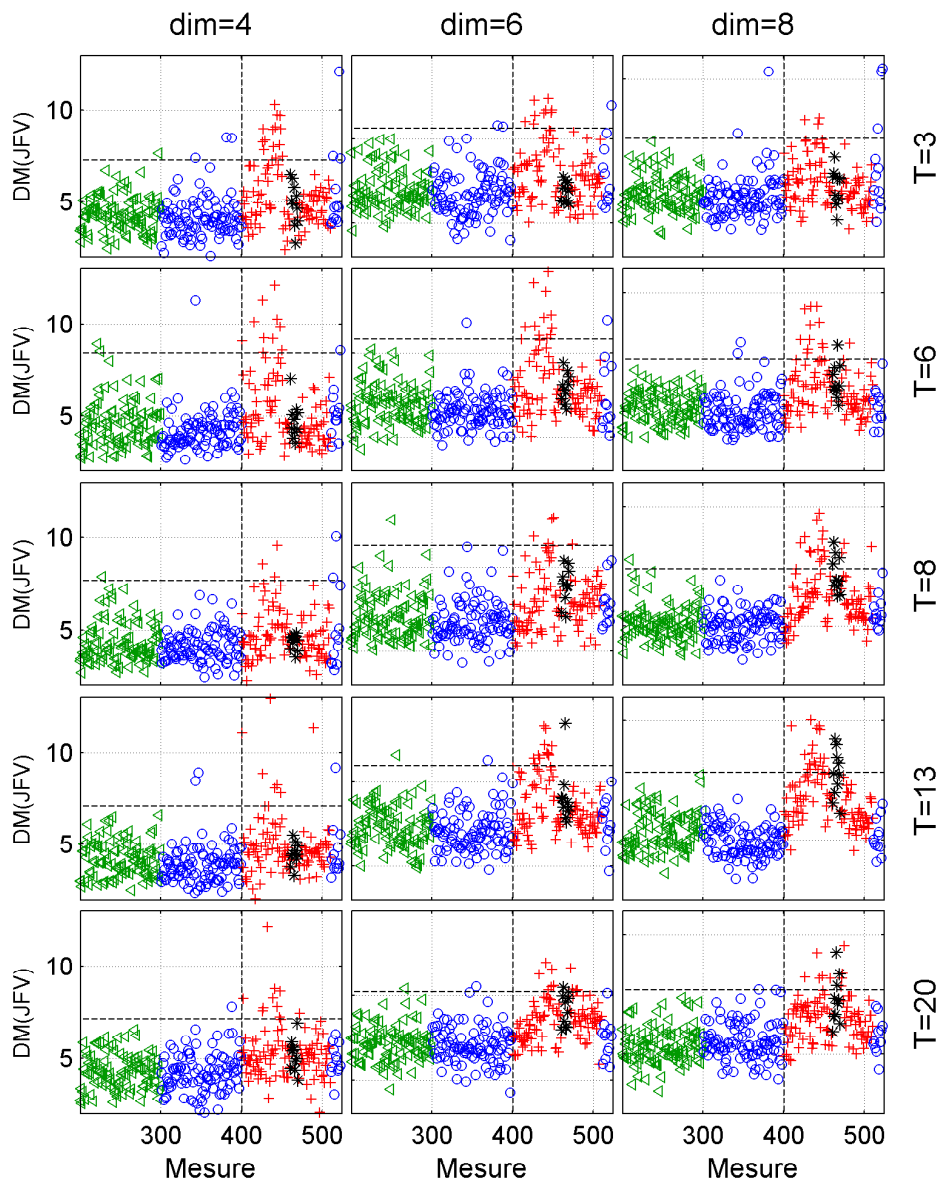


FIG. 7.10 Distance de Mahalanobis calculée sur le JFV en fonction de la mesure pour chaque jeu de paramètres, *essai 34, Excitation 3, Acc.2*; \triangleleft base de classification; \circ base de test; $+$ données mesurées sur structure endommagée; $*$ endommagement maximum.

égal à 8 ou 13. Dans ce cas, le JFV présente des performances honorables, équivalentes à celles des coefficients d'un modèle AR(30). Cependant, aucune méthode connue ne permet de fixer ces valeurs a priori, posant alors un réel problème pour une classification en mode non-supervisé.

Aucun des jeux de paramètres étudiés ici n'améliore la détection sur les autres essais. Il est cependant légitime de s'interroger sur le bénéfice potentiel d'une reconstruction en dimension supérieure à 8, bien plus coûteuse en temps de calcul.

Pour réaliser une comparaison équitable (en terme de nombre de composantes) avec le JFV calculé en dimension 8, un modèle AR(65) est ajusté sur le même essai. La figure 7.11 expose les résultats sous l'excitation 3.

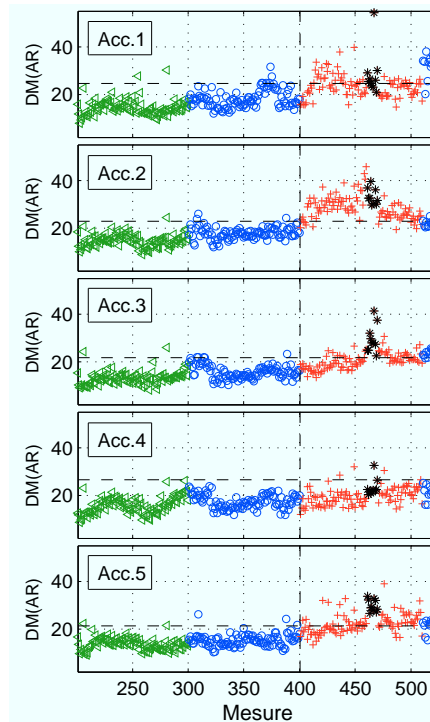


FIG. 7.11 *Distance de Mahalanobis calculée sur les coefficients AR(65) en fonction de la mesure, essai 34, endommagement D_{24} ; \triangleleft base de classification; \circ base de test; $+$ données endommagées; $*$ endommagement maximum.*

L'augmentation de l'ordre améliore légèrement la classification de quelques points isolés. Cependant, elle ne change pas de façon perceptible la sensibilité obtenue précédemment.

Le choix de l'ordre du modèle AR se montre moins critique. En effet, la fonction d'auto-corrélation partielle a toujours permis de définir une valeur conduisant à une sensibilité optimale, dans le sens où elle est peu améliorée avec un ordre supérieur.

7.4.4 Autre méthode : corrélation croisée

La capacité limitée des indicateurs proposés à détecter l'endommagement dans des conditions défavorables, mais réalistes, amène naturellement à envisager soit des développements orientés vers une véritable optimisation de leurs paramètres de calcul, soit la proposition d'autres indicateurs. Un moyen de s'affranchir partiellement de l'aléa lié à l'excitation peut consister à coupler l'information par paire de capteurs.

Ainsi, plusieurs indicateurs vectoriels sont formés :

- Calcul de modèles AR sur le signal *différence de deux mesures*,
- Calcul de modèles AR de prédiction croisée (modèle ajusté sur la mesure d'un accéléromètre et utilisé pour prédire la mesure d'un autre accéléromètre),
- Coefficients de la fonction d'auto-corrélation partielle croisée (corrélation partielle entre une valeur du capteur i et les valeurs décalées du capteurs j),

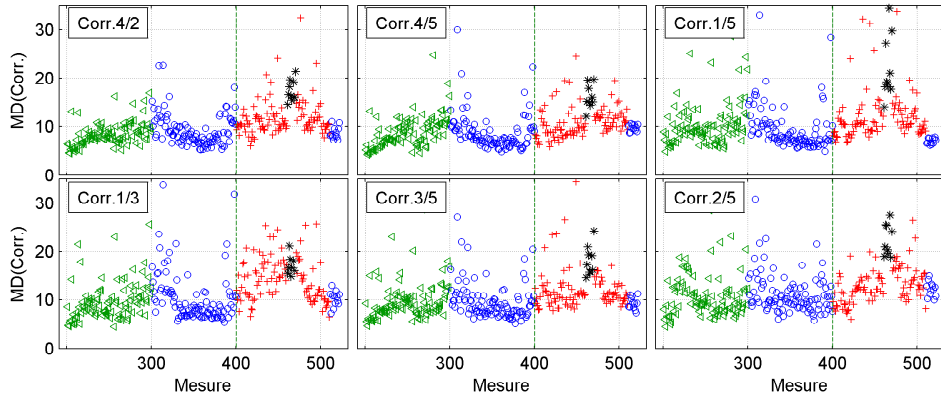
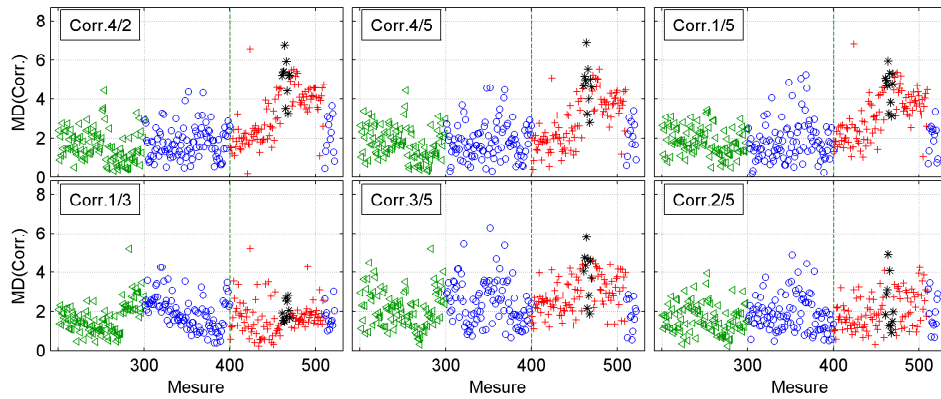
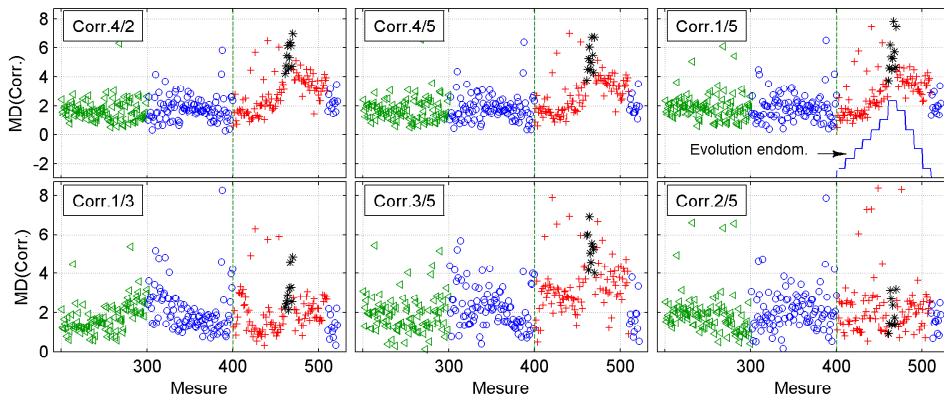


FIG. 7.12 Distance de Mahalanobis calculée sur *toutes les valeurs* de la fonction d'auto-corrélation croisée pour une sélection de paires d'accéléromètres ; **essai 34**, **endommagement D_{24}** , **Excitation 3** ; \triangleleft base de classification ; \circ base de test ; $+$ données endommagées ; $*$ endommagement maximum.



(a) Excitation 2



(b) Excitation 3

FIG. 7.13 Distance de Mahalanobis calculée sur *trois valeurs* de la fonction d'auto-corrélation croisée pour une sélection de paires d'accéléromètres ; **essai 34**, **endommagement D_{24}** ; \triangleleft base de classification ; \circ base de test ; $+$ données endommagées ; $*$ endommagement maximum.

- Coefficients de la fonction de corrélation croisée,
- Coefficients de la fonction d'information mutuelle croisée,
- Moyenne de l'erreur de prédiction par le modèle AR le plus proche (les coefficients AR_t de la mesure à tester sont estimés ; cette mesure est ensuite prédite à l'aide des coefficients AR_{ref} de la base d'apprentissage ayant la plus petite distance euclidienne avec AR_t),
- Erreur de reconstruction des AR après projection et déprojection dans l'espace des composantes principales des coefficients de la base d'apprentissage.

Parmi ces approches, peu conjuguent une bonne détection de l'endommagement et le filtrage des variations environnementales. Seule l'exploitation des valeurs de la fonction d'auto-corrélation croisée (FACC) offre des perspectives intéressantes. Le travail présenté ici étant exploratoire, cet avis n'est cependant pas définitif dans la mesure où des améliorations peuvent probablement être apportées dans la mise en œuvre des différentes méthodes.

La figure 7.12 présente la distance de Mahalanobis estimée sur des vecteurs formés par les 20 premières valeurs de la FACC, pour une sélection de couples d'accéléromètres sous l'excitation 3, lors de l'essai 34. La structure de chaque graphique conserve les mêmes codes que ceux utilisés jusqu'ici dans ce mémoire.

L'apparition de l'endommagement est assez clairement détectée. Cependant, la dispersion importante des valeurs de la distance de Mahalanobis et les lentes variations présentes sur les points relatifs à la structure saine empêchent une bonne discrimination statistique de la dégradation. Pour tenter de pallier ces deux problèmes, seules les trois valeurs de la FACC présentant la plus faible variance sur la base d'apprentissage sont conservées.

Les résultats obtenus avec les coefficients *filtrés* sont présentés dans les figures 7.13(a) et 7.13(b), correspondant respectivement aux excitations 2 et 3. Le graphique du couple 1/5 de la figure 7.13(b) indique également la séquence d'endommagement réversible (desserrage, puis serrage de l'écrou tendeur par pas de $\frac{1}{6}^{ème}$ de tours).

La comparaison des figures 7.12 et 7.13(b) permet de mettre en évidence l'intérêt du filtrage en variance des coefficients. La dispersion et l'effet des variations environnementales sont réduits. De plus lorsqu'une paire de capteurs détecte l'endommagement, l'évolution de la distance de Mahalanobis est bien corrélée aux niveaux de serrage et, par conséquent, au degré d'endommagement (Fig.7.13(b)). Avec les coefficients AR, cette évolution est généralement polluée par les variations environnementales.

La comparaison entre l'excitation 2 et l'excitation 3 révèle une plus faible dispersion pour la dernière, améliorant la détection de la dégradation. Cependant, de nombreux points aberrants sont présents et exigent l'utilisation de modèles statistiques robustes pour une classification minimisant les fausses alarmes.

Finalement, les valeurs de la FACC filtrées en variance permettent de détecter correctement l'endommagement à partir de quelques mesures avant son niveau maximal. Sur les mêmes données, le JFV (calculé selon des paramètres de reconstruction optimisés, $T = 8; dim = 8$, Fig.7.10) détecte bien la partie ascendante de l'endommagement, mais moins bien la phase

de retour à l'état initial. Enfin, les coefficients AR discriminent quasiment la totalité des mesures relatives à l'état dégradé, mais subissent davantage les variations environnementales (Fig.7.8(b)).

BILAN

La part de bruit stationnaire dans le signal de l'excitation affecte davantage le JFV, déjà assez dispersé, que les coefficients AR. Les performances sont peu modifiées tant que l'excitation comporte une part de bruit, ce qui est le cas pour des ouvrages en service. Ainsi, les problèmes liés aux variations environnementales sont prépondérants avec l'excitation 2. Si tout le bruit est supprimé, la dispersion dans les valeurs des indicateurs augmente, masquant en partie les fluctuations induites par les facteurs extérieurs, mais annulant totalement la capacité à discriminer de faibles endommagements, tels que la suppression d'un assemblage. Toutefois, la détection du retrait d'une diagonale reste possible avec les coefficients AR.

Si le choix de l'ordre du modèle AR n'apparaît pas critique pour la détection de l'endommagement lorsqu'il est suffisamment grand (plusieurs dizaines), celui des paramètres de reconstruction s'avère plus sensible. La méthode usuelle exploitant les fonctions d'auto-corrélation et d'information mutuelle n'indique pas nécessairement le délai optimal permettant la meilleure discrimination de l'endommagement. La courbe du taux de faux voisins est également insuffisante pour déterminer la dimension optimale de reconstruction. Cependant, des performances proches des coefficients AR peuvent être atteintes à conditions d'améliorer la méthode de choix de ces paramètres.

Devant ce problème, et l'impossibilité de déceler la suppression des assemblages lorsque les variations environnementales et l'excitation instationnaire sont combinées, des indicateurs exploitant l'information fournie par des paires de capteurs sont testés.

Seul un indicateur basé sur les coefficients de corrélation croisée permet d'obtenir des résultats intéressants. Bien qu'il soit impossible de déceler l'endommagement dans les cas où les AR et le JFV échouent, cet indicateur détecte correctement la diminution de tension dans une diagonale. Dans ce cas, malgré l'excitation impulsionnelle, l'indicateur filtre correctement les variations environnementales et présente une évolution bien corrélée à la tension du câble.

MODÉLISATION STATISTIQUE ET SEUIL DE CLASSIFICATION

8

SOMMAIRE	
INTRODUCTION	156
8.1 LES DIFFÉRENTES MÉTHODES TESTÉES	157
8.2 DONNÉES UTILISÉES	158
8.3 RÉSULTATS	159
BILAN	165
PERSPECTIVES	171

INTRODUCTION

LES chapitres précédents analysent exclusivement la sensibilité des indicateurs aux effets, seuls ou combinés, de l'endommagement, des variations environnementales et de la nature de l'excitation. L'analyse est réalisée de façon assez qualitative, sans réellement quantifier la capacité à détecter les dégradations imposées aux dispositifs expérimentaux. En effet, un seuil est simplement fixé de façon indicative à 3 écarts-types au-dessus de la moyenne de la base de classification. Or, l'étape de modélisation statistique est essentielle dans la démarche de détection de nouveauté, dans la mesure où elle conditionne la prise de la décision finale sur la classe d'appartenance (structure saine ou endommagée) de la mesure testée.

Outre les outils classiques tels que les diagrammes de contrôle et les tests statistiques, d'autres approches apparaissent de plus en plus fréquemment dans la littérature, telles les statistiques des valeurs extrêmes (§1.3.3). Au-delà de cette variété de modèles disponibles, les choix envisageables sont multiples :

- raisonner sur les valeurs individuelles ou sur des échantillons de valeurs,
- détecter de la nouveauté après le dépassement du seuil par une valeur ou par une séquence de valeurs.

Ces différentes options représentent plusieurs alternatives dont les performances peuvent être assez différentes.

L'objet de ce chapitre est de comparer un certain nombre de ces méthodes sur 5 jeux de données extraits des chapitres précédents, afin de proposer quelques critères de choix.

8.1 LES DIFFÉRENTES MÉTHODES TESTÉES

Les modèles statistiques exploités sont présentés dans la section 1.3.3. La population de référence évoquée est composée par la base de classification (25% des données collectées sur structure saine).

Chaque méthode est numérotée de 1 à 7. Les trois premières prennent en compte chaque valeur individuellement pour la construction du modèle. Les 4 autres méthodes s'appuient sur des échantillons de valeurs.

Méthode 1 : Modèle gaussien individuel

Les valeurs sont supposées indépendantes et de distribution gaussienne. La moyenne et la variance sont estimées par les équations 3.29 et 3.30, la limite de contrôle par l'équation 3.31.

Méthode 2 : Modèle des excédents

Une distribution exponentielle est ajustée sur les excédents de la base de classification pour estimer la limite de contrôle (Eq.3.34). Le nombre d'excédents est fixé à 15, quantité un peu faible pour ajuster le modèle, mais représentant une proportion déjà importante de la base de classification. Parmi les jeux de données testés, la plus petite base de classification compte 45 points.

Méthode 3 : Hotelling T^2

La base d'apprentissage est utilisée dans un premier temps pour calculer T^2 . Puis, la limite est fixée par l'équation 3.39.

Méthode 4 : Échantillons gaussiens - Population mère

La limite de contrôle des échantillons est fixée à partir de la moyenne et de la variance estimées sur la population mère (Eq.3.40).

Méthode 5 : Échantillons gaussiens - Moyenne des variances

La variance de la population mère est estimée par la moyenne des variances de chaque échantillon. La limite est fixée par l'équation 3.41.

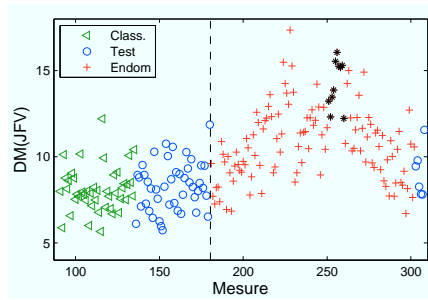
Méthode 6 : Échantillons gaussiens - Variance des moyennes

La limite est fixée par rapport à la variance de la moyenne des échantillons (Eq.3.43).

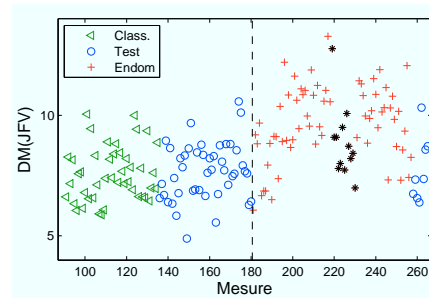
Méthode 7 : Maxima des échantillons

Une distribution des valeurs maximales extrêmes est tout d'abord ajustée sur les maxima de chaque échantillon. Puis la limite de contrôle est fixée par inversion de la fonction de distribution cumulée (Eq.3.45).

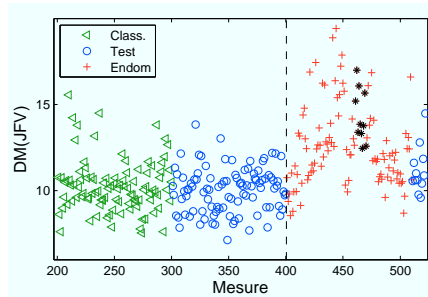
Pour limiter le nombre de fausses alarmes, il est courant de classer des valeurs comme nouvelles après le dépassement du seuil par une séquence de s points successifs, plutôt qu'une valeur individuelle. Dans l'étude proposée, s varie entre 1 et 5, de même que la taille des échantillons désignée par N_e . Dans le cas traité ici, afin de favoriser une détection précoce, il ne faut pas envisager plus de 5 mesures pour prendre une décision sur l'état de la structure. En supposant qu'un cycle de variations environnementales imposées à



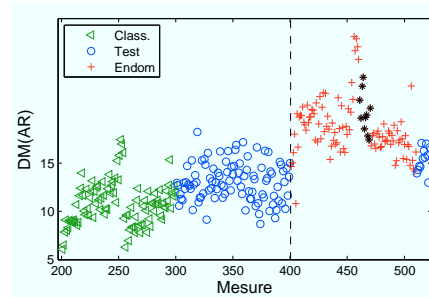
(a) J_1 : essai 24, Acc.3, Exc.1, JFV, $dim = 6$, $T = 6$



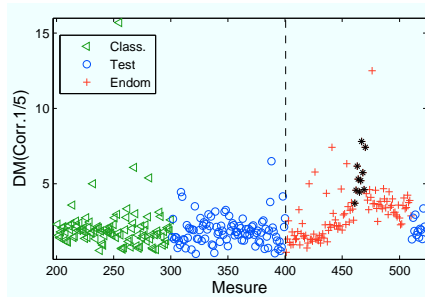
(b) J_2 : essai 25, Acc.1, Exc.1, JFV, $dim = 6$, $T = 6$



(c) J_3 : essai 34, Acc.2, Exc.3, JFV, $dim = 8$, $T = 8$



(d) J_4 : essai 34, Acc.2, Exc.2, AR(30)



(e) J_5 : essai 24, Acc.2, Exc.3, JFV, Corrélation 1/5

FIG. 8.1 Distance de mahalanobis des indicateurs pour les 5 jeux de données étudiés.

la maquette en bois correspond à une année de suivi, 5 mesures représentent 10 jours de suivi.

Dans la suite de l'exposé, les différentes méthodes sont désignées par M_i , $i = 1, \dots, 7$.

8.2 DONNÉES UTILISÉES

Cinq jeux de données formés de valeurs de distances de Mahalanobis sont employés pour tester les méthodes précédentes. Ils sont tous représentés sur les figures 8.1(a) à 8.1(e).

Les jeux 1, 2 et 3 présentent un accroissement net de la distance dès le début de l'endommagement, et les points successifs semblent assez décorrélés (Fig.8.1(a), 8.1(b) et 8.1(c)). L'enjeu est de déterminer à partir de quelle valeur de distance la dégradation est détectée. Le jeu n°4, obtenu

avec les coefficients AR, affiche des distances fluctuant avec les variations climatiques, ce qui est peu compatible avec les modèles statistiques proposés (Fig.8.1(d)). Néanmoins, leurs performances sont analysées. Enfin, le dernier jeu porte sur la corrélation croisée. Ces données sont décorréélées, mais la base de classification comporte un certain nombre de points aberrants, souvent nuisibles à l'estimation des limites de contrôle (Fig.8.1(e)). Il apparaît évident, compte tenu de la répartition des points, qu'au moins 30% des points acquis sur structure endommagée ne sont pas discernables de la base de classification dans ce dernier jeux.

Dans la suite de l'exposé, les jeux de données sont désignés par J_i , $i = 1, \dots, 5$. Le nombre de points, composant la base de classification utilisée pour l'ajustement des modèles statistiques, est indiqué dans le tableau 8.1.

$J1$	$J2$	$J3$	$J4$	$J5$
45	45	100	100	100

TAB. 8.1 *Taille de la base de classification*

8.3 RÉSULTATS

Les pourcentages de fausses alarmes (FA) et de vraies détections (VD) sont calculés pour chaque méthode. L'objectif de chacune d'elles est de parvenir à s'approcher de la performance idéale : 0% de FA et 100% de VD. Elles sont ainsi comparées par le biais de la distance euclidienne entre leurs points figuratifs de coordonnées (FA_i, VD_i) et le point idéal $(0, 100)$. Le tableau de l'annexe A.3 présente l'ensemble des résultats.

n	$J1$		$J2$		$J3$		$J4$		$J5$	
	FA	VD	FA	VD	FA	VD	FA	VD	FA	VD
1	3	40	1	23	2	31	38	84	2	15
2	8	67	6	57	1	57	19	95	3	31
3	3	76	2	67	6	62	26	98	4	44
4	7	82	9	80	7	70	25	98	15	58
5	4	78	8	72	9	75	40	99	19	57
6	6	86	7	82	4	68	53	100	0	51

TAB. 8.2 *Moyenne des pourcentages de fausses alarmes et de vraies détections pour les 5 jeux de données en fonction du nombre de points utilisés ($n = N_e \times s$).*

Rang	$J1$		$J2$		$J3$		$J4$		$J5$	
	Méth.	FA/VD	Méth.	FA/VD	Méth.	FA/VD	Méth.	FA/VD	Méth.	FA/VD
1	6 ₄₁	0/90	4 ₂₃	0/84	3 ₁₅	21/91	7 ₅₁	5/100	3 ₁₄	8/66
2	7 ₂₃	4/89	5 ₂₂	8/84	4 ₂₁	16/83	7 ₃₁	3/94	3 ₁₅	13/66
3	4 ₅₁	0/83	2 ₁₅	0/83	5 ₃₂	0/72	6 ₂₁	7/93	2 ₁₃	11/65
4	6 ₃₁	0/80	1 ₁₃	6/70	3 ₁₃	6/70	6 ₄₁	11/100	3 ₁₂	5/52

TAB. 8.3 *Pourcentages de fausses alarmes et de vraies détections pour les 4 meilleurs résultats sur toutes les méthodes, toutes les tailles d'échantillons, et toutes les longueurs de séquences ; Dans la rubrique Méth., 6₄₁ signifie, méthode 6, $N_e = 4$ et $s = 1$.*

Deux méthodes sont considérées comme équivalentes lorsque leurs valeurs de FA et VD diffèrent respectivement de moins de 10%. Si une méthode se montre systématiquement meilleure qu'une autre sur toutes les données, elles ne sont cependant pas considérées équivalentes, même si leurs taux de FA et de VD diffèrent de moins de 10%. L'objectif est ici de dégager des tendances générales sur les outils étudiés, des différences plus fines n'étant probablement pas significatives.

Nombre de points

Les différentes méthodes sont comparées pour un même nombre de points requis pour la décision ($n = N_e \times s$).

Le tableau 8.2 présente la moyenne des FA et des VD de toutes les méthodes, pour chaque valeur de n prise entre 1 et 6. Il apparaît qu'augmenter le nombre de points utilisés pour la décision ne permet pas de réduire significativement l'erreur de classification. L'utilisation de 4 mesures affiche des performances très correctes sur les 5 jeux de données. L'utilisation d'une seule mesure engendre une diminution systématique de VD, comprise entre 24% (J1, J2 et J3) et 15% (J5).

Les 4 meilleurs résultats par jeu de données, sont regroupés dans le tableau 8.3. Pour chaque rang et chaque jeu, est indiqué le pourcentage de

Rang	J1		J2		J3		J4		J5	
	Méth.	FA/VD	Méth.	FA/VD	Méth.	FA/VD	Méth.	FA/VD	Méth.	FA/VD
$s = 1$										
1	3	4/49	1	4/39	3	4/52	1	13/89	3	5/39
2	1	4/48	3	0/19	1	2/27	2	1/64	1	1/5
3	2	0/22	2	0/10	2	0/15	3	100/100	2	0/2
$s = 2$										
1	3	10/73	1	4/55	3	2/68	1	30/99	3	5/52
2	1	4/61	2	4/49	1	0/47	2	31/99	2	4/35
3	2	4/61	3	4/45	2	0/47	3	100/100	1	0/9
$s = 3$										
1	2	4/74	1	6/70	3	23/81	1	38/99	2	11/65
2	3	10/74	2	6/70	2	0/63	2	50/99	3	8/62
3	1	4/70	3	0/66	1	0/55	3	100/100	1	0/35
$s = 4$										
1	2	10/78	1	0/68	2	16/83	1	55/99	3	8/66
2	3	16/80	2	0/68	3	19/85	2	62/99	1	0/40
3	1	0/70	3	0/62	1	0/63	3	100/100	2	61/79
$s = 5$										
1	2	12/80	2	0/73	3	21/91	1	64/99	3	13/66
2	1	10/73	1	0/62	2	21/85	2	74/99	1	0/47
3	3	22/80	3	0/62	1	4/66	3	100/100	2	75/98

TAB. 8.4 *Pourcentages de fausses alarmes et de vraies détections pour les méthodes 1, 2 et 3 (valeurs individuelles).*

FA, de VD et la méthode associée. Cette dernière est désignée par trois chiffres : 6_{41} signifie la méthode 4 avec $N_e = 6$ et $s = 1$ ¹. Aucune méthode ne se distingue des autres. La distinction en fonction de J, fait néanmoins apparaître quelques différences. Si les données 1, 2 et 3 sont correctement classées avec l'ensemble des méthodes proposées, les méthodes basées sur échantillons sont plus adaptées aux données de J4, et celles basées sur les valeurs individuelles aux données de J5.

Jeux de données 1, 2 et 3

Le tableau 8.4 classe, pour une même valeur de s et pour chaque jeu de données, les méthodes individuelles (M1, M2 et M3). Il indique sur chaque ligne le rang, la méthode et les pourcentages de FA et VD. En dehors de $s = 1$, les résultats sont très homogènes sur J1 et J2, avec moins de 10% de différences, et souvent égaux entre M1 et M2. Des variations plus conséquentes, dans une fourchette de 25%, sont observées sur J3. Dans la majorité des cas, la M1 est classée au dernier rang en raison d'un taux moins important de VD, mais également de FA (surtout sur J3).

L'observation des données montre que la principale différence entre J3 et les deux autres jeux est un nombre plus important de points extrêmes dans la base de classification que dans la base de test. La méthode M2, basée sur les excédents, modélise correctement ces valeurs et place la limite relative à $s = 1$ assez haute, lui valant la dernière place du classement de performance. En revanche, la méthode basée sur le modèle gaussien (M1) calcule seulement la variance et place la limite plus bas. Au fur et à mesure que s augmente, la limite diminue. Dans la mesure où les excédents sont assez étalés, la limite descend plus rapidement avec M2. Il est résulte davantage de fausses alarmes.

La méthode M3 exploite directement la base d'apprentissage à la place de la base de classification, il est alors difficile et hasardeux de commenter ses performances.

Un tableau similaire est créé au sujet des méthodes par échantillonnage (Tab.8.5). En dehors de M7 (maxima), les performances obtenues avec un échantillonnage sont du même ordre de grandeur que celles basées sur les valeurs individuelles. Sur J3, le nombre de FA est réduit, mais celui de VD également (comme M1). Même si les écarts sont faibles, M5 est la méthode qui génère le moins d'erreurs sur les trois jeux de données, avec un taux minimum de FA faible. A l'inverse, M7 est très clairement moins performante que les trois autres. L'écart est plus prononcé sur J1 et J3 à mesure que s augmente.

Les mauvais résultats de M7 apparaissent lorsque le nombre de points d'ajustement du modèle est faible (15 pour J1, et 25 pour J3), et lorsque la base de classification présente des valeurs aberrantes. Étant donné que les maxima conservent ces valeurs aberrantes, elles prennent plus d'importance si la taille de la base de classification est réduite. Dans le cas de J3, ces valeurs sont assez étalées. Les fonctions de densité cumulée modélisée et empirique estimées sur les valeurs de J3 sont tracées sur la figure 8.2(b). Le modèle tend alors lentement vers 1, impliquant une limite de classification

¹Les méthodes individuelles ($N_e = 1$) apparaissent en bleu et celles basées sur échantillons en rouge.

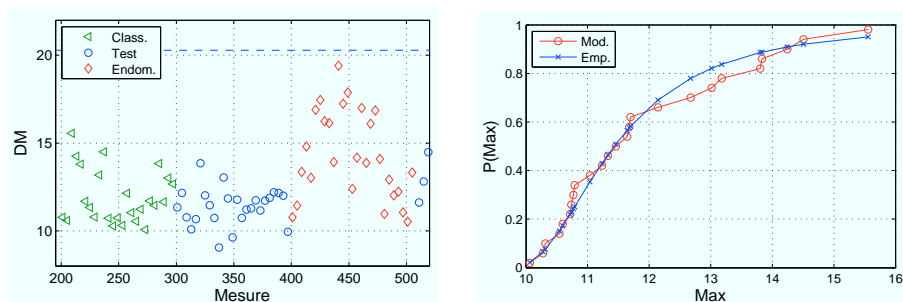
beaucoup trop élevée. La même explication est valable pour J1 en présence d'un seul point aberrant.

Prendre une décision basée sur 4 points offre la possibilité, soit de composer des échantillons de 4 points, soit de 2 points en association avec une séquence de deux valeurs. Les résultats sont équivalents pour les deux approches.

Les moyennes des taux de FA et de VD sont présentées dans le tableau 8.6 pour chaque méthode sur les jeux de données 1 à 3. En résumé, sur ces jeux, les méthodes 5 et 6 fournissent les meilleurs résultats, avec un taux de fausses alarmes moyen de 2.3% (M6), assez proche de la valeur cible de 1%. Seuls se démarquent sensiblement le modèle des maxima et le test d'*Hotelling T*². Le premier, très pénalisé par les points aberrants, affiche un taux de VD trois fois inférieurs aux autres méthodes, tandis que le second détecte presque deux fois plus de FA, en raison de limites fixées trop basses.

	J1		J2		J3		J4		J5	
Rang	Méth.	FA/VD	Méth.	FA/VD	Méth.	FA/VD	Méth.	FA/VD	Méth.	FA/VD
<i>s</i> = 2										
1	5 ₂₁	8/69	5 ₂₁	8/66	5 ₂₁	2/58	6 ₂₁	7/93	5 ₂₁	4/44
2	6 ₂₁	12/70	6 ₂₁	4/63	6 ₂₁	2/56	7 ₂₁	4/85	4 ₂₁	2/11
3	4 ₂₁	4/61	4 ₂₁	4/53	4 ₂₁	0/55	4 ₂₁	25/100	6 ₂₁	2/11
4	7 ₂₁	0/34	7 ₂₁	8/50	7 ₂₁	0/29	5 ₂₁	75/100	7 ₂₁	0/4
<i>s</i> = 3										
1	6 ₃₁	0/80	5 ₃₁	0/64	4 ₃₁	3/56	7 ₃₁	3/94	5 ₃₁	3/44
2	5 ₃₁	0/78	6 ₃₁	0/64	5 ₃₁	3/56	6 ₃₁	8/97	4 ₃₁	0/14
3	4 ₃₁	0/73	4 ₃₁	0/60	6 ₃₁	3/56	4 ₃₁	32/100	6 ₃₁	0/14
4	7 ₃₁	0/0	7 ₃₁	6/56	7 ₃₁	0/33	5 ₃₁	78/100	7 ₃₁	0/0
<i>s</i> = 4										
1	6 ₄₁	0/90	5 ₂₂	8/84	4 ₂₂	0/60	6 ₄₁	11/100	5 ₂₂	0/55
2	5 ₄₁	0/80	6 ₂₂	8/82	6 ₂₂	0/60	7 ₄₁	4/89	5 ₄₁	4/48
3	5 ₂₂	12/82	4 ₂₂	8/79	5 ₂₂	4/60	7 ₂₂	32/100	4 ₂₂	0/27
4	6 ₂₂	12/82	5 ₄₁	15/79	4 ₄₁	0/59	6 ₂₂	34/100	6 ₂₂	0/27
5	4 ₂₂	0/74	4 ₄₁	8/74	6 ₄₁	0/59	4 ₂₂	45/100	7 ₂₂	0/18
6	4 ₄₁	0/73	6 ₄₁	8/74	5 ₄₁	7/59	4 ₄₁	50/100	4 ₄₁	0/15
7	7 ₂₂	0/72	7 ₄₁	15/68	7 ₂₂	0/42	5 ₄₁	79/100	6 ₄₁	0/15
8	7 ₄₁	0/3	7 ₂₂	0/61	7 ₄₁	0/0	5 ₂₂	84/100	7 ₄₁	0/0
<i>s</i> = 5										
1	4 ₅₁	0/83	4 ₅₁	10/73	5 ₅₁	0/68	7 ₅₁	5/100	5 ₅₁	5/50
2	5 ₅₁	0/83	5 ₅₁	10/73	4 ₅₁	0/64	6 ₅₁	9/95	4 ₅₁	0/23
3	6 ₅₁	0/71	6 ₅₁	0/67	6 ₅₁	0/64	4 ₅₁	50/100	6 ₅₁	0/18
4	7 ₅₁	0/21	7 ₅₁	20/73	7 ₅₁	0/0	5 ₅₁	86/100	7 ₅₁	0/0

TAB. 8.5 Pourcentages de fausses alarmes et de vraies détections pour les méthodes 4 à 7 (par échantillonnage); Dans la rubrique Méth., 6₄₁ signifie, méthode 6, N_e = 4 et *s* = 1.



(a) Maxima des échantillons en fonction de la mesure

(b) Densités cumulées modélisée et empirique des valeurs maximales sur la base de données de classification.

FIG. 8.2 *Jeu de données 3; méthode 4; $N_e = 4$*

	M1	M2	M3	M4	M5	M6	M7
FA	2.6	6.4	10.5	2.3	4.4	2.3	4.1
VD	63.4	69.1	72.3	65.2	69.4	67.8	30.7

TAB. 8.6 *Moyenne des taux de fausses alarmes et de vraies détections sur J1, J2 et J3, pour $n = 1, \dots, 5$.*

Jeu de données 4

Ce jeu est caractérisé par une corrélation non nulle entre deux points successifs, corrélation liée aux variations environnementales. De ce fait, le découpage en temps des différentes bases conditionne le fait qu'elles contiennent, ou non, l'ensemble des états normaux de la structure. Concernant J4, ce n'est pas le cas. La moyenne des valeurs de la base de test est supérieure à celle de la base de classification. Pour détecter l'endommagement, il faut que la limite soit placée légèrement au-dessus de la valeur maximale de la base de classification.

Dans ces conditions, seules les méthodes 6 et 7 parviennent à discriminer l'endommagement avec un taux de FA inférieur à 9%, alors que les autres ne parviennent pas à descendre au-dessous de 25% (Tab.8.4 et 8.5).

Les méthodes 1 et 2 ne peuvent fonctionner que dans le cas où $s = 1$. Dans cette configuration, elles placent la limite à proximité des valeurs maximales (15,93 et 17,4 pour respectivement M1 et M2, avec une valeur maximale de 17,4). A mesure que s augmente, cette limite descend augmentant systématiquement le taux de fausses alarmes.

Du fait de la difficulté à disposer d'une base d'apprentissage parfaitement représentative de toutes les conditions normales, la valeur de T^2 augmente nécessairement sur la base de test, présentant un taux de fausses alarmes de 100%. La méthode M3 n'est donc pas adaptée pour des variables affectées par les variations environnementales.

Il convient également d'expliquer que M4 et M5 présentent systématiquement au moins 4 fois plus de FA que M6, alors que leurs performances sont équivalentes sur J1, J2 et J3. Étant donné qu'un échantillon est composé de points corrélés, sa moyenne est supérieure à celle obtenue sur un échantillon de valeurs indépendantes, et sa variance est réduite. C'est pourquoi, lorsque la limite se base sur la population mère (M4), elle est sous-estimée.

L'utilisation de la moyenne des variances est également biaisée (M5). Seule l'estimation directe de l'écart-type des moyennes est satisfaisante (M6).

Enfin, l'utilisation des modèles de maxima fixe une limite suffisamment haute pour distinguer l'endommagement avec une quantité négligeable de fausses alarmes (moins de 5%).

Pour des données aussi corrélées, le meilleur moyen de fixer la limite ne repose pas l'utilisation d'un modèle statistique fondé sur l'indépendance des variables. Le choix de la valeur maximale peut apporter une solution simple, mais peu fiable. Le recours à d'autres outils, tels que l'erreur de prédiction d'un modèle AR ajusté sur les points de la base de classification, permettrait de supprimer la corrélation, mais nécessiterait une base de données supplémentaire.

Jeu de données 5

Au-delà de 4 points dont les distances s'élèvent à environ 2,5 fois la moyenne, la base de classification présente en plus un point aberrant de distance égale à 8 fois la moyenne.

Ce point aberrant explique les mauvais résultats de M7, pour les mêmes raisons que précédemment. M2, qui fixe initialement la limite plus haut que M1, voit ses performances être très dépendantes de s , avec 61% de FA à partir de $s = 3$. Au contraire, le taux de FA demeure nul et le pourcentage de VD augmente en utilisant M1. Cependant, les meilleurs résultats sont obtenus selon la méthode M5. En effet, en moyennant les variances des échantillons, M5 subit moins les conséquences du point aberrant, comparativement à M4 et M6. Le cas de M3, n'utilisant pas la base de classification, n'est pas évoqué.

Lorsque le point aberrant est supprimé, les observations formulées sur J1, J2 et J3 sont confirmées. Les méthodes 5 et 3 génèrent de nouveau la meilleure classification.

BILAN

Sur l'ensemble des données analysées, la meilleure performance est produite par la méthode basée sur un échantillonnage et consistant à calculer la limite à partir de la moyenne des variances de chaque échantillon. Si les données corrélées sont exclues, la variance des moyennes (M5) offre le meilleur classement en raison de sa plus grande robustesse aux points aberrants. Le test d'*Hotelling* vient ensuite, mais ne tolère cependant pas de fluctuations au cours du temps.

Un autre enseignement de cette étude est l'échec des modèles aux valeurs extrêmes. Ces modèles sont plus adaptés aux cas où les valeurs extrêmes, ou rares, constituent réellement des individus de la population testée, et non des points aberrants. Dans le cas des mesures étudiées ici, ces valeurs représentent effectivement davantage des points indésirables, que les modèles doivent tenter d'ignorer.

Il est préférable d'utiliser plusieurs mesures, soit par séquence de dépassement de la limite pour le test d'*Hotelling*, soit en travaillant sur des échantillons de valeurs. Sur les données étudiées, un nombre optimal de 4 mesures est observé. Aucune méthode n'est cependant proposée pour déterminer rigoureusement cette valeur a priori, valeur dépendant de la variance et de la proportion de points aberrants des données de classification. Des développements scientifiques complémentaires sont nécessaires pour éclaircir ce point.

La gestion de la décision dans le cas de données corrélées n'est pas traitée. Des approches telles que les cartes de contrôle avec moyenne exponentielle pondérée ou l'erreur de prédiction de modèle AR sont à tester.

CONCLUSION GÉNÉRALE

LES développements scientifiques décrits dans ce mémoire s'inscrivent dans le contexte du suivi de l'état de santé des structures de génie civil par l'analyse de mesures vibratoires. L'approche étudiée repose sur la mise en œuvre de la détection de nouveauté pour déceler l'endommagement d'une structure par le biais de signaux vibratoires. Plus précisément, le travail présenté poursuit deux objectifs. Le premier objectif vise à analyser l'impact de fortes variations environnementales et de sollicitations fortement instationnaires sur la capacité de l'approche proposée à déceler l'endommagement d'une structure. Le second consiste à explorer le potentiel de la théorie des systèmes dynamiques non linéaires, et plus particulièrement de la reconstruction de l'espace des phases, pour la caractérisation des séries temporelles et l'extraction d'indicateurs sensibles à l'endommagement.

Pour réaliser ces objectifs, plusieurs types de données vibratoires relatives à différents cas d'études expérimentaux sont analysés :

- données simulées issues d'un modèle masse/ressort amorti,
- données extraites d'une étude du Laboratoire National de Los Alamos (*LANL*) sur une maquette de structure métallique,
- données mesurées sur une poutre en béton fibré endommagée progressivement par des cycles de chargement mécanique et soumise à des conditions environnementales variables (cycles de chauffage-refroidissement),
- données collectées sur une maquette de structure en bois pour différents scénarios d'endommagement, sous variations environnementales complexes (cycles thermiques et hydriques) et sollicitée par des sources de natures diverses (bruit blanc, impulsions, sollicitation hybride).

Les deux derniers dispositifs expérimentaux (poutre en béton fibré et maquette de structure en bois), réalisés dans le cadre de la thèse, permettent de contrôler la sollicitation dynamique et d'imposer des variations de conditions environnementales (lampes à rayonnement infrarouge, climatiseur/ventilateur portable). De nombreux essais sont réalisés, combinant ou non les diverses formes de sollicitations dynamiques et/ou les variations de conditions climatiques. La maquette en bois est systématiquement soumise à des séquences d'endommagement réversible, caractérisées par un pic d'endommagement et un retour de la structure à un état proche de l'état initial (non endommagé) à la fin de chaque essai.

Un indicateur d'endommagement (JFV) formé par les composantes de la matrice *jacobienne* de la dynamique estimée dans l'espace des phases reconstruit est proposé. Il s'inspire de l'algorithme de calcul des *exposants de Lyapunov* mais présente un temps de calcul considérablement réduit. Le JFV et les exposants de Lyapunov représentent des indicateurs dont le calcul

est aisément automatisable, condition indispensable pour un suivi continu d'ouvrage. Ces 2 indicateurs sont confrontés lors d'une première étude afin d'appréhender le potentiel du JFV. Ils présentent des performances comparables en terme de sensibilité à l'endommagement sur les données simulées issues du système masse/ressort. En revanche, le JFV surpasse très nettement les exposants de Lyapunov sur les données réelles collectées sur la maquette de structure métallique et fournies par le *LANL*. Une autre étude permet de fixer certains paramètres de l'algorithme de calcul du JFV, optimisant sa performance. Ces paramètres impactent également le protocole de reconstruction de l'espace des phases.

Le JFV est ensuite confronté à un indicateur vectoriel très utilisé dans le champ de la caractérisation des séries temporelles et formé par les coefficients du modèle auto-régressif (AR) ajusté sur la mesure. La comparaison est réalisée, dans un premier temps, sans variations environnementales et sous un bruit blanc stationnaire. Les 2 indicateurs sont ensuite éprouvés dans un contexte rendu plus défavorable par l'introduction progressive des effets des variations environnementales et de sollicitations impulsives, effets étudiés séparément dans un premier temps, puis combinés. L'étape de normalisation, qui situe une nouvelle mesure par rapport à la base de données de référence associée à l'état sain de la structure, est réalisée par le calcul de la distance de *Mahalanobis*.

Sous conditions environnementales constantes, les deux indicateurs détectent correctement l'endommagement pour les différents scénarios imposés à la maquette de structure en bois. Il est même possible de suivre l'évolution des dégradations, ce qui constitue une mesure relative de l'état de la structure. Les coefficients AR se montrent très sensibles, associant même les mesures finales (retour à l'état initial) à l'état endommagé du fait de l'impossibilité d'assurer une réversibilité parfaite. Le JFV est plus dispersé, et légèrement moins sensible, lui permettant de reconnaître le retour à l'état initial dans la majorité des cas.

L'introduction des variations environnementales fait apparaître un problème souvent dissimulé dans la littérature, mais d'importance capitale. Pour mettre en place la démarche de détection de nouveauté de façon rigoureuse, les mesures de référence (relatives à la structure saine) doivent être divisées en trois bases de données : la base d'apprentissage, dédiée à la normalisation des mesures (ici, le calcul de la distance de *Mahalanobis*), la base de classification, servant à la détermination de la limite de classification et la base de test, permettant de mesurer la performance de l'outil par le biais du taux de fausses alarmes. L'enregistrement des données de référence cours généralement sur une période suffisamment longue pour intégrer, autant que possible, l'ensemble des variations environnementales et opérationnelles normales éprouvées par la structure. L'affectation de ces données dans les trois bases précédentes est souvent réalisée de façon aléatoire. Il en résulte une distance de *Mahalanobis* artificiellement basse pour les données mélangées et proches temporellement de la base d'apprentissage. Dans un cas réel, la mesure testée est décorrélée en temps des mesures de référence. Il doit donc en être de même pour les données des bases de test et de classification. Ainsi, il apparaît plus rigoureux d'effectuer la division des mesures de référence par blocs continus.

Observés individuellement, les coefficients AR présentent de fortes variations liées à la température. Une régression entre les coefficients AR et la température ne semble cependant pas envisageable. Ainsi, la distance de *Mahalanobis* calculée sur les coefficients AR varie continûment au cours de l'essai. Il est difficile, dans ces conditions, de déterminer un seuil de classification en raison de la corrélation existant entre les valeurs successives. Le JFV subit également l'influence de la température, mais de façon moins marquée en raison d'une dispersion plus importante. Le problème de la corrélation des valeurs successives de la distance de Mahalanobis est beaucoup moins présent pour cet indicateur, car moins sensible.

L'endommagement apparaît ici détectable, si la distance dépasse la valeur maximale obtenue sur la base de classification. Sur les trois essais relatifs à la maquette en bois en l'absence de variations environnementales, deux consistent à supprimer de façon réversible et progressive l'un des assemblages. Le troisième essai porte sur la modification progressive et réversible de la tension d'une diagonale de contreventement. Seul ce dernier essai, qui implique un changement plus important du comportement dynamique, révèle une augmentation de la distance de Mahalanobis suffisamment significative pour qu'elle puisse être imputée avec certitude à l'endommagement. Le JFV décèle des changements sur les mêmes accéléromètres que les coefficients AR. Cependant, ces derniers détectent ces changements sur un plus grand nombre de mesures, témoignant d'une meilleure sensibilité.

L'effet de la nature de l'excitation, est étudié par le biais de trois types de sollicitations : un bruit blanc de bande de fréquence de 800Hz, des séquences aléatoires d'impulsions associées à un bruit blanc en fond, et enfin des séquences aléatoires d'impulsions seules. Sans variations environnementales et en présence de sollicitations impulsives, en dépit d'un accroissement de la dispersion, l'endommagement imposé à la maquette est détecté dans les tous cas par les deux indicateurs, à la condition que l'excitation contiennent une part de bruit blanc. Cette capacité de détection de nouveauté est considérablement dégradée lorsque l'excitation est constituée de sollicitations impulsives seules. Dans ce cas, le JFV ne détecte plus que quelques points maximaux, alors que les coefficients AR peuvent classer correctement au moins 70% des points relatifs à la structure endommagée.

Lorsque les sollicitations stationnaires sont combinées aux variations environnementales, les difficultés s'additionnent. Si une part de bruit blanc subsiste dans la sollicitation, les résultats obtenus sont similaires à ceux relatifs à une excitation stationnaire, avec cependant une dispersion des points légèrement supérieure. En l'absence de bruit dans la sollicitation (impulsions seules), seuls les coefficients AR parviennent encore à déceler la suppression de la diagonale, et ce, sur un seul accéléromètre.

Devant cette capacité limitée à déceler les dégradations induites dans la structure, le recours à d'autres indicateurs semble nécessaire. Afin de limiter l'effet de la nature de l'excitation, plusieurs méthodes s'appuyant sur le couplage des accéléromètres par paires sont testées de façon exploratoire. Seule la corrélation croisée entre deux capteurs fournit des résultats qui égalent pratiquement ceux obtenus avec les coefficients AR, surclassant même ces derniers sur le plan de l'indépendance des points des bases de classification et de test.

Un autre moyen d'améliorer les résultats consiste à utiliser une méthode différente de normalisation. Or, la distribution des indicateurs vectoriels montre une répartition en nuage unique, sans forme particulièrement complexe et sans séparation nette des mesures endommagées avec les mesures saines. Dans ce cas, l'utilisation de méthodes non linéaires telles que les réseaux de neurones auto-associatifs, ou les machines à vecteurs de support n'est pas justifiée et n'apporterait probablement pas d'amélioration.

L'étude montre également que les méthodes traditionnellement dédiées à l'estimation des paramètres de reconstruction de l'espace des phases (dimension et délai) ne permettent pas toujours d'opérer un choix déterminé ou unique. Ces méthodes sont effectivement conçues pour des systèmes parfaitement déterministes, alors que les mesures analysées sont en partie stochastiques. De plus, lorsqu'un choix est possible, les valeurs indiquées n'aboutissent pas obligatoirement à la meilleure option en terme de détection de l'endommagement. Ainsi, une étude paramétrique montre ici qu'avec un jeu de paramètres différent, le JFV parvient à détecter sans ambiguïté la suppression de la diagonale malgré les variations environnementales et sous l'excitation purement *impulsionnelle*. Par ailleurs, la détermination de l'ordre des modèles AR ne pose pas de difficultés. L'utilisation de la fonction d'auto-corrélation partielle fournit un ordre de grandeur suffisant pour capter les changements de la dynamique. A condition que la valeur initiale soit suffisamment élevée (plusieurs dizaines), les résultats sont peu affectés par une augmentation ou une diminution d'ordre de 10 unités.

Sur l'ensemble des essais analysés, la localisation de l'endommagement est impossible. Aucune corrélation significative n'apparaît entre l'augmentation des distances de Mahalanobis pour chaque capteur et la position de l'assemblage altéré. Dans certains cas, l'accéléromètre le plus proche peut ne présenter aucune variation significative.

La dernière partie du travail présenté concerne la détermination de la limite de classification, ou seuil de contrôle. C'est un point essentiel de la détection de nouveauté, dans la mesure où il conditionne la prise de décision. Un modèle statistique est ajusté sur les mesures de la base de classification, puis la valeur associée à la probabilité de fausse alarme acceptée (1%) est utilisée comme seuil. Sept méthodes sont testées. Trois de ces méthodes consistent à ajuster le modèle sur les valeurs individuelles (modèle gaussien, modèle basé sur les excédents, test multidimensionnel d'*Hotelling T²*). Les autres méthodes s'appuient sur des échantillons de valeurs (trois méthodes basées sur trois manières différentes d'estimer l'écart-type des moyennes des échantillons et un modèle basé sur les maximums). Dans le premier cas, il est possible de déclarer les mesures comme *nouvelles*, si une séquence de s points successifs dépasse la limite. L'objectif est de limiter les fausses alarmes (FA). Les taux de FA et de vraies détections (VD) sont calculés pour une longueur de séquence variant de 1 à 5 et pour une taille d'échantillon de 2 à 5.

Sur des points indépendants (ne présentant pas de corrélation temporelle), condition satisfaite préférentiellement par le JFV, les résultats des différentes méthodes sont en moyenne assez similaires, à l'exception du modèle des valeurs maximales. Ce dernier se montre très sensible à la présence de points aberrants et fixe alors des limites beaucoup trop hautes. Au contraire, la détermination du seuil de classification à l'aide de la variance

des moyennes des échantillons se montre plus robuste que le calcul de la variance sur la population mère ou de la moyenne des variances. Le test *Hottelling* T^2 génère de bons résultats, à la condition que la base de référence représente exactement les mesures testées. Dans le cas contraire, la valeur du test est fortement biaisée.

Enfin, le nombre de points nécessaires pour prendre une décision (taille d'un échantillon ou longueur de la séquence dépassant le seuil) impliquant un minimum de fausses alarmes et un maximum de vraies détections, est égal à 4 sur l'ensemble des essais analysés. Cette valeur est probablement spécifique aux données traitées, mais elle montre que l'augmentation du nombre de points n'est pas nécessairement un gage de meilleurs résultats.

PERSPECTIVES

Les conclusions de ces travaux de recherche montrent qu'aucun des deux indicateurs étudiés n'allie les deux qualités essentielles à la détection de l'endommagement dans des structures de génie civil. Les coefficients AR sont très sensibles à tous les changements du comportement dynamique, rendant difficile la distinction entre les effets des conditions environnementales et ceux de l'endommagement. Au contraire, le JFV présente des valeurs peu corrélées temporellement rendant plus facile l'étape de classification. Il se montre cependant moins sensible pour les paramètres de calcul utilisés.

Les difficultés rencontrées dans la détermination des paramètres de reconstruction de l'espace des phases rendent l'utilisation du JFV hasardeuse. Une étude paramétrique sur des mesures de caractéristiques différentes devrait permettre d'apporter des critères de sélection.

Compte tenu des résultats obtenus avec le JFV, l'utilisation de l'espace des phases reconstruit à partir de mesures issues d'un système stochastique semble déstructurer de façon trop importante l'information temporelle pour bénéficier d'une bonne sensibilité à l'endommagement. Avant de rejeter définitivement cette approche, il convient de travailler sur d'autres cas d'études et sur la possibilité d'automatisation du choix des paramètres de reconstruction optimaux lorsque le système présente une composante stochastique forte.

La recherche de l'indicateur idéal passe probablement par une exploitation plus importante de l'information entre des couples de capteurs, de façon à s'affranchir d'une partie de la variabilité due à l'excitation. Si, comme le supposent [Worden et Manson \(2007\)](#), il n'est pas possible de définir un indicateur sensible seulement à l'endommagement, alors les efforts doivent se porter sur la normalisation et la modélisation statistique.

Les gains attendus avec des méthodes de normalisation non linéaires sont faibles, compte tenu de la distribution des points dans l'espace des indicateurs vectoriels testés. L'effort doit alors porter sur la modélisation statistique.

Plusieurs pistes sont à étudier. Dans le cas où la classification est réalisée à partir d'une séquence de s dépassements, la limite de contrôle est la valeur correspondant à la probabilité $(1 - \alpha^{\tau/s})$. Dans ce travail, les points successifs sont supposés indépendants, correspondant à une valeur $\tau = 1$.

Dans le cas d'une corrélation totale des points successifs, τ doit être égal à s . Il est donc nécessaire de travailler sur la détermination d'une valeur de τ intermédiaire s'appuyant sur le degré de dépendance des points, de façon à réduire le taux de fausses alarmes.

Réduire la corrélation entre les points peut se faire par l'ajustement d'un modèle AR ou par la méthode de la moyenne mobile exponentiellement pondérée (EWMA) (Montgomery 2005). Une solution pourrait également résider dans la création d'une limite plus locale à l'aide d'un ajustement sur une fenêtre glissante. La taille de la fenêtre doit être très inférieure aux cycles climatiques pour augmenter l'indépendance des points. Cependant, le problème de la détection d'une déviation lente devrait être solutionné. Dans tous les cas, la prise en compte de la corrélation passe nécessairement par l'augmentation du nombre de points dans la base de classification. Une autre possibilité réside dans l'utilisation d'une base d'apprentissage mobile formée, par exemple, par les k dernières mesures. La détection de nouveauté se focalise alors sur des changements assez soudains du comportement.

L'analyse de données réelles collectées sur un ouvrage en service s'avère également indispensable. Idéalement, il faudrait disposer d'une base de mesures collectées durant plusieurs années. Une telle base de données permet notamment de quantifier la présence de nouveauté d'une année sur l'autre, causée par exemple par une forte canicule ou un hiver plus rigoureux. Au moyen de mesures réalisées toutes les heures, la dispersion imposée par les cycles climatiques quotidiens (non envisagés dans ce travail de thèse) peut être analysée. L'autre avantage d'une fréquence de suivi élevée réside dans la possibilité de séparer les mesures sous sollicitations intenses (trafic routier aux heures de grande affluence) des mesures réalisées en périodes calmes (bruit ambiant en milieu de la nuit). Ce type de données permettrait de mieux appréhender la nature et la variabilité de l'excitation afin d'éprouver le caractère réaliste des types de sollicitations générées sur des maquettes de laboratoire.

Enfin, les nombreux essais réalisés dans le cadre de la thèse, notamment sur la maquette de structure en bois, ont permis de générer plusieurs bases de données de signatures vibratoires associées à :

- des séquences diverses d'endommagement réversible,
- des variations sévères des conditions environnementales,
- des sollicitations dynamiques de natures diverses.

Afin de valoriser davantage le travail expérimental, les données et les détails expérimentaux seront rendus accessibles sur demande² pour l'ensemble de la communauté scientifique (des universités ou des organismes de recherche d'état) impliquée dans le développement du suivi vibratoire des structures de génie civil. A l'instar des données relatives à la maquette métallique du LANL, les essais réalisés dans le cadre de cette thèse pourront ainsi permettre d'éprouver d'autres développements scientifiques conduits par ailleurs. Enfin, un projet de recherche collaborative pourra être proposé par le biais d'un benchmark s'appuyant sur les essais réalisés.

² *Antoine.clement@insa-toulouse.fr* ou *Stephane.laurens@insa-toulouse.fr*

ANNEXES

A

SOMMAIRE

A.1	TABLEAU CHARGEMENT DE LA POUTRE EN BÉTON FIBRÉ	174
A.2	RÉSULTATS MAQUETTE EN BOIS	175
A.2.1	Coefficients AR	175
A.2.2	Composantes JFV	181
A.2.3	Corrélation croisée	187
A.3	MODÉLISATION STATISTIQUE	189

A.1 TABLEAU CHARGEMENT DE LA POUTRE EN BÉTON FIBRÉ

N° séquence	Cycles supplémentaires		Observations
	Nbre Cycles	Force [kN]	
44	2	0.8	
45	3	0.8	
46	3	0.8	
47	3	0.8	
48	3	0.8	
49	3	0.8	
50	3	0.8	
51	6	0.8	
52	6	0.8	
53	6	0.8	
54	6	0.8	
55	6	0.8	
56	6	0.8	
57	10	0.8	
58	10	0.8	
59	10	0.8	
60	10	0.8	
61	4	1	Température uniforme
62	4	1	
63	4	1	
64	3	1.1	
65	3	1.1	
66	1	1.2	Première fissure visible
67	2	1.2	
68	2	1.2	
69	3	1.3	
70	3	1.3	
71	3	1.4	
72	3	1.45	
73	3	1.6	
74	3	1.6	
75	1	1.65	
76	3	1.8	
77	0	0	
78	3	1.9	
79	3	2	Gros craquement
80	-	-	Pas de mesure
81	3	2.1	
82	3	2.2	Craquement
83	3	2.3	Craquement
84	3	2.4	3 gros craquements
85	1	2.4	
86	1	2.2	Plastification

TAB. A.1 *Historique de chargement*

A.2 RÉSULTATS MAQUETTE EN BOIS

A.2.1 Coefficients AR

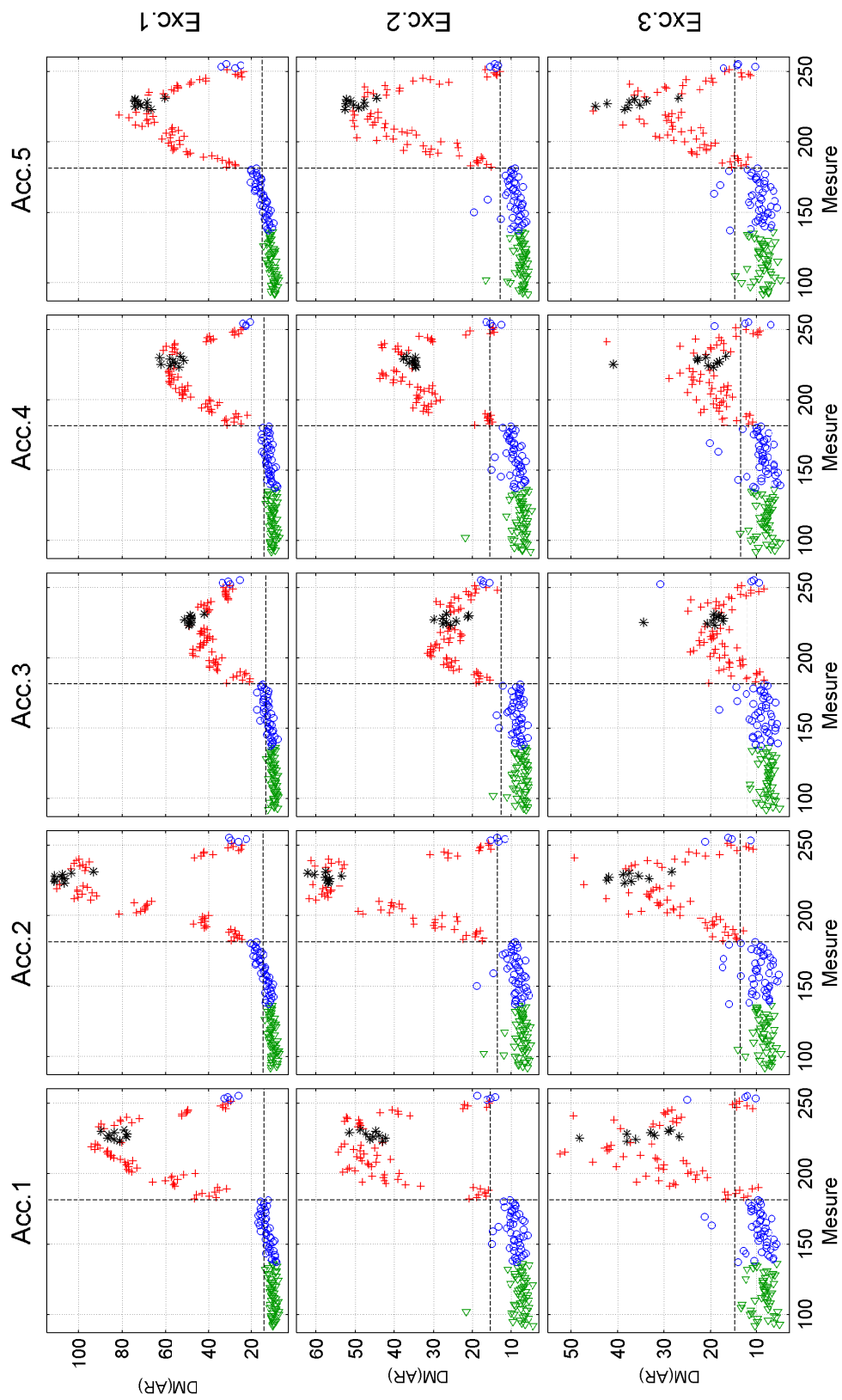


FIG. A.1 Distance de Mahalanobis calculée à partir des coefficients AR, en fonction de la mesure, essai 22, endommagement A_2 ; \circ base de test ; $+$ données endommagées ; $*$ endommagement maximum.

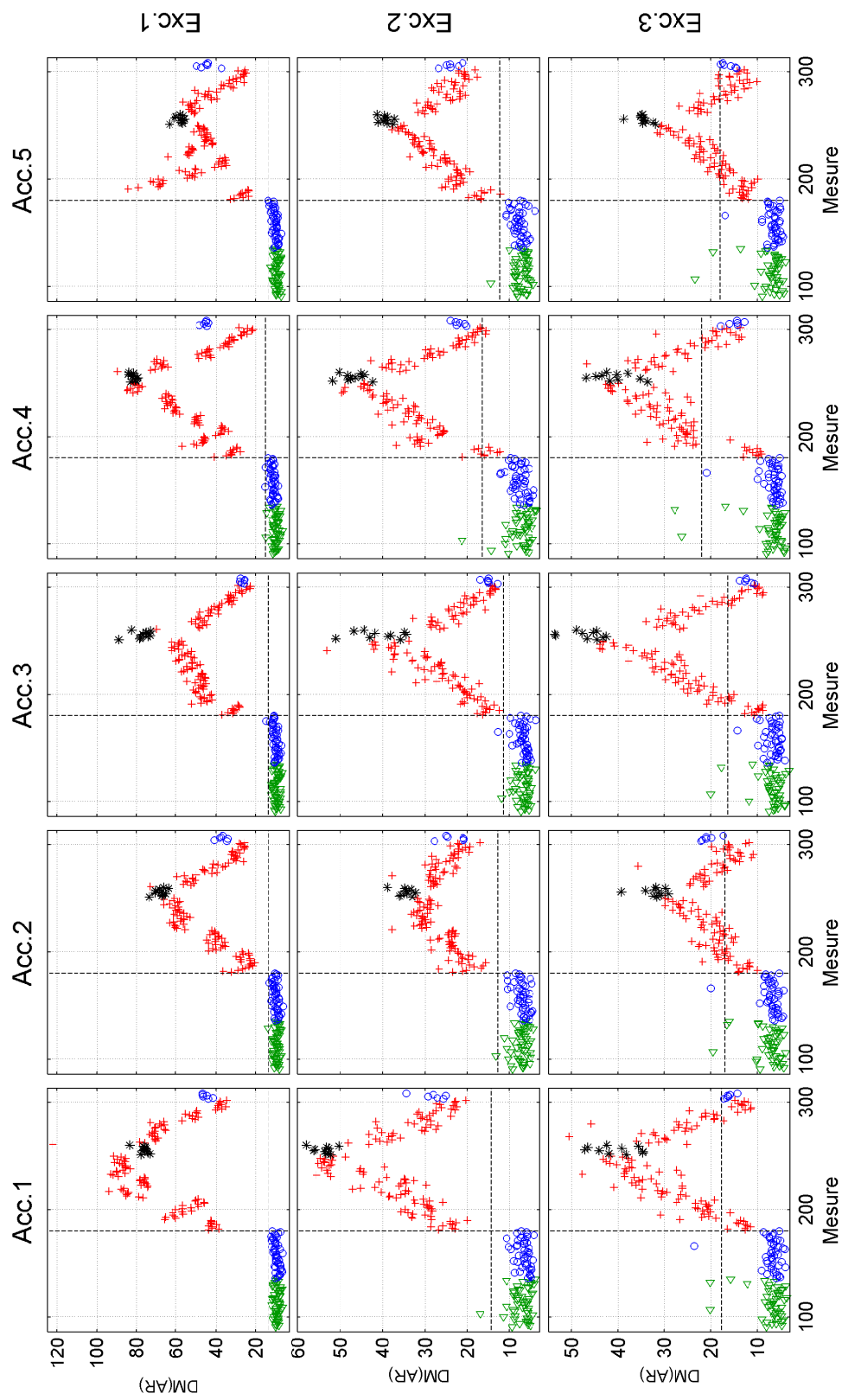


FIG. A.2 Distance de Mahalanobis calculée à partir des coefficients AR , en fonction de la mesure, essai 24, endommagement D_{24} ; \circ base de test ; \triangle données endommagées ; $*$ endommagement maximum.

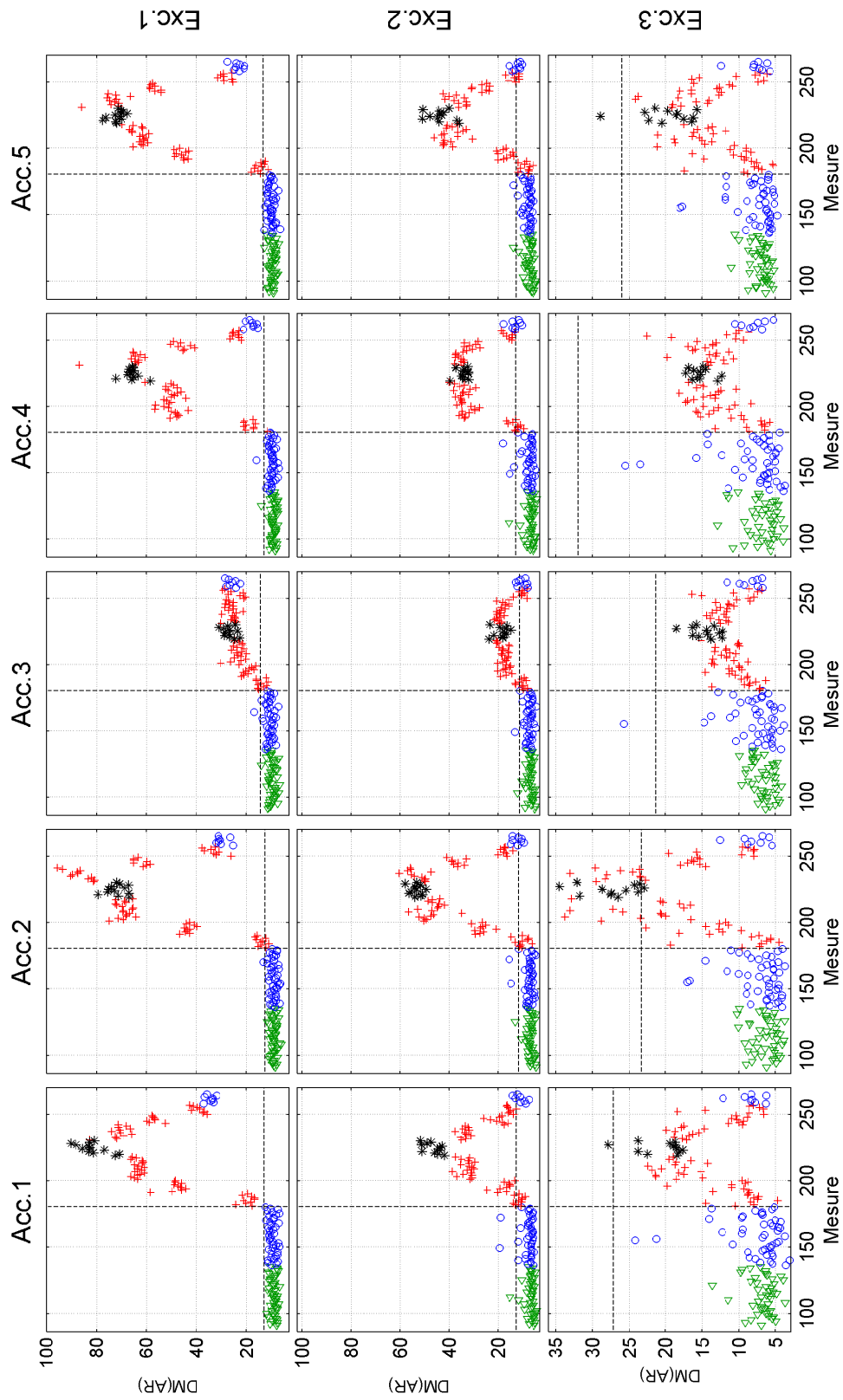


FIG. A.3 Distance de Mahalanobis calculée à partir des coefficients AR, en fonction de la mesure, essai 25, endommagement A1 ; o base de test ; + données endommagées ; * endommagement maximum.

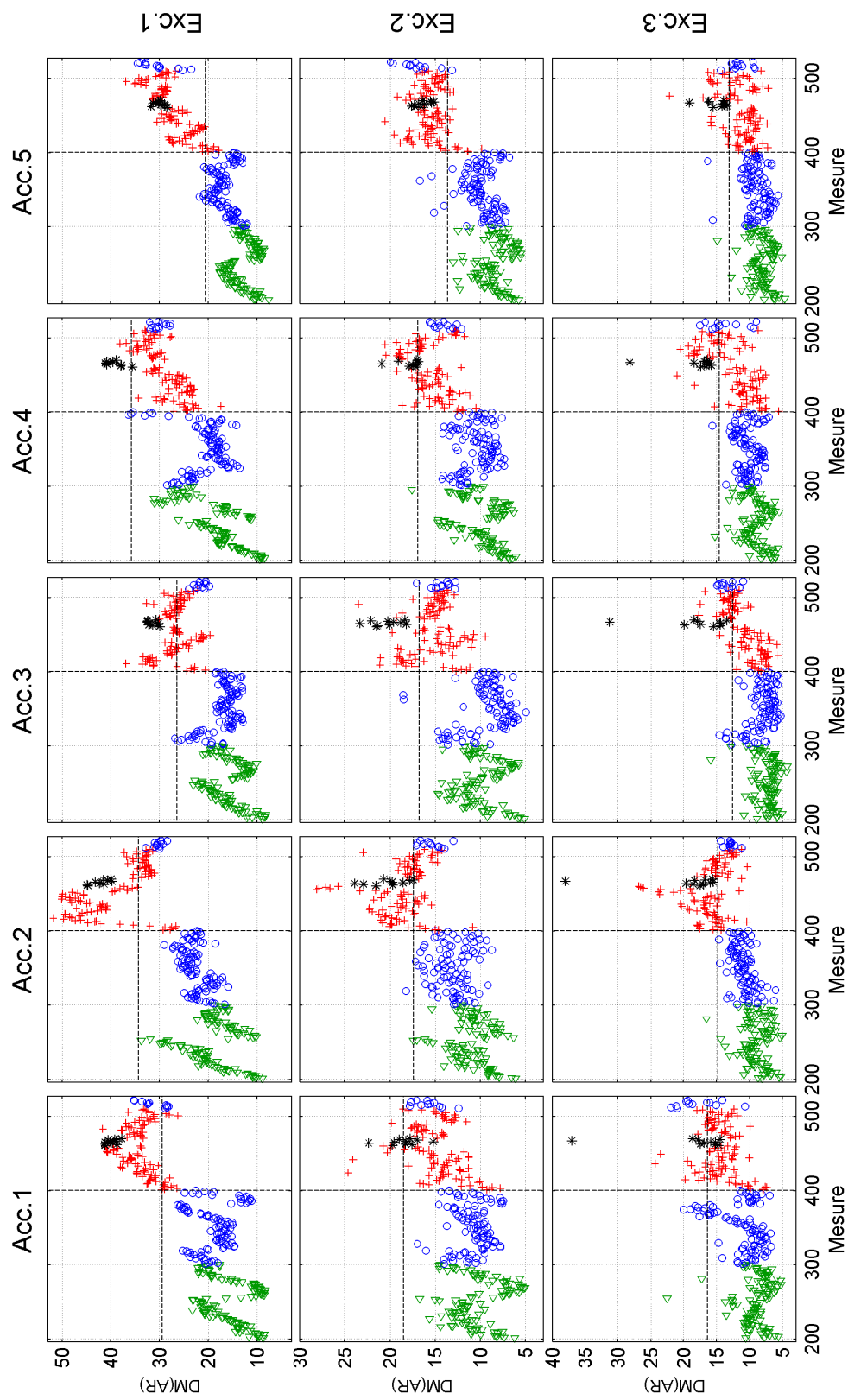


FIG. A.4 Distance de Mahalanobis calculée à partir des coefficients AR, en fonction de la mesure, essai 34, endommagement D_{24} ; \circ base de test ; $+$ données endommagées ; * endommagement maximum.

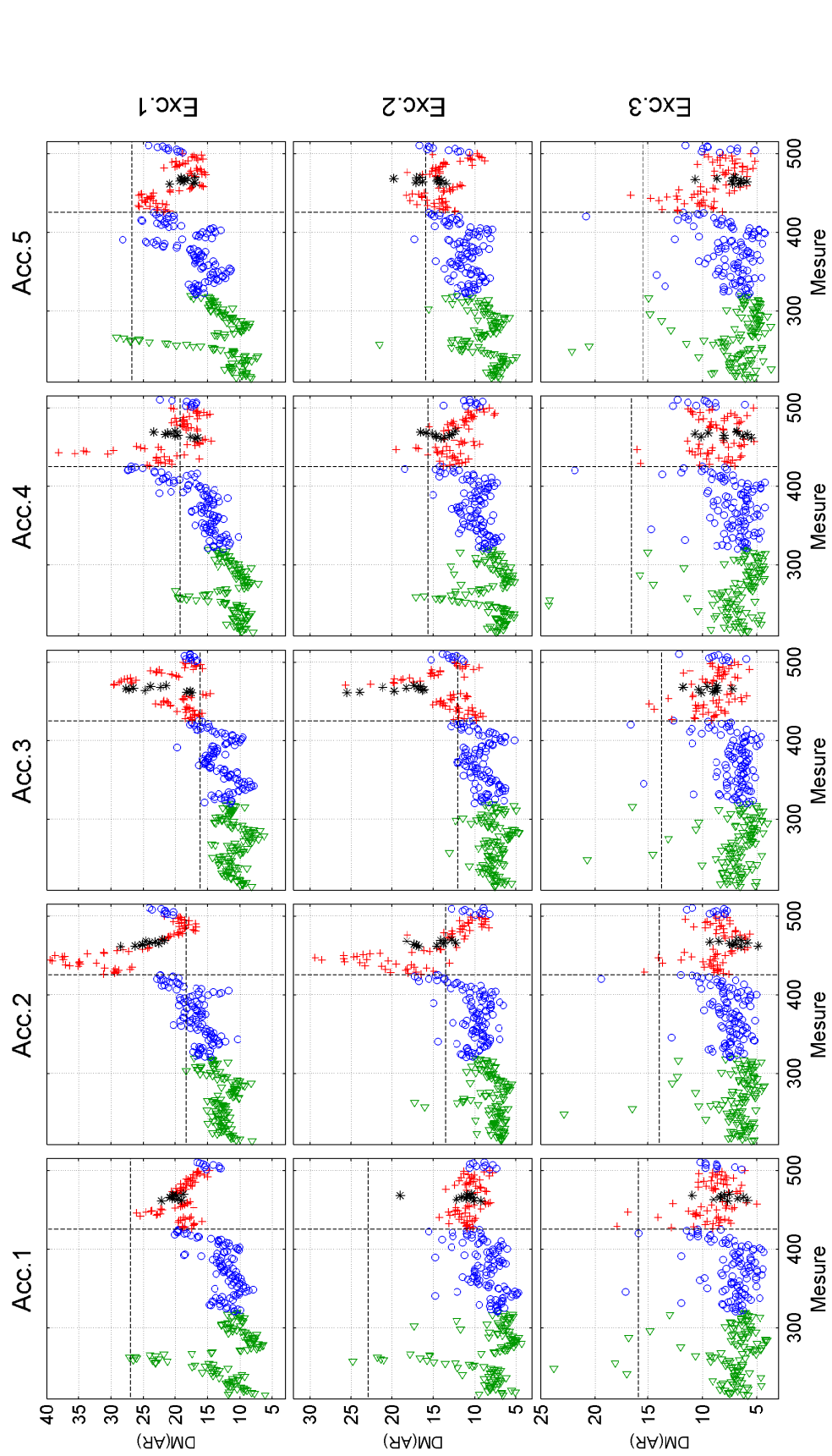


FIG. A.5 Distance de Mahalanobis calculée à partir des coefficients AR , en fonction de la mesure, essai 35, endommagement A_3 ; \circ base de test ; $+$ données endommagées ; * endommagement maximum.

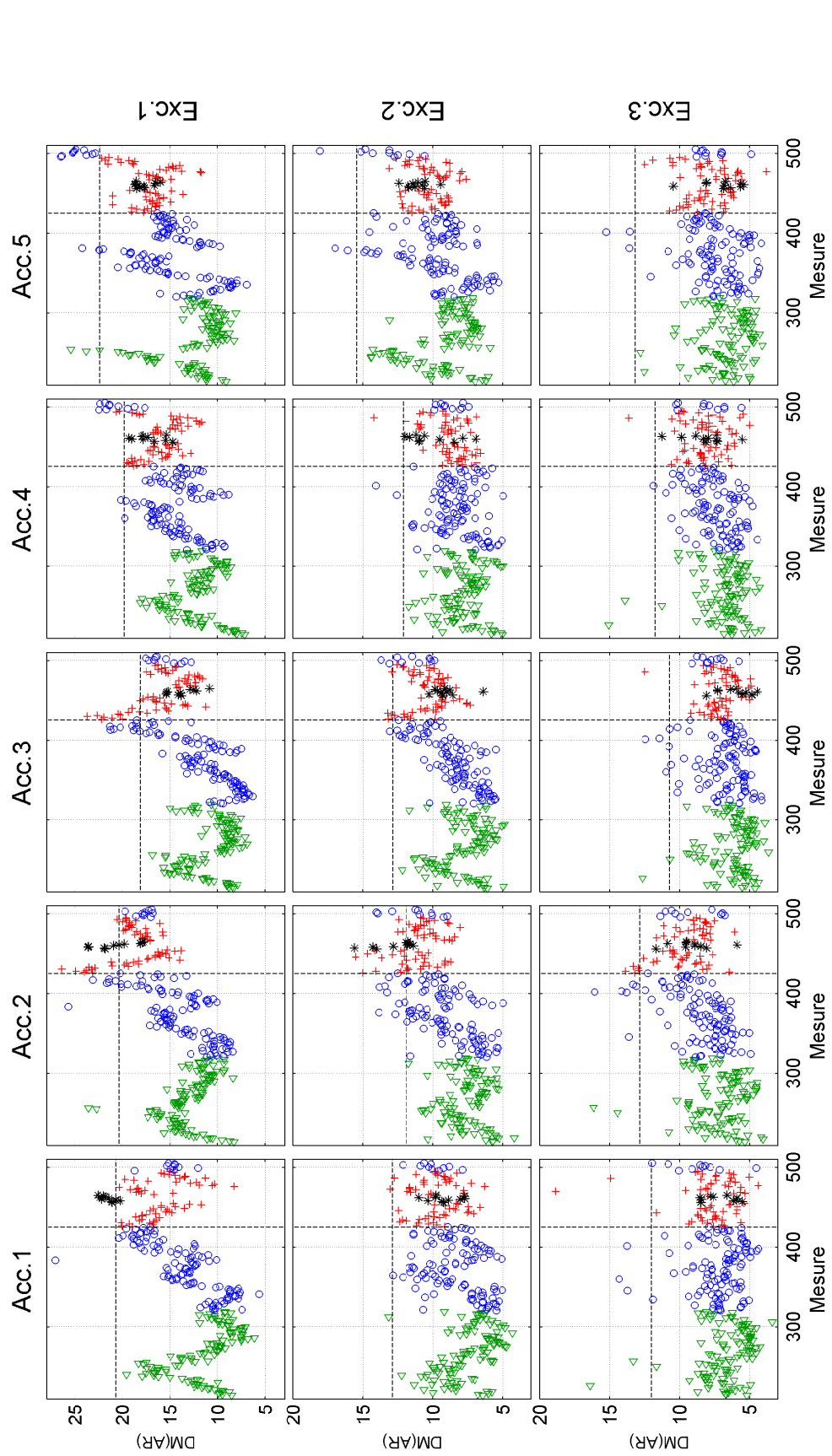


FIG. A.6 Distance de Mahalanobis calculée à partir des coefficients AR , en fonction de la mesure, essai 36, endommagement A_1 ; \circ base de test ; $+$ données endommagées ; * endommagement maximum.

A.2.2 Composantes jfv

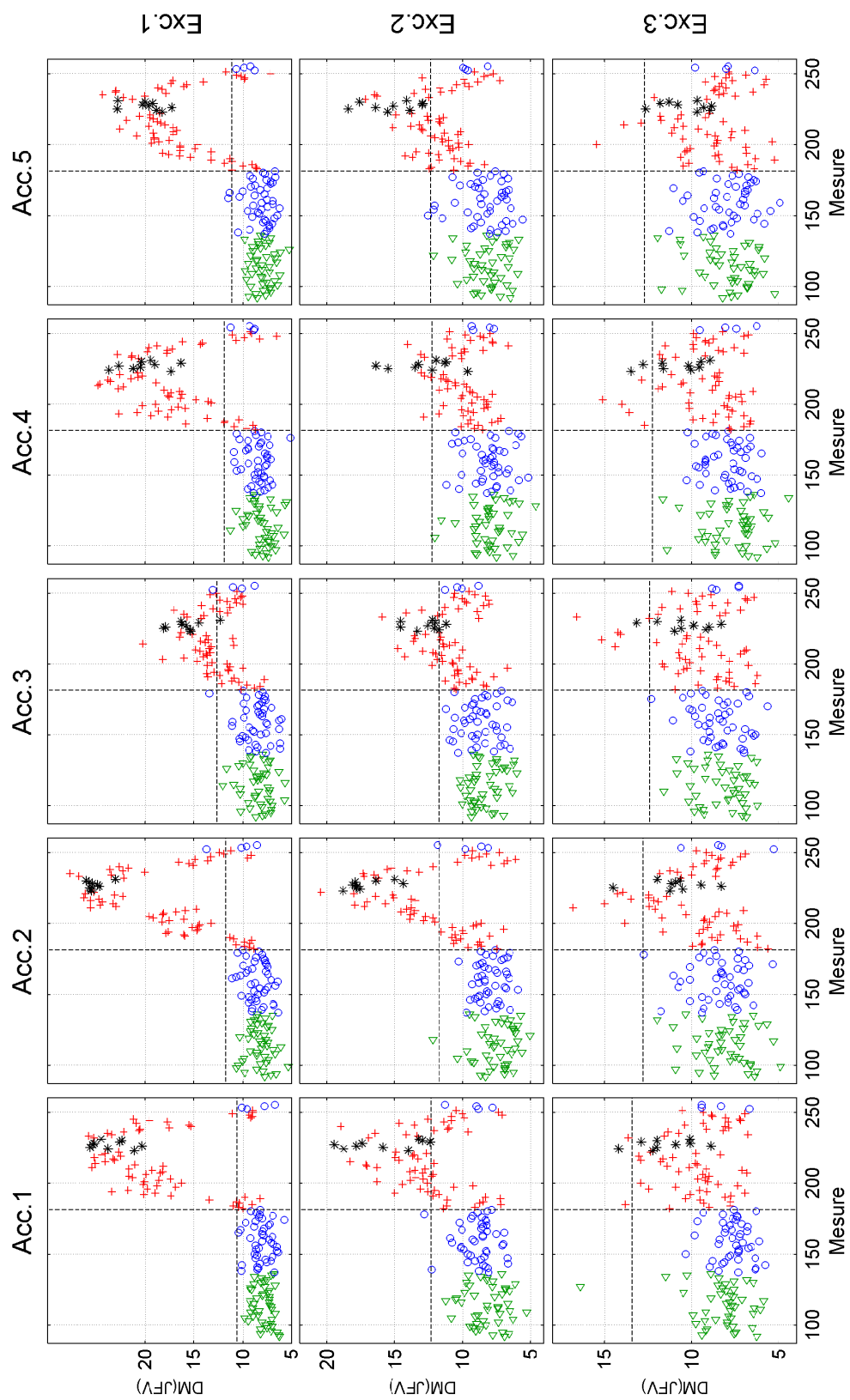


FIG. A.7 Distance de Mahalanobis calculée à partir des composantes du JFV, en fonction de la mesure, *essai 22*, endommagement A2 ; o base de test ; + données endommagées ; * endommagement maximum.

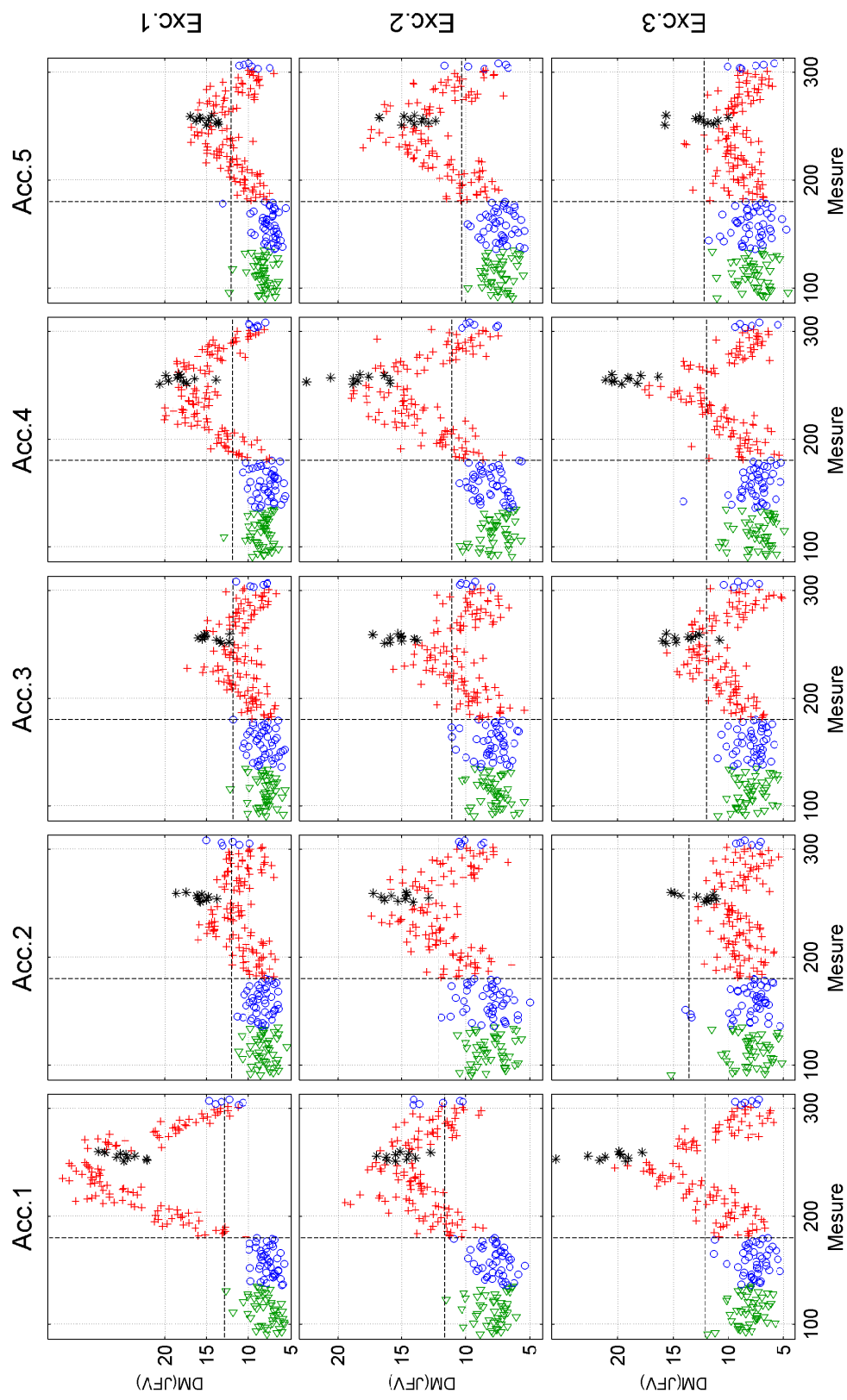


FIG. A.8 Distance de Mahalanobis calculée à partir des composantes du JFV, en fonction de la mesure, essai 24, endommagement D_{24} ; \circ base de test ; $+$ données endommagées ; $*$ endommagement maximum.

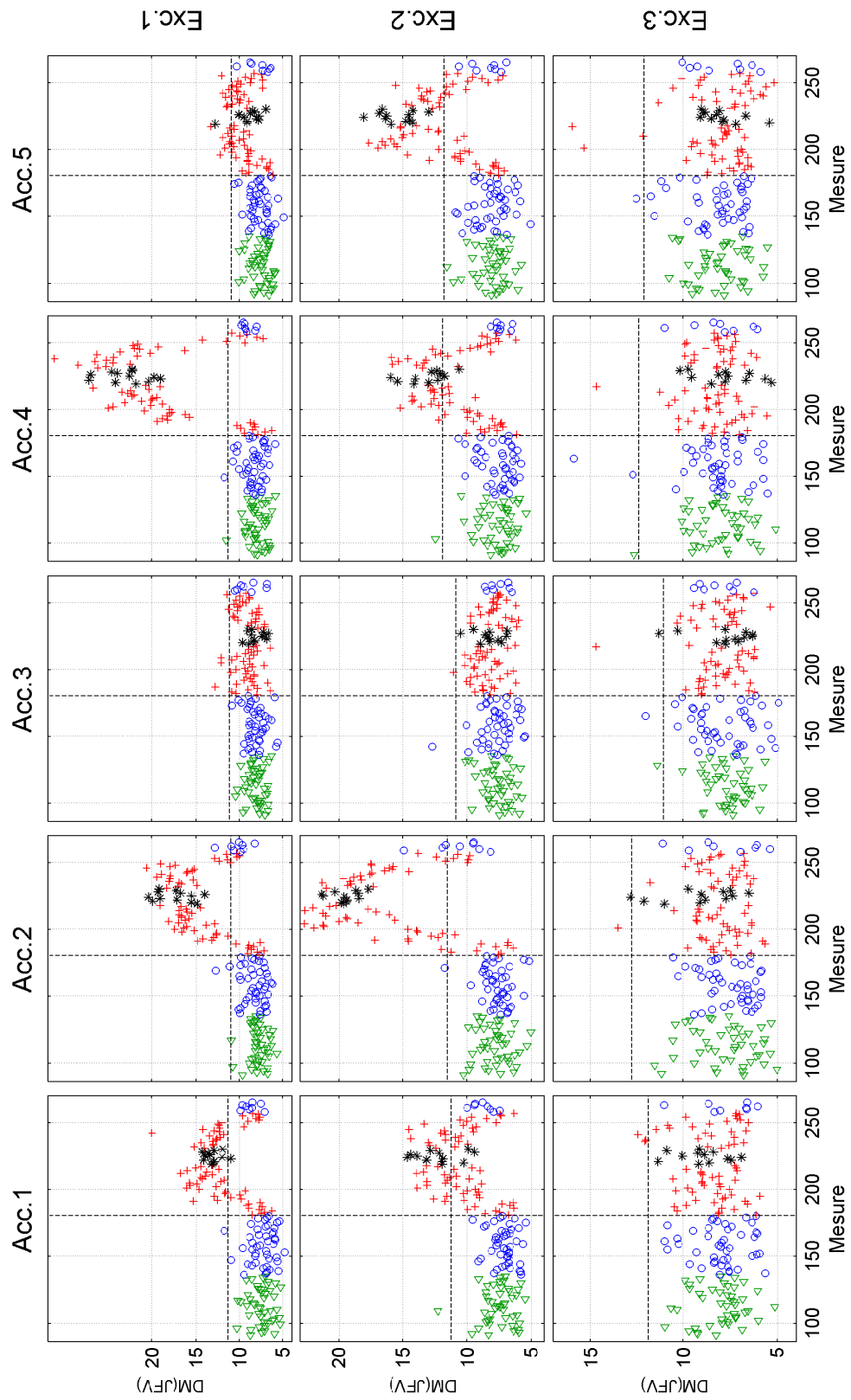


FIG. A.9 Distance de Mahalanobis calculée à partir des composantes du JFV, en fonction de la mesure, *essai 25*, endommagement A_1 ; \circ base de test ; $+$ données endommagées ; $*$ endommagement maximum.

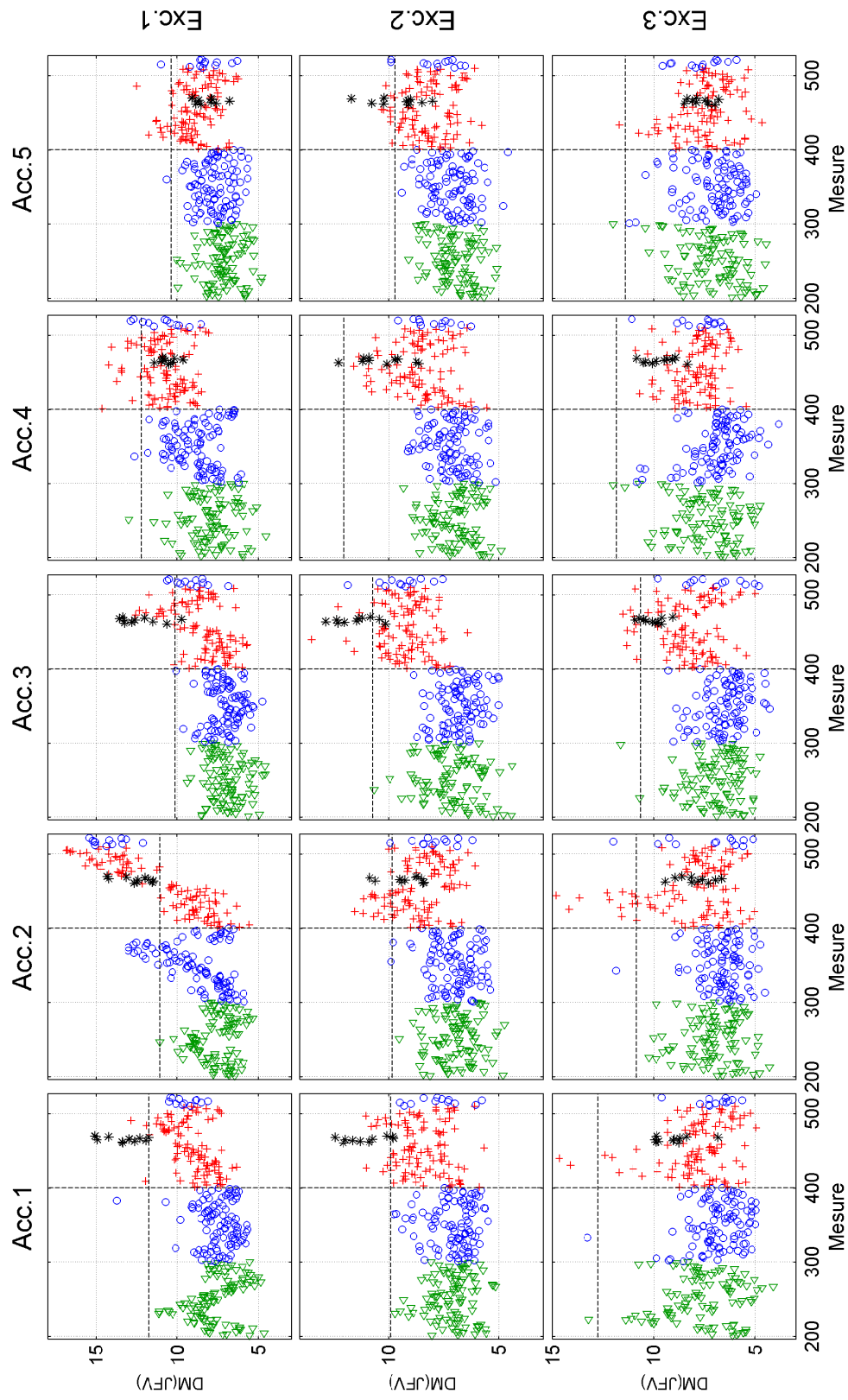


FIG. A.10 Distance de Mahalanobis calculée à partir des composantes du JFV, en fonction de la mesure, essai 34, endommagement D_{24} ; \circ base de test ; $+$ données endommagées ; $*$ endommagement maximum.

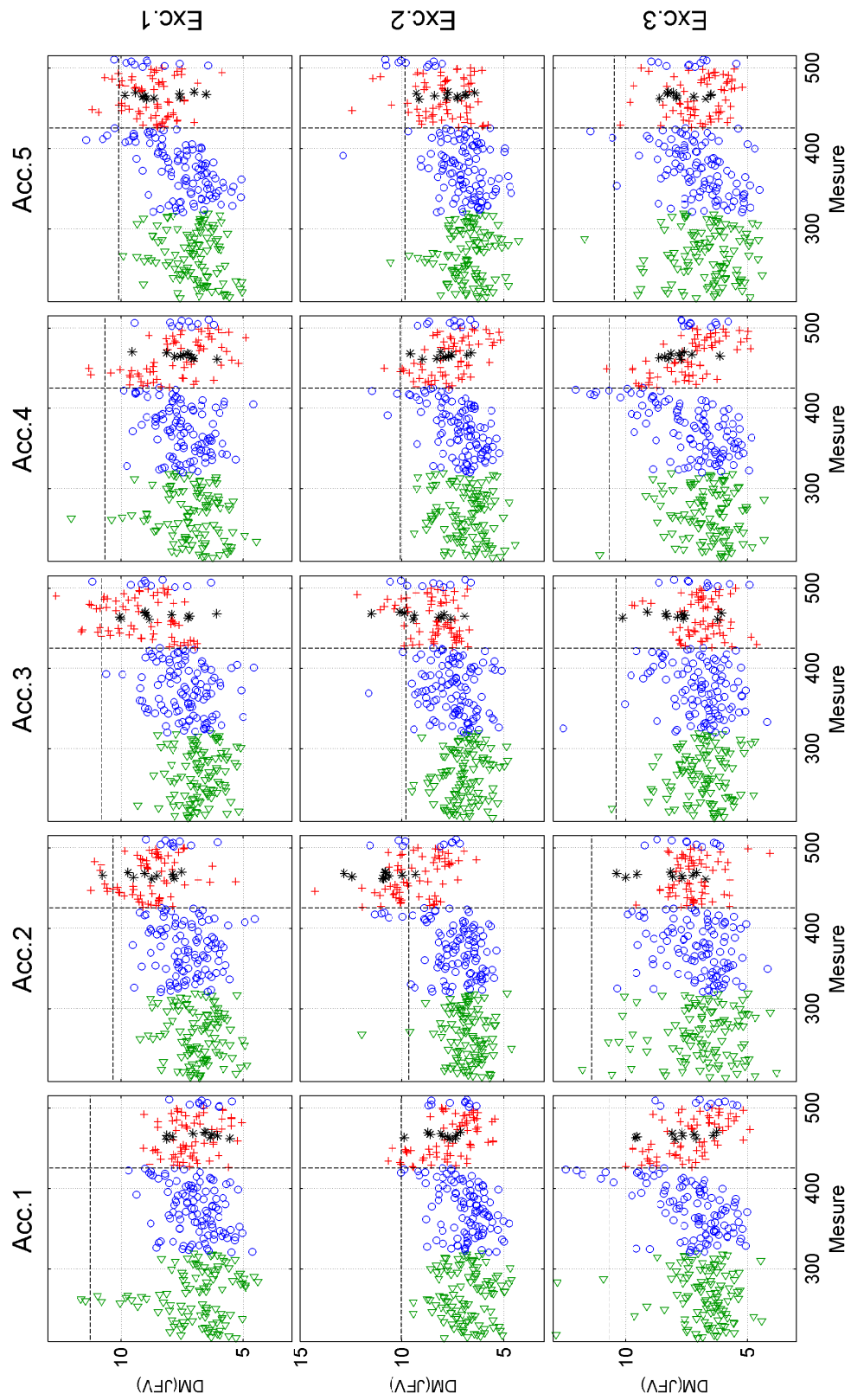


FIG. A.11 Distance de Mahalanobis calculée à partir des composantes du JFV, en fonction de la mesure, essai 35, endommagement A_3 ; \circ base de test ; $+$ données endommagées ; $*$ endommagement maximum.

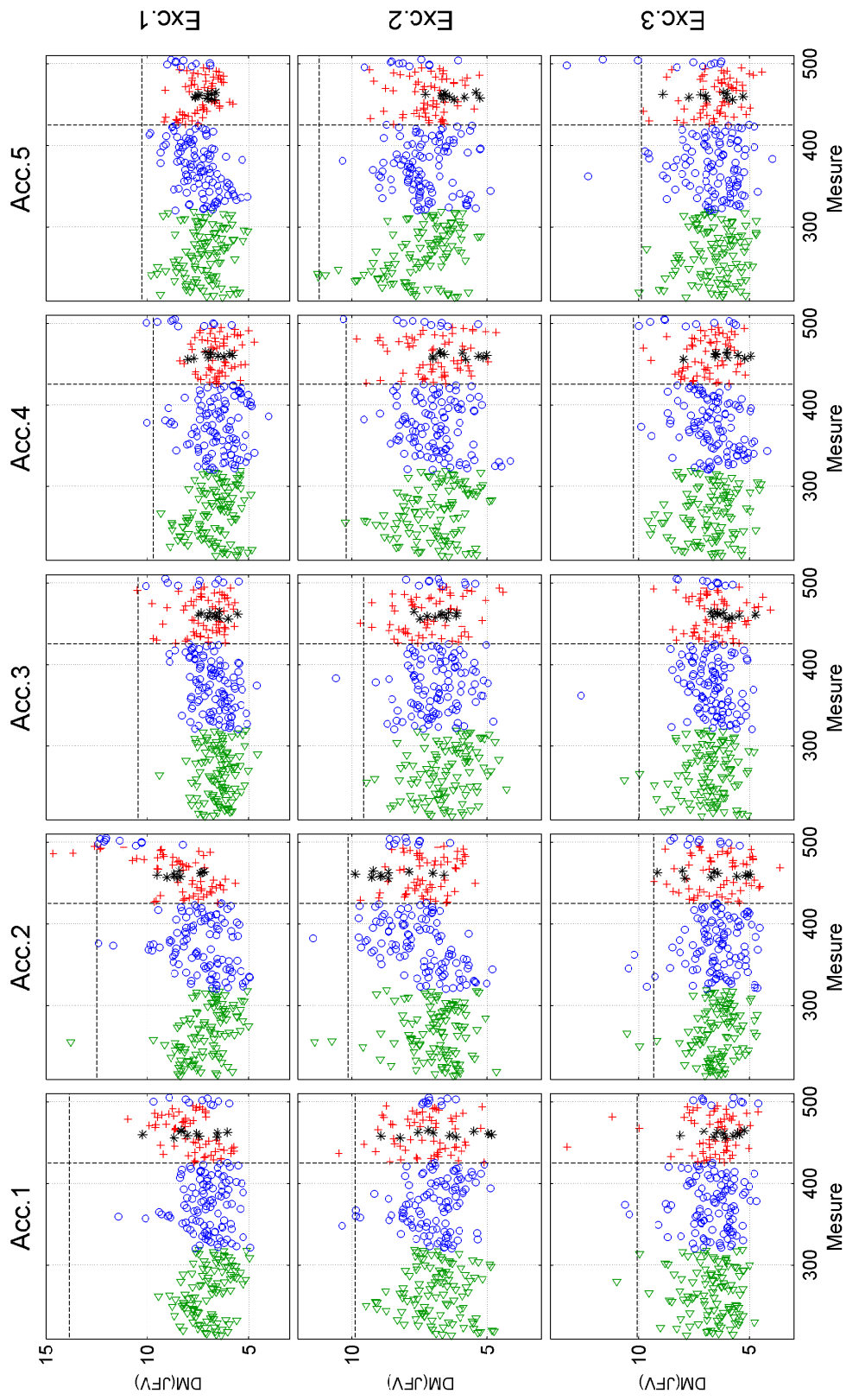


FIG. A.12 Distance de Mahalanobis calculée à partir des composantes du JFV, en fonction de la mesure, essai 36, endommagement A_1 ; ○ base de test ; + données endommagées ; * endommagement maximum.

A.2.3 Corrélation croisée

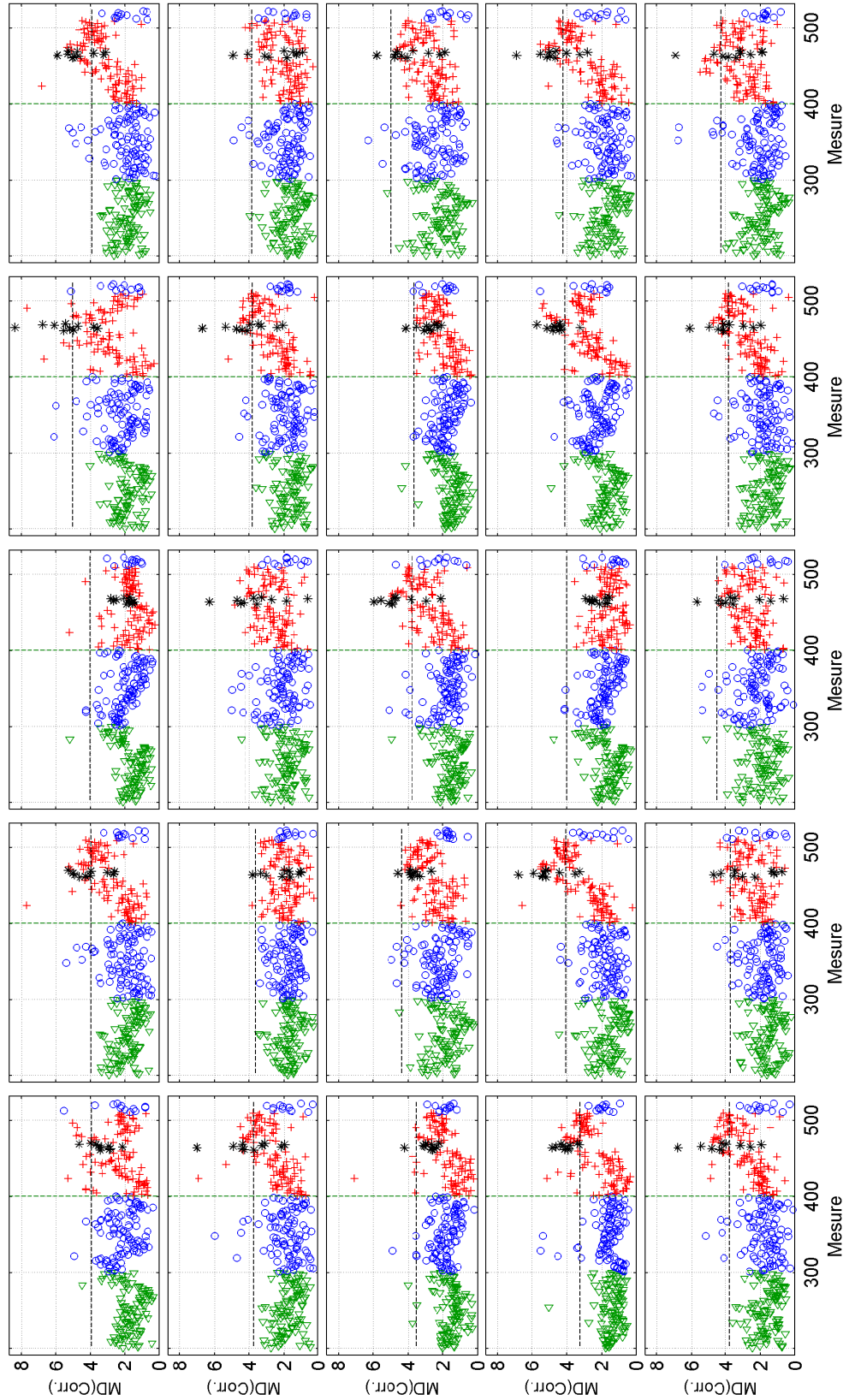


FIG. A.13 Distance de Mahalanobis calculée à partir des fonctions d'auto-corrélation croisées, en fonction de la mesure, *essai 34, excitation 2, endommagement D₂₄*; \circ base de test; $+$ données endommagées; $*$ endommagement maximum.

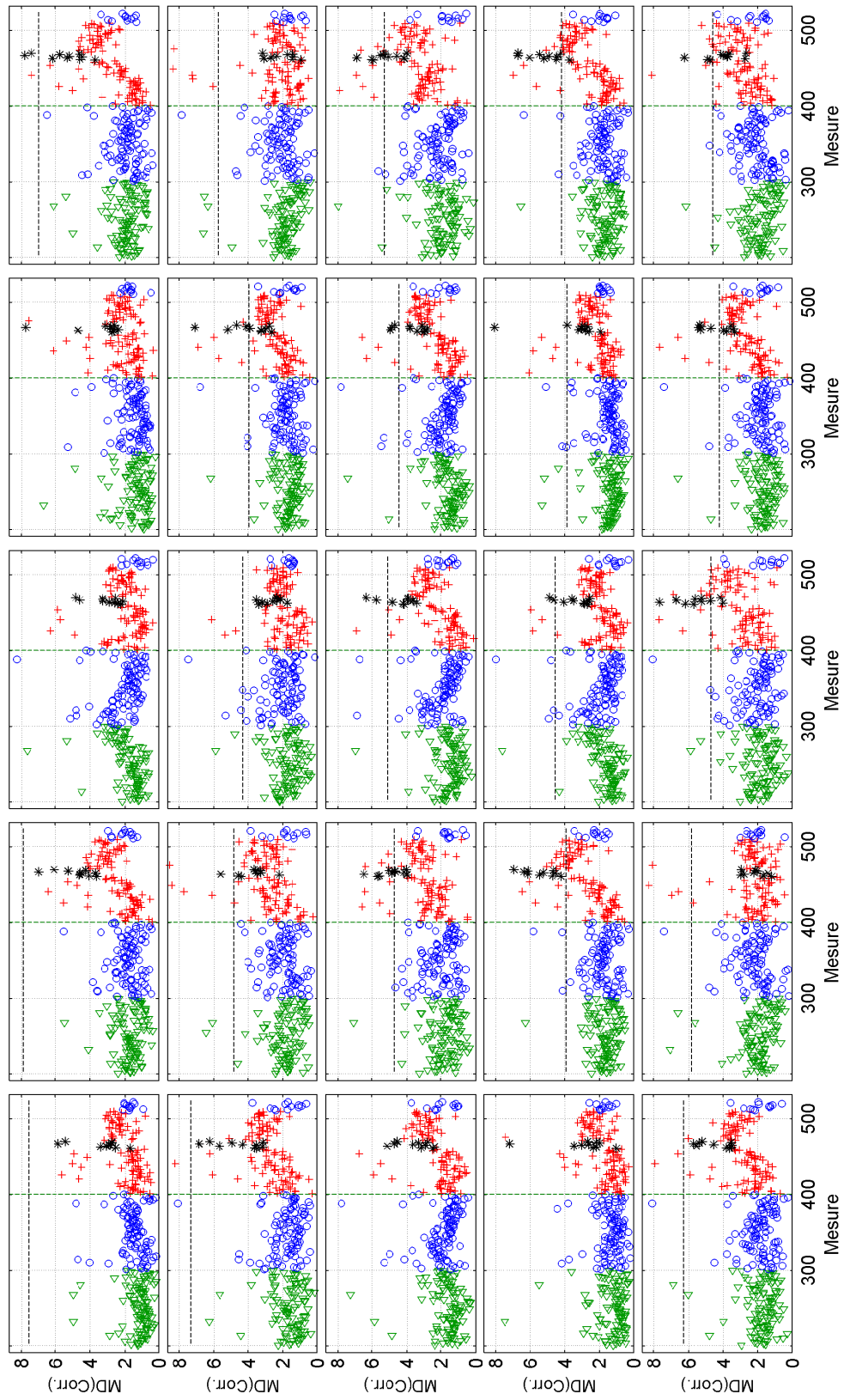


FIG. A.14 Distance de Mahalanobis calculée à partir des fonctions d'auto-corrélation croisées, en fonction de la mesure, *essai 34, excitation 3, endomagement 34*; \circ base de test; $+$ données endommagées; * endommagées maximum.

A.3 MODÉLISATION STATISTIQUE

Rang	J1		J2		J3		J4		J5	
	Méth.	FA/VD	Méth.	FA/VD	Méth.	FA/VD	Méth.	FA/VD	Méth.	FA/VD
$s = 1$										
1	3 ₁₁	4/49	1 ₁₁	4/39	3 ₁₁	4/52	1 ₁₁	13/89	3 ₁₁	5/39
2	1 ₁₁	4/48	3 ₁₁	0/19	1 ₁₁	2/27	2 ₁₁	1/64	1 ₁₁	1/5
3	2 ₁₁	0/22	2 ₁₁	0/10	2 ₁₁	0/15	3 ₁₁	100/100	2 ₁₁	0/2
$s = 2$										
1	3 ₁₂	10/73	5 ₂₁	8/66	3 ₁₂	2/68	6 ₂₁	7/93	3 ₁₂	5/52
2	5 ₂₁	8/69	6 ₂₁	4/63	5 ₂₁	2/58	7 ₂₁	4/85	5 ₂₁	4/44
3	6 ₂₁	12/70	1 ₁₂	4/55	6 ₂₁	2/56	4 ₂₁	25/100	2 ₁₂	4/35
4	1 ₁₂	4/61	4 ₂₁	4/53	4 ₂₁	0/55	1 ₁₂	30/99	4 ₂₁	2/11
5	2 ₁₂	4/61	7 ₂₁	8/50	1 ₁₂	0/47	2 ₁₂	31/99	6 ₂₁	2/11
6	4 ₂₁	4/61	2 ₁₂	4/49	2 ₁₂	0/47	5 ₂₁	75/100	1 ₁₂	0/9
7	7 ₂₁	0/34	3 ₁₂	4/45	7 ₂₁	0/29	3 ₁₂	100/100	7 ₂₁	0/4
$s = 3$										
1	6 ₃₁	0/80	1 ₁₃	6/70	3 ₁₃	23/81	7 ₃₁	3/94	2 ₁₃	11/65
2	5 ₃₁	0/78	2 ₁₃	6/70	2 ₁₃	0/63	6 ₃₁	8/97	3 ₁₃	8/62
3	2 ₁₃	4/74	3 ₁₃	0/66	4 ₃₁	3/56	4 ₃₁	32/100	5 ₃₁	3/44
4	4 ₃₁	0/73	5 ₃₁	0/64	5 ₃₁	3/56	1 ₁₃	38/99	1 ₁₃	0/35
5	3 ₁₃	10/74	6 ₃₁	0/64	6 ₃₁	3/56	2 ₁₃	50/99	4 ₃₁	0/14
6	1 ₁₃	4/70	4 ₃₁	0/60	1 ₁₃	0/55	5 ₃₁	78/100	6 ₃₁	0/14
7	7 ₃₁	0/0	7 ₃₁	6/56	7 ₃₁	0/33	3 ₁₃	100/100	7 ₃₁	0/0
$s = 4$										
1	6 ₄₁	0/90	5 ₂₂	8/84	2 ₁₄	16/83	6 ₄₁	11/100	3 ₁₄	8/66
2	5 ₄₁	0/80	6 ₂₂	8/82	3 ₁₄	19/85	7 ₄₁	4/89	5 ₂₂	0/55
3	5 ₂₂	12/82	4 ₂₂	8/79	1 ₁₄	0/63	7 ₂₂	32/100	5 ₄₁	4/48
4	6 ₂₂	12/82	5 ₄₁	15/79	4 ₂₂	0/60	6 ₂₂	34/100	1 ₁₄	0/40
5	2 ₁₄	10/78	4 ₄₁	8/74	6 ₂₂	0/60	4 ₂₂	45/100	2 ₁₄	61/79
6	3 ₁₄	16/80	6 ₄₁	8/74	5 ₂₂	4/60	4 ₄₁	50/100	4 ₂₂	0/27
7	4 ₂₂	0/74	1 ₁₄	0/68	4 ₄₁	0/59	1 ₁₄	55/99	6 ₂₂	0/27
8	4 ₄₁	0/73	2 ₁₄	0/68	6 ₄₁	0/59	2 ₁₄	62/99	7 ₂₂	0/18
9	7 ₂₂	0/72	7 ₄₁	15/68	5 ₄₁	7/59	5 ₄₁	79/100	4 ₄₁	0/15
10	1 ₁₄	0/70	3 ₁₄	0/62	7 ₂₂	0/42	5 ₂₂	84/100	6 ₄₁	0/15
11	7 ₄₁	0/3	7 ₂₂	0/61	7 ₄₁	0/0	3 ₁₄	100/100	7 ₄₁	0/0
$s = 5$										
1	4 ₅₁	0/83	2 ₁₅	0/73	3 ₁₅	21/91	7 ₅₁	5/100	3 ₁₅	13/66
2	5 ₅₁	0/83	4 ₅₁	10/73	2 ₁₅	21/85	6 ₅₁	9/95	5 ₅₁	5/50
3	2 ₁₅	12/80	5 ₅₁	10/73	5 ₅₁	0/68	4 ₅₁	50/100	1 ₁₅	0/47
4	1 ₁₅	10/73	6 ₅₁	0/67	1 ₁₅	4/66	1 ₁₅	64/99	2 ₁₅	75/98
5	6 ₅₁	0/71	7 ₅₁	20/73	4 ₅₁	0/64	2 ₁₅	74/99	4 ₅₁	0/23
6	3 ₁₅	22/80	1 ₁₅	0/62	6 ₅₁	0/64	5 ₅₁	86/100	6 ₅₁	0/18
7	7 ₅₁	0/21	3 ₁₅	0/62	7 ₅₁	0/0	3 ₁₅	100/100	7 ₅₁	0/0

TAB. A.2 Pourcentage de fausse alarme et de vraie détection pour toutes les méthodes; Dans la rubrique Méth., 6₄₁ signifie, méthode 6, $N_e = 4$ et $s = 1$.

BIBLIOGRAPHIE

- H.D.I. Abarbanel, R. Brown, J.J. Sidorowich, et L.S. Tsimring. The analysis of observed chaotic data in physical systems. *Rev. Mod. Phys.*, 65(4) : 1331–1392, Octobre 1993. (Cité pages 48, 63 et 66.)
- A. Alvandi. *Contribution à l'utilisation pratique de l'évaluation dynamique pour la détection d'endommagements dans les ponts*. PhD thesis, Ecole Nationale Des Ponts Et Chaussées, October 2003. (Cité pages 10 et 20.)
- A. Alvandi, J. Bastien, E. Gregoire, et M. Jolin. Bridge integrity assessment by continuous wavelet transforms. *International Journal of Structural Stability and Dynamics*, 9(1) :11–43, 2009. (Cité page 21.)
- M. Basseville, A. Benveniste, M. Goursat, L. Hermans, L. Mevel, et H. Van Der Auweraer. Output-only subspace-based structural identification : From theory to industrial testing practice. *Journal of Dynamic Systems, Measurement, and Control*, 123(4) :668–676, 2001. (Cité pages 10 et 12.)
- L. Bornn, C.R. Farrar, et G. Park. Damage detection in initially nonlinear systems. *International Journal of Engineering Science*, 48(10) :909–920, 2010. Structural Health Monitoring in the Light of Inverse Problems of Mechanics. (Cité pages 36 et 44.)
- L. Bornn, C.R. Farrar, G. Park, et Farinholt K. Structural health monitoring with autoregressive support vector machines. *Journal of Vibration and Acoustics*, 131(2) :021004, 2009. (Cité page 19.)
- G.E.P. Box et G.M. Jenkins. *Time series analysis : forecasting and control*. Prentice Hall PTR Upper Saddle River, NJ, USA, 1994. (Cité page 61.)
- P. J. Brockwell et R. A. Davis. *Time Series : Theory and Methods*. Springer, 1991. (Cité page 17.)
- R. Brown, P. Bryant, et H.D.I. Abarbanel. Computing the lyapunov spectrum of a dynamical system from an observed time series. *Phys. Rev. A*, 43(6) :2787, Mars 1991. (Cité pages 47, 51, 65 et 66.)
- C.J.C. Burges. A tutorial on support vector machines for pattern recognition. *Data Mining and Knowledge Discovery*, 2 :121–167, 1998. (Cité page 31.)
- E.P. Carden et P. Fanning. Vibration based condition monitoring : a review. *Structural Health Monitoring*, 3(4) :355, 2004. (Cité pages 10 et 20.)
- F. Casciati et S. Casciati. Structural health monitoring by lyapunov exponents of non-linear time series. *Structural Control and Health Monitoring*, 13(1) :132–146, 2006. (Cité pages 48 et 51.)

- E. Castillo, A.S. Hadi, J.M.S. Alegría, et N. Balakrishnan. *Extreme value and related models with applications in engineering and science*. Wiley, 2005. (Cité page 37.)
- F. Necati Catbas et A. Emin Aktan. Condition and damage assessment : Issues and some promising indices. *Journal of Structural Engineering*, 128 (8) :1026–1036, 2002. (Cité page 11.)
- V. Chandola, A. Banerjee, et V. Kumar. Anomaly detection : A survey. *ACM Computing Surveys (CSUR)*, 41(3) :15, 2009. (Cité page 16.)
- V. Cherkassky et Y. Ma. Practical selection of svm parameters and noise estimation for svm regression. *Neural Networks*, 17(1) :113–126, 2004. (Cité page 31.)
- A. Cheung, C. Cabrera, P. Sarabandi, KK Nair, A. Kiremidjian, et H. Wenzel. The application of statistical pattern recognition methods for damage detection to field data. *Smart Materials and Structures*, 17, 2008. (Cité page 18.)
- C. Craig, R. D. Neilson, et J. Penman. The use of correlation dimension in condition monitoring of systems with clearance. *Journal of Sound and Vibration*, 231(1) :1–17, Mars 2000. (Cité page 50.)
- C. Crémona. Surveillance de santé structurale, la démarche du projet s3. Dans *Proceedings of GC'2009, Paris*, 2009. (Cité page 9.)
- A. Cury et C. Crémona. Pattern recognition of structural behaviors based on learning algorithms and symbolic data concepts. *Structural Control and Health Monitoring*, 2010. (Cité pages 15, 27 et 31.)
- A. Cury, C. Crémona, et E. Diday. Application of symbolic data analysis for structural modification assessment. *Engineering Structures*, 32 :762–775, 2010. (Cité page 15.)
- G. Cybenko. Approximation by superpositions of a sigmoidal function. *Mathematics of Control, Signals, and Systems (MCSS)*, 2(4) :303–314, 1989. (Cité page 26.)
- S. Das, A.N. Srivastava, et A. Chattopadhyay. Classification of damage signatures in composite plates using one-class svms. Dans *Aerospace Conference, 2007 IEEE*, pages 1–19. IEEE, 2007. (Cité page 32.)
- S.W. Doebling, C.R. Farrar, et M.B. Prime. A summary review of vibration-based damage identification methods. *Los Alamos National Laboratory*, pages 1–34, 1997. (Cité pages 10 et 20.)
- S.W. Doebling, C.R. Farrar, M.B. Prime, et D.W. Shevitz. Damage identification and health monitoring of structural and mechanical systems from changes in their vibration characteristics : A literature review. Rapport technique, Los Alamos National Lab., NM (United States), May 1996. (Cité pages 10 et 20.)
- J.P. Eckmann, S.O. Kamphorst, D. Ruelle, et S. Ciliberto. Liapunov exponents from time series. *Physical Review A*, 34(6) :4971–4979, 1986. (Cité page 68.)

- C.R. Farrar, WE Baker, TM Bell, KM Cone, TW Darling, TA Duffey, A. Eklund, et A. Migliori. Dynamic characterization and damage detection in the i-40 bridge over the rio grande. Rapport technique, Los Alamos National Lab., NM (United States), 1994. (Cité page 78.)
- C.R. Farrar, P.J. Cornwell, S.W. Doebling, et M.B. Prime. Structural health monitoring studies of the alamosa canyon and i-40 bridges. Report, Los Alamos National Laboratory, 2000. (Cité pages 11 et 12.)
- C.R. Farrar et S.W. Doebling. Lessons learned from applications of vibration-based damage identification methods to a large bridge structure. 1997. (Cité page 32.)
- C.R. Farrar et K. Worden. An introduction to structural health monitoring. *Philosophical Transactions of the Royal Society A : Mathematical, Physical and Engineering Sciences*, 365(1851) :303, 2007. (Cité page 14.)
- C.R. Farrar, K. Worden, M.D. Todd, G. Park, D.E. Nichols, J. and Adams, M.T. Bement, et K.M. Farinholt. Nonlinear system identification for damage detection. Report, Los Alamos National Laboratory, 2007. (Cité page 44.)
- E. Figueiredo, G. Park, J. Figueiras, C. Farrar, et K. Worden. Structural health monitoring algorithm comparisons using standard data sets. Rapport technique, Los Alamos National Laboratory, 2009. (Cité pages 17, 18 et 21.)
- E. Figueiredo, M.D. Todd, C.R. Farrar, et E. Flynn. Autoregressive modeling with state-space embedding vectors for damage detection under operational variability. *International Journal of Engineering Science*, 2010. (Cité page 62.)
- A.M. Fraser et H.L. Swinney. Independent coordinates for strange attractors from mutual information. *Physical Review A*, 33(2) :1134–1140, February 1986. (Cité pages 46 et 64.)
- M.L. Fugate, H. Sohn, et C.R. Farrar. Unsupervised learning methods for vibration-based damage detection. Dans *Proceedings of 18th International Modal Analysis Conference-IMAC*. Citeseer, 2000. (Cité pages 18, 23 et 24.)
- G.V. Garcia et R. Osegueda. Combining damage index method and arma method to improve damage detection. Dans *IMAC-XVIII : A Conference on Structural Dynamics, San Antonio, TX*, pages 668–673, 2000. (Cité pages 17 et 20.)
- K. Geist, U. Parlitz, et W. Lauterborn. Comparison of different methods for computing lyapunov exponents. *Progress of Theoretical Physics*, 83 : 875–893, Mai 1990. (Cité page 51.)
- S. Ghafari, F. Golnaraghi, et F. Ismail. Effect of localized faults on chaotic vibration of rolling element bearings. *Nonlinear Dynamics*, 53(4) :287–301, 2008. (Cité pages 44 et 50.)

- D.F. Giraldo, S.J. Dyke, et J.M. Caicedo. Damage detection accommodating varying environmental conditions. *Structural Health Monitoring*, 5(2) : 155, 2006. (Cité pages 12 et 20.)
- S. Girard. *Contribution à l'inférence statistique semi- et non-paramétrique*. Habilitation à diriger des recherches, Université Joseph Fourier, Grenoble I, 2004. (Cité page 36.)
- P. Grassberger et I. Procaccia. Characterization of strange attractors. *Physical Review Letters*, 50(5) :346–349, January 1983. (Cité page 49.)
- M. Gul et F.N. Catbas. Statistical pattern recognition for structural health monitoring using time series modeling : Theory and experimental verifications. *Mechanical Systems and Signal Processing*, 23(7) :2192– 2204, 2009. (Cité pages 17 et 22.)
- N. Haritos et J.S. Owen. The use of vibration data for damage detection in bridges : a comparison of system identification and pattern recognition approaches. *Structural Health Monitoring*, 3(2) :141, 2004. (Cité page 13.)
- L.M. Hively, P.C. Gailey, et V.A. Protopopescu. Detecting dynamical change in nonlinear time series. *Physics Letters A*, 258(2-3) :103–114, Juillet 1999. (Cité page 44.)
- V. Hodge et J. Austin. A survey of outlier detection methodologies. *Artificial Intelligence Review*, 22 :85–126, 2004. (Cité pages 15 et 16.)
- J.R.M. Hosking et J.R. Wallis. Parameter and quantile estimation for the generalized pareto distribution. *Technometrics*, 29(3) :339–349, 1987. (Cité page 72.)
- JRM Hosking, J.R. Wallis, et EF Wood. Estimation of the generalized extreme-value distribution by the method of probability-weighted moments. *Technometrics*, pages 251–261, 1985. (Cité page 37.)
- T.-Y. Hsu et C.-H. Loh. Damage detection accommodating nonlinear environmental effects by nonlinear principal component analysis. *Structural Control and Health Monitoring*, 17(3) :338–354, 2010. (Cité pages 20 et 27.)
- J.N. Juang et M.Q. Phan. *Identification and Control of Mechanical Systems*. Cambridge University Press, 2001. (Cité page 12.)
- S. Kacimi et S. Laurens. The correlation dimension : A robust chaotic feature for classifying acoustic emission signals generated in construction materials. *Journal of Applied Physics*, 106(2) :024909–024909, 2009. (Cité page 50.)
- H. Kantz et T. Schreiber. *Nonlinear Time Series Analysis*. Cambridge University Press, 2004. (Cité pages 48 et 51.)
- J. Kaplan et J. Yorke. Chaotic behavior of multidimensional difference equations. *Functional Differential equations and approximation of fixed points*, pages 204–227, 1979. (Cité page 51.)

- C.-Y. Kim, D.-S. Jung, N.-S. Kim, S.-D. Kwon, et M. Feng. Effect of vehicle weight on natural frequencies of bridges measured from traffic-induced vibration. *Earthquake Engineering and Engineering Vibration*, 2 :109–115, 2003. 10.1007/BF02857543. (Cité page 13.)
- S. Kotz et S. Nadarajah. *Extreme value distributions : theory and applications*. World Scientific Publishing Company, 2000. (Cité page 37.)
- M.A. Kramer. Nonlinear principal component analysis using autoassociative neural networks. *AIChE journal*, 37(2) :233–243, 1991. (Cité page 27.)
- C. Krämer, C. A. M. De Smet, et G. De Roeck. Z24 bridge damage detection tests. Dans *In Proceedings of the International Modal Analysis Conference, IMAC XVII*, pages —, Kissimmee, FL, USA,, 1999. (Cité page 13.)
- LANL. The bookshelf structure. <http://institute.lanl.gov/ei/>, 2000. (Cité pages 78 et 80.)
- B. LeBaron. Chaos and nonlinear forecastability in economics and finance. *Philosophical Transactions : Physical Sciences and Engineering*, pages 397–404, 1994. (Cité page 44.)
- H. Li, S. Li, J. Ou, et H. Li. Modal identification of bridges under varying environmental conditions : Temperature and wind effects. *Structural Control and Health Monitoring*, 2009. (Cité page 11.)
- R.A. Livingston, S. Jin, et D. Marzougui. Application of nonlinear dynamics analysis to damage detection and health monitoring of highway structures. Dans *Proceedings of SPIE - The International Society for Optical Engineering*, volume 4337, pages 402–410, Newport Beach,CA, 2001. Society of Photo-Optical Instrumentation Engineers. (Cité page 51.)
- D.B. Logan et J. Mathew. Using the correlation dimension for vibration fault diagnosis of rolling element bearings–ii. selection of experimental parameters. *Mechanical Systems and Signal Processing*, 10(3) :251–264, Mai 1996. (Cité page 50.)
- E.N. Lorenz. Deterministic nonperiodic flow1. *Atmos J Sci*, 20 :130–141, 1963. (Cité page 46.)
- Y. Lu et J.E. Michaels. State space feature extraction applied to diffuse ultrasonic signals using simulated chaotic excitations. Dans *AIP Conference Proceedings*, volume 820, page 625, 2005. (Cité page 48.)
- B.B. Mandelbrot. *The fractal geometry of nature*. New York ; San Francisco , W.H. Freeman and Co, 1983. (Cité page 49.)
- M. Markou et S. Singh. Novelty detection : a review–part 1 : statistical approaches. *Signal Processing*, 83(12) :2481–2497, 2003a. (Cité page 16.)
- M. Markou et S. Singh. Novelty detection : a review–part 2 : : : neural network based approaches. *Signal Processing*, 83(12) :2499–2521, 2003b. (Cité page 16.)

- A. Mita et H. Hagiwara. Quantitative damage diagnosis of shear structures using support vector machine. *KSCE Journal of Civil Engineering*, 7(6) : 683–689, 2003. (Cité page 32.)
- L. Moniz, J.M. Nichols, C.J. Nichols, M. Seaver, S.T. Trickey, M.D. Todd, L.M. Pecora, et L.N. Virgin. A multivariate, attractor-based approach to structural health monitoring. *Journal of Sound and Vibration*, 283(1-2) : 295–310, 2005. (Cité pages 47 et 54.)
- D.C. Montgomery. *Introduction to Statistical Quality Control, 5th edition*. John Wiley & Sons, Inc., 2005. (Cité pages 33, 35 et 172.)
- K.K. Nair, A.S. Kiremidjian, et K.H. Law. Time series-based damage detection and localization algorithm with application to the asce benchmark structure. *Journal of Sound and Vibration*, 291(1-2) :349–368, 2006. (Cité page 17.)
- J.M. Nichols. Structural health monitoring of offshore structures using ambient excitation. *Applied Ocean Research*, 25(3) :101–114, Juin 2003. (Cité pages 53 et 54.)
- J.M. Nichols, M.J. Hunther, M. Seaver, et S.I. Trickey. The role of excitation in attractor-based structural health monitoring. Dans *Proceedings of the 4th International Workshop on Structural Health Monitoring*, Stanford University, Stanford, CA, Septembre 2003a. DEStech Publications, Inc. (Cité pages 47 et 54.)
- J.M. Nichols et J.D. Nichols. Attractor reconstruction for non-linear systems : a methodological note. *Mathematical Biosciences*, 171(1) :21–32, Mai 2001. (Cité pages 63 et 64.)
- J.M. Nichols, M.D. Todd, et J.R. Wait. Using state space predictive modeling with chaotic interrogation in detecting joint preload loss in a frame structure experiment. *Smart Materials and Structures*, 12(4) :580–601, 2003b. (Cité pages 47 et 54.)
- J.M. Nichols, S.T. Trickey, et M. Seaver. Damage detection using multivariate recurrence quantification analysis. *Mechanical Systems and Signal Processing*, 20(2) :421–437, Février 2006. (Cité page 47.)
- J.M. Nichols, L.N. Virgin, M.D. Todd, et J.D. Nichols. On the use of attractor dimension as a feature in structural health monitoring. *Mechanical Systems and Signal Processing*, 17(6) :1305–1320, 2003c. (Cité pages 47 et 50.)
- C.K. Oh et H. Sohn. Damage diagnosis under environmental and operational variations using unsupervised support vector machine. Dans *Journal of Sound and Vibration* (Sohn et al. 2002), pages 224–239. (Cité pages 19, 28 et 196.)
- C.K. Oh, H. Sohn, et I.-H. Bae. Statistical novelty detection within the yeongjong suspension bridge under environmental and operational variations. Dans *Structural health monitoring* (Oh et Sohn 2009), pages 224–239. (Cité pages 26 et 38.)

- C.C. Olson, L.A. Overbey, et M.D. Todd. Sensitivity and computational comparison of state-space methods for structural health monitoring. Dans *Proceedings of SPIE*, volume 5768, page 241, 2005. (Cité pages 53 et 54.)
- V.I. Oseledec. A multiplicative ergodic theorem : Lyapunov characteristic numbers for dynamical systems. *Trans. Moscow Math. Soc.*, 19 :197–231, 1968. (Cité page 67.)
- L.A. Overbey, C.C. Olson, et M.D. Todd. A parametric investigation of state-space-based prediction error methods with stochastic excitation for structural health monitoring. *Smart Materials and Structures*, (5) :1621, 2007. (Cité pages 48 et 53.)
- H.W. Park et H. Sohn. Parameter estimation of the generalized extreme value distribution for structural health monitoring. *Probabilistic Engineering Mechanics*, 21(4) :366–376, 2006. (Cité pages 37 et 198.)
- E. Parzen. On estimation of a probability density function and mode. *The annals of mathematical statistics*, 33(3) :1065–1076, 1962. (Cité page 23.)
- L.M. Pecora, H.P. Geffert, S. Boccaletti, D.L. Valladares, et T. Carroll. Reconstructing embedding spaces of coupled dynamical systems from multivariate data. *Physical Review E. Statistical Physics, Plasmas, Fluids, and Related Interdisciplinary Topics*, 65(3 2A), 2002. (Cité page 47.)
- B. Peeters. *System identification and damage detection in civil engineering*. PhD thesis, Katholieke Universiteit Leuven, 2000. (Cité page 12.)
- B. Peeters et G. De Roeck. One year monitoring of the z 24-bridge : environmental influences versus damage events. Dans *PROC INT MODAL ANAL CONF IMAC.*, volume 2, pages 1570–1576, 2000. (Cité page 78.)
- B. Peeters et G. De Roeck. Stochastic system identification for operational modal analysis : A review. *Journal of Dynamic Systems, Measurement, and Control*, 123(4) :659–667, 2001. (Cité pages 10, 11, 12, 17 et 32.)
- M. Piera-Martinez. *Modélisation des comportements extrêmes en ingénierie, EVS*. PhD thesis, Univ Paris-Sud- Paris XI, 2008. (Cité page 72.)
- M. Pillet. *Appliquer la maîtrise statistique des processus MSP/SPC, 4ème ed.* Eyrolles, 2008. (Cité page 33.)
- A. Rytter. *Vibration based inspection of civil engineering structures*. PhD thesis, Department of Building Technology and Structural Engineering, University of Aalborg, Danemark, 1993. (Cité pages 10 et 16.)
- M. Sano et Y. Sawada. Measurement of the lyapunov spectrum from a chaotic time series. *Physical Review Letters*, 55(10) :1082–1085, 1985. (Cité page 68.)
- B. Schölkopf, A. Smola, et K.R. Müller. Nonlinear component analysis as a kernel eigenvalue problem. *Neural computation*, 10(5) :1299–1319, 1998. (Cité page 26.)

- B. Schölkopf et A.J. Smola. *Learning with kernels : support vector machines, regularization, optimization, and beyond*. the MIT Press, 2002. (Cité page 31.)
- B. Scölkopf, R.C. Williamson, A.J. Smola, J. Shawe-Taylor, et J. Platt. Support vector method for novelty detection. *Advances in neural information processing systems*, 12(3) :582–588, 2000. (Cité page 31.)
- D.W. Scott et S.R. Sain. Multidimensional density estimation. *Handbook of Statistics*, 24 :229–261, 2005. (Cité page 24.)
- A.J. Smola et B. Schölkopf. A tutorial on support vector regression. *Statistics and computing*, 14(3) :199–222, 2004. (Cité page 19.)
- H. Sohn. Effects of environmental and operational variability on structural health monitoring. *A Special Issue of Philosophical Transaction of Royal Society A on Structural Health Monitoring*, 365 :539–560, Feb 2007. (Cité page 13.)
- H. Sohn, D.W. Allen, K. Worden, et C.R. Farrar. Statistical damage classification using sequential probability ratio tests. *Structural Health Monitoring*, 2(1) :57–74, Mars 2003. (Cité pages 39 et 40.)
- H. Sohn et C.R. Farrar. Statistical process control and projection techniques for damage detection. Dans *Proceedings of European COST F3 Conference on System Identification and Structural Health Monitoring*, pages 105–114, 2000. (Cité page 34.)
- H. Sohn et C.R. Farrar. Damage diagnosis using time series analysis of vibration signals. *Smart Material Structures*, 10 :446–451, jun 2001. (Cité pages 18 et 19.)
- H. Sohn, K. Worden, et C.R. Farrar. Statistical damage classification under changing environmental and operational conditions. *Journal of Intelligent Material Systems and Structures*, 13(9) :561–574, Septembre 2002. (Cité pages 19, 28 et 196.)
- Hoon Sohn, David W. Allen, Keith Worden, et Charles R. Farrar. Structural damage classification using extreme value statistics. Dans *Journal of Dynamic Systems, Measurement, and Control* (Park et Sohn 2006), pages 125–132. (Cité pages 19, 36 et 38.)
- Z. Sun. Structural damage assessment based on wavelet packet transform. *Journal of Structural Engineering*, 128 :1354, 2002. (Cité page 27.)
- M.M.R. Taha, A. Noureldin, J.L. Lucero, et T.J. Baca. Wavelet transform for structural health monitoring : a compendium of uses and features. *Structural Health Monitoring*, 5(3) :267–296, 2006. (Cité pages 20 et 21.)
- F. Takens. Detecting strange attractors in turbulence. *Dynamical systems and turbulence, Warwick 1980*, pages 366–381, 1980. (Cité page 46.)
- D.M.J. Tax. One-class classification. phd, Delft University of Technology, Delft, June 2001. (Cité pages 22 et 32.)

- J. Theiler. Spurious dimension from correlation algorithms applied to limited time-series data. *Phys. Rev. A*, 34(3) :2427, Septembre 1986. (Cité pages 49 et 67.)
- M. Todd, J. Nichols, M. Bement, C.R. Farrar, et B. Bakers. Experimental demonstration of local attractor variance as a damage indication feature. Dans *Proceedings of SPIE - The International Society for Optical Engineering*, volume 4753 II, pages 1292–1298, Los Angeles, CA, United States, 2002. The International Society for Optical Engineering. (Cité page 47.)
- M.D. Todd, K. Erickson, L. Chang, K. Lee, et J.M. Nichols. Using chaotic interrogation and attractor nonlinear cross-prediction error to detect fastener preload loss in an aluminum frame. *Chaos*, 14(2) :387–399, Juin 2004. (Cité pages 47, 53 et 54.)
- M.D. Todd, J.M. Nichols, L.M. Pecora, et L.N. Virgin. Vibration-based damage assessment utilizing state space geometry changes : local attractor variance ratio. *Smart Materials and Structures*, 10(5) :1000–1008, 2001. (Cité pages 47 et 54.)
- J. Toivola, M. Prada, et J. Hollmén. Novelty detection in projected spaces for structural health monitoring. *Advances in Intelligent Data Analysis IX*, pages 208–219, 2010. (Cité page 24.)
- C.H. Tsou. *A statistical learning framework for data mining of large-scale systems : algorithms, implementation, and applications*. PhD thesis, Massachusetts Institute of Technology, 2007. (Cité pages 29 et 32.)
- W.H. Woodall. Controversies and contradictions in statistical process control. *Journal of Quality Technology*, 32(4) :341–350, 2000. (Cité page 34.)
- K. Worden, W.A. Allen, H. Sohn, W. Stinemates, et C.R. Farrar. Extreme value statistics for damage detection in mechanical structures, evs. Rapport technique, Los Alamos National Lab., NM (United States), 2002a. (Cité page 38.)
- K. Worden et J.M. Dulieu-Barton. An overview of intelligent fault detection in systems and structures. *Structural Health Monitoring*, 3(1) :85, 2004. (Cité page 9.)
- K. Worden, C.R. Farrar, G. Manson, et G. Park. The fundamental axioms of structural health monitoring. *Proceedings of the Royal Society A : Mathematical, Physical and Engineering Science*, 463(2082) :1639, 2007. (Cité pages 15 et 21.)
- K. Worden et G. Manson. The application of machine learning to structural health monitoring. *Philosophical Transactions of the Royal Society A : Mathematical, Physical and Engineering Sciences*, 365(1851) :515, 2007. (Cité pages 27, 31 et 171.)
- K. Worden, G. Manson, et N. R. J. Fieller. Damage detection using outlier analysis. *Journal of Sound and Vibration*, 229(3) :647–667, Janvier 2000. (Cité pages 15 et 22.)

- K. Worden, H. Sohn, et C.R. Farrar. Novelty detection in a changing environment : regression and interpolation approaches. *Journal of Sound and Vibration*, 258(4) :741–761, 2002b. (Cité pages 22 et 33.)
- A.-M. Yan, G. Kerschen, P. De Boe, et J.-C. Golinval. Structural damage diagnosis under varying environmental conditions—part i : A linear analysis. *Mechanical Systems and Signal Processing*, 19(4) :847–864, Juillet 2005. (Cité page 25.)
- J. Yang, Y. Zhang, et Y. Zhu. Intelligent fault diagnosis of rolling element bearing based on svms and fractal dimension. *Mechanical Systems and Signal Processing*, 21(5) :2012–2024, 2007. (Cité page 50.)
- G.G. Yen et K.C. Lin. Wavelet packet feature extraction for vibration monitoring. *IEEE Transactions on Industrial Electronics*, 47(3) :650–667, 2000. (Cité page 21.)
- J.L. Zapico-Valle, M. García-Diéguez, M.P. González-Martínez, et K. Worden. Experimental validation of a new statistical process control feature for damage detection. *Mechanical Systems and Signal Processing*, In Press, Corrected Proof, 2011. (Cité page 39.)

RÉSUMÉ

Le suivi vibratoire de l'état des ouvrages de génie civil vise à anticiper une défaillance structurale par la détection précoce d'endommagement. Dans ce contexte, la détection de nouveauté constitue une approche particulièrement adaptée à l'analyse des signaux compte tenu des difficultés à modéliser une structure unique et soumise à de nombreux facteurs extérieurs influant sur la dynamique vibratoire. Le premier objectif du travail de thèse consiste à observer dans quelle mesure la détection de nouveauté parvient à détecter un endommagement dans un contexte fortement perturbé par des variations environnementales d'une part, et par une excitation de nature impulsionnelle, d'autre part. Le deuxième objectif est de proposer et d'étudier un nouvel indicateur vectoriel, désigné par JFV (pour Jacobian Feature Vector). Le calcul du JFV s'appuie sur la reconstruction de la trajectoire du système dynamique observé dans son espace des phases. Cette approche exploite les développements scientifiques récents réalisés en théorie des systèmes dynamiques non linéaires, parfois qualifiée de théorie du chaos. Le JFV est comparé aux coefficients de modèles auto-régressifs (AR), couramment utilisés en analyse des séries temporelles. Pour réaliser ce travail de thèse, plusieurs cas d'études expérimentaux sont utilisés dont notamment une maquette de structure en bois sur laquelle l'excitation est contrôlée et des variations environnementales sévères sont imposées. Enfin, différentes approches de modélisation statistique des indicateurs normalisés sont mises en oeuvre dans le but de comparer leurs aptitudes respectives à la définition d'un seuil de classification robuste.

MOTS CLÉS

détection de nouveauté, structure, analyse vibratoire, détection de l'endommagement, théorie des systèmes dynamiques non-linéaires, espace des phases, attracteur, variations environnementales, sollicitation instationnaire